Analysis of set-up time models - a metric perspective

Eitan Bachmat *   Tao Kai Lam †   Avner Magen ‡

June 26, 2006

*Dept. of Computer Science, Ben Gurion University, email: ebachmat@cs.bgu.ac.il
†EMC Corporation, email: tlam@emc.com
‡Dept. of Computer Science, University of Toronto, email: avner@cs.toronto.edu

Abstract

We consider model based estimates for set-up time. The general setting we are interested in is the following: given a disk and a sequence of read/write requests to certain locations, we would like to know the total time of transitions (set-up time) when these requests are served in an orderly fashion. The problem becomes nontrivial when we have, as is typically the case, only the counts of requests to each location rather then the whole input, and we can only hope to estimate the required time. Models that estimate the set-up time have been suggested and heavily used as far back as the sixties. However, not much theory exists to enable a qualitative understanding of such models. To this end we introduce several properties such as (i) *super-additivity* which means that the set-up time estimate decreases as the input data is refined (ii) *monotonicity* which means that more activity produces more set-up time, and (iii) an approximation guarantee for the estimate with respect to the worst possible time, by which we can study different models.

We provide criteria for super-additivity and monotonicity to hold for popular models such as the Partial Markov model (PMM). The criteria show that the estimate produced by these models will be monotone for any reasonable system. We also show that the independent reference model (IRM) based estimate functions as a worst case estimate in the sense that the estimate is guaranteed to be at least half of the actual set-up time. Using our criteria we prove that PMM based estimates are always super additive when applied to the special metrics that correspond to seek times of disk drives.

To establish our theoretical results we use the theory of finite metric spaces, and en route show a result of independent interest in that theory, which is a strengthening of a theorem of J.B. Kelly [5] about the properties of metrics that are formed by concave functions on the line.

[1]

[1] A preliminary version of this paper appears in the proceedings of COCOON 2006

1 Introduction

Set-up times which are associated with moving a system from one state to another play a major role in the performance analysis of systems. Perhaps the most glaring example is provided by disk based storage systems in which the states correspond to locations on the disk. In this case the total duration of the movements of the disk's head (from one location to another or from one disk track to another), aka the *set-up time* is the dominant feature in the total service time, and hence a lot of effort is put in order to minimize it by means of reordering the disk's content. Interestingly enough, in this and in other real world applications this becomes a problem with *partial input*. The reason is quite simple: to collect all transition information will be too costly and will render the original optimization useless as the set-up time will be second to the input collection time. Instead, the only information typically available is the state counts, ie the number of times that each state was requested. In graph language this is like finding out the length of a path in a weighted graph where we only know the number of times that each node was visited.

In order to estimate the set-up time, researchers have used stochastic *models*, in other words a stochastic process with parameters that are inherited from the observed count. The simplest of these models, the *Independent Reference Model* (IRM) is very intuitive: the requests at any time are drawn (independently of the previous state) from a distribution proportional to the count vector. This simple model and its generalization, the Partial Markov models (PMM) in which there is a bias toward "staying put", are the most popular models for the analysis of storage system performance; see for for example [1, 3, 4, 8, 10, 12, 13, 16, 17, 18] among many.

In this paper we consider new and basic properties of set-up time estimates and check whether they hold for models such as the IRM and PMM. These properties relate the set-up time estimates to the worst case case and examine the changes in the estimate due to a different way of collecting the data. The applicability of these properties to various models is an evidence to their quality, and moreover they allow for a rigorous study of models that are heavily used, often with not enough underlying rationale. To put things in perspective, it is interesting to note that while the IRM and PMM are some of the oldest models of user access patterns, dating back to the sixties, the basic properties considered above have never been explored. What follows is a brief description of these properties.

Given time intervals $I \subset J$ it is obvious that a system suffers at least as much set-up time during J as it does during I. The *monotonicity* property simply says that the set-up time estimate of the model reflects that fact, ie it gives an estimate for J which is at least as big as the one for I. A model is said to be *super-additive* if the addition of input information (by means of higher resolution of measurements) does not increase the set-up time estimate. It is almost immediate that super-additivity implies monotonicity and that it applies to the worst case time which provides the largest possible set-up time consistent with a given input data. The last property compares the set-up time estimate with the worst case estimate (which is NP hard to compute). Showing that the estimate of a model does not deviate much from the worst case estimate is tantamount to showing that is not over optimistic.

Our results: We show that monotonicity applies to IRM and its extension PMM, *regardless of the metric involved*. We further show that IRM set-up time estimate is a 1/2 approximation to the worst case and that PMM based estimates also approximate the worst case but with smaller approximation constants. Our results concerning super-additivity have the following curious feature: Super additivity holds in the IRM and PMM models *provided* that the "time-metric", ie the times associated with the transition times between pair of states, belongs to the well studied class of metric spaces known as *square Euclidean metrics*. Not all metric spaces belong to this class, but surprisingly, the physical features of the motion of disk drives allows one to show that their time-metrics are members of this class, whence providing a proof of super additivity for IRM and PMM to these I/O systems. These results show that the IRM and in certain cases the PMM can be used to produce reliably conservative estimates which are easy to calculate and that easily lend themselves to compactness-of-input/accuracy tradeoff. Following these observations the first and second authors used the IRM and PMM set-up time estimates as a central ingredient in a commercially available application which dynamically reconfigures data in a disk array. Details of the application and successful results from real production environments are to be presented elsewhere.

