Batched disk scheduling with delays

Eitan Bachmat [1] Vladimir Braverman [2]

1 Introduction

One of the important performance enhancing capabilities of modern disk drives, is the ability to permute the order of service of incoming I/O requests in order to minimize total access time. Given a batch (set) of I/O requests, the problem of finding the optimal order of service is known as the *Batched Disk Scheduling Problem* (BDSP). BDSP is a well known instance of the Asymmetric Traveling Salesman Problem (ATSP), in fact it has been used as one of a few principal test cases for the examination of heuristic algorithms for the ATSP, [4], [12]. To specify an instance of BDSP amounts to a choice of a model for the mechanical motion of the disk and a choice of locations and lengths of the requested I/O in the batch. The distance between requests is the amount of time needed by the disk to move from the end of one request to the beginning of the other, thus the amount of time needed to read the data itself, *Transfer time*, is not counted since it is independent of the order of the requests, only the order dependent *Access time* is computed.

Apart from the algorithmic problem of finding an optimal ordering of request the other natural question regarding traveling salesman problems is to estimate the length of the optimal tour given some assumptions on the problem instances. In the context of BDSP one needs to specify probability distributions for the location and length of the I/O requests. In this paper we will consider the case of fixed I/O length.

The idealized case of zero length data was considered in [2]. In particular it is shown in [2] that when the radial speed of the disk head is assumed to be constant (ignoring acceleration and deceleration effects) the length of the optimal tour concentrates around a value of the form $c\sqrt{n}$, where n is the number of requests and c is a constant which depends on the speed of the disk head and on the distribution of requests.

When data length is a non zero constant α, then the optimal tour length will be at least αn, which is the time needed to read/write the requests. Since this unavoidable time penalty is independent of the tour we will consider only the excess tour length beyond αn. Our main result is that with constant radial speed the excess length of the optimal tour is with probability approaching 1 $O(N^{2/3}\log n)$ when one makes (the realistic) assumption that request distribution only depends on the radial position of the data.

In the (unrealistic) case that the request distribution also depends on the angular position of the data the excess length of the optimal tour is with probability approaching 1 of size $\Theta(N)$.

1.1 Related work

Much work has been done on the disk scheduling problem. early studies, [5, 8, 10], only considered seek times and not rotational latency since a complete track rotation was required in early disk drives. Rotational latency was first discussed in the experimental papers, [11, 17, 6], which examined many heuristics for the problem using simulations. The papers [14, 15] consider the issue of starvation in conjunction with total service time and stress the tension between these concepts. The first analytical work on disk scheduling is by Andrews, Bender and Zhang, [1]. They showed that BDSP is in general NP complete. In the idealized case in which there are no delays (the time needed to read or write data is zero). They produced a 3/2 approximation algorithm (with an additive constant) for tour length. Note however that such an algorithm does not provide an approximation for excess tour length. In addition, they provided a polynomial time algorithm for the optimal tour assuming that data length is zero and the seek function is linear. The case of zero data length was further analyzed in [2, 3]. In [2] it was shown that the optimal tour length for seek functions of the form $f(\theta) = \theta^a$, $a \geq 1$ is $\Theta(N^{\frac{a}{a+1}})$. More precise formulas for linear seek functions ($a = 1$), where presented in [3]. In particular it is shown that BDSP with zero length data and a linear seek function is best understood in terms of 2 dimensional Lorentzian (space-time) geometry. The paper [16] considers non zero data length, but only takes rotational latency into account. While we strongly believe that the results of [16] are correct it contains an error in a central argument in the proof.

