The +1 Method Model-Free Adaptive Repositioning Policies for Robotic Multi-Agent Systems

The +1 Method
Model-Free Adaptive Repositioning Policies for Robotic Multi-Agent Systems.

Claudio Ruch, Joel Gächter, Jan Hakenberg and Emilio Frazzoli [1]

Abstract
Robotic multi-agent systems can efficiently handle spatially distributed tasks in dynamic environments. Problem instances of particular interest and generality are the dynamic vehicle routing problem and the dynamic traveling repairman problem. Operational policies for robotic fleets solving these two problems take decisions in an online setting with continuously arriving dynamic demands to optimize system time and efficiency. They can be classified along several lines. First, some require a model of the demand, e.g., based on historical information, while others work model-free. Second, they are designed for different operating conditions from light to heavy system load. Third, they work in a time-invariant or time-varying setting. We present a novel class of model-free operational policies for time-varying demands, with performance independent of the load factor and applicable to any number of dimensions, a combination of properties not achieved by any other operational policy in the literature. The underlying principle of the introduced policies is to send available robots to recent realizations of the stochastic process that generates service requests. In simple terms, the strategies rely on sending more than one robot for every service request arriving to the system. This leads to an advantage in scenarios where demand is non-uniformly distributed and correlated in space an time. We provide theoretical stability and performance guarantees for both the time-invariant and the time-varying cases as well as for correlated demand. We verify our theoretical results numerically. Finally, we apply our operational policy to the problem of mobility-on-demand fleet operation and demonstrate that it outperforms model-based and complex algorithms across all load ranges despite of its simplicity.

Introduction

Groups of robots have been considered for exploration in space, underwater and on land Leitner (2009); Sutantyo et al. (2013); Popa et al. (2004); Savla et al. (2007). They were used for entertainment purposes in Hauri et al. (2014) as well as considered for defence applications, e.g., in Browning et al. (2005). Lately, multi-robot systems have gained additional interest with the rise of the fully autonomous vehicle which could soon populate cities in large fleets as shown in Spieser et al. (2014). Two important classes of problems in this setting are the N-Vehicle Dynamic Traveling Repairman Problem (N-DTRP) and the N-Vehicle Dynamic Vehicle Routing Problem (N-DVRP). In the N-DTRP, a fleet of N robots need to serve requests that appear in a stochastic environment with minimal wait time and with highest possible fleet efficiency. Service requests are not known a priori but appear while time continues. The N-DVRP is closely related except that service requests are travel requests in which a customer wants to be transported from an origin to a destination in the operational domain. After completing a service, the robot or vehicle is at a different location in the operational domain as when it started servicing a request.

An algorithm or operational policy guiding the behavior of a fleet of robots in a N-DTRP or N-DVRP setting must guarantee system stability and optimize relevant performance metrics, e.g., the expected wait time for new service requests for all possible problem configurations. As such, it must be able to react to changes in input. A static optimization of the routes of all robots is in general not sufficient as information gets available during run-time and is not known before, see Bullo et al. (2011) for details.

In this work, we focus on four key differentiations of operational policies. First, a policy can be based on a (statistical) model of the demand. Such models can be derived from fundamental problem properties, e.g., constant distribution over time or build from historical data. This category of policies is called *model-based* which is opposed to *model-free*. Second, operational policies can be designed for different system loads that range from light load to heavy load. In light load, almost no requests occur compared to

[1] The authors are with the Institute for Dynamic Systems and Control at ETH Zürich

Corresponding author:
Claudio Ruch
Institute for Dynamic Systems and Control
ETH Zürich
Sonneggstrasse 3
8092 Zürich
Switzerland
Email: clruch@idsc.mavt.ethz.ch

the number of robots. In heavy load, operation takes place close to the system's stability limit. Operational policies can be suitable for light-load, heavy-load or a range of loads. Third, operational policies can be designed for time-invariant scenarios or be able to cover time-varying load scenarios. Fourth, operational policies can be suitable or optimized for correlated demand.

Next to these main categories, the problems can be extended in many ways, e.g., by adding time constraints on service, adding service types and priorities, including moving service requests or recharging constraints for the robots. Good overviews of the field and the methods can be found in Gendreau and Potvin (1998); Ritzinger et al. (2016); Bullo et al. (2011).

An early result belonging to the category of model-based operational policies was presented in Waisanen et al. (2008) where both the N-DVRP and N-DTRP are considered for known and uniform service request distributions and destination distributions on a square operational domain. Theoretical bounds for optimal performance are provided as well as an algorithm that can reach these bounds. An important member of this family for more general cases is the m-Stochastic Queue Median Policy presented in Bullo et al. (2011). If the time-invariant distribution is known, then for a fleet of m robots the positions of m stochastic medians can be computed. The policy of placing one robot at one of these locations which serves all requests appearing in the Voronoi partition generated by its position is provably optimal under light load, i.e., when the rate of newly generated requests λ approaches zero. However, the policy is unstable for higher arrival rates. Different algorithms for the same problem have been derived with game theoretic formulations, e.g., in Savla and Frazzoli (2010). An important result covering many load cases is presented in Frazzoli and Bullo (2004). The proposed Multi-Vehicle Receding Horizon Median/TSP policy performs optimally under light-load conditions and maintains high performance under heavy-load conditions. In Pavone et al. (2011b) and Pavone et al. (2011a), the approaches for the time-invariant case with known spatial distribution of requests were formalized to general problems in two steps. First, an algorithm is used to partition the operational domain into partitions which are simultaneously equitable with respect to the request distribution and its square root, then each robot executes the single-vehicle "Divide & Conquer" policy in its partition which is based on consecutive solutions of the traveling salesman problem in the partition. Several publications attempt to generalize the setting to cases in which the requests are generated by time-varying density functions. In Lee et al. (2015), the authors present an algorithm to keep the robots on the generators of a Centroidal Voronoi Tesselation for a known, time-varying request density function. Stochastic gradient descent methods for cases with known distribution of service requests are considered in Le Ny and Pappas (2010). Other examples of gradient descent based methods for coverage can be found in Schwager et al. (2008) and Cortes et al. (2004). Finally, also model predictive control methods have been applied to solve N-DTRP and N-DVRP problems, mostly in the setting of transportation. Examples can be found in Miao et al. (2015), Zhang et al. (2016) and Iglesias et al. (2018).

Model-free operational policies do not require any past data or knowledge of the service request generating stochastic process to operate. Two important model-free policies are presented in Arsie et al. (2009a). Both are based on the principle that every robot either moves towards its closest open request or the location minimizing the average to all requests it has served in the past. They are decentralized and are demonstrated to converge to a saddle point of the cost function for the light-load under time-invariant demand. While convergence to a local minimum is not guaranteed, it is obtained in many cases. The policy's performance under time-varying demand is not investigated. However, its design clearly does not favor this scenario as the robots converge to one single position as time continues. For time-varying demand, the optimal position of robots is a function of time.

Another model-free policy which is adaptive to time-varying demand is the "Nearest Neighbor" policy mentioned in Bullo et al. (2011). The policy directs each robot towards the closest open service request. If there is no open service request, every robot stops at the current position. To our knowledge, no theoretical guarantees for this operational policy are available at this point.

Finally, two operational policies for the N-DTRP problem are presented in Le Ny and Pappas (2013). The first policy is a variation of the policy presented in Arsie et al. (2009a) which provides the same convergence guarantees in light load and is described to be simpler. The second provably stabilizes the system as long as it is stabilizable. No guarantees on performance or robustness are provided.

Statement of Contributions: In this work, we propose novel class of operational policies which can solve both the N-DTRP and the N-DVRP. The policies exploit that past and recent requests locations are good predictors of future request locations. They utilize this principle by sending more than one available robot to the locations where a request appears. The policies are model-free, i.e., they do not require any knowledge about the stochastic process that generates new requests. Furthermore, we prove that the policies stabilize the system in all operating conditions from light to heavy load as long as the robot fleet size exceeds the minimal number of robots required for stability. We provide performance bounds for the time-invariant setting in which the arrival rate and distribution of requests is constant. Then, we extend these results and show that if change in request distribution is reasonably small compared to the arrival rate of new requests, the performance guarantees hold also for general time-varying cases. We also show that our proposed operational policies perform better with increased levels of correlation in the demand. Then, we computationally demonstrate the correctness of the obtained guarantees. Finally, we apply the proposed operational policies to the mobility-on-demand fleet operation case in simulation. We do so using the high-fidelity simulation environment Ruch and Frazzoli (2018) and demonstrate that our operational policies outperform previous work including model-based, complex algorithms based on mathematical optimization.

Paper Organization: First, the problem setting is described in Section Problem Formulation and the proposed operational policies are described in Section Operational Policies. In Section Analysis, a detailed analysis of performance bounds is presented for a time-invariant

setting and then extended to general time-varying settings and correlated demand. The results are verified with computational methods in Section Numerical Verification. Then the operational policies are applied to the mobility-on-demand fleet operation problem in Application to Mobility-on-Demand and a conclusion is presented in Conclusions and Future Work.

Problem Formulation

The *operating domain* $\Omega \subset \mathbb{R}^n$ is a non-empty, bounded and simply connected domain. To facilitate understanding, we present the entire analysis for the case $n = 2$ but as the number of dimensions is never used in the derivations, the results hold for spaces with *any number of dimensions*.

In Ω, a fleet of N robots with positions $\mathbf{p}(t) = \{p_1(t), ..., p_N(t)\} \in \Omega^N$ is operating during the time $t \in [0, \infty)$. In the operating domain a stochastic process generates *service requests* r_i over time that need to be serviced by the N robots. Request r_i is issued at time t_i. Then, w_i is the wait time of service request r_i, i.e., the time interval between t_i and its service begin. Furthermore, let m be the number of requests that have arrived to the system up to and including time t and $A(t) = (r_1, r_2, ..., r_m)$ be the list of these requests sorted according to their arrival time. We denote $R(t) \in \mathbb{N}_0$ as the number of outstanding requests at time t.

