

FrogWild! – Fast PageRank Approximations on Graph Engines

Ioannis Mitliagkas

ECE, UT Austin

ioannis@utexas.edu

Alexandros G. Dimakis

ECE, UT Austin

dimakis@austin.utexas.edu

Michael Borokhovich

ECE, UT Austin

michaelbor@utexas.edu

Constantine Caramanis

ECE, UT Austin

constantine@utexas.edu

ABSTRACT

We propose FROGWILD, a novel algorithm for fast approximation of high PageRank vertices, geared towards reducing network costs of running traditional PageRank algorithms. Our algorithm can be seen as a quantized version of power iteration that performs multiple parallel random walks over a directed graph. One important innovation is that we introduce a modification to the GraphLab framework that only partially synchronizes mirror vertices. This partial synchronization vastly reduces the network traffic generated by traditional PageRank algorithms, thus greatly reducing the per-iteration cost of PageRank. On the other hand, this partial synchronization also creates dependencies between the random walks used to estimate PageRank. Our main theoretical innovation is the analysis of the correlations introduced by this partial synchronization process and a bound establishing that our approximation is close to the true PageRank vector.

We implement our algorithm in GraphLab and compare it against the default PageRank implementation. We show that our algorithm is very fast, performing each iteration in less than one second on the Twitter graph and can be up to $7\times$ faster compared to the standard GraphLab PageRank implementation.

1. INTRODUCTION

Large-scale graph processing is becoming increasingly important for the analysis of data from social networks, web pages, bioinformatics and recommendation systems. Graph algorithms are difficult to implement in distributed computation frameworks like Hadoop MapReduce and Spark. For this reason several in-memory graph engines like Pregel, Giraph, GraphLab and GraphX [16, 15, 23, 21] are being developed. There is no full consensus on the fundamental abstractions of graph processing frameworks but certain pat-

terns such as vertex programming and the Bulk Synchronous Parallel (BSP) framework seem to be increasingly popular.

PageRank computation [18], which gives an estimate of the importance of each vertex in the graph, is a core component of many search routines; more generally, it represents, de facto, one of the canonical tasks performed using such graph processing frameworks. Indeed, while important in its own right, it also represents the memory, computation and communication challenges to be overcome in large scale iterative graph algorithms.

In this paper we propose a novel algorithm for fast *approximate calculation* of high PageRank vertices. Note that even though most previous works calculate the complete PageRank vector (of length in the millions or billions), in many graph analytics scenarios a user wants a quick estimation of the most important or relevant nodes – distinguishing the 10^{th} most relevant node from the $1,000^{th}$ most relevant is important; the $1,000,000^{th}$ from the $1,001,000^{th}$ much less so. A simple solution is to run the standard PageRank algorithm for fewer iterations (or with an increased tolerance). While certainly incurring less overall cost, the per-iteration cost remains the same; more generally, the question remains whether there is a more efficient way to approximately recover the heaviest PageRank vertices.

In this paper we address this problem. Our algorithm (called FROGWILD for reasons that will become subsequently apparent) significantly outperforms the simple reduced iterations heuristic in terms of *running time*, *network communication* and *scalability*. We note that, naturally, we compare our algorithm and reduced-iteration-PageRank within the same framework: we implemented our algorithm in GraphLab PowerGraph and compare it against the built-in PageRank implementation. A key part of our contribution also involves the proposal of what appears to be simply a technically minor modification within the GraphLab framework, but nevertheless results in significant network-traffic savings, and we believe may nevertheless be of more general interest beyond PageRank computations.

Contributions: We consider the problem of fast and efficient (in the sense of time, computation and communication costs) computation of the high PageRank nodes, using a graph engine. To accomplish this we propose and analyze a new PageRank algorithm specifically designed for the graph engine framework, and, significantly, we propose a modification of the standard primitives of the graph engine framework (specifically, GraphLab PowerGraph), that

enables significant network savings. We explain in further detail both our objectives, and our innovations.

Rather than seek to recover the full PageRank vector, we aim for the top k PageRank vertices (where k is considered to be approximately in the order of $10 - 1000$). Given an output of a list of k vertices, we define two natural accuracy metrics that compare the true top- k list with our output. The algorithm we propose, FROGWILD operates by starting a small (sublinear in the number of vertices n) number of random walkers (*frogs*) that jump randomly on the directed graph. The random walk interpretation of PageRank enables the frogs to jump to a completely random vertex (teleport) with some constant probability (set to 0.15 in our experiments, following standard convention). After we allow the frogs to jump for time equal to the mixing time of this non-reversible Markov chain, their positions are sampled from the invariant distribution π which is normalized PageRank. The standard PageRank iteration can be seen as the continuous limit of this process (*i.e.*, the frogs become water), which is equivalent to power iteration for stochastic matrices.

The main *algorithmic contributions* of this paper are comprised of the following three innovations. First, we argue that discrete frogs (a quantized form of power iteration) is significantly better for distributed computation when one is interested only in the large entries of the eigenvector π . This is because each frog produces an independent sample from π . If some entries of π are substantially larger and we only want to determine those, a small number of independent samples suffices. We make this formal using standard Chernoff bounds (see also [20, 9] for similar arguments). On the contrary, during standard PageRank iterations, vertices pass messages to all their out-neighbors since a non-zero amount of water must be transferred. This tremendously increases the network bandwidth especially when the graph engine is over a cluster with many machines.

One major issue with simulating discrete frogs in a graph engine is teleportations. Graph frameworks partition vertices to physical nodes and restrict communication on the edges of the underlying graph. Global random jumps would create dense messaging patterns that would increase communication. Our second innovation is a way of obtaining an identical sampling behavior without teleportations. We achieve this by initiating the frogs at uniformly random positions and having them perform random walks for a life span that follows a geometric random variable. The geometric probability distribution depends on the teleportation probability and can be calculated explicitly.

Our third innovation involves a simple proposed modification for graph frameworks. Most modern graph engines (like GraphLab PowerGraph [11]) employ vertex-cuts as opposed to edge-cuts. This means that each vertex of the graph is assigned to multiple machines so that graph edges see a local vertex mirror. One copy is assigned to be the master and maintains the master version of vertex data while remaining replicas are mirrors that maintain local cached read-only copies of the data. Changes to the vertex data are made to the master and then replicated to all mirrors at the next synchronization barrier. This architecture is highly suitable for graphs with high-degree vertices (as most real-world graphs are) but has one limitation when used for a few random walks: imagine that vertex v_1 contains one frog that wants to jump to v_2 . If vertex v_1 has very high degree, it is very

likely that multiple replicas of that vertex exist, possibly one in each machine in the cluster. In an edge-cut scenario only one message would travel from $v_1 \rightarrow v_2$, assuming v_1 and v_2 are located in different physical nodes. However, when vertex-cuts are used, the state of v_1 is updated (*i.e.*, contains no frogs now) and this needs to be communicated to all mirrors. It is therefore possible that *a single random walk can create a number of messages equal to the number of machines in the cluster*.

We modify PowerGraph to expose a scalar parameter p_s per vertex. By default, when the framework is running, in each super-step all masters synchronize their programs and vertex data with their mirrors. Our modification is that for each mirror we flip an independent coin and synchronize with probability p_s . Note that when the master does not synchronize the vertex program with a replica, that replica will not be active during that super-step. Therefore, we can avoid the communication and CPU execution by performing limited synchronization in a randomized way.

FROGWILD is therefore executed asynchronously but relies on the Bulk Synchronous execution mode of PowerGraph with the additional simple randomization we explained. The name of our algorithm is inspired by HogWild [19], a lock-free asynchronous stochastic gradient descent algorithm proposed by Niu *et al.*. We note that PowerGraph does support an asynchronous execution mode [11] but we implemented our algorithm by a small modification of synchronous execution. As discussed in [11], the design of asynchronous graph algorithms is highly nontrivial and involves locking protocols and other complications. Our suggestion is that for the specific problem of simulating multiple random walks on a graph, simply randomizing synchronization can give significant benefits while keeping design simple.

While the parameter p_s clearly has the power to significantly reduce network traffic – and indeed, this is precisely born out by our empirical results – it comes at a cost: the standard analysis of the Power Method iteration no longer applies. The main challenge that arises is the theoretical analysis of the FROGWILD algorithm. The model is that each vertex is separated across machines and each connection between two vertex copies is present with probability p_s . A single frog performing a random walk on this new graph defines a new Markov Chain and this can be easily designed to have the same invariant distribution π equal to normalized PageRank. The complication is that the trajectories of frogs are *no longer independent*: if two frogs are in vertex v_1 and (say) only one mirror v'_1 synchronizes, both frogs will need to jump through edges connected with that particular mirror. Worse still, this correlation effect increases, the more we seek to improve network traffic by further decreasing p_s . Therefore, it is no longer true that one obtains independent samples from the invariant distribution π . Our theoretical contribution is the development of an analytical bound that shows that these dependent random walks still can be used to obtain $\hat{\pi}$ that is provably close to π with high probability. We rely on a coupling argument combined with an analysis of pairwise intersection probabilities for random walks on graphs. In our convergence analysis we use the *contrast bound* [7] for non-reversible chains.

