The Dynamic Traveling Repair Problem: Examination of an Asymptotically Optimal Algorithm

UCI-ITS-LI-WP-01-2

Xiangwen Lu ¹
Amelia C. Regan ²
Sandra Irani ³

¹ Institute of Transportation Studies University of California, Irvine

² Department of Civil and Environmental Engineering and Institute of Transportation Studies, University of California, Irvine, aregan@uci.edu

³ Department of Information and Computer Science and Institute of Transportation Studies, University of California, Irvine, ssirani@uci.edu

July 2001

Institute of Transportation Studies
University of California, Irvine
Irvine, CA 92697-3600, U.S.A.
http://www.its.uci.edu
The Dynamic Traveling Repair Problem:

Examination of an Asymptotically Optimal Algorithm

Xiangwen Lu
Institute of Transportation Studies
University of California, Irvine
xlu@uci.edu

Amelia Regan
Department of Civil and Environmental Engineering
And, Graduate School of Management
University of California, Irvine
aregan@uci.edu

Sandra Irani
Department of Information and Computer Science
University of California, Irvine

Abstract

The dynamic traveling repair problem involves providing service to customers whose locations are uniformly distributed over a bounded area in the Euclidean plane. We assume that customer requests arrive according to a Poisson point process. Earlier research provided a conjecture that the asymptotically optimal algorithm for this problem under very heavy traffic intensity involves the following: partition the bounded area into sub-regions, wait for sufficient demand to accumulate in the sub-regions, serve the demands in the sub-regions according to the optimal TSP tour, and visit the sub-regions in first-come first-served order as in a GI/G/m queue. Further, the researchers conjectured that the optimal algorithm for the single server case can be extended to the m-server problem by simply partitioning the service region into m sub-regions of the same size and assigning one vehicle to a sub-region. In this paper we define a class of algorithms which includes the above algorithm. We then demonstrate the asymptotic optimality of an algorithm in this class and that the above algorithm is optimal among the class. Therefore, we prove the first conjecture made by the researchers. Finally, we argue that the conjecture made about the multiple server case is also true.

Keywords: Dynamic and Stochastic Routing and Scheduling, Dynamic Traveling Repair Problems, Probabilistic Analysis of Algorithms, Asymptotic Optimality

July, 2001, Under Review with Networks, Preliminary version can be found in the proceedings of the 2002 meeting of the Transportation Research Record
1. Introduction

The dynamic traveling repair problem (DTRP) is the following: \( m \) mobile servers are positioned within a bounded region \( A \) in the Euclidean plane. The servers travel at a fixed, constant speed \( v \) per time unit. Service requests arrive over time according to a Poisson process with arrival rate \( \lambda \). When requests arrive, they are distributed to the bounded region \( A \) independently according to a uniform distribution. The server must spend some time traveling to the customer locations and it must spend some time providing on-site service. The on-site service time for each customer is independently and identically distributed according to a distribution with mean \( \bar{\tau} \) and variance \( \tau'^2 \). The goal is to minimize the average waiting time of all customers. In this paper we examine this problem and address a conjecture about the asymptotic optimality of a partitioning algorithm for sequencing service to customers made by earlier researchers.

The examination of the asymptotic performance of heuristic algorithms for vehicle routing and scheduling is a common analysis technique. Asymptotic performance is thought to provide better insight into a heuristics typical performance than worst case analysis (See for example, Bramel and Simchi-Levi, 1997).

In a paper titled “Stochastic and Dynamic Vehicle Routing with General Demand and Interarrival Time Distributions”, Bertsimas and van Ryzin (1993, p. 962) examined the following \( G I / G / m \) service policy, which we refer as the \( BvR^n,k \) algorithm:
For a fixed integer $k$, the service area is partitioned arbitrarily into $k$ areas of equal size. Service requests are distributed uniformly over the whole area. When batches of $\frac{n}{k}$ customer requests accumulate in a partition, these are deposited into a queue in a first-come first-served manner as in a $GI/G/m$ queue. In each partition, customer requests are served according to the optimal TSP tour across their locations.

We define the expected fraction of time the vehicle spends providing on-site service as follows: $\frac{n}{m}$. The researchers conjectured that there exists a function $g(k)$ which determines $n$, such that as $k \to \infty$, the $BvR^n g(k)$ algorithm is asymptotically optimal. We refer to $BvR^n g(k)$ as the optimal $BvR^n$ algorithm and use $BvR^n(k)$ to denote it.

We consider a class of General Partition Algorithms which include the $BvR^n(k)$ algorithms as sub-class and develop a lower bound on the average wait time under this class of algorithms. We develop the lower bound for two different systems configurations. In the first, the size of the partitions depends upon the number of customers in the system. In the other, the partitions are fixed a priori. We refer to these as small partition and fixed partition cases. We obtain exactly the same bound for general partition algorithms applied to these systems. This lower bound matches the upper bound on the average waiting time provided by $BvR^n(k)$. Therefore, we show
that $BvR\,n,k\,\!^*$ is optimal among algorithms in this class. Finally, we identify an algorithm falling into this class whose asymptotic (heavy-traffic) performance is close to optimal. Therefore, we also prove the asymptotic optimality of $BvR\,n,k\,\!^*$. 

The dynamic traveling repair problem falls into a more general class of stochastic vehicle routing problems. This class includes the dynamic traveling salesman problem (DTSP), the probabilistic traveling salesman problem (PTSP), the probabilistic traveling salesman location problem and other problems such as the probabilistic shortest path problem and dynamic vehicle allocation problem. Powell, Jaillet and Odoni (1995) provide an excellent review of these problems. In the DTSP, customer locations are known in advance and service requests arrive according to a Poisson point process at each node. The objective is to determine a dispatching strategy which minimizes customers’ expected waiting time. That problem is discussed in Psaraftis (1988). In the PTSP, an a priori tour must be constructed for a network in which each node has a given probability of requiring a visit (Jaillet, 1988). The probabilistic traveling salesman facility location problem involves identifying the optimal location for a depot node in a network in which the probability that customers will require a visit is known (Berman and Simchi-Levi, 1988, Bertsimas, 1989).
2. Algorithms and Properties for the Single Server Case

2.1. Definition and Notation

Bertsimas and van Ryzin showed that when \( \varepsilon \) is less than one, there exists a function of \( g(k, \varepsilon) \) which determines \( n, \varepsilon \) such that such that the \( BvR \) algorithm can satisfy all service requests. This implies that \( N \), the expected number of requests in queue is finite.

Assume for now that there is just a single server and that the service region is a unit square. We number the demands according to the order in which they are served. Let \( d_i \) be the distance traveled from \( i \)th demand to \( i \)th demand. Let \( s_i \) be the on-site service time for demand \( i \). The total service time includes the travel time \( \frac{d_i}{v} \) and the on-site service time \( s_i \). If, for all times \( t \), the number of waiting requests in the system is bounded almost surely under a specific policy, we call this a stable policy. Using the definitions and notation presented by Bertsimas and van Ryzin (1991, 1993), for a stable policy, we let \( W_i \) denote the waiting time for demand \( i \). The waiting time is the time between the arrival of demand \( i \) and the arrival of the server at the location of demand \( i \).

The limiting expected values of these random variables are defined as \( \overline{W} = \lim_{n \to \infty} E[W_i] \), and, \( \overline{d} = \lim_{n \to \infty} E[d_i] \). \( N \), the expected number of requests in the queue, is equal to \( \overline{W} \).

As do Bertsimas and van Ryzin, we assume these limits exist.
2.2. Policies of Interest

Let \( W^r x^r \) be the expected waiting time for a randomly selected customer located at point \( x \). And, let \( W \) be the average waiting time under the algorithm of interest. We only consider algorithms that satisfy the following condition: there exists \( \overline{\gamma} \) and \( \underline{\gamma} \), such that

\[
0 < \underline{\gamma} \leq \frac{W^r x^r}{W} \leq \overline{\gamma}
\]  

(1)

This is a technical requirement for our proof. If \( \overline{\gamma} \) grows large and \( \underline{\gamma} \) grows small, constraint (1) becomes progressively less tight and the class of algorithms satisfying the constraint increases.

