Optimal Cover Time for a Graph-Based Coupon Collector Process

Nedialko B. Dimitrov* and C. Greg Plaxton**

University of Texas at Austin,
1 University Station C0500,
Austin, Texas 78712–0233
{ned, plaxton}@cs.utexas.edu

Abstract. In this paper we study the following covering process defined over an arbitrary directed graph. Each node is initially uncovered and is assigned a random integer rank drawn from a suitable range. The process then proceeds in rounds. In each round, a uniformly random node is selected and its lowest-ranked uncovered outgoing neighbor, if any, is covered. We prove that if each node has in-degree $\Theta(d)$ and out-degree $O(d)$, then with high probability, every node is covered within $O(n + \frac{n \log n}{d})$ rounds, matching a lower bound due to Alon. Alon has also shown that, for a certain class of d-regular expander graphs, the upper bound holds no matter what method is used to choose the uncovered neighbor. In contrast, we show that for arbitrary d-regular graphs, the method used to choose the uncovered neighbor can affect the cover time by more than a constant factor.

1 Introduction

One of the most commonly discussed stochastic processes in computer science is the so-called coupon collector process [7]. In that process, there are n distinct coupons and we proceed in rounds, collecting one uniformly random coupon (with replacement) in each round. Are $O(n)$ rounds sufficient to collect all of the coupons? Put differently, is picking coupons with replacement as efficient, to within a constant factor, as picking them without replacement? No, it is a well-known fact that with high probability the number of rounds required to collect all of the coupons is $\Theta(n \log n)$.

This shortcoming has motivated Adler et al. [1] and Alon [2] to study a similar graph-based covering process. The nodes of the graph nodes represent the coupons and covering a node represents collecting a coupon. In each round, a uniformly random node w is selected. If an uncovered neighbor of w exists, choose one such uncovered neighbor and cover it. We refer to this process as process CC.

Process CC can use a variety of different *covering methods* to decide which uncovered neighbor to cover. If our ultimate goal is to minimize cover time, certainly the most powerful covering method available is an offline method with knowledge of the entire

* Supported by an MCD Fellowship from the University of Texas at Austin.
** Supported by NSF Grants CCR–0310970 and ANI–0326001. Also affiliated with Akamai Technologies, Inc., Cambridge, MA 02142.

L. Caires et al. (Eds.): ICALP 2005, LNCS 3580, pp. 702–716, 2005.

sequence of node selections and with infinite computing power. We refer to this powerful cover time minimizing version of process CC as process MIN. To achieve our $O(n)$ goal, it is natural to consider $\log n$-regular graphs since the work of Alon implies process MIN has an expected cover time of $\Omega(n + \frac{n \log n}{d})$ rounds on d-regular graphs [2].

1.1 Logarithmic-Degree Graphs

Another natural version of process CC — in which the covering method chooses a uniformly random uncovered neighbor, if any — was studied by Adler et al. [1] and by Alon [2]. We refer to this version of process CC as process UNI. Alon shows that for logarithmic-degree Ramanujan expander graphs, process UNI completes in $O(n)$ time, matching the lower bound for process MIN.

Adler et al. show that for the hypercube, which has a weak expansion property but is not an expander, process UNI takes $O(n)$ time, also matching the lower bound for process MIN [1]. They also show that for arbitrary logarithmic-degree graphs, process UNI completes in $O(n \log \log n)$ time. Furthermore, Adler et al. present an application of process UNI to load balancing in a hypercubic distributed hash table (DHT).

A process that is intuitively similar to process UNI is one where we initially assign a rank to each node using a uniformly random permutation of the nodes, and the covering method covers the minimum-rank uncovered neighbor, if any. We refer to this permutation-based version of process CC as process P-RANK. In this paper, we show that process P-RANK completes in $O(n)$ time on arbitrary logarithmic-degree graphs.

In fact, we analyze a more general and local version of process CC in which each node initially chooses a uniformly random rank in a suitable range, and the covering method covers the minimum-rank uncovered neighbor of the selected node. (We assume that the nodes are numbered from 1 to n, and that ties in rank are broken in favor of the lower-numbered node.) We refer to this random rank version of process CC as process R-RANK. We show that the more general and local process R-RANK completes in $O(n)$ time on arbitrary logarithmic-degree graphs.

1.2 Results for General Graphs

Alon shows that process MIN on any d-regular graph has expected cover time at least $n - \frac{n}{d} + \frac{n}{d} \ln(\frac{n}{d})$ [2]. Alon also shows that process UNI completes in time $n + (1 + o(1))\frac{n \ln n}{d}$ for random nearly d-regular graphs. Alon further shows that on any (n, d, λ)-expander graph the expected cover time of process UNI is at most $n + n(\frac{\lambda}{d})^2(\ln n + 1)$. In particular, this implies that on Ramanujan graphs process UNI completes in $(1 + o(1))n$ time, matching the lower bound for process MIN.

If our goal is to maximize cover time, certainly the most powerful covering method available is an offline adversary with knowledge of the entire sequence of node selections and with infinite computing power. We refer to this powerful cover time maximizing version of process CC as process MAX. Alon notes that the upper bounds for expanders hold even if after every round an adversary "is allowed to shift the uncovered nodes to any place he wishes, keeping their number." In particular, this shows that on Ramanujan graphs, the cover time for process MAX matches the cover time for process MIN, up to constant factors. In effect, the covering method does not matter for this class of graphs.

Another previously studied variant of process CC favors covering the selected node. In this variant, we check — immediately after selecting a uniformly random node — if the selected node is uncovered. If it is, we cover it and move to the next selection. Only otherwise do we consider the neighbors of the selected node. We refer to the selection-biased variants of processes UNI, P-RANK, and R-RANKas UNI′, P-RANK′, and R-RANK′, respecively.

Adler et al. show that for all d-regular graphs, processes UNI and UNI′ finish in $O(n + n(\log n)(\log d)/d)$ time[1]. They also show that for random d-regular graphs only $O(n+\frac{n \log n}{d})$ steps are needed. Furthermore, they exhibit an application of process UNI′ to load balancing in DHTs.

All of the results matching Alon's lower bound for process MIN presented prior to this work have used some expansion properties of the underlying graph. In contrast, our proof techniques do not require the underlying graph to have any particular structure. Thus, we show the following general result: for directed graphs, with self-loops but no parallel edges, where each node has in-degree at least δ_{in} and at most Δ_{in}, and out-degree at most Δ_{out}, both process R-RANK and process R-RANK′ cover all nodes in $O(n \max(\Delta_{\text{in}}\Delta_{\text{out}}/\delta_{\text{in}}^2, (\log n)/\delta_{\text{in}}))$ rounds with high probability. This result matches Alon's lower bound for $\delta_{\text{in}} = \Delta_{\text{in}} = \Delta_{\text{out}} = \Theta(d)$, and is thus optimal under these conditions.

