Proceedings of the Twenty-Third International Conference on Automated Planning and Scheduling

Autonomous Search and Tracking via Temporal Planning*

Sara Bernardini, Maria Fox, Derek Long
Department of Informatics
King's College London
London, UK, WC2R 2LS
firstname.lastname@kcl.ac.uk

John Bookless
Advanced Technology Centre
BAE Systems
Bristol, UK, BS34 7QW
John.Bookless@baesystems.com

Abstract

Search And Tracking (SAT) is the problem of searching for a mobile target and tracking it after it is found. As this problem has important applications in search-and-rescue and surveillance operations, recently there has been increasing interest in equipping unmanned aerial vehicles (UAVs) with autonomous SAT capabilities. State-of-the-art approaches to SAT rely on estimating the probability density function of the target's state and solving the search control problem in a greedy fashion over a short planning horizon (typically, a one-step lookahead). These techniques suffer high computational cost, making them unsuitable for complex problems. In this paper, we propose a novel approach to SAT, which allows us to handle big geographical areas, complex target motion models and long-term operations. Our solution is to track the target reactively while it is in view and to plan a recovery strategy that relocates the target every time it is lost, using a high-performing *automated planning* tool. The planning problem consists of deciding where to search and which search patterns to use in order to maximise the likelihood of recovering the target. We show experimental results demonstrating the potential of our approach.

1 Introduction

The use of unmanned aerial vehicles (UAVs), or drones, has recently attracted a great deal of media attention due to their increasing role in military operations. However, military UAVs are currently remotely piloted and will not necessarily be the first class of UAVs to fly fully autonomously (ASTRAEA Programme 2012). There are many potential civil applications of UAVs, including surveillance, forest fire watch, agricultural missions and search operations. In this paper we consider the use of a UAV in a SAT mission, in which the target of the operation is a land-based vehicle that is to be followed on its course through various terrain (both urban and rural). We examine the way that such a task decomposes into separate phases and explore the role of planning amongst these phases. We identify a planning problem that arises in this mission and show how it can be modelled and solved using generic planning technology. We comment on the benefits of generic planning in this role and demonstrate its performance in simulation, showing results that indicate its potential effectiveness.

The development and production of UAVs is a growing sector. One of the significant challenges in moving the technology from purely military function to a wider exploitation of autonomous vehicles is the demand for certification within the legal framework of civil aviation rules. This challenge is somewhat paradoxical in that certification (which is currently built around a presumption that a human pilot will be responsible for an aircraft) will depend on building trust in autonomous UAV technology, which demands logging many hours of safe autonomous flight, while logging hours requires permission to fly in interesting air space. A part of the process of breaking through this apparent log-jam is to build autonomous flight control systems that have a high degree of predictable behaviour and that can be easily understood and verified. A major element in this is the development of robust and predictable controllers (Léauté and Williams 2005; Mathews and Durrant-Whyte 2007; Doherty, Kvarnström, and Heintz 2009), but as the control level is increasingly seen as a well- understood problem, attention is moving to the integration of control with higher-level decision making and flight planning.

2 Search and Tracking Missions

The SAT mission we consider in this paper is one in which the target is a vehicle (a car, for example) being sought and tracked through a mixed urban and rural landscape. We consider a single UAV and we assume that the vehicle controllers are robust and can be relied on to provide basic flight control to maintain level flight, turning and localising.

The UAV is assumed to be equipped with imaging systems that allow the target to be observed. Observation is assumed to be susceptible to error and interference from terrain. In our model the probability of observing the target on each observation cycle (which can be considered as a 'frame capture' by the imager) depends on how recently the target was last observed, the distance between the actual position of the target and the predicted position of the target, the speed of the target, the terrain and the mode of the imaging system. We assume that the imager has two modes: wide-angle mode used to increase the area being scanned when

*The authors gratefully acknowledge support from EPSRC grant EP/J012157/1 and from BAE Systems Ltd.

the target is not currently observed, at the cost of a lower probability of successfully observing the target in any specific part of the image, and narrow-angle mode in which the viewing area is reduced, but the probability of detecting the target is higher. The effect of terrain is to reduce the probability of spotting the target in urban, suburban, forested and mountainous areas, while in rough or open rural areas the probability is higher. A faster moving target in the viewing zone is considered easier to spot.

In this work we do not consider significant evasive action by the target — the target becomes slightly more erratic in speed (both faster and slower than typical for the road) once it is aware of the observer, but we do not consider attempts to leave the road, to use terrain for concealment or deviation from the originally intended path. These, of course, are all reasonable possibilities for the behaviour of the target and we intend to explore them in future work.

The objective of the mission is to follow the target to its destination. In general, a SAT mission proceeds in two phases, which interleave until the target stops or until the observer acknowledges it has lost the target irrevocably. These phases are *tracking*, where the UAV flies in a standard pattern over the target, observing its progress, and *search* in which the UAV has lost the target and flies a series of manoeuvres intended to rediscover the target. Once the target is rediscovered the UAV switches back to tracking mode.

