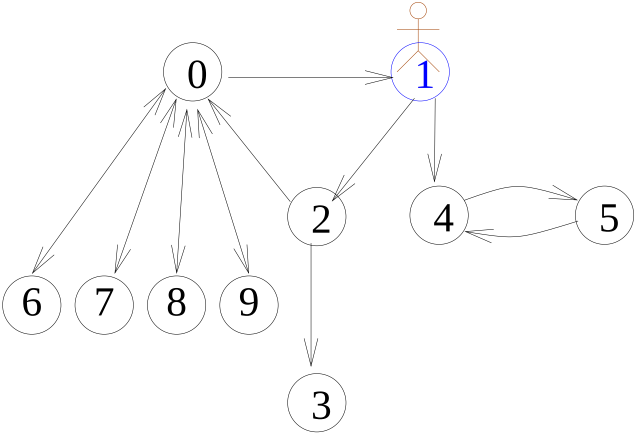
PageRank as a Function of the Damping Factor

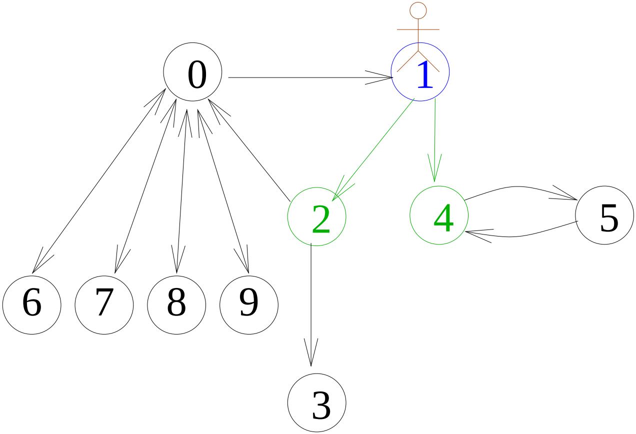
Paolo Boldi Sebastiano Vigna **Massimo Santini**
Dipartimento di Scienze dell'Informazione
Università degli Studi di Milano

PageRank [Page et al., '98]: the Web-Surfer Metaphor



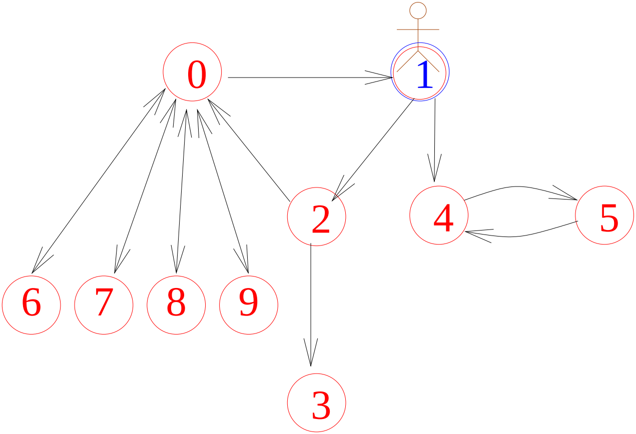
A surfer is wandering about the web. . .

PageRank [Page et al., '98]: the Web-Surfer Metaphor



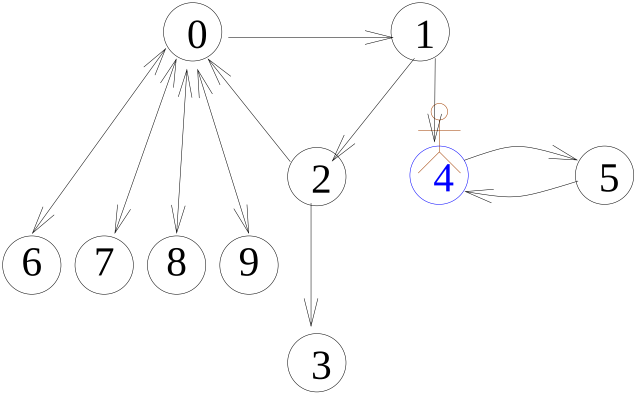
At each step, with probability α (s)he chooses the next page by clicking on a random link. . .

PageRank [Page et al., '98]: the Web-Surfer Metaphor

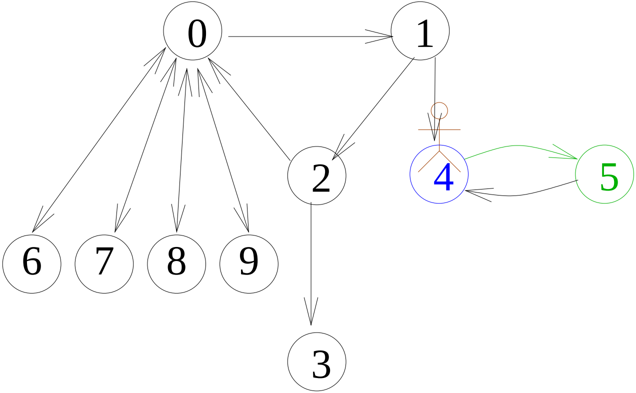


...with probability $1 - \alpha$, (s)he jumps to a random node (chosen uniformly or according to a fixed distribution, the *preference vector*)

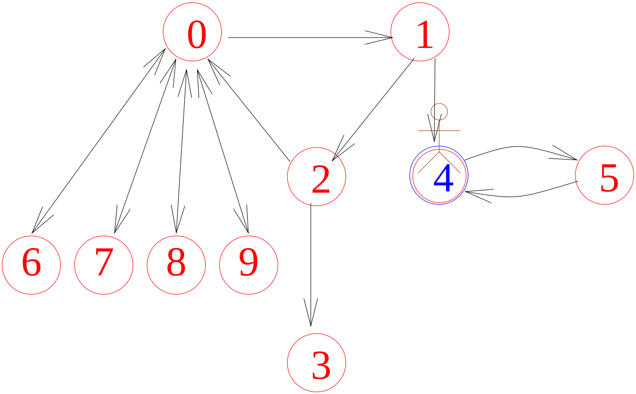
PageRank [Page et al., '98]: the Web-Surfer Metaphor



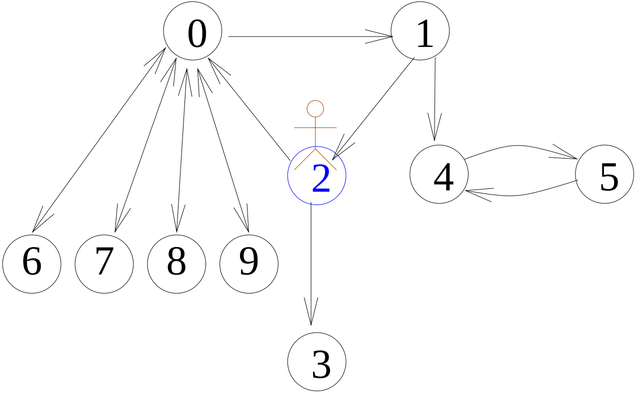
PageRank [Page et al., '98]: the Web-Surfer Metaphor