Techniques: Our results are first proven for the IRM model. We later provide a formula which expresses the PMM estimate in terms of the IRM estimate by varying the model parameters. Consequently, several properties of the IRM generalize to the PMM. The formula also provides a fast method of computing the PMM estimate directly without computing the associated stationary distribution. Naturally, much of the notions and proofs come from and use the theory of metric spaces. The classes of interest in this discussion are ℓ_1-metrics and square Euclidean metrics, as well as the general class of metrics. In the process of establishing our results we extend a result of Kelly on the properties of invariant metrics on the real line coming from concave functions.

Organization: The rest of the paper is organized as follows. Section 2 Introduces set-up times and discusses some basic definitions and facts from the theory of metric spaces relevant to our discussion. Section 3 describes the basic models which we will study and introduces the concepts of monotonicity, super additivity, dominance and approximation. In section 4 we prove criteria for monotonicity and super additivity of the IRM estimate in terms of metric properties of the set-up time function. We also discuss the relation between the IRM set-up time estimate and those of other models such as the PPM. Finally, Section 5 discusses properties of metric arising from the seek times in disk drives.

2 Preliminaries

2.1 Set-up time

Throughout the paper we let X represent the states of a system. In this section we let n denote the number of states in X. Following [1] section 6.2, we let the function $d : X \times X \longrightarrow \mathbf{R}^+$, be the set-up time function; namely, for $i, j \in X$, $d(i,j) = d_{i,j}$ represents the amount of time which is required to switch the system from state i to state j.

The abstract notion of a state can acquire many different meanings in different applications. For example, the states can refer to different tasks that the system needs to accomplish as in production systems and processors, or, to physical locations where tasks should be conducted as in storage systems. We assume that there is some process which generates a sequence of requests for the states of X.

Given a time interval I let $\mathbf{x}^I = \mathbf{x} = x_1, ..., x_m$ be the sequence of requests for states of X during I. The *Total set-up time* during time interval I is simply the sum of the set-up times between consecutive requests

$$T(\mathbf{x}) = \sum_{j=1}^{m-1} d(x_j, x_{j+1})$$

In some cases we are not given the sequence of requests (a trace) but rather some partial information about the sequence $\mathbf{x}$. We wish to estimate the total set-up time of the sequence using the information available to us. In this paper we shall assume that the partial information available to us is the activity vector $\mathbf{a} = \mathbf{a}_I = (a_1, ..., a_n)$, where a_i is the number of requests for state i during time interval I. We will assume that in general $\mathbf{a}$ can be any vector with integer nonnegative entries. We let $a = \sum_i a_i$ be the total number of requests.

2.2 Metric Spaces

The theory of finite metric spaces will be used in the statements and proofs of our results. The following section provides some basic definitions and facts about metric spaces which will be needed later on.

We continue with a few standard definitions. A pair (X, d) where X is a set and d is a function $d : X \times X \longrightarrow \mathbf{R}^+$ is called a *metric-space* if (i) $d(x, x) = 0$ for all $x \in X$ and $d(x, y) > 0$ for $x \neq y$, (ii) $d(x, y) = d(y, x)$ for all $x, y \in X$ and (iii) $d(x, y) + d(y, z) \geq d(x, z)$ for all $x, y, z \in X$. If instead of property (i) we only require that $d(x, x) = 0$ we say that (X, d) form a *semi-metric*. If we do not require the symmetry property, we say that (X, d) form a *Pseudometric*. One can "symmetrize" such an object by taking $d^*(x, y) = (d(x, y) + d(y, x))/2$. It can be easily seen that d^* satisfies 1' and 3' if d does. Set-up time functions can be reasonably assumed to satisfy the triangle inequality since one way to switch from state x to state z is to first switch from x to y and then from y to z. Set-up time functions cannot always be assumed to be symmetric as can be seen from rotational latency in disk drives.

Certain metric spaces are induced by norms. The ℓ_p norm on $\mathbf{R}^n$ is $\|\mathbf{x}\|_p = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$ where $\mathbf{x} = (x_1, \ldots, x_n)$. For two vectors in $\mathbf{R}^n$ $\mathbf{x}$ and $\mathbf{y}$ this defines a distance $d_p(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_p$. A metric space (X, d) is called an ℓ_p-metric if there exists a mapping $\phi : X \longrightarrow \mathbf{R}^n$ such that $d(x, y) = \|\phi(x) - \phi(y)\|_p$ for all $x, y \in X$. We sometimes say *Euclidean metric* instead of ℓ_2-metric. A space (X, d) is *square Euclidean* if $(X, \sqrt{d})$ is Euclidean.

2.2.1 Some basic results about metric spaces

Assume (X, d) is a finite metric space, i.e., $X = \{x_1, \ldots, x_n\}$. There are two classical criteria for it to be Euclidean.

- Schoenberg's criterion: (X, d) is Euclidean if and only if for all n real numbers $v_1, \ldots, v_n$ with $\sum_i v_i = 0$ we have $\sum_{i,j} v_i v_j d_{i,j}^2 \leq 0$.
- Cayley's criterion: Consider the order $n-1$ matrix M with entries $M_{i,j} = d^2(x_i, x_n) + d^2(x_j, x_n) - d^2(x_i, x_j)$, $i, j = 1, \ldots, n-1$. Then (X, d) is Euclidean if and only if the matrix M is positive semi definite, ie, all of its eigenvalues are nonnegative.