Tracking is managed by a reactive controller: the problem is simply one of finding a route that maximises observations of the target. In general, UAVs fly much faster than road vehicles drive, but fixed-wing UAVs cannot hover. Therefore, the flight path of a fixed-wing UAV in tracking mode is a circle of fixed radius centred on the target. The radius depends on the capabilities of the UAV: it cannot be greater than the range of the imaging equipment, nor can it be shorter than the turning radius of the UAV at current speed. We assume that the UAV flies in a mid-range circle between these extremes (which are parameters of our simulation). In fact, as the target moves the circle moves with it, so the flight path of the UAV describes a spiralling pattern over the ground.

If the UAV fails to observe the target, it must attempt to rediscover it. How this is achieved depends on the time since the observer last observed the target. For a short period after losing the target the UAV can simply track the predicted location of the target, since the target cannot move fast enough to significantly deviate from this prediction. However, after a period (whose length depends on the speed of the target, the arc of the imaging equipment, the observation probability and, possibly, the terrain), it will be necessary to make a more systematic effort to rediscover the target by directing the search into specific places.

At this point, the problem can be seen as follows: the current state includes information about the last known location and velocity of the target, the average velocity over the period the target has been tracked, the map of the terrain and the current position of the UAV. The UAV can fly across the terrain in search of the target. It is possible to view this task as a planning problem (see Figure 1) in which a search plan is constructed from a set of candidate search patterns that can be arranged in a sequence to attempt to optimise the

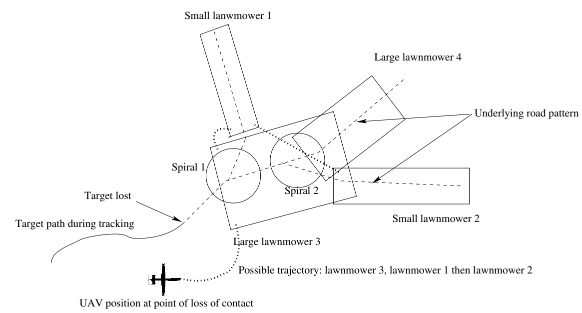

Figure 1: Initial state and search plan for the SAT mission.

likelihood of rediscovering the target. If, while flying this search plan, the target is rediscovered, the observer switches back to tracking mode.

3 Related Work

Over the past ten years there has been growing interest in efficient SAT, which has practical applications in search-and-rescue, surveillance, boarder patrol and reconnaissance operations. An approach that has been extensively explored relies on the use of Bayesian techniques. Research originally considered the two areas of searching and tracking separately and focused on scenarios with a single target and a single vehicle (Bourgault, Furukawa, and Durrant-Whyte 2006). However, the field has rapidly evolved. A unified approach to SAT has been proposed (Furukawa et al. 2006) and a number of different settings have been explored, including: (i) single/multiple targets; (ii) single/multiple vehicles; and (iii) centralised/decentralised strategies.

The probabilistic approach to SAT relies on the use of Recursive Bayesian Estimation (RBE) techniques that recursively update and predict the Probability Density Function (PDF) of the target's state with respect to time, under the assumption that the prior distribution and the probabilistic motion model of the target are known. The target's motion model that is usually adopted is represented by a simple discrete-time equation establishing that the state of the target at time step k+1 is a function of the state of the target at step k, the target's control input at step k and the system noise at step k. The system noise corresponds to environmental influences on the target. Although Bourgault et al. (2004) discuss a number of possible hard and soft constraints that can impact the target motion model, such as obstacles, force fields and terrain topology, the target is usually assumed to be subjected to external disturbances and not to move on the basis of its own intentions. For example, UAVs might search for and track life-rafts drifting with wind and current.

RBE techniques work in two stages, the update and the prediction. The update stage calculates the posterior distribution of the current state given a prior estimation of the state (based on the sequence of previous observations) and a new observation at present time. The prediction stage calculates the PDF of the next state using the posterior density of the current state and the target's motion model. From a computational point of view, the implementation of these two stages of RBE essentially requires the evaluation of a function at an arbitrary point in the target space and the integration of a function over the target space. As these operations are computationally expensive, a number of different approaches have been explored to compute them efficiently, including grid-based methods (Bourgault, Furukawa, and Durrant-Whyte 2006), particle filters (Chung and Furukawa 2006), element-based techniques (Furukawa, Durrant-Whyte, and Lavis 2007) and hybrid particle-element approaches (Lavis and Furukawa 2008).

Based on the calculated PDF, the search control problem is solved in a greedy fashion over a very short planning horizon; typically, a one-step lookahead horizon is adopted. This myopic planning approach is used to control the computational cost of the technique, which quickly becomes intractable as the number of lookahead steps, the size of the search area, or the number of dimensions of the search space, increases. A unified objective function is used for both search and tracking, allowing a vehicle to switch from one mode to the other while maintaining information gained during the previous phases.

Probabilistic-based SAT has proven successful for problems involving stationary targets or targets moving in small geographical areas, simple motion models in which the targets do not show any evading or intentional behaviour, static search spaces (although a technique to dynamically reconfigure the state space has been recently proposed by Lavis & Furukawa (2008)), and short-term missions. Whenever these assumptions are not satisfied, especially for SAT operations over large geographical areas, complex target motion models and long-term operations, RBE techniques perform poorly due to the high computational cost of accurately maintaining a large state space that includes all the possible positions of the moving targets.