1.1 Notation

Lowercase letters denote scalars or vectors. Uppercase letters denote matrices. The (i, j) element of a matrix A

is A_{ij} . We denote the conjugate transpose of a matrix A by A^\dagger and the transpose by A' . For a time-varying vector x , we denote its value at time t by x^t . When not otherwise specified, $\|x\|$ denotes the l_2 -norm of vector x . We use Δ^{n-1} for the probability simplex in n dimensions, and $e_i \in \Delta^{n-1}$ for the indicator vector for item i . For example, $e_1 = [1, 0, \dots, 0]$. For the set of all integers from 1 to n we write $[n]$.

2. PROBLEM AND MAIN RESULTS

We now make precise the intuition and outline given in the introduction. We first define the problem, giving the definition of PageRank, the PageRank vector, and therefore its top elements. We then define the algorithm, and finally state our main analytical results.

2.1 Problem Formulation

Consider a directed graph $G = (V, E)$ with n vertices ($|V| = n$) and let A denote its adjacency matrix. That is, $A_{ij} = 1$ if there is an edge from j to i . Otherwise, the value is 0. Let $d_{\text{out}}(j)$ denote the number of *successors* (out-degree) of vertex j in the graph. We assume that all nodes have at least one successor, $d_{\text{out}}(j) > 0$. Then we can define the transition probability matrix P as follows:

$$P_{ij} = A_{ij}/d_{\text{out}}(j). \quad (1)$$

The matrix is left-stochastic, which means that each of its rows sums to 1. We call $G(V, E)$ the *original graph*, as opposed to the PageRank graph which includes a probability of transitioning to any given vertex. We now define this transition probability matrix, and the PageRank vector.

DEFINITION 1 (PAGERANK [18]). *Consider the matrix*

$$Q \triangleq (1 - p_T)P + p_T \frac{1}{n} \mathbf{1}_{n \times n}.$$

where $p_T \in [0, 1]$ is a parameter, most commonly set to 0.15. The PageRank vector $\pi \in \Delta^{n-1}$ is defined as the principal right eigenvector of Q . That is, $\pi \triangleq v_1(Q)$. By the Perron-Frobenius theorem, the corresponding eigenvalue is 1. This implies the fixed-point characterization of the PageRank vector, $\pi = Q\pi$.

The PageRank vector assigns high values to *important* nodes. Intuitively, important nodes have many important predecessors (other nodes that point to them). This recursive definition is what makes PageRank robust to manipulation, but also expensive to compute. It can be recovered by exact eigendecomposition of Q , but at real problem scales this is prohibitively expensive. In practice, engineers often use a few iterations of the power method to get a "good-enough" approximation.

The definition of PageRank hinges on the left-stochastic matrix Q , suggesting a connection to Markov chains. Indeed, this connection is well documented and studied [1, 10]. An important property of PageRank from its random walk characterization, is the fact that π is the invariant distribution for a Markov chain with dynamics described by Q . A non-zero p_T , also called the *teleportation probability*, introduces a uniform component to the PageRank vector π . We see in our analysis that this implies ergodicity and faster mixing for the random walk.

2.1.1 Top PageRank Elements

Given the true k vertices of highest PageRank (breaking ties randomly for simplicity) and the output list of k vertices by an approximate PageRank algorithm we define accuracy using two metrics.

DEFINITION 2 (MASS CAPTURED). *Given a distribution $v \in \Delta^{n-1}$, the true PageRank distribution $\pi \in \Delta^{n-1}$ and an integer $k \geq 0$, we define the mass captured by v as follows.*

$$\mu_k(v) \triangleq \pi(\arg\max_{|S|=k} v(S))$$

For a set $S \in [n]$, $v(S) = \sum_{i \in S} v(i)$ denotes the total mass ascribed to the set by distribution $v \in \Delta^{n-1}$.

Put simply, the set S^* that gets the most mass according to v out of all sets of size k , is evaluated according to π and that gives us our metric. It is maximized by π itself, i.e. the optimal value is $\mu_k(\pi)$.

The second metric we use is the exact identification probability, i.e. the fraction of elements in the output list that are also in the true top- k list. Note that the second metric is limited in that it does not give partial credit for high PageRank vertices that were not in the top- k list. In our experiments in Section 4, we mostly use the normalized captured mass accuracy metric but also report the exact identification probability for some cases – typically the results are similar.

We subsequently describe our algorithm. We attempt to approximate the heaviest elements of the invariant distribution of a Markov Chain, by simultaneously performing multiple random walks on the graph. The main modification to PowerGraph, is the exposure of a parameter, p_s , that controls the probability that a given master node synchronizes with any one of its mirrors. Per step, this leads to a proportional reduction in network traffic. The main contribution of this paper is to show that we get results of comparable or improved accuracy, while maintaining this network traffic advantage. We demonstrate this empirically in Section 4.

2.2 Algorithm

During setup, the graph is partitioned using GraphLab's default ingress algorithm. At this point each one of N frogs is born on a vertex chosen uniformly at random. Each vertex i carries a counter initially set to 0 and denoted by $c(i)$. Scheduled vertices execute the following program.

Incoming frogs from previously executed vertex programs, are collected by the `init()` function. At `apply()` every frog dies with probability $p_T = 0.15$. This, along with a uniform starting position, effectively simulates the 15% uniform component from Definition 1.

A crucial part of our algorithm is the change in synchronization behaviour. The `<sync>` step only synchronizes a p_s fraction of mirrors leading to commensurate gains in network traffic (cf. Section 4). This patch on the GraphLab codebase was only a few lines of code. Section 4 contains more details regarding the implementation.

The `scatter()` phase is only executed for edges e incident to a mirror of i that has been synchronized. Those edges draw a binomial number of frogs to send to their other endpoint. The rest of the edges perform no computation. The frogs sent to vertex j at the last step will be collected at the `init()` step when j executes.

FrogWild! vertex program

Input parameters: $p_s, p_T = 0.15$

apply(i) $K(i) \leftarrow [\# \text{ incoming frogs}]$

For every incoming frog:

With probability p_T , frog dies:

$c(i) = c(i) + 1$,

$K(i) = K(i) - 1$.

<sync> For every *mirror* m of vertex i :

With probability p_s :

Synchronize state with mirror m .

scatter($e = (i, j)$) [Only on synchronized mirrors]

Generate Binomial number of frogs:

$$x \sim \text{Bin}\left(K(i), \frac{1}{d_{\text{out}}(i)p_s}\right)$$

Send x frogs to vertex j : **signal**(j, x)

Parameter p_T is the teleportation probability from the random surfer model in [18]. To get PageRank using random walks, one could adjust the transition matrix P as described in Definition 1 to get the matrix Q . Alternatively, the process can be replicated by a random walk following the original matrix P , and teleporting at every time, with probability p_T . The destination for this teleportation is chosen uniformly at random from $[n]$. We are interested in the position of a walk at a predetermined point in time as that would give us a sample from π . This holds as long as we allow enough time for mixing to occur.

Due to the inherent markovianity in this process, one could just consider it starting from the last teleportation before the predetermined stopping time. When the mixing time is large enough, the number of steps performed between the last teleportation and the predetermined stopping time, denoted by X , is geometrically distributed with parameter p_T . This follows from the time-reversibility in the teleportation process: inter-teleportation times are geometrically distributed, so as long as the first teleportation event happens before the stopping time, then $X \sim \text{Geom}(p_T)$.

This establishes that, the FROGWILD! process – where a frog performs a geometrically distributed number of steps following the original transition matrix P – closely mimics a random walk that follows the adjusted transition matrix, Q . Hence, to show our main result, Theorem 1, we analyze the latter process.

Using a binomial distribution to independently generate the number of frogs in the **scatter**() phase closely models the effect of random walks. The marginal distributions are correct, and the number of frogs, that did not die during the **apply**() step, is preserved in expectation. For our implementation we resort to a more efficient approach. Assuming $K(i)$ frogs survived the **apply**() step, and M mirrors where picked for synchronization, then we send $\lceil \frac{K(i)}{M} \rceil$ frogs to $\min(K(i), M)$ mirrors. If the number of available frogs is less than the number of synchronized mirrors, we pick $K(i)$ arbitrarily.

2.3 Main Result

Our analytical results essentially provide a high probability guarantee that our algorithm produces a solution that

approximates well the PageRank vector. Recall that our main modification of our algorithm involves limiting the synchronization of master nodes and its mirrors. In theory we introduce the following broad model to deal with partial synchronization.

DEFINITION 3 (EDGE ERASURE MODEL). *An edge erasure model is a process that is independent from the random walks (up to time t) and temporarily erases a subset of all edges at time t . The event $E_{i,j}^t$ represents the erasure of edge (i, j) from the graph for time t . The edge is not permanently removed from the graph, it is just disabled and considered again in the next step. The edge erasure models we study satisfy the following properties.*

1. Edges are erased independently for different vertices,

$$\mathbb{P}(E_{i,j}^t, E_{i,k}^t) = \mathbb{P}(E_{i,j}^t)\mathbb{P}(E_{i,k}^t)$$

and across time,

$$\mathbb{P}(E_{i,j}^t, E_{i,j}^s) = \mathbb{P}(E_{i,j}^t)\mathbb{P}(E_{i,j}^s).$$

2. Each outgoing edge is preserved (not erased) with probability at least p_s .

$$\mathbb{P}(\overline{E_{i,j}^t}) \geq p_s$$

3. Erasures do not exhibit significant negative correlation. Specifically,

$$\mathbb{P}(\overline{E_{i,j}^t} | \overline{E_{i,k}^t}) \geq p_s.$$

4. Erasures in a neighbourhood are symmetric. Any subset of out-going edges of vertex i , will be erased with exactly the same probability as another subset of the same cardinality.

