Purdue University

Purdue e-Pubs

Department of Computer Science Technical Reports

Department of Computer Science

1989

On the Optimality of the Counter Scheme for Dynamic Linear Lists

Micha Hofri

Hadas Shachnai

Report Number:

89-927

ON THE OPTIMALITY OF THE COUNTER
SCHEME FOR DYNAMIC LINEAR LISTS

Micha Hofri
Hadas Shachnai

CSD-TR-927
November 1989

ON THE OPTIMALITY OF THE COUNTER SCHEME FOR DYNAMIC LINEAR LISTS

Micha Hofri† and Hadas Shachnai

Computer Science Department
Technion - Israel Institute of Technology

November 1989

ABSTRACT

We consider policies that manage fixed-size dynamic linear lists, when the references follow the *independent reference model*. We define the *counter scheme*, a policy that keeps the records sorted by their access frequencies, and prove that among all deterministic policies it produces the least expected cost of access, at any time.

1. Introduction

We consider a linear list of n records, $n \geq 2$. An access to R_i requires a sequential search of the list starting at the header, till R_i is encountered. The cost of a single access is defined to be the number of keys examined in the search.

Assumption: The reference history is a series of independent multinomial trials, with fixed but unknown reference-probability vector (*rpv*) $p = (p_1, \cdots, p_n)$. This is the *independent reference model (irm)*.

The problem of minimizing the expected access cost, using dynamic reorganization of the list, has been widely studied. Most of the permutation rules which incur no storage overhead, at times called *memory free*, are variations of two basic methods:

Move To the Front (*MTF*), which places an accessed record at the head of the list, leaving the other elements untouched.

Transposition Rule (*TR*), which advances the referenced record one step ahead by an interchange with its preceding neighbour.

† Currently at the Department of Computer Science, Purdue University, W. Lafayette, IN 47907, USA.

Rules that use extra storage are naturally less appealing compared with the previous methods. However, their relative efficiency in the list reorganization process might compensate for their space complexity. We focus on *Counter Scheme* (CS), which handles the list in the following manner:

A frequency counter c_i stores the number of accesses to each of the records R_i, $1 \le i \le n$, throughout the reference history. The list is preserved in nonincreasing order of the counter values.

When asymptotic (expected) cost is considered, the CS achieves the optimum; in this sense it bests all other common permutation rules. It is also known to have advantages in the finite horizon case, when the average access cost following a *finite* sequence of requests to the list is considered. This was shown by Lam *et al.* (1981) when analyzing their *Generalized Counter Scheme*, a special instance of which is the above CS. They proved that—based on the last criterion—CS is better than any other possible *counter based* method.

In the following discussion we strengthen their result and prove that CS is optimal among **all** realizable policies with respect to the average cost at the mth request, $m \ge 1$.

From a statistical point of view this is hardly surprising: the optimal order only depends on the ranking of the probabilities $\{p_i\}$; the counters $\{c_i\}$ are known to be sufficient statistics for the $\{p_i\}$. A-priori they should then suffice to compute an optimal policy.

2. Proof of Optimality

Assume the initial state of the list is random, with equal probability for each of the $n!$ orderings. The arrangement of the records after the mth request, also known as "at time m", is represented as

$$\sigma_m = \begin{pmatrix} 1 & 2 & \cdots n \\ \sigma_m(1) & \sigma_m(2) & \cdots \sigma_m(n) \end{pmatrix}$$

with $\sigma_m(i)$ = the position of R_i .

We define a history of the list at time m, under the policy H, as the vector

$$\nu_m = (I^{(m)}, S^{(m)}), \qquad I^{(m)} \equiv (i_1, \cdots, i_m), \quad S^{(m)} \equiv (\sigma_0, \cdots, \sigma_{m-1}),$$

where i_k denotes the record accessed at the kth request. $I^{(m)}$ is called the reference history vector (*rhv*). The use of the policy H is left tacit in the notations σ_m and ν_m.