From a theoretical point of view, our formulation of SAT as a planning problem (described in Section 4.1) resembles the *Orienteering Problem with Time Windows* (OPTW) (Kantor and Rosenwein 1992). In a classical orienteering problem (OP), a set of vertices with associated rewards is given as well as a starting and an ending vertex. For all the vertices, the amount of time t_{ij} needed to travel from vertex v_i to vertex v_j is known. Finally, a maximum time budget T_{max} is established. The goal of the OP is to determine a path that visits a subset of the vertices in the available time T_{max} in order to maximise the total collected reward. In the time window variant of the OP, each vertex is associate with a time window and a visit to that vertex can only start during that window. The OPTW is a hard combinatorial problem because three types of decisions are involved in it: (i) allocation of vertices to be visited; (ii) sequencing of vertices to be visited; and (iii) scheduling of the visits to the chosen set of vertices. Considering our planning problem, the set of search patterns corresponds to the set of vertices of the OPTW problem, whereas the time slots in which the search patterns are active correspond to the OPTW time windows. As in the OPTW, we also want to maximise the total reward in the available amount of time (limited by the window of the latest possible search). However, in our case and differently from the OPTW, the planner can choose to visit each location more than once and needs to decide which search pattern to use at each location.

In the context of planning, the OP has been used to provide suitable heuristic advice on the goals and the goal order that should be considered by a planner that deals with oversubscription problems (Smith 2004).

4 Search as Planning

If the UAV observer loses the target beyond the short period for which it tracks the predicted location of the target, it must follow a search strategy to attempt to rediscover the target. As described in Section 3, Furukawa et al. (2012) have proposed and explored an approach to searching for a target based on modelling the predicted position of the target and then controlling the UAV to maximise the expectation of rediscovery. Their approach generates local control instructions for the UAV, responding to relatively fine-grained predicted behaviour of the target. Although, in principle, the probabilistic model of the predicted behaviour of the target might be constructed within a confined area (Mathews, Waldock, and Paraskevaides 2010), there may be scalability issues when this technique is applied to larger, more sophisticated problems in which the highly constrained and complex structure of a road network is taken in consideration. Instead, we propose that a more tractable approach is to exploit standard search patterns and use these as the building blocks for a search plan that attempts to maximise the expectation of rediscovering the target.

We consider two standard search patterns: a spiral, starting near the centre of a large circular area and spiralling outwards to some maximum radius, and a lawnmower pattern over a rectangular area. The former pattern is effective for covering an area of high density road network, particularly in urban or suburban terrain, while the latter is useful when attempting to search over a more elongated stretch covering a major road and including some possible side roads. A significant benefit of using these standard patterns is that they are well-recognised search patterns and the UAV flight is predictable and recognisable while it is following them, supporting the building of trust and confidence in the behaviour of autonomous UAVs.

The challenge, then, is to decide where these search patterns should be deployed. One possibility is to use a fixed strategy of simply flying some standard configuration of these patterns over the area where the target was lost. A simple example is to fly a spiral centred over the area where the target is predicted to be currently, which will steadily expand to include all the area the target might have reached during the search, followed by a succession of spirals or lawnmowers extending out in the direction the target appeared to be heading, starting with spirals over urban or suburban areas and switching to lawnmowers in more rural areas, centred over roads. This fixed policy was proposed to us by our industrial collaborators and is used for comparison.

A more interesting strategy is to see the problem of selection of search patterns as a planning problem: each search pattern can be assigned a value corresponding to the expectation of finding the target in a search of that area. We discuss how this value can be estimated shortly, but for the moment we note that it is a function of time, since the target will not be found in an area that is far from its last known location until sufficient time has passed for the target to have reached the area, while it is unlikely that the target will be found in the area once sufficient time has passed for it to have driven through the area (unless the area includes the destination of the target). The UAV can select a sequence of search patterns, linking them together with a series of flight paths, in order to maximise the accumulated expectation of rediscovery. This problem is very similar to an orienteering problem with time-dependent rewards (Kantor and Rosenwein 1992), as described in Section 3, although there are two important differences: one is that the reward for an area can be claimed more than once (although with some provisos further discussed below) and, unlike a plan for the orienteering problem, the execution of the search plan terminates once the target is rediscovered.

This planning problem has some interesting features: despite the inherent uncertainty in the situation, the problem is deterministic, since the uncertainty arises in the position of the target and, if the target is found, the plan ceases to be relevant. Therefore, the plan is constructed entirely under the assumption that the target remains undiscovered. Somewhat counter-intuitively, "plan-failure" corresponds to the situation in which the target is found, counter to the assumption on which the plan is based. However, on failure of this assumption, the plan becomes irrelevant and the UAV switches back to tracking mode.

We exploit the period in which the UAV tracks the predicted location of the target to perform planning. The problem has no goal, but the plan metric measures the value of the plan in terms of the accumulated expectation of finding the target. A few examples of problems of this sort have been considered before (for example, one variant of the satellite observation problem used in the 3rd International Planning Competition (Long and Fox 2003) had this character) and it is typically the case that bench-mark planners generate empty plans, ignoring the metric. We discuss below our management of this problem.

