

PageRank

Presenting Theoretical Computer Science Seminar FS26

Ben Kracht, Thorben Klabunde

April 25, 2026

1 Motivation

2 Formalization

3 Background: Markov Chains

4 Algorithm: Correctness & Convergence

5 Implementation, Complexity & Alternatives

6 Modifications

Motivation

Why PageRank?

December 2000

Top Global Web Properties

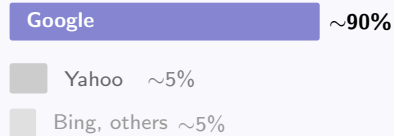
1. Yahoo! 49.9% reach
 2. AOL 43.9% reach
 3. MSN 41.9% reach
 4. Lycos 26.7% reach
 5. Microsoft 25.8% reach
- ... Google not in the top 10.

PageRank



2009

Search Engine Market Share



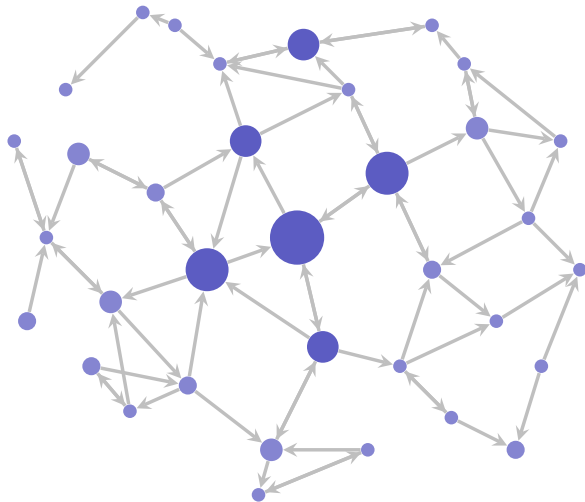
Yahoo's dominance: gone.

Source: Nielsen/NetRatings, Dec. 2000

Source: StatCounter Global Stats, 2009

In 2000, analysts predicted Yahoo would dominate the web.

The Web as a Graph

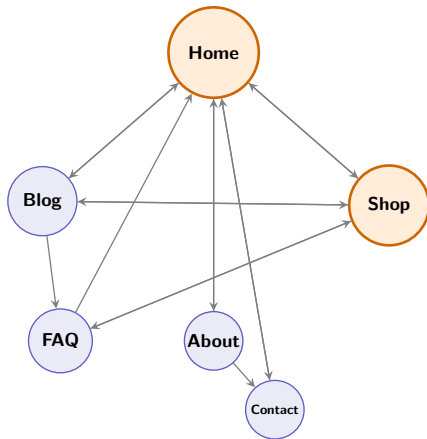


The Web can be modelled as a directed graph $\mathcal{G} = (V, E)$:

- Nodes V = web pages
- Edges E = hyperlinks
- $(u, v) \in E$ if page u links to v

Question: Given $\mathcal{G}$, how do we determine which pages are most *relevant*?

The Key Idea



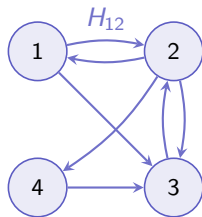
Random surfer who:

- 1 With probability α : follows a random outgoing hyperlink
- 2 With probability $1 - \alpha$: “teleports” to a random page

The **PageRank** of a page represents the time a random surfer spends there.

Formalization

Hyperlink Matrix H

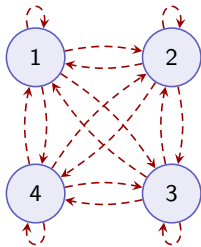


Hyperlinks only

H_{ij} = probability of following a hyperlink from page i to page j .

$$H = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Teleportation Matrix T



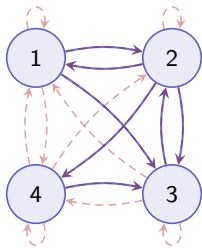
Teleportation (fully connected)

The surfer can “teleport” to *any* page at random.

With uniform teleportation:

$$T = \frac{1}{n} \mathbf{1}\mathbf{1}^\top = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$

The Google Matrix G



→ Hyperlink (α)

- - -> Teleport ($1-\alpha$)

The Google matrix combines both:

$$G = \alpha H + (1 - \alpha) T$$

where $\alpha \in (0, 1)$ is the *damping factor*.