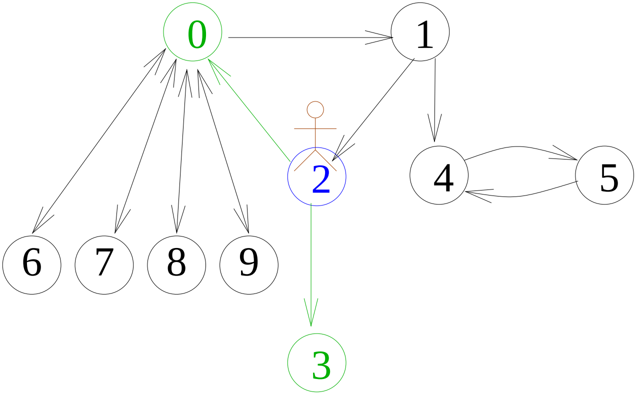
PageRank [Page et al., '98]: the Web-Surfer Metaphor



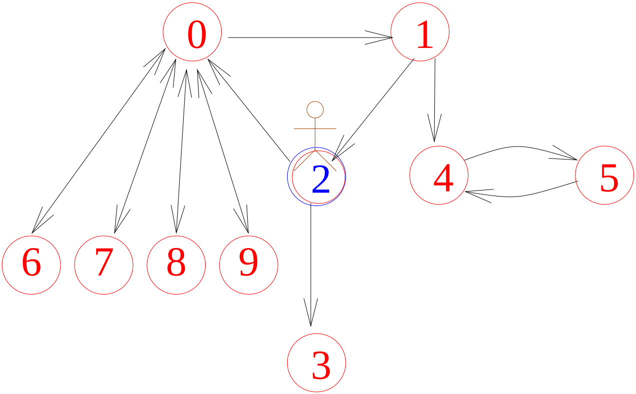
PageRank [Page et al., '98]: the Web-Surfer Metaphor



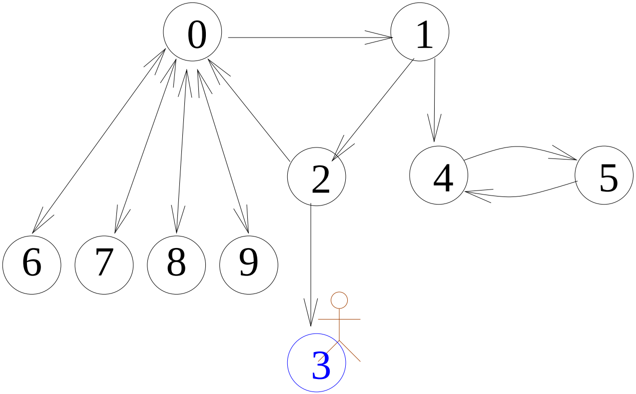
PageRank [Page et al., '98]: the Web-Surfer Metaphor



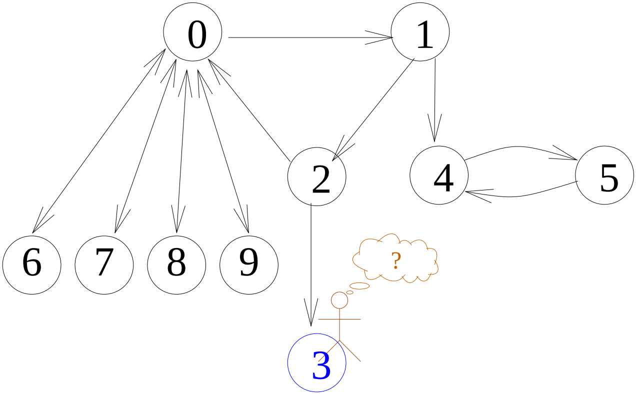
PageRank [Page et al., '98]: the Web-Surfer Metaphor



PageRank [Page et al., '98]: the Web-Surfer Metaphor

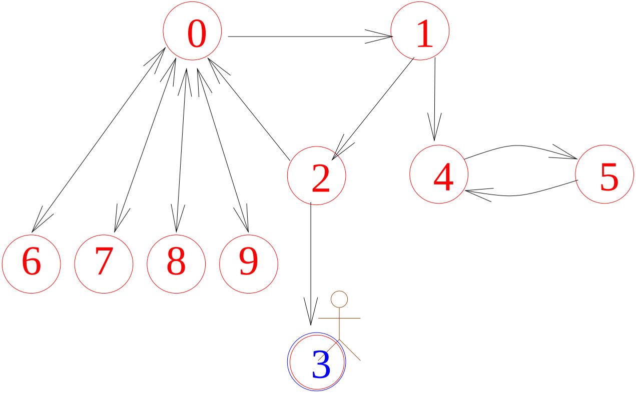


PageRank [Page et al., '98]: the Web-Surfer Metaphor



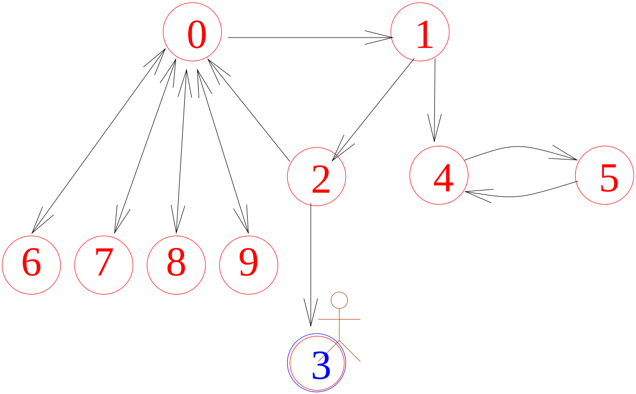
What if (s)he reaches a node with no outlinks (a *dangling node*)?

PageRank [Page et al., '98]: the Web-Surfer Metaphor



In that case, (s)he jumps to a random node *with probability 1*.

PageRank [Page et al., '98]: the Web-Surfer Metaphor



The PageRank of a page is the average fraction of time spent by the surfer on that page.

What does PageRank depends on?

PageRank is the limit distribution of a stochastic process whose states are Web pages.

What does PageRank depends on?