4.1 The Planning Problem

The domain model for the search problem has a very straightforward basic structure: there is a flight action that allows the UAV to fly from one waypoint to another and there are actions allowing it to perform each of the basic search patterns. We use spiral searches and small and large lawnmowers, although other patterns can easily be added. The search pattern actions all have similar forms: they each have an entry waypoint and an exit waypoint and the effect, other than to move the UAV, is to increase the reward (which is the accumulated expectation of finding the target). The actions are durative and their duration is fixed in the problem instance to be the correct (computed) value for the execution of the corresponding search. The search patterns can only be executed so that they coincide with a period during which the target could plausibly be in the area the pattern covers. This is calculated by considering the minimum and maximum reasonable speeds for the target and the distance from where the target was last observed. The reward is more complicated and is discussed in detail below, but the problem instance associates with the pattern a reward, using a time-dependent function.

As an example of how search actions are modelled in PDDL2.1, the following is the description of the action `doSpiral`, which specifies the location and availability of a spiral search pattern, the time it takes to complete it and the reward available.

```
(:durative-action doSpiral
:parameters (?from ?to - waypoint
                    ?p - spiral )
:duration (=?duration (timefor ?p))
:condition (and
        (at start (beginAt ?from ?p))
        (at start (endAt ?to ?p))
        (at start (at ?from))
        (at end (active ?p)))
:effect (and
        (at end (at ?to))
        (at start (not (at ?from)))
        (at end (increase (reward)
                  (rewardof ?p)))))
```

If the search pattern is flown and the target is not found there are two possible explanations: the target was not in the pattern or the search failed to find the target despite its presence. The second case can occur because of poor synchronisation of the turning search arc of the observer with the movement of the target, or because the imager failed to spot the target despite it passing through the search arc. However, the expected reward for searching the pattern a second time should be reduced to reflect the fact that the conditional probability that the target can be found in the area given that the first search failed to find it is lower than the probability that the target can be found in the area before that. In most cases, the time taken to search the pattern will lead to a reduction in the reward for the pattern if it is repeated in any case, due to the time-dependent form of the reward function, but a modified domain model can be used to account for the reduced conditional probability:

```
(:durative-action doSpiral
:parameters (?from ?to - waypoint
                    ?p - spiral )
:duration (=?duration (timefor ?p))
:condition (and
        (at start (beginAt ?from ?p))
        (at start (endAt ?to ?p))
        (at start (at ?from))
        (at end (active ?p)))
:effect (and
     (at end (at ?to))
     (at start (not (at ?from)))
     (at end (increase (reward)
      (+ (rewardof ?p) (basereward ?p))))
     (at end (decrease (basereward ?p)
      (* (rewardof ?p) (discountFactor))))))
```

The reason for the slightly complicated mechanism for discounting the reward for a pattern is that the time-dependent reward function is managed by timed-initial fluents in the problem specification that change the reward of the patterns as time progresses. The shape of the function is constructed to represent an approximate lifted Gaussian distribution, with no reward until the target could plausibly have arrived at the search area and no reward after the target is unlikely to be still present in the area (driving as slowly as reasonable throughout the approach and passage through the area). Between these extremes, the reward peaks at the point where the target would be in the centre of the search pattern if driving at average speed. The model can be modified by adding more or fewer intermediate points in the step function approximation of the distribution. We have also experimented with a continuous linear approximation, using additional actions to create the necessary reward function, but this significantly complicates the planning problem without adding any real benefits (the actions used to model the reward function are artificial actions that the planner is forced to put into the plan in order to achieve the necessary reward functions and the planner is forced to discover the organisation of these artificial actions alongside the real actions used to perform the search).

To ensure that the planner does not exploit search patterns when there is no reward associated with them, the patterns are only made active during the period when the distribution is positive, using timed initial literals (TILs) that are asserted and retracted at the appropriate times. Reward is therefore modelled as a series of n times, t_0 to t_n. At each t_i a TIL asserts the value of the reward function for the interval $[t_i, t_{i+1}]$, with reward being set to 0 in the initial state and reset to 0 by the TIL at t_n. The pattern is only considered action during this period, so a TIL `(active ?p)` is added for the pattern `?p` at t_0 and deleted at t_n.

4.2 Search Patterns and Reward Estimates

To create the initial states for our planning problems, we have to manage two tasks. First, we have to identify candidate search patterns and second we have to assign appropriate values functions to them. The first task is made difficult by the fact that there are infinitely many patterns that could be used, while the second is made difficult because of the lack of knowledge about the intentions of the target.

