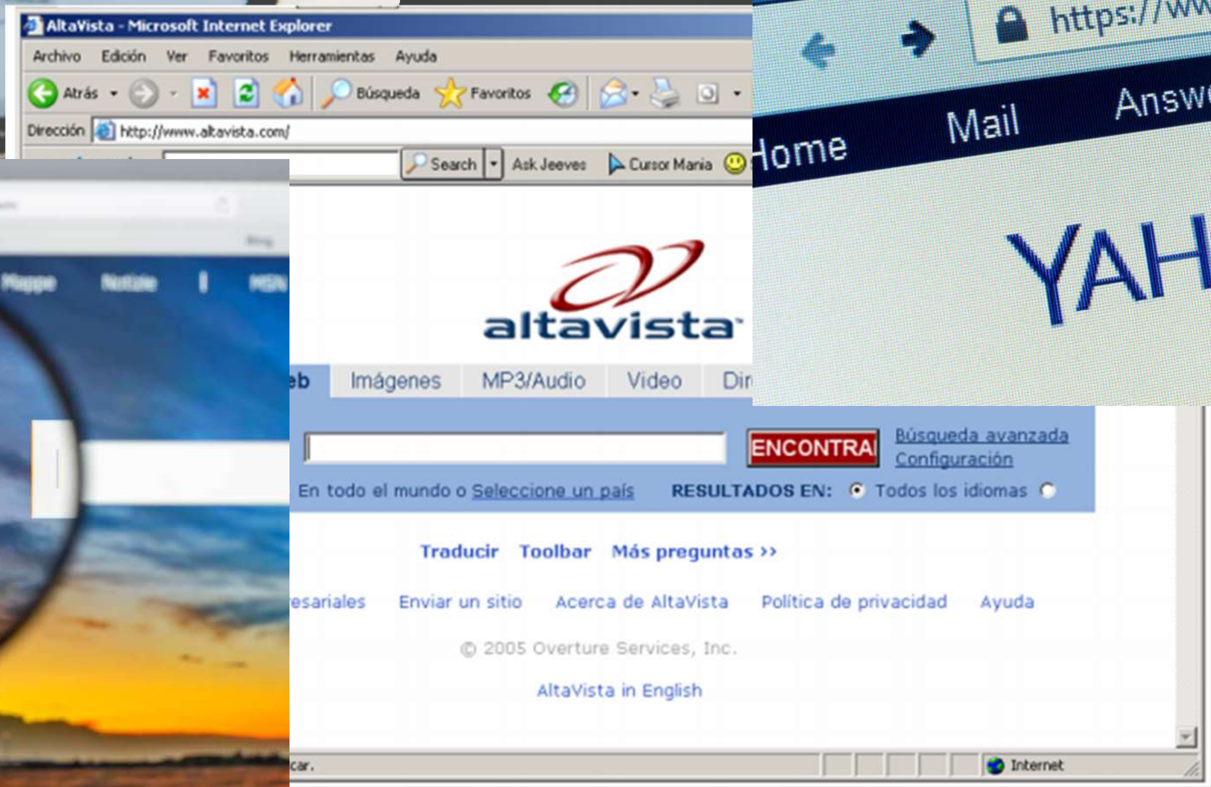
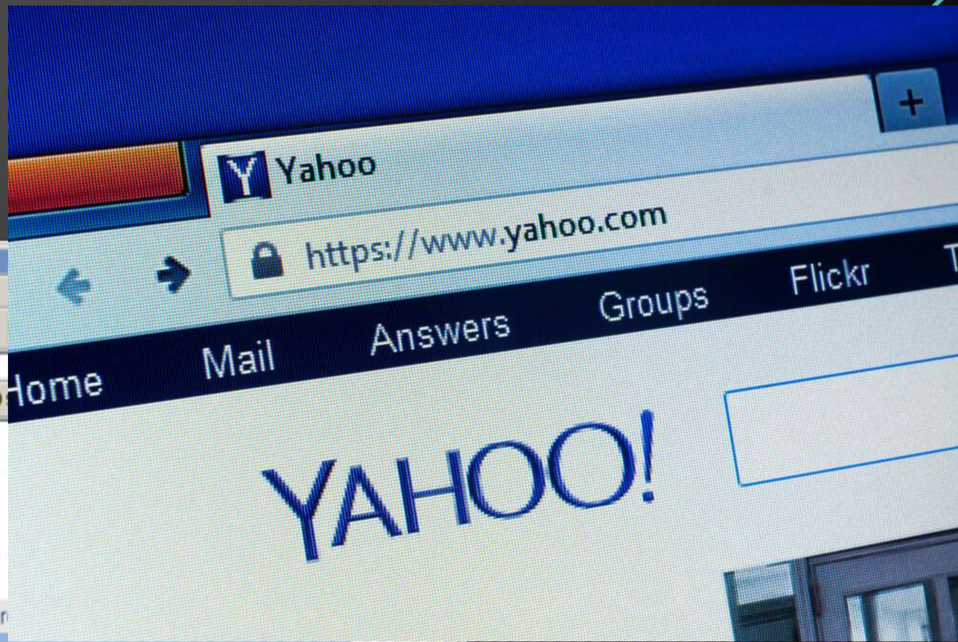