The main two edge erasure models we consider are described here. They both satisfy all required properties. Our theory holds for both¹, but in our implementation and experiments we use "At Least One Out-Edge Per Node."

EXAMPLE 4 (INDEPENDENT ERASURES). *Every edge is preserved independently with probability p_s .*

EXAMPLE 5 (AT LEAST ONE OUT-EDGE PER NODE). *This edge erasure model, decides all erasures for node i independently, like Independent Erasures, but if all out-going edges for node i are erased, it draws and enables one of them uniformly at random.*

Our results tell us that partial synchronization does not change the distribution of a single random walk. To make this and our other results clear, we need the simple definition.

DEFINITION 6. *We denote the state of random walk i at its t^{th} step by s_i^t .*

Then, we see that $\mathbb{P}(s_1^{t+1} = i | s_1^t = j) = 1/d_{\text{out}}(j)$, and $x_1^{t+1} = Px_1^t$. This follows simply by the symmetry assumed in Definition 3. Thus if we were to sample in serial, the modification of the algorithm controlling (limiting) synchronization would not affect each sample, and hence would not

¹*Independent Erasures* can lose some walkers, when it temporarily leads to some nodes having zero out-degree.

affect our estimate of the invariant distribution. However, we start multiple (all) random walks simultaneously. In this setting, the fundamental analytical challenge stems from the fact that any set of random walks with intersection are now correlated. The key to our result is that we can control the effect of this correlation, as a function the parameter p_s and the *pairwise probability* that two random walks intersect. We define this formally.

DEFINITION 7. Suppose two walkers l_1 and l_2 perform t steps under the edge erasure model. The probability that they meet is defined as follows.

$$p_{\cap}(t) \triangleq \mathbb{P}(\exists \tau \in [0, t], \text{ s.t. } s_{l_1}^{\tau} = s_{l_2}^{\tau}) \quad (2)$$

DEFINITION 8 (ESTIMATOR). Given the positions of N random walks at time t , $\{s_i^t\}_{i=1}^N$, we define the following estimator for the invariant distribution π .

$$\hat{\pi}_N(i) \triangleq \frac{|\{l : l \in [N], s_l^t = i\}|}{N} = \frac{c(i)}{N} \quad (3)$$

Here $c(i)$ refers to the tally maintained by the FROGWILD! vertex program.

Now we can state the main result. Here we give a guarantee for the quality of the solution furnished by our algorithm.

THEOREM 1 (MAIN THEOREM). Consider N frogs following the adjusted transition matrix Q of Definition 1, under the erasure model of Definition 3. The frogs start at independent locations, distributed uniformly and stop after t steps. The estimator $\hat{\pi}_N$ (Definition 8), captures mass close to the optimal. Specifically, with probability at least $1 - \delta$,

$$\mu_k(\hat{\pi}_N) \geq \mu_k(\pi) - \epsilon,$$

where

$$\epsilon < 2C\sqrt{k\|\pi\|_{\infty}}[(1 - p_T)\lambda_2]^t + \sqrt{\frac{k}{\delta} \left[\frac{1}{N} + (1 - p_s^2)p_{\cap}(t) \right]}, \quad (4)$$

for some constant $C > 0$ and $\lambda_2 < 1$.

The proof is deferred to Appendix A.2. The guaranteed accuracy via this result depends on the probability that two walkers will intersect. Via a simple argument, that probability is the same as the meeting probability for independent walks. The next theorem calculates this probability.

THEOREM 2 (INTERSECTION PROBABILITY). Consider two independent random walks obeying the same ergodic transition probability matrix, Q with invariant distribution π , as described in Definition 1. Furthermore, assume that both of them are initially distributed uniformly over the state space of size n . The probability that they meet within t steps, is bounded as follows,

$$p_{\cap}(t) \leq \frac{1}{n} + \frac{t\|\pi\|_{\infty}}{p_T},$$

where $\|\pi\|_{\infty}$, denotes the maximal element of the vector π .

The proof is based on the observation that the l_{∞} norm of a distribution controls the probability that two independent samples coincide. We show that for all steps of the random walk, that norm is controlled by the l_{∞} norm of π . We defer the full proof to Appendix A.3.

A number of studies, give experimental evidence (e.g. [4]) suggesting that PageRank values for the web graph follow a power-law distribution with parameter approximately $\theta = 2.2$. That is true for the tail of the distribution – the largest values, hence of interest to us here – regardless of the choice of p_T . The following proposition bounds the value of the heaviest PageRank value, $\|\pi\|_{\infty}$.

PROPOSITION 9 (MAX OF POWER-LAW DISTRIBUTION). Let $\pi \in \Delta^{n-1}$ follow a power-law distribution with parameter θ and minimum value p_T/n . Its maximum element, $\|\pi\|_{\infty}$, is at most $n^{-\gamma}$, with probability at least $1 - cn^{\gamma - \frac{1}{\theta-1}}$, for some universal constant c .

Assuming $\theta = 2.2$ and picking, for example, $\gamma = 0.5$, we get that

$$\mathbb{P}(\|\pi\|_{\infty} > 1/\sqrt{n}) \leq cn^{-1/3}.$$

This implies that with probability at least $1 - cn^{-1/3}$ the meeting probability is bounded as follows.

$$p_{\cap}(t) \leq \frac{1}{n} + \frac{t}{p_T\sqrt{n}}.$$

One would usually take a number of steps t that are either constant or logarithmic with respect to the graph size n . This implies that for many reasonable choices of set size k and acceptable probability of failure δ , the meeting probability vanishes as n grows. Then we can make the second term of the error in (4) arbitrarily small by controlling the number of frogs, N . The proof for Proposition 9 is deferred to Appendix A.4.

2.4 Prior Work

There is a very large body of work on computing and approximating PageRank on different computation models (e.g. see [5, 8, 20, 9, 3] and references therein). To the best of our knowledge, our work is the first to specifically design an approximation algorithm for high-PageRank nodes for graph engines. Another line of work looks for *Personalized PageRank* (PPR) scores. This quantifies the influence an arbitrary node i has on another node j , cf. recent work [14] and discussion therein. PPR is not applicable in our case, as we are looking for an answer close to a *global* optimum. Our base-line comparisons come from the graph framework papers since PageRank is a standard benchmark for running-time, network and other computations. Our implementation is on GraphLab (PowerGraph) and significantly outperforms the built-in PageRank algorithm. This algorithm is already shown in [11, 21] to be significantly more efficient compared to other frameworks like Hadoop, Spark, Giraph etc.

3. NEW ANALYSIS

In this section we provide a complete proof of our main results. The proof goes through a sequence of related random processes proving equivalence or quantifying the error introduced at every hop. We describe these processes as they come up in the proof; an overview in tabular form in Table 1. All processes are random walks, but the transition probability matrix (TPM), independence of different walkers (frogs), number of steps before sampling and number of samples vary. We start from an exact but impractical estimator and slowly move towards FROGWILD!, getting guarantees for our algorithm as described in Section 2.2.

PROCESS 10 (EXACT). *This process consists of independent random walks on the vertices of the graph following the augmented transition probability matrix Q , as given in Definition 1. We let the number of steps taken before sampling and the number of samples taken tend to infinity.*

This process is not implementable but we use it as our starting point. The following, informally stated, lemma tells us that it is exact in the limit.

LEMMA 11. *Using the samples from Process 10 and the estimator of Definition 8, gives us the PageRank vector π exactly.*

PROOF. A single random walk with TPM Q is ergodic and π is its unique invariant distribution. Using basic consistency results for random walks (cf. [7]), we get that every sample acquired has distribution that tends to π . Then the strong law of large numbers implies that the estimate tends to π . \square

The implementation of Process 10 is not possible for it uses an infinite number of steps and samples. Furthermore, the transition probability matrix Q dictates that a walker performs a random global jump (teleportation) with probability p_T at every step. This last impracticality is the first one that we lift.

PROCESS 12 (GEOMETRIC). *Independent walkers start on a node selected uniformly at random and perform random walks on the original graph. This means that there is no teleportation and the walk is described by the TPM P as defined in Section 2.1. Each walker performs a random number of steps before yielding a sample. Specifically, the number of steps follows a geometric distribution with parameter p_T . The number of walkers tends to infinity.*

LEMMA 13. *Using the samples from Process 12, the estimator in Definition 8 converges almost surely to the exact PageRank vector π as the number of walkers tends to infinity as long as $p_T > 0$.*

PROOF. First we show that from each walker, we get a sample distributed according to π . To start, we note that a walk according to the augmented TPM Q behaves like a walk according to P , with the exception that at every time step, with probability p_T , the walker will teleport anywhere on the graph uniformly. Consider a walk from starting distribution $v \in \Delta^{n-1}$ evolving according to Q for T steps, let $\{Z_i\}_{i=1}^J$ denote the sequence of those teleportation times and set $Z_0 = 0$. Here J denotes the number of teleportations that happened up to and including time T , when the walk stops; in the case $J = 0$, the walk finishes after T steps without any teleportations. The teleportation times mark different epochs: the walk is strictly local between two teleportations, evolving according to P . Since this is a Markov process, we can just study the last epoch, i.e. $t \in [Z_J, T]$.