We say that a metric (X, d) is L_1 if there exist functions f_x, $x \in X$ such that $d_{x,y} = \int_{\mathbf{R}} |f_x(t) - f_y(t)| dt$. It is known that a finite metric space is L_1 iff it is ℓ_1. Another well known fact we later use is that every ℓ_1-metric is square Euclidean [7]. Square Euclidean distances do not necessarily satisfy the triangle inequality. They do, however, have many desirable properties, and have found much use in the theory of algorithms, especially in the approximation of NP-hard problems; this is since it is possible to optimize linear objective functions over these distances. The most celebrated example is the recent paper of Arora, Rao and Vazirani [2] where an $O(\sqrt{\log n})$ approximation algorithm to sparsest-cut is presented. See also [9] for the way square-Euclidean distances are harnessed to generate good algorithms. In [6] it is shown, that some metric spaces are very far from any square Euclidean metrics.

A distance function can be defined on the line given a real positive function F with certain properties. We define the distance d_F between i and j as $d_F(i, j) = F(|i - j|)$. This framework was introduced in [11, 14], and is utilized in this paper. We note that if F is convex then d_F satisfies the triangle inequality and thus provides a metric.

3 Models and their properties

Recall that our input is an activity vector, that is the count of requests to the different states; however, in order to know the total set-up time we need to know the actual sequence of requests. A model to estimate a set-up time is an interpretation of an activity vector as a distribution over sequences, and the resulting estimate is then the expected set-up time for a random sequence drawn from this distribution. For example some models will interpret an activity vector $(100, 100)$ as a uniform distribution of sequences that visit either location 1 or 2, while other will consider the distribution in which either all first 100 requests are for the first location or all of them were for the other; clearly the two different models in the above example will produce very different time estimates. Formally, a model M is a map which assigns to each activity vector $\mathbf{a}$ a probability measure $\mu(\mathbf{a})$ on sequences of requests of length $a = \sum_{i \in x} a_i$. Given a model M, a set-up time function d and a time interval I with activity vector $\mathbf{a}_I$, the total set-up time during I is

$$T(\mathbf{a}, d : M) = \mathbf{E}(T) = \int_{\mathbf{x}} T(\mathbf{x}) d\mu(\mathbf{x})$$

We will refer to $T(\mathbf{a}, d; M)$ as the model (or model based) estimate. [2].

3.1 Examples of models and estimates

We now describe several models M and their associated set-up time estimates.

The IRM (independent reference model) The IRM models independent random requests to states in X, taking into account that the different states are not uniformly popular. The model is parameterized by a probability distribution $p = p_i$ on the set of states X. The model itself is then given by the product measure on X^a. The product measure reflects an underlying assumption that requests are generated independently of each other. To be compatible with the observed activity vector we set the request probability for state i to be $p_i = a_i/a$ and the length of the generated sequence to be a. For this model the expected total set-up time is

$$T(\mathbf{a}, d; IRM) = a \sum_{i,j} p_i p_j d_{i,j} = \frac{1}{a} \sum_{i,j} a_i a_j d_{i,j}$$

We will refer to $T(\mathbf{a}, d; IRM)$ as the IRM estimate. For ease of notation we will sometimes use $T(\mathbf{a}, d)$ instead of $T(\mathbf{a}, d; IRM)$.

The worst case (supremum) model In the worst case model W we assume that the sequence of states during time interval I was the sequence which maximizes the total set-up time among all sequences which are consistent with the vector $\mathbf{a}$. The measure is thus a δ measure on the worst

[2] During time interval I there were $a - 1$ transitions between states. It turns out to be more convenient to estimate the total time needed for a transitions, thus we add a "virtual" transition between the last state and the first state. It is important to note that in all cases of practical interest, a is a large number and the addition of the last "virtual" transition has negligible effect.

case sequence. Consequently,

$$T(\mathbf{a}, d; W) = \max \sum_{i=1}^{a} d_{x_i, x_{i+1}}$$

where the maximum is over all sequences of states in X, of length a that agree with the frequency vector $\mathbf{a}$ and $x_1 = x_{a+1}$. We refer to $T(\mathbf{a}, d; W)$ as the worst case estimate.

The PMM (Partial Markov models) A *Partial Markov model*, or PMM for short is a Markov chain which models a "lazy" walk of the IRM. In other words, at a state i there is a probability r_i of not moving to another state, and in the event of a move, the next state is j with probability q_j, independent of the current requested state. Consequently, the transition probabilities of moving from i to j are $p_{i,j} = (1 - r_i)q_j$ for $i \neq j$ and $p_{i,i} = r_i + (1 - r_i)q_i$. Here $0 \leq r_i, q_i \leq 1$. We call the vector $\mathbf{r} = (r_i)$ the *locality* vector of the model. Given a locality vector $\mathbf{r}$ and an observed activity vector $\mathbf{a}$ for some time interval I there exists a unique partial Markov model P which is compatible with $\mathbf{r}$ and $\mathbf{a}$. By compatibility we mean that $\mathbf{r}$ is the locality vector of P and $\mathbf{a}/a$ is the stationary distribution of P which expresses the expected reference probabilities in the model P. Fix the vector $\mathbf{r} = (r_i)$. We let $P^{\mathbf{r}}$ denote the partial Markov model which for each interval I uses the model P compatible with $\mathbf{r}$ and $\mathbf{a}_I$ to model the request stream during I (note that $P^{\mathbf{0}}$ is simply the IRM). The $P^{\mathbf{r}}$ estimate is

$$T(\mathbf{a}, d; P^{\mathbf{r}}) = a(\sum_{i,j} (a_i/a) P^{\mathbf{r}}_{i,j} d_{i,j})$$