PageRank is the limit distribution of a stochastic process whose states are Web pages.
What does this distribution depend on?

What does PageRank depends on?

PageRank is the limit distribution of a stochastic process whose states are Web pages.
What does this distribution depend on?

- the *graph* G

What does PageRank depends on?

PageRank is the limit distribution of a stochastic process whose states are Web pages.
What does this distribution depend on?

- the *graph* G
- the *preference vector* $\mathbf{v}$

What does PageRank depends on?

PageRank is the limit distribution of a stochastic process whose states are Web pages.
What does this distribution depend on?

- the *graph* G
- the *preference vector* $\mathbf{v}$
- the *damping factor* α.

What does PageRank depends on?

PageRank is the limit distribution of a stochastic process whose states are Web pages.
What does this distribution depend on?

- the *graph* G
- the *preference vector* $\mathbf{v}$
- the *damping factor* α.

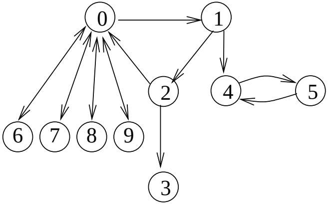
We will focus on the

damping factor

Notation

G = the (adjacency matrix of the) graph

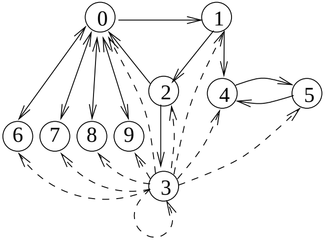
$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$



Notation

$\bar{G}$ = the modified graph to eliminate dangling nodes

$$\begin{pmatrix}
0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}$$



Notation

P = the row-normalized version of $\bar{G}$

$$
\begin{pmatrix}
0 & \frac{1}{5} & 0 & 0 & 0 & 0 & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\
0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
$$

Notation

$$A(\alpha) = \alpha P + (1-\alpha)\mathbf{1}^T\mathbf{v}$$

$$\begin{pmatrix}
\frac{1-\alpha}{10} & \frac{\alpha+1}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{\alpha+1}{10} & \frac{\alpha+1}{10} & \frac{\alpha+1}{10} & \frac{\alpha+1}{10} \\
\frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{4\alpha+1}{10} & \frac{1-\alpha}{10} & \frac{4\alpha+1}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} \\
\frac{4\alpha+1}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{4\alpha+1}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} \\
\frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} \\
\frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{9\alpha+1}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} \\
\frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{9\alpha+1}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} \\
\frac{9\alpha+1}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} \\
\frac{9\alpha+1}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} \\
\frac{9\alpha+1}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} \\
\frac{9\alpha+1}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10} & \frac{1-\alpha}{10}
\end{pmatrix}$$

(Here, we assumed $\mathbf{v} = (1/N)\mathbf{1}$)

PageRank: formal definition

- Is the definition of PageRank well-given?

PageRank: formal definition

- Is the definition of PageRank well-given?
- More formally: is $A(\alpha)$ guaranteed to have *a unique stationary distribution*?

PageRank: formal definition

- Is the definition of PageRank well-given?
- More formally: is $A(\alpha)$ guaranteed to have *a unique stationary distribution*?

Answer: $A(\alpha)$ is irreducible and aperiodic *for all* $\alpha \in [0, 1)$.

PageRank: formal definition

- Is the definition of PageRank well-given?
- More formally: is $A(\alpha)$ guaranteed to have *a unique stationary distribution*?

Answer: $A(\alpha)$ is irreducible and aperiodic *for all* $\alpha \in [0, 1)$. Hence, in that case, it has a unique limit distribution $\mathbf{r}(\alpha)$:

$$\mathbf{r}(\alpha)A(\alpha) = \mathbf{r}(\alpha)$$

PageRank: formal definition

- Is the definition of PageRank well-given?
- More formally: is $A(\alpha)$ guaranteed to have *a unique stationary distribution*?

Answer: $A(\alpha)$ is irreducible and aperiodic *for all* $\alpha \in [0, 1)$. Hence, in that case, it has a unique limit distribution $\mathbf{r}(\alpha)$:

$$\mathbf{r}(\alpha)A(\alpha) = \mathbf{r}(\alpha)$$

$\mathbf{r}(\alpha)$ is the PageRank vector, as a (vector) function of α:

$$\mathbf{r} : [0, 1) \rightarrow [0, 1]^n.$$

The magic value $\alpha = 0.85$

One usually computes and considers only $\mathbf{r}(0.85)$. Why 0.85?

The magic value $\alpha = 0.85$

One usually computes and considers only $\mathbf{r}(0.85)$. Why 0.85?

- "the smart guys at Google use 0.85" (???)

The magic value $\alpha = 0.85$

One usually computes and considers only $\mathbf{r}(0.85)$. Why 0.85?

- "the smart guys at Google use 0.85" (???)
- it works pretty well

The magic value $\alpha = 0.85$

One usually computes and considers only $\mathbf{r}(0.85)$. Why 0.85?

- "the smart guys at Google use 0.85" (???)
- it works pretty well
- iterative algorithms that approximate PageRank converge quickly if $\alpha = 0.85$: larger values would require more iterations; moreover...

The magic value $\alpha = 0.85$

One usually computes and considers only $\mathbf{r}(0.85)$. Why 0.85?

- "the smart guys at Google use 0.85" (???)
- it works pretty well
- iterative algorithms that approximate PageRank converge quickly if $\alpha = 0.85$: larger values would require more iterations; moreover...
- ...numeric instability arises when α is too close to 1...

Is that it?

Yet, we believe that understanding how $\mathbf{r}(\alpha)$ changes when α is modified is important.

Is that it?

Yet, we believe that understanding how $\mathbf{r}(\alpha)$ changes when α is modified is important.

- Basic knowledge about PageRank nature.

Is that it?

Yet, we believe that understanding how $\mathbf{r}(\alpha)$ changes when α is modified is important.

- Basic knowledge about PageRank nature.
- Variations of PageRank to obtain better/alternative ranking techniques (such as TotalRank [Boldi, '05]).

Is that it?

Yet, we believe that understanding how $\mathbf{r}(\alpha)$ changes when α is modified is important.

- Basic knowledge about PageRank nature.
- Variations of PageRank to obtain better/alternative ranking techniques (such as TotalRank [Boldi, '05]).
- Tool for link-spam detection [Zhang et al., '04].

