Jointly Optimal Paging and Registration for a Symmetric Random Walk

Bruce Hajek[1]
Department of Electrical and Computer Engineering
and the Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
e-mail: b-hajek@uiuc.edu

Abstract — **Jointly optimal paging and registration policies are identified for a cellular network composed of a linear array of cells. Motion is modeled as a random walk with a symmetric, unimodal step size distribution. Minimization of the discounted, infinite-horizon average cost is addressed. The jointly optimal pair of paging and registration policies is found. The optimal registration policy is a distance threshold type: the mobile station registers whenever its distance from the previous reporting point exceeds a threshold. The paging policy is ping-pong type: cells are searched in an order of increasing distance from the cell in which the previous report occurred.**

I. INTRODUCTION

The structure of jointly optimal paging and registration policies is investigated in [1], for motion of a mobile station (MS) modeled by a Markov process. The reader is referred to that paper for background information. The model presented in this paper is a special case. The motion of an MS is modeled by a discrete-time random walk $(X(t) : t \geq 0)$ on the integers Z, such that the displacement of the walk at each step has distribution b. Equivalently, $(X(t) : t \geq 0)$ is a discrete time Markov process on Z with one-step transition probability matrix P given by $p_{ij} = b_{j-i}$. For any probability vector w, $wP = w * b$. It is assumed that b_i is a nonincreasing function of $|i|$, or in other words, b is symmetric about zero and unimodal. In [1] multiple states can correspond to the same cell, corresponding to different velocities for example. In this paper, however, each integer state i corresponds to a distinct cell or paging area in which the MS can be paged. It is assumed that the network knows the initial state x_0.

The state of the MS from one time t to the next evolves according to the given transition probabilities. The order of the possible events at a particular integer time instant $t \geq 1$ are as follows. First, it is announced whether the MS is to be paged, and the answer is "yes" with probability λ_p, independently of the state of the MS and all past events. The cost of the paging at time t is $\mathcal{P}N_t$, where $\mathcal{P}$ is the cost of searching one state and N_t is the number of states that are searched until the MS is found. As a result of being paged, the state of the MS is reported to the cellular network.

If the MS is not paged, then the MS decides whether to register. The cost of registration is $\mathcal{R}$ and the benefit of registration is that the cellular network learns the state of the MS. We say that a *report* occurs whenever either a paging or registration occurs, because in either case, the cellular network learns the state of the MS.

Let P_t denote the event that the MS is paged at time t, and let R_t denote the event that the MS registers at time t. For any set A, let I_A denote the indicator function of A, which is one on A and zero on the complement A^c. Probability vectors are considered to be row vectors.

A. *Paging policy notation*

For simplicity we consider only serial paging policies, so that states are searched one at a time until the MS is located. Let $\mathcal{N}_t$ denote the σ algebra representing the information available to the network by time t after the paging and registration decisions have been made and carried out. Thus, for $t \geq 0$,

$$\begin{aligned} \mathcal{N}_t &= \sigma((I_{P_s}, N_s, I_{R_s} : 1 \leq s \leq t), \\ &\quad (X(s) : 1 \leq s \leq t \text{ and } I_{P_s \cup R_s} = 1)) \end{aligned}$$

The initial state x_0 is treated as a constant, so even though it is known to the network it is not included in the definition of $\mathcal{N}_t$.

A paging policy u is a collection $u = (u(t) : t \geq 1)$ such that for each $t \geq 1$, $u(t)$ is an $\mathcal{N}_{t-1}$ measurable random variable with values in the set of paging orders for the set of states Z. Given a paging order a, the *paging rank vector*, $r^a = (r^a_l : l \in Z)$ is defined as follows. If states are searched sequentially in the order specified by a, then r^a_l is the number of single-state searches required to find the MS if it is in state l.

B. *Registration policy notation*

Let $\mathcal{M}_t$ denote the σ algebra representing the information available to the MS by time t, after the paging and registration decisions for time t have been made and

[1]This work was supported by the National Science Foundation under Grant NSF ITR 00-85929.

0-7803-7629-3/02/$10.00 © 2002 IEEE

carried out. Thus,

$$\mathcal{M}_t = \sigma(X(s), I_{P_s}, N_s, I_{R_s} : 1 \le s \le t).$$

The MS also knows the initial position x_0, which is treated as a constant. We assume without loss of optimality that the MS does not register at time t if it is paged at time t. A registration policy v is a collection $v = (v(t) : t \ge 1)$ such that for each $t \ge 1$, $v(t)$ is an $\mathcal{M}_{t-1}$ measurable random vector with values in $[0,1]^S$ with the following interpretation. Given the information $\mathcal{M}_{t-1}$, if $X(t) = l$ and if the MS is not paged at time t, then the MS registers with probability $v_l(t)$.