To address the first problem we observe that the planner can only consider a finite subset of search patterns and, since we want to perform planning in real time, is limited to being able to consider a reasonably small number of candidates. Therefore, we generate a sample of possible search patterns by randomly selecting a set of shapes (circles for spirals and large and small rectangular areas for lawnmowers) and placing them onto the general search area. There are several biasing factors in this. First, we use as our general search area a circular sector that is centred on the last known location of the target and extends outwards with its symmetry axis aligned with the average bearing of the target over the period the target has been observed. The sector extends outwards for a distance of several kilometres (the exact distance is a parameter — we comment on the choice of value in our report on our simulation below), whose value depends on the total area included in the sector and the relative time required to fly a search pattern of a given area. There is a point at which the area where the target could be present is so much larger than the area that the UAV can search, during a period when the target could be present, that the expectation of finding the target diminishes to a negligible value. The angle subtended by the sector is also a parameter and reflects the degree of uncertainty in the heading of the target. In general, a target will follow the direction that is forced on it by the roads it uses, but the average bearing will converge, over time, on the direction from the origin to the destination. The longer the target is observed, the closer will this convergence become, although the target's path will always remain subject to the constraints of the road network.

Once the relevant sector is identified, we then sample points using a probability distribution laid over the sector. This distribution is based on the density of roads across the sector, which is measured by using a fine-mesh grid and counting the number of significant roads within each grid cell, the terrain type (urban, suburban, mountainous, forested, rough or open rural ground) and the distance from the symmetry axis and from the last known location of the target. The distribution decays linearly with distance from the origin, linearly away from the symmetry axis and is weighted by values for terrain type and road density. Again, all of these values are parameters that can be adjusted and we have adopted values we consider appropriate. Although the density of patterns decays away from the origin, the effect is muted because the relative areas available for selection are proportional to the distance from the origin.

We then decide, for each point, the type of pattern to use: we favour spirals in the part of the search closest to the origin, where spirals give good coverage of the area in which the target might be found, and lawnmowers in rural areas or areas of lower road density, where spirals are likely to cover significant areas of little value in the search. For spirals we additionally select a radius based on the width of the sector at that point and the road network density. For lawnmowers we also select an orientation and then width and length. The orientation is based on the road network and is aligned to follow major roads or high densities of roads, while the width and length are determined by examining the underlying road network and probability distribution.

To assign a value function to each pattern we compute a shortest and longest time of arrival for the target by considering an average speed and variation in speed over the path from the origin to the pattern. In principle, this mechanism should use the road map to identify shortest paths, but this is too costly to compute in real time, so we instead sample the terrain along the straight line from the origin to the leading and far edges of the pattern. This is used to as a guide to the likely speed of the target on this path. In practice, if the straight line path traverses rural areas then the target will either have to use smaller roads or else deviate from the direct path in order to exploit more major roads. In either case, the target will arrive at the target later than if the direct path is through suburban terrain. On the other hand, if the terrain is urban then speed will be constrained by traffic laws and other road users. The earliest and latest times are used to set up a value function, with these as the limits of the reward (outside this range the pattern is awarded no value). The peak reward is calculated as a proportion of the probability density in the distribution across the intersection of the sector and the annulus centred at the same origin and with edges coinciding with the boundaries of the search pattern. This represents a surrogate for the total available probability density across the time period covered by the search pattern, although it is clearly an approximation.

Once the initial state is prepared, we can plan.

4.3 Planning Searches

We use a version of POPF (Coles et al. 2010), called OPTIC (Benton, Coles, and Coles 2012), designed to perform anytime, cost-improving search. We use a time-bounded search (because we are in a time-critical situation) limited to 10 seconds. The planner will typically find a first solution very easily, since the empty plan is already a feasible solution, but it will then spend the additional time improving on this by adding further search patterns to the plan, or trying different collections of patterns. The search uses a weighted-$A^\star$ scheme with steadily changing weights in a tiered fashion (see (Benton, Coles, and Coles 2012) for details). The plans produced in this way are monotonically improving, so the final plan produced is the one we select for execution.

The plan is dispatched via a simple controller, action by action. At the conclusion of execution of the plan (including a trivial empty plan) the UAV enters a holding pattern, flying in a fixed circle. In principle, two alternatives are available at this time, depending on how long has passed since the target was last seen: spend more time generating a new plan, or abandon the search. We always presume to abandon the search at this point in our current implementation.

We use OPTIC because it is very fast at producing its first solution and provides an any-time improvement behaviour. We do not currently exploit continuous reasoning in OPTIC as we use the discrete representation of the problem. We have found that, on average, OPTIC produces around 6 plans in its 10 second window per problem instance, and the last of these is selected for execution.

5 A Search and Tracking Simulation

In order to evaluate the behaviour of our planned search approach we have developed a simulation. This was built in consultation with our industrial collaborators and is intended to provide an appropriately abstracted view of the problem. The key abstraction is of the control problem for the UAV, which we assume to be solved to provide us with the level of command we exploit in our search plans. Our dispatcher identifies waypoints and turning circles for the UAV according to the flight path and search pattern being executed, but we do not consider control of flight surfaces or altitude. We also ignore the problem of managing no-fly zones in either simulation or in planning.

The simulation defines the area of operations, which for our experiments is part of Scotland about 100 kilometres square, with Glasgow and Edinburgh approximately defining its lower corners. Terrain types were defined by hand, along with an approximate road network for the major roads and rural minor roads. Figure 2 shows the map being explored, with dark regions representing urban areas. The observer can be seen circling the target. The figure also shows part of the road network highlighted on the map. These are the roads that the observer considers the most likely ones for the target to be following, and planning problem instances will be constructed on the assumption that the areas covered by these roads are the highest reward areas for searching.