Previous related work

Previous related work

- PageRank (values and rankings) change significantly when α is modified [Pretto, '02; Langville & Meyer., '04]

Previous related work

- PageRank (values and rankings) change significantly when α is modified [Pretto, '02; Langville & Meyer., '04]
- Convergence rate of the Power Method is α [Haveliwala & Kamvar, '03]

Previous related work

- PageRank (values and rankings) change significantly when α is modified [Pretto, '02; Langville & Meyer., '04]
- Convergence rate of the Power Method is α [Haveliwala & Kamvar, '03]
- The condition number of the PageRank problem is $(1+\alpha)/(1-\alpha)$ [Haveliwala & Kamvar, '03]

Explicit formula for PageRank

Explicitly [Haveliwala et al., '03]:

$$\mathbf{r}(\alpha) = (1-\alpha)\mathbf{v}(I - \alpha P)^{-1}$$

Explicit formula for PageRank

Explicitly [Haveliwala et al., '03]:

$$\mathbf{r}(\alpha) = (1-\alpha)\mathbf{v}(I-\alpha P)^{-1}$$

Hence, $\mathbf{r}(\alpha)$ is a rational function of α, with no singularities on $[0,1)$...

Explicit formula for PageRank

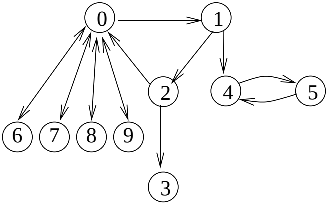
Explicitly [Haveliwala et al., '03]:

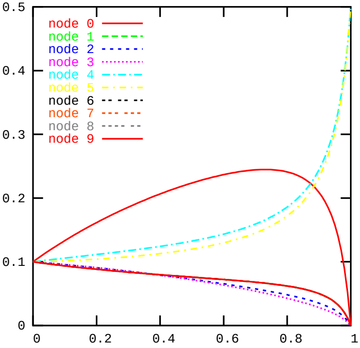
$$\mathbf{r}(\alpha) = (1-\alpha)\mathbf{v}(I-\alpha P)^{-1}$$

Hence, $\mathbf{r}(\alpha)$ is a rational function of α, with no singularities on $[0, 1)$...

...it can be extended by continuity to $[0, 1]$.

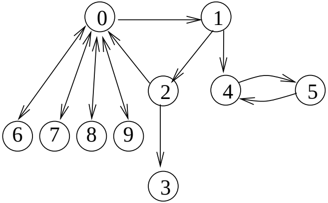
Our example

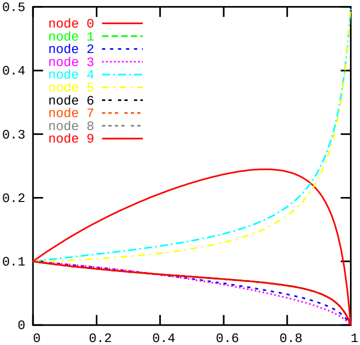




$$r_0(\alpha) = -5\,\frac{(-1+\alpha)\left(\alpha^2+18\,\alpha+4\right)}{8\,\alpha^4+\alpha^3-170\,\alpha^2-20\,\alpha+200}$$

Our example





$$r_1(\alpha) = -2\,\frac{(-1+\alpha)\left(\alpha^2+2\,\alpha+10\right)}{8\,\alpha^4+\alpha^3-170\,\alpha^2-20\,\alpha+200}$$

Limit cases: $\alpha \to 1$

$$\lim_{\alpha \to 1} A(\alpha) = P.$$

Limit cases: $\alpha \to 1$

$$\lim_{\alpha \to 1} A(\alpha) = P.$$

The "uniform" part added to P vanishes, whereas the part due to $\bar{G}$ becomes larger: some interpret this fact as a hint that $\mathbf{r}(\alpha)$ becomes "more faithful to reality" when $\alpha \to 1$.

Limit cases: $\alpha \longrightarrow 1$

$$\lim_{\alpha \to 1} A(\alpha) = P.$$

The "uniform" part added to P vanishes, whereas the part due to $\bar{G}$ becomes larger: some interpret this fact as a hint that $\mathbf{r}(\alpha)$ becomes "more faithful to reality" when $\alpha \to 1$.
Is this true?

Limit cases: $\alpha \to 1$

$$\lim_{\alpha \to 1} A(\alpha) = P.$$

The "uniform" part added to P vanishes, whereas the part due to $\bar{G}$ becomes larger: some interpret this fact as a hint that $\mathbf{r}(\alpha)$ becomes "more faithful to reality" when $\alpha \to 1$.
Is this true?
Since $\mathbf{r}(-)$ can be extended by continuity to $[0, 1]$, let:

$$\mathbf{r}^* = \lim_{\alpha \to 1^-} \mathbf{r}(\alpha).$$

Limit cases: $\alpha \to 1$

$$\lim_{\alpha \to 1} A(\alpha) = P.$$

The "uniform" part added to P vanishes, whereas the part due to $\bar{G}$ becomes larger: some interpret this fact as a hint that $\mathbf{r}(\alpha)$ becomes "more faithful to reality" when $\alpha \to 1$.
Is this true?
Since $\mathbf{r}(-)$ can be extended by continuity to $[0, 1]$, let:

$$\mathbf{r}^* = \lim_{\alpha \to 1^-} \mathbf{r}(\alpha).$$

$\mathbf{r}^*$ is a stationary distribution of P, i.e.,

$$\mathbf{r}^* P = \mathbf{r}^*$$

Limit cases: $\alpha \to 1$

$$\lim_{\alpha \to 1} A(\alpha) = P.$$

The "uniform" part added to P vanishes, whereas the part due to $\bar{G}$ becomes larger: some interpret this fact as a hint that $\mathbf{r}(\alpha)$ becomes "more faithful to reality" when $\alpha \to 1$.
Is this true?
Since $\mathbf{r}(-)$ can be extended by continuity to $[0, 1]$, let:

$$\mathbf{r}^* = \lim_{\alpha \to 1^-} \mathbf{r}(\alpha).$$

$\mathbf{r}^*$ is a stationary distribution of P, i.e.,

$$\mathbf{r}^* P = \mathbf{r}^*$$

It is not *unique*, though, unless P itself is aperiodic and irreducible.

From the point of view of the Markov chain

We would like to compute

$$\lim_{\alpha \to 1^-} \lim_{n \to \infty} (\alpha P + (1-\alpha)\mathbf{1}^T \mathbf{v})^n$$

From the point of view of the Markov chain

We would like to compute

$$\lim_{\alpha \to 1^-} \lim_{n \to \infty} (\alpha P + (1 - \alpha)\mathbf{1}^T \mathbf{v})^n$$

It could be tempting to exchange the limits (*i.e.* hope for uniform convergence), but...

