Analysis and Computation of Google’s PageRank

Ilse Ipsen

North Carolina State University

Joint work with Rebecca M. Wills
The heart of our software is PageRank™, a system for ranking web pages developed by our founders Larry Page and Sergey Brin at Stanford University.

And while we have dozens of engineers working to improve every aspect of Google on a daily basis, PageRank continues to provide the basis for all of our web search tools.

PageRank

An objective measure of the citation importance of a web page [Brin & Page 1998]

- Assigns a rank to every web page
- Influences the order in which Google displays search results
- Based on link structure of the web graph
- Topic independent
Overview

- Simple Web Model
- Google Matrix
- Stability of PageRank
- Eigenvalue Problem: Power Method
- Linear System: Jacobi Method
- Dangling Nodes
- Krylov Spaces
Simple Web Model

Construct matrix $S$

- Page $i$ has $d \geq 1$ outgoing links:
  If page $i$ has link to page $j$ then $s_{ij} = 1/d$
  else $s_{ij} = 0$

- Page $i$ has 0 outgoing links: (dangling node) $s_{ij} = 1/n$

$s_{ij}$: probability that surfer moves from page $i$ to page $j$
Matrix $S$

$S$ is stochastic: $0 \leq s_{ij} \leq 1$ \quad $S1 = 1$

Left eigenvector: $\omega^T S = \omega^T$ \quad $\omega \geq 0$ \quad $||\omega||_1 = 1$

Ranking: $\omega_i$ is probability that surfer visits page $i$

But:

- $S$ does not model surfing behaviour properly
- Rank sinks, and pages with zero rank
- Several eigenvalues with magnitude 1
  $\Rightarrow$ power method does not converge

Remedy: Change the matrix
Google Matrix

Convex combination

\[ G = \alpha S + (1 - \alpha)1 v^T \]

**Stochastic matrix** \( S \)

Damping factor 0 < \( \alpha \) < 1, e.g. \( \alpha = .85 \)

Personalization vector \( v > 0 \) \( \|v\|_1 = 1 \)

Properties of \( G \):

- stochastic \( \Rightarrow \) \( G \) has eigenvalue 1
- primitive \( \Rightarrow \) spectral radius 1 unique
Page Rank

Unique left eigenvector:

$$\pi^T G = \pi^T \quad \pi > 0 \quad \|\pi\|_1 = 1$$

Power method converges to $\pi$

$i$th entry of $\pi$: PageRank of page $i$

PageRank $\overset{\triangle}{=} \text{largest left eigenvector of } G$
Stability of PageRank

How sensitive is PageRank $\pi$ to

- Round off errors
- Changes in damping factor $\alpha$
- Changes in personalization vector $v$
- Addition/deletion of links
Perturbation Theory

For Markov chains

Schweizer 1968, Meyer 1980
Haviv & van Heyden 1984
Funderlic & Meyer 1986
Seneta 1988, 1991
Ipsen & Meyer 1994
Kirkland, Neumann & Shader 1998
Cho & Meyer 2000, 2001
Kirkland 2003, 2004
Perturbation Theory

For Google matrix

Chien, Dwork, Kumar & Sivakumar 2001
Ng, Zheng & Jordan 2001
Bianchini, Gori & Scarselli 2003
Boldi, Santini & Vigna 2004
Langville & Meyer 2004
Golub & Greif 2004
Kirkland 2005
Chien, Dwork, Kumar, Simon & Sivakumar 2005
Changes in the Matrix $S$

**Exact:**

$$\pi^T G = \pi^T \quad G = \alpha S + (1 - \alpha) 1v^T$$

**Perturbed:**

$$\tilde{\pi}^T \tilde{G} = \tilde{\pi}^T \quad \tilde{G} = \alpha (S + E) + (1 - \alpha) 1v^T$$