2.3 General Partition Algorithms

We now define a class of algorithms for the single vehicle DTRP which we refer to as General Partition Algorithms. These work as follows: using a grid, divide the area \( A \) into \( k \) partitions of equal size\(^1\). A general partition algorithm will, for each region, partition time into periods. Let \( t_{i,j} \) denote the end of region \( i \)'s \( j^{th} \) period. \( t_{i,0} \) is defined to be 0 for each region \( i \). The server then serves the requests in a sequence of visits. In each visit, the server selects a region to serve based on the accumulated demand in each region. This is the only information considered in the sequencing decision. In the \( j^{th} \) visit to region \( i \), the server will serve exactly those requests which arrive in region \( i \) in the interval from \( t_{i,j-1} \) through \( t_{i,j} \). The order in which the regions are visited and the way
that time is partitioned for each region will depend on the particular partition algorithm applied. If the server travels across regions in which there are no waiting customers, en route to a region in which there are waiting customers, we say that the empty region has also been visited.

$BvR^?n,k^?i$ is a general partition algorithm in which a period ends for a region when $\frac{n}{k}$ new requests have accumulated in that region. The regions are visited in first-come-first-serve order according to when each period ends.

Note that it is not required that the current period for a region be completed when the server arrives in that region. For example, the class of Exhaustive Partition Algorithms will visit a region, will serve all waiting customers, and will also serve those customers that arrive while service is being provided to waiting customers. Thus, the stopping criterion for a period is when there are no outstanding requests in the region.

2.4 Properties of General Partition Algorithms

We will now show how a general partition algorithm gives rise to a distribution function $f$ which describes the spatial distribution of consecutively served requests. Fix a general partition algorithm and positive integers $M$ and $i$. We are interested in the $i^{th}$ request through the $?i^?M^?1^{th}$ request. However, for convenience, we would like to focus on a

---

1 Note that Bertsimas and van Ryzin used a sweep algorithm rather than a grid. Any method to develop equal partitions will do. We use a grid to facilitate the development of our proof.
sequence of requests which start and end at the boundaries of visits. Suppose that the $i^{th}$ request is served in the middle of the $i^{th}$ visit. Let $r_b$ be the first request served in visit $i$. (We use $b$ for 'begin'). Now suppose that the $j^{th} M^{th}$ request is served in the middle of the $p^{th}$ visit. Let $r_e$ be the last request served in visit $p$. (We use $e$ for 'end'). We will focus on requests $r_b, r_e$.

Suppose that requests $r_b, r_e$ comprise $q$ consecutive visits. Suppose that $r$ distinct regions are visited during these $q$ visits. We call these $R_1, R_q$. Fix these $q$ visits, let $v_1, v_2, ..., v_r$ be the first visit to region $R_1, R_q$ and $v_1', v_2', ..., v_r'$ be the last visit to region $R_1, R_q$ among these $q$ visits. Remembering that each visit consists of an arrival period and that each visit clears only the customers arriving during that period. For each $j \leq 1, r$, let $t_{j, begin}$ refer to the beginning of $v_j$ and $t_{j, end}$ refer to the end of $v_j$.

Therefore, $t_{j, end} - t_{j, begin}$ refers to the time interval associated with the Poisson arrival process for region $R_j$. Let $Z_j$ denote the number of requests served in region $R_j$ during the interval $t_{j, begin}$ to $t_{j, end}$. $Z_j$ is exactly the number of requests that arrive in $R_j$ during the interval $t_{j, begin}$ to $t_{j, end}$. We use $|R_j|$ to denote the area of partition $R_j$. $Z_j$ is therefore a Poisson random variable with mean $|R_j| \cdot (t_{j, end} - t_{j, begin})$. Since the partitions are of equal size, $|R_j|$ is equal to $\frac{1}{k}$, where $k$ is the number of partitions. Thus, $Z_j$ is a Poisson random variable with mean $\frac{|R_j| \cdot (t_{j, end} - t_{j, begin})}{k}$.

Further, $Z_j \sim 1, r$.
are independent Poisson random variables. The independence results from the fact that the regions are disjoint.

Next we generate a random permutation \( x_1, \ldots, x_n \) of the set of consecutively served requests \( r_1, \ldots, r_n \) where \( n \) is the number of requests. We observe the following: We have \( r \) independent Poisson random variables, each representing the number of customers in region \( R_j \), with mean \( \frac{t \cdot t_{i, \text{end}} - t_{i, \text{begin}}}{k} \). Further, the locations of these requests are uniformly distributed in region \( R_j \). Let \( \mu_j = \begin{cases} 1 & \text{if } x^j \subseteq R_j \\ 0 & \text{otherwise} \end{cases} \) define an index function over \( R_j \). The locations of the served requests, \( x_1, \ldots, x_n \), are a realization of i.i.d. random variables with probability density function of the form

\[
 f^x \begin{pmatrix} \frac{t \cdot t_{i, \text{end}} - t_{i, \text{begin}}}{j^{rr}} \xi_j \end{pmatrix} \quad \left( \begin{pmatrix} \frac{t \cdot t_{i, \text{end}} - t_{i, \text{begin}}}{j^{rr}} \xi_j \end{pmatrix} \sum_{j^{rr}} \right) \quad \left( \begin{pmatrix} \frac{t \cdot t_{i, \text{end}} - t_{i, \text{begin}}}{j^{rr}} \xi_j \end{pmatrix} \sum_{j^{rr}} \right) \end{pmatrix} \left( \begin{pmatrix} \frac{t \cdot t_{i, \text{end}} - t_{i, \text{begin}}}{j^{rr}} \xi_j \end{pmatrix} \sum_{j^{rr}} \right) \end{pmatrix} \left( \begin{pmatrix} \frac{t \cdot t_{i, \text{end}} - t_{i, \text{begin}}}{j^{rr}} \xi_j \end{pmatrix} \sum_{j^{rr}} \right)
\]

(2)

2.5 The Small Partition Case

Let \( N \) represent the expected number of customers in queue. As before, let \( R \) be the size of the partitions. For any fixed \( \gamma \) and \( \gamma \), we consider only those general partition algorithms under which \( \gamma \) is less than or equal to \( \gamma \). For any \( \gamma > 0 \), let \( M \geq N \) and assume that \( M \) is an integer. For a fixed general partition algorithm \( \gamma \), when the system is in steady-state, for any randomly selected request \( i \), we are interested in requests from request \( i \) through the next \( M \) consecutively served requests. We expand the sequence of requests as described in the previous sections so that this sequence begins and ends at the boundaries
of visits. Let $Q$ denote the random variable that indicates the number of requests in this expanded sequence. Note that if the area of each partition is proportional to $\frac{1}{\pi^2}$, this ensures that the probability that the number of request in a randomly selected region is more than $N$ is negligible. With very high probability, $N \sim Q \sim 3N$, as $\lambda$ approaches one. Let $x_1, x_2, ..., x_Q$ represent the locations of $Q$ consecutively served customers. Let the index represent the order in which service is performed. Now let $y_1, y_2, ..., y_Q$ be a random permutation of $x_1, x_2, ..., x_M$. Therefore, $y_1, y_2, ..., y_M$ are distributed independently according to a distribution $f$, where $f$ is a piece-wise (discontinuous) function. Let $W(x)$ be the expected waiting time for a random selected customer $i$ that is located at point $x$. We present three propositions about $f$, $W(x)$ and $W(x_1)$.

Let $W$ be the average waiting time for fixed algorithm $\lambda$, we show that that the expected waiting time for any $M$ demands can be bounded by $W$. This leads to a constraint (5) on $f$, $W(x)$, the distribution of customer locations.