(Standard choice: $\alpha = 0.85$ (Brin & Page, 1998))

The Stationary Distribution

Given an initial distribution $x^{(0)}$, the state after k steps is:

$$(x^{(k)})^\top = (x^{(0)})^\top G^k$$

since at each step:

$$\Pr[\text{page } i \text{ after } k \text{ clicks}] = \sum_{j=1}^n \Pr[\text{page } j \text{ after } k-1 \text{ clicks}] \cdot G_{ji}$$

Central Question

Does this process converge to a unique stationary distribution x^* satisfying

$$(x^*)^\top = (x^*)^\top G \quad ?$$

Two Formulations (Page & Brin, 1998)

- 1 **Power method:** Compute the stationary distribution iteratively.
- 2 **Eigenvector problem:** Find the left eigenvector of G with eigenvalue 1.

Both can be used to compute the **PageRank vector**, but the **power method** remains dominant in practice due to computational efficiency.

Background

Modelling PageRank as a Markov Chain

Definition 3.1 (Markov Chain)

A *Markov Chain* is a stochastic process where the probability of transitioning to the next state depends **only** on the current state, and is independent of the past.

PageRank Model

- Web pages V
- Google Matrix G
- Surfer's distribution $x^{(k)}$
- PageRank vector x^*



Markov Chain

- State space $\mathcal{S}$
- Transition Matrix P
- State probability at k
- Stationary distribution π

Required Properties for Convergence

1. Irreducibility

Guarantees the *uniqueness* of x^*

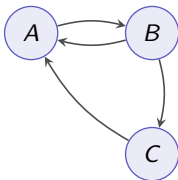
2. Aperiodicity

Guarantees *convergence* to x^*

Irreducibility: The trapped Surfer

Irreducibility (Def 3.3)

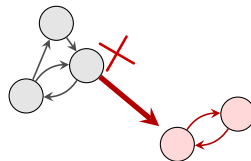
A chain is *irreducible* if the corresponding graph is strongly connected ($i \leftrightarrow j \forall i, j$).



Every state can reach every other state.

Problem (Sinks/Leaking Mass)

If pages form a dead-end, the random surfer gets permanently trapped, draining all probability mass from the rest of the web.



Probability Leaks Here

Aperiodicity: Endless Surfer Oscillations

Definition 3.4 (Periodicity)

The *period* of state i is
$$d(i) = \gcd\{k \geq 1 \mid (G^k)_{ii} > 0\}.$$

Aperiodicity (Def 3.4/3.5)

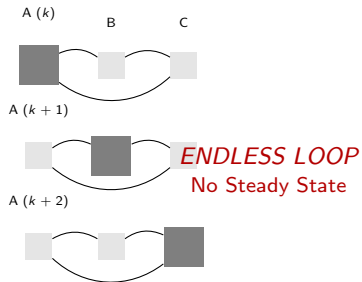
A state is *aperiodic* if the period of the state is 1. A sufficient condition is if the return time is 1. Periodicity forbids convergence.



Period = 2. Return times: $\{2, 4, 6, \dots\}$

Problem: Oscillation

If a graph structure forces the surfer into a rigid cycle, the probability distribution will alternate endlessly instead of converging.

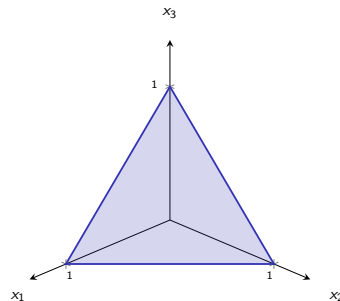


Probability Simplex

Definition 3.6 (Probability Simplex)

The *probability simplex* over n states is the set:

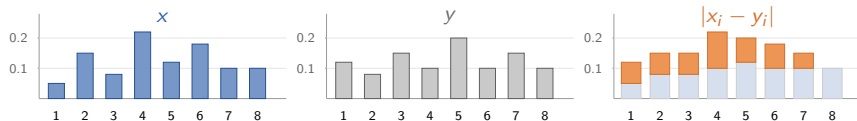
$$\mathbb{P}^n = \left\{ x \in [0, 1]^n \mid \sum_{i=1}^n x_i = 1 \right\}$$



Definition 3.7 (Total Variation)

For $x, y \in \mathbb{P}^n$, the *total variation distance* is:

$$\|x - y\|_{TV} = \frac{1}{2} \sum_{i=1}^n |x_i - y_i|$$



Definition 3.9 (Stochastic Matrix)

Let P be an $n \times n$ matrix with entries p_{ij} . We call P a (row-)stochastic matrix if:

- ① **Non-negativity:** $p_{ij} \geq 0$ for all $i, j \in \{1, \dots, n\}$
- ② **Row sums equal 1:** $\sum_{j=1}^n p_{ij} = 1$ for all $i \in \{1, \dots, n\}$

In a Markov Chain, p_{ij} is the transition probability from state i to state j .