The following notation is convenient in our proof method:

$$\bar{\sigma}_m = \begin{pmatrix} 1 & \cdots & n \\ \sigma_m(\sigma_0^{-1}(1)) & \cdots & \sigma_m(\sigma_0^{-1}(n)) \end{pmatrix}.$$

$\bar{\sigma}_m$ denotes the *canonical ordering* of the list after the mth request: given an initial state σ_0, each record is identified by its original position in the list. (We could formulate this as a transformation on the relevant *name space*). For any initial order, $\bar{\sigma}_0$ is the identity permutation, and $\bar{\sigma}_m$ describes the list-order at time m in terms of the initial position of the records.

We denote by $\bar{v}_m$ the *canonical history vector* :

$$\bar{v}_m = (\bar{I}^{(m)}, \bar{S}^{(m)}), \qquad \bar{I}^{(m)} \equiv (\bar{\imath}_1, \cdots, \bar{\imath}_m), \quad \bar{S}^{(m)} \equiv (\bar{\sigma}_0, \cdots, \bar{\sigma}_{m-1}) ,$$

where $\bar{\imath}_k = \sigma_0(i_k)$ and $\bar{\sigma}_k$ is the canonical ordering of the list after the kth reference, $1 \le k \le m$. The vector $\bar{I}^{(m)}$ will be naturally called the *crhv*. Denote by $f_H(v)$ the list resulting from using policy H with history v; then $\sigma_m = f_H(v_m)$, and $\bar{\sigma}_m = f_H(\bar{v}_m)$.

With some abuse of our notation we may also write

$$\sigma_m = f_H(\bar{v}_m \mid \sigma_0) ,$$

since evidently, σ_0 determines a one-to-one mapping from $\bar{\sigma}_m$ to σ_m.

We introduce now two classes of policies:

H is *key-ignoring* if for every pair of initial orderings σ_{01}, σ_{02} and any arrangement σ_{m1} such that

$$\sigma_{m1} = f_H(\bar{v}_m \mid \sigma_{01}) ,$$

let g be the permutation that carries σ_{01} to σ_{02} – that is, $\sigma_{01} \xrightarrow{g} \sigma_{02}$, then there exists an ordering σ_{m2}, such that

$$\sigma_{m1} \xrightarrow{g} \sigma_{m2} ,$$

and

$$\mathrm{Prob}_H(\sigma_{m1} \mid \sigma_{01}, \bar{v}_m) = \mathrm{Prob}_H(\sigma_{m2} \mid \sigma_{02}, \bar{v}_m) .$$

Considering the general case, in which H is not necessarily deterministic – the essence of the last requirement is that the sequence $\{\sigma_m\}$ has to be measurable with respect to the increasing σ-algebra generated by the sequence $\{\bar{v}_m\}$, which is key-ignorant.

Let H_{KI} be the set of all *key-ignoring* policies.

We use the notation H_D for the class of *deterministic* permutation rules, such that under an initial ordering σ_0 and a given reference history v_m, the outcome σ_m is defined by H uniquely.

Let $H_{DK} = H_D \cap H_{KI}$. The next Lemma shows that we may restrict our attention to H_{DK}:

Lemma 1: Within the class of H_{KI}, there exists a policy $H \in H_{DK}$ which minimizes the average access cost at time m, $m \ge 1$.

We leave out the proof; it uses induction on m to show that any non-deterministic rule in H_{KI} cannot do better than the best strategy in H_{DK}.

Consider two initial orderings σ_{01}, σ_{02} which differ only by the interchange of two records R_i, R_j:

$$\begin{aligned} &\sigma_{01}(i) = k , \quad \sigma_{01}(j) = l \qquad k < l \\ &\sigma_{02}(i) = l , \quad \sigma_{02}(j) = k \qquad \sigma_{01}(s) = \sigma_{02}(s) \qquad 1 \le s \le n, \quad s \ne i, j. \end{aligned} \tag{1-0}$$

Two observations about this notation, formulated as lemmata, provide the tools for the main result.

Lemma 2: For all $H \in H_{DK}$, σ_{0i} as specified in equation (1-0), a canonical history vector $\bar{v}_m$ and the final states

$$\sigma_{m1} = f_H(\bar{v}_m \mid \sigma_{01}) \quad \text{and} \quad \sigma_{m2} = f_H(\bar{v}_m \mid \sigma_{02}),$$

we find

$$\sigma_{m1}(i) = \sigma_{m2}(j), \quad \sigma_{m1}(j) = \sigma_{m2}(i); \quad \sigma_{m1}(s) = \sigma_{m2}(s), \quad \forall\, s \neq i, j. \tag{1-m}$$

The proof is immediate: since $H \in H_{DK}$, it effects for a specified sequence of references, expressed in terms of initial position, the same deterministic permutation g, (regardless of the "actual" labels of the records), which preserves the relation (1-0).