Furthermore, Alon's results for Ramanujan graphs raise the question whether there is any separation between the cover times for process MAX and process MIN. In other words, are there any graphs for which the choice of covering method matters? We define a weakly adversarial process, process A-RANK, that is similar to process P-RANK. In process A-RANK, instead of picking a uniformly random permutation, an adversary is initally allowed to fix the permutation used to assign ranks to the nodes. We then proceed as in process P-RANK. In addition, we define the selection-biased variant of process A-RANK as process A-RANK′. We establish that there exists a logarithmic-degree graph on which process A-RANK and process A-RANK′ each take $\omega(n)$ rounds to complete. This implies that in general there is separation between the cover times of process MIN and process MAX. In other words, the covering method does matter. Due to space limitations, the proofs of our $\omega(n)$ lower bounds for processes A-RANKand A-RANK′ are omitted from this paper; these proofs may be found in [4].

1.3 Proof Outline

The proof of our theorem is inspired by the delay sequence argument used by Ranade for the analysis of a certain packet routing problem on the butterfly [8] (see also [6]). In a delay sequence argument, we identify certain combinatorial structures that exist whenever the random process lasts for a long time. Then, we show that the probability any of these structures exist is small. This in turn implies an upper bound on the running time of the random process.

There are significant differences between our proof and that of Ranade. For example, in our problem, the connection between the running time and the length of a delay sequence is not clear-cut, while in the butterfly routing problem analyzed by Ranade, the length of the delay sequence is equal to the running time. But let us begin by giving the notion of a delay sequence in our problem.

Consider the node that was covered last, say w_1. Why wasn't w_1 covered earlier? It was not covered earlier because at the last opportunity to cover w_1 — that is, the last selection in w_1's neighborhood — we covered some other node, w_2, instead. In such a case we consider w_1 to be delayed by w_2. Similarly, w_2 may be delayed by some node w_3, et cetera, until finally we reach a node w_k that is not delayed, i.e., w_k is covered at the first opportunity. The sequence of nodes $w_1, \ldots, w_k$ corresponds to our notion of a delay sequence.

In analyzing process R-RANK, we find it useful to first analyze a much simpler process, process SELECT, in which we repeatedly select a uniformly random node, never covering anything. After establishing several lemmas for the simpler process, we proceed to analyzing process R-RANK. This is the bulk of the proof, and includes a technical lemma to work around the difficulties in linking cover time to delay sequence length. Finally, we reduce process R-RANK$'$ to process R-RANK to show that the same bounds hold.

Due to space limitations, we omit our analysis of process R-RANK$'$ from the present paper. In [4], we analyze process R-RANK$'$ via a reduction from process R-RANK; in addition, we establish the existence of a logarithmic-degree graph on which processes A-RANKand A-RANK$'$ each take $\omega(n)$ rounds to complete, establishing that the covering method does matter.

The rest of this paper is structured as follows. In Section 2, we present a number of useful definitions and lemmas related to standard probability distributions. In Section 3, we analyze process SELECT. In Section 4, we analyze process R-RANK using the results in Section 3.

2 Preliminaries

We use the term ℓ-sequence to refer to a sequence of length ℓ. For any ℓ-sequence σ of elements of a given type, and any element x of the same type, we let $\sigma : x$ denote the $(\ell + 1)$-sequence obtained by appending element x to σ.

For any nonnegative integer n and probability p, we use the notation $X \sim \mathrm{Bin}\,(n, p)$ to denote that the random variable X has a binomial distribution with n trials and success probability p. Similarly, we write $X \sim \mathrm{Geo}\,(p)$ to indicate that the random variable X has a geometric distribution with success probability p, and we write $X \sim \mathrm{NegBin}\,(r, p)$ to indicate that the random variable X has a negative binomial distribution with r successes and success probability p. Due to space limitations, we include proofs for only two of the lemmas stated in this section. The other lemmas follow from simple arguments involving independence of random variables or tail bounds for the binomial distribution [4].

Lemma 1. *Let p denote an arbitrary probability, let ℓ denote an arbitrary nonnegative integer, and let $X \sim \mathrm{NegBin}\,(\ell, p)$. For any integer j such that $1 \leq j \leq \ell$, let p_j denote an arbitrary probability such that $p_j \geq p$, let $Y_j \sim \mathrm{Geo}\,(p_j)$, and let $Y = \sum_{1 \leq j \leq \ell} Y_j$. Then for any nonnegative integer i, $\Pr(X \geq i) \geq \Pr(Y \geq i)$.*

Proof. Note that if $p_j = p$ for all j, then the random variables X and Y have the same distribution. Furthermore, increasing any of the p_j's can only decrease Y. □

Lemma 2. *For any nonnegative integers r and n, and any probability p, we have $\Pr(X < r) = \Pr(Y > n)$, where $X \sim \mathrm{Bin}(n, p)$ and $Y \sim \mathrm{NegBin}(r, p)$.*

Proof. The random variables X and Y can be seen as different views of the same experiment where we successively flip coins with probability of success p. With Y, we ask "How many flips are required for r successes?" With X, we ask "How many successes are in the first n flips?" In this experiment, the event of seeing less than r successes in the first n flips ($X < r$) corresponds to the event that we have to wait more than n flips for the first r successes ($Y > n$). This gives the result. □

Lemma 3. *For any integer $r \geq 2$, $\Pr(X \geq 2E[X]) = \Pr(X \geq 2r/p) \leq \exp(-r/8)$, where $X \sim \mathrm{NegBin}(r, p)$.*

Proof. Let $j = \left\lfloor \frac{2r}{p} \right\rfloor - 1$ and let $Y \sim \mathrm{Bin}(j, p)$. By Lemma 2, we know that $\Pr(X \geq \frac{2r}{p}) \leq \Pr(X \geq \left\lfloor \frac{2r}{p} \right\rfloor) = \Pr(X > \left\lfloor \frac{2r}{p} \right\rfloor - 1) = \Pr(Y < r) = \Pr(Y \leq r - 1)$.

$$\begin{aligned}\Pr\left(Y \leq \frac{jp}{2}\right) &= \Pr\left(Y \leq r - (\eta+1)\frac{p}{2}\right) \\ &= \Pr(Y \leq r - 1)\end{aligned}$$

where $\frac{2r}{p} = \left\lfloor \frac{2r}{p} \right\rfloor + \eta$ and the last equality holds because $0 < (\eta+1)\frac{p}{2} < 1$.