From the point of view of the Markov chain

We would like to compute

$$\lim_{\alpha \to 1^-} \lim_{n \to \infty} (\alpha P + (1 - \alpha)\mathbf{1}^T \mathbf{v})^n$$

It could be tempting to exchange the limits (*i.e.* hope for uniform convergence), but...

- $A(\alpha) = \alpha P + (1 - \alpha)\mathbf{1}^T \mathbf{v}$ is irreducible for every $\alpha \in [0, 1)$,

From the point of view of the Markov chain

We would like to compute

$$\lim_{\alpha \to 1^-} \lim_{n \to \infty} (\alpha P + (1 - \alpha)\mathbf{1}^T \mathbf{v})^n$$

It could be tempting to exchange the limits (*i.e.* hope for uniform convergence), but. . .

- $A(\alpha) = \alpha P + (1 - \alpha)\mathbf{1}^T \mathbf{v}$ is irreducible for every $\alpha \in [0, 1)$, so $\lim_{n \to \infty} A(\alpha)^n$ is *stable*. . .

From the point of view of the Markov chain

We would like to compute

$$\lim_{\alpha \to 1^-} \lim_{n \to \infty} (\alpha P + (1-\alpha)\mathbf{1}^T \mathbf{v})^n$$

It could be tempting to exchange the limits (*i.e.* hope for uniform convergence), but. . .

- $A(\alpha) = \alpha P + (1-\alpha)\mathbf{1}^T \mathbf{v}$ is irreducible for every $\alpha \in [0,1)$, so $\lim_{n\to\infty} A(\alpha)^n$ is *stable*. . .
- on the other hand, $P = \lim_{\alpha \to 1^-} A(\alpha)$ can be reducible,

From the point of view of the Markov chain

We would like to compute

$$\lim_{\alpha \to 1^-} \lim_{n \to \infty} (\alpha P + (1-\alpha)\mathbf{1}^T \mathbf{v})^n$$

It could be tempting to exchange the limits (*i.e.* hope for uniform convergence), but...

- $A(\alpha) = \alpha P + (1-\alpha)\mathbf{1}^T \mathbf{v}$ is irreducible for every $\alpha \in [0, 1)$, so $\lim_{n\to\infty} A(\alpha)^n$ is *stable*...
- on the other hand, $P = \lim_{\alpha \to 1^-} A(\alpha)$ can be reducible, so $\lim_{n\to\infty} P^n$ needs not to be stable.

Some technicalities...

For a generic Markov chain P, the limit behaviour of P^n is

$$\lim_{n\to\infty} (P^{(nd_j+r)})_{i,j} = \pi^{(r)}_{i,j}$$

where d_j is the *period* of node j.

Some technicalities...

For a generic Markov chain P, the limit behaviour of P^n is

$$\lim_{n\to\infty} (P^{(nd_j+r)})_{i,j} = \pi^{(r)}_{i,j}$$

where d_j is the *period* of node j.

At least componentwise, these are converging subsequences, hence it is possible to define the *Cesaro* limit:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{0<t\le n} P^t = \Pi, \qquad (\Pi)_{i,j} = \frac{1}{d_j} \sum_{1\le r\le d_j} \pi^{(r)}_{i,j}.$$

Back to the limit

Two easy Lemmas:

Back to the limit

Two easy Lemmas:

- $\lim_{n\to\infty} A(\alpha)^n = \mathbf{1}^T \mathbf{v}(1-\alpha)\sum_{n\geq 0}(\alpha P)^n$

Back to the limit

Two easy Lemmas:

- $\lim_{n\to\infty} A(\alpha)^n = \mathbf{1}^T \mathbf{v}(1-\alpha)\sum_{n\geq 0}(\alpha P)^n$
- $(1-\alpha)\sum_{n\geq 0}\alpha^n a_n \sim a$ for every $a_n \to a$ (as $\alpha \to 1^-$)

Back to the limit

Two easy Lemmas:

- $\lim_{n\to\infty} A(\alpha)^n = \mathbf{1}^T\mathbf{v}(1-\alpha)\sum_{n\geq 0}(\alpha P)^n$
- $(1-\alpha)\sum_{n\geq 0}\alpha^n a_n \sim a$ for every $a_n \to a$ (as $\alpha \to 1^-$)

By the limiting behaviour of subsequences and the Cesaro limit

$$\lim_{\alpha\to 1^-}\lim_{n\to\infty} A(\alpha)^n = \mathbf{1}^T\mathbf{v}\Pi$$

Back to the limit

Two easy Lemmas:

- $\lim_{n\to\infty} A(\alpha)^n = \mathbf{1}^T \mathbf{v}(1-\alpha)\sum_{n\geq 0}(\alpha P)^n$
- $(1-\alpha)\sum_{n\geq 0}\alpha^n a_n \sim a$ for every $a_n \to a$ (as $\alpha \to 1^-$)

By the limiting behaviour of subsequences and the Cesaro limit

$$\lim_{\alpha\to 1^-}\lim_{n\to\infty} A(\alpha)^n = \mathbf{1}^T \mathbf{v}\Pi$$

or, from the point of view of the limit distribution,

$$\mathbf{r}^* = \lim_{\alpha\to 1^-}\lim_{n\to\infty} \mathbf{p}_0 A(\alpha)^n = \mathbf{v}\Pi.$$

(whatever the initial $\mathbf{p}_0$).

So what are the consequences on the limit probability?

To understand what happens when $\alpha \to 1$, we can consider an (arbitrary) stationary distribution $\mathbf{t}$ of P.

So what are the consequences on the limit probability?

To understand what happens when $\alpha \to 1$, we can consider an (arbitrary) stationary distribution $\mathbf{t}$ of P.

Of course $t_x = 0$ for all transient nodes x of P.

So what are the consequences on the limit probability?

To understand what happens when $\alpha \to 1$, we can consider an (arbitrary) stationary distribution $\mathbf{t}$ of P.

Of course $t_x = 0$ for all transient nodes x of P.

A node x is *recurrent* iff its strongly connected component in $\bar{G}$ has no outgoing arcs (except possibly for loops).

So what are the consequences on the limit probability?

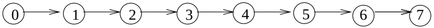
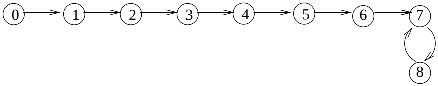
To understand what happens when $\alpha \to 1$, we can consider an (arbitrary) stationary distribution $\mathbf{t}$ of P.

Of course $t_x = 0$ for all transient nodes x of P.