C. Cost function

Let β be a number with $0 < \beta < 1$, called the discount factor. Given a paging policy u and registration policy v, the expected infinite horizon discounted cost $C(u,v)$ is defined as

$$C(u,v) = E\left[\sum_{t=1}^{\infty} \beta^t \{\mathcal{P} I_{P_t} N_t + \mathcal{R} I_{R_t}\}\right]. \quad (1)$$

The pair (u,v) is *jointly optimal* if $C(u,v) \le C(u',v')$ for all other paging policies u' and registration policies v'.

D. Stationary Reduced Complexity Laws

As shown in [1], jointly optimal control policies can be expressed as reduced complexity laws (RCLs). A paging RCL has the form $f = (f(i_0,k) : i_0 \in Z, k \ge 0)$ such that $f(i_0,k)$ is an ordering of the states Z with the following meaning. If the MS just enters state l at time t, if k is the amount of time elapsed since the last report, if i_0 is the last reported state, and if the mobile is paged at time t, then the paging order is given by $u(t) = f(i_0,k)$. A registration RCL has the form $g = (g(i_0,k) : i_0 \in Z, k \ge 0)$ such that $g(i_0,k) = (g_l(i_0,k) : l \in Z)$ has the following meaning. If the MS just enters state l at time t, if the last report occurred k time units earlier in state i_0, and if the MS is not paged at time t, then the MS registers at time t with probability $g_l(i_0,k)$. Equivalently, $v(t) = g(i_0,k)$. There is no loss in optimality [1] to require $g_l(i_0,k) \in \{0,1\}$.

Due to the translation invariance of P for the example of this paper, both the paging and registration RCLs can be taken to be translation invariant. Full justification can be provided by the same dynamic programming argument used in [1] to show that jointly optimal policies can be expressed as RCLs. Thus, we write the RCLs as $f = (f(k) : k \ge 1)$ and $g = (g(k) : k \ge 1)$. These RCLs give the control decisions if the last reported state is $i_0 = 0$, and hence for other values of i_0 by translation in space.

E. Optimal Policies

More specific RCLs will be specified in this section. These policies do not depend on the time k elapsed since last report, so the argument k is suppressed. For the paging policy f^* we take the ping-pong policy, given by $f^* = (0,1,-1,2,-2,3,-3,\ldots)$. Thus, if the MS is to be paged and if it was last reported to be at state i_0, then the states are searched in the order $i_0, i_0+1, i_0-1, i_0+2, i_0-2, \ldots$. The registration policy g^* is given by $g_l^* = I_{\{l \ge d_r \text{ or } l \le -d_l\}}$ where the two distance thresholds $d_l, d_r \ge 1$ are such that either $d_l = d_r$ or $d_l = d_r - 1$. The following proposition is the main result of this paper.

Proposition I.1 *There is a choice of the distance thresholds d_l and d_r such that the ping-pong paging policy and the distance-threshold registration policy given by g^* are jointly optimal.*

The related work of Madhow, Honig and Steiglitz [2] finds the optimal registration policy assuming that the paging policy is fixed to be the ping-pong policy. Also, it is not difficult to show that for the distance threshold registration policy g^*, the optimal paging policy is the ping-pong paging policy. However, even if a pair of RCLs is individually optimal, meaning f is optimal for g fixed, and g is optimal for f fixed, a simple example in [1] shows that (f,g) is not necessarily jointly optimal. Moreover, the main task in the proof of Proposition I.1 seems to be the proof that the ping-pong paging policy can be used without loss of optimality (if the registration policy is suitably adjusted).

The remainder of this summary will give the main points of the proof of the proposition.