In the case when no teleportation happens, i.e. $J = 0$, then the last epoch is just $[0, T]$. This happens with probability $(1 - p_T)^T$. Otherwise, we have at least one teleportation, and from the memory-less property of geometric random variables, the duration of the last epoch is geometrically distributed with parameter p_T . Furthermore, after the teleportation the walk starts from the uniform distribution,

denoted by u . Hence, for the walk according to Q , we get that its distribution after T steps is

$$Q^T v = (1 - p_T)^T P^T v + \left[1 - (1 - p_T)^T\right] \sum_{i=0}^{\infty} (1 - p_T)^i p_T P^i u.$$

Using this along with the consistency of the random walk, i.e. $\pi = \lim_{T \rightarrow \infty} Q^T v$ and taking the limit we get for $p_T > 0$,

$$\pi = \sum_{i=0}^{\infty} (1 - p_T)^i p_T P^i u. \quad (5)$$

The right hand side describes exactly Process 12, which proves the statement. \square

PROCESS 14 (TRUNCATED GEOMETRIC). *This is the same as Process 12 with the difference that any walkers still active after t steps are stopped and their positions acquired as samples. This means that the number of steps is given by the minimum of t and a geometric random variable with parameter p_T .*

LEMMA 15. *Let $\pi_t \in \Delta^{n-1}$ denote the distribution of the masses acquired through Process 14. The mass it captures is lower-bounded as follows.*

$$\mu_k(\pi_t) \geq \mu_k(\pi) - 2C\sqrt{k\|\pi\|_{\infty}}[(1 - p_T)\lambda_2(\tilde{P}P)]^t$$

where $\tilde{P} = D^{-1}P'D$, for $D = \text{diag}(\pi)$.

PROOF. We start by getting a bound on the χ^2 divergence of π_t with respect to π .

$$\chi^2(\pi_t; \pi) \leq 2C[(1 - p_T)\lambda_2(\tilde{P}P)]^t$$

\square

4. EXPERIMENTS

In this section we compare the performance of our algorithm to the PageRank algorithm shipped with GraphLab v2.2 (PowerGraph) [15]. The fact that GraphLab is the fastest distributed engine for PageRank has been established experimentally in [21]. We focus on two algorithms: the basic built-in algorithm provided as part of the GraphLab *graph analytics toolkit*, referred to here as GRAPHLAB PR, and FROGWILD. Since we are looking for a top- k approximation and GRAPHLAB PR is meant to find the entire PageRank vector, we only run it for a small number of iterations (usually 2 are sufficient). This gives us a good top- k approximation and is much faster than running the algorithm until convergence. We also fine tune the algorithm's tolerance parameter to get a good but fast approximation.

We compare several performance metrics, namely: running time, network usage, and accuracy. The metrics do not include time and network usage required for loading the graph into GraphLab (known as the *ingress time*). They reflect only the execution stage.

4.1 The Systems

We perform experiments on two systems. The first system is a cluster of 20 virtual machines, created using VirtualBox 4.3 [22] on a single physical server. The server is based on an Intel[®] Xeon[®] CPU E5-1620 with 4 cores at 3.6 GHz, and 16 GB of RAM. The second system, comprises of a cluster of up to 24 EC2 machines on AWS (Amazon web services) [2]. We use m3.xlarge instances, based on Intel[®] Xeon[®] CPU E5-2670 with 4 vCPUs and 15 GB RAM.

Table 1: Processes used in our proof

PROCESS NAME	KIND	# STEPS	INDEP.	# SAMPLES	RESULT	INTRODUCED ERROR	STATEMENT
EXACT	RANDOM WALK, Q	∞	\checkmark	∞	π	0	LEMMA 1
GEOMETRIC	RW, P	$\text{Geom}(p_T)$	\checkmark	∞	π	0	
TRUNC. GEOM.	RW, P	$\min\{t, \text{Geom}(p_T)\}$	\checkmark	∞	π_t	$2C\sqrt{k\ \pi\ _\infty}[(1-p_T)\lambda_2]^t$	
BLOCKING	BLOCKING RW, P	$\min\{t, \text{Geom}(p_T)\}$	\times	N	$\hat{\pi}_N$	$\sqrt{\frac{k}{\delta}} \left[\frac{1}{N} + (1-p_s^2)p_{\cap}(t) \right]$	
FROGWILD!	BRIDGED RW, P	$\min\{t, \text{Geom}(p_T)\}$	\times	N	$\hat{\pi}_N$	0	

4.2 The Data

For the VirtualBox system, we use the LiveJournal graph [13] with 4.8M vertices and 69M edges. For the AWS system, in addition to the LiveJournal graph, we use the Twitter graph [12] which has 41.6M nodes and 1.4B edges.

4.3 Implementation

FROGWILD is implemented on the standard GAS (gather, apply, scatter) model. We implement `init()`, `apply()`, and `scatter()`. The purpose of `init()` is to collect the random walks sent to the node by its neighbors using `scatter()` in the previous iteration. In the first iteration, `init()` generates a random fraction of the initial total number of walkers. This implies that the initial walker locations are randomly distributed across nodes. FROGWILD requires the length of random walks to be geometrically distributed (see Section 2.2). For the sake of efficiency, we impose an upper bound on the length of random walks. The algorithm is executed for the constant number of iterations (experiments show good results with even 3 iterations) after which all the random walks are stopped simultaneously. The `apply()` function is responsible for keeping track of the number of walkers that have stopped on each vertex and `scatter()` distributes the walkers still alive to the neighbors of the vertex. The `scatter()` phase is the most challenging part of the implementation. In order to reduce information exchange between machines, we use a couple of ideas.

First, we notice that random walks do not have identity. Hence, random walks destined for the same neighbor can be combined into a single message. The second optimization and significant part of our work is modifying the GraphLab engine. The recent versions of GraphLab (since PowerGraph) partition the graph by splitting vertices. As a consequence, the engine will need to synchronize all the mirrors of a vertex over the network a number of times during each GAS cycle.

When running a few random walks, only a handful of neighbours end up receiving walkers. For this reason, synchronizing all mirrors can be very wasteful. We deal with that by implementing *randomized synchronization*. We expose parameter $p_s \in [0, 1]$ to the user as a small extension to the GraphLab API. It describes the fraction of replicas that will be synchronized. Replicas not synchronized remain idle for the upcoming scatter phase. The above edits in the engine are only a matter of a few (about 10) lines of code. Our modification of the GraphLab engine as well as the FROGWILD vertex program can be found in [6].

4.4 Results

FROGWILD is significantly faster and uses less network and CPU compared to GRAPHLAB PR. Let us start with

the Twitter graph and the AWS system. In Figure 1(a) we see that, while GRAPHLAB PR takes about 7.5 seconds per iteration (for 12 nodes), FROGWILD takes less than 1 sec, achieving more than a 7x speedup. Reducing the value of p_s decreases the running time. We see a similar picture when we study the total running time of the algorithms in Figure 1(b).

We plot network performance in Figure 1(c). We get a 1000x improvement compared to the exact GRAPHLAB PR, and more than 10x with respect to doing 1 or 2 iterations of GRAPHLAB PR. In Figure 1(d) we can see that the CPU usage reported by the engine is also much lower for FROGWILD.

We now turn to compare the approximation metrics for the PageRank algorithm. For various k , we check the two accuracy metrics: Mass captured (Figure 2(a)) and the Exact identification (Figure 2(b)). Mass captured – is the total PageRank that the reported top- k vertices worth in the exact ranking. Exact identification – is the number of vertices in the intersection of the reported top- k and the exact top- k lists. We can see that the approximation achieved by the FROGWILD for $p_s = 1$ and $p_s = 0.7$ always outperforms the GRAPHLAB PR with 1 iteration. The approximation achieved by the FROGWILD with $p_s = 0.4$ is relatively good for the both metrics, and with $p_s = 0.1$ is reasonable for the *Mass captured* metrics.

In Figure 3 we can see the tradeoff between the accuracy, total running time, and the network usage. The performance of FROGWILD is evaluated for various number of iterations and the values of p_s . The results show that with the accuracy comparable to GRAPHLAB PR, FROGWILD has much less running time and network usage. Figure 4 illustrate how much network traffic we save using FROGWILD. The area of each circle is proportional to the number of bytes sent by each algorithm.

Finally, we plot some results for the LiveJournal graph on the VirtualBox system. Figure 5 shows the effect of the number of walkers, N , and the number of iterations for FROGWILD on the achieved accuracy. Good accuracy and running time (see Figure 6) is achieved for 800K initial random walks and 4 iterations of FROGWILD. Also for the LiveJournal graph we can see, in Figure 7, that our algorithm is faster and uses much less network, while still maintaining good PageRank accuracy.

5. REFERENCES

- [1] A. Agarwal and S. Chakrabarti. Learning random walks to rank nodes in graphs. In *Proceedings of the 24th international conference on Machine learning*, pages 9–16. ACM, 2007.
- [2] Amazon web services. <http://aws.amazon.com>, 2014.