Example ($n = 3$):

$$P = \begin{pmatrix} 0.2 & 0.5 & 0.3 \\ 0.1 & 0.7 & 0.2 \\ 0.4 & 0.1 & 0.5 \end{pmatrix} \quad \begin{array}{l} \rightarrow 0.2 + 0.5 + 0.3 = 1 \checkmark \\ \rightarrow 0.1 + 0.7 + 0.2 = 1 \checkmark \\ \rightarrow 0.4 + 0.1 + 0.5 = 1 \checkmark \end{array}$$

Convergence

Algorithm 1 *

PageRank via Power Iteration

Require: Stochastic matrix $G \in \mathbb{R}^{n \times n}$, tolerance $\varepsilon > 0$

1: $x \leftarrow \left(\frac{1}{n}, \dots, \frac{1}{n}\right)^\top$

▷ we start with uniform

2: **repeat**

3: $x' \leftarrow G^\top x$

4: $\delta \leftarrow \|x' - x\|_{TV}$

5: $x \leftarrow x'$

6: **until** $\delta < \varepsilon$

7: **return** x

Convergence Theorem

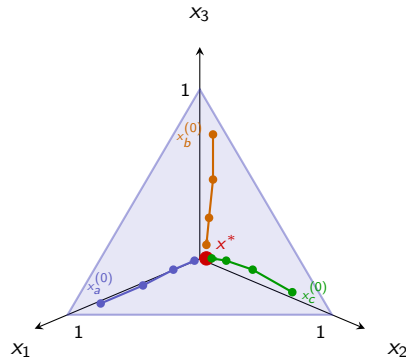
Theorem 4.1

Let G be a stochastic matrix that is irreducible, aperiodic, and has only strictly positive entries. Then:

- 1 G has a **unique** stationary distribution x^* satisfying $(x^*)^\top = (x^*)^\top G$
- 2 For **any** initial distribution $x^{(0)} \in \mathbb{P}^n$, the iteration

$$(x^{(k)})^\top = (x^{(0)})^\top G^k$$

converges to x^*



Convergence: Proof Strategy (1/2)

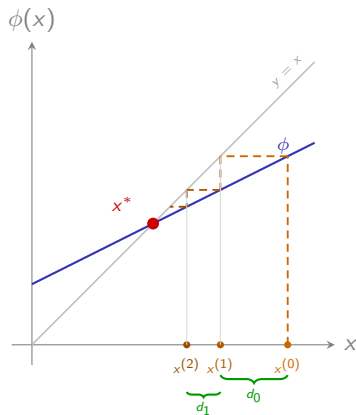
Goal: Show that the power iteration always converges to a *unique* stationary distribution x^*

Plan:

- 1 Decompose G using **Doeblin's decomposition**:

$$G = (1 - \delta) Q + \delta \Pi$$

- 2 Show that the mapping $\phi: x^\top \mapsto x^\top G$ is a **contraction** on $(\mathbb{P}^n, \|\cdot\|_{TV})$



Goal: Show that the power iteration always converges to a *unique* stationary distribution x^*

Plan:

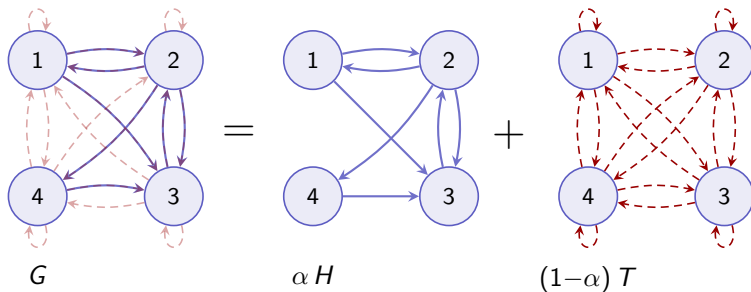
- 1 Decompose G using **Doeblin's decomposition**:

$$G = (1 - \delta) Q + \delta \Pi$$

- 2 Show that the mapping $\phi: x^\top \mapsto x^\top G$ is a **contraction** on $(\mathbb{P}^n, \|\cdot\|_{TV})$
- 3 Apply **Banach's Fixed Point Theorem** to obtain existence, uniqueness, and convergence