2 Problem formulation and preliminaries

A computer hard disk has the shape of an annulus, which geometrically can be thought of as a cylinder C. For convenience we normalize the radial distance between

[1] Dept. of Computer Science, Ben-Gurion U., Beer-Sheva, Israel, 84105, email: ebachmat@cs.bgu.ac.il
[2] Dept. of Computer Science, UCLA, Los Angeles, email: vova@cs.ucla.edu

the inner and outer circles to be 1. Each point on the disk (cylinder) is then represented by polar coordinates, $R = (r, \theta)$ where $0 \leq r \leq 1$ is the radial distance from the inner circle and $0 \leq \theta \leq 1$ is the angle relative to an arbitrary but fixed ray. A complete circular angle is chosen to be of 1 unit instead of 2π as in [1] for the convenience of future computations. The points $(r, 0)$ and $(r, 1)$ are identified.

A disk rotates at a constant speed in a fixed direction, hence we may measure time in units of disk rotations or via our normalization in angular units.

The head of the disk is an arm which moves in and out in the radial direction. To reach a certain location on the disk, the head moves radially while the disk rotates, until it reaches the correct radial position, it then waits for the disk to further rotate until the desired location passes under the head, the head can then read/write data at the location.

Associated with the disk is a *seek function* $f(\theta)$. The function $f(\theta)$ represents the maximum radial distance the disk head can travel (starting and ending with no radial velocity) in time θ. The acceleration and deceleration involved in disk head motion dictates that f is in general a convex function. In this paper we will assume that the seek function has the form

$$f(\theta) = \alpha + c\theta$$

for some constants $c, \alpha > 0$. The constant term α in the seek function can be thought of as representing the time needed to read/write the data item and also some unavoidable overhead associated with seeking such as head positioning time. Thus, we are implicitly assuming that all data requests have approximately equal size.

Given a pair of locations $P_1 = (r_1, \theta_1)$, $P_2 = (r_2, \theta_2)$ on the disk, the time it takes the head of the disk to reach P_2 starting at P_1 is composed of two parts, *seek time* and *rotational latency*. Seek time is the time it takes the head of the disk to move from radius r_1 to radius r_2. By definition of the seek function f the seek time is simply $f^{-1}(|r_2 - r_1|)$, where f^{-1} refers to the inverse function which in our case is simply $\alpha + \frac{|r_2 - r_1|}{c}$. Starting from P_1 the head will reach the desired radius r_2 at the angle $\theta = \theta_1 + \alpha + |r_2 - r_1|/c \ Mod\ 1$. Rotational latency is the time the head has to wait until the disk has rotated from angle θ to the desired angle θ_2.

For $0 \leq x, y \leq 1$, let $x \ominus y = x - y$ if $x \geq y$, and $x \ominus y = 1 - y + x$ otherwise. $x \ominus y$ represents the distance in a unit length circle between angles x and y in the counter clockwise direction. Rotational latency is thus given by $\theta_2 \ominus \theta$. The *access time* $a(P_1, P_2)$ is the sum of seek time and rotational latency time and represents the total amount of time needed to access point P_2 starting from point P_1. Note that access time is asymmetrical since rotational latency time is asymmetrical. By our definitions the access time is always at least α. This constant factor represents unavoidable access time which includes the time required to read/write data, known as *transfer time* as well the time needed for various basic operations related to disk arm movement such as head switching time. We define the excess time $d(P_1, P_2) = a(P_1, P_2) - \alpha$. It is interesting to note that while the access time satisfies the triangle inequality the excess time may not satisfy it. The excess time is the tour dependent part of the access time.

Given a batch of requests $\bar{P} = \{P_0, P_1, ..., P_{n-1}\}$ we define the *total excess time* of the batch $T(\bar{P})$ to be the minimum over all permutations π of $\sum_{i=0}^{N-2} d(P_{\pi(i)}, P_{\pi(i+1)})$.

To study the asymptotics of $T(\bar{P})$ we need to specify a distribution for the location of requests. The distribution of request locations can be given by a density of the form $p(r, \theta) dr d\theta$ where $p(r, \theta)$ is an arbitrary continuous density function. The most realistic densities are independent of θ, namely, $p(r, \theta) = p(r)$. The reason is that files are laid out radially on the disk, hence, variations in file popularities lead to variations in the popularity of tracks or radial positions. On the other hand it is hard to imagine that an application will prefer data which resides at a certain angle over data at other angles. As will be seen angle independent density distributions lead to substantially lower total excess times.