Definition 1: (System Time) *The system time is defined as the long-term average of wait times* $\bar{W} := \lim_{t \to \infty} \frac{1}{m} \sum_{i=1}^{m} w_i$.

All N robots in the system can move in an unconstrained way in Ω with constant speed $v > 0$. Stability of the system is defined as follows according to Pavone et al. (2010):

Definition 2: (Stable Policy) *An operational policy π guiding the movements of the N robots is stable if the expected number of demands in the system is uniformly bounded at all times.*

Stability also implies that the system time under policy π does not diverge, i.e., $\bar{W}(\pi) < \infty$. We can now define the first of the problems considered in this work.

Definition 3: (N-DTRP) *The service request r_i is a tuple $(t_i, \mathbf{x}_i, s_i)$ in which t_i is the time at which the request is generated, $\mathbf{x}_i \in \Omega$ is the location of the request and $s_i \geq 0$ is the service time of the request. A service request r_i is accounted as serviced if after t_i one of the N robots moves to $\mathbf{x}_i$ and stays there during a period of s_i.*

The N-DTRP is to find a stabilizing operating policy π that minimizes $\bar{W}(\pi)$.

The N-DVRP is a generalization of the N-DTRP in which every request r has an additional destination to which it must be transported. Formally:

Definition 4: (N-DVRP) *The service request r_i is a tuple $(t_i, \mathbf{x}_i, \mathbf{y}_i)$ in which t_i is the time at which the request is generated, $(\mathbf{x}_i, \mathbf{y}_i) \in \Omega \times \Omega$ are the origin and destination of the request r_i. A service request r_i is serviced if after t_i one of the N robots moves to $\mathbf{x}_i$ and then directly to $\mathbf{y}_i$.*

The N-DVRP is to find a stabilizing operating policy π that minimizes $\bar{W}(\pi)$.

The difference between the N-DTRP and the N-DVRP is that in the latter, a robot servicing request r_i is at a different location $\mathbf{y}_i$ after ending the service.

In this work, we distinguish three different statuses of the robots at a given time t. They can be busy serving a target, in a repositioning task or waiting at the location determined by the operational policy. For a fleet size of N robots we partition:

$$N = N_b(t) + N_r(t) + N_w(t) \qquad \forall t$$

Where $N_b(t)$ robots busy at time t, $N_r(t)$ are repositioning and $N_w(t)$ are waiting at the location determined by the operational policy.

The used variables are summarized in Table 1.

N	Number of robots
v	Constant speed of robots
$\Omega \subset \mathbb{R}^2$	Operating domain of robots
$|\Omega|$	$|\Omega| = \int_\Omega 1 dA$, area of Ω.
$N_b(t)$	Number of busy robots at time t
$N_r(t)$	Number of robots in a repositioning task at time t
$N_w(t)$	Number of available robots that have reached their determined waiting destination at time t
$p_i(t) \in \Omega$	Position of robot i at time t
$R(t) \in \mathbb{N}_0$	Number of unserved requests at time t.
$r_i = (t_i, \mathbf{x}_i, s_i)$	Request r_i with submission time t_i, submission location $\mathbf{x}_i$ and service time s_i
π	Operational policy
$\lambda(t)$	Request arrival rate at time t
$\bar{\lambda}$	Long-term average arrival rate.

Table 1. Important variables and notation.

Operational Policies

The presented policy is based on the fact that recent locations of past service requests are good estimators of future request locations. As it will be shown in the analysis section this holds for both time-invariant and time-varying processes and especially if individual requests are correlated. Therefore, the proposed operational policy relocates available robots to the locations where requests have arrived to the system most recently. Specifically, if a request appears it is serviced by one robot, another idling robot is sent to the location of the request if the status of the fleet allows for it. In simple terms, the principle of the policy can be described with the idea of sending one "+1" robots to a new request.

We first introduce variables to conveniently describe this process. A set L is used to to record past request locations: $L := \{(i, \Lambda(i)) : i \in \{1, ..., m\}\}$ where the function Λ maps indices to locations in the operating domain: $\Lambda : \{1, ..., m\} \to \Omega$. The auxiliary function $\Phi(k) : \mathbb{N} \to \Omega^k$ returns the $k \leq m$ elements of L with the highest index. In an implementation, L would simply be map of time instances and locations sorted according to the time instances. We can now define the operational policy for a scenario with m requests arriving in the order $r_1, ..., r_m$ for a fleet of N robots initially located at positions $\{p_1(0), ..., p_N(0)\} \in \Omega^N$.

Operational Policy Model-free Anticipative Repositioning

Initialize $L \leftarrow \{(j, p_j(0))\} : j \in \{1, ..., N\}\}$.
while operating **do**
   $L \leftarrow L \cup \{N + j, r_j\} \forall$ new requests j
   $N_{busy} \leftarrow \min(R(t) + N_b(t), N)$
   $N_{idle} \leftarrow \max(N - N_{busy}, 0)$
   $N_{serve} \leftarrow \min(R(t), N_{idle})$.
   `sendToRequests(`N_{serve}`)`
   `repositionIdle(`$\Phi(N_{idle})$`)`

The operational policy can be executed with varying sampling times depending on what problem must be solved and what hardware can be used to carry out the computations. The abstract operation `sendToRequests(`N_{serve}`)` denotes the process of sending N_{serve} robots to N_{serve} open requests and possibly selecting the requests to serve if $N_{serve} < R(t)$. The abstract operation `repositionIdle(`$\Phi(N_{idle})$`)` denotes the process of repositioning N_{idle} robots to the last N_{idle} past request locations. These are tracked in L and accessed with $\Phi(N_{idle})$.

Both operations can be carried out in different ways in a manner adapted to the problem specifications. One possibility is to complete `sendToRequests(`N_{serve}`)` solving a bipartite matching problem with Euclidean or network distances as a cost function between the robot locations and the open requests. Such a choice will minimize the sum of distances to reach all open requests at every time step. Similarly, the operation `repositionIdle(`$\Phi(N_{idle})$`)` can be implemented by solving a bipartite matching problem based with Euclidean or network distances between the robot locations and the past request locations $\Phi(N_{idle})$. Bipartite matching problems are in general solved with the Hungarian method as published in Kuhn (1955) or adaptions of it. If the cost function is a metric, faster algorithms that exploit its properties are available, e.g., in Agarwal and Varadarajan (2004). In this work, this specific implementation based on bipartite matching is used for the computation of numeric results. However, bipartite matching was not required to derive theoretical performance guarantees. Having stated the operational policies, we analyze its properties in the next section for the case of the N-DTRP.

Analysis

In this section, we derive theoretical results that characterize Model-free Anticipative Repositioning. We do so for the N-DTRP but note that the operational policy could also be applied to the N-DVRP successfully. However, in that case, the available information on where trips end is not fully utilized but could be included in augmented versions.

We demonstrate stability for both the cases of a constant demand profile and a time-varying demand profile, i.e., that the expected number of unserviced demands in the system is bounded at all times. Then, for both constant and time-varying demand profiles, theoretic guarantees on performance of the operational policies are presented. For the sake of brevity, we present a detailed analysis of the theoretic guarantees for constant request distributions. Then, instead

of repeating the same arguments in the time-varying case, we show that for requests distributions which change relatively slow in comparison to the arrival rate of new requests, the results for time-invariant settings can be fully recovered in the time-varying case. This first part of the analysis is done for requests which are identically and independently distributed (i.i.d.). To conclude the theoretical analysis, we drop the assumption of i.i.d. service requests and present insight into the behavior of the proposed operational policy for the case when service request locations are correlated. Specifically we look at the case in which a realization of an arrival at some location in the system leads to another arrival at the same location with certain probability. We start with the analysis of stability.

Stability Analysis

Lemma 1: *Let $\bar{S}$ be the average time needed for the completion of a service request, including the service time and displacement of the robot to the request location. Let $\bar{\lambda}$ be the average arrival rate of requests. Model-free Anticipative Repositioning is stable if the fleet size N is larger than $N_{stability} := \bar{S}\bar{\lambda}$.*

Proof. We follow the procedure presented in Treleaven et al. (2013). Let $\bar{R}$ be the long-term average number of unserviced demands, let $A(t)$ be the number of requests arrived in the system up to time t and S_i the service time necessary to serve request i, then

$$\bar{R} = \lim_{t \to \infty} \frac{1}{t} \int_0^t R(\tau) d\tau$$
$$= \lim_{t \to \infty} \frac{\sum_{i=0}^{A(t)} S_i}{A(t)} \lim_{t \to \infty} \int_0^t \lambda(\tau) d\tau$$
$$= \bar{S}\bar{\lambda} = N_{stability}$$

For the second equality, Little's theorem was used. We now see that the long-term average of busy robots is $\bar{S}\bar{\lambda} = N_{stability}$. If $N < N_{stability}$, the robots cannot service all request and the number of open requests grows unbounded. If $N \geq N_{stability}$ Model-free Anticipative Repositioning is stable as it continuously assigns robots to service requests. As the policy does not carry out any repositioning tasks unless there are idle robots, it is sufficient to consider only the assignment part of the policy. ∎

Performance under Constant Demand

In this section, we analyze the operational policy introduced in Section Operational Policies for a constant demand profile, i.e., new requests arrive at unknown rate $\lambda > 0$ and at a location sampled from the spatial density $f(\mathbf{x})$ which is constant in time and for which $f(\mathbf{x}) = 0, \forall \mathbf{x} \notin \Omega$.