Partial Markov models are useful in capturing locality of reference phenomenon, [1, 3], which means that a request to state i is likely to be followed by another request to state i within a short time span. Many applications naturally exhibit this type of behavior. The larger the entries of the locality vector $\mathbf{r}$, the more likely states are to repeat in succession. In the partial Markov model the number of repetitive successions is distributed geometrically.

3.2 Properties of models

We introduce notions which will allow us to examine the behavior of model based estimates with regards to changes in the input data and to compare estimates for different models.

super additivity Let I be a time interval and let $I_1, ..., I_k$ be a subdivision of I into subintervals. Accordingly, we have $\mathbf{a}_I = \sum j = 1^k \mathbf{a}_{I_j}$. A model M is said to be *super additive* with respect to a set-up time function d if the inequality

$$T(\mathbf{a}_I, d; M) \geq \sum_{j=1}^{k} T(\mathbf{a}_{I_j}, d; M) \qquad (1)$$

always holds. Super additivity may be interpreted as stating that the addition of input information, namely, $\mathbf{a}_{I_j}$ instead of $\mathbf{a}_I$, never increases the estimate.

monotonicity We say that a vector $\mathbf{a} = (a_i)$ dominates a vector $\mathbf{b} = (b_i)$ if for all i, $a_i \geq b_i$. We use the notation $\mathbf{a} \geq \mathbf{b}$ to denote dominance. A model M is said to be *monotone* with respect to d if for any pair of time intervals $I \subset J$ we have $T(a_I, d; M) \leq T(a_J, d; M)$, or stated otherwise, for any pair of vectors $\mathbf{a}, \mathbf{b}$ with nonnegative entries and such that $\mathbf{a} \geq \mathbf{b}$ we have

$$T(\mathbf{a}, d; M) \geq T(\mathbf{b}, d; M) \qquad (2)$$

dominance We say that a model M_1 *dominates* a model M_2 with respect to a set-up time function d, if for all activity vectors $\mathbf{a}$ we have

$$T(\mathbf{a}, d; M_1) \geq T(\mathbf{a}, d; M_2) \qquad (3)$$

approximation Let $0 < \alpha < 1$. Given a set up function d, a model M_1 is said to be provide an α *approximation* to a model M_2 (and vice versa) if for any activity vector $\mathbf{a}$ we have

$$\alpha \leq \frac{T(\mathbf{a}, d; M_1)}{T(\mathbf{a}, d; M_2)} \leq \frac{1}{\alpha} \qquad (4)$$

We say that a model M is *conservative* if it α approximates the worst case model W for some $\alpha > 0$.

4 Metric space criteria for properties of models

In this section we establish criteria for monotonicity and super additivity of the IRM and PMM estimates in terms of metric properties of the set-up time function d. We also establish a criterion for the IRM estimate to be a 1/2 approximation to the worst case estimate and study the relation between the IRM estimate and the PMM estimate.

Theorem 1 *(A criterion for Super additivity of the IRM) The IRM estimate is super additive with respect to d if and only if d is square Euclidean*

proof: It is enough to establish super additivity for a subdivision of I into two subintervals, that is to show that for all nonnegative vectors $\mathbf{a} = (a_i), \mathbf{b} = b_i$,

$$T(\mathbf{a}+\mathbf{b}, d) \geq T(\mathbf{a}, d) + T(\mathbf{b}, d) \tag{5}$$

Let $a = \sum_i a_i$ and $b = \sum_i b_i$. Then

$$\begin{aligned}
&T(\mathbf{a}+\mathbf{b}, d) - T(\mathbf{a}, d) - T(\mathbf{b}, d) \\
&= \sum_{i\neq j} \frac{(a_i+b_i)(a_j+b_j)d_{ij}}{a+b} - \sum_{i\neq j} \frac{a_i a_j d_{ij}}{a} - \sum_{i\neq j} \frac{b_i b_j d_{ij}}{b} \\
&= \frac{1}{ab(a+b)} \sum_{i\neq j} d_{ij}(a_i b_j ab + a_j b_i ab - a_i a_j b^2 - b_i b_j a^2) \\
&= \frac{1}{ab(a+b)} \sum_{i\neq j} d_{ij}(a_i b - b_i a)(b_j a - a_j b) \\
&= -\frac{ab}{a+b} \sum_{i\neq j} d_{ij} \left(\frac{a_i}{a} - \frac{b_i}{b}\right)\left(\frac{a_j}{a} - \frac{b_j}{b}\right).
\end{aligned}$$