**Proposition 1** If $\lambda$ is NR $\geq \lambda$, the expected waiting time for a customer located in $x$ is

$$W(x) \leq \frac{Qf}{4 \pi} \cdot \frac{1}{2 \pi R_i}.$$  \hspace{1cm} (3)
Proposition 2 If we focus on policies under which \( 0 ? ? \frac{W ? x_i^e}{W} ? \frac{2}{?} \), when \( ? \) is large enough, then \( f, \int_?^? \frac{4?}{?} \frac{2}{?} \frac{2}{?} \). (4)

Proposition 3 For any \(? > 0\), when time is sufficiently large, \( f, \int_?^? x_i^e \) satisfies the following constraint \( \int_?^? \frac{\frac{2 ? N ? l^l ?}{Q}}{?} \). (5)

where \( ? \frac{1}{? 2 ? R_i^l ? \sqrt{2 / \nu_i^l} \). To obtain our lower bound for \( W \), we analyze the average distance between consecutive demands served. First, we obtain a lower bound for the distance traveled to serve all of these selected \( M \) demands expressed in \( f, \int_?^? x_i^e \) which leads to the average distance traveled per customer served. This is shown in lemma 1. We obtain lemma 1 by generalizing the classical TSP result of Beardwood et al. (1962) and applying a smoothing technique to the distribution function \( f, \int_?^? x_i^e \). Next, minimizing the lower bound obtained under constraint (5) leads to our lower bound for the average distance traveled to serve each customer (6.a). We use lemma 3 to obtain the lower bound (6.a). We then provide a lower bound on the average waiting time for service (6.b) based on (6.a). For the proof of theorem 1 and the related lemmas, please see section 5 and the attached appendix.

Lemma 1 Let \( \overline{d} \) be the expected average distance traveled per demand served and let \( N \) be the average number of customers awaiting service, \( R \) be the size of the partition and
\( Q \) represent the number of consecutively served customers. We consider algorithms for which for any \( ? > 0 \), we fix \( \bar{r} \) \( \gamma \) and \( \underline{r} \). For any general partition algorithms under which \( \bar{r} \) \( NR \) \( \underline{r} \), for any \( \gamma \), \( ?_0 \), \( ?_0 \), such that when \( ? \gamma ? \), 
\[ \sqrt{Qd} \gamma ? \gamma ? \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \gamma \ga
Finally, as we find that if we want to minimize the average waiting time, under the optimal algorithm among the class of algorithms considered, the distribution function \( f \hat{x} \) that describes the spatial distribution of \( M \) consecutively served customers will be almost uniform in a small area and zero over all other areas. In fact, the lower bounds (equations 6.a and 6.b) hold for a class of algorithms that satisfying the following condition.

For any \( N \), we consider \( M = N \) consecutively served customers and let \( y_1, y_2, \ldots, y_M \) be a random permutation of the locations. Consider algorithms under which \( y_1, y_2, \ldots, y_M \) are distributed independently according to a distribution \( f \hat{x} \), where \( f \hat{x} \) is required to satisfy the following condition:

there exists \( r \) such that \( |f \hat{x} - f \hat{y}| \leq r \|x - y\| \), where \( \|x - y\| \) is the distance between \( x \) and \( y \). Note that \( r \) does not depend on \( N \).

### 2.6 The Fixed Partition Case

Up to this point, we have examined algorithms for which the number and size of the partitions depends upon \( N \). Now we consider algorithms for which the partitions are pre-determined.

First use a grid to divide the unit square into fixed small partitions. Let \( \alpha \) be the area of each partition. When the server selects a partition to serve it will either finish all customers waiting in the queue at the selected partition or finish all the customers
arriving during some arbitrarily selected period. We show that as \( \frac{r}{m^2} \) approaches zero, the average waiting time for service will be bounded from below by \( \frac{A}{2m^2v^2(1-\frac{r}{m})^2} \).

**Theorem 2** For any algorithm falling into the category mentioned above, when \( \frac{r}{m} \) is small enough, \( \lim_{r \to 0} W \frac{r}{m^2} \) is \( \frac{A}{2m^2v^2} \). Note that this exactly matches equation (6.b), the bound for the small partition case.

### 2.7 An Asymptotically Optimal General Partition Algorithm

In this section we show the asymptotic optimality of a specific general partition algorithm. Let \( P_j \) be the partition which divides the area into \( j \times j \) squares and \( OPT \) denote the optimal algorithm. For a specific arrival sequence, let \( \frac{r_1, r_2, \ldots}{m} \), denote the order in which the requests are satisfied by \( OPT \). Given a partition \( P \) of the area, we will devise a General Partition Algorithm called \( A_p \) based on the behavior of \( OPT \). We take the sequence \( \frac{r_1, r_2, \ldots}{m} \) and remove some of the requests to obtain another sequence \( \frac{r_1, r_2, \ldots}{m} \) as follows: Examine each \( r_i \) in turn. Suppose that request \( r_i \) is located in region \( R_i \). Suppose that at the time that \( r_i \) is reached by \( OPT \)'s server, there are other outstanding requests in \( R_i \). Remove these additional outstanding requests from \( \frac{r_1, r_2, \ldots}{m} \) and continue. The requests that were removed from \( \frac{r_1, r_2, \ldots}{m} \) will be called *extra* requests. We denote the sequence \( \frac{r_1, r_2, \ldots}{m} \). The algorithm \( A_p \) will work as follows: for each request \( r_j \) in \( \frac{r_1, r_2, \ldots}{m} \), visit region \( R_j \) and satisfy \( r_j \) and any extra requests which are waiting...
in $R_i$ at the time \( OPT \) serves $r_j$. In other words, the periods are chosen so that when

\( OPT \) serves a request $r_j$ from $\mathcal{R}$, the current period for $R_i$ ends.

**Proposition 4** (Bertsimas and van Ryzin) \( \bar{d}^* \geq \frac{\sqrt{A}}{\sqrt{N} \cdot m/2} \) where $A$ is a constant,

\( a \geq \frac{2}{3 \sqrt{2} m}, \) $m$ is the number of servers.

**Proposition 5** For a fixed arrival rate $a$, let $\bar{d}$ denote the average distance traveled per customer for the optimal algorithm. For any $a$, for the partition $P_j$, when $j$ is large enough, the average distance traveled per customer served for the algorithm defined above is at most $\bar{d} + \frac{a}{\sqrt{N}}$.

Proof: The maximum diameter of a region under a given partition $P_j$ is $\frac{\sqrt{2}}{\sqrt{j}}$, then each extra request introduces a distance of at most $\frac{2\sqrt{2}}{\sqrt{j}}$ to $A_p$. When $j \nmid N$, we know

\( \frac{2\sqrt{2}}{\sqrt{j}} \geq \frac{1}{\sqrt{N}}. \) Thus, as the partitions become more fine, the additional distance introduced by an extra request decreases. Furthermore, as the partitions become more fine, the probability that any given request is an extra request also becomes smaller. Using these two facts, we conclude that for any arrival rate $a$, when $j$ big enough, the
average distance traveled per customer served for the algorithm defined above is at most
$$\tilde{d}^* + \frac{?}{\sqrt{N}}.$$  

If we let $\tilde{d}_{\gamma,p}$ be the average distance traveled per customer served, combining
propositions 4 and 5, we show that $\tilde{d}_{\gamma,p} = \tilde{d}^* + \frac{0}{\sqrt{d}}.$

2.8 The Optimality of $BvR \beta_n,k$ Among the General Partition Class

Lemma 3 (Bertsimas and van Ryzin) $W^* \leq \frac{?^2 A}{2m^2 \nu^2 \tilde{d}^*} + \frac{?}{\tilde{d}^*} \frac{1}{?} \frac{?}{?} \frac{?}{?} \frac{?}{?} \frac{?}{?} \frac{?}{?}$, where $m$ is
the number of servers, $\frac{?}{m}.$

Lemma 3 provides an upper bound for the average waiting time under
the $BvR \beta_n,k$ algorithm. Theorem 1 provides a lower bound for all general partition
algorithms. As $\gamma$ these bounds converge. Therefore, we show that under high traffic
intensity, $BvR \beta_n,k$ is asymptotically optimal among the General Partition class. Since
we have shown the asymptotic optimality of one algorithm in this class as we show that
$BvR \beta_n,k$ is asymptotically optimal among the whole class, then we have also
demonstrated the asymptotic optimality of $BvR \beta_n,k.$
3. The M-Server Case

Now we assume that instead of a single server that there are m mobile servers in the Euclidean service region. Comparing the single server and multiple server cases we find the following interesting result. If we divide the area into m sub-regions of equal size and assign each server to a single sub-region, if each server works independently, the algorithm that is asymptotically optimal for the single server case is also asymptotically optimal for m server case.