3 Single track computations

In this section we estimate the excess time for a batch $\bar{P}$ of n requests which are assumed to be uniformly distributed on a single disk track, namely $r_i = r$ for all $i = 0, ..., n-1$ and some fixed r. This unrealistic scenario will serve as a basis for the computation of excess time for a uniformly distributed batch of requests on the whole disk which will be given in the next section. Since $r_i = r_j$ for all i, j the seek time is always equal to α and thus the excess time is equal to rotational latency time. Consider n uniformly distributed requests. We may shift all the angles by a fixed constant and reorder them so that $0 = \theta_0 < \theta_1 < \theta_2 < ... < \theta_{n-1}$.

Let θ_{b_i} be a tour, i.e., the b_i form a permutation. Let θ_{b_i} and $\theta_{b_{i+1}}$ be two consecutive requests on the tour. Let

$$V_i = \theta_{b_{i+1}} \ominus \theta_{b_i}$$

and let $d_i = \theta_{b_{i+1}} \ominus (\theta_{b_i} + \alpha)$ be the excess time. We claim that if $V_i \geq \alpha$ then $d_i = V_i - \alpha$. We first note that for any pair of angles x, y and $z \leq x \ominus y$ we have $x \ominus y = (x \ominus y + z) + (y + z \ominus y)$. This can be seen either from the definition or from the interpretation of $x \ominus y$ as a distance in the counterclockwise direction. It also follows easily from the definition that for $z < 1$ we have

$y + z \ominus y = z$. Combining these observations and setting $x = \theta_{b_{i+1}},\ y = \theta_{b_i}, z = \alpha$ yields the claim. The total excess time is $T = \sum_{i=0}^{n-1} d_i$, where we set $b_n = b_0$.

We construct a tour among the θ_i in steps of size $s = \lceil \alpha n \rceil + \lceil \sqrt{n} \log n \rceil$ indices at a time. To see why we expect such a tour to be relatively short we estimate the length of such a tour as follows. Given a request at θ_i the request at $\theta_{i+s (\bmod n)}$ is expected to be at a circular distance of $\alpha + \log n/\sqrt{n}$ from θ_i. and so d_i for such a step is expected to be $\log n/\sqrt{n}$, summing over n requests we get an expected excess time of $\sqrt{n}\log(n)$. The factor of $\log n$ is taken to probabilistically assure us that only on rare occasions we will need more than a disk rotation to reach $\theta_{i+s (\bmod n)}$ from θ_i. We note that when s is not prime to n, taking steps of size s will not provide a tour. We will overcome this problem by slightly altering the steps. In what follows we make the above heuristic arguments precise.

Let n be an arbitrary positive integer. Put:

$$m = \left\lceil \sqrt{n} \log n \right\rceil, \quad s = \lceil \alpha n \rceil + m, \quad g = \gcd(n, s), \quad h = \frac{n}{g}.$$

For $i = 0, \ldots, n-1$, write $i = ph + q$ with $0 \le p < g$ and $0 \le q < h$, and denote $b_i = (qs + p) (\bmod n)$. Our tour will be

$$\theta_0, \theta_{b_1}, \theta_{b_2}, \ldots, \theta_{b_{n-1}}$$

We note that when s is prime to n, we have, $g = 1$, $i = q$ and $b_i = si (\bmod n)$. The first lemma shows that this is indeed a tour.