II. Preliminaries for the Proof

Given a probability distribution μ on a discrete set $\mathcal{S}$, the *ordered probability distribution corresponding to* μ, denoted $\tilde{\mu}$, is the probability distribution on $\{1,2,\ldots\}$ defined as follows. If the numerical values $(\mu_i : i \in \mathcal{S})$ are arranged in nonincreasing order (with multiplicity), then $\tilde{\mu}_j$ is the j^{th} number in the sequence. For example, if μ assigns probabilities $0.3, 0.3$, and 0.4 to three particular points, then $\tilde{\mu} = (0.4, 0.3, 0.3, 0, 0, \ldots)$. Given two such probability distributions μ and ν (possibly on different discrete sets) distribution μ is *more concentrated* than ν, written $\mu \prec \nu$, if for any $r \ge 1$,

$$\sum_{i=1}^{r} \tilde{\mu}_i \ge \sum_{i=1}^{r} \tilde{\nu}_i.$$

Write $\mu \equiv \nu$ to denote that both $\mu \prec \nu$ and $\nu \prec \mu$. For any probability distribution, $\mu \equiv \tilde{\mu}$.

Let $s(\mu)$ denote the mean number of states that must be searched to find the MS given that the MS has distribution μ and the optimal search order for μ is used. The optimal search order is maximum likelihood search [3], under which states are searched in order of decreasing probability. Summation by parts yields

$$s(\mu) = \sum_{i=1}^{\infty} i\tilde{\mu}_i = 1 + \sum_{r=1}^{\infty}(1 - \sum_{i=1}^{r} \tilde{\mu}_i),$$

which immediately implies the following lemma.

Lemma II.1 *If $\mu \prec \nu$ then $s(\mu) \le s(\nu)$.*

A function or probability distribution μ on Z is said to be *neat* if $\mu_0 \geq \mu_1 \geq \mu_{-1} \geq \mu_2 \geq \mu_{-2} \geq \ldots$. Throughout this section, b is a symmetric, unimodal probability distribution on Z.

Lemma II.2 *If μ is a neat probability distribution, then the convolution $\mu * b$ is neat.*

Proof. For $i \geq 0$ let $b^{(i)}$ denote the uniform probability distribution over the interval of integers $[-i, i]$. The conclusion is easy to verify in case b has the form $b^{(i)}$ for some i. In general, b is a convex combination of such $b^{(i)}$'s. Convex combinations of neat distributions are neat, so $\mu * b$ is indeed neat. ■

The next three lemmas lead up to the fourth lemma below, which is the key lemma in the proof of Proposition I.1. The notion of $\tilde{\mu}$ and the ordering $(\prec)$ extend naturally to summable nonnegative functions on discrete sets, though two such functions will be compared in the $\prec$ order only if their total sum is the same. We use such extension for notational convenience, rather than normalizing the appropriate functions to be probability distributions.

Lemma II.3 *Consider two monotone sequences of some finite length n: $a_1 \geq a_2 \geq \ldots \geq a_n = 0$ and $0 = b_1 \leq b_2 \leq \ldots \leq b_n$. Let $c_i = a_i + b_i$ for all i and let $d_i = a_i + b_{i+1}$ for $0 \leq i \leq n-1$ and $d_n = 0$. Then $d \prec c$.*

Proof. Note that $d_i \geq c_i$ and $d_i \geq c_{i+1}$ for each $1 \leq i \leq n-1$, and of course the sum all the c_i is equal to the sum of all the d_i. Therefore, for any subset A of $\{1, 2, \ldots, n\}$, there is another subset A' with $|A| = |A'|$ and $\sum_{i \in A} c_i \leq \sum_{i \in A'} d_i$. That proves the lemma. ■

Lemma II.4 *Let r and L be positive integers. Consider the convolution $F * G$ of two binary valued functions on Z, such that the support of F has cardinality r, and the support of G is a set of L consecutive integers. Then the convolution is minimal (most concentrated) in the $(\prec)$ order if the support of F is a set of r consecutive integers.*

Proof. Suppose without loss of generality that $G = I_{\{0 \leq i \leq L-1\}}$. If the support of F is not an interval of integers, let j_{max} be the largest integer in the support of F and let j_0 be the smallest integer such that the support of F contains the interval of integers $[j_o, j_{\max}]$. Then $F = F^a + F^b$, such that $F_i^a = 0$ for $i \geq j_0 - 1$ and the support of F^b is the interval of integers $[j_0, j_{\max}]$. Let F' be the new function defined by $F'_i = F_i^a + F_{i+1}^b$. The graph of F' is obtained by sliding the rightmost portion of the graph of F to the left one unit.