4. Abbreviated Proof of the Theorems (details in the appendix)

4.1. Proof of the Small Partition Case

Proof of Theorem 1. From lemma 1, for any $\gamma, \alpha_0$, when $m \geq \gamma$, 

$$\sqrt{Qd} \leq \gamma \sqrt{\int_0^R f(x) dx} \leq \sqrt{Nd} \leq \frac{\sqrt{\int_0^R f(x) dx}}{\sqrt{Q}}.$$ 

From proposition 3, for the above fixed $\gamma, \alpha_0$, we have, 

$$\frac{\int_0^R f(x)^2 dx}{\frac{2N\gamma^2}{1 + 3 \gamma^2}} \frac{\alpha_0}{Q}.$$ 

Where $\gamma \geq \frac{1}{\sqrt{2 + 2R}} \frac{\gamma}{\sqrt{2}} \gamma$ as $\gamma \geq 1$.

Using lemma 2, when $\gamma \geq \alpha_0$, 

$$\sqrt{Nd} \leq \frac{1}{\sqrt{2\frac{1}{\gamma^2} + 3 \frac{1}{\gamma^2}}} = \frac{1}{\sqrt{2\frac{1}{\gamma^2} + 3 \frac{1}{\gamma^2}}}.$$

Letting $\gamma \to 0$, we have $\lim_{\gamma \to 0} \sqrt{Nd} \to \frac{1}{\sqrt{2\frac{1}{\gamma^2} + 3 \frac{1}{\gamma^2}}}.$

Let $\gamma \to 0$, we have (6.a): $\lim_{\gamma \to 0} \frac{2Nd}{2} \to \frac{1}{\sqrt{2}}$. 

16
Now we show (6.b) based on (6.a).

Recall that \( \frac{d}{m} \) is the average arrival rate for each server and that \( \frac{d}{v} \) is the actual average service time for each demand. In a stable system \( \frac{d}{m} \frac{d}{v} \) must be less than 1.

We know that \( \frac{d}{m} \frac{d}{v \sqrt{2N}} \approx \alpha \left( \frac{1}{\sqrt{2N}} \right) < 1 \). Recalling that \( N \approx W \) and that \( \frac{d}{m} \approx s \), we have, \( \lim_{N \to \infty} (1 - \frac{d}{v})^2 W \approx \frac{\alpha^2 A}{2m^2 v^2} \).

4.2. Proof of the Fixed Partition Case

Proof of theorem 3: We observe that in these systems, a server may arrive at a region, and provide continuous service to customers until there are no customers in the region.

We call this the initial busy period. The server may then remain idle at the current region until a new customer arrives. At that time it enters into what we refer to as a subsequent busy period. In principle, a server may have many of these subsequent busy periods (later we show that only poor algorithms will allow the server to remain idle). Let \( p \) represent the fraction of customers that arrive during the initial busy periods and \( 1 - p \) represent the fraction that arrive during the either the idle periods or the subsequent busy periods.

Let \( Z_i \) be the number of customers served during the initial busy period for region \( R_i \).

Define \( X_i \) to be the number of customers served during the subsequent busy periods for region \( R_i \). Let \( p_i \) represent the fraction of customers served at region \( R_i \) during the
initial busy period and \( ?! ? \) \( p_i \), represent the fraction of customers served in region \( R_i \) during subsequent busy periods. At steady state we know that

\[
\frac{? ? X_i}{? ? Z_i} = 1 \frac{P_i}{P_i} \frac{1 ? P_i}{? ? X_i} ? \frac{1 ? P_i}{? ? Z_i}.
\]

From now on, we only consider the case in which \( ? \) is relatively large (\( ? ? \frac{3}{4} \)).

**Observation 1:** (The constraints on \( ? Z_i \) and \( ? Z_i X_i \))

We show in the appendix that when we have at least two partitions (\( ? \frac{1}{2} \)), a necessary condition for the system to be stable is \( p_i \frac{1}{2} \).

If we only consider the algorithms which satisfy the constraints that \( 0 \frac{W ? X_i}{W} \frac{3}{4} \), we know that when \( ? \frac{3}{4} \), \( ? Z_i \frac{2 ? N ? P_i}{1 ? P_i} \frac{1}{3} \frac{4 \frac{N}{2} ? 1 \frac{1}{2}} \)

Let \( g \) be the ratio of the average waiting time for partition \( R_i \) to \( W \). \( Z_i X_i \) is the total number of customer served during one visit to region \( R_i \).

\[
\frac{? Z_i}{? X_i} = \frac{2 g ? N}{p_i} \frac{1}{? ? ?}.
\]

**Observation 2:** (A Lower bound for the average distance traveled per customer served)

18
Given that the locations of demands in any given area is uniformly distributed, from Beardwood, et al. (1962), we know, for any given \( n \), there exists \( n_0 \), such that when \( n \geq n_0 \),

\[
\frac{L_{\text{TSP}}}{\sqrt{n}} X_1 \leq X_n \leq \frac{L_{\text{TSP}}}{\sqrt{n}} X_1.
\]

If the demand is from region \( R_i \), The average travel distance per demand served is

\[
\frac{L_{\text{TSP}}}{\sqrt{n}} Z_i X_i \leq \frac{L_{\text{TSP}}}{\sqrt{n}} Z_i X_i.
\]

When \( Z_i \geq n_0 \), we know \( \frac{L_{\text{TSP}}}{\sqrt{n}} Z_i X_i \leq \frac{L_{\text{TSP}}}{\sqrt{n}} Z_i X_i \leq \frac{L_{\text{TSP}}}{\sqrt{n}} Z_i X_i \).

Let \( \frac{PZ_i X_i}{n} \geq \frac{PZ_i X_i}{n} \),

\[
A \text{ lower bound for the average distance traveled per customer served is}
\]

\[
\frac{L_{\text{TSP}}}{\sqrt{n}} Z_i X_i \leq \frac{L_{\text{TSP}}}{\sqrt{n}} Z_i X_i \leq \frac{L_{\text{TSP}}}{\sqrt{n}} Z_i X_i.
\]

From observation 1, we know \( \frac{2g_i^2 N}{p_i} \) and \( \frac{4}{3} \).

From the definition of \( g_i \), we know \( g_i \leq 1 \).

Minimizing under above constraints leads to \( \lim_{\gamma \to 1} \sqrt{2N \bar{d}} \geq \sqrt{p_1} \).

Let \( \gamma \to 0 \), we know \( \lim_{\gamma \to 1} \sqrt{2N \bar{d}} \geq \sqrt{p_1} \).

\[ (7.a) \]
Observation 3: (the cost of remaining idle when the system is not empty)

Now we make the following observation: the average idle time per demand served during each subsequent busy period(s) is bounded from below by \( \frac{1}{\bar{\mu}} \). This comes from dividing the average interarrival time by the average number of customers served during a single busy period in an M/G/1 queue. So the average extra-cost due to idle periods per customer served for customers region \( R_i \) is equal to \( \frac{1}{\bar{\mu}} \). The average extra-cost due to idle periods per overall demand served is bounded by

\[
\frac{1}{\bar{\mu}} \leq \frac{1}{\bar{\mu}} + \frac{1}{\bar{\mu}}.
\]

From (7.a) and (7.b), we know that the average extra-cost due to switching and idling is at least

\[
\frac{\sqrt{p} \bar{\mu}}{\sqrt{2N}} + \frac{1}{\bar{\mu}} \geq \frac{1}{\sqrt{2N}} + \frac{1}{\bar{\mu}}.
\]

Because our system is in steady-state

\[
\frac{\sqrt{p} \bar{\mu}}{\sqrt{2N}} + \frac{1}{\bar{\mu}} \rightarrow \frac{\sqrt{p} \bar{\mu}}{\sqrt{2N}} + \frac{1}{\bar{\mu}}
\]

Algebraic manipulation leads to

\[
\lim_{\mu} \frac{\bar{\mu} \bar{\mu}}{p \bar{\mu} \bar{\mu} + \bar{\mu} \bar{\mu} + \bar{\mu} \bar{\mu}}
\]

We can show that when \( p = 1 \), \( \frac{\bar{\mu} \bar{\mu}}{p \bar{\mu} \bar{\mu} + \bar{\mu} \bar{\mu} + \bar{\mu} \bar{\mu}} \) is minimized.

