COMBINATORICA
Akadémiai Kiadó – Springer-Verlag

COMBINATORICA 16 (1) (1996) 1–18

BIASED RANDOM WALKS

YOSSI AZAR, ANDREI Z. BRODER, ANNA R. KARLIN,
NATHAN LINIAL† and STEVEN PHILLIPS‡

Received December 22, 1992
Revised December 9, 1994

How much can an imperfect source of randomness affect an algorithm? We examine several simple questions of this type concerning the long-term behavior of a random walk on a finite graph. In our setup, at each step of the random walk a "controller" can, with a certain small probability, fix the next step, thus introducing a bias. We analyze the extent to which the bias can affect the limit behavior of the walk. The controller is assumed to associate a real, nonnegative, "benefit" with each state, and to strive to maximize the long-term expected benefit. We derive tight bounds on the maximum of this objective function over all controller's strategies, and present polynomial time algorithms for computing the optimal controller strategy.

1. Introduction

Ever since the introduction of randomness into computing, people have been studying how imperfections in the sources of randomness affect the outcome of the computation. A number of authors [17, 19, 4, 14] consider the same problem from the source's viewpoint. There are many ways for the source to deviate from perfect randomness, but for most models the source is an adversary who can select a strategy from a given repertoire, and whose goal is to derail the randomized algorithm. While the randomized algorithm is searching for a witness, the source tries to fail this search. Lichtenstein *et al.* [14] point out the control-theoretic flavor of this problem, namely, we are studying a stochastic process (the randomized algorithm) on which an agent (source, controller, adversary...) can exert some bias. The strategies available to the controller are given by some rules, and there are certain goals that he strives to achieve. Some qualitative information about the

Mathematics Subject Classification (1991): 60 J 15, 90 C 40, 68 R 10

† This research was supported in part by the Alon Fellowship and the Israel Science Foundation administered by the Israel Academy of Sciences.This work was supported in part by a grant from the Israeli Academy of Sciences. A portion of this work was done while the author was visiting DEC Systems Research Center.

‡ A portion of this work was done while the author was in the Computer Science Department at Stanford and while visiting DEC Systems Research Center. Partially supported by NSF Grant CCR-9010517 and an OTL grant.

optimal strategy can be obtained using Markov decision theory, but we are mainly interested in quantitative questions: how much can the controller's bias affect the stochastic process, and which stochastic processes are the hardest to influence? The only instances of this problem for which answers are available [17, 19, 4, 14, 6, 2, 3, 10] are (either originally thus stated, or easily translated into these terms) random walks on a finite tree, which is directed from the root to the leaves. The decision which branch to take is determined locally and the controller can bias this decision with a certain probability. The present paper concerns similar questions pertaining to random walks on finite graphs, which, in contrast, are time-infinite. This instance of the general problem is particularly appealing in the context of weakly-random sources, since witness-searching through random walks (mostly on expander graphs) has proved highly successful [1, 8, 12].

An instance of our problem is specified by an n-vertex connected graph G and a fixed $1 > \varepsilon > 0$. We will be concerned mostly with regular graphs, and use d to denote their degree. We consider the following variation of the standard random walk on the vertices of G: each step of the walk is preceded by an $(\varepsilon, 1-\varepsilon)$-coin being flipped. With probability $1-\varepsilon$ one of the d neighbors is selected uniformly at random, and the walk moves there. With probability ε the controller gets to select which neighbor to move to. The selection can be probabilistic, but it is time independent. In other words if the original transition probability matrix of the random walk was Q, then the modified transition probability matrix is

$$P = (1-\varepsilon)Q + \varepsilon B,$$

where B is an arbitrary stochastic matrix chosen by the controller, with support restricted to the edges of G. The interesting situation is when ε is not substantially larger than $1/d$; otherwise, the process is dominated by the controller's strategy.

This setup is a special case of a Markov decision process, where a controller is selecting from a set of available actions to bias the behavior of a Markov chain. Much is known about Markov decision processes (see e.g. [9]). For example, we could have defined our problem in terms of time-dependent strategies. However, for every Markov decision process there is a time independent optimal strategy, and hence all of our definitions and results consider only time independent strategies. Further conclusions from Markov decision theory are presented in Section 4 below. Our problem differs from the classic setup in that we desire quantitative information: how much can the controller affect characteristics of the random walk such as its stationary distribution? Which are the hardest graphs to influence? These type of questions are not addressed in the general theory.

The problem considered here can also be viewed as a special case of the question: "How do perturbations in a matrix affect the eigenvectors?". There is a rich literature on this subject (see e.g. [13, 16, 11]). Nevertheless, the general results are not very useful for our case — typically the estimate of the change depends on the eigenvalues of the original matrix. For instance, for general matrices (see [13]), the change is inversely proportional to the difference between the first and the second eigenvalue, which may be unbounded. Here, we consider a very restricted situation: the original matrix is stochastic and describes a random walk on a graph, and the support of the perturbation matrix is restricted to the support of the original matrix. This enables us to derive uniform upper and lower bounds (that is, independent of the original matrix or its eigenvalues), whereas the general theory provides only non-uniform upper bounds and no lower bounds.