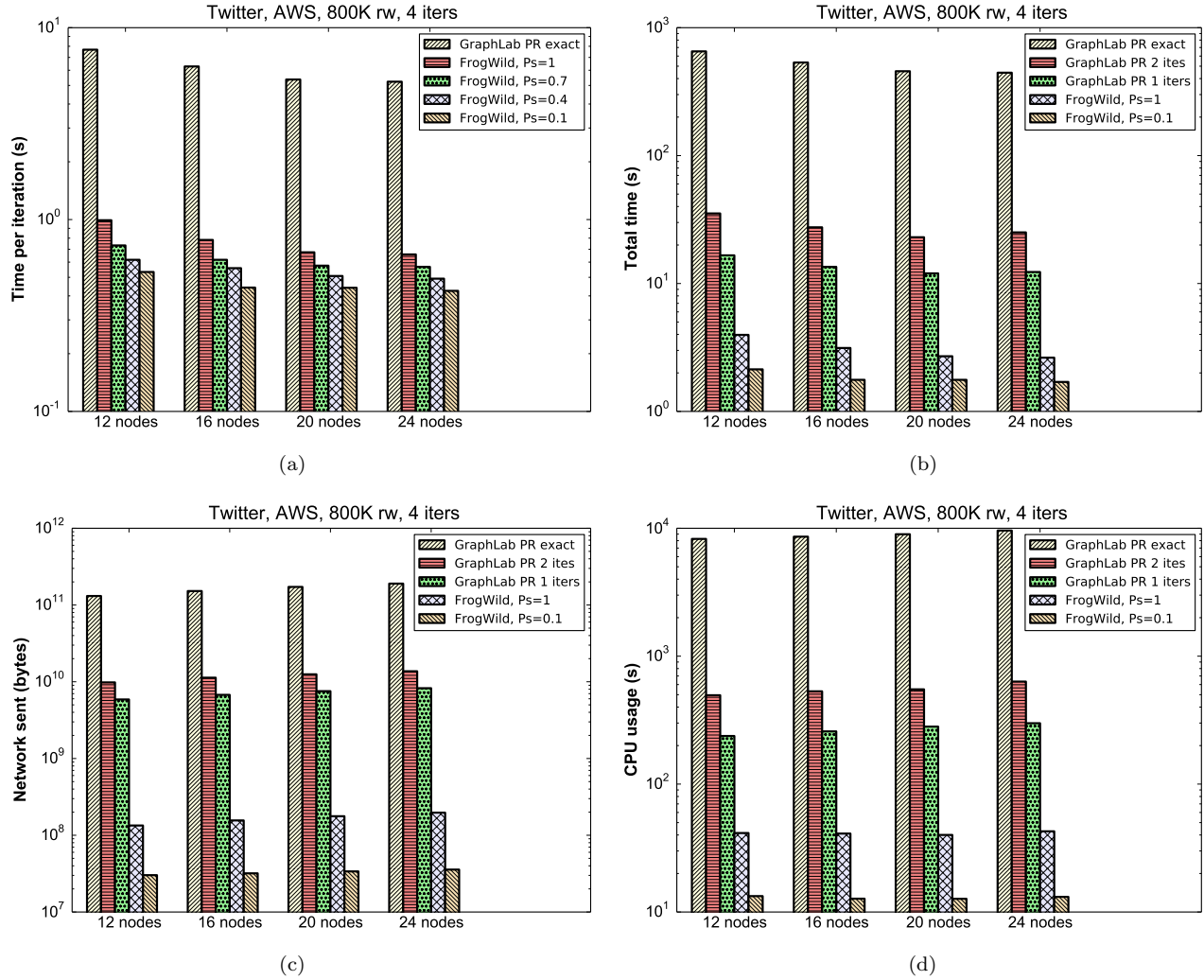
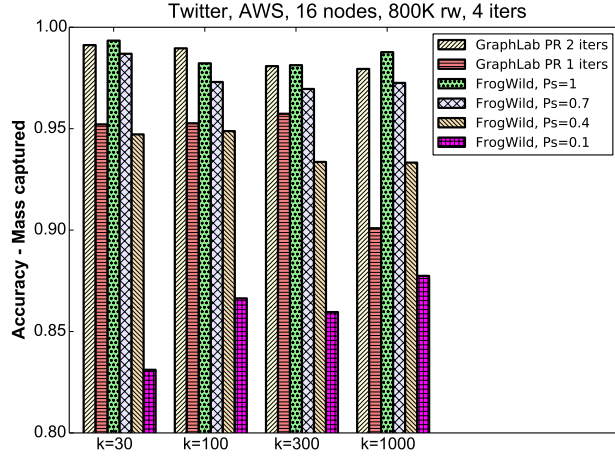
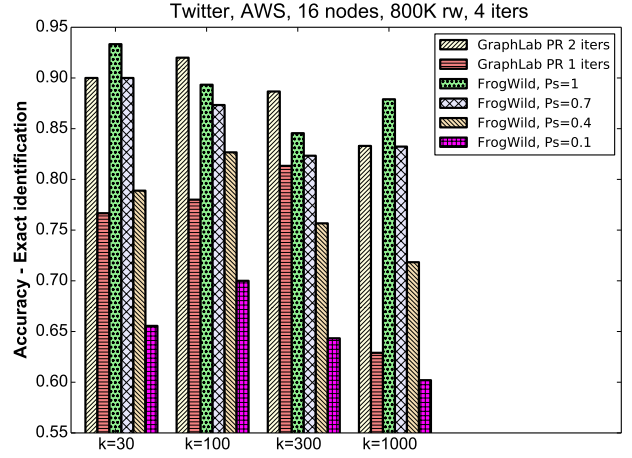


Figure 1: PageRank performance for various number of nodes. (a) – Running time per iteration. (b) – Total running time of the algorithms. (c) – Total network bytes sent by the algorithm during the execution (does not include ingress time). (d) – Total CPU usage time. Notice, this metric may be larger than the total running time since many CPUs run in parallel.

- [3] R. Andersen, C. Borgs, J. Chayes, J. Hopcraft, V. S. Mirrokni, and S.-H. Teng. Local computation of pagerank contributions. In *Algorithms and Models for the Web-Graph*, pages 150–165. Springer, 2007.
- [4] L. Becchetti and C. Castillo. The distribution of pagerank follows a power-law only for particular values of the damping factor. In *Proceedings of the 15th international conference on World Wide Web*, pages 941–942. ACM, 2006.
- [5] P. Berkhin. A survey on pagerank computing. *Internet Mathematics*, 2(1):73–120, 2005.
- [6] M. Bhorokhovich and I. Mitliagkas. FrogWild! code repository. <https://github.com/michaelbor/frogwild>, 2014. Accessed: 2014-10-30.
- [7] P. Bremaud. *Markov chains: Gibbs fields, Monte Carlo simulation, and queues*, volume 31. springer, 1999.
- [8] A. Z. Broder, R. Lempel, F. Maghoul, and J. Pedersen. Efficient pagerank approximation via graph aggregation. *Information Retrieval*, 9(2):123–138, 2006.
- [9] A. Das Sarma, D. Nanongkai, G. Pandurangan, and P. Tetali. Distributed random walks. *Journal of the ACM (JACM)*, 60(1):2, 2013.
- [10] S. Fortunato and A. Flammini. Random walks on directed networks: the case of pagerank. *International Journal of Bifurcation and Chaos*, 17(07):2343–2353, 2007.
- [11] J. E. Gonzalez, Y. Low, H. Gu, D. Bickson, and C. Guestrin. Powergraph: Distributed graph-parallel computation on natural graphs. In *OSDI*, volume 12, page 2, 2012.
- [12] H. Kwak, C. Lee, H. Park, and S. Moon. What is Twitter, a social network or a news media? In *WWW '10: Proceedings of the 19th international conference on World wide web*, pages 591–600, New York, NY, USA, 2010. ACM.
- [13] J. Leskovec and A. Krevl. SNAP Datasets: Stanford

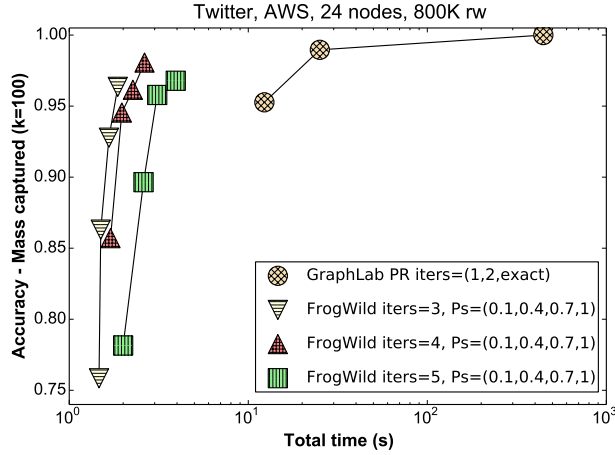


(a)

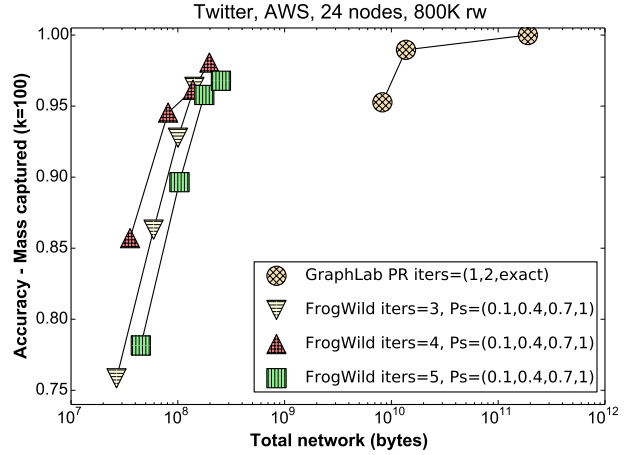


(b)

Figure 2: PageRank approximation accuracy for various number of top- k PageRank vertices. (a) – Mass captured. The total PageRank that the reported top- k vertices worth in the exact ranking. (b) – Exact identification. The number of vertices in the intersection of the reported top- k and the exact top- k lists.