ESTRUTURA
DE DADOS
PAGE RANK
AM

BUSCADORES





Professor
Douglas Maioli

professor douglas maioli

Todas | Notícias | Imagens | Vídeos | Maps | Mais | Configurações | Ferrament

Aproximadamente 34.300 resultados (0,49 segundos)

https://www.youtube.com › professordouglasmaioli

Professor Douglas Maioli - YouTube

Canal com aulas e dicas do **Professor Douglas Maioli**.

https://professordouglasmaioli.com.br

Professor Douglas Maioli: Início

Professor Douglas Maioli. Seja bem-vindo ao nosso site ...

https://br.linkedin.com › professordouglasmaioli

Douglas Silva Maioli - Univesp Oficial - Campinas, São Paulo ...

Atuei como **professor** das disciplinas de matemática e física do ciclo básico das engenharias. Gráfico E. E. **Prof**. Antonio Vilela Júnior. **Professor** de Matemática e ...

https://www.instagram.com › professordouglasmaioli

Matemática Prof Douglas Maioli (@professordouglasmaioli ...

902 Followers, 466 Following, 37 Posts - See Instagram photos and videos from Matemática **Prof Douglas Maioli** (**@professordouglasmaioli**)

https://www.facebook.com › ... › Campinas, Sao Paulo

Matemática com o Prof. Douglas Maioli - Home | Facebook

Matemática com o **Prof. Douglas Maioli**, Campinas. 532 likes · 27 talking about this. Vídeo aulas gratuitas de Matemática, Física e Computação no Canal...

★★★★★ Avaliação: 5 · 5 votos

PAGE RANK

PAGE RANK

The Anatomy of a Large-Scale Hypertextual Web Search Engine

Sergey Brin and Lawrence Page
{sergey, page}@cs.stanford.edu
Computer Science Department, Stanford University, Stanford, CA 94305

Abstract

In this paper, we present Google, a prototype of a large-scale search engine which makes heavy use of the structure present in hypertext. Google is designed to crawl and index the Web efficiently and produce much more satisfying search results than existing systems. The prototype with a full text and hyperlink database of at least 24 million pages is available at http://google.stanford.edu/

To engineer a search engine is a challenging task. Search engines index tens to hundreds of millions of web pages involving a comparable number of distinct terms. They answer tens of millions of queries every day. Despite the importance of large-scale search engines on the web, very little academic research has been done on them. Furthermore, due to rapid advance in technology and web proliferation, creating a web search engine today is very different from three years ago. This paper provides an in-depth description of our large-scale web search engine -- the first such detailed public description we know of to date.

Apart from the problems of scaling traditional search techniques to data of this magnitude, there are new technical challenges involved with using the additional information present in hypertext to produce better search results. This paper addresses this question of how to build a practical large-scale system which can exploit the additional information present in hypertext. Also we look at the problem of how to effectively deal with uncontrolled hypertext collections where anyone can publish anything they want.