Intuition for the Contraction Argument

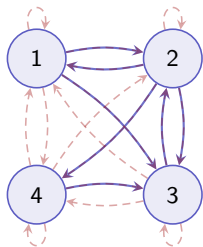
Recall the set-up of the Google matrix as: $G = \alpha H + (1 - \alpha) T$



Doebelin Decomposition: Google Matrix

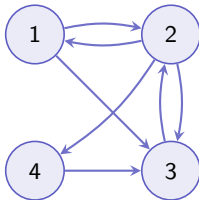
Surfer performs a biased coin flip:

- **(Prob α):** Follow the state-dependent probabilities (H).
- **(Prob $1 - \alpha$):** Forget current state entirely. Jump to a new page using teleportation (T).



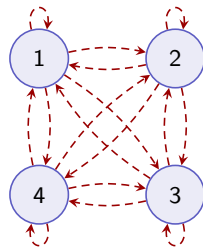
G

$=$



αH

$+$



$(1-\alpha) T$

T positive $\Rightarrow G$ positive: $G_{ij} > 0 \ \forall i, j$

Doeblin Decomposition: 2×2 Example

Let $G = \begin{pmatrix} 0.65 & 0.35 \\ 0.25 & 0.75 \end{pmatrix}$. We want $G = (1 - \delta)Q + \delta\Pi$, where $\Pi = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}$.

Decomposition:

$$\underbrace{\begin{pmatrix} 0.65 & 0.35 \\ 0.25 & 0.75 \end{pmatrix}}_G = \begin{pmatrix} 0.4 & 0.1 \\ 0 & 0.5 \end{pmatrix} + \begin{pmatrix} 0.25 & 0.25 \\ 0.25 & 0.25 \end{pmatrix}$$
$$= \underbrace{0.5}_{(1-\delta)} \underbrace{\begin{pmatrix} 0.8 & 0.2 \\ 0 & 1 \end{pmatrix}}_Q + \underbrace{0.5}_{\delta} \underbrace{\begin{pmatrix} 0.50 & 0.50 \\ 0.50 & 0.50 \end{pmatrix}}_{\Pi}$$

Tie to Contraction Mapping

We have decomposed the transition: $G = (1 - \delta)Q + \delta\Pi$.

Imagine tracking two different probability distributions, x and $y \in \mathbb{P}^n$, after one step:

$$x^\top G = (1 - \delta)x^\top Q + \delta x^\top \Pi$$

$$y^\top G = (1 - \delta)y^\top Q + \delta y^\top \Pi$$

The Magic of the Uniform Baseline

$$x^\top \Pi = y^\top \Pi = \mu^\top$$

2x2 example ($\mu = (\frac{1}{2}, \frac{1}{2})$)

$$\Pi = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}, \quad x^\top = (1, 0), \quad y^\top = (\frac{1}{2}, \frac{1}{2})$$

$$x^\top \Pi = (\frac{1}{2}, \frac{1}{2}) = y^\top \Pi = \mu^\top$$

Step 1: Decompose G as follows.

For simplicity, let $\mu \in \mathbb{P}^n$ denote the uniform distribution ($\mu_i = \frac{1}{n}$, $1 \leq i \leq n$) and let

$$\Pi = \mathbf{1} \mu^\top = \begin{pmatrix} - & \mu^\top & - \\ - & \mu^\top & - \\ & \vdots & \\ - & \mu^\top & - \end{pmatrix}$$

Since G is positive, $\exists \delta > 0$ such that:

$$G_{ij} \geq \delta \cdot \Pi_{ij} \quad \forall i, j \quad \Longrightarrow \quad \boxed{G = (1-\delta) Q + \delta \Pi}$$

such that Q is stochastic.