Remark 1: Another phrasing of Lemma 2 is that for policies in H_{DK}, the vector $S^{(m)}$ is determined uniquely in terms of the initial ordering and $I^{(m)}$. Moreover, $\bar{S}^{(m)}$ is determined uniquely in terms of $\bar{I}^{(m)}$ alone.

Remark 2: Clearly, when $p_i \neq p_j$, the two histories induced by $\bar{v}_m$ and σ_{01}, or by $\bar{v}_m$ and σ_{02}, need not (will not) have equal probabilities.

Let $\bar{C}^{(m)} = (\bar{c}_1^{(m)}, \cdots, \bar{c}_n^{(m)})$ be the frequency vector accumulated after a sequence of m references, where $\bar{c}_i^{(m)}$ is the counter of the record positioned ith in the initial order.

Lemma 3: For any frequency vector $\bar{C}^{(m)}$ and $\sim H \in H_{DK}$,

$$\text{Prob}_H(\sigma_{m1}(i) < \sigma_{m1}(j) \mid \bar{C}^{(m)}, \sigma_{01}) = \text{Prob}_H(\sigma_{m2}(j) < \sigma_{m2}(i) \mid \bar{C}^{(m)}, \sigma_{02}).$$

Proof: For H as given, $\bar{C}^{(m)}$ determines $\bar{I}^{(m)}$ up to the order of the references. There are $\binom{m}{\bar{C}^{(m)}}$ (a multinomial coefficient) such arrangements, providing as many canonical history vectors $\bar{v}_m$ that fit the frequency vector $\bar{C}^{(m)}$, for any initial permutation of the records. Under the *irm* they are all equi-probable, and since H is deterministic, we find that the probability of each $\bar{v}_m$ which fits $\bar{C}^{(m)}$, for a fixed initial order, is given by

$$\text{Prob}_H(\bar{v}_m \mid \bar{C}^{(m)}) = \binom{m}{\bar{C}^{(m)}}^{-1}.$$

From relation (1-m) it follows that for any $\bar{v}_m$,

$$\sigma_{m1}(i) < \sigma_{m1}(j) \iff \sigma_{m2}(j) < \sigma_{m2}(i).$$

Let $|A|$ denote the cardinality of the set A. Then,

$$\mathrm{Prob}_H(\sigma_{m1}(i) < \sigma_{m1}(j) \mid \overline{C}^{(m)}, \sigma_{01}) = \frac{|\overline{v}_m : \sigma_{m1}(i) <_H \sigma_{m1}(j),\ \overline{v}_m \text{ fits } \overline{C}^{(m)}|}{\binom{m}{\overline{C}^{(m)}}}$$

$$= \frac{|\overline{v}_m : \sigma_{m2}(j) <_H \sigma_{m2}(i),\ \overline{v}_m \text{ fits } \overline{C}^{(m)}|}{\binom{m}{\overline{C}^{(m)}}}$$

$$= \mathrm{Prob}_H(\sigma_{m2}(j) < \sigma_{m2}(i) \mid \overline{C}^{(m)}, \sigma_{02}). \qquad \square$$

Let $E_{PR}^{(m)}(C)$ and $E_{PR}(C)$ denote the expected access cost to the list after the mth request and in the limiting state respectively, under the permutation rule PR. Our main result is

Theorem 4: If $H \in H_{DK}$, then

$$E_{CS}^{(m)}(C) \le E_H^{(m)}(C)$$

for all $m \ge 1$.

