UAV ROUTING IN A STOCHASTIC, TIME-VARYING ENVIRONMENT

John J. Enright and Emilio Frazzoli¹

Mechanical and Aerospace Engineering University of California, Los Angeles Los Angeles, CA 90095 {jenright, frazzoli}@ucla.edu

Abstract: In this paper we consider the following problem. An Uninhabited Aerial Vehicle (UAV), modeled as a vehicle moving at unit speed along paths of bounded curvature, must visit stochastically-generated targets in a convex, compact region of the plane. Targets are generated according to a spatio-temporal Poisson process, uniformly in the region. It is desired to minimize the expected waiting time between the appearance of a target, and the time it is visited. We present algorithms for UAV routing, and compare their performance with respect to asymptotic performance bounds, in the light and heavy load limits. Simulation results are presented and discussed. *Copyright* $\bigcirc 2005$ IFAC

Keywords: Algorithms; Aircraft control; Autonomous vehicles; Mobile robots.

1. INTRODUCTION

One of the prototypical missions for Uninhabited Aerial Vehicles, e.g., in environmental monitoring, security, or military setting, is wide-area surveillance. A UAV in such a mission must provide coverage of a certain region and detect, locate, and investigate events of interest ("targets") as they manifest themselves. In particular, we are interested in cases in which close-range information is required on the targets, i.e., cases in which the UAV must proceed to the location of the targets to gather on-site information.

Variations of problems falling in this class have been studied in a number of papers in the recent past, e.g., (Schumacher *et al.* 2003, Beard *et al.* 2002, Richards et al. 2002, Gil et al. 2003). In these papers, the problem is set up in such a way that the location of targets is known a priori; a strategy is computed that attempts to optimize the cost of servicing the known targets. In (Li and Cassandras 2003) a stable receding-horizon strategy is proposed, but its performance in not characterized. In (Frazzoli and Bullo 2004), we addressed the case in which new targets are generated continuously by a stochastic process: we provided algorithms for minimizing the expected waiting time between the appearance of a target and the time it is serviced by one of the vehicles. A limitation of the results presented in (Frazzoli and Bullo 2004) is the fact that omni-directional vehicles were considered in the problem formulation: as such, the results are not applicable to many vehicles of interest, including aircraft and car-like robots.

In this paper, we wish to extend the results available in the literature to address non-holonomic vehicle dynamics. In particular, we will consider

¹ The research leading to this document was partially supported by AFOSR, through grant F49620-02-1-0325. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of the supporting organization.

paths with bounded curvature, which provide a good approximation of feasible trajectories for aircraft. The main contributions of the paper are: (i) the design of a UAV control policy achieving a level of performance in light load that is provably within a constant additive factor from optimality, and (ii) the establishment of a new lower bound for the achievable performance in heavy load. Finally, we present simulation results that suggest that a greedy policy can approximate the optimal policy in heavy load.

2. PROBLEM FORMULATION

The basic version of the problem we wish to study in this paper is known as the Dynamic Traveling Repairperson Problem (DTRP), and was introduced by Bertsimas and van Ryzin in (Bertsimas and van Ryzin 1991). Our problem is different from the original DTRP since we consider a vehicle that is constrained to move at unit speed along paths of bounded curvature, i.e., we impose a non-holonomic constraint on the vehicle's dynamics. In the remainder of the section, we define the details of the problem and its components.

Let the environment $\mathcal{Q} \subset \mathbb{R}^2$ be a convex, compact set with unit area, and let $\|\cdot\|$ denote the Euclidean norm in \mathbb{R}^2 . Consider a single UAV, modeled as a nonholonomic vehicle constrained to move at unit speed along a path with bounded curvature, and let $1/\rho$ be the maximum curvature. In other words, let the configuration $g \in SE(2)$ of the vehicle be given in coordinates by $g = (x, y, \theta)$, where x, y are respectively the projections of the vehicle's position along fixed orthogonal axes, and θ is the orientation of the vehicle's longitudinal axis with respect to the y = 0 axis; then the dynamics of the vehicle are described by the differential equations

$$\begin{aligned} \dot{x} &= \cos(\theta), \\ \dot{y} &= \sin(\theta), \\ \dot{\theta} &= \omega, \quad \omega \in [-1/\rho, 1/\rho]. \end{aligned}$$
(1)