Lemma 3.1 *$b_i \ne b_j$ for $0 \le i < j \le n-1$.*

Proof: Suppose $b_i = b_j$. Write

$$i = p_1 h + q_1, \qquad 0 \le p_1 < g,\ 0 \le q_1 < h,$$

$$j = p_2 h + q_2, \qquad 0 \le p_2 < g,\ 0 \le q_2 < h.$$

Then

$$q_1 s + p_1 \equiv q_2 s + p_2 (\bmod g),$$

and therefore

$$p_1 \equiv p_2 (\bmod g).$$

Due to the fact that $p_1, p_2 < g$, we conclude that $p_1 = p_2$. Suppose, say, that $q_1 < q_2$. Then

$$(q_2 - q_1)s \equiv 0 (\bmod n).$$

Thus $q_2 - q_1 = 0 (\bmod h)$. Since $q_1, q_2 < h$, we have $q_1 = q_2$ and $i = j$. ■

The next lemma shows that the steps in the sequence are nearly of size s.

Lemma 3.2 *For $0 \le i < n$:*

$$s \le (b_{i+1} - b_i) (\bmod n) \le s + 1.$$

Proof: For $0 \le i < n-1$, let:

$$i = ph + q, \qquad 0 \le q < h.$$

If $q < h - 1$, then

$$b_{i+1} - b_i = (q+1)s + p - qs - p = s,$$

whereas if $q = h - 1$, then:

$$b_{i+1} - b_i = p + 1 - (h-1)s - p = 1 + s - hs = s + 1 (\bmod n).$$

For $i = n-1$ we have $n - 1 = (g-1)h + h - 1$ and

$$b_0 - b_{n-1} = 0 - (h-1)s - (g-1) = s + 1 (\bmod n).$$

■

In order to analyze this sequence probabilistically we consider first the properties of the sorted sequence $\theta_0, \ldots, \theta_{n-1}$. For integers j and k with $0 \le j < k < n$, the conditional distribution of θ_k given θ_j is the distribution of $\theta_j \oplus Y^{(k-j)}$, where $Y^{(l)}$ is the l-th order statistic (l-th number in size) of a sequence $(Y_i)_{i=1}^{n-1}$ of independent $U[0,1)$-distributed random variables. Similarly, for $0 \le k < j < n$, the conditional distribution of θ_k given θ_j is the distribution of $\theta_j \oplus Y^{(n+k-j)}$. We conclude that for two consecutive requests θ_{b_i} and $\theta_{b_{i+1}}$ the distribution of V_i is given by $Y^{b_{i+1} - b_i (\bmod n)}$. In particular we have

$$EV_i = \frac{b_{i+1} - b_i (\bmod n)}{n}, \qquad 0 \le i \le n-1,$$

We recall the following inequality, due to Hoeffding, which will be helpful in analyzing the behavior of V_i.

Theorem A [9, Theorem 2] *If $(W_j)_{j=0}^{n-1}$ are independent random variables such that $c \le W_j \le d$, $(j = 0, 1, \ldots n-1)$, then for $t > 0$*

$$P\left(|\sum_{j=0}^{n-1} (W_j - EW_j)| \ge nt \right) \le 2e^{-2nt^2/(e-f)^2}.$$

Lemma 3.3 *Let $Y^{(s)}$ be the s-th order statistics of $n-1$ independent $U[0,1)$-distributed random variables, X_j, representing the non sorted sequence of angles. Then:*

$$P\left(Y^{(s)} < \alpha \right) \le 2n^{-2\log n} \tag{1}$$

and

$$P\left(Y^{(s+1)} > \alpha + 2\frac{\log n + 1}{\sqrt{n}} \right) \le 2n^{-2\log n} \tag{2}$$