There are a number of goals for the controller which are worth studying. The present article concentrates on affecting the limit, or long-term behavior of the walk. For example, we may want to make the visit of a particular vertex as likely as possible, or as unlikely as possible. More generally, if we define a weight vector w that associates with every vertex x the benefit w_x of visiting it, it is of interest to find out how to maximize the expected benefit in the limit, i.e., $\sum_{x \in V} \pi_x w_x$, where π is the limit distribution of the biased walk.

We say that a strategy is *simple* if there is a function $b : V \to V$, such that whenever the controller gets to decide where the walk goes, if the walk is at x, the controller forces it to go to $b(x)$.

Here are our main results:

1. In the special case when w equals 1 at one particular vertex (called *root* or *favored vertex*) and 0 everywhere else, that is, the controller tries to maximize the limit frequency of a certain vertex, we can show that under the controller's best strategy, the root's limit probability goes up from $\frac{1}{n}$ to $\left(\frac{1}{n}\right)^{1-c\varepsilon}$, where c depends only on d. (Clearly the result is meaningful only for ε small enough.) For instance, in a d-regular tree, for n and d fixed, as $\varepsilon \to 0$ the root's limit probability is less than
$$n^{-1+\varepsilon\left(\frac{d}{\ln(d-1)}+f(n,d)\right)+O(\varepsilon^2)},$$
where, as $n \to \infty$ and d is fixed, $f(n,d) = O(1/\log n)$. Furthermore the best strategy has a simple characterization in terms of the hitting time to the root. ("Go closer to the root" is not necessarily the best strategy!)
2. The aforementioned bound is tight and is achieved on any expander. That is, if G is an expander, then any strategy for the controller cannot make the limit probability of any vertex larger than $\left(\frac{1}{n}\right)^{1-c\varepsilon}$.
3. Analogous results hold for w which is the characteristic function of a set of vertices $S \subseteq V$, and the limit probability can be raised from $\frac{|S|}{n}$ to $\left(\frac{|S|}{n}\right)^{1-c\varepsilon}$.
4. We can also find an optimal strategy in polynomial time when the goal is to minimize any weighted average of the expected hitting times to a set $S \subseteq V$.

The first three results are presented in Section 3, while the fourth appears in Section 5. For completeness we present the relationship of our problem to Markov decision theory in Section 4, where we draw the following conclusions from the general theory:

1. For any benefit function $w : V \to \mathbf{R}$, we can find an optimal strategy which maximizes $\sum_{x \in V} \pi_x w_x$. This strategy is found in time polynomial in n and in the number of bits needed to represent w.
2. There is always a simple optimal strategy.
3. When the controller wants to maximize the stationary probability of a target vertex v, a strategy is optimal if and only if it has the property that at each vertex the bias is toward a neighbor with minimal biased hitting time to v.

Regarding the increase in limit probability in part (1), there is an interesting similarity that we would like to point out. Ben-Or and Linial [6] study the following question: Given a full binary tree, where a fraction p of the leaves are considered "success", we perform a walk, starting at the root, until we reach a leaf. Suppose the decision where to proceed at internal vertices of the tree is determined by a set of n players. "Honest" players are supposed to take random moves, while the "dishonest" ones may play an optimal strategy to maximize the chance of hitting a success leaf. Ben-Or and Linial show that for every $\varepsilon > 0$, there is a choice of εn dishonest players who have a strategy that increases the probability of success from p to $p^{1-\varepsilon}$. We conjecture that this result is also true in our case:

Conjecture 1.1. *In any graph, a controller can increase the stationary probability of any vertex from p to $p^{1-\varepsilon}$.*

We have proved that the stationary probability can be raised to $p^{1-O(\varepsilon)}$ only for regular, bounded-degree graphs.

A preliminary version of this paper has appeared in [5].

2. Example: The d-regular tree

Consider a d-regular tree of depth l; by this we mean a rooted $(d-1)$-ary tree with a self-loop at the root and $d-1$ self-loops at each leaf. The number of vertices is $n=((d-1)^{l+1}-1)/(d-2)$.

Suppose that the goal of the controller is to maximize the stationary probability of the root. It is clear in this case that the best bias strategy is for each vertex to bias towards its neighbor closest to the root. Let $L(i)$ be the set of vertices at distance i from the root. Let p_i be the limit probability of being at level i in the tree, that is $p_i = \sum_{v \in L(i)} \pi_v$. Let $\alpha = (1-\varepsilon)/d + \varepsilon$ be the probability on edges pointing to the root. Then we have:

$$\begin{aligned} p_0 &= \alpha p_0 + \alpha p_1, \\ p_i &= (1-\alpha)p_{i-1} + \alpha p_{i+1} \text{ for } 0 < i < l, \\ p_l &= (1-\alpha)p_{l-1} + (1-\alpha)p_l. \end{aligned}$$

One can check the solution to these equations is $p_i = p_0(\frac{1-\alpha}{\alpha})^i$, and hence

$$\pi_{\text{root}} = p_0 = \left(\sum_i \frac{1-\alpha}{\alpha}\right)^{-1}.$$

For n and d fixed, as $\varepsilon \to 0$ we have

$$p_0 \le n^{-1+\varepsilon\left(\frac{d}{\ln(d-1)}+f(n,d)\right)+O(\varepsilon^2)},$$

where, as $n \to \infty$ and d is fixed, $f(n,d) = O(1/\log n)$.