**Error:**

$$\tilde{\pi}^T - \pi^T = \alpha \tilde{\pi}^T E (I - \alpha S)^{-1}$$

$$\|\tilde{\pi} - \pi\|_1 \leq \frac{\alpha}{1 - \alpha} \|E\|_\infty$$

[Kirkland 2005]
Changes in Damping Factor $\alpha$

Exact:

$$\pi^T G = \pi^T \quad G = \alpha S + (1 - \alpha) 1 v^T$$

Perturbed:

$$\tilde{\pi}^T \tilde{G} = \tilde{\pi}^T \quad \tilde{G} = (\alpha + \mu) S + (1 - (\alpha + \mu)) 1 v^T$$

Error:

$$\| \tilde{\pi} - \pi \|_1 \leq \frac{2}{1 - \alpha} \mu$$

[Langville & Meyer 2004]
Changes in Vector $\nu$

Exact:

$$\pi^T G = \pi^T \quad G = \alpha S + (1 - \alpha)1 \nu^T$$

Perturbed:

$$\tilde{\pi}^T \tilde{G} = \tilde{\pi}^T \quad \tilde{G} = \alpha S + (1 - \alpha)1 (\nu + f)^T$$

Error:

$$\| \tilde{\pi} - \pi \|_1 \leq \| f \|_1$$
Sensitivity of PageRank $\pi$

$$\pi^T G = \pi^T \quad G = \alpha S + (1 - \alpha) 1 v^T$$

Changes in

- $S$: condition number $\alpha/(1 - \alpha)$
- $\alpha$: condition number $2/(1 - \alpha)$
- $f$: condition number 1

$\alpha = .85$: condition numbers $\leq 14$
$\alpha = .99$: condition numbers $\leq 200$

PageRank insensitive to perturbations
Adding an In-Link

Adding an in-link can only **increase** PageRank (monotonicity)

Removing an in-link can only decrease PageRank

[Chien, Dwork, Kumar & Sivakumar 2001]
[Chien, Dwork, Kumar, Simon & Sivakumar 2005]
Adding an Out-Link

\[ \tilde{\pi}_3 = \frac{1 + \alpha + \alpha^2}{3(1 + \alpha + \alpha^2/2)} < \pi_3 = \frac{1 + \alpha + \alpha^2}{3(1 + \alpha)} \]

Adding an out-link may decrease PageRank
PageRank Computation

Eigenvector problem: \( \pi^T G = \pi^T \)

Power method:

Pick \( x^{(0)} > 0 \) \( \| x^{(0)} \|_1 = 1 \)

Repeat \[ x^{(k+1)} \] \( T = [ x^{(k)} ] T G \)

until \( \| x^{(k+1)} - x^{(k)} \| \leq \tau \)

\[ [ x^{(k+1)} ] T - [ x^{(k)} ] T = [ x^{(k)} ] T G - [ x^{(k)} ] T \] residual
Why is an Iteration Cheap?

Google matrix \( G = \alpha S + (1 - \alpha)1v^T \)

\[
S = H + \underbrace{dw^T}_{\text{dangling nodes}} \quad \quad w \geq 0 \quad \|w\|_1 = 1
\]

often: \( w = \frac{1}{n} 1 \)

Matrix \( H \): models webgraph
substochastic
dimension: several billion
very sparse
Matrix Vector Multiplication

Vector $x > 0 \quad \|x\|_1 = 1$

$$x^T G = x^T \left[ \alpha (H + d w^T) + (1 - \alpha)1v^T \right]$$
$$= \alpha x^T H + \alpha x^T d w^T + (1 - \alpha)v^T$$

Cost: # non-zeros in $H$
Asymptotic Convergence Rate

\[ G = \alpha S + (1 - \alpha)1v^T \]

Power method convergence rate:

\[ \frac{|\lambda_2(G)|}{|\lambda_1(G)|} \leq \alpha \]

Eigenvalues of \( G \):

\[ \lambda_1(G) = 1 \quad \lambda_i(G) = \alpha \lambda_i(S) \quad i > 1 \]

