43rd IEEE Conference on Decision and Control
December 14-17, 2004
Atlantis, Paradise Island, Bahamas

Decentralized Algorithms for Vehicle Routing in a Stochastic Time-Varying Environment

Emilio Frazzoli
Mechanical and Aerospace Engineering
University of California at Los Angeles
Los Angeles, CA 90095, USA
frazzoli@ucla.edu

Francesco Bullo
Mechanical and Environmental Engineering
University of California at Santa Barbara
Santa Barbara, CA 93106, USA
bullo@engineering.ucsb.edu

Abstract— **In this paper we present decentralized algorithms for motion coordination of a group of autonomous vehicles, aimed at minimizing the expected waiting time to service stochastically-generated targets. The vehicles move within a convex environment with bounded velocity, and target generation is modeled by a spatio-temporal Poisson process. The general problem is known as the m-vehicle Dynamic Traveling Repairperson Problem (m-DTRP); the best previously known control algorithms rely on centralized a-priori task assignment and locational optimization, and are of limited applicability in scenarios involving ad-hoc networks of autonomous vehicles. In this paper, we present a new class of algorithms for the m-DTRP problem that: (i) are spatially distributed, scalable to large networks, and adaptive to network changes, (ii) are provably locally optimal in the light load case, and (iii) achieve the same performance as the best known centralized algorithms in the heavy-load, single-vehicle case. Simulation results are presented and discussed.**

I. Introduction

Advances in computer technology, wireless communications, and miniaturization of electromechanical systems, coupled with new perceived critical needs of our society, motivate the rapidly increasing interest in the design and deployment of large networks of mobile devices capable of sensing spatially distributed phenomena, and/or of interacting directly with the environment [1]. As the size, complexity, and pervasiveness of such networks increase, the emphasis is shifting from operator-mediated or operator-initiated actions to completely autonomous operations, in which the network interacts directly, with minimal or no human supervision, with the physical environment.

For example, sensing abilities of the most disparate forms are at the core of the rapidly growing interest in sensor networks [2]; in many cases of interest, effective usage of sensors is enabled by mobile platforms, physically carrying sensors to the vicinity of events of interest. In a prototypical mission in a military or security setting, teams of Unmanned Aerial Vehicles (UAVs) can be used for wide-area surveillance, by detecting, locating, and identifying assets or threats in a region of interest. Similar consideration can be made for networks of ground or underwater vehicles.

In a surveillance mission, the UAVs must ensure continued coverage of a certain area; as events occur, i.e., as new targets are detected by on-board sensors or other assets (e.g., intelligence, high-altitude or orbiting platforms, etc.), UAVs must proceed to the location of the new event and provide close-range information about the target. Variations of problems falling in this class have been studied in a number of papers in the recent past, e.g., [3], [4], [5], [6], [7]. In these papers, the problem is set up in such a way that the location of targets is fixed and known a priori; a strategy is computed that attempts to optimize the cost of servicing the known targets. In the present work, we wish to address the case in which new targets are generated continuously by a stochastic process: we will provide algorithms for minimizing the expected waiting time between the appearance of a target and the time it is serviced by one of the vehicles.

In this paper, we present a new class of algorithms for the m-DTRP problem that: (i) are spatially distributed, scalable to large networks, and adaptive to network changes, (ii) are provably locally optimal in the light load case, and (iii) achieve the same performance as the best known centralized algorithms in the heavy-load, single-vehicle case. Here, by network changes we mean changes in the number of vehicles, in the environment boundaries, and in the characterization of the target generation process. Our receding horizon control policies combine algorithms for the Euclidean Traveling Salesperson Problem and for the continuous multi-median problem. We establish asymptotic performance results for our policies in the light load and heavy load regimes. Simulation results are presented and discussed.

The paper is structured as follows. In Section II we introduce some notation and formulate the problem we wish to address. In Section III we provide some background on the Euclidean Traveling Salesman Problem and on the continuous multi-median problem. In Section IV we consider the vehicle routing problem in stochastic, time-varying environments: we review known policies and we design novel decentralized ones. In Section V we present results from numerical experiments, and finally, in Section VI we draw some conclusions and discuss some directions for future work.

II. Notation and 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 [8]. The m-vehicle version of the problem, m-DTRP, was first studied by the same authors in [9]. In this section, we define the problem and its components.

Let the environment $\mathcal{Q} \subset \mathbb{R}^d$ be a convex, compact set with unit volume, and let $\|\cdot\|$ denote the Euclidean norm in $\mathbb{R}^d$. For simplicity, in this paper we will consider mainly the planar case, i.e., $d = 2$, with the understanding that extensions to higher dimensions are possible.

Consider m omnidirectional vehicles, modeled as point masses, and let

$$p(t) = (p_1(t), \ldots, p_m(t)) \in \mathcal{Q}^m$$

describe the locations of the vehicles at time t. The vehicles are free to move, with bounded velocity, within the environment $\mathcal{Q}$; without loss of generality, we will assume that the maximum velocity magnitude is unitary, i.e., $\|\dot{p}_i(t)\| \leq 1$, $i \in \{1, \ldots, m\}$, where the dot represents differentiation with respect to time. The vehicles are identical, and have unlimited fuel and target-servicing capacity.