3. Bounds on the controller's influence

3.1. Preliminaries

We review the following elementary definitions and theorem from Markov chain theory.

Let Q be the transition probability matrix of a finite irreducible, aperiodic Markov chain. Let π be the vector of stationary probabilities of Q (i.e. $\pi Q = \pi$). Let h_{ij}^t be the probability that the first visit to state j starting at state i occurs at the t'th transition. ($h_{ii}^0 = 1$.) Let $h_i(j)$ be the expected hitting time to state j starting at state i. ($h_i(j) = \sum_{t\ge 1} t h_{ij}^t$, $h_i(i) = 0$.) For $j \ne i$, $h_i(j)$ clearly satisfies the relation

$$h_i(j) = 1 + \sum_k Q_{i,k} h_k(j).$$

We will use h_i instead of $h_i(j)$, whenever the target vertex j is clear from the context.

Let f_i^t be the probability that the first return to state i starting at state i occurs at the t'th transition. Let $R_i = \sum_{t\ge 1} t f_i^t$. The basic limit theorem of Markov chains states that

$$\lim_{n\to\infty} Q_{ji}^n = \pi_i = \frac{1}{R_i}.$$

We need the following elementary lemma.

Lemma 3.1. *Let Q denote the transition probability matrix of an irreducible Markov chain with a finite state space S, and say $|S| = n$. Consider $h = h(0) = (h_0(0) = 0, h_1(0), \ldots, h_{n-1}(0))$, the vector of hitting times to state 0. If $h' = (h'_0, h'_1, \ldots, h'_{n-1})$*

satisfies $h'_0 = 0$ *and for* $i \neq 0$, $h'_i \leq 1 + \sum_{j \in S} Q_{i,j} h'_j$, *then* $h' \leq h$. *Similarly if for each vertex except the origin,* $h'_i \geq 1 + \sum_{j \in S} Q_{i,j} h'_j$, *then* $h' \geq h$.

Proof. We have for $c = 1 + \sum_j Q_{0,j} h_j$ and $c' = 1 + \sum_j Q_{0,j} h'_j$

$$Qh = h - \vec{1} + c e_0$$
$$Qh' \geq h' - \vec{1} + c' e_0$$

and therefore

$$Q(h' - h) \geq (h' - h) + (c' - c) e_0.$$

Hence $h' - h$ is sub-harmonic: at all states except state 0, $h' - h$ is bounded above by a weighted average of $h' - h$ at the "neighboring states" (states to which there is a non-zero transition probability). By irreducibility it follows that if $h' - h$ is not constant, it must have a maximum at the origin. Thus $h' \leq h$. ∎

We will also need the following theorem due to Metropolis *et al.* [15]:

Theorem 3.1. (Metropolis *et al.*) *Let* $G(V, E)$ *be a graph, and let* π *be a strictly positive probability distribution on* V. *Let* d_x *denote the degree of vertex* x. *For each edge* $(x, y) \in E$, *let*

$$M_{x,y} = \begin{cases} \frac{1}{d_x} & \text{if } \pi_x / d_x \leq \pi_y / d_y \\ \frac{1}{d_y} \frac{\pi_y}{\pi_x} & \text{otherwise,} \end{cases}$$

and add a self-loop at each vertex with $M_{x,x} = 1 - \sum_{y \in N(x)} M_{x,y}$. *Then* π *is the stationary distribution of the Markov chain with transition probabilities* $M_{x,y}$.

Proof. For each edge $(x, y) \in E$,

$$\pi_x M_{x,y} = \min\left\{ \frac{\pi_x}{d_x}, \frac{\pi_y}{d_y} \right\} = \pi_y M_{y,x}.$$

Therefore

$$\pi_x M_{x,x} + \sum_{y \in N(x)} \pi_y M_{y,x} = \pi_x M_{x,x} + \sum_{y \in N(x)} \pi_x M_{x,y} = \pi_x,$$

so π is the stationary distribution of the Markov chain given by the transition probabilities $M_{x,y}$. ∎

3.2. Lower bound

Theorem 3.2. *Let $G=(V,E)$ be a connected graph, $S\subset V$, $v\in S$ and $x\in V$. Let $\Delta(x,v)$ be the length of the shortest path between vertices x and v in G and $\Delta(x,S)=\min_{v\in S}\Delta(x,v)$. Let $\beta=1-\varepsilon$. There is a bias strategy for which the stationary probability at S (i.e. the sum of the stationary probabilities of $v\in S$) is at least*

$$\frac{\sum\limits_{v\in S} d_v}{\sum\limits_{v\in S} d_v+\sum\limits_{x\notin S}\beta^{\Delta(x,S)-1}d_x}.$$

Proof. Define a probability distribution π^M by $\pi^M_v=\gamma d_v$ for $v\in S$, and $\pi^M_x=\gamma\beta^{\Delta(x,v)-1}d_x$ for each $x\notin S$, where

$$\gamma=\frac{1}{\sum\limits_{v\in S} d_v+\sum\limits_{x\notin S}\beta^{\Delta(x,S)-1}d_x}.$$

Let M be the transition probability matrix given by the Metropolis theorem, so that π^M is the stationary distribution for M. For all $x\in V$ let $P_{x,x}=0$ and for all $y\neq x$ $P_{x,y}=M_{x,y}/(1-M_{x,x})$ ($M_{x,x}<1$ by construction). Note that P is a transition probability matrix. We claim (1) that P is the transition probability matrix of a ε-biased random walk, and (2) that for each state in S its stationary probability under P is at least its stationary probability under M, which leads to the desired bound.