The vehicle has unlimited range and targetserving capacity. In the following, we will indicate by p = (x, y) the position of the vehicle. We note that the above kinematic model of an airplane's dynamics is very common in the literature on UAV motion planning; the model is very similar to the one studied in (Dubins 1957), with the difference that the vehicle we consider is constrained to move at constant speed. Results in terms of minimum-length paths for Dubins' vehicle hold for our model, where they assume an additional connotation of being minimum-time paths as well. In the following, we will often use the expressions "Dubins vehicle" and "Dubins paths" to indicate a vehicle modeled by (1) and paths that are feasible with respect to the same model.

Information on outstanding targets—the demand at time t is summarized as a finite set of target positions $D(t) \subset Q$, with $n(t) := \operatorname{card}(D(t))$. Targets are generated, and inserted into D, according to a homogeneous (i.e., time-invariant) spatio-temporal Poisson process, with time intensity $\lambda > 0$, and uniform spatial density. In other words, given a set $S \subseteq Q$, the expected number of targets generated in S within the time interval [t, t'] is

$$E \left[\operatorname{card}(D(t') \cap S) - \operatorname{card}(D(t) \cap S) \right] =$$

= $\lambda(t' - t) \operatorname{Area}(S).$

Servicing of a target $e_j \in D$, and its removal from the set D, is achieved when the UAV moves to the target's position.

A static feedback control policy for the system is a map $\pi : SE(2) \times 2^{\mathcal{Q}} \rightarrow [-1/\rho, 1/\rho]$, assigning a control input to the vehicle, as a function of the current state of the system, i.e., $\omega(t) = \pi(g(t), D(t))$. The policy π is stable if, under its action,

$$n_{\pi} := \lim_{t \to +\infty} \operatorname{E}\left[n(t) | \dot{p} = \pi(p, D)\right] < +\infty,$$

that is, if the UAV is able to service targets at a rate that is—on average—at least as fast as the rate at which new targets are generated.

Let T_j be the time that the *j*-th target spends within the set D, i.e., the time elapsed from the time e_j is generated to the time it is serviced. If the system is stable, then we can write the balance equation (known as Little's formula (Larson and Odoni 1981))

$$n_{\pi} = \lambda T_{\pi}$$

where $T_{\pi} := \lim_{j \to +\infty} \operatorname{E}[T_j]$ is the steady-state system time under the policy π . Our objective is to minimize the steady-state system time, over all possible static feedback control policies, i.e.,

$$T^* = \inf T_{\pi}.$$

In the following, we are interested in designing control policies that provide constant-factor approximations of the optimal achievable performance. In particular, as done in (Bertsimas and van Ryzin 1991), we will analyze the asymptotic cases of light load, as $\lambda \to 0$, and heavy load, as $\lambda \to \infty$; in the light load case, we will derive lower and upper bounds on the achievable system time T^* . The upper bound is given in terms of system times achieved by an explicit, readily implementable algorithm, and is therefore constructive. In the heavy load case, we present a novel lower bound for the system time, and show simulation results showing that a greedy policy provides a good approximation to the achievable performance.

2.1 Some preliminary results

Before we address the design of control policies, let us state a few preliminary results on the length of minimum-length Dubins paths, which we will use in the following. Note that a full characterization of optimal paths is given in (Dubins 1957); a further classification is given in (Shkel and Lumelsky 2001); our purpose in this section is to provide bounds on the length of optimal paths given the boundary conditions.

Let us consider first the problem of point-to-point steering with the heading angle fixed both at the initial and final condition. Let $d : SE(2) \times$ $SE(2) \rightarrow \mathbb{R}_+$ be a function returning the Euclidean distance between vehicles in two given configurations, i.e.,

$$d: (x_1, y_1, \theta_1, x_2, y_2, \theta_2) \mapsto \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

Define $L_{\rho}(g) : SE(2) \times SE(2) \to \mathbb{R}_+$ as the function returning the cost of the minimum-length Dubins path between two configurations.