Finally, we know

\[
\lim_{\mu} \frac{\bar{\mu} \bar{\mu}}{p \bar{\mu} \bar{\mu} + \bar{\mu} \bar{\mu} + \bar{\mu} \bar{\mu}}
\]
Because $\varepsilon$, $\delta$, $\gamma$, when $\delta$ is small enough, we have $\lim_{\delta \to 0} W \varepsilon \gamma \delta^2 \frac{\gamma^2}{2\nu^2m^2}$.

4. Conclusion

We construct a class of algorithms and demonstrate that $BvR n, k^\gamma$ is in this class and that it is optimal among algorithms in this class. Then we show that an algorithm in this class is asymptotically optimal. Therefore $BvR n, k^\gamma$ is asymptotically optimal. If it is $BvR n, k^\gamma$ is asymptotically optimal for the single vehicle case, it is also asymptotically optimal for the multiple vehicle case. Our results demonstrate the robustness of partition algorithms for routing and scheduling problems. These results mirror those developed earlier for the traveling salesman problem (Karp, 1985).

Acknowledgements

This research was partially supported by a grant from the University of California Transportation Center (UCTC) and by the US National Science Foundation (NSF) under Grants No. CMS-9875675 and No. CCR-9625844. The authors gratefully acknowledge this generous support. Any opinions, findings and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the UCTC or the NSF.
References


Appendix

First, in section A.1, we introduce a smoothing technique used to prove lemma 1. In section A.2, we introduce the lemmas needed to prove the lemma 1. We use lemma 1.1 and 1.2 to prove lemma 1. Using lemma 1.3 to prove lemma 1.1. In section A.3, we provide the proof for the propositions and in the last section, we provide proof of the lemmas.

A.1 Smoothing technique:

To prove lemma 1, we are interested in a smoothed version of $f$. A variety of smoothing techniques will work. We will choose one to make our discussion precise.

We select a parameter $\Delta$, for each area $A_i$ on which $f(x_i)$ is not zero, we define a new area $\hat{A}_i$ which contains the original one with the same center, the ratio of the parameter of new area to the original one is $\Delta = \frac{1}{2}$.

First we define a $g(x_i)$ based on $f(x_i)$ and $\Delta$: $g(x) = \frac{f(x)}{\Delta} \int_{x_{i}}^{x_{i}+\Delta} f(x) dx$. When $x \in A_i$.

We define $f_{\Delta}(x_i)$ to be a smoothed version of $f(x_i)$ as following: $f_{\Delta}(x) = \int_{x_{i} - \Delta}^{x_{i} + \Delta} g(x) dx$.

Note that as $\Delta$ gets small, $f_{\Delta}$ approaches $f$ in the limit. Furthermore, for any fixed value of $\Delta$, we can find a constant $C$ such that for any two points $x$ and $y$ in the area $A$,

$$|f_{\Delta}(x) - f_{\Delta}(y)| \leq C \|x - y\|,$$
where \( \| x - y \| \) is the Euclidean distance from \( x \) to \( y \) and 
\[
\| x \| = \max_{y} \left\| f \left( x \right) - f \left( y \right) \right\|.
\]
Note that 
\( f \) is bounded and so after applying this smoothing technique, the smoothed version of \( f \) will satisfy the Lipschitz \(-\)-condition.

### A.2 Lemmas

**Lemma 1.1** Let \( y_{1,\gamma}, y_{2,\gamma}, ..., y_{n,\gamma} \) be the i.i.d. random variables with distribution \( f_{\gamma} \), where 
\[
\gamma = \left\| f \left( x \right) - f \left( y \right) \right\| \left\| x - y \right\|. \]

Let \( L_{\text{TSP}} \gamma_{1,\gamma}, y_{2,\gamma}, ..., y_{n,\gamma} \) be the length of optimal TSP tour over \( y_{1,\gamma}, y_{2,\gamma}, ..., y_{n,\gamma} \), for any \( \gamma > 0 \), we can find \( N_{0} \), when \( n > N_{0} \), we have for any \( \gamma \), 
\[
\frac{L_{\text{TSP}} \gamma_{1,\gamma}, y_{2,\gamma}, ..., y_{n,\gamma}}{\sqrt{n}} \leq \sqrt{\int_{A} \gamma_{1,\gamma} \gamma \, dx} \leq \frac{L_{\text{TSP}} \gamma_{1,\gamma}, y_{2,\gamma}, ..., y_{n,\gamma}}{\sqrt{n}},
\]
where \( \gamma \) is the TSP constant defined in Beardwood et al. (1962).

**Lemma 1.2** Assume \( y_{1,i}, y_{M} \) are i.i.d. random variables with common distribution \( f \) and \( z_{1,i}, z_{M} \) are i.i.d. random variables with distribution \( f_{\gamma} \). Let \( L_{\text{TSP}} \gamma_{1,i}, y_{1,i}, y_{M} \) and 
\( L_{\text{TSP}} \gamma_{1,i}, z_{1,i}, z_{M} \) be the length of the optimal TSP tour over \( y_{1,i}, y_{M} \) and \( z_{1,i}, z_{M} \) respectively. For any \( \gamma_{i} \), when \( \gamma_{i} \) is sufficient small and \( M \) is sufficient large, we have,
\[
\frac{\frac{1}{M} \frac{1}{M} L_{\text{TSP}} \gamma_{1,i}, y_{1,i}, y_{M} \frac{1}{M} L_{\text{TSP}} \gamma_{1,i}, z_{1,i}, z_{M}}{\frac{1}{M}} \leq \gamma_{i}.
\]

**Lemma 1.3** Let \( \left\{ B_{i} \right\}_{i=1}^{m} \) be a grid partition over \( A \) such that each \( B_{i} \) has the same area.

Let \( f \) be constant within \( B_{i} \) and \( X_{1}, X_{2}, ..., X_{n} \) be i.i.d. random variables distributed
according to \( f'x' \). For any \( \varepsilon > 0 \), we can find \( N_1 \), \( N_2 \), \( N_3 \), \( N_4 \), when \( n \geq N_1 \), \( n \geq N_2 \), \( n \geq N_3 \), \( n \geq N_4 \), we have

\[
\alpha \sum_{k=1}^{n} \frac{L_{\text{TSP}}(X_1, X_2, \ldots, X_n)}{\sqrt{n}} - \frac{\sigma}{\sqrt{n}} \sqrt{\int_{\Delta} f(X) dX} \leq \alpha' \quad \text{for any } f'x'.
\]

To prove lemma 1.1 and lemma 1.3, we borrow the method from the classical paper by Beardwood, Halton and Hammersley (1959). To prove lemma 2, we rely on optimization methods and algebraic techniques. First we state the main result from the paper of Beardwood, Halton and Hammersley (1959) that we will use later in our proof.

**Theorem BHM:** Assume that \( X_1, X_2, \ldots, X_n \) are i.i.d. random variables with distribution \( f'x' \) and let \( L_{\text{TSP}}(X_1, X_2, \ldots, X_n) \) be the length of TSP tour over \( X_1, X_2, \ldots, X_n \).

\[
\lim_{n \to \infty} \frac{L_{\text{TSP}}(X_1, X_2, L, X_n)}{\sqrt{n}} = \alpha \sqrt{\int_{\Delta} f(X) dX} \quad \text{a.s.} \quad (a.1)
\]

**A.3 Proof of propositions**

**Proof of proposition 1**

We have \( Q \) customers, the locations of these customers are i.i.d. random variables with distribution \( f'x' \). We fix a region first, let it be \( R_i \). Assume \( f'x' = c_i \) when \( x \in R_i \).