A node x is *recurrent* iff its strongly connected component in $\bar{G}$ has no outgoing arcs (except possibly for loops).

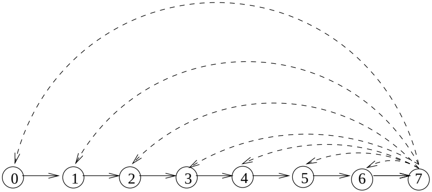
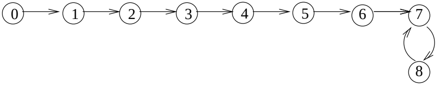
So what are the components of $\bar{G}$ with no outgoing arcs?

Components of $\bar{G}$

G_1	
G_2	

Two graphs: they look quite similar, but. . .

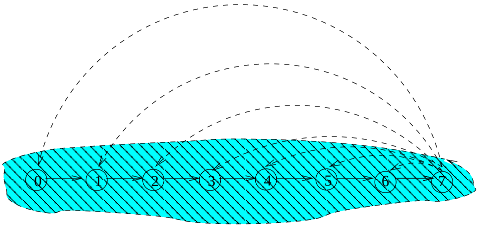
Components of $\bar{G}$

$\bar{G}_1$	
$\bar{G}_2$	

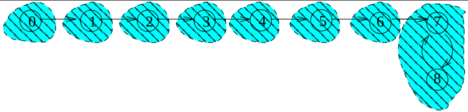
... G_1 has a dangling node, whereas G_2 has none

Components of $\bar{G}$

components of $\bar{G}_1$



components of $\bar{G}_2$



All nodes are recurrent in G_1, whereas all nodes except for 7 and 8 are transient in G_2

A general statement

Theorem

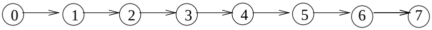
Except for degenerate cases a node is recurrent iff it belongs to a component of G that is nontrivial and has no arcs going out of it.

A general statement

Theorem

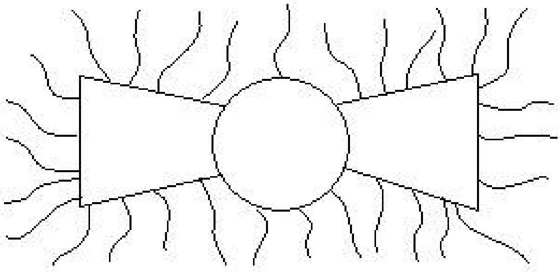
Except for degenerate cases a node is recurrent iff it belongs to a component of G that is nontrivial and has no arcs going out of it.

The degenerate cases happen when there are no components satisfying the above conditions, in which case *every node is recurrent*:



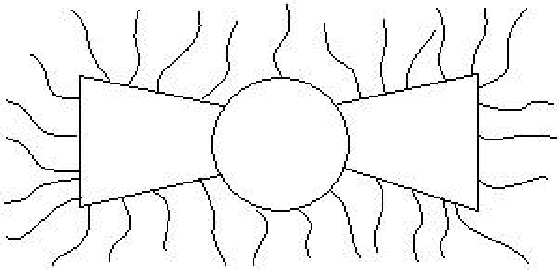
Bowtie

As a consequence, when $\alpha \to 1$, all PageRank concentrates in a bunch of pages that live in the rightmost part of the bowtie [Kumar et al., '00]:



Bowtie

As a consequence, when $\alpha \to 1$, all PageRank concentrates in a bunch of pages that live in the rightmost part of the bowtie [Kumar et al., '00]:



$\mathbf{r}(\alpha)$ becomes meaningless as $\alpha \to 1$!

General behaviour

What about the general behaviour of $\mathbf{r}(\alpha)$?

General behaviour

What about the general behaviour of $\mathbf{r}(\alpha)$?
The following theorem gives an explicit formula for derivatives of PageRank:

Theorem

For every $k > 0$

$$\mathbf{r}^{(k)}(\alpha) = k!\,\mathbf{r}(\alpha)\left(Q(\alpha) - \frac{1}{1-\alpha}I\right)Q(\alpha)^{k-1}.$$

where

$$Q(\alpha) = P(I - \alpha P)^{-1}.$$

Computing the derivatives. . .

. . . is possible using a modification of the standard Power Method algorithm.

Computing the derivatives...

...is possible using a modification of the standard Power Method algorithm.

Theorem

There exists an algorithm that simultaneously approximates $\mathbf{r}(\alpha)$, $\mathbf{r}'(\alpha)$, $\mathbf{r}''(\alpha)$, $\ldots$, $\mathbf{r}^{(k)}(\alpha)$ *(for a fixed value of* $\alpha \in [0, 1)$*).*

Computing the derivatives. . .

. . . is possible using a modification of the standard Power Method algorithm.

Theorem

There exists an algorithm that simultaneously approximates $\mathbf{r}(\alpha)$, $\mathbf{r}'(\alpha)$, $\mathbf{r}''(\alpha)$, $\ldots$, $\mathbf{r}^{(k)}(\alpha)$ *(for a fixed value of* $\alpha \in [0, 1)$*).*

After t iterations, the difference (in norm) between $\mathbf{r}^{(j)}(\alpha)$ and its approximation is $O\left(t^j \alpha^t\right)$.

Computing the derivatives. . .

. . . is possible using a modification of the standard Power Method algorithm.

Theorem

There exists an algorithm that simultaneously approximates $\mathbf{r}(\alpha)$, $\mathbf{r}'(\alpha)$, $\mathbf{r}''(\alpha)$, $\dots$, $\mathbf{r}^{(k)}(\alpha)$ *(for a fixed value of* $\alpha \in [0,1)$*).*

After t iterations, the difference (in norm) between $\mathbf{r}^{(j)}(\alpha)$ and its approximation is $O\left(t^j \alpha^t\right)$.

This algorithm provides *pointwise approximation for arbitary derivatives* with precision guarantee.

Observe that...

By definition, $Q(\alpha) = P(I - \alpha P)^{-1}$, so $Q(0) = P$.

Observe that...

By definition, $Q(\alpha) = P(I - \alpha P)^{-1}$, so $Q(0) = P$.
Hence, since

$$\mathbf{r}^{(k)}(\alpha) = k!\,\mathbf{r}(\alpha)\left(Q(\alpha) - \frac{1}{1-\alpha}I\right)Q(\alpha)^{k-1}.$$

Observe that...