[Haveliwala & Kamvar 2003] [Langville & Meyer 2003] [Elden 2003]
Error in Power Method

\[ \pi^T G = \pi^T \quad G = \alpha S + (1 - \alpha)1v^T \]

Error in iteration \( k \)

- **Forward:** \( e_k \equiv x^{(k)} - \pi \)
  \[ e_k^T = \alpha^k e_0^T S^k \quad \|e_k\|_1 \leq 2 \alpha^k \]

  [Bianchini, Gori & Scarselli 2003]

- **Backward:** \( r_k^T = [x^{(k)}]^T G - [x^{(k)}]^T \)
  \[ r_k^T = \alpha^k r_0^T S^k \quad \|r_k\|_1 \leq 2 \alpha^k \]
Termination

Residual norm $\|r_k\|_1 \leq 2\alpha^k$
Stop when $\|r_k\|_1 \leq 10^{-8}$

For $\alpha = .85$: $k \geq 119$

<table>
<thead>
<tr>
<th>$n$</th>
<th>2293</th>
<th>2947</th>
<th>3468</th>
<th>5757</th>
<th>281903</th>
<th>683446</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k$</td>
<td>86</td>
<td>85</td>
<td>90</td>
<td>90</td>
<td>91</td>
<td>92</td>
</tr>
</tbody>
</table>

Thanks to Chen Greif!

One-norm too stringent?
Infinity-Norm Bounds

\[ \pi^T G = \pi^T \quad G = \alpha S + (1 - \alpha) \mathbf{1} v^T \]

Error in iteration \( k \)

- **Forward:**
  \[ \max_i |\pi_i - x_i^{(k)}| \leq \alpha^k \max_i |\pi_i - x_i^{(0)}| \]

- **Backward:**
  \[ \max_i |x_i^{(k+1)} - x_i^{(k)}| \leq \alpha^k \max_i |x_i^{(1)} - x_i^{(0)}| \]
## Iteration Counts

<table>
<thead>
<tr>
<th>$n$</th>
<th>1-N</th>
<th>$\infty$-N</th>
<th>Disagrees</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>2293</td>
<td>86</td>
<td>75</td>
<td>19</td>
<td>.8</td>
</tr>
<tr>
<td>2947</td>
<td>85</td>
<td>76</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3468</td>
<td>90</td>
<td>83</td>
<td>3</td>
<td>.1</td>
</tr>
<tr>
<td>5757</td>
<td>90</td>
<td>79</td>
<td>18</td>
<td>.3</td>
</tr>
<tr>
<td>281903</td>
<td>91</td>
<td>69</td>
<td>24064</td>
<td>8.5</td>
</tr>
<tr>
<td>683446</td>
<td>92</td>
<td>65</td>
<td>50292</td>
<td>7.4</td>
</tr>
</tbody>
</table>

Disagrees: # pages with different rankings

$$\|r_k\|_{1,\infty} \leq 10^{-8}, \alpha = .85$$
Termination Criterion

Old: \[ \| x^{(k+1)} - x^{(k)} \|_1 \leq \tau \]

- Bound becomes more stringent as \( n \) grows

New: \[ \| x^{(k+1)} - x^{(k)} \|_\infty \leq \tau \]

- Reduces iteration count
- Disagreements in ranking \( \leq 10\% \)
Iteration Counts for Different $\alpha$ 

$$G = \alpha S + (1 - \alpha)1v^T$$

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$n = 281903$</th>
<th>$n = 683446$</th>
<th>bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>.85</td>
<td>69</td>
<td>65</td>
<td>119</td>
</tr>
<tr>
<td>.90</td>
<td>107</td>
<td>102</td>
<td>166</td>
</tr>
<tr>
<td>.95</td>
<td>219</td>
<td>220</td>
<td>415</td>
</tr>
<tr>
<td>.99</td>
<td>1114</td>
<td>1208</td>
<td>2075</td>
</tr>
</tbody>
</table>