Information on outstanding targets—the demand—at time t is summarized as a finite set of target positions $D(t) \subset \mathcal{Q}$, with $n(t) := \mathrm{card}(D(t))$. Targets are generated, and inserted into D, according to a homogeneous (i.e., time-invariant) spatio-temporal Poisson process, with time intensity λ, and spatial density $\varphi : \mathcal{Q} \to \mathbb{R}_+$. In other words, given a set $\mathcal{S} \subseteq \mathcal{Q}$, the expected number of targets generated in $\mathcal{S}$ within the time interval $[t, t + \Delta t]$ is

$$\mathrm{E}\left[\mathrm{card}(D(t + \Delta t) \cap \mathcal{S}) - \mathrm{card}(D(t) \cap \mathcal{S})\right] = \varphi(\mathcal{S})\lambda\Delta t,$$

where $\varphi(\mathcal{S}) := \int_{\mathcal{S}} \varphi(q) \ dq$. The spatial density φ is assumed normalized so that $\varphi(\mathcal{Q}) = 1$.

Servicing of a target $e_j \in D$, and its removal from the set D, is achieved when one of the vehicles moves to the target location and spends an additional on-site servicing time $s_j \geq 0$; the on-site servicing times are independently and identically distributed, with $\mathrm{E}[s_i] = \bar{s}$, $\mathrm{E}\left[s_i^2\right] = \bar{s^2}$.

A static feedback control policy for the system is a map $\pi : \mathcal{Q}^m \times 2^{\mathcal{Q}} \to \mathbb{R}^{d \times m}$, assigning a commanded velocity to each of the m vehicles, as a function the current state of the system: $\dot{p}(t) = \pi(p(t), D(t))$. The policy π is stable if, under its action,

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

that is, if the vehicles are able to service targets at a rate that is—on average—at least as fast as the rate at which new targets are generated. For a stable system, the product $\rho = \lambda\bar{s}$ represents the average time spent on on-site servicing.

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 [10])

$$n_\pi = \lambda T_\pi,$$

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

$$T^* = \inf_\pi T_\pi.$$

In the following, we are interested in designing control policies that provide constant-factor approximations of the optimal achievable performance; a policy π is said to provide a constant-factor approximation of κ if $T_\pi \leq \kappa T^*$. Moreover, we are interested in decentralized, scalable, adaptive control policies, that rely only on local exchange of information between neighboring vehicles, and do not require explicit knowledge of the global structure of the network.

III. The continuous multi-median and the traveling salesperson problems

The Traveling Salesperson Problem (TSP), the multi-median problem, and their variations are some of the most widely known combinatorial and geometric optimization problems. While extensively studied in the literature, these problems continue to attract great interest from a wide range of fields, including Operations Research, Mathematics and Computer Science.

A. The continuous multi-median problem

Given a set $\mathcal{Q} \subset \mathbb{R}^d$ and a vector $P = (p_1, \ldots, p_m)$ of m distinct points in $\mathcal{Q}$, the expected distance between a random point q, generated according to a probability density function φ, and the closest point in P is given by

$$\begin{aligned} H_m(P, \mathcal{Q}) &:= \mathrm{E}\left[\min_{i \in \{1, \ldots, m\}} \|p_i - q\|\right] \\ &= \sum_{i=1}^{m} \int_{\mathcal{V}_i(P, \mathcal{Q})} \|p_i - q\| \varphi(q) dq, \end{aligned}$$

where $\mathcal{V}(P, \mathcal{Q}) = (\mathcal{V}_1(P, \mathcal{Q}), \ldots, \mathcal{V}_m(P, \mathcal{Q}))$ is the Voronoi partition of the set $\mathcal{Q}$ generated by the points P. In other words, $q \in \mathcal{V}_i(P, \mathcal{Q})$ if $\|q - p_i\| \leq \|q - p_k\|$, for all $k \in \{1, \ldots, m\}$. The set $\mathcal{V}_i$ is referred to as the Voronoi cell of the generator p_i. The function H_m is known in the locational optimization literature as the continuous Weber function or the continuous multi-median function; see [11], [12] and references therein.