Our measure of service level is the long term fraction of requests waiting longer than $w_{\max}$ denoted $\bar{F}_{w_{\max}}$. As the robot can move in the operating domain without constraints with velocity v, also a maximum distance to a service request can be defined to have a wait time of smaller than $w_{\max}$:

Definition 5: *$\bar{F}_{w_{\max}}$ is the long-term fraction of requests waiting longer than $w_{\max}$ to receive service. The maximal distance of the closest available robot from a request to have wait time less than $w_{\max}$ is $r_{\max} := v \, w_{\max}$.*

For convenience, we introduce additional notation to describe the ball containing all points within a certain radius from a location $\mathbf{x}_c$:

Definition 6: *The set $B(\mathbf{x}_c, r)$ denotes as a ball with radius r centered at $\mathbf{x}_c$. Furthermore let $\Gamma(\mathbf{x}_c, r)$ be the integral of $f(\mathbf{x})$ on $B(\mathbf{x}_c, r)$, i.e.,*

$$\Gamma(\mathbf{x}_c, r) := \int_{B(\mathbf{x}_c, r)} f(\mathbf{x}) \, dA$$

As a first step, we show that the operational policy can drive the wait times arbitrarily low given a sufficiently high number of robots.

Lemma 2: *For any $\epsilon > 0$ and for any $w_{\max} > 0$ there exists a number of robots $N \in \mathbb{N}$ such that $\bar{F}_{w_{\max}} < \epsilon$.*

Proof. Let $N_w(t)$ be the number of available robots that have reached their waiting destination at time t. Model-free Anticipative Repositioning sets the waiting destinations at time t for the N_{idle} available robots as the last N_{idle} points in Ω where requests occurred before or at t.

After the system has run for a sufficiently long time to reach steady state, consider the probability that request r_j arriving to the system at time t waits longer than $w_{\max}$:

$$Pr(w_j > w_{\max}) =$$
$$\int_\Omega Pr(w_j > w_{\max} \mid \mathbf{x}_j = \mathbf{x}) \, Pr(\mathbf{x}_j = \mathbf{x}) \, dA =$$
$$\int_\Omega Pr(w_j > w_{\max} \mid \mathbf{x}_j = \mathbf{x}) \, f(\mathbf{x}) \, dA$$

where $Pr(w_j > w_{\max} \mid \mathbf{x}_j = \mathbf{x})$ is the conditioned probability that the waiting time of request j is longer then $w_{\max}$, given that the request r_j arrives at location $\mathbf{x}$.

Given sufficient idle robots, this probability is equal to the probability that none of the previous $N_w(t)$ requests arrived in $B(\mathbf{x}, r_{\max})$ which is exactly $(1 - \Gamma(\mathbf{x}, r_{\max}))^{N_w(t)}$. The situation is illustrated in Figure 1.
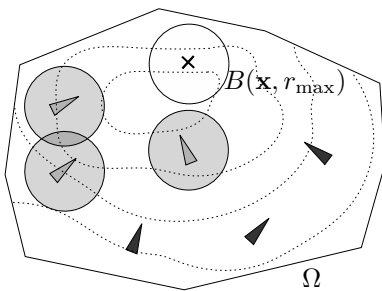

Figure 1. Contour lines of f are indicated with dotted lines. Three dark robots are busy or repositioning. Three light robots ($N_w(t) = 3$) are waiting at the determined location. A request arrives at $\mathbf{x}$, it is served in a time smaller than $w_{\max}$ if one of the $N_w(t)$ robots is positioned within $B(\mathbf{x}, r_{\max})$.

Therefore,

$$Pr(w_j > w_{\max}) = \int_\Omega (1 - \Gamma(\mathbf{x}, r_{\max}))^{N_w(t)} f(\mathbf{x}) \, dA \tag{1}$$

Note that $\Gamma(\mathbf{x}_c, r) \in [0, 1]$ because f is a probability density function. Only values $(1 - \Gamma(\mathbf{x}, r_{\max})) \in [0, 1)$ will

contribute to the integral as $\Gamma(\mathbf{x}, r_{\max}) = 0 \Rightarrow f(\mathbf{x}) = 0 \in B(\mathbf{x}, r_{\max})$. Therefore, $(1 - \Gamma(\mathbf{x}, r_{\max}))$ is strictly decreasing with growing $N_w(t)$ and approaching 0 in the limit. $\bar{F}_{w_{\max}}$ is the long-term average of $Pr(w_j > w_{\max})$, i.e,

$$\bar{F}_{w_{\max}} := \lim_{t \to \infty} \frac{1}{t} \int_0^t Pr(w_j > w_{\max}) \, d\tau$$
$$= \int_\Omega (1 - \Gamma(\mathbf{x}, r_{\max}))^{\bar{N}_w} f(\mathbf{x}) \, dA \tag{2}$$

Where $\bar{N}_w := \lim_{t \to \infty} \frac{1}{t} \int_0^t N_w(\tau) d\tau$ is the long term average number of robots waiting at the location determined by the operational policy. If the number of robots that are busy serving requests and which are repositioning to their wait location is finite, for any $\epsilon > 0$ and for any $w_{\max} > 0$ we can ensure large enough N such that $\bar{N}_w$ is sufficiently large to reduce $\bar{F}_{w_{\max}}$ to a value smaller than ϵ. ∎

An interesting consequence of this result is that for discrete time-invariant request distributions, the algorithm will either not meet a performance requirement $\bar{F}_{w_{\max}} < \epsilon$ or it will directly drive the wait times to zero.

Corollary 1: *Let all request arrive at $\{1, ..., Q\} \in \Omega^Q$ distinct points with probabilities $p_1, ..., p_Q$ such that $\sum_{i=1}^Q p_i = 1$. Furthermore, let all points be pairwise more distant than $r_{\max}$. Then, the number of robots N for a certain level of performance ϵ is independent of $w_{\max}$.*

Proof. We evaluate eq. 1 for the discrete distribution and see that

$$Pr(w_j > w_{\max}) = \sum_{i=1}^Q (1 - p_i)^{N_w(t)} p_i$$

independently of $r_{\max}$. Therefore, for a stringent enough level of performance $w_{\max}$ it only matters how many points of action Q there are and how the distribution of requests on these points is. ∎

This results highlights one of the strong properties of the presented operational policy. If the demand pattern is discrete with certain locations of increased demand and the performance requirement $w_{\max}$ is strict enough, the policy will not gradually reduce the wait times but directly drive them to zero. Many real-world problem instances can be modeled as such a distinct set of "centers of action", e.g, in surveillance applications, there will be a distinct number of points of interest, e.g., water ponds in animal surveillance or gates in security surveillance. Also in mobility-on-demand applications, requests often originate from certain hot spots, e.g., a shopping centers or schools.

In the analysis up to this point, we have shown that with a sufficiently large number of robots, any performance requirements can be met. Next, we derive lower bounds on the service levels achieved with the operational policy. First, we demonstrate that uniformly distributed requests will result in the poorest performance of the policy.

Lemma 3: *Under Model-free Anticipative Repositioning and for $\bar{N}_w \geq 1$, $\bar{F}_{w_{\max}}$ is maximal for the uniform spatial distribution $f(\mathbf{x}) = \frac{1}{|\Omega|}, \forall \mathbf{x} \in \Omega$.*

The idea of the proof is to apply a transformation to some density function f which decreases its uniformity. Then, we show that under this transformation $\bar{F}_{w_{\max}}$ will increase.

Proof. Let f be some probability density function on the domain Ω. We transform f to obtain the perturbed density function $\tilde{f}$ by increasing f by a factor $\delta > 1$ in a ball $B(\mathbf{x_c}, r_c)$ at $\mathbf{x}_c$ with distance $> r_{\max}$ from the boundary of Ω. The radius of the ball r_c is chosen such that $r_c << w_{\max} \; v$. To maintain the property $\int_{\Omega} \tilde{f}(\mathbf{x}) dA = 1$, f is uniformly decreased everywhere else:

$$\tilde{f}(\mathbf{x}) = \begin{cases} \delta f(\mathbf{x}) & \mathbf{x} \in B(\mathbf{x}_c, r_c) \\ f(\mathbf{x}) - \frac{1}{|\Omega| - |B(\mathbf{x}_c, r_c)|}(\delta - 1) \int_{B(\mathbf{x}_c, r_c)} f(\mathbf{x}) dA & \text{otherwise} \end{cases}$$

We assume that f is not exactly zero in the operating domain for balls with radius larger than r_c:

$$\int_{B(\mathbf{x}, r_c)} f(\mathbf{x}) \, dA > 0, \forall \mathbf{x} \in \Omega \qquad (3)$$

This technical assumption does not restrict generality and can be guaranteed by assigning an arbitrarily small and positive value to f instead of the value zero. This implies that the possibility of a service request is almost zero but not exactly zero and corresponds to the nature of the problem setting.

The transformed density function $\tilde{f}$ is identical to f except that some density was globally "subtracted" and added at one specific location. Going from f to $\tilde{f}$, a less uniform distribution $\tilde{f}$ was created for which

$$\int_{\Omega} \left(f - \frac{1}{|\Omega|} \right)^2 dA < \int_{\Omega} \left(\tilde{f} - \frac{1}{|\Omega|} \right)^2 dA \qquad (4)$$

Special cases can be found in which eq. 4 is not satisfied. For these a finite number of identical transformations will yield a distribution $\tilde{f}$ satisfying eq. 4. We compare the two expressions for the same number $\bar{N}_w$:

$$\bar{F}_{w_{\max}} = \int_{\Omega} (1 - \Gamma(\mathbf{x}, r_{\max}))^{\bar{N}_w} f(\mathbf{x}) \, dA \qquad (5)$$

$$\bar{\tilde{F}}_{w_{\max}} = \int_{\Omega} \left(1 - \tilde{\Gamma}(\mathbf{x}, r_{\max}) \right)^{\bar{N}_w} \tilde{f}(\mathbf{x}) \, dA \qquad (6)$$

where $\tilde{\Gamma} = \int_{B(\mathbf{x}_c, r_c)} \tilde{f}(\mathbf{x}) dA$.

For the comparison, the operating domain is divided into the three areas $B(\mathbf{x}_c, r_c)$, $\Omega_B := B(\mathbf{x}_c, r_{\max}) \setminus B(\mathbf{x}_c, r_c)$ and $\Omega_C =: \Omega \setminus \{\Omega_A \cup \Omega_B\}$ as shown in Figure 2.