Keywords: World Wide Web, Search Engines, Information Retrieval, PageRank, Google

1. Introduction

(Note: There are two versions of this paper -- a longer full version and a shorter printed version. The full version is available on the web and the conference CD-ROM.)
The web creates new challenges for information retrieval. The amount of information on the web is growing rapidly, as well as the number of new users inexperienced in the art of w
its link graph, often starting with high quality human maintained indices such as Yahoo! or with search engines. Human maintained lists cover popular topics effectively but are subj
improve, and cannot cover all esoteric topics. Automated search engines that rely on keyword matching usually return too many low quality matches. To make matters worse, some
taking measures meant to mislead automated search engines. We have built a large-scale search engine which addresses many of the problems of existing systems. It makes especial
hypertext to provide much higher quality search results. We chose our system name, Google, because it is a common spelling of googol, or 10^{100} and fits well with our goal of build

1.1 Web Search Engines -- Scaling Up: 1994 - 2000

Search engine technology has had to scale dramatically to keep up with the growth of the web. In 1994, one of the first web search engines, the World Wide Web Worm (WWWW) [
web accessible documents. As of November, 1997, the top search engines claim to index from 2 million (WebCrawler) to 100 million web documents (from Search Engine Watch). I
comprehensive index of the Web will contain over a billion documents. At the same time, the number of queries search engines handle has grown incredibly too. In March and April

PAGE RANK

The Anatomy of a Large-Scale Hypertextual Web Search Engine

Sergey Brin and Lawrence Page
{sergey, page}@cs.stanford.edu
Computer Science Department, Stanford University, Stanford, CA 94305

Abstract

In this paper, we present Google, a prototype of a large-scale search engine which makes heavy use of the structure present in hypertext. Google is designed to crawl and index the Web efficiently and produce much more satisfying search results than existing systems. The prototype with a full text and hyperlink database of at least 24 million pages is available at http://google.stanford.edu/

To engineer a search engine is a challenging task. Search engines index tens to hundreds of millions of web pages involving a comparable number of distinct terms. They answer tens of millions of queries every day. Despite the importance of large-scale search engin ... oliferation, creating a web search engine today is very different from three years ago. This ... f to date.

Apart from the problems of scaling traditional ... esent in hypertext to produce better search results. This paper addresses this question of how ... em of how to effectively deal with uncontrolled hypertext collections where anyone c ...

Sergey Brin and Lawrence Page
{sergey, page}@cs.stanford.edu

1. Introduction

(Note: There are two versions of this paper -- a longer full version and a shorter printed version. The full version is available on the web and the conference CD-ROM.)
The web creates new challenges for information retrieval. The amount of information on the web is growing rapidly, as well as the number of new users inexperienced in the art of w ... its link graph, often starting with high quality human maintained indices such as Yahoo! or with search engines. Human maintained lists cover popular topics effectively but are subj ... improve, and cannot cover all esoteric topics. Automated search engines that rely on keyword matching usually return too many low quality matches. To make matters worse, some ... taking measures meant to mislead automated search engines. We have built a large-scale search engine which addresses many of the problems of existing systems. It makes especial ... hypertext to provide much higher quality search results. We chose our system name, Google, because it is a common spelling of googol, or 10^{100} and fits well with our goal of build ...

1.1 Web Search Engines -- Scaling Up: 1994 - 2000

Search engine technology has had to scale dramatically to keep up with the growth of the web. In 1994, one of the first web search engines, the World Wide Web Worm (WWWW) ... web accessible documents. As of November, 1997, the top search engines claim to index from 2 million (WebCrawler) to 100 million web documents (from Search Engine Watch). ... comprehensive index of the Web will contain over a billion documents. At the same time, the number of queries search engines handle has grown incredibly too. In March and April ...