The m-median of the set $\mathcal{Q}$, with respect to the measure induced by φ, is the global minimizer

$$P_m^*(\mathcal{Q}) = \operatorname*{argmin}_{P \in \mathcal{Q}^m} H_m(P, \mathcal{Q}).$$

We let $H_m^*(\mathcal{Q}) = H_m(P_m^*(\mathcal{Q}), \mathcal{Q})$ be the global minimum of H_m. It is straightforward to show that the map $P \mapsto H_1(P, \mathcal{Q})$ is differentiable and strictly convex on $\mathcal{Q}$. Therefore, it is a simple computational task to compute $P_1^*(\mathcal{Q})$. It is convenient to refer to $P_1^*(\mathcal{Q})$ as the median of $\mathcal{Q}$. On the other hand, the map $P \mapsto H_m(P, \mathcal{Q})$ is differentiable (whenever $(p_1, \ldots, p_m)$ are distinct) but not convex, thus making the solution of the continuous m-median problem hard in the general case. It is known [11],

[13] that the discrete version of the m-median problem is NP-hard for $d \geq 2$. Gradient algorithms for the continuous m-median problems can be designed [14] by means of the equality

$$\frac{\partial H_m(P, \mathcal{Q})}{\partial p_i} = \int_{\mathcal{V}_i(P,\mathcal{Q})} \frac{p_i - q}{\|p_i - q\|} \varphi(q) dq. \qquad (1)$$

The set of critical points of H_m contains all configurations $(p_1, \ldots, p_m)$ with the property that each vehicle p_i is the generator of the Voronoi cell $\mathcal{V}_i(P, \mathcal{Q})$ as well as the median of $\mathcal{V}_i(P, \mathcal{Q})$. Finally, let Z_{H_m} denote the set of local minima of H_m and define

$$H^*_{m,\text{local}}(\mathcal{Q}) = \sup_{P \in Z_{H_m}} H_m(P, \mathcal{Q}).$$

B. The Euclidean Traveling Salesperson Problem

The Euclidean TSP is formulated as follows: given a set D of n points in R^d, find the minimum-length tour of D. Let $\text{TSP}(D)$ denote the minimum length of a tour through all the points in D; by convention, $\text{TSP}(\emptyset) = 0$.

The asymptotic behavior of stochastic TSP problems for large n exhibits the following interesting property. Assume that the locations of the n target are independent random variables, uniformly distributed in a compact set $\mathcal{Q}$; in [15] it is shown that there exists a constant $\beta_{\text{TSP},2}$ such that, almost surely,

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

In other words, the optimal cost of stochastic TSP 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 (2) is $\beta_{\text{TSP},2} = 0.7120 \pm 0.0002$, see [16], [17], [18]. Similar results hold in higher dimensions, and for non-uniform point distributions: from [19], the limit (2) takes the general form

$$\lim_{n \to +\infty} \frac{\text{TSP}(D)}{n^{1-1/d}} = \beta_{\text{TSP},d} \int_{\mathcal{Q}} \bar{\varphi}(q)^{1-1/d}\, dq \;\; \text{a.s.}, \qquad (3)$$

where $\bar{\varphi}$ is the density of the absolutely continuous part of the distribution φ. Notice that the bound (3) holds for all compact sets: the shape of the set only affects the convergence rate to the limit. According to [10], if $\mathcal{Q}$ is a “fairly compact and fairly convex” set in the plane, (2) provides an adequate estimate of the optimal TSP tour length for values of n as low as 15. Remarkably, the asymptotic cost of the stochastic TSP for uniform point distributions is an upper bound on the asymptotic cost for general point distributions, i.e.,

$$\lim_{n \to +\infty} \frac{\text{TSP}(D)}{n^{1-1/d}} \leq \beta_{\text{TSP},d}.$$

This follows directly from an application of Jensen's inequality for concave functions to the right hand side of (3):

$$\int_{\mathcal{Q}} \bar{\varphi}(q)^{1-\frac{1}{d}}\, dq \leq \left(\int_{\mathcal{Q}} \bar{\varphi}(q)\, dq \right)^{1-\frac{1}{d}} \leq \varphi(\mathcal{Q})^{1-\frac{1}{d}} = 1.$$

C. Tools for solving TSPs

The TSP is known to be NP-complete, which suggests that there is no general algorithm capable of finding the optimum tour in an amount of time polynomial in the size of the input. Even though the exact optimal solutions of a large TSP can be very hard to compute, several exact and heuristic algorithms and software tools are available for the numerical solution of Euclidean TSPs.

The most advanced TSP solver to date is arguably `concorde` [20]. Heuristic polynomial-time algorithms are available for constant-factor approximations of TSP solutions, among which we mention Christofides' [21]. On a more theoretical side, Arora proved the existence of polynomial-time approximation schemes, providing a $(1+\varepsilon)$ constant-factor approximation for any $\varepsilon > 0$ [22].

A modified version of the Lin-Kernighan heuristic [23] is implemented in `linkern`; this powerful solver yields approximations in the order of 5% of the optimal tour cost very quickly for many instances. For example, in our numerical experiments on a 2.4 GHz Pentium machine, approximations of random TSPs with 1,000 points typically required about two seconds of CPU time.[1]

In the following, we will present algorithms that require on-line solutions of large TSPs. Practical implementations of the algorithms will rely on heuristics, such as Lin-Kernighan's or Christofides'. If a constant-factor approximation algorithm is used, the effect on the asymptotic performance guarantees of our algorithms can be simply modeled as a scaling of the constant $\beta_{\text{TSP},d}$.