Without loss of generality, choose a $\delta > 1$ for the transformation such that

$$\delta \left(1 - \int_{B(\mathbf{x}, r_{\max})} f(\mathbf{x}) dA \right) < 1, \forall \mathbf{x} \in \Omega \qquad (7)$$

Such as choice is always possible since f satisfies eq. 3.

The proof is completed by contradiction. Assume $\bar{\tilde{F}}_{w_{\max}} > \bar{F}_{w_{\max}}$, then $\bar{\tilde{F}}_{w_{\max}} - \bar{F}_{w_{\max}} > 0$. As r_c is very small, it follows that $|\Omega| >> |B(\mathbf{x}_c, r_c)|$. Therefore, the contributions to $\bar{\tilde{F}}_{w_{\max}}$ and $\bar{F}_{w_{\max}}$ from Ω_C are equal as in Ω_C it holds that $f \approx \tilde{f}$ and $\Gamma(\mathbf{x}, r_{\max}) \approx \tilde{\Gamma}(\mathbf{x}, r_{\max})$.
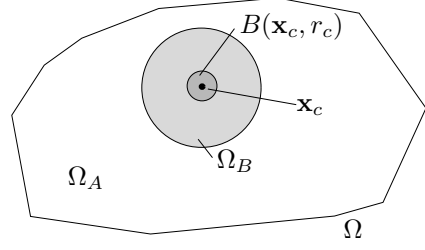

Figure 2. operating domaindivided into the three sets $B(\mathbf{x}_c, r_c), \Omega_B := B(\mathbf{x}_c, r_{\max}) \setminus B(\mathbf{x}_c, r_c)$ and $\Omega_C =: \Omega \setminus \{\Omega_A \cup \Omega_B\}$

In the areas $B(\mathbf{x}_c, r_c)$ and Ω_B, $\tilde{f}$ increases linearly in δ but the terms $\left(1 - \tilde{\Gamma}(\mathbf{x}, r_{\max}) \right)^{\bar{N}_w}$ and $\left(1 - \tilde{\Gamma}(\mathbf{x}, r_{\max}) \right)^{\bar{N}_w}$ decay to zero with higher order in $\bar{N}_w$. Therefore, consider the worst case $\bar{N}_w = 1$. The expression $\bar{\tilde{F}}_{w_{\max}} > \bar{F}_{w_{\max}}$ then reduces to

$$\int_{B(\mathbf{x}_c, r_c)} \left((1 - \tilde{\Gamma})\delta - (1 - \Gamma) \right) f \, dA$$
$$+ \int_{\Omega_B} (\Gamma - \tilde{\Gamma}) f dA > 0$$

Approximate f with the constant value $\bar{f}_B$ in $B(\mathbf{x}_c, r_c)$. Additionally, note that the second integral is strictly negative and the first integral is strictly positive. Thus, the first integral is replaced with the maximum possible value which is attained for $\Gamma, \tilde{\Gamma} \to 0^+$. Denoting the area of $B(\mathbf{x}_c, r_c)$ as B_c, the previous equation further reduces to:

$$(\delta - 1)\bar{f}_B B_c - \delta \bar{f}_B B_c \int_{\Omega_B} f dA > 0 \qquad (8)$$

Due to Condition 3, the term $\bar{f}_B B_c$ is strictly positive and is removed, eq. 8 then can be rearranged to $\delta \left(1 - \int_{\Omega_B} f dA \right) > 1$ which contradicts the choice of δ made in eq. 7. $\blacksquare$

The result allows to formulate a lower bound on the performance of the operational policy for some $w_{\max}$ as a function of the number of robots N:

Theorem 1: *For a given waiting time $w_{\max}$ and any constant distribution f:*

$$\bar{F}_{w_{\max}}(N) \leq \left(1 - \pi(v \; w_{\max})^2 \frac{1}{|\Omega|} \right)^{N - N_{stable} - \bar{N}_r}$$

when using Model-free Anticipative Repositioning.

Proof. We use the formulation from eq. 2 and insert the worst case distribution f using Lemma 3. Then we compute the long-term average number of robots which have reached their wait location $\bar{N}_w$ as the number of robots N subtracting the long-term average number of robots which are busy with servicing a task N_{stable} and the long-term average number of robots which are in a repositioning task $\bar{N}_r$, which yields the claim. $\blacksquare$

The formulation in Theorem 1 illustrates why one possible approach to implement the abstract operations

`sendToRequests(` N_{serve} `)` and
`repositionIdle(` $\Phi(N_{idle})$ `))` of Model-free Anticipative Repositioning is to solve bipartite matching problems. This implementation can be justified as an efficient way to minimize N_{stable} and $\bar{N}_r$. Theorem 1 allows to conveniently determine the median fleet size to have median waiting times below $w_{\max}$ by setting $\epsilon = \frac{1}{2}$:

Corollary 2: *The minimum fleet size to have median waiting time below $w_{\max}$ is*

$$N = \frac{log(\frac{1}{2})}{log\left(1 - \pi(v\, w_{\max})^2 \frac{1}{|\Omega|}\right)} + N_{stable} + \bar{N}_r$$

As an example, consider an area of $10 \times 10\ [km]$ on which robots drive with a speed of $5\ [m/s]$ and and the median time-to-service should be $3\ [min]$. On average, 10 robots are servicing requests and 10 are in a repositioning task. Then the formula yields a minimum fleet size of ≈ 57 robots.

Performance under Time-Varying Demand

In this section, we analyze the case of time-varying demand. The spatial density of origins is now an unknown function of time: $f(\mathbf{x}, t),\ t \geq 0$. Requests arrive at time-varying rate $\lambda(t)$. As in the constant demand case $f(\mathbf{x}, t) = 0\ \forall \mathbf{x} \notin \Omega,\ \forall t$.

First, we show that with reasonable assumptions, the performance of the time-invariant case can be recovered. Then we present an illustrative example in which Model-free Anticipative Repositioning performs almost optimally while existing methods designed for time-invariant demand, e.g., the algorithms proposed in Arsie et al. (2009b) provide less performance. Finally, we present an example to demonstrate the rapid increase in performance with every robot added to the fleet.

Recovering the Time-Invariant Case First, we introduce a definition to quantify the change in demand with respect to time based on Lipschitz continuity:

Definition 7: *A demand distribution $f(\mathbf{x}, t)$ is Lipschitz continuous with respect to time with constant $\gamma \geq 0$ if $|f(\mathbf{x}, t_1) - f(\mathbf{x}, t_2)| \leq \gamma |t_1 - t_2|,\ \forall t_1, t_2 \geq 0, \forall \mathbf{x} \in \Omega$.*

We are now able to state the conditions under which Model-free Anticipative Repositioning recovers the performance guarantees of the time-invariant case under time-varying demand.

Lemma 4: *Let r be a request arriving at time t^* under time-varying demand $f(\mathbf{x}, t)$ and w_r its wait time. The minimum arrival rate in the temporal neighborhood of t^* is λ and the time between arrivals is exponentially distributed. Let s be a request arriving under time-invariant demand $f(\mathbf{x}, t^*)$ and w_s its wait time. Then, for an identical number of waiting robots $N_w(t^*)$,*

$$Pr(w_r > w_{\max}) \leq Pr(w_s > w_{\max}) + \tau c$$

where c is a constant depending only on $r_{\max}$ and $f(\mathbf{x}, t^*)$ and τ has expected value $\frac{\gamma}{\lambda}$ and variance $\frac{\gamma}{\lambda^2}$.

Proof. From the proof of Lemma 2, recall the formula for $Pr(w_s > w_{\max})$:

$$Pr(w_s > w_{\max}) = \int_{\Omega} (1 - \Gamma(\mathbf{x}, r_{\max}))^{N_w(t^*)} f(\mathbf{x}, t^*)\, dA$$

For request r arriving under time-varying demand, we assume that the $N_w(t^*)$ relevant requests arrived at times $t_1, t_2, ..., t_{N_w(t)} < t^*$. Using the property that the demand distribution f is Lipschitz continuous, we can then reformulate $Pr(w_r > w_{\max})$. The arguments of f are omitted for the sake of brevity:

$$Pr(w_r > w_{\max})$$
$$= \int_{\Omega} \prod_{k=1}^{N_w(t^*)} (1 - \Gamma(\mathbf{x}_k, r_{\max}, t_k))\, f\, dA$$
$$\leq \int_{\Omega} \prod_{k=1}^{N_w(t)} \left(1 - \left((\Gamma(\mathbf{x}_k, r_{\max}, t) - r_{\max}^2 \pi \gamma(t^* - t_k))\right)\right) f\, dA$$
$$\leq \int_{\Omega} \left(1 - (\Gamma(\mathbf{x}, r_{\max}, t) - r_{\max}^2 \pi \gamma(t^* - t_1))\right)^{N_w(t^*)} f\, dA$$
$$\leq \int_{\Omega} \left(1 - \Gamma(\mathbf{x}, r_{\max}, t)\right)^{N_w(t^*)} + \frac{r_{\max}^2 \pi \gamma(t^* - t_1)}{\Gamma(\mathbf{x}, r_{\max}, t)}\right) f\, dA$$
$$= Pr(w_s > w_{\max}) + \gamma(t^* - t_1) \int_{\Omega} \frac{r_{\max}^2 \pi}{\Gamma(\mathbf{x}, r_{\max}, t)} f(\mathbf{x}, t^*)\, dA$$
$$= Pr(w_s > w_{\max}) + \tau c$$

The first inequality was obtained by using the property of Lipschitz-continuity, the second by taking the largest time interval and the third by introducing an upper bound for the polynomial valid for $N_w(t^*) \geq 1$ and $\Gamma \leq 1$. $\tau := \gamma(t - t_1)$ is the highest possible rate of request density change multiplied by the time interval between the current time and he arrival time of the earliest request that lead to a robot reposition command. As time intervals were assumed to be exponentially distributed with rate λ:

$$\text{Exp}[\tau] = \frac{\gamma}{\lambda} \qquad\qquad \text{Var}[\tau] = \frac{\gamma}{\lambda^2}$$

$\blacksquare$

Corollary 3: *In the limit $\frac{\gamma}{\lambda} \to 0^+$, the performance $\bar{F}_{w_{\max}}$ of Model-free Anticipative Repositioning is independent of spatial or temporal changes in the distribution of service requests.*

This result demonstrates that for $\frac{\gamma}{\lambda} \to 0^+$ the introduced operational policy offers the same performance guarantees for time-varying demand as for time invariant demand despite of its simplicity. Especially, it recovers the results presented in Lemma 2, Corollary 1, Lemma 3 and Theorem 1 as well as the results regarding fleet sizing in Corollary 2.