**bound**: $k$ such that $2\alpha^k \leq 10^{-8}$

Fewer iterations than predicted by bound
Properties of Power Method

- Converges to unique vector
- Convergence rate $\alpha$
- Vectorizes
- Storage for only a single vector
- Matrix vector multiplication with very sparse matrix
- Accurate (no subtractions)
- Simple (few decisions)

But: can be slow
PageRank Computation

- **Power method**
  Page, Brin, Motwani & Winograd 1999

- **Acceleration of power method**
  Brezinski & Redivo-Zaglia 2004
  Kamvar, Haveliwala, Manning & Golub 2003
  Haveliwala, Kamvar, Klein, Manning & Golub 2003

- **Aggregation/Disaggregation**
  Ipsen & Kirkland 2004
PageRank Computation

- **Methods that adapt to web graph**
  - Broder, Lempel, Maghoul & Pedersen 2004
  - Kamvar, Haveliwala & Golub 2004
  - Haveliwala, Kamvar, Manning & Golub 2003
  - Lee, Golub & Zenios 2003
  - Lu, Zhang, Xi, Chen, Liu, Lyu & Ma 2004

- **Krylov methods**
  - Golub & Greif 2004
PageRank from Linear System

Eigenvector problem:

\[
\pi^T (\alpha S + (1 - \alpha)1v^T) = \pi^T \pi \geq 0 \quad \|\pi\|_1 = 1
\]

Linear system:

\[
\pi^T (I - \alpha S) = (1 - \alpha)v^T
\]

\(I - \alpha S\) nonsingular M-matrix

[Arasu, Novak, Tomkins & Tomlin 2002]
[Bianchini, Gori & Scarselli 2003]
Stationary Iterative Methods

\[ \pi^T (I - \alpha S) = (1 - \alpha) v^T \]

- Can be faster than power method
- Can be faster than Krylov space methods
- Predictable, monotonic convergence
- Can converge even for \( \alpha \approx 1 \)
- Less storage than Krylov space methods
- Accurate (no subtractions)

[Gleich, Zhulov & Berkhin 2005]
Example

[Gleich, Zhukov & Berkhin 2005]

Web graph: 1.4 billion nodes
6.6 billion edges
Beowulf cluster with 140 processors
Stopping criterion: residual norm $\leq 10^{-7}$

BiCGSTAB: 28.2 minutes (preconditioner?)

Power method: 35.5 minutes
Jacobi Method

Assume no page has a link to itself

\[ \pi^T (I - \alpha S) = (1 - \alpha)v^T \quad I - \alpha S = D - O \]

\[ [x^{(k+1)}]^T = [x^{(k)}]^T OD^{-1} + (1 - \alpha)v^T D^{-1} \]

- \( I - \alpha S \) is M-matrix
- Jacobi converges
- No dangling nodes: \( D = I \quad O = \alpha S \)

Jacobi method = power method
Dangling Nodes

\[ S = H + dw^T \] is dense

What to do about dangling nodes?

- **Remove** [Brin, Page, Motwani & Winograd 1998]
  No PageRank for dangling nodes
  Biased PageRank for other nodes

- **Lump into single state** [Lee, Golub & Zenios 2003]
  As above

- **Remove** \( dw^T \) [Langville & Meyer 2004]
  [Arasu, Novak, Tomkins & Tomlin 2002]
  \( H \) is not stochastic

What is being computed?