We claim that $F' * G \prec F * G$. To see this, note that $F * G = F^a * G + F^b * G$. The idea of the proof is to focus on the interval of integers $I = [j_0 - 1, j_0 + r - 2]$ and appeal to Lemma II.3. The function $F^a * G$ is nonincreasing on I, it takes value zero at the right endpoint of I, and it is also zero everywhere to the right of I. The function $F^b * G$ is nondecreasing on I, it takes value zero at the left endpoint of I, and it is also zero everywhere to the left of I. The convolution $F' * G$ is the same as $F * G$ except the second function $F^b * G$ is shifted one unit to the right. Lemma II.3 thus implies that $F' * G \prec F * G$. This procedure can be repeated until F is reduced to a function with support being a set of r consecutive integers. The lemma is proved. ■

Lemma II.5 *Let $r \geq 1$ and consider the convolution $F * b$ such that F is a binary valued function on the integers with support of cardinality r. Then the convolution is minimal (most concentrated) in the $(\prec)$ order if the support of F consists of r consecutive integers.*

Proof. For $i \geq 0$ let $b^{(i)}$ denote the uniform probability distribution on the interval of $L = 2i + 1$ integers $[-i, i]$. The lemma is true if $b = b^{(i)}$ for some i by Lemma II.4. Furthermore, if $b = b^{(i)}$ is convolved with the neat binary-valued function F^* with support of cardinality r, then the result $F^* * b^{(i)}$ is neat. In general, b can be written as a convex combination of such probability distributions $b^{(i)}$. Therefore, for any binary F with support of cardinality r:

$$\begin{aligned} b * F &= \sum_i \lambda_i b^{(i)} * F \overset{(a)}{\succ} \sum_i \lambda_i \widetilde{b^{(i)} * F} \\ &\succ \sum_i \lambda_i \widetilde{b^{(i)} * F^*} \overset{(b)}{\equiv} b * F^*. \end{aligned}$$

Here (a) follows from the fact that ordering probability distributions before adding them increases the concentration of the sum, and (b) follows from the fact that since all the distributions $b^{(i)} * F^*$ are neat, their sum is equally concentrated if they are ordered before being added. ■

Lemma II.6 *If μ and ν are probability distributions such that $\mu \prec \nu$ and μ is neat, then $\mu * b \prec \nu * b$.*

Proof. Fix $r \geq 1$, let F range over all binary valued functions on Z with support of cardinality r, and let F^* denote the unique choice of F that is neat. Use "<>" to denote inner products.

$$\begin{aligned} \sum_{i=1}^{r} \widetilde{\nu * b_i} &= \max_F < \nu * b, F > \; = \; \max_F < \nu, b * F > \\ &\overset{(a)}{\leq} \max_F < \tilde{\nu}, \widetilde{b * F} > \overset{(b)}{=} < \tilde{\nu}, \widetilde{b * F^*} > \\ &\leq < \tilde{\mu}, \widetilde{b * F^*} > \overset{(c)}{=} < \mu, b * F^* > \\ &= < \mu * b, F^* > \overset{(d)}{=} \sum_{i=1}^{r} \widetilde{\mu * b_i} \end{aligned}$$

Above (a) follows since ordering two distributions increases their inner product, (b) follows from Lemma II.5 and the monotonicity of $\tilde{\nu}$, (c) follows from the fact that both μ and $b * F^*$ are neat, so their innner product is the same as the inner product of their ordered probability distributions, and (d) follows from the fact that $\mu * F^*$ is neat. ∎

Let μ be a probability measure on Z and let $0 \le \lambda \le 1$. Let $\mathcal{T}(\mu, \nu)$ be the set of probability measures ν such that $(1-\lambda)\nu \le \mu$, pointwise. Intuitively, such a measure ν is obtained from μ by trimming away from μ probability mass λ and re-normalizing the remaining mass. The following lemma, which is easy to prove, means that given μ and λ, in order that ν be as concentrated as possible, the mass of μ should be trimmed from the states with the smallest probability under μ.

Lemma II.7 *(Minimum likelihood trimming maximizes concentration) There exists $\nu \in \mathcal{T}(\mu, \lambda)$ such that for some k,*

$$(1-\lambda)\tilde{\nu}_j = \begin{cases} \tilde{\mu}_j & \text{if } j < k \\ 0 & \text{if } j > k. \end{cases}$$

Furthermore, for any other $\nu' \in \mathcal{T}(\mu, \lambda)$, $\nu \prec \nu'$.