An important aspect of this result is that performance guarantees can be recovered even for fast changing demands as long as the arrival rate is high enough. Therefore, Model-free Anticipative Repositioning is able to cover most multi-robot systems operating under time-varying demand as many operate under a large number of service requests. Consider for instance a mobility-on-demand system, e.g., a conventional taxi system. The 536 San Francisco taxis for which traces were recorded and published in Piorkowski et al. (2009) have more than $20,000$ requests on most days, resulting in an average arrival rate of $0.23\ Hz$. The typical

onset period of a demand peak is not less than 30 min which results in the ratio $\frac{\gamma}{\lambda^-} \approx 0.0024$. As shown in Section Numerical Verification such a ratio is sufficiently small to offer the same performance levels as in a time-invariant case.

Example: Time-Varying Demand

For situations with reasonably smooth changes in the demand profile or smaller inter-arrival periods, many examples can be found that illustrate the advantages of the proposed method. Consider Ω to be a ball in which an invisible request generator moves with small speed ν on a random walk. It generates one request at random intervals with time-varying rate $\lambda(t)$ that are served by the robot on spot with small service time. In such a case, the robot will arrive at every request within time $\frac{\nu dt}{v} \approx 0$. The policies presented in Arsie et al. (2009b) will converge to a waiting position in the middle of the ball and wait for new request there. On average, they will thus take $\frac{r}{2v}$ of time to reach a request. The situation is illustrated in Figure 3.
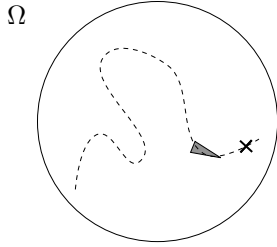

Figure 3. A scenario in which small wait times are achieved under time-varying demand: requests are generated by a generator moving on a random walk in a circular operational domain Ω.

Example: Fleet Sizing

As an example to demonstrate how quickly additional robots in the fleet improve performance, consider the case $N = 1$ with only one robot and the series of requests $\{r_p, r_q, r_p, r_q, ...\}$ arriving at p, q respectively with very long intervals between each of the requests. $p, q \in \Omega$ are chosen such that $||p - q|| = \max_{s,t \in \Omega} ||s - t||$. Using the presented operational policy, a request will be served at p and then position the idle robot at p, thus at the most distant location to the next request q in the operational domain. The setting is illustrated in Figure 4. However, adding a second robot would already recover near-optimal service levels.
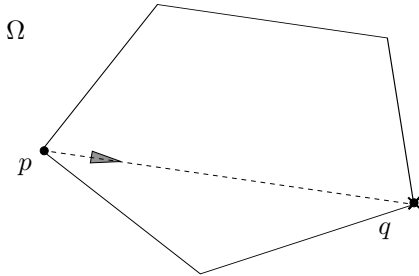

Figure 4. An example demonstrating the rapid increase of performance with additional robots: requests arrive at p and q in a sequence with long time intervals between arrivals. One single robot will always wait at the point where the request does not arrive, two robots will drive the wait times to zero.

Performance under Correlated Demand

In literature, the arrivals are typically considered identically and independently distributed (i.i.d), e.g., in Treleaven et al. (2013); Arsie et al. (2009b); Bullo et al. (2011). This viewpoint covers many real-world problems and facilitates analysis, but it excludes the important case of correlated samples. Specifically, the case when the realization of a certain arrival at some location $\mathbf{x}$ increases the probability of an arrival at the same location within a limited amount of time. In real-world system instances, specifically in the case of a N-DVRP, this will occur with high frequency. Take for example the case of a mobility-on-demand transportation system. If a transportation request at location $\mathbf{x}$ is caused by a customer who is leaving an event, e.g., a dinner party, then it is likely that another transportation request with the same origin will occur within a short amount of time. The realization of the first request increases the probability of a realization at the same location. Interestingly, the operational policy presented in this work offers better performance with increasing levels of correlation in the demand generating process. In this subsection, we present a brief analysis of the case of correlated demand. Specifically, we introduce a general formulation to model correlated demand. Then, we show that our policy increases the service level in the system with rising degree of correlation in the stochastic process that generates the arrivals.

The correlated demand model samples requests from a service request distribution or from past realizations depending on a parameter ω. Specifically, the service request is generated from some time-varying distribution $f(\mathbf{x}, t)$ as in the previously analyzed cases with a probability $\omega \in (0, 1)$. With probability $1 - \omega$ the arrival generated at one of the past M request locations with uniform distribution. The number M is an integer for which $M > N$. In this model, the probability ω is the likelihood of drawing a sample independent from previous realizations.

The comparison of correlated and uncorrelated demand is based on the following assumptions and notation:

- $f(\mathbf{x}, t)$ is Lipschitz continuous with respect to time with constant γ. λ^- is the smallest arrival rate in the period of interest with $\frac{\gamma}{\lambda^-}$ being small.
- $\bar{F}_{w_{max}}$ is the long-term fraction of requests waiting longer than $w_{\max}$ for some time-varying distribution of arrivals $f(\mathbf{x}, t)$ for $\omega \to 1^-$
- $\bar{G}_{w_{\max}}(\omega)$ is the long-term fraction of requests waiting longer than $w_{\max}$ for the same distribution $f(\mathbf{x}, t)$ and $\omega \in (0, 1)$.

The first result demonstrates that such a setting of correlated arrivals will not decrease the performance of the operational policy.

Lemma 5: $\bar{G}_{w_{\max}}(\omega) \leq \bar{F}_{w_{\max}}$

Proof. By Lemma 4, $\bar{F}_{w_{\max}}$ is approximately equal for both the time-invariant and the time-varying system. By Lemma 2 $\bar{F}_{w_{\max}}$ depends only on the past $\bar{N}_w$ realizations of the arrival process, but not on their order. As realizations characterizing $\bar{G}_{w_{\max}}(\omega)$ and $\bar{F}_{w_{\max}}$ are identically distributed, the claim follows. ∎

The results shows that correlated demands will not *decrease* the service level of the system in any case. However, there are many cases in which the level of correlation in the demand *increases* the service level of the system as the next result shows. We consider the limit cases of stringent wait time requirements $w_{\max}$, small robot speeds and of a large operating domain:

Lemma 6: *Let $\bar{N}_w \geq 1$. Then, in the limit cases*

1. $w_{\max} \to 0^+$
2. $v \to 0^+$
3. $|\Omega| \to \infty$

the waiting times under correlated demand are strictly smaller than under uncorrelated demand, i.e., $\bar{G}_{w_{\max}}(\omega) < \bar{F}_{w_{\max}}$.

Proof. By Theorem 1 $\bar{F}_{w_{\max}}$ is bounded by $\left(1 - \pi(vw_{\max})^2 \frac{1}{|\Omega|}\right)^{\bar{N}_w}$. As $\bar{G}_{w_{\max}}(\omega)$ is composed by a combination of independent samples and correlated samples, $\bar{G}_{w_{\max}}(\omega)$ can be bounded by

$$\bar{G}_{w_{\max}}(\omega) \leq \omega \left(1 - \pi(vw_{\max})^2 \frac{1}{|\Omega|}\right)^{\bar{N}_w} + (1-\omega)\left(1 - \frac{\bar{N}_w}{M}\right)$$

The expression $\bar{G}_{w_{\max}}(\omega) \leq \bar{F}_{w_{\max}}$ can thus be rewritten:

$$1 - \frac{\bar{N}_w}{M} \leq \left(1 - \pi(vw_{\max})^2 \frac{1}{|\Omega|}\right)^{\bar{N}_w}$$

In the cases $w_{\max} \to 0^+, v \to 0^+$ and $|\Omega| \to \infty$, the limit of the expression on the right hand side is one. As $\bar{N}_w \leq N < M$, the left-hand expression is smaller than one and $\bar{G}_{w_{\max}}(\omega) < \bar{F}_{w_{\max}}$. ∎

The analysis shows that under correlated demand, Model-free Anticipative Repositioning performs equally well as under i.i.d. arrivals. Furthermore, the fraction of very small wait times will be increased with higher degrees of correlation. Finally, for very large operational domains, e.g., a mobility-on-demand problem in a rural area, correlated demand will also lead to improved service levels of the system. This is also the case for small maximum wait times or small velocities. Interestingly, these are exactly the cases in which limited data is available to implement model-based policies and in which a system designer may have to rely on model-free policies.

Numerical Verification

The previous theoretical results were verified numerically. To do so, an instance of the N-DTRP was implemented in software and Model-free Anticipative Repositioning was used to control the fleet of robots. The N-DTRP was executed in a unit square operational domain $1[\mathrm{m}] \times 1[\mathrm{m}]$ with a constant service time of $10[\mathrm{s}]$. Requests are generated by a homogeneous Poisson process and constant arrival rate of $0.005 \left[\frac{1}{\mathrm{s}}\right]$. Robots move freely at speed $v = 0.05 \left[\frac{\mathrm{m}}{\mathrm{s}}\right]$.