Professor
Douglas Maioli

PAGE RANK

The Anatomy of a Large-Scale Hypertextual Web Search Engine

Sergey Brin and Lawrence Page
{sergey, page}@cs.stanford.edu
Computer Science Department, Stanford University, Stanford, CA 94305

Abstract

In this paper, we present Google, a prototype of a large-scale search engine which makes heavy use of the structure present in hypertext. Google is designed to crawl and index the Web efficiently and produce much more satisfying search results than existing systems. The prototype with a full text and hyperlink database of at least 24 million pages is available at http://google.stanford.edu/

To engineer a search engine is a challenging task. Search engines index tens to hundreds of millions of web pages involving a comparable number of distinct terms. They answer tens of millions of queries every day. Despite the importance of large-scale search engines on the web, very little academic research has been done on them. Furthermore, due to rapid advance in technology and web proliferation, creating a web search engine today is very different from three years ago. This paper provides an in-depth description of our large-scale web search engine -- the first such detailed public description we know of to date.

Apart from the problems of scaling traditional search techniques to data of this magnitude, there are new technical challenges involved with using the additional information present in hypertext to produce better search ... k at the problem of how to effectively deal with

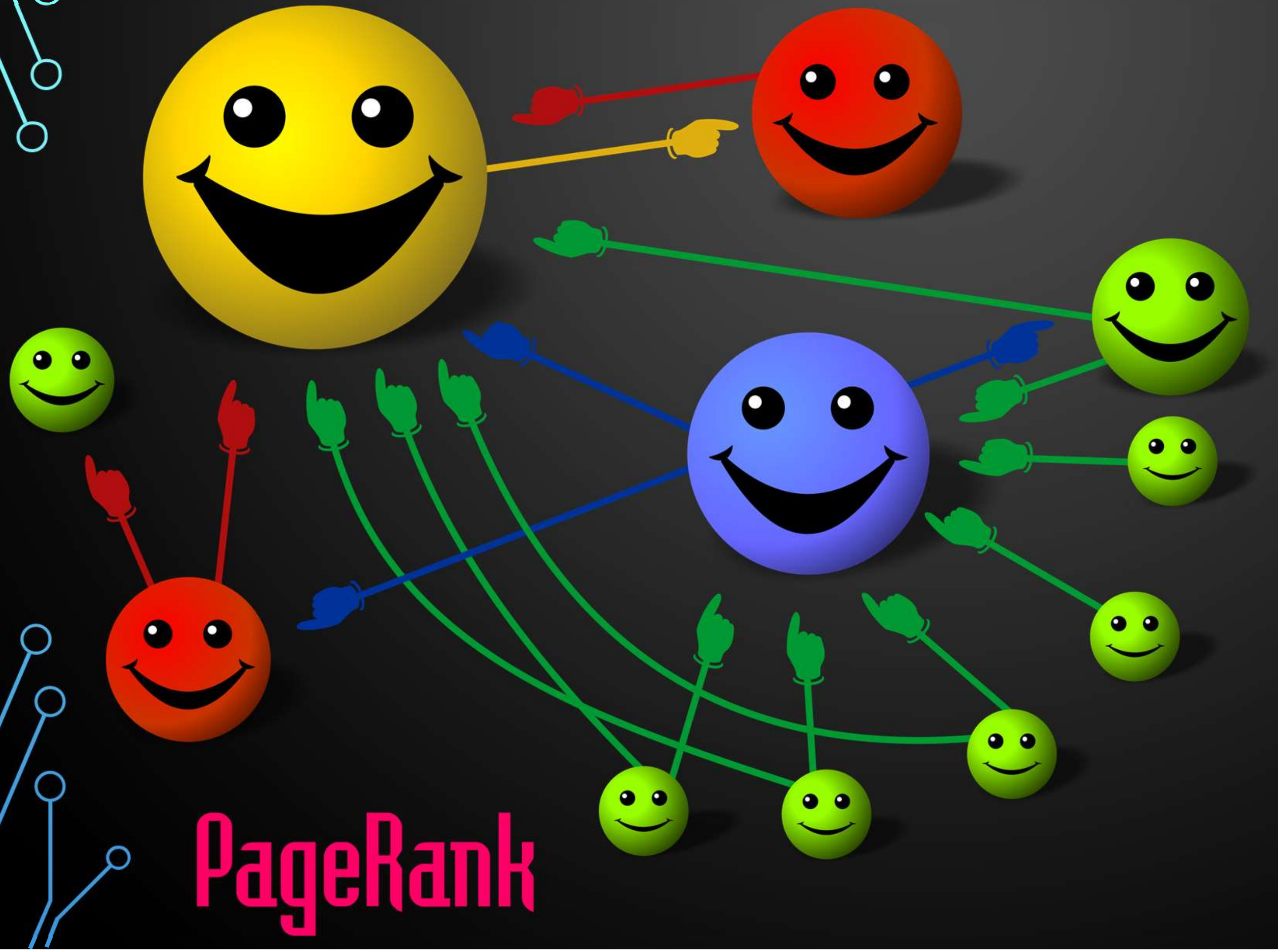
Neste artigo, apresentamos o Google, um protótipo de um motor de busca de larga escala que faz uso da estrutura presente nos hipertextos. O Google foi projetado para rastrear e indexar a web de forma eficiente e produzir resultados de pesquisa muito mais satisfatórios do que os sistemas existentes.