Proof: Let X_j be the j-th (non sorted) angle which was chosen uniformly. Let

$$W_j = \mathbf{1}\left(X_j \in [0,\alpha]\right), \qquad 1 \le j \le n-1.$$

The W_j's are independent $B(1,\alpha)$-distributed random variables. By definition $\sum_{j=0}^{n-1} W_j$ counts the number of elements X_j which are at a distance of at most α from $0 = \theta_0$, hence,

$$P\left(Y^{(s)} < \alpha\right) = P\left(\sum_{j=1}^{n-1} W_j \ge s\right).$$

Therefore:

$$P\left(Y^{(s)} < \alpha\right) \quad (3)$$

$$= P(\sum_{i=1}^{n-1}(W_i - \alpha) \ge \lceil \alpha n\rceil + \lceil \sqrt{n}\log n\rceil - \alpha(n-1)) \quad (4)$$

$$\le P\left(\sum_{i=1}^{n-1}(W_i-\alpha) \ge \sqrt{n}\log n\right). \quad (5)$$

The variables W_i satisfy the assumptions of Theorem A, with $e=0, f=1$ for each i, and hence:

$$P\left(Y^{(s)} < \alpha\right) \le 2e^{\frac{-2n(\log n)^2}{(n-1)}} \le 2e^{-2\log n^2} = 2n^{-2\log n}. \quad (6)$$

The second assertion is proved in the same manner after defining

$$W_j = \mathbf{1}\left(X_j \in \left[0, \alpha + 2\frac{\log n + 1}{\sqrt{n}}\right]\right), \qquad 1 \le j \le n-1. \quad (7)$$

■

We have the following easy consequence of the previous lemma.

Corollary 3.4 *With probability at least* $1 - 4n^{1-2\log n}$ *we have* $T \le (\log n + 1)\sqrt{n}$

Proof: Since V_i has the same distribution as either $Y^{(s)}$ or $Y^{(s+1)}$ we see from inequality 1 and the union bound that with probability greater than $1 - 2n^{-2\log n + 1}$ we have $V_i \ge \alpha$ and hence $d_i = V_i - \alpha$ for all i. The desired result is now obtained from inequality 1 and the union bound. ■

4 Entire disk computations

We now consider scheduling for a set $P_0, ..., P_{n-1}$ of independently chosen requests on a disk, sampled with respect to a density of the form $p(r)$. In polar coordinates we have $P_i = ((r_i, \theta_i))_{i=0}$, Since the density p depends only on r the angles θ_i are i.i.d. random variables, uniformly distributed on $[0,1)$. Consider the random sets

$$I_j = \left\{i : r_i \in \left[jn^{-1/3}, (j+1)n^{-1/3}\right)\right\}$$

for $j = 0, ..., \lceil n^{-1/3}\rceil$. Put $i_j = \min I_j$ and $k_j = |I_j|$. Define the two sets of indices

$$A = \left\{j : \ k_j \le n^{1/3}\right\}$$

and

$$B = \left\{j : \ k_j > n^{1/3}\right\}.$$

For each $j \in B$ we construct a tour on I_j as follows. Consider the sorted sequence of angles $\theta_{0,j}, ..., \theta_{k_j-1,j}$. Apply to this sequence the tour construction of the previous section, replacing α by

$$\beta = \alpha + \frac{n^{-1/3}}{c}.$$

For each $j \in A$ we construct an arbitrary tour through I_j We then patch the tours into a single tour by moving from the last request on the tour of I_j to the first request of the tour on I_{j+1}. We let T_j denote the excess time of the tour on I_j and by T the total excess time of the tour. Let S denote the maximal access time between any two disk locations. By construction, we have

$$T \le (\sum_j T_j) + Sn^{1/3} \quad (8)$$

We also have

$$\sum_{j\in A} T_j \le \sum_{j\in A} k_j S \le \sum_{j\in A} n^{1/3} S \le n^{2/3} S \quad (9)$$

Our main result is the following theorem

Theorem 4.1 *If requests* $\bar{P} = P_0, ..., P_{n-1}$ *are chosen independently on the disk w.r.t a density* $p(r)$ *then with high probability one can construct in polynomial time a tour through* $\bar{P}$ *whose excess time satisfies* $T = O(n^{2/3}\log(n))$.