Setting $v_i = \dfrac{a_i}{a} - \dfrac{b_i}{b}$, we get

$$T(\mathbf{a}+\mathbf{b}, d) - T(\mathbf{a}, d) - T(\mathbf{b}, d) = -\frac{ab}{a+b} \sum_{i\neq j} d_{ij} v_i v_j. \tag{6}$$

We note that $\sum_i v_i = 0$, hence by Schoenberg's criterion the IRM estimate is super additive if d is square Euclidean. Conversely if the IRM estimate is super additive then

$$\sum_{i,j} d_{ij} v_i v_j \leq 0$$

for all $\mathbf{v}$ of the form $\mathbf{a}/a - \mathbf{b}/b$ where $\mathbf{a}, \mathbf{b}$ are vectors with integer non negative entries. After scaling we may deduce that the property holds whenever $\mathbf{a}, \mathbf{b}$ have rational non negative entries and by density of the rationals for all $\mathbf{a}, \mathbf{b}$ with non negative entries. Every vector $\mathbf{v} = (v_1, \ldots, v_h)$ such that $\sum_i v_i = 0$ has a multiple of the form $\frac{1}{a}\mathbf{a} - \frac{1}{b}\mathbf{b}$, where $\mathbf{a}, \mathbf{b}$ have non negative entries. Indeed if $a_i = \max\{v_i, 0\}$ and $b_i = \max\{-v_i, 0\}$, then $a = b$ and $\frac{1}{a}\mathbf{a} - \frac{1}{b}\mathbf{b} = \frac{1}{a}\mathbf{v}$, hence Schoenberg's criterion holds and d is square Euclidean. *q.e.d*

Theorem 2 *(Comparison of the IRM and PMM estimates and criteria for monotonicity)*

1. *If* $\mathbf{a}$ *is an activity vector and* $\mathbf{r}$ *a locality vector, let* $\mathbf{a}^{\mathbf{r}}$ *be the vector with entries* $\mathbf{a}^{\mathbf{r}}_i = a_i(1 - r_i)$*, then*

$$T(\mathbf{a}, d; P^{\mathbf{r}}) = T(\mathbf{a}^{\mathbf{r}}, d; IRM)$$

2. *The IRM estimate is monotone with respect to* d *if and only if for any pair of locality vectors* $\mathbf{r}$ *and* $\mathbf{s}$*, such that* $\mathbf{r} \leq \mathbf{s}$*,* $P^{\mathbf{r}}$ *dominates* $P^{\mathbf{s}}$ *and in particular the IRM estimate dominates the set of PMM estimates with respect to* d*.*

3. *The IRM estimate is monotone with respect to* d *if and only if the matrices* $B(k, d)_{i,j} = d_{i,k} + d_{k,j} - d_{i,j}$ *define a nonnegative quadratic form when restricted to vectors with nonnegative entries. In particular, if* d *is a pseudo metric or square Euclidean then the IRM estimate is monotone with respect to* d*.*

4. *If the IRM estimate is monotone with respect to* d *then* $\sqrt{d}$ *satisfies the triangle inequality.*

5. *All PMM* $P^{\mathbf{r}}$ *are super additive with respect to square Euclidean metrics.*

proof: Let $\mathbf{a} = \mathbf{a}_I$ be an activity vector. Recall that in a PMM, each state has two parameters r_i, q_i and the transition matrix $M_{i,j}$ is $(1 - r_i)q_j$ if $i \neq j$ and $r_i + (1 - r_i)q_i$ if $i = j$. Let π be the stationary distribution of the chain, namely $\pi M = \pi$. By the requirement of compatibility with the observed activity vector we have $\pi = \mathbf{a}/a$. We can now express the q_i-s in terms of π and the r_i-s :

$$\pi_j = \sum_i \pi_i M_{ij} = \pi_j M_{jj} + \sum_{i \neq j} \pi_i M_{ij} = \pi_j r_j + \sum_i \pi_i (1 - r_i) q_j.$$

Note that $a_i^{\mathbf{r}} = a\pi_i(1-r_i)$. We get that $q_j = a_j^{\mathbf{r}}/a^{\mathbf{r}}$ for all j where $a^{\mathbf{r}} = \sum_i a_i^{\mathbf{r}}$. Using the fact that $d_{ii} = 0$ we get that

$$T(\mathbf{a}; d, P^{\mathbf{r}}) = a(\sum_{i,j} \pi_i P^{\mathbf{r}}_{ij} d_{ij}) = a(\sum_{i,j} \pi_i (1-r_i) q_j d_{ij})$$

$$= \sum_{i,j} \frac{a_i^{\mathbf{r}} a_j^{\mathbf{r}}}{a^{\mathbf{r}}} \cdot d_{ij} = T(\mathbf{a}^{\mathbf{r}}, d; IRM)$$

proving the first statement of the theorem.

It is easy to see that if $\mathbf{b} \leq \mathbf{a}$ then there is a locality vector $\mathbf{r}$ such that $b = \mathbf{a}^{\mathbf{r}}$. Conversely, if $\mathbf{r} \geq \mathbf{s}$ then $\mathbf{a}^{\mathbf{r}} \geq \mathbf{a}^{\mathbf{s}}$. These simple observations together with part 1 imply part 2.