The simulation determines success in spotting and tracking the target, according to terrain, speed and discrepancy between anticipated and actual target positions. The target follows a path acquired using Google Maps, using a selected (configurable) origin and destination. This information is also used to decide what speed is appropriate for the target, based on distance between waypoints in the route proposed by Google Maps and terrain type.

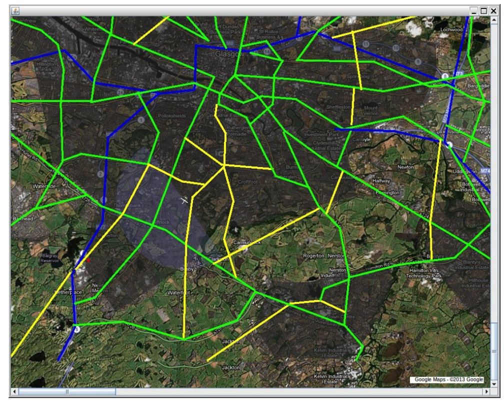

Figure 2: A screenshot showing terrain (dark regions are urban areas and grey regions are suburban) and the road network. The target (the red dot to the South-West of the observer) is undetected as indicated by the wide 180° beam search. When detected, the beam has a 90° arc.

Figure 3: A screenshot showing an initial state: rectangles and circles are search patterns that the planner will consider.

The simulation integrates the planner and displays the stages of the planning process. Figures 3 and 4 show an example of an initial state and a plan, respectively. The intensity of red used in the plan indicates repetitions of a search pattern (more intense red implies more executions).

The spiral becomes active at a future timepoint (after the start of the plan), specified in the initial state description, which has been chosen to maximise the chance of interception with the target. The duration of the action is computed from information supplied in the problem description, and the effect of the action is to increase the reward by a step function at the end of the execution. The problem description provides other possible actions also, such as `doSmallLawnMower`, which can be executed in different places and offer different rewards. The goal of the problem is to maximise reward, as this maximises the expectation of

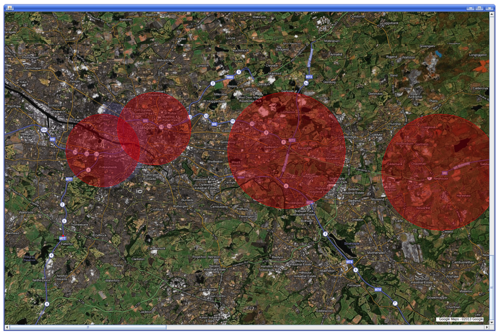

Figure 4: The search patterns that have been selected by the planner ready for execution.

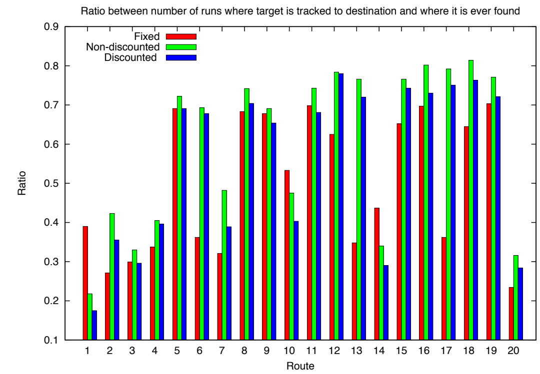

Figure 5: Proportion of successfully tracked targets.

intercepting the target.

Search patterns tend to be placed in areas that avoid the highest urban concentrations. This is because the probability of seeing the target in such areas is much lower, so the rewards associated with patterns over these areas are correspondingly low. The planner prefers to execute a search pattern on the other side of urban areas, in the direction in which it predicts the target will head. The success of this is dependent on careful timing. The estimated times of entry and exit of a target into a pattern are difficult to compute accurately and quickly and this remains an area where we hope to improve our performance.

The simulation tool offers various opportunities for interaction, including redirecting the target, repositioning the observer, speeding and slowing the simulation and modifying parameters that govern spotting probabilities, flight dynamics, target behaviour and so on. These values were chosen to broadly represent characteristics of a prototypical UAV as described to us by our industrial collaborators. One of the ways in which these parameters affect the planning process is in the choice of the radius of the search area sector described in Section 4.2. Based on the speed of the observer and the areas of the sector considered, we used a search radius of 20 kilometres as the outer edge of the possible search patterns. This is a configurable parameter, but not one we have experimented with at this stage.

6 Experiments and Results

We conducted experiments to compare plan-based search with a fixed policy selected as a baseline for evaluating the benefits of a plan-based approach. The static policy works as follows: the observer tracks the target until it has lost sight of the target. It continues to track the predicted location of the target for about three minutes. If it has not rediscovered the target, the observer then executes a fixed sequence of search patterns. It first performs a spiral search around the point where it lost the target and then executes a large lawnmower pattern over a 20 kilometre square area.