Proof: Given some fixed $j \in B$, let $b_{i,j}$ $i = 0, ..., k_j - 1$ be the order of the tour on I_j and let $V_{i,j} = \theta_{b_{i+1},j} \ominus \theta_{b_i,j}$ as in the previous section. By lemma 3.3 and the union bound we have with probability at least $1 - 4k_j^{-2\log k_j + 1}$ that $\beta + \log k_j/\sqrt{k_j} \ge V_{i,j} \ge \beta$. The seek time $s_{i,j}$ between requests $P_{b_i,j}$ and $P_{b_{i+1},j}$ is

$$s_{i,j} = \alpha + \frac{|r_{b_i,j} - r_{b_{i+1},j}|}{c}.$$

Since $P_{b_i,j}, P_{b_{i+1},j} \in I_j$ we have $|r_{b_i,j} - r_{b_{i+1},j}| \le n^{1/3}$, hence, the seek time is less than β. Let $a_{i,j}$ denote the access time, then assuming $V_{i,j} \ge \beta \ge s_{i,j}$ we have

$$d_{i,j} = a_{i,j} - \alpha = s_{i,j} + \theta_{b_{i+1},j} \ominus (\theta_{b_i} + s_{i,j}) - \alpha$$

$= s_{i,j} + v_{i,j} - s_{i,j} - \alpha = v_{i,j} - \alpha.$

We conclude that with probability at least $1 - 4k_j^{-2\log k_j + 1}$ we have

$$d_{i,j} \le \beta + \log k_j/\sqrt{k_j} - \alpha = \log k_j/\sqrt{k_j} + n^{-1/3}/c.$$

This leads to $T_j \le \sqrt{k_j}\log k_j + k_j n^{-1/3}/c$. Summing over all j and using the union bound we have

$$\sum_{j\in B} T_j \le (\sum_j \sqrt{k_j}\log k_j) + n^{2/3}/c \tag{10}$$

with probability at least

$$1 - (\sum_j 4k_j^{-2\log k_j+1}) \ge 1 - 4n^{-(2/9)\log n + 2/3} \tag{11}$$

We have $\sum_{j\in B}\sqrt{k_j}\log k_j \le \sum_j \sqrt{k_j}\log k_j$. The latter is a symmetric and convex in the variables k_j and hence, subject to $\sum_j k_j = n$ is maximized when all k_j are equal. Thus, with high probability

$$\sum_{j\in B}\sqrt{k_j}\log k_j \le n^{1/3}(\sqrt{n^{2/3}}\log n^{2/3}) \tag{12}$$

which implies together with equations 8, 9 and 10 that that $T = O(n^{2/3}\log n)$. ∎

5 Variants

We can consider using the same methods the more general case in which the seek function has the form $f(\theta) = \alpha + c\theta^a$ for $a \ge 1$.

Theorem 5.1 *If requests $\bar{P} = P_0, ..., P_{n-1}$ are chosen independently on the disk with respect to a density $p(r)$ and the seek function has the form $f(\theta) = \alpha + c\theta^a$ then with high probability one can construct in polynomial time a tour through $\bar{P}$ whose excess time satisfies $T = O(n^{\frac{a+1}{a+2}}\log(n))$.*

Proof: The proof follows the exact same outline as the proof of theorem 4.1, the only difference is the choice of a constant γ_a so that I_j is defined as

$$I_j = \{i : r_i \in [jn^{-\gamma_a}, (j+1)n^{-\gamma_a})\}$$

We show how to optimize the choice of γ_a. We will use the same notation as in the proof of 4.1 for the corresponding definitions. Given γ_a we let

$$\beta_a = \alpha + f^{-1}(n^{-\gamma_a}) = \alpha + \frac{n^{-\gamma_a/a}}{c^{1/a}}$$

and construct tours through I_j in the same manner as in 4.1. The choice of β guarantees that $\beta \ge s_{i,j}$ which in turn leads to