IV. Algorithms for the m-DTRP

In this section, we will discuss and analyze algorithms for the solution of the single- and multiple-vehicle DTRP. We will first present the key existing results, and discuss the best available control policies. Then, we will introduce our proposed policies and analyze their performance.

A. Existing results

The main reference on dynamic vehicle routing problems to date is the work of Bertsimas and van Ryzin [8], [9]. As discussed in these works, a key idea about this problem is its formulation as a minimization of waiting time—as opposed to travel cost. In [8], lower bounds are derived for the optimal system time in the single-vehicle DTRP, both in the light load case (i.e., $\lambda \to 0^+$), and in the heavy load case (i.e., $\lambda \to +\infty$). Subsequently, policies are designed for the two cases, and their performance is compared to the lower bounds. A similar approach is taken in [9] to extend the single-vehicle results to the multiple-vehicle case. These results are obtained through techniques drawn from combinatorial optimization, queueing theory, and geometrical probability. In what follows we consider the $d = 2$ case, and use the shorthand $\beta = \beta_{\text{TSP},2}$.

[1] Both `concorde` and `linkern` are written in ANSI C and are freely available for academic research use at `http://www.math.princeton.edu/tsp/concorde.html`.

1) Lower bounds: In the light load case, the lower bound on the m-DTRP system time is strongly related to the solution of the m-median problem:

$$T^* \geq H_m^*(\mathcal{Q}) + \bar{s}, \qquad \text{as } \lambda \to 0^+. \tag{4}$$

This bound is tight and we present asymptotically optimal policies for the light load case below.

For the heavy load case, the lower bound takes the form

$$T^* \geq \frac{\gamma^2 \lambda}{m^2(1-\rho)^2} + \frac{2\rho - 1}{2\lambda}, \qquad \text{as } \lambda \to (1/\bar{s})^-, \rho \to 1^-,$$

with $\gamma = \frac{2}{3\sqrt{2\pi}} \approx 0.266$. The bound is not known to be tight.

2) An optimal policy for the light-load case: In the light load case, some policies are known to achieve the lower bound (4) and, hence, to be optimal. Such an optimal policy was introduced in [8] for a single vehicle, and extended to the multiple-vehicle case in [9].

> **Stochastic Queue Median (SQM) policy —** Place one vehicle at each of the m-median locations of the region $\mathcal{Q}$. When targets arise, assign them to the nearest median location, and the corresponding vehicle. Each vehicle services its assigned targets in a First Come–First Served (FCFS) order, returning to its median after each service is completed.

Under the SQM policy, the system time approaches the established light-load lower bound (4), i.e., $T_{\text{SQM}} \to T^*$ as $\lambda \to 0^+$. On the other hand, the SQM policy is not able to stabilize the system as the load increases.

3) A good policy for the heavy-load case: The lower bound for the heavy load case is not known to be tight, and the best known policy for this case provably achieves only a constant-factor approximation. The best known heavy-load policy was introduced in [9]:

> **The Modified G/G/m policy** — For some fixed integer $k > 1$, divide $\mathcal{Q}$ into k subregions of equal measure, e.g., using radial cuts centered at a common depot. Within each region, form sets of targets of size l/k and, as sets are formed, deposit them in a queue. Service the queue in a FCFS order with the first available vehicle, by following optimal TSP tours, starting and ending at the depot. Optimize over l.

The modified G/G/m policy (so called because of the connection to queueing systems with general inter-arrival and service times, and m servers [10]) achieves the best known constant-factor approximation for the system time in heavy load, in the sense that

$$T_{\text{modG/G/m}} \leq \frac{\beta_{\text{TSP},2}^2}{2\gamma^2} T^*, \qquad \text{as } \lambda \to +\infty$$

The number k of subregions must be very large for the bound to hold. In [24] it is conjectured that the upper bound is in fact tight, and the policy is asymptotically optimal; this conjecture is strengthened by the recent work in [25], where the conjecture is proved true under certain additional assumptions.

The policies outlined in this section rely on the centralized *a priori* computation of the m-median of $\mathcal{Q}$ and of a partition of $\mathcal{Q}$ into regions of equal measure. As a consequence, these policies are not scalable to very large networks of vehicles, and are not adaptive to changes in the environment and in the network composition, e.g., due to failures or to the addition of new resources. In the following sections, we introduce novel policies for vehicle routing, which are decentralized and spatially distributed.

B. A novel policy for the single-vehicle DTRP

In this section, we propose a new policy for the single-vehicle DTRP, which achieves the same performance as the best known policies in the heavy load case, while maintaining optimal performance in the light load case. In what follows, given a tour T of D, a fragment of T is a connected subgraph of T. We now introduce our first policy.

> **Single-Vehicle Receding Horizon Median/TSP (sRH) policy** — While the set of targets D is empty, move at unit speed toward $P_1^*(\mathcal{Q})$ if $p \neq P_1^*(\mathcal{Q})$, otherwise stop. While D is not empty, do the following: (i) for a given $\eta \in (0, 1]$, find a path that maximizes the number of targets reached within $\tau = \max\{\text{diam}(\mathcal{Q}), \eta \,\text{TSP}(D)\}$ time units; (ii) service from the current location this optimal fragment. Repeat.