HH05 – p.35
Use $\nu$ for Dangling Nodes

\[
\pi^T (I - \alpha S) = (1 - \alpha) \nu^T \quad S = H + d \omega^T
\]

Choose $w = \nu$

\[
\pi^T (I - \alpha H) = (1 - \alpha + \alpha \pi^T d) \nu^T
\]

multiple of $\nu^T$

Solve $\delta^T (I - \alpha H) = \text{multiple of } \nu^T$

Then $\delta$ is multiple of $\pi$

[Gleich, Zhukov & Berkhin 2005]
## Iteration Counts for $w = v$

<table>
<thead>
<tr>
<th>$n$</th>
<th>Power</th>
<th>Jacobi</th>
</tr>
</thead>
<tbody>
<tr>
<td>2293</td>
<td>75</td>
<td>74</td>
</tr>
<tr>
<td>2947</td>
<td>76</td>
<td>74</td>
</tr>
<tr>
<td>3468</td>
<td>83</td>
<td>82</td>
</tr>
<tr>
<td>5757</td>
<td>79</td>
<td>78</td>
</tr>
<tr>
<td>281903</td>
<td>69</td>
<td>69</td>
</tr>
<tr>
<td>683446</td>
<td>65</td>
<td>65</td>
</tr>
</tbody>
</table>

$\|r_k\|_\infty \leq 10^{-8}$, $\alpha = .85$

Jacobi: same # iterations as power method
Extension to Arbitrary $\omega$

\[ \pi^T(I - \alpha S) = (1 - \alpha)v^T \quad S = H + dw^T \]

Rank-one update: $I - \alpha S = (I - \alpha H) - \alpha dw^T$

1. Solve $\delta^T(I - \alpha H) = (1 - \alpha)v^T$
2. Solve $\omega^T(I - \alpha H) = w^T$
3. Update $\pi^T = \delta^T + \alpha \frac{\delta^T d}{1 - \alpha \omega^Td} \omega^T$

This requires only two sparse solves
Solve with $I - \alpha H$

After similarity permutation:

$$H = \begin{pmatrix} H_1 & H_2 \\ 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} x_1^T \\ x_2^T \end{pmatrix} \begin{pmatrix} I - \alpha H_1 & -\alpha H_2 \\ 0 & I \end{pmatrix} = \begin{pmatrix} b_1^T \\ b_2^T \end{pmatrix}$$

1. Sparse solve
   $$x_1^T (I - \alpha H_1) = b_1$$
2. Set
   $$x_2^T = \alpha x_1^T H_2 + b_2^T$$
PageRank via Linear System

Rank one update: \[ S = \begin{pmatrix} H_1 & H_2 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} w^T \]

- **General dangling node vector** \( w \geq 0, \|w\|_1 = 1 \)

  Traditional: \( w = \frac{1}{n} \mathbf{1}, \quad w = v \)

- **Cost:**
  - Two sparse solves with \( I - \alpha H_1 \) via Jacobi
  - Two matrix vector multiplications with \( H_2 \)
  - Inner products and vector additions

- More dangling nodes \( \Rightarrow \) cheaper
Krylov Spaces

$$G = \alpha S + (1 - \alpha) 1 v^T$$

- Eigen problem: $$\pi^T G = \pi^T$$ \quad $$\|\pi\|_1 = 1$$
- Linear system: $$\pi^T (I - \alpha S) = (1 - \alpha) v^T$$
- Solution:

$$\pi^T = (1 - \alpha) \sum_{j=0}^{\infty} \alpha^j v^T S^j$$

in $$\mathcal{K}_\infty(v^T, S) \equiv \text{span}\{v^T, v^T S, v^T S^2, \ldots\}$$
Contributions of \( \pi \) in \( K_k(v^T, S) \):

\[
\pi_k^T \equiv (1 - \alpha) \sum_{j=0}^{k} \alpha^j v^T S^j
\]

Error: \( \| \pi_k - \pi \|_1 \leq \alpha^{k+1} \)

Iteration \( k \) of Power method with \( x_0 = v \):

\[
[x^{(k)}]^T = \pi_k^T + \alpha^{k+1} v^T S^k \quad \text{in} \quad K_k(v^T, S)
\]