$$T_j \le \sqrt{k_j}\log k_j + k_j\frac{n^{-\gamma_a/a}}{c^{1/a}}$$

and hence

$$\sum_j T_j \le \sum_j \sqrt{k_j}\log k_j + O(n^{1-\gamma_a/a})$$

As in the proof of 4.1 we may eventually assume that $k_j = n^{1-\gamma_a}$ for all j. Given this value, the tour length T will satisfy w.h.p

$$T \le \sum_j T_j + O(n^{\gamma_a}) \le O(n^{(1+\gamma_a)/2}\log n) + O(n^{1-\gamma_a/a}) + O(n^{\gamma_a})$$

To minimize the sum we equate the first two exponents $\frac{1+\gamma_a}{2}$ and $1 - \frac{\gamma_a}{a}$. This leads to $\gamma_a = \frac{a}{a+2}$ which in turn gives the term $O(n^{\frac{a+1}{a+1}})\log n$. We notice that this term majorizes the third term $n^{\frac{a}{a+2}}$ and we obtain the required result. ∎

When the distribution $p(r,\theta)$ does depend on θ the asymptotic behavior of the excess time changes radically.

Theorem 5.2 *Let $f(\theta) = \alpha + \theta^a$, $a \ge 1$ be a seek function. Let $p(r,\theta)$ be a continuous density such that there exists a point (r_0,θ_0) for which $p(r_0,\theta_0) \ne p(r_0,\theta_0+\alpha)$, then, w.h.p the excess length of the optimal tour through requests $\bar{P}$ chosen independently w.r.t p is at least cn for some constant c which depends on p.*

Proof: We assume that $x = p(r_0,\theta_0) > p(r_0,\theta_0+\alpha) > 0$, the other case being similar. Since p is continuous there is an open neighborhood $p \in U$ such that for all $(r,\theta) \in$U we have $p(r,\theta) - p(r,\theta+\alpha) > x/2$. Let $U_\alpha = U + (0,\alpha)$ be the set obtained from U by a rotation of angle α. Let $\mu(U) = \int_U drd\theta$ be the Lebesgue measure of U. For a set W let $\nu(W) = \int_W p(r,\theta)drd\theta$ be it's measure w.r.t p. Since Lebesgue measure is translation invariant we have $\nu(U) - \nu(U_\alpha) = \int_U p(r,\theta) - p(r,\theta+\alpha) \ge \mu(U)x/2 > 0$. For a set W let W_δ consist of all points (r,θ) whose standard Euclidean distance from W is at most δ. Choose $\delta > 0$ such that $\gamma = \nu(U) - \nu(U_\alpha(\delta)) > 0$. Given $\varepsilon > 0$ let $A(U,\varepsilon)$ be the set of points whose access time from some point in U is less than $\alpha + \varepsilon$. By continuity of the seek function we can choose $\varepsilon(\delta) > 0$ such that $A(U,\varepsilon) \subset U_\alpha(\delta)$. Given any $\gamma > 0$ w.h.p there will be at least $(\nu(U) - \eta)n$ requests in U and at most $(\nu(U_\alpha(\delta)) + \eta)n$ requests in $A(U,\varepsilon)$. Given any tour on $\bar{P}$, each point in U which is not followed on the tour by a point in $A(U,\varepsilon)$ contributes by definition at least ε to the excess time. Choosing η such that $2\eta < \gamma$ we see that with high probability there are at least $\gamma - 2\eta$ such points. taking $\eta to 0$ we see that w.h.p the length of an optimal tour is at least $\gamma\varepsilon n$. ∎

6 conclusions and future work

We have studied the total service time for n requests under realistic assumptions of a non zero unavoidable service time. We have shown that when we make the realistic assumption that requests are uniformly spread in the angular direction the avoidable service time has size $O(n^{2/3} \log n)$. When we assume that the distribution is not uniform in the angles we obtain a linear size contribution to the total service time.