To prove part 3, we check the sign of the partial derivatives of $T(\mathbf{a}, d)$ with respect to a_k (where $k \in X$ is an arbitrary element).

$$\begin{aligned} &\frac{\partial}{\partial a_k} T(\mathbf{a}, d) \\ &= \frac{a(\sum_i a_i d_{ik} + \sum_j a_j d_{jk}) - \sum_{i,j} a_i a_j d_{ij}}{a^2} \\ &= \frac{1}{a^2} \sum_{i,j} a_i a_j (d_{ik} + d_{jk} - d_{ij}) = \frac{1}{a^2} \mathbf{a} B \mathbf{a}^t \end{aligned}$$

where $B = B(k, d)$ is the matrix with ij entry $d_{ik} + d_{jk} - d_{ij}$. Assume that for all k, $B(k, d)$ is positive semi definite on vectors with nonnegative entries then $\frac{\partial}{\partial a_k} T(\mathbf{a}, d) \geq 0$ for all k and all activity vectors $\mathbf{a}$. It follows from the Mean-value Theorem that if $\mathbf{a} \geq \mathbf{b}$ then $T(\mathbf{a}, d) \geq T(\mathbf{b}, d)$. Conversely if there are $\mathbf{a} \geq 0$ and k such that $\mathbf{a} B(k, d) \mathbf{a}^t < 0$ then taking $\mathbf{b}$ which is identical to $\mathbf{a}$ except that b_k is slightly smaller than a_k we get $T(\mathbf{a}, d) < T(\mathbf{b}, d)$, which proves the first statement of part 3.
If d is a semi-metric then B has nonnegative entries and so $\mathbf{a} B(k, d) \mathbf{a}^t \geq 0$ and if d is square Euclidean then by Cayley's criterion $\mathbf{a} B(k, d) \mathbf{a}^t \geq 0$ which completes part 3.

To prove part 4, we assume that $\sqrt{d}$ is not a metric and show that the IRM estimate is not monotone with respect to d. Without loss of generality $\sqrt{d_{12}} + \sqrt{d_{13}} < \sqrt{d_{23}}$. Let $a_1 = \sqrt{d_{23}}, a_2 = \sqrt{d_{13}}, a_3 = \sqrt{d_{12}}$ and $a_i = 0$ for $i > 3$. We claim that the transition-time in the IRM estimate is strictly smaller than that of a PMM with $r_1 > 0$ and $r_i = 0$ for $i > 1$. Indeed,

$$T(P^{\mathbf{r}}, \mathbf{a}; d)$$

$$
\begin{aligned}
&= \frac{a_1(1-r_1)a_2d_{12}+a_1(1-r_1)a_3d_{13}+a_2a_3d_{23}}{a_1(1-r_1)+a_2+a_3} \\
&= \sqrt{d_{12}d_{13}d_{23}} \cdot \frac{(1-r_1)(\sqrt{d_{12}}+\sqrt{d_{13}})+\sqrt{d_{23}}}{(1-r_1)\sqrt{d_{23}}+\sqrt{d_{13}}+\sqrt{d_{12}}} \\
&= \sqrt{d_{12}d_{13}d_{23}} \cdot \frac{(\sqrt{d_{12}}+\sqrt{d_{13}}+\sqrt{d_{23}})-r_1(\sqrt{d_{12}}+\sqrt{d_{13}})}{(\sqrt{d_{12}}+\sqrt{d_{13}}+\sqrt{d_{23}})-r_1\sqrt{d_{23}}} \\
&> \sqrt{d_{12}d_{13}d_{23}} \\
&= T(\mathbf{a};d).
\end{aligned}
$$

The inequality is due to $\sqrt{d_{12}}+\sqrt{d_{13}} < \sqrt{d_{23}}$ and $r_1 > 0$. Finally, part 5 is proven by using Theorem 1, part 1 of the current theorem, and the fact that $\mathbf{a^r}+\mathbf{b^r}=(\mathbf{a}+\mathbf{b})^{\mathbf{r}}$. *q.e.d*

We next show that the IRM estimate is a 1/2 approximation to the worst case.

Theorem 3 *(Comparison of the IRM estimate and worst case estimate) If d satisfies the triangle inequality then for all activity vectors* $\mathbf{a}$ *we have*

$$2T(\mathbf{a},d;IRM) \geq T(\mathbf{a},d;W) \tag{7}$$

where W is the worst case model.

proof: Assume first that the activity vector is the vector $(1,1,\ldots,1)$. The IRM estimate here is $\frac{1}{n}\sum_{i,j} d_{i,j}$, while the worst case estimate is the length of the longest Hamiltonian cycle in the complete graph on X with edge weights given by d. Assume without loss of generality that the longest Hamiltonian path in X is $1,2,\ldots,n$. Since d satisfies the triangle inequality we have for $1 \leq i < n$ and for $j \in X$ $d_{i,i+1} \leq d_{i,j}+d_{j,i+1}$ (the $n+1$ point coincides with the first point). Summing over all i,j we get

$$n\sum_{i=1}^{n} d_{i,i+1} \leq 2\sum_{i,j} d_{i,j}.$$

Therefore $2T(\mathbf{a},d;IRM) \geq T(\mathbf{a},d;W)$. To complete the proof we need to consider a general activity vector $(a_1,\ldots,a_n)$. Let X' be the metric space with a points that is composed of groups of a_j points of type j. Given d on X we induce a metric on X' by letting the distance between a point of type i and a point of type j be $d_{i,j}$. Clearly X' also satisfies the triangle inequality. We have thus reduced the problem to the case of the activity vector $(1, 1, \ldots, 1)$ and are done. *q.e.d*