Recall the Chernoff bounds in the form $\Pr(Y \leq (1-\lambda)jp) \leq \exp(-\lambda^2 jp/2)$ for $0 < \lambda < 1$ (see [3, 5]).

We apply this bound with $\lambda = \frac{1}{2}$ to get

$$\begin{aligned}\Pr(Y \leq r - 1) &= \Pr(Y \leq jp/2) \\ &\leq \exp(-jp/8) \\ &\leq \exp\left(\frac{-2r + (\eta+1)p}{8}\right) \\ &\leq \exp(-r/8)\end{aligned}$$

where η is as previously defined and the last inequality holds because $r \geq 2$. □

Lemma 4. *Let p be an arbitrary probability and let X be the sum of n independent Bernoulli variables $X_1, \ldots, X_n$, where X_j has success probability $p_j \geq p$. Then $\Pr(X \leq np/2) \leq \exp(-np/12)$.*

Proof. The result follows from Chernoff bounds (see, e.g., [3, 5]). □

Lemma 5. *Suppose we repeatedly throw balls independently and uniformly at random into n bins, and let the random variable X denote the number of throws required for every bin to receive at least n balls. Then X is $O(n^2)$ with high probability, that is, with failure probability that is an arbitrary inverse polynomial in n.*

Proof. The result follows from Lemma 4. □

3 Process SELECT

Throughout the remainder of the paper, we fix an arbitrary directed graph $G = (V, E)$ where $|V| = n > 0$. We say that an event holds "with high probability" if the probability that it fails to occur is upper bounded by an arbitrary inverse polynomial in n. We let δ_{in}, Δ_{in}, and Δ_{out} denote the minimum in-degree, maximum in-degree, and maximum out-degree of any node, respectively. For ease of exposition, we assume throughout the paper that $\delta_{\text{in}} > 0$. The edge set E is allowed to contain loops but not parallel edges. For any node v, we define $\Gamma_{\text{in}}(v)$ as $\{w \mid (w, v) \in E\}$. For any sequence of edges $\sigma = (u_1, v_1), \ldots, (u_\ell, v_\ell)$, we define the two sequences of nodes $\text{src}(\sigma) = u_1, \ldots, u_\ell$ and $\text{dst}(\sigma) = v_1, \ldots, v_\ell$.

In this section, we analyze a simple stochastic process, process SELECT, defined as follows. Initially, we fix a positive integer r and independently assign each node in V a uniformly random integer rank between 1 and r. Process SELECT then proceeds in an infinite number of rounds, indexed from 1. In each round, one node is selected uniformly at random, with replacement. The following definitions are central to our analysis of this process.

A node sequence is said to be *rank-sorted* if the associated sequence of node ranks is nondecreasing.

For any node sequence σ, we inductively define a nonnegative integer $\text{duration}(\sigma)$ and a node sequence $\text{select}(\sigma)$ as follows. If σ is empty, then $\text{duration}(\sigma)$ is 0 and $\text{select}(\sigma)$ is empty. Otherwise, σ is of the form $\tau : v$ for some shorter node sequence τ and node v. Let i denote the the least i such that $i > \text{duration}(\tau)$ and the node selected in round i belongs to $\Gamma_{\text{in}}(v)$. Let u denote the node selected in round i. Then we define $\text{duration}(\sigma)$ as i, and $\text{select}(\sigma)$ as $\text{select}(\tau) : u$.

Lemma 6. *For any ℓ-sequence of distinct nodes σ, $\Pr(\sigma$ is rank-sorted$) = \binom{\ell+r-1}{\ell} r^{-\ell}$.*

Proof. There are $\binom{\ell+r-1}{\ell}$ ways that ranks can be assigned to the ℓ distinct nodes so that the resulting ℓ-sequence is rank-sorted. The result follows since each such assignment occurs with probability $r^{-\ell}$. □

Lemma 7. *For any ℓ-sequence of nodes $\sigma = v_1, \ldots, v_\ell$ and any nonnegative integer i, we have*
$\Pr(\text{duration}(\sigma) = i) \leq \Pr(X \geq i)$*, where* $X \sim \text{NegBin}\left(\ell, \frac{\delta_{\text{in}}}{n}\right)$.

Proof. We proceed by proving that

$$\Pr(\text{duration}(\sigma) = i) = \Pr\left(\sum_{k=1}^{\ell} Y_k = i\right)$$

where $Y_k \sim \text{Geo}\left(\frac{d_k}{n}\right)$ and d_k denotes the in-degree of v_k. The desired bound then follows by Lemma 1.

We prove the foregoing claim by induction on ℓ. If $\ell = 0$, the claim holds since $\text{duration}(\sigma) = \sum_{k=1}^{\ell} Y_k = 0$.

For $\ell > 0$, we let τ denote the node sequence $v_1, \ldots, v_{\ell-1}$ and assume inductively that

$$\Pr(\text{duration}(\tau) = i) = \Pr\left(\sum_{k=1}^{\ell-1} Y_k = i\right).$$

Thus,

$$\begin{aligned}
\Pr(\text{duration}(\sigma) = i) &= \sum_{j=0}^{i-1} \Pr(\text{duration}(\tau) = j) \cdot \Pr(\text{duration}(\sigma) - \text{duration}(\tau) \\
&= i - j \mid \text{duration}(\tau) = j) \\
&= \sum_{j=0}^{i-1} \Pr(\text{duration}(\tau) = j) \cdot \Pr(\text{duration}(\sigma) - \text{duration}(\tau) \\
&= i - j) \\
&= \sum_{j=0}^{i-1} \Pr(\text{duration}(\tau) = j) \cdot \Pr(Y_\ell = i - j) \\
&= \sum_{j=0}^{i-1} \Pr\left(\sum_{k=1}^{\ell-1} Y_k = j\right) \cdot \Pr(Y_\ell = i - j) \\
&= \Pr\left(\sum_{k=1}^{\ell-1} Y_k = i\right).
\end{aligned}$$

The second equality holds because each selection is independent of previous selections. The third equality holds because the waiting time to obtain a selection in $\Gamma_{\text{in}}(v_\ell)$ is distributed as Y_ℓ. □

Lemma 8. *For any ℓ-sequence of edges σ, $\Pr(\text{select}(\text{dst}(\sigma)) = \text{src}(\sigma)) \leq \delta_{\text{in}}^{-\ell}$.*

Proof. We proceed by induction on ℓ. For $\ell = 0$, $\Pr(\text{select}(\text{dst}(\sigma)) = \text{src}(\sigma)) = 1 = \delta_{\text{in}}^0$ since we have assumed that $\delta_{\text{in}} > 0$.