In other words, if $D \neq \emptyset$, the vehicle looks for a maximum-reward path starting from the current vehicle position and with an appropriate duration τ. After this path is completed, the algorithm is repeated taking into account the targets that have appeared during the execution of the previous step. In general, the performance of the system will depend on the choice of the horizon length η, which can be seen as a trade-off between computation requirements and achievable service rate. Note that the time horizon is not fixed, but is adjusted according to the cost of the outstanding demand. The following two results describe the asymptotic performance of the sRH policy.

Theorem 4.1: The sRH policy is asymptotically optimal in the light load case, that is,

$$T_{\text{sRH}} \to T^*, \qquad \text{as } \lambda \to 0^+.$$

Proof: Consider a generic initial condition for the vehicle's position in $\mathcal{Q}$ and for the outstanding target positions $D(0)$, with $n_0 = \text{card}(D(0))$. An upper bound to the time needed to service all of the initial targets is $n_0(\text{diam}(\mathcal{Q}) + s_{\max})$, where $s_{\max}$ is the maximum time for on-site servicing of targets in $D(0)$. When there are no targets outstanding in the target set D, the vehicle moves at unit speed toward the median point $P_1^*(\mathcal{Q})$. The vehicle reaches $P_1^*(\mathcal{Q})$ in at most $\text{diam}(\mathcal{Q})$ units of time.

The time needed to service the initial targets and go to the median is hence bounded by $t_{\text{ini}} \leq (n_0 + 1)\,\text{diam}(\mathcal{Q}) +$

$n_0 s_{\max}$. The probability that at the end of this initial phase the number of targets is reduced to zero is

$$P[n(t_{\text{ini}})=0] = \exp(-\lambda t_{\text{ini}}) \geq \exp(-\lambda((n_0+1)\operatorname{diam}(\mathcal{Q}) + n_0 s_{\max}),$$

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

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

$$T_{\text{sRH}} = \lim_{j\to+\infty} \mathrm{E}[T_j] = H_1^*(\mathcal{Q}) + \bar{s}.$$

This time equals the lower bound (4), thus establishing the optimality of the policy. ■

Theorem 4.2: An upper bound on the system time of the sRH policy in heavy load is

$$T_{\text{sRH}} \leq \frac{\beta_{\text{TSP},2}^2}{(2-\eta)\gamma^2} T^*, \qquad \text{as } \lambda \to +\infty.$$

We refer the reader to the appendix for the proof of Theorem 4.2.

C. A decentralized policy for the multiple-vehicle DTRP

Here we design a decentralized policy for the DTRP problem applicable to multiple-vehicle systems. Our design combines the sRH policy discussed in the previous section with distributed algorithms for locational optimization discussed in [14]. Here we refer to the sRH policy defined for a single vehicle in the environment $\mathcal{Q}$ as the sRH($\mathcal{Q}$) policy.

We shall assume that each vehicle has sufficient information available to determine: (1) its Voronoi cell, and (2) the locations of all outstanding events in its Voronoi cell. Any control policy that relies on information (1) and (2), is spatially distributed in the sense that the behavior of each vehicle depends only on the location of all other vehicles with contiguous Voronoi cells. A spatially distributed algorithm for the local computation and maintenance of Voronoi cells is provided in [14].

> **Multi-Vehicle Receding Horizon Median/TSP (mRH) policy** — For all $i \in \{1,\dots,m\}$, the i-th vehicle computes its Voronoi cell $\mathcal{V}_i(P,\mathcal{Q})$ and executes the sRH($\mathcal{V}_i(P,\mathcal{Q})$) policy with the single following modification. While the vehicle is servicing targets in an optimal fragment of $D \cap \mathcal{V}_i(P,\mathcal{Q})$, it will shortcut all targets already serviced by other vehicles.

In what follows we characterize the asymptotic properties of this policy in light load via the following theorem and provide simulation results for its behavior in heavy load.

Theorem 4.3: The mRH policy is locally asymptotically optimal in the light load case, that is,

$$T_{\text{mRH}} \to H_{m,\text{local}}^*(\mathcal{Q}) + \bar{s}, \qquad \text{as } \lambda \to 0^+.$$

Proof: The proof follows the same line as that of Theorem 4.1 on the performance of the sRH policy in light load. However, after the initial transient period, the vehicles follow a gradient flow for a non-convex cost function and therefore they will only reach the set of critical points of H_m. Because the targets are generated randomly, almost surely the vehicles will converge to the set of local minima of H_m (as opposed to the set of critical points of H_m). ■

We conclude this section by comparing, in terms of their asymptotic performance, our decentralized policies with the centralized policies proposed by Bertsimas and van Ryzin [8], [9]. In the light load limit for a single vehicle, the performance of the sRH policy is optimal and identical to the performance of the SQM policy. In the light load limit for a multi-vehicle network, the mRH policy is locally optimal, whereas the SQM policy is optimal (provided a global minimum for the continuous multi-median problem can be computed). In the heavy load limit for a single vehicle, the performances of the sRH and of the modified G/G/1 policies are identical; simulation results suggest that the mRH policy achieves the same performance of the modified G/G/m policy. No analytic results are available yet on the mRH policy in the heavy load limit.