Worst-Case Performance

In the first trial, the fraction of requests receiving service after a period longer than $w_{\max} = 5[\mathrm{s}]$ denoted $\bar{F}_{w_{\max}}$ was determined for a total of 3,000 uniformly distributed service arrivals. The results were compared to the predicted lower bound on performance stated by Theorem 1. As it can be seen in Figure 5, the theoretical bounds are confirmed by the trial. The gap between the theoretical bound and the simulation at $N = 1$ is given by $\frac{(w_{\max} v)^2 \pi}{|\Omega|}$ and increases with larger $|\Omega|$.
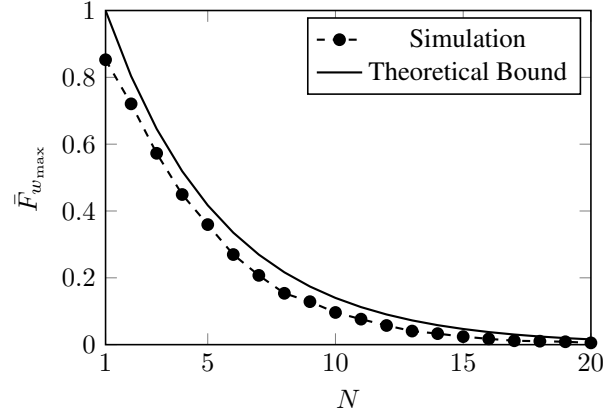
Figure 5. Comparison of simulated values $\bar{F}_{w_{\max}}(N)$ and theoretical bounds determined in Theorem 1 for Model-free Anticipative Repositioning.

Uniformity of Request Distributions

According to the result provided in Lemma 3, the case of a uniform distribution of requests leads to the worst possible performance of the introduced operational policy. This relation, proven in theory, is confirmed by analyzing the results for arrival distributions on a discrete set of points. In the trial, a discrete set of random points p in the operation domain of size $M = \{2, 4, 8, 16, 32\}$ was chosen and to each of them a random probability p_i was assigned such that $\sum_{i=1}^{M} p_i = 1$. Then, a total of 5,000 arrivals was sampled on the distribution and the N-DTRP was solved using Model-free Anticipative Repositioning. The results are shown in Figure 6. They confirm the theoretical statement that the results improve with increasing non-uniformity of the distribution.

Time-Varying Demand

Next, the results presented in Lemma 4 were verified. For this, a trapezoidal distribution of arrivals in one dimension of the unit square operational domain was chosen. The distribution is constant in the other dimension. Then, this trapezoidal distribution was moved back and forth with velocity ν on the unit square resulting in a time-varying request density. If the slope of the trapezoid distribution is chosen to be η, then the change of the distribution with respect to time for any given point is given by $\gamma \in \{-\nu\eta, 0, \nu\eta\}$ and is thus Lipschitz continuous with respect to time with constant $\gamma = \nu\eta$. The setting is illustrated in Figure 7. In the trial, a total of 6,000 requests were served by a fleet of 20 robots, $w_{\max}$ was chosen to be $2[\mathrm{s}]$. The trapezoidal distribution was reduced to a triangle with base
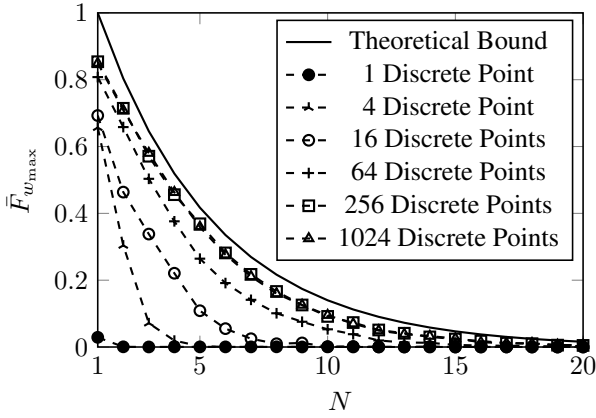
Figure 6. Comparison of simulated values $\bar{F}_{w_{\max}}(N)$ for discrete distributions on sets of points of different cardinality and theoretical bounds determined in Theorem 1 for Model-free Anticipative Repositioning.
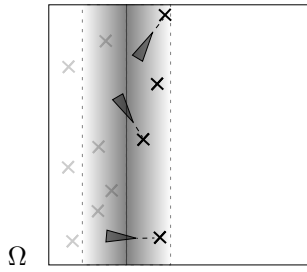

Figure 7. Visualization of the trapezoidal case to demonstrate the performance of Model-free Anticipative Repositioning in the setting of time-varying demand.
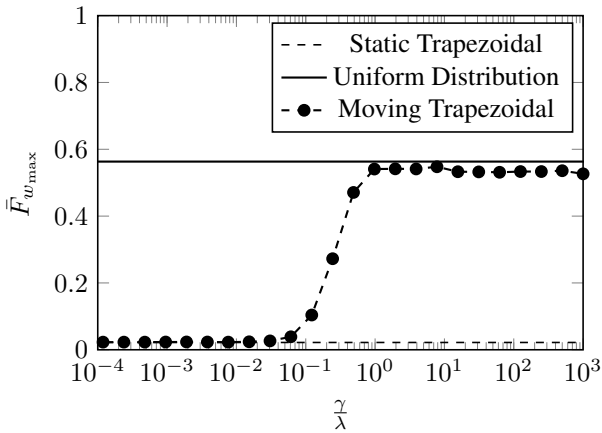

Figure 8. Comparison of simulated values $\bar{F}_{w_{\max}}(N)$ for a uniform distribution, a static trapezoidal distribution and a moving trapezoidal distribution for Model-free Anticipative Repositioning.

length of $0.1[m]$. The arrival rate was fixed at $0.005[\frac{1}{s}]$ and ν as varied in the range $[10^{-10}, 0.5]\ [\frac{m}{s}]$, i.e., at the maximum speed, the period of the time-varying distribution is $4[s]$. The results are shown in Figure 8. It can be seen that for ratios $\frac{\gamma}{\lambda}$ up to ≈ 0.05, the performance of the under time-varying demand is identical as for a static trapezoidal distribution. For faster velocities ν, $\bar{F}_{w_{\max}}$ increases and finally reaches the value of the uniform distribution.

Correlated Demand

Finally, the results presented in Subsection Performance under Correlated Demand are verified. As described in the derivations, a request distribution is considered in which the arrivals are sampled from some distribution $f(\mathbf{x})$ with probability $\omega \in (0, 1]$ and are sampled uniformly from the last $M \in \mathbb{M}$ realizations of the stochastic process otherwise. Figure 9 shows the influence of ω on the result for $M = 30$ for 6,000 service requests. The uncorrelated requests are drawn from a uniform distribution on a unit square. The service level of the system increases with decreasing ω and it always stays below the service level of the nonuniform case.
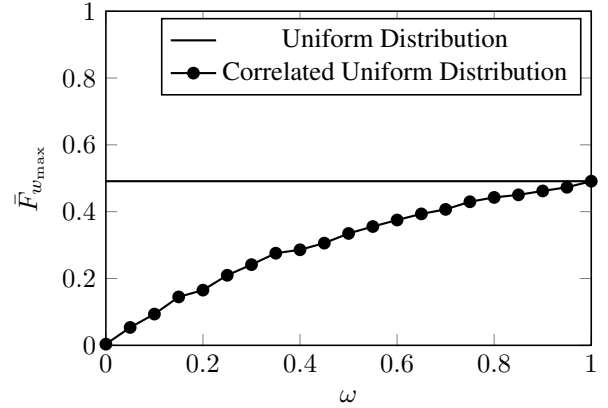

Figure 9. Comparison of simulated values $\bar{F}_{w_{\max}}(N)$ for a uniform distribution with different levels of correlation ω for Model-free Anticipative Repositioning.

In addition to the numerical verification provided in this section, a study of a possible application of the proposed policy to the problem of mobility-on-demand is presented in the next section.

Application to Mobility-on-Demand

In a one-way mobility-on-demand (MoD) system, customers request transportation from their current position to some destination and are served in an on-demand fashion. An example of an existing MoD system is a conventional taxi system or a modern ride-hailing service, e.g., Uber or Lyft. Although these systems play an important role in today's transportation infrastructure, their significance in comparison to public transportation (buses, trains, subways etc.) and the privately owned motorized vehicle is limited. As an example, the number of active Uber drivers in the United States end of 2015 was less than 500,000 as described in Hall and Krueger (2018). Also, car-sharing schemes where users drive by themselves are limited in size. As an example, in the period 2006 – 2010, the estimated worldwide fleetsize of shared vehicles was estimated to be 33,665 by Shaheen and Cohen (2013). A negligible number compared to the total number of private vehicles which exceeded 1 billion in 2011 according to Sousanis (2011).

The availability and legal acceptance of fully self-driving vehicles would change these dynamics in three principal ways. First, the costs and limitations of human drivers, e.g., resting hours, would vanish. Second, the systems could be centrally coordinated and controlled as there is

no more autonomy of the driver, this allows to increase system efficiency as shown in Ruch and Frazzoli (2019). Finally, true two-way mobility can be offered as vehicles can be repositioned efficiently to meet future anticipated demand. For these reasons, with the fully autonomous vehicle, a new mode of transportation called autonomous mobility-on-demand (AMoD) would emerge. This novel mode of transportation would include the benefits of public transportation such as a lower carbon footprint, affordable cost and the high share of variable cost for the user. On the other hand, it would also offer the main benefits of private transportation, namely on-demand point-to-point transportation in a private vehicle without taking into account schedules. Prices of an AMoD system are estimated to be almost as low as the ones of public transport as shown in Hörl et al. (2017), therefore it is possible that AMoD systems will take a mode share possibly as large as the one of private vehicles today, e.g., between 52% and 66% in Germany and between 83% and 89% in the United States as estimated in Buehler (2011).

Such large fleets of vehicles will need to be operated as efficiently as possible while offering the highest possible service level to users. Appropriate design, e.g., utilizing the right fleet sizes is important to achieve an optimal trade-off between costs, emissions and service level Spieser et al. (2014). Apart from such static design decisions, most gains can be achieved by implementing the best operational policies that guide the behavior of the fleet. An operational policy must take the decision which robotic taxi serves an upcoming request and it must handle the problem of unbalancedness. MoD system have been shown to become unbalanced due to the spatio-temoral variations in the demand profile, see, e.g., Pavone et al. (2011c). This means that some locations get depleted of vehicles, while others are overfilled and the service level is impacted negatively. An operational policy must solve this issue by replacing an appropriate choice and amount of vehicles. The problem of designing operational policies for MoD systems to address these issues has been a subject of recent research. Some solutions are proposed based on network flow optimization, e.g., Pavone et al. (2011c); Zhang et al. (2015); Salazar et al. (2018). Others approaches apply model predictive control (Morari and Lee (1999)) methods, e.g., Zhang et al. (2016); Iglesias et al. (2018); Seccamonte (2019). Finally, some methods also rely on heuristics, e.g., Bischoff and Maciejewski (2016) and most recently also end-to-end reinforcement learning has been proposed in Fluri Christian (2019).