Contraction Mapping

Step 2: To show that $\phi: x^\top \mapsto x^\top G$ is a contraction on $(\mathbb{P}^n, \|\cdot\|_{TV})$, let $x, y \in \mathbb{P}^n$ be arbitrary:

$$\begin{aligned}\|x^\top G - y^\top G\|_{TV} &= \|(1-\delta)(x^\top Q - y^\top Q) + \underbrace{\delta(x^\top \Pi - y^\top \Pi)}_{=\mu^\top - \mu^\top = 0}\|_{TV} \\ &= (1-\delta) \|(x^\top - y^\top) Q\|_{TV} \\ &\leq (1-\delta) \|x^\top - y^\top\|_{TV} \quad (\text{non-expansiveness of } Q)\end{aligned}$$

Proposition 3.11 (Non-Expansiveness under Stochastic Matrices)

Let $Q \in \mathbb{R}^{n \times n}$ be a stochastic matrix. Then for any probability distributions $x, y \in \mathbb{P}^n$,

$$\|x^\top Q - y^\top Q\|_{TV} \leq \|x^\top - y^\top\|_{TV}.$$

Multiplying by a stochastic matrix can never increase the TV between two distributions.

Proposition 3.11 (Non-Expansiveness under Stochastic Matrices)

Let $Q \in \mathbb{R}^{n \times n}$ be a stochastic matrix. Then for any probability distributions $x, y \in \mathbb{P}^n$,

$$\|x^\top Q - y^\top Q\|_{TV} \leq \|x^\top - y^\top\|_{TV}.$$

Multiplying by a stochastic matrix can never increase the total variation distance between two distributions.

Non-Expansiveness: Proof

Let $z = x^\top - y^\top$. We show $\|zQ\|_{TV} \leq \|z\|_{TV}$.

$$\begin{aligned}\|zQ\|_{TV} &= \frac{1}{2} \sum_{i=1}^n \left| \sum_{j=1}^n z_j Q_{ji} \right| \\ &\leq \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n |z_j| Q_{ji} \quad (\text{triangle inequality}) \\ &= \frac{1}{2} \sum_{j=1}^n |z_j| \underbrace{\sum_{i=1}^n Q_{ji}}_{=1} \quad (\text{swap sums; } Q \text{ row-stochastic}) \\ &= \frac{1}{2} \sum_{j=1}^n |z_j| = \|z\|_{TV}.\end{aligned}$$

Applying Banach's Fixed Point Theorem

Step 3: We can show:

- $\phi: x^\top \mapsto x^\top G$ is a **contraction** with factor $(1-\delta) < 1$ (Step 2)
- $(\mathbb{P}^n, \|\cdot\|_{TV})$ is a **complete** metric space (Proposition 3.8)

Banach's Fixed Point Theorem

Let (X, d) be a non-empty complete metric space with a contraction mapping $T: X \rightarrow X$.
Then:

- 1 T admits a **unique** fixed point $x^* \in X$, i.e. $T(x^*) = x^*$
- 2 For any $x_0 \in X$, the sequence $x_n = T(x_{n-1})$ satisfies $\lim_{n \rightarrow \infty} x_n = x^*$

As all prerequisites are satisfied, we can apply the theorem to ϕ on $(\mathbb{P}^n, \|\cdot\|_{TV})$.

Applying Banach's Fixed Point Theorem to $\phi: x^\top \mapsto x^\top G$ yields:

Result

- 1 There exists a **unique** fixed point $x^* \in \mathbb{P}^n$ satisfying

$$(x^*)^\top = (x^*)^\top G$$

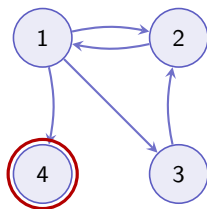
- 2 For **any** starting distribution $x^{(0)} \in \mathbb{P}^n$, the power iteration converges:

$$x^{(k)} \rightarrow x^* \quad \text{as } k \rightarrow \infty$$

This x^* is the **PageRank vector**, the unique stationary distribution of the random surfer model. □

Computation

Constructing the Hyperlink Matrix H



Node 4: **dangling node**
(no outgoing links)

Row-normalize the adjacency matrix A :

$$H_{ij} = \frac{A_{ij}}{\deg_{\text{out}}(i)}, \quad A_{ij} \in \{0, 1\}$$

Replace dangling rows with a distribution $\mathbf{v} \in \mathbb{P}^n$:

$$\tilde{H} = H + a\mathbf{v}^\top$$

$$\text{where } a_i = \begin{cases} 1, & \text{if } \deg_{\text{out}}(i) = 0 \\ 0, & \text{otherwise} \end{cases}$$

Canonical choice: $\mathbf{v} = \frac{1}{n} \mathbf{1}$

(Langville & Meyer, 2006)

Constructing the Teleportation Matrix T

Two central design choices for the teleportation matrix T :