By definition, $Q(\alpha) = P(I - \alpha P)^{-1}$, so $Q(0) = P$.
Hence, since

$$\mathbf{r}^{(k)}(\alpha) = k!\, \mathbf{r}(\alpha) \left(Q(\alpha) - \frac{1}{1-\alpha} I \right) Q(\alpha)^{k-1}.$$

we have

$$\mathbf{r}^{(k)}(0) = k!\, \mathbf{v}\, (P - I)\, P^{k-1}.$$

Observe that...

By definition, $Q(\alpha) = P(I - \alpha P)^{-1}$, so $Q(0) = P$.
Hence, since

$$\mathbf{r}^{(k)}(\alpha) = k!\, \mathbf{r}(\alpha) \left(Q(\alpha) - \frac{1}{1-\alpha} I \right) Q(\alpha)^{k-1}.$$

we have

$$\mathbf{r}^{(k)}(0) = k!\, \mathbf{v}\, (P - I)\, P^{k-1}.$$

or, equivalently,

$$\mathbf{r}^{(k)}(0) = k!\, \mathbf{v} \left(P^k - P^{k-1} \right).$$

Maclaurin expansion

Hence, the Maclaurin expansion of $\mathbf{r}(\alpha)$, that is $\mathbf{r}(\alpha) = \sum_{k=0}^{\infty} \frac{\mathbf{r}^{(k)}(0)}{k!}\alpha^k$ can be written as

$$\mathbf{r}(\alpha) = \sum_{k=0}^{\infty} \mathbf{v}\left(P^k - P^{k-1}\right)\alpha^k.$$

Maclaurin expansion

Hence, the Maclaurin expansion of $\mathbf{r}(\alpha)$, that is $\mathbf{r}(\alpha) = \sum_{k=0}^{\infty} \frac{\mathbf{r}^{(k)}(0)}{k!}\alpha^k$ can be written as

$$\mathbf{r}(\alpha) = \sum_{k=0}^{\infty} \mathbf{v}\left(P^k - P^{k-1}\right)\alpha^k.$$

It can be shown that

Theorem

Let $\mathbf{R}_0, \mathbf{R}_1, \ldots$ be the approximations of PageRank computed by the Power Method for a certain value, say α_0, of the damping factor.

Maclaurin expansion

Hence, the Maclaurin expansion of $\mathbf{r}(\alpha)$, that is $\mathbf{r}(\alpha) = \sum_{k=0}^{\infty} \frac{\mathbf{r}^{(k)}(0)}{k!}\alpha^k$ can be written as

$$\mathbf{r}(\alpha) = \sum_{k=0}^{\infty} \mathbf{v}\left(P^k - P^{k-1}\right)\alpha^k.$$

It can be shown that

Theorem

Let $\mathbf{R}_0, \mathbf{R}_1, \ldots$ be the approximations of PageRank computed by the Power Method for a certain value, say α_0, of the damping factor.

Then, for all k,

$$\mathbf{v}\left(P^k - P^{k-1}\right) = \frac{R_k - R_{k-1}}{\alpha_0^k}.$$

(by convention, $R_{-1} = \mathbf{0}$)

As a consequence

Using a simple variant of the standard Power Method, one can compute the coefficients of the Maclaurin expansion of $\mathbf{r}(\alpha)$:

As a consequence

Using a simple variant of the standard Power Method, one can compute the coefficients of the Maclaurin expansion of $\mathbf{r}(\alpha)$:

- t iterations $\Longrightarrow$ Maclaurin polynomial of degree t

As a consequence

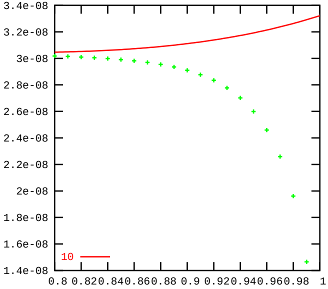
Using a simple variant of the standard Power Method, one can compute the coefficients of the Maclaurin expansion of $\mathbf{r}(\alpha)$:

- t iterations $\Longrightarrow$ Maclaurin polynomial of degree t
- you can approximate *simultaneously* PageRank on all possible values of $\alpha \in [0, 1)$.

As a consequence

Using a simple variant of the standard Power Method, one can compute the coefficients of the Maclaurin expansion of $\mathbf{r}(\alpha)$:

- t iterations $\implies$ Maclaurin polynomial of degree t
- you can approximate *simultaneously* PageRank on all possible values of $\alpha \in [0, 1)$.

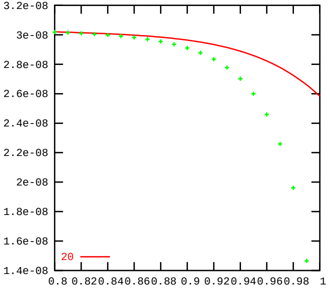


degree 10

As a consequence

Using a simple variant of the standard Power Method, one can compute the coefficients of the Maclaurin expansion of $\mathbf{r}(\alpha)$:

- *t* iterations $\Longrightarrow$ Maclaurin polynomial of degree *t*
- you can approximate *simultaneously* PageRank on all possible values of $\alpha \in [0, 1)$.

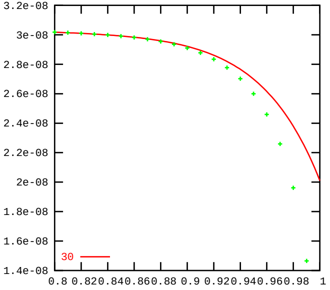


degree 20

As a consequence

Using a simple variant of the standard Power Method, one can compute the coefficients of the Maclaurin expansion of $\mathbf{r}(\alpha)$:

- t iterations $\Longrightarrow$ Maclaurin polynomial of degree t
- you can approximate *simultaneously* PageRank on all possible values of $\alpha \in [0, 1)$.

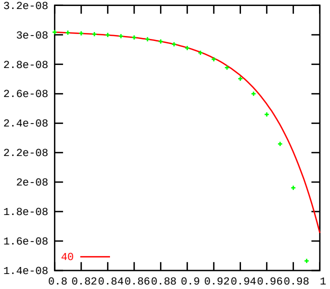


degree 30

As a consequence

Using a simple variant of the standard Power Method, one can compute the coefficients of the Maclaurin expansion of $\mathbf{r}(\alpha)$:

- t iterations $\Longrightarrow$ Maclaurin polynomial of degree t
- you can approximate *simultaneously* PageRank on all possible values of $\alpha \in [0, 1)$.