For $\ell > 0$, σ can be written in the form $\tau : (u, v)$, where we inductively assume that the claim of the lemma holds for τ. Let A denote the event that the first node selected in $\Gamma_{\text{in}}(v)$ after round $\text{duration}(\text{dst}(\tau))$ is u. We have

$$\begin{aligned}
\Pr(\text{select}(\text{dst}(\sigma)) = \text{src}(\sigma)) &= \Pr(\text{select}(\text{dst}(\tau)) = \text{src}(\tau)) \cdot \Pr(A \mid \text{select}(\text{dst}(\tau)) = \text{src}(\tau)) \\
&= \Pr(\text{select}(\text{dst}(\tau)) = \text{src}(\tau)) \cdot \Pr(A) \\
&\leq \delta_{\text{in}}^{-\ell}.
\end{aligned}$$

The second step follows from the independence of the events A and $\text{select}(\text{dst}(\sigma')) = \text{src}(\sigma')$. (These two events are independent since each selection is independent of previous selections.) The third step follows from the induction hypothesis and the observation that $\Pr(A)$ is equal $1/\Gamma_{\text{in}}(v)$, which is at most $1/\delta_{\text{in}}$. □

Lemma 9. *For any ℓ-sequence of edges σ and nonnegative integer i, the events $A =$ "$\mathrm{dst}(\sigma)$ is rank-sorted", $B =$ "$\mathrm{duration}(\mathrm{dst}(\sigma)) = i$", and $C =$ "$\mathrm{select}(\mathrm{dst}(\sigma)) = \mathrm{src}(\sigma)$" are mutually independent.*

Proof. Note that event A depends only on the rank assignments, while events B and C depend only on the selections. Thus event A is independent of events B and C. Below we argue that events B and C are independent.

Let $\sigma = (u_1, v_1), \ldots, (u_\ell, v_\ell)$ and let σ_j denote the length-j prefix of σ, $0 \leq j \leq \ell$. Define a selection to be *j-special*, $1 \leq j \leq \ell$, if it is the first selection after round $\mathrm{duration}(\sigma_{j-1})$ in $\Gamma_{\mathrm{in}}(v_j)$. A selection is *special* if it is j-special for some j. Note that event B depends only on the timing of the special events; in particular, B occurs if and only if the ℓ-special selection occurs in round i. Suppose we run process SELECT, but at each step, instead of revealing the selected node, we reveal only whether the selection is special. This information is sufficient to determine the unique i for which B occurs, but does not bias the distribution of $\mathrm{select}(\mathrm{dst}(\sigma))$. Since event C only depends on $\mathrm{select}(\mathrm{dst}(\sigma))$, it is independent of B. □

Lemma 10. *Let σ be an ℓ-sequence of edges so that the nodes of $\mathrm{dst}(\sigma)$ are distinct, let $X \sim \mathrm{NegBin}\left(\ell, \frac{\delta_{\mathrm{in}}}{n}\right)$, let i be a nonnegative integer, and let events A, B, and C be defined as in the statement of Lemma 9. Then $\Pr(A \cap B \cap C) \leq \binom{\ell+r-1}{\ell} \cdot \Pr(X \geq i) \cdot (r\delta_{\mathrm{in}})^{-\ell}$.*

Proof. By Lemma 6, $\Pr(A) \leq \binom{\ell+r-1}{\ell} r^{-\ell}$. By Lemma 7, $\Pr(B) \leq \Pr(X \geq i)$. By Lemma 8, $\Pr(C) \leq \delta_{\mathrm{in}}^{-\ell}$. The claim then follows by Lemma 9. □

4 Process R-RANK

In the section we analyze an augmented version of process SELECT, referred to as Process R-RANK, in which we maintain a notion of a "covered subset" of the nodes. Initially, all of the nodes are uncovered. Process R-RANK then proceeds in rounds in exactly the same manner as process SELECT, except that in any given round, if one or more outgoing neighbors of the selected node are uncovered, we cover the uncovered outgoing neighbor with minimum rank. (As indicated in Section 1, ties are broken according to some arbitrary numbering of the nodes.)

Note that process R-RANK simply augments process SELECT by also covering nodes; rank assignment and selections are performed in exactly the same manner in the two processes. Thus all of the definitions and lemmas presented in Section 3 are applicable to process R-RANK. The following additional definitions are useful for our analysis of process R-RANK.

The *cover time* of process R-RANK is defined as the number of rounds required to cover all of the nodes.

We inductively define the notion of a *linked* sequence of edges. For ℓ equal to 0 or 1, any ℓ-sequence of edges is linked. For $\ell > 1$, an ℓ-sequence of edges of the form $\sigma : (u, v) : (u', v')$ is linked if the $(\ell - 1)$-sequence $\sigma : (u, v)$ is linked and (u, v') belongs to E.

For any node v, we define $\mathrm{parent}(v)$ as follows. Let i denote the round in which node v is covered. If i is the first round in which some node in $\Gamma_{\mathrm{in}}(v)$ is selected, then $\mathrm{parent}(v)$ is defined to be nil. Otherwise, $\mathrm{parent}(v)$ is the node covered in the first round prior to round i in which the selected node belongs to $\Gamma_{\mathrm{in}}(v)$.

We inductively define the notion of a *chronological* sequence of nodes as follows. Any ℓ-sequence of nodes with $\ell \leq 1$ is chronological. An ℓ-sequence of nodes of the form $\sigma : v : v'$ is chronological if $\sigma : v$ is chronological and node v is covered before node v'.

We inductively define the notion of an *active* node sequence as follows. The empty node sequence is active. A singleton node sequence consisting of the node v is active if $\mathrm{parent}(v) = \mathrm{nil}$. An ℓ-sequence of nodes of the form $\sigma : v : v'$ is active if $\sigma : v$ is active and $\mathrm{parent}(v') = v$.

We call an ℓ-sequence of edges σ *active* if $\mathrm{dst}(\sigma)$ is active and $\mathrm{select}(\mathrm{dst}(\sigma)) = \mathrm{src}(\sigma)$.

We call an ℓ-sequence of edges σ *i-active* if it is active and either $\ell = i = 0$ or $\ell > 0$, σ is of the form $\sigma : (u, v)$, and v is the node covered in round i.