One very interesting open question is to find a tighter bound in the uniform angle case. It is easy to show that the optimal tour length is at least $\Theta(n^{1/2})$. The question is what is the correct power of n in the optimal tour length. In the non uniform angle case it is of interest to compute the actual coefficient of n in the size of the optimal tour. This problem is closely related to the Monge-Kantorovitch transportation problem which is studied in [7, 13] among others. The main difference is that in this case transportation is in a unique direction, that of disk rotation. We hope to return to this issue in a future publication.

Acknowledgments: We thank D. Berend for many helpful conversations and advice concerning this work

References

[1] M. Andrews, M. A. Bender, and L. Zhang, New Algorithms for the Disk Scheduling Problem, *Algorithmica, 32(2)*, 277-301, 2002.

[2] E. Bachmat, Average case analysis for batched disk scheduling and increasing subsequences, *Proceedings of the Symposium on Theory Of Computing (STOC)*, Montreal, 2002.

[3] E. Bachmat, Analysis of disk scheduling with linear seek functions, increasing subsequences and space-time geometry, Submitted, 2005. Available at www.cs.bgu.ac.il/ ebachmat.

[4] J. Cirasella, D. S. Johnson, L. A. McGeoch, and W. Zhang, The Asymmetric Traveling Salesman Problem: Algorithms, Instance Generators, and Tests, *ALENEX 2001 Proceedings, Springer Lecture Notes in Computer Science 2153*, A. L. Buchsbaum and J. Snoeyink (eds.), 32-59.

[5] E.G. Coffman, L.A. Klimko and B.Ryan, Analysis of scanning policies for reducing disk seek times, *SIAM Journal of computing*, 1(3), 1972.

[6] G. Gallo, F. Malucelli and M. Marre, Hamiltonian paths algorithms for disk scheduling, *Universita of Pisa technical report*, 20/94, 1994.

[7] W. Gangbo and R. McCann, The geometry of optimal transportation, *Acta Math.* 177, 1996, 113-161.

[8] C.C. Gotlieb and H.McEwan, Performance of a movable-head disk storage device, *Journal of the ACM*, 20(4):604–623, 1973.

[9] W. Hoeffding. Probability inequalities for sums of bounded random variables, *Journal of the American Statistical Assosiation*, 301, 1963, 13–30.

[10] M. Hofri, Disk scheduling: FCFS vs SSTF revisited. *Communications of the ACM*, 23(11), 1980.

[11] D. Jacobson and J. Wilkes, Disk scheduling algorithms based on rotational position, *HP labs technical report*, HPL-CSP-91-7rev1, 1991.

[12] D. S. Johnson, G. Gutin, L. A. McGeoch, A. Yeo, W. Zhang, and A. Zverovich, Experimental Analysis of Heuristics for the ATSP, *The Traveling Salesman Problem and its Variations*, G. Gutin and A. Punnen, Editors, Kluwer Academic Publishers, Dordrecht, 2002, 445-487.

[13] R. McCann, Exact solutions to the transportation problem on the line, *Proc. Royal Soc. London Ser. A*, 455, 1999, 1341-1380.

[14] A. Riska, E. Riedel and S. Iren, Managing Overload Via Adaptive Scheduling, *in Proceedings of the 1st Workshop on Algorithms and Architecture for Self-Managing Systems*, pages 23-24, San-Diego, CA, June 2003.

[15] Alma Riska, Erik Riedel and Sami Iren, Adaptive disk scheduling for overload management, *in Proceedings of the 1st International Conference on Quantitative Evaluation of Systems, (QEST)*, pages 176-186, The Netherlands, September, 2004.

[16] P. Sanders and B. Vcking, Tail bounds and expectations for random arc allocation and applications, *Combinatorics Probability and Computing*, 12(3), 2003, 301-318.

[17] M. Seltzer, Chen P. and J. Ousterhout, Disk scheduling revisited, *Proceedings of the Usenix technical conference*, 313-324, winter 1990.