1.1 Web Search Engines -- Scaling Up: 1994 - 2000

Search engine technology has had to scale dramatically to keep up with the growth of the web. In 1994, one of the first web search engines, the World Wide Web Worm (WWWW) ... web accessible documents. As of November, 1997, the top search engines claim to index from 2 million (WebCrawler) to 100 million web documents (from Search Engine Watch). ... comprehensive index of the Web will contain over a billion documents. At the same time, the number of queries search engines handle has grown incredibly too. In March and April ...

Professor
Douglas Maioli

PAGE RANK



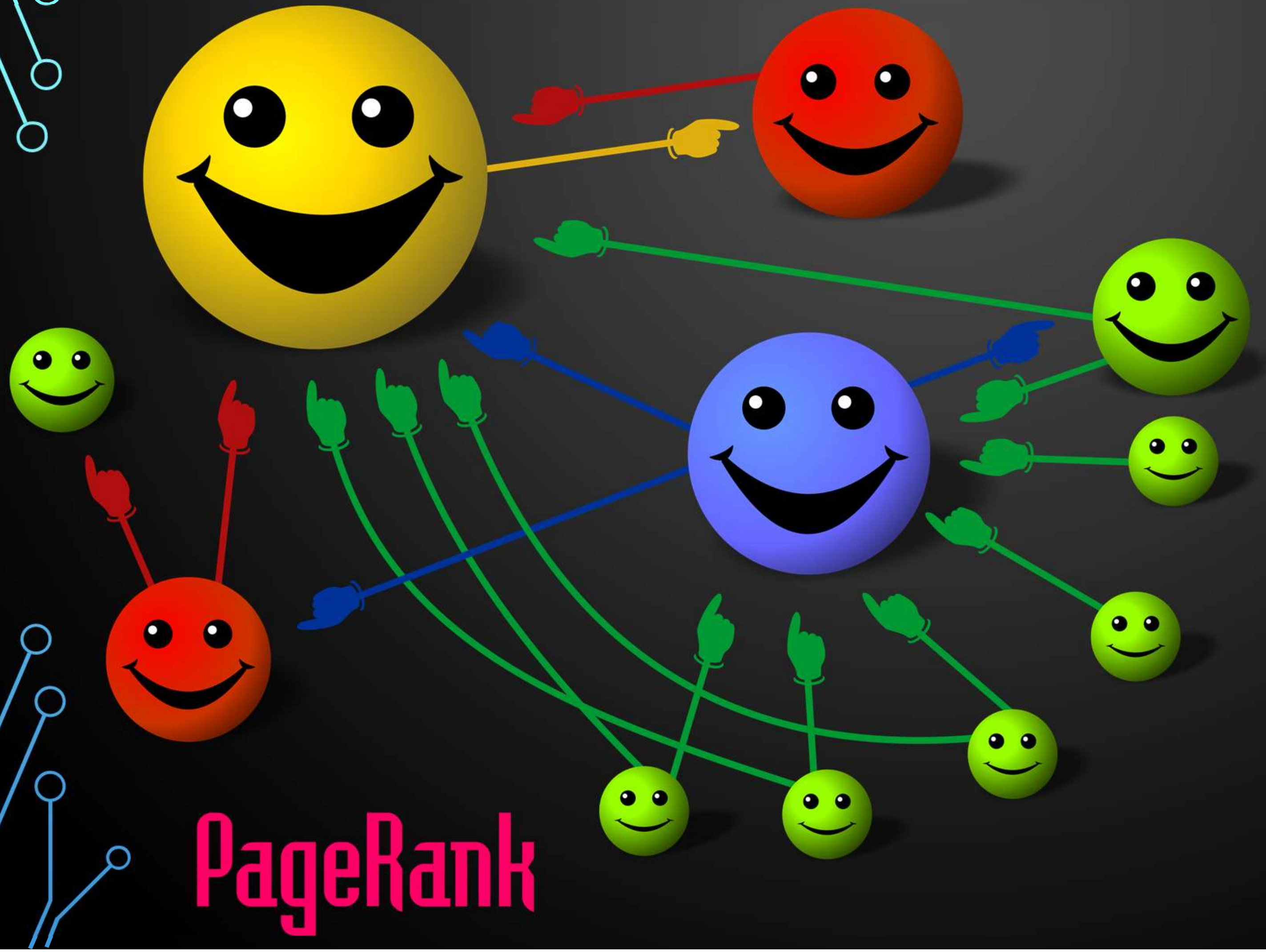
1) **Quantidade de links que recebe;**

2) **Links de sites com PageRank alto valem mais;**