We use a configuration of our plan-based search that tracks the predicted location of the target for the same period as the static policy, before planning and executing a search plan. We compare two domain models: one using non-discounted rewards and the other discounting rewards for patterns after they are searched.

We generated 20 routes and executed the simulation on each route 1000 times (the simulation has a non-deterministic spotting model and target behaviour), for each of the 3 strategies (a total of 60000 runs). The simulation begins with the target undetected, but in the search arc of the observer. In a small number of runs the observer fails to detect the target in the very early stage. Our simulation does not use a search plan in this first stage, so failure at this point leads to an early abort. We discount these runs (less than 0.5%) in our analysis.

The metrics we use to evaluate the strategies are: the proportion of runs in which the target is tracked to its destination, the proportion of the time the target spends driving that the observer successfully tracks it (whether or not the target is detected throughout this period), and the time at which the target is finally lost on those runs in which it fails to track it to its destination.

Figure 5 shows the proportion of runs in which the target was tracked to its destination. The plan-based search strategy is consistently better, using the non-discounted model, than the other strategies. The reason it is better than the discounted model appears to be that the planner struggles to find better plans when the discounting is applied, so the solutions are, on average, shorter than the corresponding plans in the non-discounted model. In general, it is better to search, even in places where search has already been performed, than to abandon the mission. The static policy has an overall success rate of under 50%, while the non-discounted model yields better than 60% success rate. There are a few cases where the static policy appears to do better than the plan-based search and we are still investigating the reason for this. In general, on shorter routes the plan-based approach appears to perform worse, which appears to be because the search patterns are biased towards an assumption that the target is driving to a distant location.

Figure 6 shows the average time that the observer tracks the target, plotted against journey duration, for the three strategies. It can be seen that the non-discounted model produces the best performance, while the static policy is generally weaker. Figure 7 shows the average time at which the

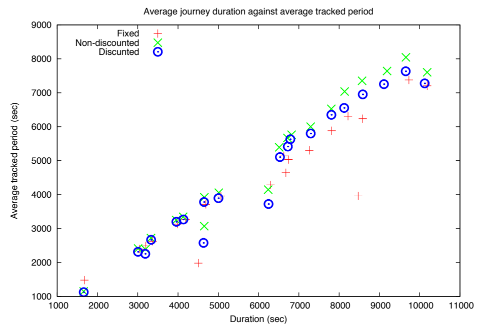

Figure 6: Average time tracked plotted against average journey length for the 20 routes and three strategies.

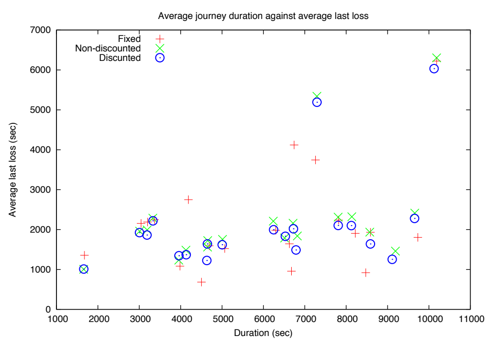

Figure 7: Average time of last loss of target plotted against average journey length for the 20 routes and three strategies.

target is lost for the last time, plotted against duration of the journey. This figure shows an interesting pattern, which is that the target is lost, on average, at about the same time in the journey (after about half-an-hour), regardless of the duration of the journey. We conjecture that this is because these journeys all start in an urban area that it takes approximately 30 minutes to cross. During this period the target has a constrained heading (due to the road network across town) and, on leaving, it often corrects its heading by a significant deviation. This indicates that there remains considerable scope for improvement in the estimation of reward distributions over the patterns. However, it is also the case that the observer typically must commit to an area of search, since once it has searched a large area, there is little chance to try searching widely diverging alternatives.

We also analysed the data to find the probability distribution, over time, of relocating the target after losing it, considering only the cases in which it was successfully relocated. Figure 8 shows how this probability changes over time. This clearly shows that the planned search is far more robust, offering significant probability of rediscovery even after half-an-hour. The static policy only finds the target after it has been lost for 10 minutes in less than 4% of cases. The figure shows the probability of finding the target at a particular time after losing it. A slightly different question is how likely it is that the static policy will find the target *once it has entered the lawnmower search*. We found that to

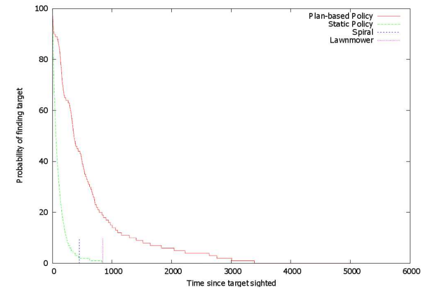

Figure 8: The probability of recapturing the target over time.

be approximately 13%, which indicates that the lawnmower is not as ineffective as Figure 8 might suggest. We found that the plan-based search very rarely finds the target after the fourth search pattern in a plan.

Our results clearly demonstrate a better performance using planning than using the static policy, but there remains scope for improvement. There appears to be weakness in our efforts to position the time windows on search patterns and we are investigating ways to improve this first.