Let \( n_i \) be the number of customers located at \( R_i \). It is easy to know that the distribution of \( n_i \) is Binomial with parameter \( c_i, R_i \). Let \( W'x' \) be the expected waiting time for a random selected customer located at \( x \).
Assume from the moment that the server enters a region and begins to provide the service and keeps working until there is no other customer in the current region, there are \( r \) customers served totally. We estimate the average waiting time in the following way:

when the server finishes the previous customer, there are \( r \) customers in the region, the average waiting time is bigger than \( \frac{1}{2} \) of the length of the total sum of \( \frac{r}{1} \) interarrival time minus the length of the total stay period, which is less than the sum of \( r \) on site service time and travel time (which is less than \( \frac{\sqrt{2}}{v} \)) to provide service.

Therefore,

\[
W \leq x \leq \frac{1}{\sqrt{2}} R_i \frac{\sqrt{2} s}{2}\frac{1}{R_i} = \frac{Qf}{2}\frac{x}{2R_i} \frac{1}{2R_i} \frac{\sqrt{2} s}{2R_i} \frac{1}{2}\frac{\sqrt{2}}{v}.
\]

Note that when \( ? ? 1 \), \( N ? \) implies \( R_i \) \( 0 \) as \( ? ? 1 \).

For any \( NR \) such that \( \frac{1}{4} ? R_i \frac{\sqrt{2}}{v} \) and \( NR ? ? 0 \), we know, when \( ? \) is big enough, we have \( W \leq x \leq \frac{Qf}{4} \frac{x}{2R_i^2} \frac{1}{2R_i} \).

Proof of proposition 2

From proposition 1, we know when \( ? \) is big enough,

\[
W \leq x \leq W \frac{Qf}{4} \frac{x}{2R_i^2} \frac{1}{2R_i} \frac{1}{W} \frac{Qf}{4N} \frac{x}{2NR_i} \frac{1}{f} \frac{x}{N} \frac{4NW}{W} \frac{x}{WQ} \frac{4N}{2NRQ}.
\]
If \( \bar{x}, \bar{y} \not\in \mathbb{R} \), and \( a \not\in \mathbb{R} \), we know \( f, g, h \) \( \frac{4\bar{z}}{\bar{y}} \), \( \frac{2}{\bar{x}^2} \).

**Proof of Proposition 3**

First we try to give a lower bound on the expected total waiting time.

Let \( Z_j \) denote the number of requests in \( R_j \). Let \( T_j \) be the total waiting time for the customers in \( R_j \). Let \( n_j \) be the total number of customers served in \( R_j \).

Using the fact that the distribution of the \( Q \) requests is \( f, g, h \), the distribution of these random variables of \( Z_j \)'s is given by the following multinomial distribution:

\[
P(Z_1 = z_1, \ldots, Z_r = z_r) = \frac{Q!}{z_1! \cdots z_r!} c_j \binom{n_j}{z_j}.
\]

\[
Q \approx \frac{1}{2} c_j R_j^2
\]

\[
\frac{\sqrt{2}}{\frac{3}{2}}
\]

We have \( \frac{Q}{2} \int_a^b f(x) \, dx \geq \frac{2}{\sqrt{2}} \int_a^b g(x) \, dx \).

Because \( \frac{W}{I} \), \( W \not\in \mathbb{R} \), \( \int_{\mathbb{R}^3} W \not\in \mathbb{R} \). So for any \( i \), \( \in \), there exists \( T \not\in \mathbb{R} \), when \( t \not\in T \), we have \( \frac{W}{I} \).
So when $t < T$, $\int_A f \, dx = \frac{2N}{Q^2} \int \frac{y}{y^2} \, dx$. 

A.4 Proof of lemmas

Proof of lemma 1.1

Observation 1. There exists $C_1 > 0$, such that $\sup_{f \in \mathcal{H}} \int \sqrt{f(x)^2} \, dx < C_1$.

For any $f \in \mathcal{H}$, let $f(x) = \min_{x \in A} f(y)$. Because $\int f(x) \, dx \leq 1$, we know: $\int f(y) \, dx \leq |D| \cdot A_1^2$ and $\int f(y) \, dx \leq \frac{1}{|D|}$. Therefore, $\int f(x) \, dx \leq 1$.

Let $\|A\|$ be the area of $A$ and $C_1 = \|A\| \sqrt{\frac{1}{|D|}}$. Therefore, $\int f(x) \, dx \leq C_1$.

Observation 2. For any given $1 > ? > 0$, let $\frac{\gamma^2}{\gamma} = \gamma$, for any for any $x$ and $y$ satisfying $\|x-y\| > ? \gamma$, assume $f(y) = ? ? x_1$. and $f(x) = ? ?$ where $? > 0, ? > 0$, note that $\sqrt{?} > ?$, therefore,

$$\left| \sqrt{f(x)} - \sqrt{f(y)} \right| \leq \frac{\gamma}{\sqrt{?} \sqrt{?} \sqrt{?}} \max_{y \in \mathcal{H}} \frac{\gamma}{\sqrt{?} \sqrt{?} \sqrt{?}} = \sqrt{?} > ?.$$
We divide $A$ into grid partitions of identical size with diameter $\delta$. As before, let $B_j$ be the $j^{th}$ partition in $A$.

Let $f_j(x) = \min_{y \in B_j} f(y), x \in B_j$.

Observation 2 leads us to observation 3.

**Observation 3.** $\int f_j \sqrt{\int x^2 dx} \rightarrow 1$, and $\int f_j \sqrt{\int x^2 dx} - \int f_j \sqrt{\int x^2 dx}$.

We place $Y_i$ into one of the two sets $(\frac{\int f_j \sqrt{\int x^2 dx}}{\int f_j \sqrt{\int x^2 dx}}$ and $\frac{\int f_j \sqrt{\int x^2 dx}}{\int f_j \sqrt{\int x^2 dx}}$ as follows:

If $\frac{\int f_j \sqrt{\int x^2 dx}}{\int f_j \sqrt{\int x^2 dx}} = 0$, let $Y_i \in \{1\}$ with probability one; otherwise, let $Y_i \in \{2\}$ with probability

$I \frac{\int f_j \sqrt{\int x^2 dx}}{\int f_j \sqrt{\int x^2 dx}}$ and $Y_i \in \{2\}$ with probability $\frac{\int f_j \sqrt{\int x^2 dx}}{\int f_j \sqrt{\int x^2 dx}}$.

Let $n_2$ be the number of requests belonging to $\{2\}$ and $L_{TSP}(\{2\})$ be the length of optimal TSP tour over all the nodes belonging to $\{2\}$.

We can show that

(I) $\frac{n_2}{n} \rightarrow 1$, almost surely.

(II) The random variables in the set of $\{2\}$ are i.i.d. random variables with the probability density function $\frac{\int f_j \sqrt{\int x^2 dx}}{\int f_j \sqrt{\int x^2 dx}}$.

Using lemma 1.3 and (II), we know that for any $\delta > 0$, there exists $N_4$, such that when $n_2 > N_4$, the following holds,
From (I), we know there exists \( n_0 \neq 0 \), when \( n \neq n_0 \),

\[
\frac{n_2}{\frac{f}{\sqrt{n}} \int f \, dx} \neq \frac{1}{n_3} \neq \frac{1}{n_3}.
\]

When \( n \neq \max n_0, \frac{N_4}{1?2?3?} \),

\[
\frac{L_{TSP}}{\sqrt{n}} \neq \frac{f}{\int f \, dx} ?? \frac{1?2?3?}{\int f \, dx}.
\]

From observation 1, observation 3, we know

\[
\frac{L_{TSP}}{\sqrt{n}} \neq \frac{f}{\int f \, dx} ?? \frac{1?2?3?}{\int f \, dx} \neq \frac{1?2?3?}{\int f \, dx}.
\]

Let \( ? = 2?3C_1 \neq ?_4 \), because \( C_1 \) and \( ? \) are constant and \( ?_3, ?_4 \) are arbitrary small number, \( ?_5 \) is arbitrary. At last, we find that

\[
\frac{L_{TSP}}{\sqrt{n}} \neq \frac{f}{\int f \, dx} \neq \frac{1?2?3?}{\int f \, dx} \neq \frac{1?2?3?}{\int f \, dx}.
\]

This concludes the proof of lemma 1.1.