For (1) note that $P_{x,y}\geq M_{x,y}\geq(1-\varepsilon)/d_x$ for $y\in N(x)$, so P is the transition probability matrix of an ε-biased walk. For (2) denote by $h^M(v)$ the hitting times to vertex $v\in S$. For each vertex $x\neq v$,

$$h^M_x(v)=1+\sum_{y\in N(x)}M_{x,y}h^M_y(v)+M_{x,x}h^M_x(v)$$

or

$$\begin{aligned}h^M_x(v)&=\frac{1}{1-M_{x,x}}+\sum_{y\in N(x)}\frac{M_{x,y}}{1-M_{x,x}}h^M_y(v)\\&\geq 1+\sum_{y\in N(x)}P_{x,y}h^M_y(v).\end{aligned}$$

Hence by Lemma 3.1, for every $v\in S$, $h^M(v)\geq h^P(v)$, where $h^P(v)$ is the vector of hitting times to v in the Markov chain P. Now let R^M_v and R^P_v be the expected return times to v in the two chains:

$$1/\pi^M_v=R^M_v$$

$$
\begin{aligned}
&= 1 + \sum_{y \in N(v)} M_{v,y} h_y^M(v) \\
&= 1 + \sum_{y \in N(v)} P_{v,y} h_y^M(v) \\
&\geq 1 + \sum_{y \in N(v)} P_{v,y} h_y^P(v) = R_v^P = 1/\pi_v^P.
\end{aligned}
$$

The last equality follows from the fact that there is no bias at $v \in S$. In particular, if $v \in S$ and $y \in N(v)$, then $M_{v,y} = P_{v,y} = 1/d_v$. Lastly we have

$$
\sum_{v \in S} \pi_v^P \geq \sum_{v \in S} \pi_v^M = \frac{\sum_{v \in S} d_v}{\sum_{v \in S} d_v + \sum_{x \notin S} \beta^{\Delta(x,S)-1} d_x}. \qquad \blacksquare
$$

Corollary 3.1. *Let $G = (V, E)$ be any connected, d-regular graph and let $S \subset V$. Then there is a bias strategy for which $\sum_{v \in S} \pi_v$ is at least $\left(\frac{|S|}{n}\right)^{1-c\varepsilon}$, for a constant $c > 0$ depending only on d.*

Proof. For d-regular graphs, the lower bound given by the theorem is just $|S|/(|S| + \sum_{x \notin S} \beta^{d(x,S)-1})$. It is minimized for graphs with $(d-1)^i |S|$ nodes at distance i from S. $\blacksquare$

3.3. Upper bound

It can be seen that the above bound for a single target vertex is tight by considering expander graphs.

Theorem 3.3. *Let $G = (V, E)$ be a d-regular expander graph, i.e., the second eigenvalue of its adjacency matrix does not exceed $d - \delta$ for some $\delta > 0$. Then for any bias strategy, the stationary probability at any vertex is at most $n^{c\varepsilon - 1}$ for some constant c depending only on d and δ.*

Proof. Let A be G's adjacency matrix and let $Q = \frac{1}{d} A$ be the transition probability matrix of the standard random walk on G. It is standard (by the spectral theorem) that $Q_{vw}^{(k)}$ (the probability of reaching w in a k-step walk starting from v) satisfies

$$
Q_{vw}^{(k)} = \frac{1}{n} + O(\lambda_2^k),
$$

where λ_2 is the second largest eigenvalue of Q. By assumption, $\lambda_2 \le 1-\delta/d$ so for some c' depending only on $\frac{\delta}{d}$

$$Q_{vw}^{(c'\ln n)} = \frac{1}{n}\left(1+O\left(\frac{1}{n}\right)\right),$$

for all v,w. Fix any bias strategy, and let $P_{vw}^{(k)}$ be the probability of reaching w in a k-step *biased* walk starting from v. We have (for $k=c'\ln n$)

$$\begin{aligned}\pi_w = \sum_{v\in V}\pi_v P_{vw}^{(k)} &\le \sum_{v\in V}\pi_v Q_{vw}^{(k)}\left(\frac{\frac{1-\varepsilon}{d}+\varepsilon}{\frac{1}{d}}\right)^k \\ &\le \sum_{v\in V}\pi_v \frac{1}{n}\left(1+O\left(\frac{1}{n}\right)\right)\left(\frac{\frac{1-\varepsilon}{d}+\varepsilon}{\frac{1}{d}}\right)^{c\ln n} \\ &\le e^{(d-1)\varepsilon k}\frac{1}{n}\left(1+O\left(\frac{1}{n}\right)\right) = n^{c\varepsilon-1},\end{aligned}$$

for some constant c depending only on d and δ. ∎

3.4. The biased return time equation

The biased return time to the target satisfies a simple equation.