Proof: From the explicit expression for the expected access cost, $E^{(m)}(C) = \Sigma_{i=1}^{n} p_i\, \sigma_m(i)$, it follows that it can be split into a sum over the relative positions of pairs of records. Hence it would be sufficient to show that for any frequency vector $\overline{C}^{(m)}$, an arbitrary policy $H \in H_{DK}$, and every pair of records R_i, R_j $1 \le i, j \le n$, with the respective access probabilities p_i, p_j, the following implication holds:

$$p_i > p_j \implies \mathrm{Prob}_{CS}(\sigma_m(i) < \sigma_m(j) \mid \overline{C}^{(m)}) \ge \mathrm{Prob}_H(\sigma_m(i) < \sigma_m(j) \mid \overline{C}^{(m)}),$$

where the two probabilities on the right-hand side are with respect to the initial permutations and the crhvs that are compatible with $\overline{C}^{(m)}$. Clearly, if the reference probabilities are equal, the order of the records in the list does not matter. Also, any possible depndence on $\overline{I}^{(m)}$ is restricted to H, since under CS the outcome σ_m is determined uniquely by cm (or $\overline{C}^{(m)}$ and the initial order of the records). The $n!$ permutations are split into two halves, differing just as the paradigmatic initial orderings σ_{01} and σ_{02} do, with respect to the records placed in locations k and l.

Consider such a particular pair R_i and R_j, and assume with no loss in generality that $p_i > p_j$. In the vector $\overline{C}^{(m)}$ we suppress the superscripts temporarily, that is, $\overline{C}^{(m)} = (\overline{c}_1^{(m)}, \cdots, \overline{c}_k^{(m)}, \cdots, \overline{c}_l^{(m)}, \cdots, \overline{c}_n^{(m)}) \equiv (c_1, \cdots, c_k, \cdots, c_l, \cdots, c_n)$.

Consider first the particular case of equality of the two counters c_k and c_l. Little reflection shows that for *either* the CS or any other $H \in H_{DK}$, with any suitable $I^{(m)}$, half the σ_0 will result in $\sigma_m(i) < \sigma_m(j)$, and the other half with the reverse order.

Without loss of generality, we assume now that $c_k > c_l$, and then

$$\mathrm{Prob}_{CS}(\sigma_m(i) < \sigma_m(j),\ \overline{C}^{(m)}, \sigma_{01}) + \mathrm{Prob}_{CS}(\sigma_m(i) < \sigma_m(j),\ \overline{C}^{(m)}, \sigma_{02}) =$$

$$= \frac{1}{n!} \binom{m}{\overline{C}^{(m)}} p_{i_1}^{c_1} \cdots p_i^{c_k} \cdots p_j^{c_l} \cdots p_{i_n}^{c_n} \equiv A\, p_i^{c_k}\, p_j^{c_l},$$

where one term is selected. Then, if $H \in H_{DK}$,

$$\text{Prob}_H(\sigma_m(i) < \sigma_m(j),\ \overline{C}^{(m)}, \sigma_{01}) + \text{Prob}_H(\sigma_m(i) < \sigma_m(j),\ \overline{C}^{(m)}, \sigma_{02}) =$$

$$= A\,(p_i^{c_k}\, p_j^{c_l} \cdot x + p_j^{c_k}\, p_i^{c_l} \cdot (1-x))\ ,$$

where

$$x \equiv \text{Prob}_H(\sigma_m(i) < \sigma_m(j) \mid \overline{C}^{(m)}, \sigma_{01}) = \text{Prob}_H(\sigma_m(j) < \sigma_m(i) \mid \overline{C}^{(m)}, \sigma_{02}),$$

with the last equality provided by Lemma 3. Now, the combination $p_i > p_j$ and $c_k > c_l$ implies

$$p_i^{c_k}\, p_j^{c_l} = p_i^{c_k}\, p_j^{c_l}\, x + p_i^{c_k}\, p_j^{c_l}\,(1-x) > p_i^{c_k}\, p_j^{c_l}\, x + p_i^{c_l}\, p_j^{c_k}\,(1-x).$$

Hence, summing over such $n!/2$ pairs of initial orders

$$\text{Prob}_{CS}(\sigma_m(i) < \sigma_m(j),\ \overline{C}^{(m)}) = \sum_{\{\sigma_{01},\sigma_{02}\}} \text{Prob}_{CS}(\sigma_m(i) < \sigma_m(j),\ \overline{C}^{(m)}, \sigma_{01})$$

$$+ \text{Prob}_{CS}(\sigma_m(i) < \sigma_m(j),\ \overline{C}^{(m)}, \sigma_{02}) \geq \text{Prob}_H(\sigma_m(i) < \sigma_m(j),\ \overline{C}^{(m)})\ .$$