V. Simulation Results

In this section we present the results of a numerical experimentation of the sRH and mRH policies. All numerical experiments were conducted on a 2.4GHz Pentium-class machine, running RedHat®Linux 9. The algorithms described in the paper were implemented in Matlab®6.1, with external calls to the program `linkern`. Because we use a heuristic TSP solver as opposed to an exact one, we expect about 5% cost errors in the computation of the TSP solution. In Figure 1 we present a summary of numerical results of sample simulations of the sRH policy. The experimental results match well with the theoretical prediction for the asymptotic upper bound on the cost. In Figure 2 we show snapshots of a simulation experiment for the mRH policy in the heavy load case. We refer to [14] for simulations of the mRH policy in the light load regime.

VI. Conclusions

In this paper we presented some initial results on the design of decentralized algorithms for vehicle routing in a stochastic time-varying environment. Our control policies are spatially distributed in the sense that the behavior of each vehicle depends only on the location of all other vehicles with contiguous Voronoi cells.

We conclude by mentioning some limitations of our approach. In our analysis, we considered omni-directional vehicles with first order dynamics: non-holonomic constraints will have to be taken into account for practical application to UAVs or other systems. In this paper, all targets have the same value and are removed from the demand queue only upon service; in some scenarios, targets

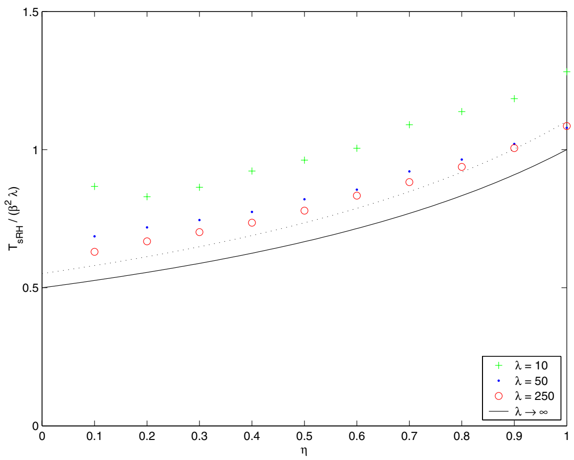

Fig. 1. Numerical experiment results: System time as a function of the parameter η using the sRH policy, for several values of λ. The results are averages over 1000 runs per point. The solid line is the theoretical asymptotic upper bound on the system time; since an exact TSP solver was not used, in favor of a fast heuristic program, we also report a correction to the upper bound allowing for a 5% sub-optimality in the TSP solution (dotted line).

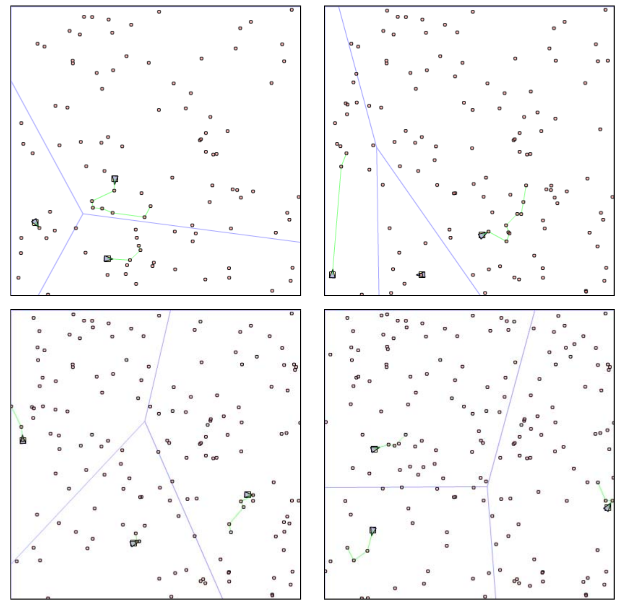

Fig. 2. Example of execution of the mRH policy ($m = 3, \lambda = 100$, $\eta = 0.1$, $\rho = 0$). The frames are 0.1 time units apart, and are presented in a left-right, top-down order.

might have different values and disappear before being serviced. These issues are under current investigation.

VII. Acknowledgments

This research has been partially supported by DARPA/AFOSR through grant F49620-02-1-0325, and by NSF through grant CCR-0325716. 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 organizations.

References

[1] D. L. Estrin *et al.* *Embedded, Everywhere: A Research Agenda for Networked Systems and Embedded Computers*. National Academy Press, 2001.

[2] S. Tilak, N. B. Abu-Ghazaleh, and W. Heinzelman. A taxonomy of wireless micro-sensor network models. *SIGMOBILE Mobile Computing and Communication Review*, 6(2):28–36, April 2002.