(a)



(b)

Figure 3: PageRank approximation accuracy with the “Mass captured” metric for top-100 vertices. (a) - Accuracy versus total running time. (b) - Accuracy versus total network bytes sent.

- large network dataset collection.
<http://snap.stanford.edu/data>, June 2014.
- [14] P. Lofgren, S. Banerjee, A. Goel, and C. Seshadhri. Fast-ppr: Scaling personalized pagerank estimation for large graphs. *arXiv preprint arXiv:1404.3181*, 2014.
 - [15] Y. Low, J. Gonzalez, A. Kyrola, D. Bickson, C. Guestrin, and J. M. Hellerstein. Graphlab: A new parallel framework for machine learning. In *Conference on Uncertainty in Artificial Intelligence (UAI)*, July 2010.
 - [16] G. Malewicz, M. H. Austern, A. J. Bik, J. C. Dehnert, I. Horn, N. Leiser, and G. Czajkowski. Pregel: a system for large-scale graph processing. In *Proceedings of the 2010 ACM SIGMOD International Conference on Management of data*, pages 135–146. ACM, 2010.
 - [17] M. E. Newman. Power laws, pareto distributions and zipf’s law. *Contemporary physics*, 46(5):323–351, 2005.
 - [18] L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: Bringing order to the web. 1999.
 - [19] B. Recht, C. Re, S. Wright, and F. Niu. Hogwild: A lock-free approach to parallelizing stochastic gradient descent. In *Advances in Neural Information Processing Systems*, pages 693–701, 2011.
 - [20] A. D. Sarma, S. Gollapudi, and R. Panigrahy. Estimating pagerank on graph streams. *Journal of the ACM (JACM)*, 58(3):13, 2011.
 - [21] N. Satish, N. Sundaram, M. A. Patwary, J. Seo, J. Park, M. A. Hassaan, S. Sengupta, Z. Yin, and P. Dubey. Navigating the maze of graph analytics frameworks using massive graph datasets.

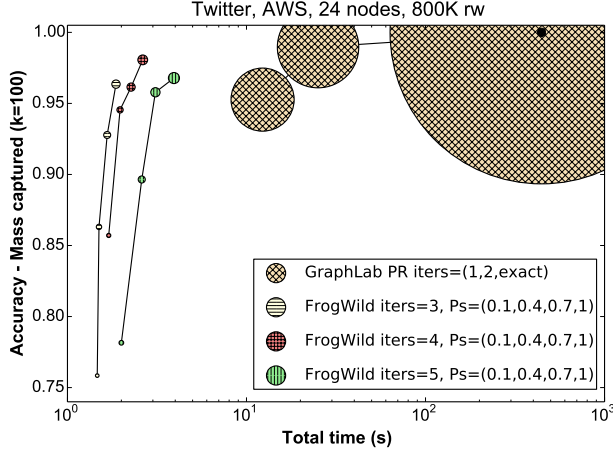


Figure 4: PageRank approximation accuracy with the “Mass captured” metric for top-100 vertices. The area of each circle is proportional to the total network bytes sent by the specific algorithm.

- [22] VirtualBox 4.3. www.virtualbox.org, 2014.
 [23] R. S. Xin, J. E. Gonzalez, M. J. Franklin, and I. Stoica. Graphx: A resilient distributed graph system on spark. In *First International Workshop on Graph Data Management Experiences and Systems*, page 2. ACM, 2013.

APPENDIX

A. THEOREM PROOFS

A.1 Proof for Theorem 1

PROOF. First we use the result of Theorem 3 to bound the distance between the PageRank vector π and the estimator $\hat{\pi}_N$. Note that, by construction, the adjusted matrix Q is always ergodic and $\min_i \pi(i) \geq c/n$ for $c = p_T$. All of the other assumptions of Theorem 3 are satisfied, hence,

$$\|\hat{\pi}_N - \pi\|_2 < C\lambda_2 + \sqrt{\frac{1}{\delta} \left[\frac{1}{N} + (1 - p_s^2)p_\tau(t) \right]},$$

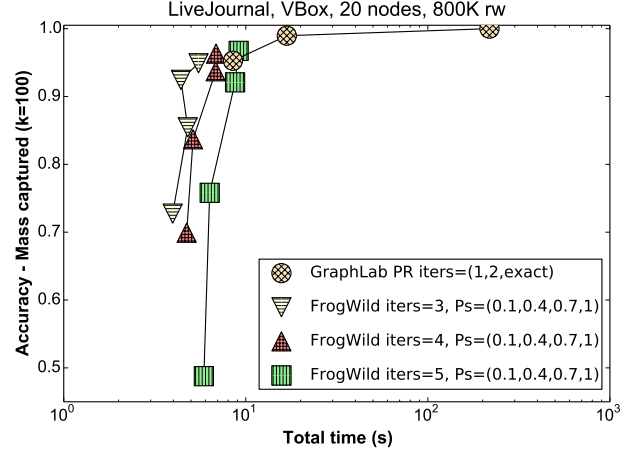
with probability at least $1 - \delta$. Here $C = \left(\frac{1-p_T}{p_T} \right)$, and $\lambda_2 = \lambda_2(\tilde{Q}Q)^t$. We define the top- k sets

$$\hat{S}^* = \operatorname{argmax}_{S \subseteq [n], |S|=k} \hat{\pi}_N(S)$$

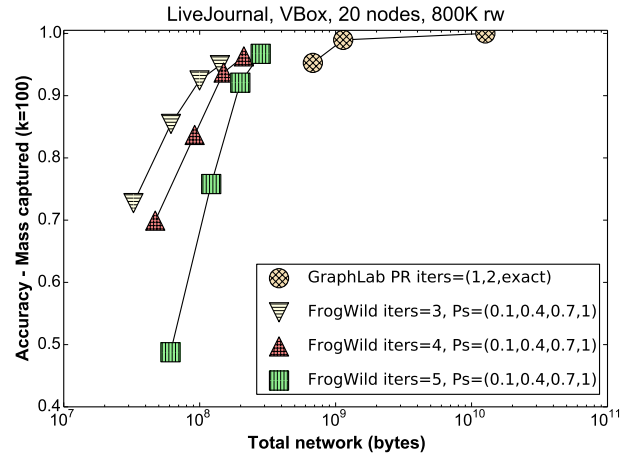
$$S^* = \operatorname{argmax}_{S \subseteq [n], |S|=k} \pi(S).$$

Let $\pi|_S$ denote the restriction of vector π to the set S . That is, $\pi|_S(i) = \pi(i)$ if $i \in S$ and 0 otherwise. Now we show that for any set S of cardinality k ,

$$\begin{aligned} |\pi(S) - \hat{\pi}_N(S)| &\leq \|(\pi - \hat{\pi}_N)|_S\|_1 \leq \sqrt{k} \|(\pi - \hat{\pi}_N)|_S\|_2 \\ &\leq \sqrt{k} \|\pi - \hat{\pi}_N\|_2 \triangleq \epsilon' \end{aligned} \quad (6)$$



(a)



(b)

Figure 7: (a) – Accuracy (fraction of nodes in the true top 100 that exists in the top 100 reported by the algorithm) of the PageRank algorithms as a function of running time. The running time varies due to the number of iterations performed by the algorithms. (b) – Accuracy as a function of network usage. The network usage varies due to the number of iterations performed by the algorithms.

Here we used the fact that for k -length vector x , $\|x\|_1 \leq \sqrt{k}\|x\|_2$ and $\|x|_S\| \leq \|x\|$. As a consequence of (6),

$$\begin{aligned} \hat{\pi}_N(\hat{S}^*) &= \max_{S \subseteq [n], |S|=k} \hat{\pi}_N(S) \\ &\geq \hat{\pi}_N(S^*) \geq \pi(S^*) - \epsilon'. \end{aligned} \quad (7)$$

Finally, using (6) and (7) in order we get

$$\begin{aligned} \mu(\hat{S}^*) &= \pi(\hat{S}^*) \geq \hat{\pi}_N(\hat{S}^*) - \epsilon' \\ &\geq \pi(S^*) - 2\epsilon' \geq \mu_{\max,k} - 2\epsilon'. \end{aligned}$$

This concludes the proof. \square

A.2 Theorem 3

THEOREM 3 (CONVERGENCE WITH ERASURES). *Let N walkers start from independent, uniformly random positions and take t steps according to the ergodic transition matrix P , under the edge erasure model. Let π denote the unique stationary distribution of P and assume that $\min_i \pi(i) \geq \frac{c}{n}$ for some constant $c \leq 1$. The estimator $\hat{\pi}_N$, described in Definition 8, falls inside a neighbourhood of π with good probability. Specifically, with probability at least $1 - \delta$,*

$$\|\hat{\pi}_N - \pi\|_2 < \left(\frac{1-c}{c}\right) \lambda_2(\tilde{P}P)^t + \sqrt{\frac{1}{\delta} \left[\frac{1}{N} + (1-p_s^2)p_{\cap}(t) \right]},$$

where $\tilde{P} = D^{-1}P'D$, for $D = \text{diag}(\pi)$.