III. Proof of Proposition I.1

Let f and g be RCLs (possibly dependent on the elapsed time k since last report). The cost $C(f,g)$ can be computed by considering the process only up until the first time τ that a report occurs (i.e. one reporting cycle). Let $\alpha(k) = P[\tau = k] = \alpha_p(k) + \alpha_r(k)$, where $\alpha_p(k)$ is the probability $\tau = k$ and the first report is a page, and $\alpha_r(k)$ is the probability $\tau = k$ and the first report is a registration. Also let $w(k)$ denote the conditional distribution of the MS given that no report occurs up to time k for the pair of policies (f,g). Then

$$\begin{aligned} C(f,g) &= \frac{E\left[\sum_{t=1}^{\tau} \beta^t \{\mathcal{P} I_{P_t} N_t + \mathcal{R} I_{R_t}\}\right]}{1 - E[\beta^\tau]} \\ &= \frac{\sum_{k=1}^{\infty} \beta^k \{\mathcal{P}\alpha_p(k) s(w(k-1) * b) + \mathcal{R}\alpha_r(k)\}}{1 - \sum_{k=1}^{\infty} \beta^k \alpha(k)} \end{aligned}$$

Note that the cost depends entirely on the α's and on the mean numbers of pages required, given by the terms $s(w(k-1) * b)$.

Lemma III.1 *(Optimality of ping-pong paging f^*) There exists a registration RCL g^o so that $C(f^*, g^o) \le C(f,g)$.*

Proof. Take the registration RCL g^o to be a distance threshold policy with time varying thresholds and possibly with randomization at the left threshold if the thresholds are equal, or at the right threshold if the right threshold is one larger than the left threshold. More precisely, for fixed k: all the values $g^o_l(k)$ are binary except for possibly one value of l, and $1 - g^o(k)$ is neat. The thresholds and randomization parameter are selected so that the α's, α_p's, and α_r's are the same for the pair (f^*, g^o) as for the originally given pair (f,g). Let $w^o(k)$ denote the conditional state distributions under policy (f^*, g^o). To complete the proof of the lemma it remains to show that $s(w^o(k-1) * b) \le s(w(k-1) * b)$ for $k \ge 1$. For that purpose apply Lemmas II.2, II.6, and II.7 to show by induction that for all $k \ge 1$: $w^o(k)$ is neat, $w^o(k) \prec w(k)$, and $w^o(k-1) * b \prec w(k-1) * b$. Thus by Lemma II.1, $s(w^o(k-1) * b) \le s(w(k-1) * b)$, completing the proof of Lemma III.1. ∎

The proof of Proposition I.1 is now completed. In view of Lemma III.1, it remains to show that if the ping-pong paging policy f^* is used, then for some choice of fixed distance thresholds d_l and d_r the registration policy g^* is optimal. This can be done be examining a dynamic program for the optimal registration policy, under the assumption that the ping-pong paging policy f^* is used. Let $V_n(j)$ denote the mean discounted cost for n time steps to go, given that the mobile is located directed distance j from its last reported state. Then

$$\begin{aligned} V_{n+1}(j) &= \beta \sum_{l \in Z} p_{jl} [\lambda_p (\mathcal{P} \tau_l^{f^*} + V_n(0)) \\ &\quad + (1-\lambda_p) \min\{V_n(l), \mathcal{R} + V_n(0)\}] \end{aligned}$$

By a contraction property of these dynamic programming equations, the limit $V_* = \lim_{n\to\infty} V_n$ exists. Argument by induction yields that the functions $-V_n$ are neat, and hence that $-V_*$ is neat. By the dynamic programming principle, an optimal registration policy g^* is given by

$$g_l^* = \begin{cases} 1 & \text{if } V_*(l) \ge \mathcal{R} + V_*(0) \\ 0 & \text{else} \end{cases}$$

As $-V_*$ is neat, the optimal registration policy g^* has the required threshold type. Proposition I.1 is proved.

IV. Possible extensions

A result similar to Proposition I.1 is probably true for Brownian motion in one or more dimensions, though the case of two or more dimensions could be difficult.

References

[1] B. Hajek, K. Mitzel, and S. Yang, "Paging and Registration in Cellular Networks: Jointly Optimal Policies and an Iterative Algorithm," Preprint available at comm.csl.uiuc.edu/~hajek

[2] U. Madhow, M. L. Honig, and K. Steiglitz, "Optimization of wireless resources for personal communications mobility tracking," *IEEE/ACM Transactions on Networking*, vol. 3, pp. 698–707, December 1995.

[3] C. Rose and R. Yates, "Minimizing the average cost of paging under delay constraints," *Wireless Networks*, vol. 1, pp. 211–219, 1995.