[3] C. Schumacher, P. R. Chandler, S. J. Rasmussen, and D. Walker. Task allocation for wide area search munitions with variable path length. In *Proc. American Control Conf.*, pages 3472–3477, Denver, CO, June 2003.

[4] R. W. Beard, T. W. McLain, M. A. Goodrich, and E. P. Anderson. Coordinated target assignment and intercept for unmanned air vehicles. *IEEE Trans. Robotics and Automation*, 18(6):911–922, 2002.

[5] A. Richards, J. Bellingham, M. Tillerson, and J. How. Coordination and control of multiple UAVs. In *Proc. AIAA Conf. on Guidance, Navigation, and Control*, Monterey, CA, August 2002.

[6] A. E. Gil, K. M. Passino, and A. Sparks. Cooperative scheduling of tasks for networked uninhabted autonomous vehicles. In *Proc. IEEE Conf. on Decision and Control*, pages 522–527, Maui, Hawaii, December 2003.

[7] W. Li and C. G. Cassandras. Stability properties of a cooperative receding horizon controller. In *Proc. IEEE Conf. on Decision and Control*, pages 492–497, Maui, Hawaii, December 2003.

[8] D. J. Bertsimas and G. J. van Ryzin. A stochastic and dynamic vehicle routing problem in the Euclidean plane. *Operations Research*, 39:601–615, 1991.

[9] D. J. Bertsimas and G. J. van Ryzin. Stochastic and dynamic vehicle routing in the Euclidean plane with multiple capacitated vehicles. *Operations Research*, 41(1):60–76, 1993.

[10] R. C. Larson and A. R. Odoni. *Urban Operations Research*. Prentice-Hall, Englewood Cliffs, NJ, 1981.

[11] P. K. Agarwal and M. Sharir. Efficient algorithms for geometric optimization. *ACM Computing Surveys*, 30(4):412–458, 1998.

[12] Z. Drezner, editor. *Facility Location: A Survey of Applications and Methods*. Springer Series in Operations Research. Springer Verlag, New York, 1995.

[13] N. Megiddo and K. J. Supowit. On the complexity of some common geometric location problems. *SIAM Journal on Computing*, 13(1):182–196, 1984.

[14] J. Cortés, S. Martínez, T. Karatas, and F. Bullo. Coverage control for mobile sensing networks. *IEEE Transactions On Robotics and Automation*, 20(2):243–255, 2004.

[15] J. Beardwood, J. Halton, and J. Hammersley. The shortest path through many points. *Proceedings of the Cambridge Philoshopy Society*, 55:299–327, 1959.

[16] G. Percus and O. C. Martin. Finite size and dimensional dependence of the Euclidean traveling salesman problem. *Physical Review Letters*, 76(8):1188–1191, 1996.

[17] D. S. Johnson, L. A. McGeoch, and E. E. Rothberg. Asymptotic experimental analysis for the held-karp traveling salesman bound. In *Proc. 7th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 341–350, 1996.

[18] N. J. Cerf, J. Boutet de Monvel, O. Bohigas, O. C. Martin, and A. G. Percus. The random link approximation for the Euclidean traveling salesman problem. *Journal de Physique I*, 7(1):117–136, 1997.

[19] J. M. Steele. Probabilistic and worst case analyses of classical problems of combinatorial optimization in Euclidean space. *Mathematics of Operations Research*, 15(4):749, 1990.

[20] D. Applegate, R. Bixby, V. Chvátal, and W. Cook. On the solution of traveling salesman problems. In *Documenta Mathematica, Journal der Deutschen Mathematiker-Vereinigung*, pages 645–656, Berlin, Germany, August 1998. Proceedings of the International Congress of Mathematicians, Extra Volume ICM III.

[21] N. Christofides. Bounds for the travelling-salesman problem. *Operations Research*, 20:1044–1056, 1972.

[22] S. Arora. Nearly linear time approximation scheme for Euclidean TSP and other geometric problems. In *Proc. 38th IEEE Annual Symposium on Foundations of Computer Science*, pages 554–563, Miami Beach, FL, October 1997.

[23] S. Lin and B. W. Kernighan. An effective heuristic algorithm for

the traveling-salesman problem. *Operations Research*, 21:498–516, 1973.
[24] D. J. Bertsimas and G. J. van Ryzin. Stochastic and dynamic vehicle routing with general interarrival and service time distributions. *Advances in Applied Probability*, 25:947–978, 1993.
[25] X. Lu, A. Regan, and S. Irani. An asymptotically optimal algorithm for dynamic traveling repair problem. In *Proc. of the 81st Meeting of the Transportation Research Board*, January 2002.
[26] W. Feller. *An Introduction to Probability Theory and Its Applications*. Wiley, New York, NY, 1968.

APPENDIX

Proof of Theorem 4.2 Due to space limitations, we limit the proof to the basic case in which targets are generated according to a spatially-uniform Poisson point process, and no on-site servicing is required, i.e., $\varphi(q) = 1$, and $s_i = 0$, for all $i \in \mathbb{N}$.