For the convergence of the associated Markov chain we rely on Proposition 18, taken from [7]. First we need the following definition.

DEFINITION 16 (χ^2 -CONTRAST). *The χ^2 -contrast $\chi^2(\alpha; \beta)$ of α with respect to β is defined by*

$$\chi^2(\alpha; \beta) = \sum_i \frac{(\alpha(i) - \beta(i))^2}{\beta(i)}.$$

LEMMA 17. *Let $\pi \in \Delta^{n-1}$ a distribution satisfying $\min_i \pi(i) \geq \frac{c}{n}$ for constant $c \leq 1$, and let $u \in \Delta^{n-1}$ denote the uniform distribution. Then, $\chi^2(u; \pi) \leq \left(\frac{1-c}{c}\right)$.*

PROOF.

$$\begin{aligned} \chi^2(u; \pi) &= \sum_i \frac{(1/n - \pi(i))^2}{\pi(i)} = \sum_i \left(\frac{1}{n^2 \pi(i)} - \frac{1}{n} \right) \\ &= \frac{1}{n} \sum_i \frac{1 - n\pi(i)}{n\pi(i)} \leq \frac{1}{n} \sum_i \frac{1-c}{c} = \frac{1-c}{c} \end{aligned}$$

Here we used the assumed lower bound on $\pi(i)$ and the fact that $(1-x)/x$ is decreasing in x . \square

PROPOSITION 18 ([7]). *Consider an ergodic transition matrix P on a finite state space, and let π denote its unique stationary distribution. Let, $\tilde{P} = D^{-1}P'D$ be the multiplicative reversibilization of P . Then for any starting distribution ν , after t steps we get $\|P^t \nu - \pi\|_1^2 \leq \lambda_2(\tilde{P}P)^t \chi^2(\nu; \pi)$.*

Now we have all the necessary tools to prove the main theorem.

PROOF OF THEOREM 3. We know, that the individual walk distributions (marginals) follow the dynamics $x_l^{t+1} = Px_l^t$, for all $l \in [N]$, i.e. $x_l^t = x_1^t$. First we show that $\|\hat{\pi}_N - x_1^t\|_2$ is small.

$$\mathbb{P}(\|\hat{\pi}_N - x_1^t\|_2 > \epsilon) \leq \frac{\mathbb{E}[\|\hat{\pi}_N - x_1^t\|_2^2]}{\epsilon^2} \quad (8)$$

Here we used the Chebyshev inequality. We use s_l^t to denote the position of walker l at time t as a vector. For example, $s_l^t = e_i$, if walker l is at state i at time t . Now let us break down the norm on the numerator of (8).

$$\begin{aligned} \|\hat{\pi}_N - x_1^t\|_2^2 &= \left\| \frac{1}{N} \sum_l (s_l^t - x_1^t) \right\|_2^2 \\ &= \frac{1}{N^2} \sum_l \|s_l^t - x_1^t\|_2^2 + \frac{1}{N^2} \sum_{l \neq k} (s_l^t - x_1^t)'(s_k^t - x_1^t) \quad (9) \end{aligned}$$

For the diagonal terms we have:

$$\begin{aligned} \mathbb{E}[\|s_l^t - x_1^t\|_2^2] &= \sum_{i \in [n]} \mathbb{E}[\|s_l^t - x_1^t\|_2^2 | s_l^t = i] \mathbb{P}(s_l^t = i) \\ &= \sum_{i \in [n]} \|e_i - x_1^t\|_2^2 x_1^t(i) = 1 - \|x_1^t\|_2^2 \leq 1 \quad (10) \end{aligned}$$

Under the edge erasures model, the trajectories of different walkers are not generally independent. For example, if they happen to meet, they are likely to make the same decision for their next step, since they are faced with the same edge erasures. Now we prove that even when they meet, we can consider them to be independent with some probability that depends on p_s .

Consider the position processes for two walkers, $\{s_1^t\}_t$ and $\{s_2^t\}_t$. At each step t and node i a number of out-going edges are erased. Any walkers on i , will choose uniformly at random from the remaining edges. Now consider this alternative process.

DEFINITION 19 (BLOCKING WALK). *A blocking walk on the graph under the erasure model, follows these steps.*

1. Walker l finds herself on node i at time t .
2. Walker l draws her next state uniformly from the full set of out-going edges.

$$w \sim \text{Uniform}(\mathcal{N}_o(i))$$

3. If the edge (i, w) is erased at time t , the walker cannot traverse it. We call this event a block and denote it by B_l^t . In the event of a block:

- Walker redraws her next step from the out-going edges of i not erased at time t .
- Otherwise, w is used as the next state.

Clearly, a blocking walk is exactly equivalent to our original process; walkers end up picking a destination uniformly at random among the edges not erased. From now on we focus on this description of our original process. We use the same notation: $\{s_l^t\}_t$ for the position process and $\{x_l^t\}_t$ for the distribution at time t .

Let us focus on just two walkers, $\{s_1^t\}_t$ and $\{s_2^t\}_t$ and consider a third process: two independent random walks on the same graph. We assume that these walks operate on the full graph, i.e. no edges are erased. We denote their positions by $\{v_1^t\}_t$ and $\{v_2^t\}_t$ and their marginal distributions by $\{z_1^t\}_t$ and $\{z_2^t\}_t$.

DEFINITION 20 (TIME OF FIRST INTERFERENCE). *For two blocking walks, τ_I denotes the earliest time at which they meet and at least one of them experiences blocking.*

$$\tau_I = \min \{t : \{s_1^t = s_2^t\} \cap (B_1^t \cup B_2^t)\}$$

We call this quantity the time of first interference.

LEMMA 21 (PROCESS EQUIVALENCE). *For two walkers, the blocking walk and the independent walk are identical until the time of first interference. That is, assuming the same starting distributions, $x_1^0 = z_1^0$ and $x_2^0 = z_2^0$, then*

$$x_1^t = z_1^t \quad \text{and} \quad x_2^t = z_2^t \quad \forall t \leq \tau_I.$$

PROOF. The two processes are equivalent for as long as the blocking walkers make *independent* decisions effectively

picking *uniformly from the full set of edges* (before erasures). From the independence in erasures across time and vertices in Definition 3, as long as the two walkers do not meet, they are making an independent choices. Furthermore, since erasures are symmetric, the walkers will be effectively choosing uniformly over the full set of out-going edges.

Now consider any time t that the blocking walkers meet. As long as neither of them blocks, they are by definition taking independent steps uniformly over the set of all outgoing edges, maintaining equivalence to the independent walks process. This concludes the proof. \square

LEMMA 22. *Let all walkers start from the uniform distribution. The probability that the time of first interference comes before time t is upper bounded as follows. $\mathbb{P}(\tau_I \leq t) \leq (1 - p_s^2)p_{\cap}(t)$*

PROOF. Let M_t be the event of a meeting at time t , $M_t \triangleq \{s_1^t = s_2^t\}$. In the proof of Theorem 2, we establish that $\mathbb{P}(M_t) \leq \rho^t/n$, where ρ is the maximum row sum of the transition matrix P . Now denote the event of an interference at time t as follows. $I_t \triangleq M_t \cap (B_1^t \cup B_2^t)$, where B_1^t denotes the event of blocking, as described in Definition 19. Now,

$$\mathbb{P}(I_t) = \mathbb{P}(M_t \cap (B_1^t \cup B_2^t)) = \mathbb{P}(B_1^t \cup B_2^t | M_t)p_{\cap}(t).$$

For the probability of a block given that the walker have met at time t ,

$$\begin{aligned} \mathbb{P}(B_1^t \cup B_2^t | M_t) &= 1 - \mathbb{P}(\overline{B_1^t} \cap \overline{B_2^t} | M_t) \\ &= 1 - \mathbb{P}(\overline{B_2^t} | \overline{B_1^t}, M_t) \mathbb{P}(\overline{B_1^t} | M_t) \leq 1 - p_s^2. \end{aligned}$$

To get the last inequality we used, from Definition 3, the lower bound on the probability that an edge is not erased, and the lack of negative correlations in the erasures.