Lemma 11. *For any nonnegative integer ℓ, there are at most $n\Delta_{\mathrm{out}}^{\ell}\Delta_{\mathrm{in}}^{\ell-1}$ linked ℓ-sequences of edges.*

Proof. We proceed by induction on ℓ, treating $\ell = 0$ and $\ell = 1$ as the base cases. For $\ell = 0$, the empty sequence is the only linked 0-sequence, and the claim holds since $n/\Delta_{\mathrm{in}} \geq 1$. (Note that Δ_{in} is at most n since we do not allow parallel edges.) For $\ell = 1$, the number of linked 1-sequences is at most $|E| \leq n\Delta_{\mathrm{out}}$.

Now let ℓ be greater than 1 and inductively assume that the number of linked $(\ell-1)$-sequences of edges is at most $n\Delta_{\mathrm{out}}^{\ell-1}\Delta_{\mathrm{in}}^{\ell-2}$. Recall that any linked ℓ-sequence of edges is of the form $\sigma : (u, v) : (u', v')$ where the $(\ell-1)$-sequence of edges $\sigma : (u, v)$ is linked and (u, v') belongs to E. Observe that for any linked $(\ell-1)$-sequence of edges $\sigma : (u, v)$, there are at most Δ_{out} nodes v' such that (u, v') belongs to E, and for each such choice of v', there are at most Δ_{in} nodes u' such that (u', v') belongs to E. Thus the number of linked ℓ-sequences is at most $\Delta_{\mathrm{out}}\Delta_{\mathrm{in}}$ times the number of linked $(\ell-1)$-sequences, and the desired bound follows from the induction hypothesis. □

Lemma 12. *Suppose we run two instances of process R-RANK in parallel using the same random ranks and the same sequence of random selections, but in the second instance, we allow an arbitrary subset of the covered nodes to be uncovered after each round. Then the cover time of the first instance is at most the cover time of the second instance.*

Proof. By a straightforward induction on the number of rounds, at all times, the set of covered nodes in the first instance contains the set of covered nodes in the second instance. The claim of the lemma follows. □

Lemma 13. *For any rank assignment, the expected cover time of process R-RANK is $O(n^2)$.*

Proof. It follows from Lemma 5 that the cover time is $O(n^2)$ with high probability since in that time each vertex is selected at least n times, implying that all of its neighbors are covered.

We can then consider a modified version of process R-RANK in which the infinite sequence of rounds is partitioned into epochs of $O(n^2)$ rounds, and where at the end of each epoch, if the nodes are not all covered, all nodes are uncovered before proceeding to the next epoch. Since each epoch covers all the nodes with high probability, the expected cover time of this modified version of process R-RANK is $O(n^2)$. By Lemma 12, for any rank assignment, the expected cover time of process R-RANK is $O(n^2)$. □

Lemma 14. *Assume that v is the node covered in round i and let u be the node selected in round i. Then there is an i-active edge sequence σ terminating in edge (u, v) and such that* $\mathrm{duration}(\mathrm{dst}(\sigma)) = i$.

Proof. Observe that u belongs to $\Gamma_{\mathrm{in}}(v)$. Furthermore, if $\mathrm{parent}(v) = \mathrm{nil}$, then the singleton node sequence v is active with $\mathrm{duration}(v) = i$. Thus the singleton edge sequence $\sigma = (u, v)$ is i-active with $\mathrm{duration}(\mathrm{dst}(\sigma)) = i$.

We prove the claim by induction on i. For $i = 1$, we have $\mathrm{parent}(v) = \mathrm{nil}$ and so the claim follows by the observations of the previous paragraph.

For $i > 1$, if $\mathrm{parent}(v) = \mathrm{nil}$, the claim once again follows from the foregoing observations. Otherwise, $\mathrm{parent}(v) = v'$ where v' is the node covered in round j with $j < i$. Let u' denote the node selected in round j. Since $j < i$, we can inductively assume that there is a j-active edge sequence, call it τ, terminating in edge (u', v') and such that $\mathrm{duration}(\mathrm{dst}(\tau)) = j$. Since τ is active, the node sequence $\mathrm{dst}(\tau)$ is active and $\mathrm{select}(\mathrm{dst}(\tau)) = \mathrm{src}(\tau)$. Let $\sigma = \tau : (u, v)$. Thus $\mathrm{src}(\sigma) = \mathrm{src}(\tau) : u$ and $\mathrm{dst}(\sigma) = \mathrm{dst}(\tau) : v$. Since $\mathrm{parent}(v) = v'$ and $\mathrm{dst}(\tau)$ is an active node sequence terminating in node v', $\mathrm{dst}(\sigma)$ is active. Since $\mathrm{duration}(\mathrm{dst}(\tau)) = j$, $\mathrm{select}(\mathrm{dst}(\tau)) = \mathrm{src}(\tau)$, u was selected in round i, and i is the least integer greater that j such that the node selected in round i belongs to $\Gamma_{\mathrm{in}}(v)$, we have $\mathrm{duration}(\mathrm{dst}(\sigma)) = i$ and $\mathrm{select}(\mathrm{dst}(\sigma)) = \mathrm{src}(\sigma)$. Since $\mathrm{dst}(\sigma)$ is active and $\mathrm{select}(\mathrm{dst}(\sigma)) = \mathrm{src}(\sigma)$, σ is active. Since σ is active and v is the node covered in round i, σ is i-active. Thus the edge sequence σ satisfies all of the requirements of the lemma. □

Lemma 15. *Any active node sequence is rank-sorted, chronological, and consists of distinct nodes.*

Proof. Note that any chronological node sequence consists of distinct nodes. Thus, in what follows, it is sufficient to prove that any active node sequence is rank-sorted and chronological.

We proceed by induction on the length of the sequence. For the base case, note that any node sequence of length 0 or 1 is rank-sorted and chronological. For the induction step, consider an active node sequence σ of the form $\tau : v : v'$. Since σ is active, $\tau : v$ is active and $\mathrm{parent}(v') = v$. Since $\tau : v$ is active, the induction hypothesis implies that it is also rank-sorted and chronological. Since $\mathrm{parent}(v') = v$, $\mathrm{rank}(v) \leq \mathrm{rank}(v')$ and v is covered before v'. Hence σ is rank-sorted and chronological. □

Lemma 16. *For any nonempty active edge sequence σ, if the last edge in σ is (u, v), then v is the node covered in round* $\mathrm{duration}(\mathrm{dst}(\sigma))$ *and node u is selected in the same round.*

Proof. We prove the claim by induction on the length of the active edge sequence σ.