Proposition 1. For any $g_1, g_2 \in SE(2)$, the minimum length $L_{\rho}(g_1, g_2)$ of a path steering a Dubins' vehicle with maximum curvature $1/\rho$ from g_1 to g_2 satisfies:

$$d(g_1, g_2) \le L_{\rho}(g_1, g_2) \le d(g_1, g_2) + \kappa \pi \rho,$$
 (2)
with $7/3 \le \kappa \le 2.658.$

PROOF. The lower bound is trivial, as it is the Euclidean distance between the two configurations. The proof of the upper bound can be found in (?) and is not reported here for lack of space. \Box

Now consider a slightly different problem. Instead of requiring the vehicle to have a prescribed heading at the final condition, we will require it only to move to a certain position, leaving the final heading unconstrained. Without loss of generality, choose the initial configuration as the identity of SE(2), i.e., $g_1 = e = (0,0,0)$. We have the following result, that can be verified numerically. (We omit the details for lack of space.)

Proposition 2. The minimum length $P_{\rho}(x, y)$ of a path steering a Dubins' vehicle with maximum curvature ρ from the identity in SE(2) to a point $(x, y) \in \mathbb{R}^2$ satisfies the following inequality:

$$P_{\rho}(x,y) \le c_1 \rho + \|(x,y) - (0,c_2\rho)\|, \qquad (3)$$

with $c_1 \approx 3.77$, and $c_2 \approx 2.93$.

3. THE LIGHT LOAD LIMIT

We begin our analysis by looking at the lightload asymptotic case, i.e., at the system times achievable as $\lambda \to 0$.

3.1 Lower bound on the system time

In order to study optimal policies in the light load case, we need to introduce a problem from geometric optimization.

Given a set $\mathcal{Q} \subset \mathbb{R}^d$ and a point $p \in \mathcal{Q}$, the expected distance between a random point q, generated according to a uniform distribution over \mathcal{Q} , and p is given by

$$H(p,\mathcal{Q}) := \operatorname{E}\left[\|p-q\|\right] = \int_{\mathcal{Q}} \|p-q\| dq. \quad (4)$$

The function H is known as the continuous Weber function or the continuous median function; see (Agarwal and Sharir 1998, Drezner 1995) and references therein. The median of the set Q is the global minimizer

$$p^*(\mathcal{Q}) = \operatorname*{argmin}_{p \in \mathcal{Q}} H(p, \mathcal{Q}).$$

It is straightforward to show that the map (4) is differentiable and strictly convex on Q. Therefore, computing the global minimizer is a simple task.

The following lower bound was derived in (Bertsimas and van Ryzin 1991) for the Euclidean case; since the length of a Dubins path between two configurations is no less than the Euclidean distance between the points, the same results holds in our case.

Theorem 3. The system time for the problem stated in Section 2 is lower bounded as:

$$T^* \ge H(p^*, \mathcal{Q}) \tag{5}$$

Theorem 3 holds for any policy, and any value of λ . However, it is most useful in the light-load case and as such it is reported in this section.

3.2 A constructive upper bound

Consider the following control policy, which we call the OFFSET MEDIAN (OM) policy. Let p^* be the median of Q, and define the *loitering station* for the UAV as a circular trajectory of radius $c_2\rho$ centered at p^* . In the OM policy, the UAV visits all targets in a greedy fashion: in other words, it always pursues the closest target in a Dubins' distance sense. When no targets are available, it returns to its loitering station; the direction in which the orbit is followed is inconsequential, and can be chosen in such a way that the station is reached in minimum time.

Theorem 4. An upper bound on the system time of the Offset Median policy in light load is

$$T_{\rm OM} \le T^* + c_1 \rho \text{ as } \lambda \to 0$$
 (6)

PROOF. Consider a generic initial condition for the UAV configuration in Q and for the outstanding target positions D(0), with $n_0 = \operatorname{card}(D(0))$. An upper bound to the time needed to service all of the initial targets is $n_0(\operatorname{diam}(Q) + \kappa \pi \rho)$. When there are no targets outstanding in the target set D, the vehicle moves at unit speed toward its loitering station, which is reached in at most $\operatorname{diam}(Q) + \kappa \pi \rho$ units of time.