First of all, we establish a connection between the cost $C(t_i)$ to service outstanding targets at the i-th decision time t_i and the system time T_{sRH}. Indicate with $\bar{C}$ an upper bound on the steady-state value of the cost $C(t_i)$, i.e.,

$$\lim_{i \to +\infty} C(t_i) \leq \bar{C},$$

and assume that $\bar{C}$ is finite. A newly-generated target waits on average at most $\eta\bar{C}/2$ before the next decision time, at which it will be first considered for service. At each decision time occurring after its generation, it has a probability at least η to be included in the set of targets to be serviced; each time it is not selected for service, its service time increases by at most $\eta\bar{C}$. Once it is selected for service, it has to wait on average at most $\eta\bar{C}/2$ before being actually serviced. Summarizing,

$$T_{\text{sRH}} \leq \eta\frac{\bar{C}}{2} + \sum_{k=1}^{+\infty}\left[(1-\eta)^k\eta\bar{C}\right] + \eta\frac{\bar{C}}{2} = \bar{C}. \qquad (5)$$

In other words, in the sRH policy, if the cost of servicing outstanding targets is eventually bounded by a constant $\bar{C}$, the system time is upper bounded by the same constant.

Now, we proceed to study of the sequence $C(t_i)$, formed by the costs of servicing all targets in queue at the decision times. In the sRH policy, and in heavy load conditions, at each decision time t_i a path is computed that services at least $\eta n(t_i)$ points in $\eta C(t_i)$ time. We want to compute an estimate of the cost $C(t_{i+1})$, based on the knowledge of the cost $C(t_i)$, and on the properties of the sRH policy.

The targets in $D(t_i)$ are sampled from a time-varying distribution with p.d.f. $\tilde{\varphi}(q, t_i)$, possibly different from the function φ, and in general unknown. (It depends on the target-servicing choices made at the previous decision times.)

In the heavy-load limit, we can write

$$\begin{aligned}\lim_{\lambda\to+\infty}\frac{C(t_i)}{\lambda} &= \lim_{\lambda\to+\infty}\frac{C(t_i)}{\int_{\mathcal{Q}}\sqrt{n(t_i)\tilde{\varphi}(q,t_i)}\,dq} \\ &\quad\cdot\frac{\int_{\mathcal{Q}}\sqrt{n(t_i)\tilde{\varphi}(q,t_i)}\,dq}{\lambda} \\ &= \beta\lim_{\lambda\to+\infty}\frac{\int_{\mathcal{Q}}\sqrt{n(t_i)\tilde{\varphi}(q,t_i)}\,dq}{\lambda}, \text{ a.s.}\end{aligned}$$

The set of targets in queue at time t_{i+1} can be partitioned into a set of "old" targets, i.e., targets already in queue at time t_i, and "new" targets, i.e., targets generated in the time interval $[t_i, t_{i+1})$. Since new targets are generated according to φ, the point distribution at time t_{i+1} is described by a p.d.f. such that:

$$n(t_{i+1})\tilde{\varphi}(q,t_{i+1}) = n_{\text{old}}(t_{i+1})\tilde{\varphi}_{\text{old}}(q,t_{i+1}) + n_{\text{new}}(t_{i+1})\varphi(q),$$

for some unknown function $\tilde{\varphi}_{\text{old}}$. The cost of servicing "old" targets is by the definition of the sRH policy, equal to $(1-\eta)$ times the cost at the previous decision time. As a consequence,

$$\lim_{\lambda\to+\infty}\frac{\sqrt{n_{\text{old}}(t_{i+1})\tilde{\varphi}_{\text{old}}(q,t_{i+1})}}{\lambda} = \lim_{\lambda\to+\infty}\frac{(1-\eta)C(t_i)}{\beta\lambda}, \text{ a.s.}$$

Using the strong law of large numbers [26], it can be shown that the number of new targets generated within the time interval $[t_i, t_{i+1})$, of duration $\eta C(t_i)$, satisfies

$$\lim_{\lambda\to+\infty}\frac{n_{\text{new}}(t_{i+1})}{\lambda} = \eta C(t_i), \text{ a.s.}$$

Hence, if the point-generation process is spatially uniform (i.e., $\varphi(q) = 1$), we can write:

$$\lim_{\lambda\to+\infty}\frac{C(t_{i+1})}{\lambda} = \sqrt{\frac{(1-\eta)^2 C(t_i)^2}{\lambda^2} + \beta^2\eta\frac{C(t_i)}{\lambda}}, \text{ a.s.} \qquad (6)$$

The above equation describes a discrete-time, nonlinear system, that converges from all positive initial conditions to the stable equilibrium:

$$\lim_{\lambda, i\to+\infty}\frac{C(t_i)}{\lambda} = \frac{\beta^2}{2-\eta}, \text{ a.s.}.$$

In this case we can conclude that in the heavy-load limit,

$$\bar{C} = \frac{\beta^2\lambda}{2-\eta}. \qquad (7)$$

From (5) and (7) we get the stated result. ■