Error: \( \| x^{(k)} - \pi \|_1 \leq 2 \alpha^{k+1} \)

Power method produces good approximation.
Modified Power Method

Compute $\pi_k^T \equiv (1 - \alpha) \sum_{j=0}^{k} \alpha^j v^T S^j$

Recursion $\pi_{k+1}^T = \alpha \pi_k^T S + \pi_0^T$

Norm $\|\pi_k\|_1 = 1 - \alpha^{k+1}$

Residual $\pi_k^T G - \pi_k^T = \pi_{k+1}^T - \pi_k^T - \alpha^{k+1} \pi_0^T$

Set $\pi_0 = (1 - \alpha)v^T$

Repeat $\pi_{k+1}^T = \alpha \pi_k^T S + \pi_0^T$ until

$\|\pi_{k+1} - \pi_k - \alpha^{k+1} \pi_0\|/(1 - \alpha^{k+1}) \leq \tau$
## Iteration Counts for $\alpha = .85$

<table>
<thead>
<tr>
<th>$n$</th>
<th>Power</th>
<th>MPower</th>
<th>Disagrees</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>2293</td>
<td>75</td>
<td>90</td>
<td>32</td>
<td>1.4</td>
</tr>
<tr>
<td>2947</td>
<td>76</td>
<td>90</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3468</td>
<td>83</td>
<td>87</td>
<td>3</td>
<td>.09</td>
</tr>
<tr>
<td>5757</td>
<td>79</td>
<td>82</td>
<td>5</td>
<td>.09</td>
</tr>
<tr>
<td>281903</td>
<td>64</td>
<td>74</td>
<td>40091</td>
<td>14.2</td>
</tr>
<tr>
<td>683446</td>
<td>65</td>
<td>72</td>
<td>130885</td>
<td>19</td>
</tr>
</tbody>
</table>

Disagrees:  # pages with different rankings

HH05 – p.44
### Iteration Counts for $\alpha = .99$

<table>
<thead>
<tr>
<th>$n$</th>
<th>Power</th>
<th>MPower</th>
<th>Disagrees</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>2293</td>
<td>1188</td>
<td>1274</td>
<td>15</td>
<td>.7</td>
</tr>
<tr>
<td>2947</td>
<td>1165</td>
<td>1253</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3468</td>
<td>1327</td>
<td>1190</td>
<td>3</td>
<td>.09</td>
</tr>
<tr>
<td>5757</td>
<td>1258</td>
<td>1162</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>281903</td>
<td>1114</td>
<td>1008</td>
<td>58967</td>
<td>21</td>
</tr>
<tr>
<td>683446</td>
<td>1208</td>
<td>910</td>
<td>284962</td>
<td>42</td>
</tr>
</tbody>
</table>

Disagrees: # pages with different rankings
Modified Power Method

- Computes contribution of $\pi$ in $\mathcal{K}_k(v^T, S)$

Compared to power method:
- Fewer iterations for $\alpha \approx 1$ and larger $n$
- High rank disagreements for "very" large $n$

Potential?
Summary

Google Matrix \( G = \alpha S + (1 - \alpha) 1v^T \)

- PageRank = left eigenvector of \( G \)
- PageRank insensitive to perturbations in \( G \)
- Adding in-links can only increase PageRank
- Adding out-links may decrease PageRank
- PageRank \( \pi \) in \( \mathcal{K}_\infty(v^T, S) \)
- Iterate \( k \) of power method in \( \mathcal{K}_k(v^T, S) \)
- Forward, backward errors \( \leq 2 \alpha^k \)
- Infinity-norm termination criterion
Summary, ctd

Modified Power Method

- Computes contribution of $\pi$ in $\mathcal{K}_k(v^T, S)$
- May be faster for large $\alpha$

Jacobi Method

- Computes PageRank from linear system
- Can be competitive with Krylov methods
- Rank one update for dangling nodes
- More dangling nodes $\Rightarrow$ cheaper