Design Choice 1: Rank of T

- ▷ Rank-One Construction:
 $T = \mathbf{1}\mathbf{v}^\top$, $\mathbf{v} \in \mathbb{P}^n$
- ▷ Keeps G decomposable for efficient computation (see Complexity)

Design Choice 2: $\mathbf{v} \in \mathbb{P}^n$

- ▷ **Uniform:** $\mathbf{v} = \frac{1}{n} \mathbf{1}$
Standard PageRank, guarantees G is positive $\implies$ irreducible and aperiodic
- ▷ **Personalized:** $\mathbf{v}$ encodes user preferences
(Haveliwala, 2002)

Standard PageRank: Constructing the Google Matrix G

Combining everything, the **Google matrix** is:

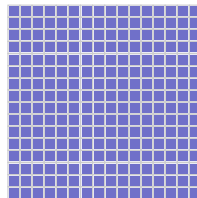
$$G = \underbrace{\alpha \tilde{H}}_{\text{Hyperlink}} + \underbrace{(1 - \alpha) T}_{\text{Teleportation}} = \alpha \underbrace{(H + a\mathbf{v}^T)}_{\tilde{H}} + (1 - \alpha) \underbrace{\mathbf{1}\mathbf{v}^T}_T$$

Per-Iteration Cost: Naive vs. Sparse Approach

Each iteration computes $(x^{(k+1)})^\top = (x^{(k)})^\top G$.

G is **completely dense** due to T

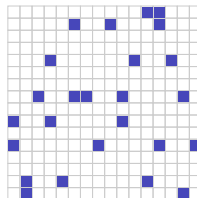
Dense G



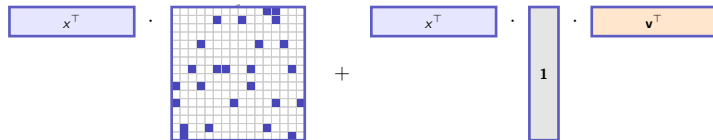
Instead, rearrange G :

$$\begin{aligned} G &= \underbrace{\alpha(H + a\mathbf{v}^\top)}_{\text{Hyperlink}} + \underbrace{(1 - \alpha)\mathbf{1}\mathbf{v}^\top}_{\text{Teleportation}} \\ &= \underbrace{\alpha H}_{\text{sparse}} + \underbrace{(\alpha a + (1 - \alpha)\mathbf{1})\mathbf{v}^\top}_{\text{rank-one correction}} \end{aligned}$$

Sparse H



Iteration Complexity: Exploiting Sparsity



Sparse Product ($x^\top \alpha H$):

Cost is $O(\text{nnz}(H)) \approx 10n \Rightarrow \mathbf{O(n)}$.

Rank-One Product ($x^\top \dots \mathbf{v}^\top$):

$\Rightarrow \mathbf{O(n)}$.

(* simplified for brevity, missing a correction)

Overall Cost per Iteration: $O(n)$

Convergence Rate

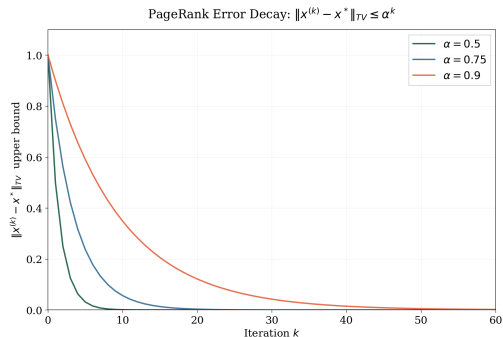
By Theorem 4.1 with $y = x^*$:

$$\|x^\top G - (x^*)^\top G\|_{TV} \leq (1 - \delta) \|x - x^*\|_{TV}$$

Applying repeatedly:

$$\begin{aligned}\|x^{(k)} - x^*\|_{TV} &= \|(x^{(k-1)})^\top G - (x^*)^\top G\|_{TV} \\ &\leq (1 - \delta) \|x^{(k-1)} - x^*\|_{TV} \\ &\leq (1 - \delta)^2 \|x^{(k-2)} - x^*\|_{TV} \\ &\vdots \\ &\leq (1 - \delta)^k \|x^{(0)} - x^*\|_{TV}\end{aligned}$$

The error decreases **exponentially**: $\|x^{(k)} - x^*\|_{TV} \leq (1 - \delta)^k \|x^{(0)} - x^*\|_{TV}$



Error upper bound α^k for different damping factors α .