The following result is easily derived from the last two

Theorem 4 *Let $P^{\mathbf{r}}$ be a PMM. Let $r_{max} = Max_i r_i$ be the maximal entry of the locality vector $\mathbf{r}$. If d satisfies the triangle inequality then $P^{\mathbf{r}}$ estimate is a $\frac{1-r_{max}}{2}$ approximation to the worst case estimate and in particular $P^{\mathbf{r}}$ is conservative.*

proof: By part 1 of theorem 1 we have $T(a, d; P^{\mathbf{r}}) = T(\mathbf{a}^{\mathbf{r}}, d; IRM)$. Let $\mathbf{b} = (1 - r_{max})\mathbf{a}$ then by definition of r_{max} $\mathbf{b} \leq \mathbf{a}^{\mathbf{r}}$. Since d satisfies the triangle inequality the IRM is monotone with respect to d by part 3 of theorem 1, Combining with theorem 3 we get

$$T(\mathbf{a}, d; P^{\mathbf{r}}) = T(\mathbf{a}^{\mathbf{r}}, d; IRM) \geq T(\mathbf{b}, d, IRM)$$

$$= (1 - r_{max})T(\mathbf{a}, d, IRM) \geq \frac{1 - r_{max}}{2} T(\mathbf{a}, d; W)$$

as desired. *q.e.d*

5 Set-up Time Functions of a Disk

5.1 Super additivity

In this section we show that the radial seek time function of a disk drive, which is the standard set-up function in storage system research is an l_1-metric and in particular is square Euclidean. From this we conclude that the IRM and PMM estimates are super additive when applied to disk seek times. Data on disk drives resides on *tracks* which form concentric circles of varying radii r around the center of a platter. To get from a track at radius r_1 to another track at radius r_2 the head of the device performs a radial motion. The time it takes the disk head to perform this radial motion is known as (radial) seek time. The head starts and ends with no radial velocity and must first accelerate, reach a maximal speed, and then decelerate towards the targeted track. The acceleration and deceleration processes are invariant under translation. Furthermore as the distance $|r_1 - r_2|$ grows the head spends more time at higher speeds and so the average velocity during the transition increases. Consequently the time it takes to seek from r_1 to r_2 has the form

$$d_F(r_1, r_2) = F(|r_1 - r_2|)$$

where F is a concave non decreasing function (note that the slope of F is inverse-proportional to the peak velocity during the transition).

The radial seek function d_F is the standard set-up time function for disk where F is determined by the physical characteristics of the drive itself. If we let X be the set of data locations on the disk then a theorem of Kelly proved in [5] can be interpreted as stating that (X, d_F) is square Euclidean. We prove a stronger result of independent interest using a much simpler proof.

Theorem 5 *Let F be a concave nondecreasing function with $F(0) = 0$ and let $X \subset \mathbf{R}$. Then (X, d_F) is an l_1 metric space.*

proof: Let $X = \{x_1, \ldots, x_n\}$. Consider

$$Y = \{|x_i - x_j| : 1 \leq i, j \leq n\}$$

the set of possible distances in X, and order the elements of Y as $0 = y_0 < y_1 < y_2 < \ldots < y_m$. let G be the piecewise linear function which

(i) coincides with F on Y

(ii) is linear on all intervals $[y_i, y_{i+1}]$ and

(iii) is constant on $[y_m, \infty)$ (that is, gets the value $F(y_m)$ there).

Obviously $(X, d_F) = (X, d_G)$ since $F = G$ on the set of all relevant values Y, so it is enough to prove the claim for G, which is also non decreasing and concave. We now define functions $H_{s,t}$ as follows.

$H_{s,t}(x) = sx$ if $x < t$ and st otherwise.

We also let $s_i = \frac{G(y_i) - G(y_{i-1})}{y_i - y_{i-1}}$ be the sequence of slopes of G. We now claim that G is a convex combination of functions of the form $H_{s,t}$.

The proof proceeds by induction on m. If $m = 0$ then $G = H_{1,0} = 0$. For $m > 0$, look at the function $\tilde{G} = G - H_{s_m, y_m}$. It is not hard to see that $\tilde{G}(0) = 0$, $\tilde{G}$ is constant beyond y_{m-1} and is piecewise linear with breakpoints $y_1, \ldots, y_{m-1}$. A piecewise linear function is concave and nondecreasing if and only if its slopes are decreasing and nonnegative, and so $s_1 \geq s_2 \geq \ldots \geq s_m \geq 0$, and similarly $s_1 - s_m \geq s_2 - s_m \geq \ldots \geq s_{m-1} - s_m \geq 0$. But, these are the slopes of G' and it is therefore concave and nondecreasing. We may now apply the induction hypothesis to $\tilde{G}$ and this proves the claim.