The time needed to service the initial targets and go to the median is hence bounded by $t_{\text{ini}} \leq (n_0 + 1) [\operatorname{diam}(\mathcal{Q}) + \kappa \pi \rho]$. The probability that at the end of this initial phase the number of targets is reduced to zero is

$$P[n(t_{\text{ini}}) = 0] = \exp(-\lambda t_{\text{ini}})$$

$$\geq \exp(-\lambda(n_0 + 1)(\operatorname{diam}(\mathcal{Q}) + \kappa \pi \rho))$$

that is, $P[n(t_{\text{ini}}) = 0] \rightarrow 1^{-}$ as $\lambda \rightarrow 0$. As a consequence, after an initial transient, all targets will be generated with the vehicle in its loitering station, and an empty demand queue.

After the initial transient, when the next target arises, say the *j*th target at location e_j , will then require $T_j \leq ||e_j - p^*|| + c_1\rho$. The system time can be computed as

$$T_{\rm OM} = \lim_{j \to +\infty} \mathbb{E}\left[T_j\right] = H^*(\mathcal{Q}) + c_1 \rho. \qquad \Box$$

In other words, we have shown that the system time achieved by the OM policy is within a constant additive factor from the optimal. The additive factor, which can be considered as a penalty due to the non-holonomic constraints imposed on the vehicle's dynamics, depends linearly on the minimum turn radius ρ . At this time, neither the lower bound nor the upper bound are known to be tight.

In Figure 1 we show simulation results that confirm our theoretical predictions. When the minimum turn radius is very small, the performance of the OM policy approximates the lower bound valid for a vehicle without kinematics constraints, i.e., as $\rho \to 0$, $T_{\text{OM}} \to H^*(\mathcal{Q})$. As ρ increases, the penalty associated to the bounded curvature constraints dominates, and the system time increases linearly with ρ .

4. THE HEAVY LOAD LIMIT

In this section, we turn our attention to the heavyload limit, in which $\lambda \to \infty$.



Fig. 1. Performance of the OM policy in light load: the system time grows at most linearly with ρ , as suggested by the upper bound (6).

4.1 Lower bound on the system time

In the heavy load case, the nature of optimal control policies is related to the well-known Traveling Salesperson Problem (TSP). We will first discuss some well-known results for the Euclidean version of the TSP, then derive a lower bound on the asymptotic cost of TSP problems for boundedcurvature vehicles. Based on this result, we will provide a lower bound on the system time in the heavy load limit.

4.2 The Euclidean Traveling Salesperson Problem

The Euclidean TSP (ETSP) is formulated as follows: given a set D of n points in \mathbb{R}^d , find the minimum-length tour of D. Let $\mathrm{ETSP}(D)$ denote the minimum length of a tour through all the points in D; by convention, $\mathrm{ETSP}(\emptyset) = 0$. The asymptotic behavior of stochastic ETSP problems for large n exhibits the following interesting property. Assume that the locations of the n targets are independent random variables, uniformly distributed in a compact set \mathcal{Q} ; in (Beardwood *et al.* 1959) it is shown that there exists a constant $\beta_{\mathrm{TSP},2}$ such that, almost surely,

$$\lim_{n \to +\infty} \frac{\mathrm{TSP}(D)}{\sqrt{n}} = \beta_{\mathrm{TSP},2}.$$
 (7)

In other words, the optimal cost of stochastic ETSP tours approaches a deterministic limit, and grows as the square root of the number of points in D; the current best estimate of the constant in (7) is $\beta_{\text{TSP},2} = 0.7120 \pm 0.0002$, see (Percus and Martin 1996, Johnson *et al.* 1996).

5. THE TRAVELING SALESPERSON PROBLEM FOR A DUBINS VEHICLE

While the ETSP has attracted a great deal of interest from the scientific community, its boundedcurvature counterpart (which we will call DTSP) has not been studied extensively. Some initial work has been done in (Savla *et al.* 2005a) mainly in terms of upper bounds for worst-case tours. Here we present a new result, which can be seen as a first step in the search of deterministic bounds similar to those available for the ETSP.