Number of Iterations

For ε -precision: $(1 - \delta)^k \|x^{(0)} - x^*\|_{TV} \leq \varepsilon$.

Since $\|x^{(0)} - x^*\|_{TV} \leq 1$, by rearranging and taking logarithms:

$$k \geq \frac{\log(1/\varepsilon)}{-\log(1 - \delta)}$$

Remark: Standard PageRank

With uniform teleportation, $\delta = 1 - \alpha$, so the contraction factor is simply α :

$$\|x^{(k)} - x^*\|_{TV} \leq \alpha^k \quad \implies \quad k \geq \left\lceil \frac{-\log \varepsilon}{\log(1/\alpha)} \right\rceil$$

For $\alpha = 0.85$: $k \approx 14\tau$ iterations for τ digits of accuracy.

Modifications

Problem: PageRank is Query-Independent

“Digital Design & Computer Arch.”

Keyword Retrieval + PageRank Reranking



#1 Architectural Digest, PR: 0.038

#2 Dezeen, PR: 0.031

#3 ArchDaily, PR: 0.025

#4 Harris & Harris: DDCA, PR: 0.009

#5 MIT OCW, Digital Systems: PR: 0.007

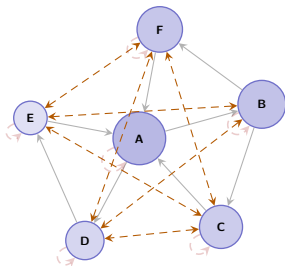
■ Off-topic, high PR ■ Relevant, low PR

Problem:

Standard PageRank is **static** for a given web graph.

Personalized PageRank: Visualized

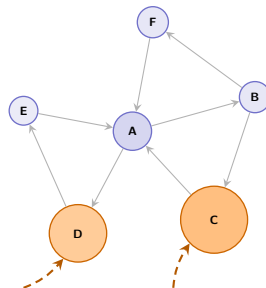
Standard PageRank (Brin & Page, 1998)



uniform teleportation

Teleportation: $\mathbf{v} = \frac{1}{n}\mathbf{1} \in \mathbb{P}^n$
Ranks are **user-independent**

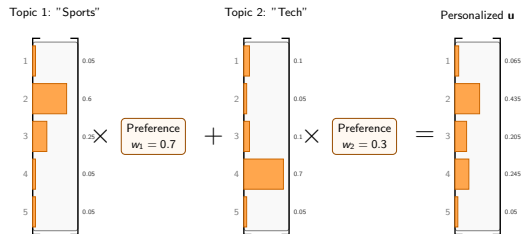
Personalized PageRank (Haveliwala, 2002)



topical teleportation

Teleportation: $\mathbf{u}_j \in \mathbb{P}^n$ concentrated on topic
Ranks are **user-dependent**

Personalized PageRank: Formalized

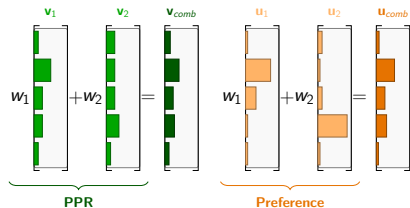


Key idea: Replace uniform $\mathbf{v} \in \mathbb{P}^n$ with *topic-specific* $\mathbf{u}_j \in \mathbb{P}^n$:

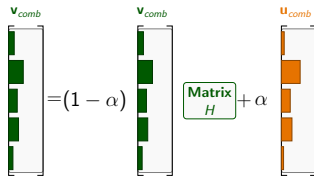
$$\begin{aligned}(\mathbf{x}_j^*)^\top &= (\mathbf{x}_j^*)^\top G \\&= \alpha (\mathbf{x}_j^*)^\top H + (1 - \alpha) (\mathbf{x}_j^*)^\top T \\&= \alpha (\mathbf{x}_j^*)^\top H + (1 - \alpha) \underbrace{(\mathbf{x}_j^*)^\top \mathbf{1}}_{=1} \mathbf{u}_j^\top \\&= \alpha (\mathbf{x}_j^*)^\top H + \underbrace{(1 - \alpha) \mathbf{u}_j^\top}_{\text{topic-specific}}\end{aligned}$$

Personalized PageRank (PPR) is Linearly Composable

1. Linear Combination



2. Fixed-Point Iteration



Theorem: Linearity of PPR (Jeh & Widom, 2003)