Theorem 3.4. *Let $G=(V,E)$ be a connected, d-regular graph and let $v_0 \in V$. Consider any simple bias strategy B. For each vertex $v\in V$, consider $h_v = h_v(v_0)$, the biased hitting time to v_0 (in a walk starting at v), and let $b(v)$ be the neighbor of v that is biased towards in B. Then R_0, the return time at v_0, satisfies*

$$R_0 = n - \varepsilon\sum_{v\in V}\left(h_v - h_{b(v)}\right).$$

Proof. For each vertex v, h_v satisfies

$$h_v = 1 + \frac{1-\varepsilon}{d}\sum_{u\in N(v)} h_u + \varepsilon h_{b(v)}$$

and the return time satisfies

$$R_0 = 1 + \frac{1-\varepsilon}{d}\sum_{u\in N(v_0)} h_u + \varepsilon h_{b(v_0)}.$$

Adding these equations for all v gives

$$R_0 + \sum_{v \in V} h_v = n + \frac{1-\varepsilon}{d} \sum_{v \in V} \sum_{u \in N(v)} h_u + \varepsilon \sum_{v \in V} h_{b(v)}$$

$$= n + (1-\varepsilon) \sum_{v \in V} h_v + \varepsilon \sum_{v \in V} h_{b(v)}$$

and therefore

$$R_0 = n - \varepsilon \sum_{v \in V} \left(h_v - h_{b(v)} \right) .$$

∎

Specializing to the case where ε is small yields the following result.

Corollary 3.2. *Let $G = (V, E)$ be a connected, d-regular graph and let $v_0 \in V$. Consider any simple bias strategy B. For each vertex $v \in V$, consider $\bar{h}_v = \bar{h}_v(v_0)$, the unbiased hitting time to v_0 (in a walk starting at v), and let $b(v)$ be the neighbor of v that is biased towards in B. Then the stationary probability at v_0 satisfies*

$$\pi_{v_0} = \frac{1}{n} \left(1 + \frac{\varepsilon}{n} \sum_{v \in V} (\bar{h}_v - \bar{h}_{b(v)}) \right) + O(\varepsilon^2), \quad \varepsilon \to 0.$$

Proof. Follows from the above theorem, noting that $\pi_{v_0} = 1/R_0$, and $\bar{h}_v = h_v + O(\varepsilon)$ for each vertex v. ∎

The assumption that the bias strategy is simple involves no loss of generality (see Theorem 4.3). Thus the optimal strategy for small ε is to bias towards the neighbor with the smallest unbiased hitting time.

4. Connections with Markov Decision Theory and the properties of the Optimal Strategy

As noted in the introduction, the problem considered in this paper is an instance of a Markov decision problem. This section presents the elements of Markov decision theory that yield useful results in this context. Wherever a theorem has a natural proof in the restricted context of biased random walks, the proof is presented.

We have considered controller strategies that maximize objective functions of the following type: $\sum_{v \in V} w_v \pi_v$. If we allow time-dependent controller strategies, then π_v may not be well defined. We use the relationship between biased random walk strategies and Markov decision theory [9] in order to show that there is a time-independent optimal strategy.

A *Markov decision process* can be described as follows. Consider a system, that at each discrete point of time ($t=0,1,2,\ldots$) is observed and classified into one of a possible number of states, denoted by I. (I is finite.) After each observation of the system, one of a set of possible actions is taken. Let K_i be the set of actions possible when the system is in state i. A (possibly randomized) policy R is a set of distributions $D_a(H_{t-1},Y_t)$, where a is an action in K_{Y_t} meaning that if H_{t-1} is the history of states and actions up to time $t-1$, and Y_t is the state at time t, then the probability that action a is taken at time t is $D_a(H_{t-1},Y_t)$. The actions can be such that they change the state of the system. We define this precisely by saying that $q_{ij}(a)$ is the probability of the system being in state j at the next instant, given that the system is in state i and action a is taken. Another way of saying this is that no matter what policy R is employed,

$$\Pr(Y_{t+1}=j|H_{t-1},Y_t=i,A_t=a)=q_{ij}(a),$$

where A_t is the action taken at time t.

An additional set of parameters associated with a Markov decision process are costs: when the chain is in state i and action a is taken, a known cost c_{ia} is incurred.

Let $S_{R,T}(i)$ be the expected cost of operating a system up to time T using the policy R, given that $Y_0=i$. In other words,

$$S_{R,T}(i)=\sum_{0\le t\le T}\sum_j\sum_a \Pr_R(Y_t=j,A_t=a)w_{ja}.$$

A standard problem in Markov decision theory is to minimize the expected average cost per unit time, i.e. to find a policy R to minimize (or maximize)

$$\limsup_{T\to\infty}\frac{S_{R,T}(i)}{T}.$$

Another standard problem is the optimal first passage problem. This problem consists of finding a policy R that minimizes $S_{R,\tau}(i)$, where τ denotes the smallest positive value of t such that $Y_t=j$.