Since a sum of ℓ_1-metrics is also an ℓ_1-metric, we are left with the task of showing that for a function $F = H_{s,y}$, the resulting metric d_F is an ℓ_1-metric. Notice that $d_F(i,j) = s \cdot \min\{|i-j|, y\}$. Let $f_i = \frac{1}{2s}\chi_{[x_i, x_i+y]}$ be the function whose value is $\frac{1}{2s}$ on the interval $[x_i, x_i + y]$ and zero otherwise. It is easy to see that for any $i, j \in \mathbf{R}$

$$d_F(i,j) = s \cdot \min\{|i-j|, y\} = \int_{\mathbf{R}} |f_i(x) - f_j(x)| dx$$

This shows that d_F is an L_1 metric and hence l_1. *q.e.d*

Combining theorem 1 part 5 with theorem 5 we get

Theorem 6 *All PMM models and in particular the IRM are super additive with respect to the seek time function d_F for any physical disk drive.*

Remark 1: Radial motion in a disk is one dimensional. It is natural to replace $\mathbf{R}$ in Theorem 5 by $\mathbf{R}^n$ and thus consider motion along straight lines in higher dimensions with acceleration and deceleration. This may describe for instance the movement of a robot in the plane. This problem has been studied in detail by Von Neumann and Schoenberg, [11, 14]. They present a complete, yet implicit, characterization of functions F which lead to square Euclidean metrics in terms of spherical functions. In particular it can be shown from their results that even for $n = 2$ the functions $H_{s,t}$ do not yield square Euclidean metrics. On the other hand the functions $F(x) = x^c$, $0 < c < 2$ do yield square Euclidean functions for all n. We conjecture that all translation invariant metrics on the line are square Euclidean.

6 Conclusions and future work

We have introduced several natural properties of set-up time estimates and studied them for the IRM and PMM. We have shown that the IRM estimate satisfies monotonicity which is a “sanity check” for set-up time estimates, and further that the IRM is an easily computable approximation to the worst case estimate. We have also related the PMM with IRM estimates showing that the first inherits many of the properties of the second. In the specific but important context of seek functions in disk drives we showed that the IRM and PMM share another formal property that holds for worst case estimates namely super additivity. It would be interesting to explore monotonicity, super additivity and various approximation and dominance relations among other models. One interesting class of examples are the renewal models which were suggested by Opderbeck and Chu in [8]. The IRM is a special case of such models where the renewal model is based on exponential inter-arrival times. It would be interesting to investigate other cases such as hyperexponential, gamma or Pareto bounded heavy tail distributions. Such an investigation will likely require refined definitions for properties such as monotonicity and super additivity since the associated models are not Markovian.

Acknowledgments: We would like to thank Timothy Chow for helpful discussions regarding a preliminary version of this paper.

References

[1] Aven O.I., Coffman E.G. and Kogan Y.A. *Stochastic analysis of computer storage*, D.Reidel publishing, 1987.

[2] Arora, S. and Rao, S. and Vazirani, U., *Expander flows, geometric embeddings and graph partitioning*, STOC '04: Proceedings of the thirty-sixth annual ACM symposium on Theory of computing, 2004.

[3] Coffman E.G. and Denning P.J. *Operating systems theory,* Prentice Hall, 1973.

[4] Grossman D.D. and Silverman H.F. Placement of records on a secondary storage device to minimize access time *J. of the ACM 20,* 429-438, 1973.

[5] Kelly J.B. Hypermetric spaces and metric transforms, in *Inequalities III*, edited by O.Shisha, 149-158, 1972.

[6] Linial, N. and Magen, A., Least-distortion Euclidean embeddings of graphs: products of cycles and expanders, *Journal of Combinatorial Theory. Series B 79*, 157–171, 2000.

[7] Lovasz L., Pyber L., Welsh D.J.A. and Ziegler G.M. Combinatorics in pure mathematics, in *Handbook of combinatorics , MIT press ,North Holland* 2039-2082, 1995.

[8] Opderbeck H. and Chu W.W., The renewal model for program behavior, *SIAM J. of computing 4,* 356–374, 1975.

[9] Schulman, L., Clustering for Edge-Cost Minimization, in *Proceedings of the Thirty Second Symposium on Theory of Computing, STOC,* 547–555, 2000.

[10] Shriver E., Performance modeling for realistic storage devices, *PhD Thesis,* New York University, 1997.

[11] Schoenberg I.J. Metric spaces and completely monotone functions, *Annals of Math. 39(4),* 811-841, 1938.

[12] Triantafillou P., Christodoulakis S. and Georgiadis C., Optimal data placement on disks: A comprehensive solution for different technologies, *IEEE Transactions on Knowledge and Data Engineering 12,* 324-330, 2000.

[13] Vanichpun S. and Makowski A., The output of a cache under the independent reference model - where did all the locality of reference go?, *Proceedings of SIGMETRICS*, 2004.

[14] Von Neumann J. and Schoenberg I.J. Fourier integrals and metric geometry, *Transactions of the AMS 50,* 226-251, 1941.

[15] Wang M., Ailamaki A. and Faloutsos C. Capturing the spatio-temporal behavior of real traffic data, in *Proc. of Performance 2002,* Rome 2002.

[16] Wong C. K. Minimizing expected head movement in one dimensional and two dimensional mass storage systems, *Computing Surveys 12,* 167-180, 1980.

[17] Wong C. K. *Algorithmic Studies in mass storage systems*, computer science press, 1983.

[18] Yue P.C. and Wong C.K., On the optimality of the probability ranking scheme in storage applications, *J. of the ACM 20,* 624-633, 1973.