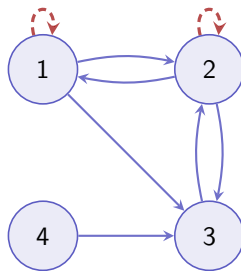
For preference vectors $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{P}^n$ with corresponding PPR vectors $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{P}^n$ and weights $w_1, w_2 \geq 0$ s.t. $w_1 + w_2 = 1$:

$$(w_1 \mathbf{v}_1 + w_2 \mathbf{v}_2)^\top = \alpha (\mathbf{v}_{comb})^\top H + (1 - \alpha) (\mathbf{u}_{comb}^\top)$$

where:

- $\mathbf{v}_{comb} = w_1 \mathbf{v}_1 + w_2 \mathbf{v}_2$ (combined PPR)
- $\mathbf{u}_{comb} = w_1 \mathbf{u}_1 + w_2 \mathbf{u}_2$ (combined preference).

Personalized PageRank: Convergence



$$(\mathbf{u}_j)_3 = (\mathbf{u}_j)_4 = 0$$

$$(\mathbf{u}_j)_1, (\mathbf{u}_j)_2 > 0$$

Personalized PageRank allows $\mathbf{u}_j$ with **zero entries** for off-topic pages.

$\Rightarrow T = \mathbf{1} \mathbf{u}_j^\top$ is no longer strictly positive.

$\Rightarrow G$ is not strictly positive.

Banach-based proof (Thm. 4.1) required positive G .

However: Convergence still holds.

- The teleportation terms cancel in the contraction estimate: $\mathbf{u}_j^\top - \mathbf{u}_j^\top = 0$, regardless of the structure of $\mathbf{u}_j$
- Only **non-expansiveness** of H (Prop. 3.11) is needed, contraction factor is α



Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. *Computer Networks*, 30, 107–117. Retrieved from <https://snap.stanford.edu/class/cs224w-readings/Brin98Anatomy.pdf>



Haveliwala, T. H. (2002). Topic-sensitive pagerank. In *Proceedings of the 11th international conference on world wide web* (pp. 517–526). doi:10.1145/511446.511513



Jeh, G., & Widom, J. (2003). Scaling personalized web search. In *Proceedings of the 12th international conference on world wide web* (pp. 271–279). doi:10.1145/775152.775191



Langville, A. N., & Meyer, C. D. (2006). The mathematics of google's pagerank. In *Google's pagerank and beyond: The science of search engine rankings* (Chap. 4, pp. 31–46). doi:10.1515/9781400830329

Completeness of the Probability Simplex

Proposition 3.8 (Completeness of the Probability Simplex)

The probability simplex $(\mathbb{P}^n, \|\cdot\|_{TV})$ is a complete metric space.

Why do we need this? Completeness is a prerequisite for Banach's Fixed Point Theorem. It ensures that the iterative process cannot “drift off” to a limit outside $\mathbb{P}^n$, so a unique stationary distribution must exist within $\mathbb{P}^n$.

Proposition 3.10 (Stochastic Decomposition)

Let G and Π be stochastic matrices such that $G_{ij} \geq \delta \Pi_{ij}$ for some $\delta \in (0, 1)$ and all $i, j \in \{1, \dots, n\}$. Then the matrix

$$Q = \frac{1}{1 - \delta} (G - \delta \Pi)$$

is also a stochastic matrix.

Equivalently, we can decompose G as:

$$G = (1 - \delta) Q + \delta \Pi$$

Stochastic Decomposition: Proof

We verify the two defining properties of a stochastic matrix for $Q = \frac{1}{1-\delta}(G - \delta \Pi)$.

1. Non-negativity. Since $G_{ij} \geq \delta \Pi_{ij}$ by assumption:

$$Q_{ij} = \frac{1}{1-\delta}(G_{ij} - \delta \Pi_{ij}) \geq 0. \quad \checkmark$$

2. Row sums equal one. Both G and Π are stochastic, so each of their rows sums to 1. Hence:

$$\sum_{j=1}^n Q_{ij} = \frac{1}{1-\delta} \sum_{j=1}^n (G_{ij} - \delta \Pi_{ij}) = \frac{1}{1-\delta} \left(\underbrace{\sum_{j=1}^n G_{ij}}_{=1} - \delta \underbrace{\sum_{j=1}^n \Pi_{ij}}_{=1} \right) = \frac{1-\delta}{1-\delta} = 1. \quad \checkmark$$

Therefore Q is a stochastic matrix. □