Our problem of determining the optimal controller strategy to maximize objective functions of the form $\sum_{v\in V} w_v\pi_v$ is a problem of the first type. A biased random walk on a graph $G=(V,E)$ can be phrased as the following Markov decision process $\{Y_t\}$: the set I is the set of vertices V of G and the set of actions K_u is the set of neighbors, $N(u)$, of u in G. The transition probabilities $q_{uv}(x)$ are then

$$q_{uv}(x)=\begin{cases}\frac{(1-\varepsilon)}{d_u}+\varepsilon & \text{if } v=x,\\ \frac{(1-\varepsilon)}{d_u} & \text{if } v\neq x \text{ and } v\in N(u),\\ 0 & \text{otherwise.}\end{cases}$$

The cost of taking action x in state u, c_{ux}, is w_u. Clearly, in the time-dependent case, for a given controller strategy R, the limit

$$\limsup_{T\to\infty} \frac{S_{R,T}(i)}{T}$$

is the natural replacement for the sum $\sum_{v\in V} w_v \pi_v$ that we used in the time-independent case. (Indeed, the limit reduces to $\sum_{v\in V} w_v \pi_v$ if the time dependence is eliminated.)

The following theorem is a basic theorem in Markov decision theory.

Theorem 4.1. ([9], p. 25) *There is an optimal policy that is memoryless, time-invariant and deterministic.*

Therefore, in our context there is a simple time-independent optimal strategy.

4.1. Computing the Optimal Strategy

It follows from results about Markov decision theory that there is an explicit linear programming algorithm for determining the optimal controller strategy. Our specialized setup allows a simpler construction. We present a natural linear program for computing a controller strategy which maximizes (resp. minimizes) $\sum_{x\in V} \pi_x w_x$ for any weight function w. We consider here the more general case of directed graphs, and do not require regularity.

We use the following notation. Let d_x denote the out-degree of node x and let $N^+(x)$ be the set of vertices that can be reached from x by a directed edge. Similarly, $N^-(x)$ is the set of vertices with an edge into x. Let E denote the set of directed edges. If the walk is at x, with probability $1-\varepsilon$ it moves to a vertex from $N^+(x)$ selected uniformly at random; with probability ε the controller selects a neighbor of x according to some distribution. So as before, if the original transition probability matrix of the random walk was Q, the modified transition probability matrix is

$$P = (1-\varepsilon)Q + \varepsilon B,$$

where B is an arbitrary stochastic matrix with support restricted to the directed edges of G, chosen by the controller.

Consider the following system of inequalities:

$$\forall x \in V: \qquad \pi_x \geq 0,$$

$$\forall x \in V: \qquad \pi_x = \sum_{y\in N^-(x)} \frac{1-\varepsilon}{d_y}\pi_y + \sum_{y\in N^-(x)} \varepsilon B_{y,x}\pi_y,$$

$$\forall (x,y) \in E: \qquad B_{x,y} \geq 0,$$
$$\forall x \in V: \qquad \sum_{y \in N^+(x)} B_{x,y} = 1,$$
$$\sum_{x \in V} \pi_x = 1.$$

Clearly the set of feasible solutions to the above system is in one to one correspondence with the set of possible strategies of the controller. The apparent difficulty suggested by the fact that these equations are quadratic can be easily overcome.

Theorem 4.2. *The controller strategy that maximizes* $\sum_{x \in V} \pi_x w_x$ *for any weight function* w *can be found in polynomial time.*

Proof. We convert the quadratic system into a linear program by defining $\forall (x,y) \in E$, $e_{x,y} = B_{x,y}\pi_x$. This yields

$$\forall x \in V: \qquad \pi_x \geq 0,$$
$$\text{(1)} \qquad \forall x \in V: \qquad \pi_x = \sum_{y \in N^-(x)} \frac{1-\varepsilon}{d_y}\pi_y + \sum_{y \in N^-(x)} \varepsilon e_{y,x},$$
$$\forall (x,y) \in E: \qquad e_{x,y} \geq 0,$$
$$\text{(2)} \qquad \forall x \in V: \qquad \sum_{y \in N^+(x)} e_{x,y} = \pi_x,$$
$$\sum_{x \in V} \pi_x = 1.$$

The objective function remains

$$\max \sum_{x \in V} w_x \pi_x \ .$$

This linear program can be solved in polynomial time to obtain an optimal strategy and its value. Note that the polytope of the feasible solutions is non-empty since the stationary distribution of an unbiased strategy is always a feasible solution. The polytope is also bounded, e.g. by the unit cube, thus an optimal strategy always exists. ∎

4.2. Properties of the Optimal Strategy

In this section we present some additional properties of the optimum controller strategy. The following theorem follows from Markov decision theory. For completeness we include a proof.

Theorem 4.3. *There exists a simple optimal bias strategy, i.e. there is an optimal solution to the linear program, such that for each vertex x there is exactly one vertex $y \in N(x)$ such that $e_{x,y} > 0$.*

Proof. It is straightforward to verify that there is a redundant equation out of the $2n+1$ equations of the linear program. (Sum the equations (1) for all $x \in V$, interchange the order of summation and substitute equations (2) to get an identity.) Therefore a vertex of the polytope can have at most $2n$ non-zero coordinates. Since n nonzero variables are accounted for by positive stationary probabilities, at most n of the variables $e_{x,y}$ can be positive. Since for each x, there is at least one y with $e_{x,y}$ positive, we conclude that there is exactly one of y with $e_{x,y}$ positive. Therefore, there is a simple optimal strategy. ∎

The following two theorems characterize the best controller strategy when there is a single target vertex. Analogous results hold when the controller wants to avoid a single vertex.