Combining the above results, we get

$$\begin{aligned} \mathbb{P}(\tau_I \leq t) &= \mathbb{P}\left(\sum_{\tau=1}^t \mathbb{I}_{\{I_\tau\}} \geq 1\right) \leq \mathbb{E}\left[\sum_{\tau=1}^t \mathbb{I}_{\{I_\tau\}}\right] = \sum_{\tau=1}^t \mathbb{P}(I_\tau) \\ &\leq \sum_{\tau=1}^t (1 - p_s^2) \mathbb{P}(M_\tau) = \frac{1 - p_s^2}{n} \sum_{\tau=1}^t \rho^\tau = (1 - p_s^2)p_{\cap}(t) \end{aligned}$$

which proves the statement. \square

Now we can bound the off-diagonal terms in (10).

$$\begin{aligned} \mathbb{E}\left[(s_l^t - x_1^t)'(s_k^t - x_1^t)\right] &= \mathbb{E}\left[(s_l^t - x_1^t)'(s_k^t - x_1^t) | \tau_I \leq t\right] \mathbb{P}(\tau_I \leq t) \\ &\quad + \mathbb{E}\left[(s_l^t - x_1^t)'(s_k^t - x_1^t) | \tau_I > t\right] \mathbb{P}(\tau_I > t) \end{aligned}$$

In the second term, the case when l, k have not interfered, by Lemma 21, the trajectories are independent and the cross-covariance is 0. In the first term, the cross-covariance is maximized when $s_l^t = s_k^t$. That is,

$$\mathbb{E}\left[(s_l^t - x_1^t)'(s_k^t - x_1^t) | \tau_I \leq t\right] \leq \mathbb{E}[\|s_l^t - x_1^t\|_2^2] \leq 1$$

From this we get

$$\mathbb{E}\left[(s_l^t - x_1^t)'(s_k^t - x_1^t)\right] \leq (1 - p_s^2)p_{\cap}(t), \quad (11)$$

and in combination with (10), we get from (9) that

$$\mathbb{E}\left[\|\hat{\pi}_N - x_1^t\|_2^2\right] \leq \frac{1}{N} + \frac{(N-1)(1-p_s^2)p_{\cap}(t)}{N}.$$

Finally, we can plug this into (8), to get

$$\mathbb{P}(\|\hat{\pi}_N - x_1^t\|_2 > \epsilon) \leq \frac{1 + (1-p_s^2)p_{\cap}(t)(N-1)}{N\epsilon^2}.$$

We can now combine this bound, with a bound on $\|x_1^t - \pi\|_2$ using the triangle inequality to get the theorem statement. We first use the fact that $\|x_1^t - \pi\|_2 \leq \|x_1^t - \pi\|_1$ and then apply Proposition 18 on the RHS, assuming a uniform starting distribution and invoking Lemma 17.

A.3 Proof of Theorem 2

PROOF. Let $u \in \Delta^{n-1}$ denote the uniform distribution over $[n]$, i.e. $u_i = 1/n$. The two walks start from the same initial uniform distribution, u , and independently follow the same law, Q . Hence, at time t they have the same marginal distribution, $p^t = Q^t u$. From the definition of the augmented transition probability matrix, Q , in Definition 1, we get that

$$\pi_i \geq \frac{p_T}{n}, \quad \forall i \in [n].$$

Equivalently, there exists a distribution $q \in \Delta^{n-1}$ such that

$$\pi = p_T u + (1 - p_T)q.$$

Now using this, along with the fact that π is the invariant distribution associated with Q (i.e. $\pi = Q^t \pi$ for all $t \geq 0$) we get that for any $t \geq 0$,

$$\begin{aligned} \|\pi\|_\infty &= \|Q^t \pi\|_\infty \\ &= \|Q^t p_T u + Q^t (1 - p_T)q\|_\infty \\ &\geq p_T \|Q^t u\|_\infty. \end{aligned}$$

For the last inequality, we used the fact that Q and q contain non-negative entries. Now we have a useful upper bound for the maximal element of the walks' distribution at time t .

$$\|p^t\|_\infty = \|Q^t u\|_\infty \leq \frac{\|\pi\|_\infty}{p_T} \quad (12)$$

Let M_t be the indicator random variable for the event of a meeting at time t .

$$M_t = \mathbb{I}_{\{\text{walkers meet at time } t\}}$$

Then, $\mathbb{P}(M_t = 1) = \sum_{i=1}^n p_i^t p_i^t = \|p^t\|_2^2$. Since p^0 is the uniform distribution, i.e. $p_i^0 = \frac{1}{n}$ for all i , then $\|p^0\|_2^2 = \frac{1}{n}$. We can also bound the l_2 norm of the distribution at other times. First, we upper bound the l_2 norm by the l_∞ norm.

$$\|p\|_2^2 = \sum_i p_i^2 \leq \sum_i p_i \|p\|_\infty = \|p\|_\infty$$

Here we used the fact that $p_i \geq 0$ and $\sum p_i = 1$.

Now, combining the above results, we get

$$\begin{aligned} p_{\cap}(t) &= \mathbb{P}\left(\sum_{\tau=0}^t M_\tau \geq 1\right) \leq \mathbb{E}\left[\sum_{\tau=0}^t M_\tau\right] = \sum_{\tau=0}^t \mathbb{E}[M_\tau] \\ &= \sum_{\tau=0}^t \mathbb{P}(M_\tau = 1) = \sum_{\tau=0}^t \|p^\tau\|_2^2 \leq \sum_{\tau=0}^t \|p^\tau\|_\infty \\ &\leq \frac{1}{n} + \frac{t\|\pi\|_\infty}{p_T}. \end{aligned}$$

For the last inequality, we used (12) for $t \geq 1$ and $\|p^0\|_2^2 = 1/n$. This proves the theorem statement. \square

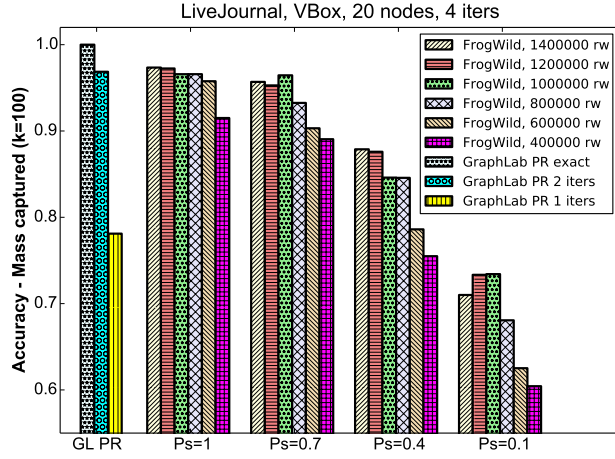
A.4 Proof of Proposition 9

PROOF. The expected maximum value of n independent draws from a power-law distribution with parameter θ , is

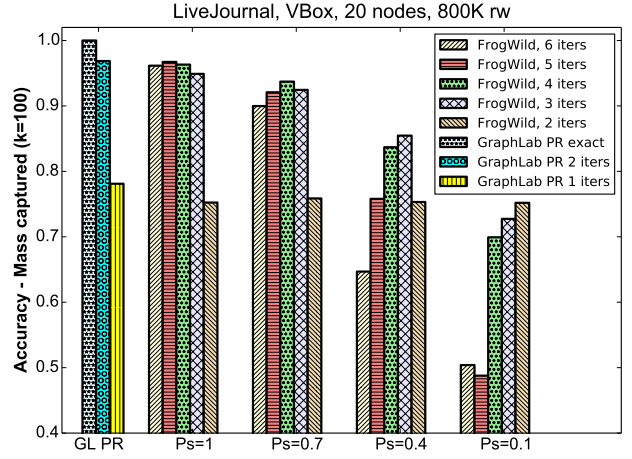
shown in [17] to be

$$\mathbb{E}x_{max} = O(n^{-\frac{1}{\theta-1}}).$$

Simple application of Markov's inequality, gives us the statement. \square

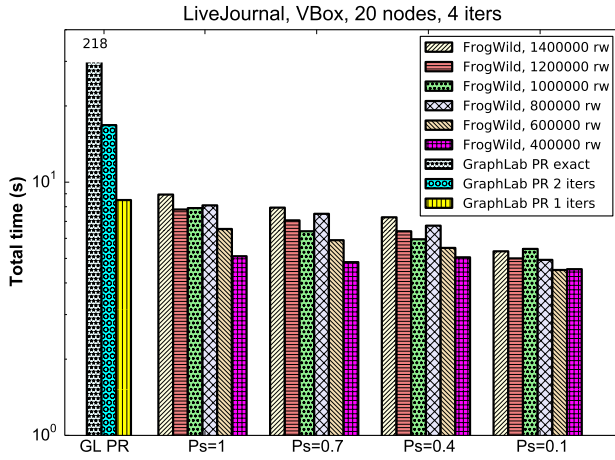


(a)

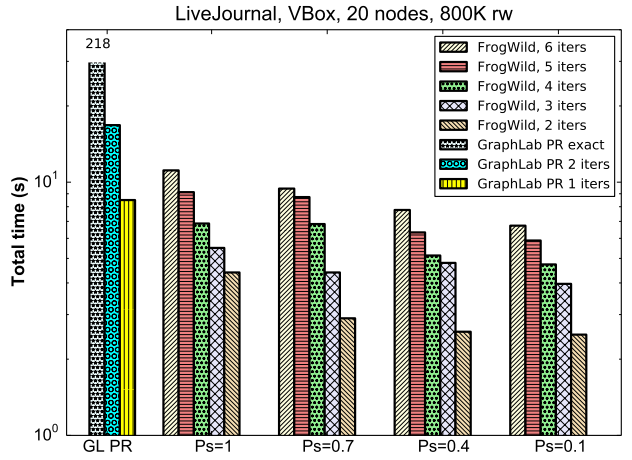


(b)

Figure 5: PageRank approximation accuracy with the “Mass captured” metric for top-100 vertices. (a) – Accuracy for various number of initial random walks in the FrogWild. (b) – Accuracy for various number of iterations of FrogWild.



(c)



(d)

Figure 6: Total running time of the PageRank algorithms. (a) – Total running time for various number of initial random walks in the FrogWild. (b) – Total running time for various number of iterations of FrogWild.