If σ consists of a single edge (u, v), then by the definition of an active edge sequence, the singleton node sequence $\mathrm{dst}(\sigma)$ is active and $\mathrm{select}(\mathrm{dst}(\sigma)) = \mathrm{src}(\sigma)$. Since $\mathrm{dst}(\sigma)$ is active, $\mathrm{parent}(v) = \mathrm{nil}$, that is, v is the node covered in the first round in which a node in $\Gamma_{\mathrm{in}}(v)$ is selected, which is round $\mathrm{duration}(\mathrm{dst}(\sigma))$. Since $\mathrm{select}(\mathrm{dst}(\sigma)) = \mathrm{src}(\sigma)$, node u is selected in the same round.

Now assume that σ is an active edge sequence of the form $\tau : (u, v)$, where τ is of the form $\tau' : (u', v')$. Since σ is active, the node sequence $\mathrm{dst}(\sigma)$ is active and $\mathrm{select}(\mathrm{dst}(\sigma)) = \mathrm{src}(\sigma)$. It follows that $\mathrm{dst}(\tau)$ is active and $\mathrm{select}(\mathrm{dst}(\tau)) = \mathrm{src}(\tau)$, that is, τ is also active. Since τ is active and shorter than σ, we can inductively assume that v' is the node covered in round $\mathrm{duration}(\mathrm{dst}(\tau))$ and node u' is selected in the same round. Since $\mathrm{dst}(\sigma)$ is active, $\mathrm{parent}(v) = v'$, that is, v is the node covered in the first round after round $\mathrm{duration}(\mathrm{dst}(\tau))$ in which a node in $\Gamma_{\mathrm{in}}(v)$ is selected. Applying the definition of $\mathrm{duration}(\mathrm{dst}(\sigma))$, we conclude that v is the node covered in round $\mathrm{duration}(\mathrm{dst}(\sigma))$. Since $\mathrm{select}(\mathrm{dst}(\sigma)) = \mathrm{src}(\sigma)$, node u is selected in the same round. □

Lemma 17. *If σ is an active sequence of edges, then σ is linked.*

Proof. We proceed by induction on the length of σ. If the length of σ is 0 or 1, then σ is linked by definition.

Now assume that σ is an edge sequence of the form $\tau : (u, v)$, where τ is of the form $\tau' : (u', v')$ and σ is active. Since σ is active, $\mathrm{dst}(\sigma)$ is active. Since $\mathrm{dst}(\sigma)$ is active, $\mathrm{dst}(\tau)$ is also active. Since $\mathrm{dst}(\tau)$ is active and τ is shorter than σ, we can inductively assume that τ is linked. Therefore, in order to establish that σ is linked, it is sufficient to prove that (u', v) is an edge. Since $\mathrm{dst}(\sigma)$ is active, $\mathrm{parent}(v) = v'$. Hence, letting i denote the round in which node v is covered, we find that v' is the node covered in the first round prior to round i in which the selected node belongs to $\Gamma_{\mathrm{in}}(v)$. By Lemma 16, v' is covered in a round in which node u' is selected. Thus u' belongs to $\Gamma_{\mathrm{in}}(v)$, that is, (u', v) is an edge, as required. □

Lemma 18. *If an edge sequence σ is i-active, then* $\mathrm{duration}(\mathrm{dst}(\sigma)) = i$.

Proof. If σ is empty, then the claim holds since $i = 0$ and $\mathrm{duration}(\mathrm{dst}(\sigma)) = 0$. Otherwise, σ is of the form $\tau : (u, v)$, and by the definition of an i-active edge sequence, v is the node covered in round i. By Lemma 16, v is the node covered in round $\mathrm{duration}(\mathrm{dst}(\sigma))$, so $\mathrm{duration}(\mathrm{dst}(\sigma)) = i$. □

Lemma 19. *For any ℓ-sequence of edges σ, and any nonnegative integer i, the probability that σ is i-active is at most $\binom{\ell+r-1}{\ell} \cdot \Pr(X \geq i) \cdot (r\delta_{\mathrm{in}})^{-\ell}$, where $X \sim \mathrm{NegBin}\left(\ell, \frac{\delta_{\mathrm{in}}}{n}\right)$.*

Proof. If the nodes in $\mathrm{dst}(\sigma)$ are not all distinct, then $\Pr(\sigma \text{ is } i\text{-active}) = 0$ by Lemma 15 and the claimed inequality holds since the right-hand side is nonnegative.

Now assume that $\mathrm{dst}(\sigma)$ consists of distinct nodes, and let events A, B, and C be as defined in the statement of Lemma 9. Below we prove that if σ is i-active, then events A, B, and C all occur. The claimed inequality then follows by Lemma 10.

Assume that σ is i-active. Thus event B occurs by Lemma 18. Furthermore, σ is active, so $\mathrm{dst}(\sigma)$ is active and event C occurs by the definition of an active edge sequence. Since $\mathrm{dst}(\sigma)$ is active, event A occurs by Lemma 15. □

Lemma 20. *For any nonnegative integers i and ℓ, the probability that some ℓ-sequence of edges is i-active is at most*

$$n\Delta_{\mathrm{out}}^{\ell}\Delta_{\mathrm{in}}^{\ell-1}\binom{\ell+r-1}{\ell}\frac{\Pr(X\geq i)}{(r\delta_{\mathrm{in}})^{\ell}}$$

where $X \sim \mathrm{NegBin}\left(\ell, \frac{\delta_{\mathrm{in}}}{n}\right)$.

Proof. By Lemma 17, if an edge sequence σ is not linked, then $\Pr(\sigma \text{ is } i\text{-active}) = 0$. A union bound then implies that the probability some ℓ-sequence of edges is i-active is at most the number of linked ℓ-sequences of edges multiplied by the maximum probability that any particular ℓ-sequence is i-active. The desired inequality then follows by Lemmas 11 and 19. □

Lemma 21. *For nonnegative integers i, ℓ, and r such that $i \geq 64n\max(\Delta_{\mathrm{out}}\Delta_{\mathrm{in}}/\delta_{\mathrm{in}}^2, (\ln n)/\delta_{\mathrm{in}})$ and $r \geq \min(\lceil 2e^2\Delta_{\mathrm{out}}\Delta_{\mathrm{in}}/\delta_{\mathrm{in}}\rceil, \ell)$, we have*

$$\Delta_{\mathrm{out}}^{\ell}\Delta_{\mathrm{in}}^{\ell-1}\binom{\ell+r-1}{\ell}\frac{\Pr(X\geq i)}{(r\delta_{\mathrm{in}})^{\ell}} \leq \exp(-i\delta_{\mathrm{in}}/(32n))$$

where $X \sim \mathrm{NegBin}\left(\ell, \frac{\delta_{\mathrm{in}}}{n}\right)$.