Theorem 4.4. *When the controller wants to maximize the stationary probability of a target vertex v, the optimal strategy has the property that at each vertex the bias is toward a neighbor with minimal biased hitting time to v.*

Proof. Fix a bias strategy, represented by a matrix B. Consider the vector h of biased hitting times to the target vertex. Assume there is a vertex x, with neighbors y and z, such that $h_y < h_z$ but $b_{x,z} > 0$. We will show that B cannot be optimal.

Construct a new bias matrix B', that differs from B in just two places: $b'_{x,z} = 0$ and $b'_{x,y} = b_{x,y} + b_{x,z}$. By Lemma 3.1 (the "similarly" part, with h and h' exchanging roles), the vector h' of hitting times in the new chain satisfies $h' \le h$. In fact we have strict inequality at x, and because the graph is connected, there must be strict inequality at some neighbor of the target. Therefore the return time to the target is smaller in the new chain. Hence, the stationary probability at the target is strictly larger in the new chain than the old chain and B is not optimal. ∎

The converse to the above theorem is also true:

Theorem 4.5. *Let B be a strategy that biases in the direction of its own biased hitting time to the favored vertex v, and let h be the corresponding hitting times to v. Let h' be the hitting times to v for any other strategy. Then $h \le h'$, and B is an optimal strategy.*

Proof. For each vertex x let $b(x)$ be the neighbor of x that has the least biased hitting time to the target: $b(x)$ must be the vertex that the controller biases towards from x in strategy B. If there are ties, we can assume without loss of generality that the controller biases completely towards just one of its neighbors with least biased hitting time.

Now consider any other simple bias strategy, represented by b', with associated hitting times h'. We have:

$$h'_x = 1 + \frac{(1-\varepsilon)}{d_x} \sum_{y \in N(x)} h'_y + \varepsilon h'_{b'(x)} \,.$$

However, the definition of b gives us

$$\begin{aligned} h_x &= 1 + \frac{(1-\varepsilon)}{d_x} \sum_{y \in N(x)} h_y + \varepsilon h_{b(x)} \\ &\leq 1 + \frac{(1-\varepsilon)}{d_x} \sum_{y \in N(x)} h_y + \varepsilon h_{b'(x)} \,. \end{aligned}$$

Hence by Lemma 3.1 we have $h \leq h'$, so $\pi_v \geq \pi'_v$, and B is optimal. ∎

Comparing the above two theorems with Corollary 3.2 we see that for small enough ε, the optimal strategy is to bias towards the neighbor with least unbiased hitting time to the target v. Under this strategy the stationary probability at v is

$$\pi_v = \frac{1}{n}\left(1 + \frac{\varepsilon}{n} \sum_{x \in V} (\bar{h}_x - \bar{h}_{b(x)})\right) + O(\varepsilon^2), \quad \varepsilon \to 0,$$

where now $b(x)$ denotes the neighbor of x with least unbiased hitting time to v.

Corollary 4.1. *The strategy that maximizes the stationary probability at a vertex v minimizes $h_x(v)$ for all x.*

Proof. Let h be the vector of hitting times to v under the strategy that maximizes the stationary probability at v, and let h' be the hitting times to v under any other bias strategy. By Theorem 4.4 and Theorem 4.5, we have $h \leq h'$. ∎

Lastly we note some properties of the value of the optimal strategy, as a function of ε.

Theorem 4.6. *The value of the optimal strategy is a non-decreasing function of ε. It is continuous and piecewise differentiable.*

Proof. The first fact follows because any legal bias strategy for ε' is also a legal strategy for ε if $\varepsilon' < \varepsilon$. Secondly, the value for any strategy is given by the solution to a set of linear equations, hence is a rational function of ε. For each ε, the optimum strategy is obtained by choosing the minimum value simple strategy. Hence the optimum value is the minimum of a finite collection of rational functions of ε, so the optimum value is a continuous and piecewise differentiable function of ε. ∎

5. The hitting time

In this section we consider the biased hitting time problem. We are given the same graph and random walk as before, but now the goal is to minimize the hitting time to some set $S \subset V$ from all vertices. More precisely, for a bias strategy B, let $h_x(S)$ be the expected biased hitting time from x to the set S. Clearly for each $x \in S$, $h_x(S) = 0$. Our objective function is to minimize over all possible bias strategies $\sum_{x \in V} l_x h_x(S)$ where l_x is the weight of vertex x. We will assume that for each vertex x, $l_x \geq 0$.