Theorem 5. The expected cost of a stochastic DTSP visiting a set D of n randomly-generated points in Q satisfies the following inequality:

$$\lim_{n \to \infty} \frac{\mathrm{E}\left[\mathrm{DTSP}(D, \rho)\right]}{n^{2/3}} \ge \frac{3}{4} (3\rho)^{1/3} \qquad (8)$$

PROOF. Choose a random point $p_i \in D$ as the initial position of the vehicle on the tour, and choose the heading randomly. We would like to compute a bound on the expected distance, according to the metric induced by the length of Dubins' paths, to the closest next point in the tour; let us call such distance δ^* .

To this purpose, consider the set R_{δ} of points that are reachable from a Dubins' vehicle with an arc of length $\delta \leq \rho$; the area of such a set is

$$\operatorname{Area}(R_{\delta}) = \frac{\delta^3}{3\rho}.$$
 (9)

In other words, the area of the set R_{δ} decreases faster than the area of a circle of radius δ as $\delta \to 0$.

Given a distance δ , the probability that $\delta^* > \delta$ is no less than the probability that there is no other target reachable with a path of length at most δ ; in other words,

$$\Pr[\delta^* > \delta] \ge 1 - n\operatorname{Area}(R_{\delta}) = 1 - n\frac{\delta^3}{3\rho}$$

In terms of expectation, defining $c = n/(3\rho)$,

$$\begin{split} \mathbf{E}\left[\delta^*\right] &= \int_0^\infty \Pr(\delta^* > \xi) \ d\xi \\ &\geq \int_0^\infty \max\{0, 1 - n\frac{\xi^3}{3\rho}\} \ d\xi \\ &= \int_0^{c^{-1/3}} (1 - c\xi^3) \ d\xi \\ &= \frac{3}{4} \left(\frac{3\rho}{n}\right)^{1/3} \end{split}$$

The expected total tour length will be no smaller than n times the expected length of the shortest path between two points, i.e.,

$$E[DTSP(D, \rho)] \ge \frac{3}{4} (3\rho n^2)^{1/3} - o(n^{2/3}).$$

Dividing both sides by $n^{2/3}$ and taking the limit as $n \to \infty$, we get the desired result. \Box

Note that the dependency of the optimal DTSP cost on the number of points in the tour is with the power 2/3; in the Euclidean case the dependency was on the power 1/2.

5.1 Lower bound on the system time

At this point we can state the desired result in terms of a lower bound on the system time for any policy in the heavy load case.

Theorem 6. The system time for the problem stated in Section 2, satisfies the following inequality, for $\lambda \to \infty$:

$$T^* \ge \frac{81}{64}\rho\lambda^2 \tag{10}$$

PROOF. Let us assume that a stabilizing policy is available. In such a case, the number of outstanding targets approaches a finite steadystate value, n^* , related to the system time by Little's formula, i.e., $n^* = \lambda T^*$. In order for the policy to be stabilizing, the average time needed to visit the next target must be no greater than the average time interval between the appearance of new targets, i.e., $\delta^* \leq 1/\lambda$. But we know that $\delta^* \geq 3/4(3\rho/n^*)^{1/3}$; rearranging, and using Little's formula, we get the desired result. \Box

Note that the system time depends quadratically on the parameter λ , whereas in the Euclidean case it depends only linearly on it. As a consequence, bounded-curvature constraints make the system time much more sensitive to increases in the target generation rate.

5.2 Towards an upper bound on the system time

A tight upper bound on the DTRP for a Dubins' vehicle is not yet available, as the DTSP problem is still largely unexplored. Very recently, a new algorithm for the stochastic DTSP was discovered (Savla *et al.* 2005*b*), which provides a tour with cost $O(n^{2/3}(\log n)^{1/3})$. While this cannot provide a tight bound, it guarantees a sub-linear increase of the cost with the number of targets, and would ensure stability of the DTRP for a Dubins' vehicle. The investigation of this algorithm in the context of the DTRP problem is the subject of current work.

In the meantime, we present simulation results (Figure 2) that suggest that a simple greedy policy does stabilize the system for any λ , and provides a (multiplicative) constant-factor approximation of the optimal system time. In the greedy policy, the vehicle always moves towards the closest (according to the Dubins' distance) outstanding target.