The operational policy proposed in this work is designed to solve the N-DTRP and the N-DVRP problems. Moreover, it is suitable across all load factors and for highly time-varying settings, which makes its application to MoD fleet control problems very promising. Therefore, we apply the policy to the problem of mobility-on-demand fleet control to demonstrate its effectiveness to non-synthetic problem instances. Using the high-fidelity transportation simulator AMoDeus (Ruch and Frazzoli (2018)), we are able to show that the proposed policy outperforms benchmark policies for the mobility-on-demand fleet management problem. It results in very high service levels across all load factors. The remainder of this section is organized as follows, in Subsection Compared Operational Policies the operational policies to be compared to are briefly outlined, then in Subsection Simulation Environment the setup of the simulations is described and finally, the results are presented in Subsection Results.

Compared Operational Policies

The following benchmark operational policies for fleet control of mobility-on-demand systems were compared to Model-free Anticipative Repositioning denoted (MFAR).

Adaptive Real-Time Rebalancing Policy (ARRP): This policy is an implementation of the results published in Pavone et al. (2011c). The city is divided into a set of virtual stations on which a complete graph is defined. A linear program formulation is used to minimize the total rebalancing distance that is necessary to reach an equilibrium for each time step. As the linear program exhibits a totally unimodular constraint matrix, the result is integral. Therefore, the problem directly delivers the number of rebalancing vehicles between any pair of virtual stations in the city and takes into account the number of unserved requests and the number of arriving customer vehicles.

Global Bipartite Matching Policy (GBM): This operational policy was described in Ruch and Frazzoli (2018). It continuously solves a bipartite matching problem between the locations of open requests and of available vehicles. The cost assigned to the edges is either Euclidean or network distance. This policy does not include any rebalancing, i.e., vehicles are only moved once requests are open.

No Communications Policy without Sensor Information (NC1): In this policy presented in Arsie et al. (2009a), every vehicle moves towards the closest open request when idle. If there are no open requests, it moves to the location minimizing the average distance to locations at which it served requests in the past. The policy is decentralized and is demonstrated to converge to a saddle point of the cost function for the light-load under time-invariant demand. While convergence to a local minimum is not guaranteed, it is obtained in many cases Arsie et al. (2009a).

No Communications Policy with Limited Sensor Information (NC2): In this adapted version of NC1 also presented in Arsie et al. (2009a), the vehicles only move towards the open request locations in the Voronoi cell generated by their position with respect to the position of other vehicles. This adaption requires local sensor information but limits the amount of empty distance driven by the vehicles.

Feedforward Fluidic Time-Varying Rebalancing Policy (FFTVRP): This policy presented in Spieser et al. (2016) represents a generalization of the policy presented in Pavone et al. (2011c) to time-varying systems. The policy computes the minimal cost dispatching using a linear program generated with a time-varying network flow problem. The implementation requires a model of hourly request arrivals, in this comparison perfect knowledge of the demand distribution was assumed.

Simulation Environment

For comparison of the policies, the simulation environment AMoDeus Ruch and Frazzoli (2018) was used. AMoDeus is a high-fidelity traffic simulator based on MATSim

(Horni et al. (2016)) specifically developed to test fleet control policies and mobility-on-demand system instances. AMoDeus is open-source and publicly available, furthermore it contains benchmark transportation scenarios which are available online. The scenario of San Francisco was chosen to conduct the simulations presented in this work. It is based on a dataset of taxi traces in the city of San Francisco (Piorkowski et al. (2009)) that were used to generate a realistic mobility-on-demand customer request profile. The dataset contains a total of 464,045 customer trips, which were recorded between May 17, 2008, 03:00:04 to June 10 2008, 02:25:34. For this study, the requests on Friday, May 30 were chosen, totally 22,479. The simulation takes place on a detailed city network covering the entire area of San Francisco.

Results

Three main system characteristics are needed to quantitatively analyze a mobility-on-demand system. First, the service level offered to customers, mainly characterized by the statistical distribution of wait times. Depending on the case, an operator might decide to minimize mean wait times or possibly minimize high wait times, e.g., the 95% quantile. Of course, larger fleets will generally have more available vehicles closer to the locations of upcoming requests. Therefore, from a service level viewpoint, higher fleet sizes are desirable. But these come at the cost of increased capital expenditures. Therefore, they have to be considered as a variable in the performance analysis. Finally, the vehicles will cover some distance without a passenger on board which contributes significantly to operational expenditures. Empty distance cannot be reduced to zero as shown in Treleaven et al. (2013). Furthermore, covering empty distance might lead to an increased service level. Therefore, the empty distance covered by a fleet of vehicles also has to be considered in the analysis. A visual comparison of different operational strategies in the evening at 23:00 pm is shown in Figure 10. It can be seen that the number of open requests for our suggested operational policy is visibly smaller than for other operational policies. While this figure already clearly demonstrates the differences of different operational policies, a view on the aggregate values of different fleet sizes gives a clearer impression of the difference in performance. Figure 11 shows the mean wait time for different operational policies. The 95% quantile wait times are shown in Figure 12. It can be seen that Model-free Anticipative Repositioning clearly outperforms existing strategies across the entire system load range, also the ones which use past data in an internal model. The total empty distances are shown in Figure 13. As it can be seen, the distances covered by the fleet when using Model-free Anticipative Repositioning are in the same range as for other policies.

Conclusions and Future Work

In this work, we have presented a novel operational policy to solve the N-Vehicle Dynamic Travelling Repairman Problem and the N-Vehicle Dynamic Travelling Vehicle Routing Problem. Our proposed policy is model-free, adaptive to different load factors and adaptive to time-varying demand. We analytically derived performance guarantees for the case
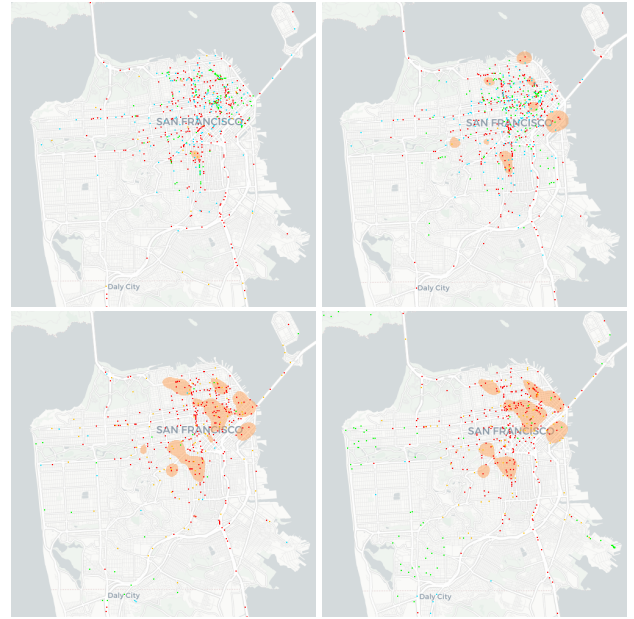

Figure 10. Visualizations of the AMoD system at 23:00 pm, the orange heatmap shows outstanding requests. Vehicles are represented by small dots. Operational policies: MFAR (top left), NC2 (top right), FFTVRP (bottom left), ARRP (bottom right).
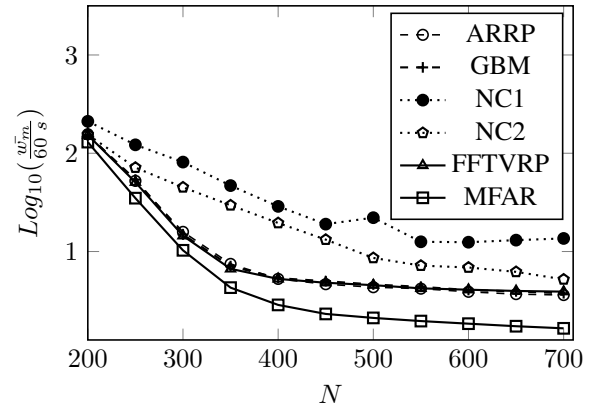

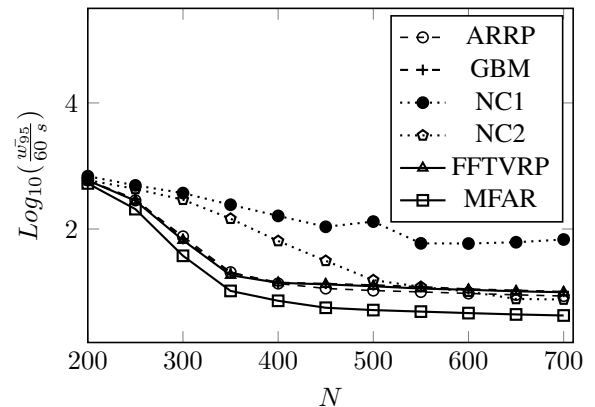**Figure 11.** Mean wait time of different fleet control policies.