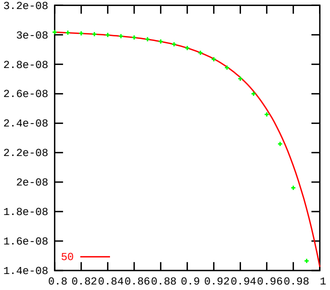


degree 40

As a consequence

Using a simple variant of the standard Power Method, one can compute the coefficients of the Maclaurin expansion of $\mathbf{r}(\alpha)$:

- t iterations $\Longrightarrow$ Maclaurin polynomial of degree t
- you can approximate *simultaneously* PageRank on all possible values of $\alpha \in [0, 1)$.

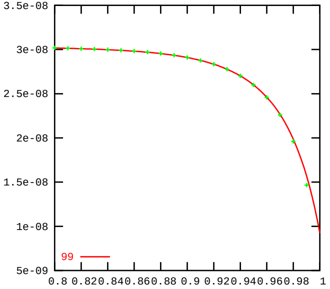


degree 50

As a consequence

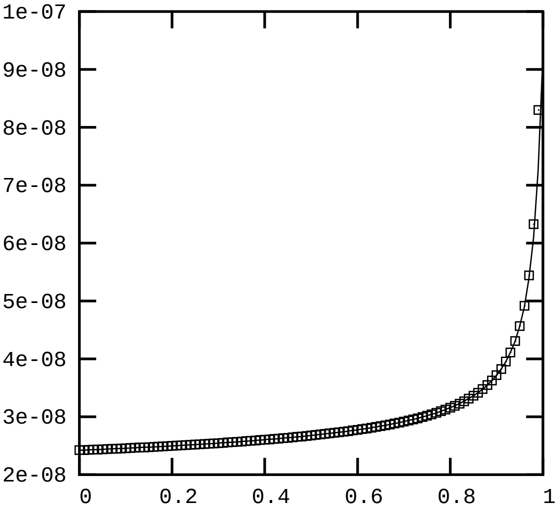
Using a simple variant of the standard Power Method, one can compute the coefficients of the Maclaurin expansion of $\mathbf{r}(\alpha)$:

- t iterations $\Longrightarrow$ Maclaurin polynomial of degree t
- you can approximate *simultaneously* PageRank on all possible values of $\alpha \in [0, 1)$.

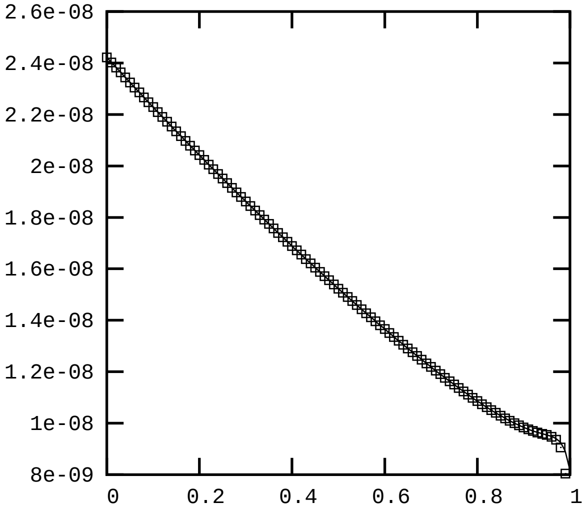


degree 99

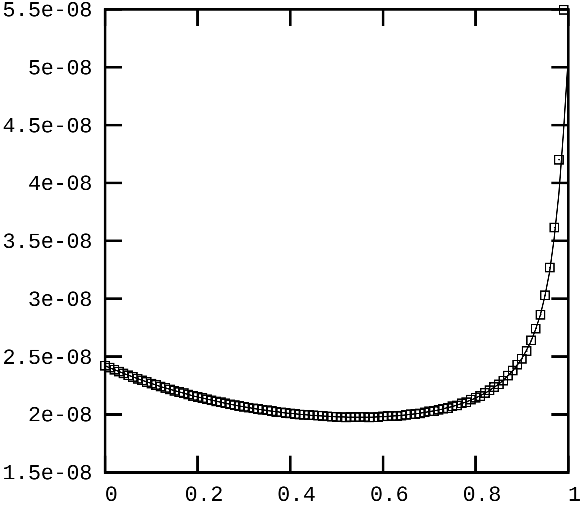
Some typical behaviours



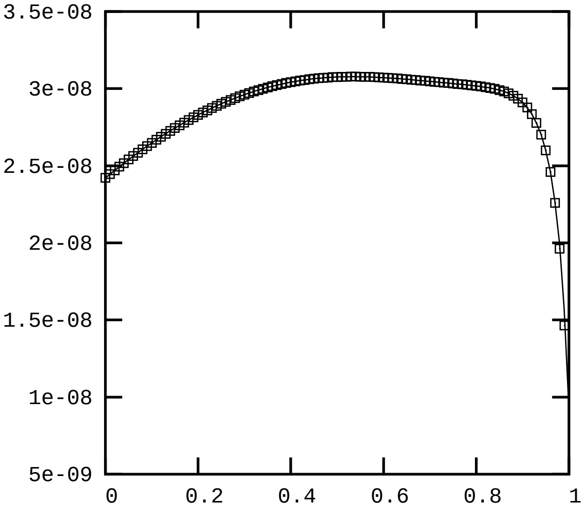
Some typical behaviours



Some typical behaviours



Some typical behaviours



Conclusions

Future work...

Conclusions

Future work...

- Practical implications (?):

Conclusions

Future work...

- Practical implications (?):
  - Is it possible to use the algorithms/properties described to detect/eliminate link spam? ...

Conclusions

Future work. . .

- Practical implications (?):
  - Is it possible to use the algorithms/properties described to detect/eliminate link spam? . . .
  - . . . or to obtain new (better?) ranking techniques?

Conclusions

Future work...

- Practical implications (?):
  - Is it possible to use the algorithms/properties described to detect/eliminate link spam? ...
  - ...or to obtain new (better?) ranking techniques?
- On a more theoretical side:

Conclusions

Future work...

- Practical implications (?):
  - Is it possible to use the algorithms/properties described to detect/eliminate link spam? ...
  - ...or to obtain new (better?) ranking techniques?
- On a more theoretical side:
  - Is there a (deep and yet undiscovered) analytical reason behind the magical value 0.85?

Conclusions

Future work. . .

- Practical implications (?):
  - Is it possible to use the algorithms/properties described to detect/eliminate link spam? . . .
  - . . . or to obtain new (better?) ranking techniques?
- On a more theoretical side:
  - Is there a (deep and yet undiscovered) analytical reason behind the magical value 0.85?

The "Laboratory for Web Algorithmics"

`http://law.dsi.unimi.it/`

provides (free) datasets and (GNU GPL) code for dealing with large web graphs and compute PageRank and friends. . .