6. CONCLUSIONS

In this paper, we have considered the problem of steering a UAV in order to minimize the expected waiting time between the appearance of



Fig. 2. Performance of the greedy policy in heavy load with $\rho = 0.05$: for fixed ρ , the system time grows quadratically with λ , as suggested by the lower bound (10).

randomly-generated targets and the time they are visited by the UAV. We have proposed a control policy that achieves a system time that is provably within a constant additive factor from the optimal, in the light load case. In the heavy load case, we have developed a lower bound on the system time, showing that it depends at least quadratically on the target generation rate, and linearly on the minimum turn radius. Future work will aim at achieving tighter bounds on the achievable performance in both cases, and in extending the present work to the multiple-vehicle, decentralized control case.

ACKNOWLEDGMENTS

The authors are grateful to Prof. Francesco Bullo and Mr. Ketan Savla of the University of California, Santa Barbara, for numerous, stimulating, and insightful discussions on the subject of this paper.

REFERENCES

- Agarwal, P. K. and M. Sharir (1998). Efficient algorithms for geometric optimization. ACM Computing Surveys 30(4), 412–458.
- Beard, R. W., T. W. McLain, M. A. Goodrich and E. P. Anderson (2002). Coordinated target assignment and intercept for unmanned air vehicles. 18(6), 911–922.
- Beardwood, J., J. Halton and J. Hammersley (1959). The shortest path through many points. Proceedings of the Cambridge Philoshopy Society 55, 299–327.
- Bertsimas, D. J. and G. J. van Ryzin (1991). A stochastic and dynamic vehicle routing problem in the Euclidean plane. *Operations Research* **39**, 601–615.

- Drezner, Z., Ed.) (1995). Facility Location: A Survey of Applications and Methods. Springer Series in Operations Research. Springer Verlag. New York.
- Dubins, L.E. (1957). On curves of minimal length with a constraint on average curvature and with prescribed initial and terminal positions and tangents. *American Journal of Mathematics* **79**, 497–516.
- Frazzoli, E. and F. Bullo (2004). Decentralized algorithms for vehicle routing in a stochastic time-varying environment. In: *Proc. IEEE Conf. on Decision and Control.* Paradise Island, Bahamas.
- Gil, A. E., K. M. Passino and A. Sparks (2003). Cooperative scheduling of tasks for networked uninhabited autonomous vehicles. In: *Proc. IEEE Conf. on Decision and Control.* Maui, Hawaii. pp. 522–527.
- Johnson, D. S., L. A. McGeoch and E. E. Rothberg (1996). Asymptotic experimental analysis for the held-karp traveling salesman bound. In: Proc. 7th Annual ACM-SIAM Symposium on Discrete Algorithms. pp. 341– 350.
- Larson, R. C. and A. R. Odoni (1981). Urban Operations Research. Prentice-Hall. Englewood Cliffs, NJ.
- Li, W. and C. G. Cassandras (2003). Stability properties of a cooperative receding horizon controller. In: *Proc. IEEE Conf. on Decision* and Control. Maui, Hawaii. pp. 492–497.
- Percus, G. and O. C. Martin (1996). Finite size and dimensional dependence of the Euclidean traveling salesman problem. *Physical Review Letters* 76(8), 1188–1191.
- Richards, A., J. Bellingham, M. Tillerson and J. How (2002). Coordination and control of multiple UAVs. In: Proc. of the AIAA Conf. on Guidance, Navigation, and Control. Monterey, CA.
- Savla, K., E. Frazzoli and F. Bullo (2005a). On the point-to-point and traveling salesperson problems for Dubins' vehicles. In: Proc. of the American Control Conference. To appear.
- Savla, K., F. Bullo and E. Frazzoli (2005b). The stochastic travelling salesperson problem for Dubins' vehicle. Submitted to the IEEE Conf. on Decision and Control.
- Schumacher, C., P. R. Chandler, S. J. Rasmussen and D. Walker (2003). Task allocation for wide area search munitions with variable path length. In: *Proc. of the American Control Conference.* Denver, CO. pp. 3472–3477.
- Shkel, A. M. and V. J. Lumelsky (2001). Classification of the Dubins set. *Robotics and Au*tonomous Systems 23, 179–202.