Proof. First, we show that the LHS of the claimed inequality is a nonincreasing function of r.

It is sufficient to prove that the expression $\binom{\ell+r-1}{\ell}r^{-\ell}$ is a nonincreasing function of r. Fix ℓ and let $f(r)$ denote the preceding expression. Note that

$$\begin{aligned}\frac{f(r+1)}{f(r)} &= \frac{r+\ell}{r}\left(\frac{r}{r+1}\right)^{\ell}\\ &= \left(1+\frac{\ell}{r}\right)\left(1+\frac{1}{r}\right)^{-\ell}\\ &\leq 1,\end{aligned}$$

where the last inequality holds since the binomial theorem implies $(1+\frac{1}{r})^{\ell} \geq 1+\frac{\ell}{r}$.

Since we have established that the LHS of the claimed inequality is a nonincreasing function of r, we can assume in what follows that $r = \min(\lceil 2e^2\Delta_{\mathrm{out}}\Delta_{\mathrm{in}}/\delta_{\mathrm{in}}\rceil, \ell)$.

Let us rewrite the LHS of the claimed inequality as $\lambda \cdot \Pr(X \geq i)$, where

$$\begin{aligned}\lambda &= \Delta_{\mathrm{out}}^{\ell}\Delta_{\mathrm{in}}^{\ell-1}\binom{\ell+r-1}{\ell}(r\delta_{\mathrm{in}})^{-\ell}\\ &\leq \Delta_{\mathrm{out}}^{\ell}\Delta_{\mathrm{in}}^{\ell}\left(\frac{e(\ell+r-1)}{\ell r\delta_{\mathrm{in}}}\right)^{\ell}\\ &\leq \left(\frac{e\Delta_{\mathrm{out}}\Delta_{\mathrm{in}}(\ell+r)}{\ell r\delta_{\mathrm{in}}}\right)^{\ell}. \qquad (1)\end{aligned}$$

We begin by establishing two useful upper bounds on λ, namely, Equations (2) and (4) below.

If $r = \lceil 2e^2 \Delta_{\text{out}} \Delta_{\text{in}} / \delta_{\text{in}} \rceil$, then since since $r = \min(\lceil 2e^2 \Delta_{\text{out}} \Delta_{\text{in}} / \delta_{\text{in}} \rceil, \ell)$, we have $r \leq \ell$. Substituting the value of r into Equation (1), we find that

$$\begin{aligned} \lambda &\leq \left(\frac{e(\ell + r)}{2e^2 \ell} \right)^{\ell} \\ &\leq \left(\frac{2e\ell}{2e^2 \ell} \right)^{\ell} \\ &\leq e^{-\ell}. \end{aligned} \tag{2}$$

If $r = \ell$, then Equation (1) implies

$$\lambda \leq \left(\frac{2e \Delta_{\text{out}} \Delta_{\text{in}}}{\ell \delta_{\text{in}}} \right)^{\ell}. \tag{3}$$

Let $h(\ell)$ denote the natural logarithm of the RHS of Equation (3), that is, $h(\ell) = \ell \ln(2e \Delta_{\text{out}} \Delta_{\text{in}} / (\ell \delta_{\text{in}}))$. Using elementary calculus, it is straightforward to prove that the derivative of $h(\ell)$ with respect to ℓ is positive for $\ell < 2\Delta_{\text{out}} \Delta_{\text{in}} / \delta_{\text{in}}$, is 0 when $\ell = 2\Delta_{\text{out}} \Delta_{\text{in}} / \delta_{\text{in}}$, and is negative for $\ell > 2\Delta_{\text{out}} \Delta_{\text{in}} / \delta_{\text{in}}$. It follows that $h(\ell) \leq h(2\Delta_{\text{out}} \Delta_{\text{in}} / \delta_{\text{in}}) = 2\Delta_{\text{out}} \Delta_{\text{in}} / \delta_{\text{in}}$. Since $\ln$ is monotonic, the RHS of Equation (3) is also maximized when $\ell = 2\Delta_{\text{out}} \Delta_{\text{in}} / \delta_{\text{in}}$. Combining this result with Equation (2), we find that for any r

$$\lambda \leq \exp(2\Delta_{\text{out}} \Delta_{\text{in}} / \delta_{\text{in}}). \tag{4}$$

(Note that $\exp(2\Delta_{\text{out}} \Delta_{\text{in}} / \delta_{\text{in}}) \geq 1$ and Equation (2) implies $\lambda \leq 1$ when $r = \lceil 2e^2 \Delta_{\text{out}} \Delta_{\text{in}} / \delta_{\text{in}} \rceil$.)

We are now ready to proceed with the proof of the lemma. We consider the two cases $\ell > \lceil i\delta_{\text{in}}/(2n) \rceil$ and $\ell \leq \lceil i\delta_{\text{in}}/(2n) \rceil$ separately.

If $\ell > \lceil i\delta_{\text{in}}/(2n) \rceil$, then $\ell > 2ec \max(\Delta_{\text{out}} \Delta_{\text{in}} / \delta_{\text{in}}, \ln n)$ where $c = 16/e > e$. Thus $\ell > \lceil 2e^2 \Delta_{\text{out}} \Delta_{\text{in}} / \delta_{\text{in}} \rceil$ and so $r = \lceil 2e^2 \Delta_{\text{out}} \Delta_{\text{in}} / \delta_{\text{in}} \rceil$. It follows from Equation (2) that $\lambda \leq e^{-\ell} \leq \exp(-i\delta_{\text{in}}/(2n)) \leq \exp(-i\delta_{\text{in}}/(64n))$, and hence the claim holds since $\Pr(X \geq i) \leq 1$.