**Proof of lemma 1.2**

Step 1. We place \( Y_j \) into one of the two sets \((?, ?)\) as follows:
If \( f_1, f_2, \ldots, f_n \) are \( f \) distributed densities, let \( Y_j \) with probability \( \frac{f_j(Y_j)}{f(Y_j)} \) and \( Y_j \) with probability \( \frac{1 - f_j(Y_j)}{f(Y_j)} \).

Step 2. Let \( \gamma = \int_{\mathbb{R}^d} f_{x_1} \ldots f_{x_k} dx \). Now we dispatch the elements in the set \( \gamma \) according to \( \frac{f_j(x)}{\gamma} \).

After these two steps, from \( y_1, y_2, \ldots, y_M \), we get \( z_1, z_2, \ldots, z_M \), where \( z_1, z_2, \ldots, z_M \) are i.i.d. random variables with distribution \( f_\gamma \).

Now we examine the number of elements in set \( \gamma \). Let \( n_2 \) be the number of requests belonging to \( \gamma \) and \( L_{\text{TSP}}(\gamma) \) be the length of optimal TSP tour over all the nodes belonging to \( \gamma \).

To calculate the difference between \( \mathbb{E}_\gamma L_{\text{TSP}}(Z_1, Z_2, \ldots, Z_M) \) and \( \mathbb{E}_\gamma L_{\text{TSP}}(Y_1, Y_2, \ldots, Y_M) \), we note the following fact,

(I) \( L_{\text{TSP}}(Z_1, Z_2, \ldots, Z_M) \) and \( L_{\text{TSP}}(Y_1, Y_2, \ldots, Y_M) \).

(II) From Karp and Steel (1985), lemma 2, we know there exists a tour whose length is less than \( 2\sqrt{\frac{\gamma M}{n_2}} \) over \( M \) nodes.

Combining these two facts we know that

\[
\mathbb{E}_\gamma L_{\text{TSP}}(Z_1, Z_2, \ldots, Z_M) \leq \mathbb{E}_\gamma L_{\text{TSP}}(Y_1, Y_2, \ldots, Y_M) \leq 2\sqrt{\frac{\gamma M}{n_2}}.
\]

It is easy to see that when \( \gamma \) goes to zero, \( \frac{\gamma M}{n_2} \) goes to zero too.
So for any $?, \text{ when } ?$ is sufficient small, $M$ is big enough, we have

$$\left| \frac{? \exists L_{TSP} : y_1, \ldots, y_M \exists \frac{? L_{TSP} : z_1, \ldots, z_M \exists \theta}{\sqrt{M}} ? ? \right|_1.$$ 

This concludes the proof of lemma 1.2.

**Proof of lemma 1.3**

We prove lemma 1.3 by induction based on the different values $f ? x^c_i$ can have.

First we assume that $f ? x^c_i$ is equal to zero or any constant over all the $B_i \text{ i.e.}$

$$f(x) ? \sum_{i} c_i \cdot \mathbb{I}_{g \cap B_i}(x), \mathbb{I}_{g \cap B_i}(x) ? 1 \text{ if } x \in B_i \text{ and } 0 \text{ otherwise} \text{ and } c_i \text{ equals to zero or } c.$$ 

Let $m$ be the number of the partitions on which $f ? x^c_i$ is $c$. The probability that any demand falls into any specific partition on which $f ? x^c_i \neq 0$ is $\frac{1}{m}$.

Let $L_{TSP}$ be the length of optimal TSP subtour over all the demands belonging to the $i^{th}$ piece of the partition on which $f ? x^c_i$ is $c$. Let $L_{TSP}$ be the optimal TSP tour over all the demands. Let $x$ is the length of the circumference of the partition $B_i$. For the optimal TSP tour over all the nodes, we are interested in the points that lie in the optimal TSP tour and the perimeter of the partition. For these points, we construct a tour through all the points and select each. After these steps, for each partition, we have a connected graph which each node has even degree. We know there exists one Euler tour that traverses all the links exactly once. Remembering that $n_0$ is the total number of
partitions, the length of this tour is less than \( L_{\text{TSP}} \leq 3n_0x \) and is as least the same length of the \( L_{\text{TSP}_i} \). Finally, we have \( L_{\text{TSP}} \geq L_{\text{TSP}_i} \geq 3n_0x \).

Now we try to estimate \( \frac{L_{\text{TSP}}(X_1, X_2, \ldots, X_n)}{\sqrt{n}} \). Because \( \frac{n_0x}{\sqrt{n}} \geq 0 \) as \( n \to \infty \), we focus on

\( \frac{L_{\text{TSP}_i}}{\sqrt{n}} \).

To express the idea clearly, from now on, we focus only on the partitions in which \( f(r)x \cdot c \) is a constant, \( c \). Assume these partitions are \( B_i, i \geq 1, 2, \ldots, m \).

Let \( n_i \) be the number of nodes in \( B_i \), let \( D_i = \frac{\sqrt{n_i}}{m} \frac{\sqrt{n_i}}{m} \),

Using Chebychev’s Inequality,

\[
P\{D_i \} \leq 1 - \frac{1}{k_i^2}
\]

(a.2)

By De Morgan’s rule,

\[
P\{\bigcap_i D_i^c \} = 1 - P\{\bigcap_i D_i\} \geq 1 - \sum_i P\{D_i\}
\]

Finally,

\[
P\{\bigcap_i D_i \} \geq 1 - \frac{m}{k_i^2} \geq 1 \quad k_i \geq \left( \frac{n_i}{m} \right)^{1/2}
\]

(a.3)

Let \(|B|\) be the size of the area of any partition and \(|A|\) be the size of the whole area, we use the result of Beardwood et al. (1959), please refer the (a.1) and combining with (a.2), (a.3), for any \( n_i \geq 2 \), there exists \( n_i \geq 2 \), when \( n_i \geq n_i \), we have

\[
P\left\{ \frac{\sqrt{n_i}}{\sqrt{|B|}} \geq \frac{\sqrt{\sqrt{|B|}}}{\sqrt{n_i}} \right\} \frac{3}{2} \leq 1 \quad n_i \geq 2
\]

(a.4)
i.e. $P^{\sqrt{n}} \frac{L_{T_{SP}}}{\sqrt{n}} \geq \frac{\sqrt{B}}{n} \frac{\sqrt{n}}{1} \frac{1}{n} \frac{m}{n}$ (a.5)

After some calculation, we know when $n \geq \max \frac{16m, m \geq n \geq \frac{3}{8}}{4^{1/3}}$, we have

$n \geq \frac{n}{m} - k_1 \sqrt{\frac{n}{m}} \geq n \geq \frac{1}{n}$. So when $n \geq \max \frac{16m, m \geq n \geq \frac{3}{8}}{4^{1/3}}$, with at least $P\{D_1\}$ probability that all $n_i$ satisfies $n_i \geq n \geq \frac{1}{n}$.

Because when $n \geq n \geq \frac{1}{n}$, we have $n_i \geq \frac{1}{m} \geq k_1 \sqrt{\frac{1}{nm}}$. (a.6)

From (a.5), (a.6) and $P\{A \cup B \cup \frac{1}{n} \geq A \cup B \cup \frac{1}{n} \geq 1$, we know,

$P^{\sqrt{n}} \frac{L_{T_{SP}}}{\sqrt{n}} \geq \frac{\sqrt{B}}{n} \frac{\sqrt{n}}{1} \frac{1}{n} \frac{m}{n}$

$P^{\sqrt{n}} \frac{L_{T_{SP}}}{\sqrt{n}} \geq \frac{\sqrt{B}}{n} \frac{\sqrt{n}}{1} \frac{1}{n} \frac{m}{n}$

i.e. $P^{\sqrt{n}} \frac{L_{T_{SP}}}{\sqrt{n}} \geq \frac{\sqrt{B}}{n} \frac{\sqrt{n}}{1} \frac{1}{n} \frac{m}{n}$

Because $k_1 \geq \frac{n}{m} \geq \frac{1}{n}$, so at last we have

$P^{\sqrt{n}} \frac{L_{T_{SP}}}{\sqrt{n}} \geq \frac{\sqrt{B}}{n} \frac{\sqrt{n}}{1} \frac{1}{n} \frac{m}{n}$