The last inequality holds for any $\overline{C}^{(m)}$, $m \geq 1$ and any pair of indices i, j, such that $p_i > p_j$. Converting the joint probabilities to the required conditional ones is immediate, since the *irm* assumes independence of the state of the list and subsequent references. The inequality in the theorem is then achieved by summing on all record pairs and frequency vectors. □

3. Further Remarks

We have shown that CS is optimal within the class of H_{DK}. One may be easily convinced, by adversary-type arguments, that any policy which is not *key ignoring*, would not be optimal under the *irm* assumptions.

The CS is the best reorganization method not only in the limiting sense, but for any finite sequence of requests. It also provides an indirect proof to the superiority of CS—when the *irm* assumption holds—over some of the well-known permutations rules, which have not yet been analyzed with respect to our measures (Transpose belongs to that category).

To avoid the allocation of huge counter fields, CS may be replaced by the *Limited Counters Scheme* (Hofri and Shachnai, 1988). This 'truncated' version of CS reduces significantly its storage requirements while still being very cost-effective. It would be of interest to examine the classes of policies which can still do better than the various versions of LCS.

We comment, that the optimality of CS holds under the following assumptions on the model :

(*i*) The set of records in the list remains fixed over time.

(*ii*) No initial information on the *rpv*.

(*iii*) Independent and time-homogeneous reference probabilities.

Permitting insertions and deletions, or having some initial knowledge of any subset of the access probabilities may lead to new conclusions concerning the existence of an optimal policy and its thusimplied characteristics.

Relaxing the independence assumption has not been considered in previous work. We believe, that for certain models of dependent references, the optimality of CS still holds, albeit with a different character. This is certainly the case when the components of p are time varying, but retain their ranking time-invariant. For a different one, assume a reference model which follows a first-order Markov chain, i.e. p_{ij} is the conditional probability of accessing R_j after a reference to R_i, $1 \le i, j \le n$. If none of those transition probabilities is known in advance, and the same cost structure holds (where key-comparisons carry a price tag but record shuffles do not), consider the following reorganization scheme :

Each of the records is associated with a frequency vector c_i, where c_{ij} counts the number of accesses to R_j immediately following a request to R_i. Then a reference to R_i (preceded by a search for R_k) would result with an increment of the appropriate counter (c_{ki}) and a new permutation of the list – in descending order of the counters c_{ij}, $1 \le j \le n$.

By the Law of Large Numbers, this rule is asymptotically optimal for the above access model. We expect it should be also the best policy for any finite sequence of requests, but we have produced no formal proof of that. However, if we charge both for comparisons and shuffles, there is little hope for an optimal policy with such a simple structure.

We conclude by pointing out, that *counter based* methods are not optimal with respect to our measure when memory of the past is of limited span.

This can be demonstrated on a model in which the relative order of the records after the mth request is determined by the reference history accumulated since the l+1st request, $1 \le l < m$.

Let $C^{(m-l)}$ be the partial frequency vector representing the last $m-l$ requests. Obviously, keeping the list in descending order of the counters in $C^{(m-l)}$ would not always minimize the expected access cost at the m+1st reference, as that would imply, for $l=m-1$, that

$$E_{MTF}^{(m)}(C) \le E_{TR}^{(m)}(C) \quad \forall m \ge 1 \ .$$

But the last inequality contradicts Rivest's proof (Rivest, 1976) that

$$E_{MTF}(C) > E_{TR}(C)$$

for all non-trivial *rpv*'s.

REFERENCES

Hofri M. and Shachnai H., Self-Organizing Lists and Independent References - a Statistical Synergy. The Department of Computer Science, the Technion TR#524, October 1988.

Lam, K., Siu, M.K., and Yu, C.T., A generalized counter scheme. *Theor. Comput. Sci.* **16**, #3, 271-278, 1981.

Rivest, R., On self-organizing sequential search heuristics. *Commun. ACM*, **19**, #9, 63-67, 1976. 9, 1976