Figure 12. 95% quantile wait time of different fleet control policies.

of time-invariant demand. Then, we demonstrated that these performance guarantees also apply to the time-varying case if the arrival rate of new requests is larger than the time constant defining the change in the demand distribution with
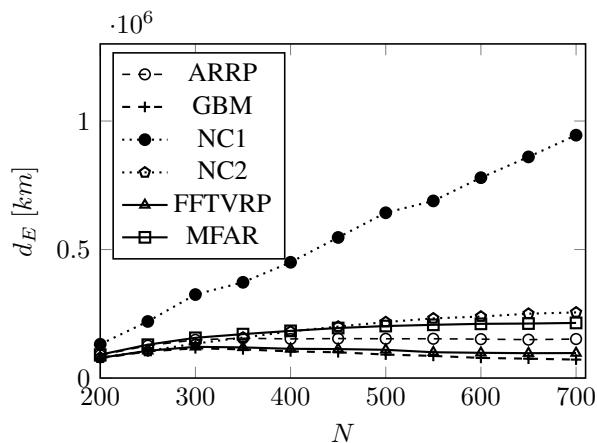
Figure 13. Empty Distances of different fleet control policies.

respect to time. Finally, we demonstrated that for the case of service requests which are not independently distributed, the operational policy performs better with higher degrees of correlation between subsequent requests. We provided numerical confirmation of our derivations. Finally, we argued that the conditions under which our policy performs well are encountered in many real-world problem instances, especially in mobility-on-demand systems. We use our operational policies to control a fleet of autonomous robotic taxis in a mobility-on-demand system on the simulation platform AMoDeus and show that the policy outperforms both model-free and model-based existing approaches in terms of service level across the entire range of system loads. As future work, we plan to develop a distributed version of the policy, to modify the current version of it to reduce the empty distance covered and to apply it to a mobility-on-demand system instance in a non-simulation environment. We also plan to extend the results such that the distribution of destinations can be used to further increase the performance in the case of the N-DVRP. Finally, we will look for related applications in other domains not directly related to multi-robot systems.

References

Agarwal P and Varadarajan K (2004) A near-linear constant-factor approximation for euclidean bipartite matching? In: *Proceedings of the twentieth annual symposium on Computational geometry*. ACM, pp. 247–252.

Arsie A, Savla K and Frazzoli E (2009a) Efficient routing algorithms for multiple vehicles with no explicit communications. *IEEE Transactions on Automatic Control* 54(10): 2302–2317.

Arsie A, Savla K and Frazzoli E (2009b) Efficient routing algorithms for multiple vehicles with no explicit communications. *IEEE Transactions on Automatic Control* 54(10): 2302–2317. DOI:10.1109/TAC.2009.2028954.

Bischoff J and Maciejewski M (2016) Simulation of city-wide replacement of private cars with autonomous taxis in berlin. *Procedia computer science* 83: 237–244.

Browning B, Bruce J, Bowling M and Veloso M (2005) Stp: Skills, tactics, and plays for multi-robot control in adversarial environments. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering* 219(1): 33–52.

Buehler R (2011) Determinants of transport mode choice: a comparison of germany and the usa. *Journal of Transport Geography* 19(4): 644–657.

Bullo F, Frazzoli E, Pavone M, Savla K and Smith SL (2011) Dynamic vehicle routing for robotic systems. *Proceedings of the IEEE* 99(9): 1482–1504.

Cortes J, Martinez S, Karatas T and Bullo F (2004) Coverage control for mobile sensing networks. *IEEE Transactions on robotics and Automation* 20(2): 243–255.

Fluri Christian CRea (2019) Learning to operate a fleet of cars. *Under Review* .

Frazzoli E and Bullo F (2004) Decentralized algorithms for vehicle routing in a stochastic time-varying environment. In: *Decision and Control, 2004. CDC. 43rd IEEE Conference on*, volume 4. IEEE, pp. 3357–3363.

Gendreau M and Potvin JY (1998) Dynamic vehicle routing and dispatching. In: *Fleet management and logistics*. Springer, pp. 115–126.

Hall JV and Krueger AB (2018) An analysis of the labor market for ubers driver-partners in the united states. *ILR Review* 71(3): 705–732.

Hauri S, Alonso-Mora J, Breitenmoser A, Siegwart R and Beardsley P (2014) Multi-robot formation control via a real-time drawing interface. In: *Field and service robotics*. Springer, pp. 175–189.

Hörl S, Ruch C, Becker F, Frazzoli E and Axhausen KW (2017) Fleet control algorithms for automated mobility: A simulation assessment for zurich. *Arbeitsberichte Verkehrs-und Raumplanung* 1270.

Horni A, Nagel K and Axhausen KW (2016) *The multi-agent transport simulation MATSim*. Ubiquity Press London.

Iglesias R, Rossi F, Wang K, Hallac D, Leskovec J and Pavone M (2018) Data-driven model predictive control of autonomous mobility-on-demand systems. In: *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 1–7.

Kuhn HW (1955) The hungarian method for the assignment problem. *Naval research logistics quarterly* 2(1-2): 83–97.

Le Ny J and Pappas GJ (2010) Adaptive robot deployment algorithms .

Le Ny J and Pappas GJ (2013) Adaptive deployment of mobile robotic networks. *IEEE Transactions on automatic control* 58(3): 654–666.

Lee SG, Diaz-Mercado Y and Egerstedt M (2015) Multirobot control using time-varying density functions. *IEEE Transactions on Robotics* 31(2): 489–493.

Leitner J (2009) Multi-robot cooperation in space: A survey. In: *Advanced Technologies for Enhanced Quality of Life, 2009. AT-EQUAL'09*. IEEE, pp. 144–151.

Miao F, Lin S, Munir S, Stankovic JA, Huang H, Zhang D, He T and Pappas GJ (2015) Taxi dispatch with real-time sensing data in metropolitan areas: A receding horizon control approach. In: *Proceedings of the ACM/IEEE Sixth International Conference on Cyber-Physical Systems*. ACM, pp. 100–109.

Morari M and Lee JH (1999) Model predictive control: past, present and future. *Computers & Chemical Engineering* 23(4-5): 667–682.

Pavone M, Arsie A, Frazzoli E and Bullo F (2011a) Distributed algorithms for environment partitioning in mobile robotic networks. *IEEE Transactions on Automatic Control* 56(8):

1834–1848.

Pavone M, Frazzoli E and Bullo F (2011b) Adaptive and distributed algorithms for vehicle routing in a stochastic and dynamic environment. *IEEE Transactions on Automatic Control* 56(6): 1259–1274.

Pavone M, Smith S, Frazzoli E and Rus D (2011c) Load balancing for Mobility-on-Demand systems. In: *Robotics: Science and Systems VII*.

Pavone M, Treleaven K and Frazzoli E (2010) Fundamental performance limits and efficient polices for Transportation-On-Demand systems. In: *49th IEEE Conf. on Decision and Control*. pp. 5622–5629.

Piorkowski M, Sarafijanovic-Djukic N and Grossglauser M (2009) A parsimonious model of mobile partitioned networks with clustering. In: *Communication Systems and Networks and Workshops, 2009. COMSNETS 2009. First International*. IEEE, pp. 1–10.

Popa DO, Sanderson AC, Komerska RJ, Mupparapu SS, Blidberg DR and Chappel SG (2004) Adaptive sampling algorithms for multiple autonomous underwater vehicles. In: *Autonomous Underwater Vehicles, 2004 IEEE/OES*. IEEE, pp. 108–118.

Ritzinger U, Puchinger J and Hartl RF (2016) A survey on dynamic and stochastic vehicle routing problems. *International Journal of Production Research* 54(1): 215–231.

Ruch H and Frazzoli (2018) Amodeus, a simulation-based testbed for autonomous mobility-on-demand systems. In: *Proc. 21th IEEE Conf. Intelligent Transportation Systems*.

Ruch S and Frazzoli (2019) The price of anarchy in mobility-on-demand systems. *Under Review* .

Salazar M, Rossi F, Schiffer M, Onder CH and Pavone M (2018) On the interaction between autonomous mobility-on-demand and public transportation systems. *arXiv preprint arXiv:1804.11278* .

Savla K, Bullo F and Frazzoli E (2007) The coverage problem for loitering dubins vehicles. In: *Decision and Control, 2007 46th IEEE Conference on*. IEEE, pp. 1398–1403.

Savla K and Frazzoli E (2010) Game-theoretic learning algorithm for a spatial coverage problem. Institute of Electrical and Electronics Engineers.

Schwager M, Bullo F, Skelly D and Rus D (2008) A ladybug exploration strategy for distributed adaptive coverage control. In: *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*. Citeseer, pp. 2346–2353.

Seccamonte F (2019) Decentralized model predictive control for amod. *Under Review* .

Shaheen SA and Cohen AP (2013) Carsharing and personal vehicle services: worldwide market developments and emerging trends. *International Journal of Sustainable Transportation* 7(1): 5–34.

Sousanis J (2011) World vehicle population tops 1 billion units. *Wards Auto* 15.

Spieser K, Samaranayake S and Frazzoli E (2016) Vehicle routing for shared-mobility systems with time-varying demand. In: *American Control Conference (ACC), 2016*. IEEE, pp. 796–802.

Spieser K, Treleaven K, Zhang R, Frazzoli E, Morton D and Pavone M (2014) Toward a systematic approach to the design and evaluation of automated mobility-on-demand systems: A case study in singapore. In: *Road vehicle automation*. Springer, pp.

229–245.

Sutantyo D, Levi P, Möslinger C and Read M (2013) Collective-adaptive lévy flight for underwater multi-robot exploration. In: *Mechatronics and Automation (ICMA), 2013 IEEE International Conference on*. IEEE, pp. 456–462.

Treleaven K, Pavone M and Frazzoli E (2013) Asymptotically optimal algorithms for One-to-One pickup and delivery problems with applications to transportation systems. *IEEE Trans. Automat. Contr.* 58(9): 2261–2276.

Waisanen HA, Shah D and Dahleh MA (2008) A dynamic pickup and delivery problem in mobile networks under information constraints. *IEEE Transactions on Automatic Control* 53(6): 1419–1433.

Zhang R, Rossi F and Pavone M (2016) Model predictive control of autonomous mobility-on-demand systems. In: *2016 IEEE international conference on robotics and automation (ICRA)*. IEEE, pp. 1382–1389.

Zhang R, Spieser K, Frazzoli E and Pavone M (2015) Models, algorithms, and evaluation for autonomous mobility-on-demand systems. In: *American Control Conference (ACC), 2015*. IEEE, pp. 2573–2587.