The biased hitting time problem is an example of an optimal first-passage problem from Markov decision theory. The model is the same as that used for our previous problem, except that now the cost of every action c_{ia} is 1 and we start from an initial distribution proportional to l_x. Consequently we once again have

Theorem 5.1. ([9], p. 29) *There is an optimal controller strategy for the biased hitting time problem which is memoryless, time-invariant and deterministic.*

The main result of this section is

Theorem 5.2. *Given S, there is a controller strategy for the hitting time problem that is optimal for all objective functions such that for all x, the weight l_x is non-negative. The strategy can be found in polynomial time.*

Proof. Consider the graph G_S in which the nodes in S are shrunk into a single node s. The graph G_S may have multiple edges: the number of edges between a vertex $v \notin S$ and s will be the number of neighbors of v in S. Now for any vertex $v \notin S$ and for any bias strategy, the hitting time from v to S in G is equal to the hitting time from v to s in G_S. We now show how to find a strategy that minimizes this hitting time, using the results on maximizing the stationary probability at a target vertex.

Firstly we can find a bias strategy for G_S that maximizes the stationary probability at vertex s using linear programming, as shown in Section 4.1. Note that the results of Section 4.1 still hold even though G_S has multiple edges. In the notation of that section, d_y is now the number of edges leaving a vertex y, and $N^-(y)$ is a multi-set, with a vertex x appearing once for each edge (y, x).

Now let h be the vector of hitting times to s under this optimum strategy. By Corollary 4.1 $h_x(s)$ is minimal for every x. (Note again that the Corollary is still valid when G_S has multiple edges.) Therefore, for any non-negative objective function l, the bias strategy that maximizes the stationary probability at vertex s is also an optimal strategy for the biased hitting time problem to S. ∎

Acknowledgments. Nati Linial wishes to acknowledge his conversations with Oded Goldreich which helped in formulating some of initial conjectures in the present research. We wish to thank an anonymous referee for many useful comments and corrections.

References

[1] M. Ajtai, J. Komlós, and E. Szemerédi: Deterministic Simulation in LOGSPACE, *19th Annual ACM STOC*, 1987, 132–140.

[2] M. Ajtai, and N. Linial: The influence of large coalitions, *Combinatorica*, **13** (1993) 129–145.

[3] N. Alon, and M. Naor: Coin-flipping Games Immune Against Linear-sized Coalitions, *SIAM J. Comput.* **22** (1993), 403–417.

[4] N. Alon, and M. O. Rabin: Biased Coins and Randomized Algorithms, in *Randomness and Computation (S. Micali ed.)* Advances in Computing Research, Vol. 5, 499–507.

[5] Y. Azar, A. Z. Broder, A. R. Karlin, N. Linial, and S. J. Phillips: Biased Random Walks, in *24th Annual ACM STOC*, 1992, 1–9.

[6] M. Ben-Or, and N. Linial: Collective coin flipping, in *Randomness and Computation (S. Micali ed.)* Academic Press, New York, 1990, 91–115.

[7] M. Ben-Or, N. Linial, and M. Saks: Collective coin flipping and other models of imperfect randomness, in *Colloq. Math Soc. János Bolyai no. 52, Combinatorics Eger*, 1987, 75–112.

[8] A. Cohen, and A. Wigderson: Dispersers, Deterministic Amplification, and Weak Random Sources, *30th Annual IEEE FOCS*, 1989, 14–19.

[9] C. Derman: *Finite State Markov Decision Processes*, Academic Press, New York, 1970.

[10] O. Goldreich, S. Goldwasser, and N. Linial: Fault-tolerant Computation in the Full Information Model, *32nd Annual IEEE FOCS*, 1991, 447–457.

[11] E. J. Hinch: *Perturbation methods*, Cambridge University Press, 1991.

[12] R. Impagliazzo, and D. Zuckerman: How to Recycle Random Bits, *30th Annual IEEE FOCS*, 1989, 248–253.

[13] T. Kato: *Perturbation theory for linear operators*, New York, Springer Verlag, 1982.

[14] D. Lichtenstein, N. Linial, and M. Saks: Some extremal problems arising from discrete control processes, *Combinatorica*, **9** (1989), 269–287.

[15] M. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and M. Teller: Equation of state calculations by fast computing machines, *Journal of Chemical Physics* **21** (1953), 1087–1092.

[16] F. Rellich: *Perturbation theory of eigenvalue problems*, New York, Gordon and Breach, 1969.

[17] M. Santha, and U. Vazirani: Generating Quasi-Random Sequences from Semi-random Sources, *J. Comput. System Sci.* **33** (1986), 75–87.

[18] U. Vazirani: Efficiency Considerations in Using Slightly-Random Sources, *19th Annual ACM STOC*, 1987, 160–168.

[19] U. VAZIRANI: Randomness, Adversaries and Computation, Ph.D. Thesis, University of California, Berkeley, 1986.

Yossi Azar

Department of Computer Science,
Tel Aviv University,
Israel
azar@math.tau.ac.il

Andrei Z. Broder

Digital Equipment Corporation
Systems Research Center,
130 Lytton Avenue, Palo Alto,
CA 94301, USA.
broder@src.dec.com

Anna R. Karlin

Department of Computer Science,
University of Washington,
Seattle, WA 98195, USA
karlin@cs.washington.edu

Nathan Linial

The Hebrew University of Jerusalem
Israel
nati@cs.huji.ac.il

Steven Phillips

AT&T Bell Laboratories
600 Mountain Avenue,
Murray Hill, NJ 07974,
U.S.A.
phillips@research.att.com