Now assume that $\ell \leq \lceil i\delta_{\text{in}}/(2n) \rceil$. Let $Y \sim \text{NegBin}\left(\lfloor \frac{i\delta_{\text{in}}}{2n} \rfloor, \frac{\delta_{\text{in}}}{n} \right)$ and $Z \sim \text{NegBin}\left(\lfloor \frac{i\delta_{\text{in}}}{2n} \rfloor - \ell, \frac{\delta_{\text{in}}}{n} \right)$. By the definition of the negative binomial distribution, $\Pr(Y \geq i) = \Pr(X + Z \geq i)$. And, since Z is nonnegative, $\Pr(X + Z \geq i) \geq \Pr(X \geq i)$. Thus

$$\Pr(X \geq i) \leq \Pr(Y \geq i). \tag{5}$$

Since $E[Y] \leq \frac{i}{2}$ and $\lfloor i\delta_{\text{in}}/(2n) \rfloor \geq \lfloor 32 \max(\Delta_{\text{out}} \Delta_{\text{in}} / \delta_{\text{in}}, \ln n) \rfloor > 2$, Lemma 3 implies $\Pr(Y \geq i) \leq \Pr(Y \geq 2E[Y]) \leq \exp\left(\frac{-i\delta_{\text{in}}}{16n} + \frac{1}{8} \right)$. The claim follows since

$$\begin{aligned}\lambda \cdot \Pr(X \geq i) &\leq \exp\left(\frac{2\Delta_{\text{out}}\Delta_{\text{in}}}{\delta_{\text{in}}}\right) \cdot \Pr(Y \geq i) \\ &\leq \exp\left(\frac{-i\delta_{\text{in}}}{16n} + \frac{1}{8} + \frac{2\Delta_{\text{out}}\Delta_{\text{in}}}{\delta_{\text{in}}}\right) \\ &\leq \exp\left(\frac{-i\delta_{\text{in}}}{32n} + \frac{1}{8}\right) \\ &\leq \exp\left(\frac{-i\delta_{\text{in}}}{64n}\right).\end{aligned}$$

(The first step follows from Equations (4) and (5). For the third step and fourth steps, note that the assumption $i \geq 64n \max(\Delta_{\text{out}}\Delta_{\text{in}}/\delta_{\text{in}}^2, (\ln n)/\delta_{\text{in}})$ implies $i\delta_{\text{in}}/(32n) \geq 2\Delta_{\text{out}}\Delta_{\text{in}}/\delta_{\text{in}}$ and $i\delta_{\text{in}}/(64n) \geq 1/8$, respectively.) □

Lemma 22. *If $r \geq \min(\lceil 2e^2\Delta_{\text{out}}\Delta_{\text{in}}/\delta_{\text{in}}\rceil, n)$, then every active edge sequence is, with high probability, $O(n \max(\Delta_{\text{out}}\Delta_{\text{in}}/\delta_{\text{in}}^2, (\log n)/\delta_{\text{in}}))$-active.*

Proof. Let c denote an arbitrary positive real greater than or equal to 1, and let i denote the positive integer $\lceil 64cn \max(\Delta_{\text{out}}\Delta_{\text{in}}/\delta_{\text{in}}^2, (\ln n)/\delta_{\text{in}}))\rceil$.

For any nonnegative integer j, let p_j denotes the probability that there is a j-active edge sequence. Any j-active edge sequence σ is active, so the associated node sequence $\text{dst}(\sigma)$ is active. It follows from Lemma 15 that any j-active sequence has length at most n. In other words, $\ell \leq n$ for any j-active ℓ-sequence of edges. Furthermore, if $j > 0$ then the length of a j-active sequence is nonzero. Since any j-active ℓ-sequence of edges satisfies $\ell \leq n$, the condition $r = \min(\lceil 2e^2\Delta_{\text{out}}\Delta_{\text{in}}/\delta_{\text{in}}\rceil, n)$ allows us to apply Lemmas 20 and 21. Applying these two lemmas, together with a union bound, we obtain $p_j \leq n^2 \exp(-j\delta_{\text{in}}/(64n))$ for $j > i$.

Let p denote the probability that there is a j-active edge sequence for some $j \geq i$. By a union bound, $p \leq \sum_{j \geq i} p_j$. Using the upper bound on p_j derived in the preceding paragraph, we find that p is upper bounded by an infinite geometric sum with initial term $n^2 \exp(-i\delta_{\text{in}}/(64n))$ and ratio $\exp(-\delta_{\text{in}}/(64n))$. Thus

$$\begin{aligned}p &= O((n^3/\delta_{\text{in}}) \exp(-i\delta_{\text{in}}/(64n))) \\ &= O(n^3 \exp(-c \max(\Delta_{\text{out}}\Delta_{\text{in}}/\delta_{\text{in}}, \log n))) \\ &= O(n^{3-c}).\end{aligned}$$

By setting c to a sufficiently large positive constant, we can drive p below any desired inverse polynomial threshold. The claim of the lemma follows. □

Lemma 23. *If $r \geq \min(\lceil 2e^2\Delta_{\text{out}}\Delta_{\text{in}}/\delta_{\text{in}}\rceil, n)$, then the cover time of process R-RANK is, with high probability, $O(n \max(\Delta_{\text{out}}\Delta_{\text{in}}/\delta_{\text{in}}^2, (\log n)/\delta_{\text{in}}))$. The same asymptotic bound holds for the expected cover time.*

Proof. The high probability claim is immediate from Lemmas 14 and 22. The bound on the expected cover time then follows by Lemma 13. □

Theorem 1. *If both Δ_{in} and Δ_{out} are $O(\delta_{\text{in}})$, then there is an r in $O(\delta_{\text{in}})$ such that the cover time of process R-RANK is $O(n + \frac{n \log n}{\delta_{\text{in}}})$ with high probability. The same asymptotic bound holds for the expected cover time.*

Proof. Immediate from Lemma 23. □

The result of Theorem 1 matches the lower bound proved by Alon for process MIN and is thus optimal [2].

Note that as r tends to infinity, the behavior of process R-RANK converges to that of process P-RANK. Thus, the bounds of Theorem 1 also hold for process P-RANK.

References

1. M. Adler, E. Halperin, R. Karp, and V. Vazirani. A stochastic process on the hypercube with applications to peer-to-peer networks. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing*, pages 575–584, 2003.
2. N. Alon. Problems and results in extremal combinatorics, II. Manuscript, 2004.
3. N. Alon and J. H. Spencer. *The Probabilistic Method*. Wiley, New York, NY, 1991.
4. Nedialko B. Dimitrov and C. Greg Plaxton. Optimal cover time for a graph-based coupon collector process. Technical Report TR–05–01, Department of Computer Science, University of Texas at Austin, January 2005.
5. Stasys Jukna. *Extremal Combinatorics*, pages 224–225. Springer, 2001.
6. F. T. Leighton. *Introduction to Parallel Algorithms and Architectures: Arrays, Trees, and Hypercubes*, pages 547–556. Morgan-Kaufmann, San Mateo, CA, 1991.
7. R. Motwani and P. Raghavan. *Randomized Algorithms*. Cambridge University Press, Cambridge, UK, 1995.
8. A. G. Ranade. How to emulate shared memory. *Journal of Computer and System Sciences*, 42:307–326, 1991.