$P^{\sqrt{n}} \frac{L_{T_{SP}}}{\sqrt{n}} \geq \frac{\sqrt{B}}{n} \frac{\sqrt{n}}{1} \frac{1}{n} \frac{m}{n}$

$P^{\sqrt{n}} \frac{L_{T_{SP}}}{\sqrt{n}} \geq \frac{\sqrt{B}}{n} \frac{\sqrt{n}}{1} \frac{1}{n} \frac{m}{n}$

34
Note that \( \int_{-\infty}^{\infty} f(x) \, dx \leq \sqrt{m |B|} \), we have

\[
? \leq \frac{L_{TSP}}{\sqrt{n}} \leq \frac{m^{3/2}}{\sqrt{n}} \leq \frac{\sqrt{m}}{\sqrt{n}} \leq \frac{\sqrt{21}}{\sqrt{n}} \leq \frac{2}{n}.
\]

Because \( \frac{m}{n} \to 0 \), \( \frac{m^{3/2}}{\sqrt{n}} \to 0 \) and \( \frac{n_0 x_i}{\sqrt{n}} \to 0 \) as \( n \to \infty \), \( L_{TSP} \leq L_{TSP_1} - 3n_0 x \) and \( \phi \) is arbitrary, so for any \( \phi \neq 0 \), we can find \( N_0 \), \( \phi \), when \( n \geq N_0 \), we have

\[
\phi \leq \frac{L_{TSP}(X_1, X_2, \ldots, X_n)}{\sqrt{n}} \leq \phi \sqrt{f(X) \, dX} - \phi.
\]

The above argument holds for any fixed \( m \), when \( m \) goes from 1 to \( n_0 \), we fix

\( n^* = \max_m \{ n^* | m \} \), we know when \( n \geq n^* \), we have

\[
\phi \leq \frac{L_{TSP}(X_1, X_2, \ldots, X_n)}{\sqrt{n}} \leq \phi \sqrt{f(X) \, dX} - \phi.
\]

The last step in this induction proof is to assume the lemma is true when \( f \) has \( k \) different values and to show that the lemma holds when \( f \) has \( k + 1 \) different values.

The proof is based on similar ideas and is rather tedious, so we omit the rest of proof here.

This concludes the proof of lemma 1.3.
Proof of lemma 1

We know that $y_{1}, y_{2}, ..., y_{M}$ are i.i.d. random variables with distribution of $f_{i} x_{i}$. After applying the smoothing technique, we know that for any $f_{i} x_{i}$ that satisfies the condition of $f_{i} x_{i} > \frac{4\sigma}{\sqrt{d}} > \frac{2\sigma}{\Phi}$, we can find common $\gamma$ such that $f_{i} x_{i}$, the smoothed version of $f_{i} x_{i}$ has the following two properties:

Property I: Its satisfies the Lipschitz $\gamma$-condition;

Property II: Let $z_{1}, z_{2} \boxplus z_{o}$ be i.i.d. r.v.s with distribution of $f_{i} x_{i}$. From lemma 1.2, we know that for any $\gamma_{1} > 0$, $\gamma_{2} > 0$, when $\gamma_{1} > \gamma_{2}$, $\gamma_{3}$, $\gamma_{4}$,

$$\frac{\gamma_{1} L_{\text{TSR}} y_{1}, y_{2} \boxplus y_{o} \gamma_{3}^{2}}{\sqrt{\gamma_{1}}} \frac{\gamma_{2} L_{\text{TSR}} z_{1}, z_{2} \boxplus z_{o} \gamma_{4}^{2}}{\sqrt{\gamma_{2}}}$$

From lemma 1.1 and property I, we know for any $\gamma_{5} > 0$, $\gamma_{6}$, when $\gamma_{5} > \gamma_{6}$, $\gamma_{7}$,

$$\frac{\gamma_{5} L_{\text{TSR}} z_{1}, z_{2} \boxplus z_{o} \gamma_{7}^{2}}{\sqrt{\gamma_{5}}} \gamma_{6} \sqrt{f_{i} x_{i} \gamma_{7}^{2} dx} \gamma_{8}$$

Because $\gamma_{5} \sqrt{f_{i} x_{i} \gamma_{7}^{2} dx} > \gamma_{1} > 4\gamma_{2} > 4\gamma_{4}^{2}$, $\gamma_{6} \sqrt{f_{i} x_{i} \gamma_{7}^{2} dx}$, and $\gamma_{2}, \gamma_{3}$ are arbitrary,

$$\sup_{f_{i} x_{i}} \gamma_{5} \sqrt{f_{i} x_{i} \gamma_{7}^{2} dx} |A| \frac{4\sigma}{\sqrt{d}} > \frac{2\sigma}{\Phi}$$

by the similar argument we made in the proof of lemma 1.1, we can show that for any $\gamma_{9} > 0$, $\gamma_{10}$, when $\gamma_{9} > \gamma_{10}$,

$$\sqrt{\gamma_{9} d} > \gamma_{10} \gamma_{11} \sqrt{f_{i} x_{i} \gamma_{12} dx}$$

This concludes the proof of lemma 1.
Proof of lemma 2

By adding some additional non-binding constraints we can show that the original problem can be translated into the following problem:

\[
\min_{x_i, A_i^*} \sqrt{x_i A_i^*} \text{ subject to: } \begin{array}{c}
\sum_{i=1}^{n} x_i A_i^* = 1, \\
\sum_{i=1}^{n} x_i^2 A_i^* = 1, \\
\sum_{i=1}^{n} x_i = ?_i, \\
\end{array}
\]

for some small value \(?_i\). When \(?_i\) the constraints \{\(x_i, ?_i\)\}_{i=1}^{n} are non-binding and hence, do not affect the solution to our problem.

For this new problem, we use a standard optimization techniques as follows:

Let

\[
L(X, ?, ?, ?) = \sum_{i=1}^{n} x_i A_i^* - \sum_{i=1}^{n} x_i A_i^* - 1 + \sum_{i=1}^{n} x_i^2 A_i^* - \sum_{i=1}^{n} x_i^2 A_i^* + \sum_{i=1}^{n} (x_i - ?_i).
\]

By considering the Kuhn-Tucker conditions we know that the optimal solution must satisfy the following:

\[
\frac{L}{x_i} = 0, \quad \frac{L}{x_i} = 0, \quad \frac{L}{x_i} = 0, \quad ?_i = 0, \quad ?_i = ?_i.
\]

Therefore, for the optimal solution \(1 \leq 2x_i^2 - 4x_i^3 \geq 0, \quad ?_i = 0, \quad ?_i = 0\) must hold.

Considering the set of equations \(1 \leq 2x_i^2 - 4x_i^3 \geq 0, \quad ?_i = 0, \quad ?_i = 0\). From the theory of algebraic equations, we know the following facts.

Case One: The equations have the same nonnegative solution i.e. \(x_i = x_j\) for any \(i, j\).

Case Two: There are at most two different position solutions.

Assume that the positive solutions are \(a\) and \(b\) respectively and that \(a > b\). Let

\[
Z = \min_{x_i, A_i^*} \sqrt{x_i A_i^*} \text{ subject to: }
\]
\[ z \leq x_i A_i^* \leq 1 \]

\[ z \leq x_i^2 A_i^* \leq 1 \]

\[ x_i \in \{a, b\} \]

\[
Z = \min \left\{ \sum_{i=1}^{n} \left[ \sqrt{a A_i^* \leq \sqrt{b A_i^*} \leq \frac{1}{2}} \right] \right\}
\]

subject to

\[ a A_i^* \leq \sqrt{b A_i^*} \leq 1 \]

\[ a^2 A_i^* \leq \sqrt{b^2 A_i^*} \leq 1 \]

So if we let \( x \leq A_i^* \) and \( y \leq A_i^* \), we obtain,

\[
Z \leq \min_{a, b} \sqrt{ax} \sqrt{by} \leq 1
\]

subject to:

\[ ax \leq by \leq 1 \]

\[ a^2 x \leq b^2 y \leq 1 \]

\[ a \leq 0, \ b \leq 0, \ b \geq a. \]

Therefore, \( Z \leq \frac{1}{\sqrt{1}} \).

This concludes the proof of lemma 2.