7 Conclusions and Future Work

We have presented a planning approach to SAT, viewing the search problem as a planning problem in which search patterns must be selected and sequenced to maximise the expectation of rediscovering the target. Our results indicate that the approach is promising and certainly outperforms static search strategies. By using this approach we have been able to tackle SAT problems on a large scale — a 100 kilometre square area represents a significant challenge to the problem of search, far beyond the capabilities of current alternatives.

Things that we have yet to consider include the effects of no-fly zones that prevent the UAV from tracking the target, but that the target might enter knowingly or unknowingly, to evade pursuit. Our reward estimation techniques can be improved to take a better account of the underlying road network and also to formulate and exploit hypotheses about the destination of the target. We have found that the model that attempts to discount reward, to account for the fact that failure to find a target in a search pattern should change the conditional probability of finding the target in a future search in the same area, remains a challenge to the planning technology. Of course, there is a possibility to construct more specialised planning for this problem, but an important benefit of the use of a generic planner is that we can readily modify the collection of search actions, add alternative actions and otherwise extend the domain model. This flexibility is particularly important during prototyping and development.

One of the advantages we see in the use of a plan-based approach to the search problem is that the behaviour of the UAV is predictable and well understood. A plan can be used as a common medium of exchange between the UAV and human observers, allowing safer interaction between the UAV and other air traffic.

References

Benton, J.; Coles, A.; and Coles, A. 2012. Temporal Planning with Preferences and Time-Dependent Continuous Costs. In *Proceedings of the Twenty Second International Conference on Automated Planning and Scheduling (ICAPS-12)*.

Bourgault, F.; Furukawa, T.; and Durrant-Whyte, H. F. 2004. Process Model, Constraints, and the Coordinated Search Strategy. In *Proceedings of the 2004 IEEE International Conference on Robotics and Automation (ICRA 2004)*, 5256–5261.

Bourgault, F.; Furukawa, T.; and Durrant-Whyte, H. F. 2006. Optimal Search for a Lost Target in a Bayesian World. In *Field and Service Robotics*, volume 24 of *Springer Tracts in Advanced Robotics*. Springer Berlin. 209–222.

Chung, C. F., and Furukawa, T. 2006. Coordinated Search-and-Capture Using Particle Filters. In *Proceedings of the 9th International Conference on Control, Automation, Robotics and Vision (ICARCV '06)*, 1–6.

Coles, A. J.; Coles, A. I.; Fox, M.; and Long, D. 2010. Forward-Chaining Partial-Order Planning. In *Proceedings of the 20th International Conference on Automated Planning and Scheduling (ICAPS-10)*.

Doherty, P.; Kvarnström, J.; and Heintz, F. 2009. A temporal logic-based planning and execution monitoring framework for unmanned aircraft systems. *Autonomous Agents and Multi-Agent Systems* 19:332–377.

Furukawa, T.; Bourgault, F.; Lavis, B.; and Durrant-Whyte, H. F. 2006. Recursive Bayesian Search-and-tracking using Coordinated UAVs for Lost Targets. In *Proceedings of the 2006 IEEE International Conference on Robotics and Automation (ICRA 2006)*, 2521–2526.

Furukawa, T.; Mak, L. C.; Durrant-Whyte, H. F.; and Madhavan, R. 2012. Autonomous Bayesian Search and Tracking, and its Experimental Validation. *Advanced Robotics* 26(5-6):461–485.

Furukawa, T.; Durrant-Whyte, H. F.; and Lavis, B. 2007. The element-based method — theory and its application to Bayesian search and tracking. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007)*, 2807–2812.

Kantor, M. G., and Rosenwein, M. B. 1992. The Orienteering Problem with Time Windows. *Journal of the Operational Research Society* 43(6):pp. 629–635.

Lavis, B., and Furukawa, T. 2008. HyPE: Hybrid Particle-Element Approach for Recursive Bayesian Searching and Tracking. In *Proceedings of the 2008 Robotics: Science and Systems Conference*, 135–142.

Léauté, T., and Williams, B. C. 2005. Coordinating Agile Systems through the Model-based Execution of Temporal Plans. In *Proceedings of the 2005 AAAI Conference on Artificial Intelligence*, 114–120.

Long, D., and Fox, M. 2003. The 3rd International Planning Competition: Results and Analysis. *Journal of Artificial Intelligence Research (JAIR)* 20:1–59.

Mathews, G., and Durrant-Whyte, H. 2007. Decentralised Optimal Control for Reconnaissance. In *Proceedings of the 2007 Conference on Information, Decision and Control (IDC '07)*, 314–319.

Mathews, G.; Waldock, A.; and Paraskevaides, M. 2010. Toward a Decentralised Sensor Management System for Target Acquisition and Track. In *Proceedings of the Systems Engineering for Autonomous Systems (SEAS) Defence Technology Centre (DTC) Technical Conference*.

ASTRAEA Programme. 2012. *www.astraea.aero – Accessed February 21, 2013*.

Smith, D. E. 2004. Choosing Objectives in Over-Subscription Planning. In *Proceedings of the 14th International Conference on Automated Planning & Scheduling (ICAPS 2004)*.