3) **Quanto mais links envia, menor o peso de cada.**

Professor
Douglas Maioli

PAGE RANK



- **Não será levado em consideração links para a própria página (sem laços);**
- **Vários links de uma página para outra, será considerado como uma (sem arestas múltiplas);**
- **Grafo Simples**

Professor
Douglas Maioli

PAGE RANK
MODELO SIMPLIFICADO

Em um dado um grafo G com n vértices, temos que o Page Rank do vértice p na iteração t ($PR(p,t)$) é dado por:

$$\boldsymbol{PR(p,0) = \frac{1}{n}}$$

$$\boldsymbol{PR(p,t) = \sum_{p' \in C(p)} \frac{PR(p',t-1)}{NumLinks(p')}}$$

Em que:

$C(p)$ é o conjunto de página que tem links para p

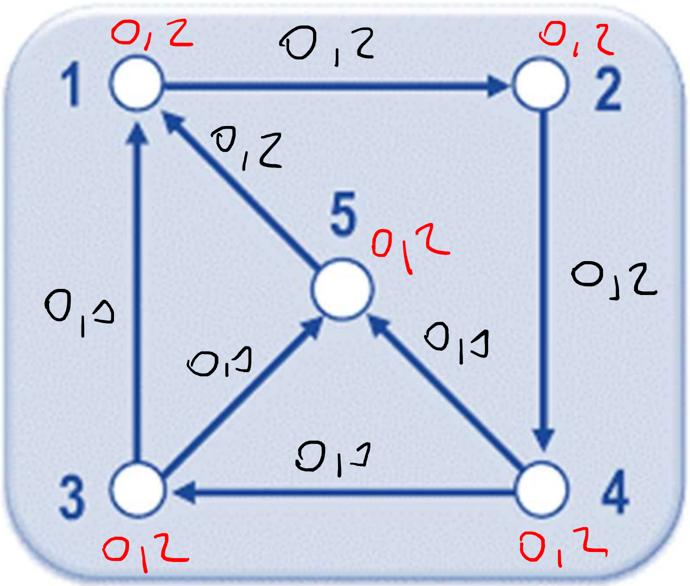
$NumLinks(p')$ é a quantidade de links na página p'

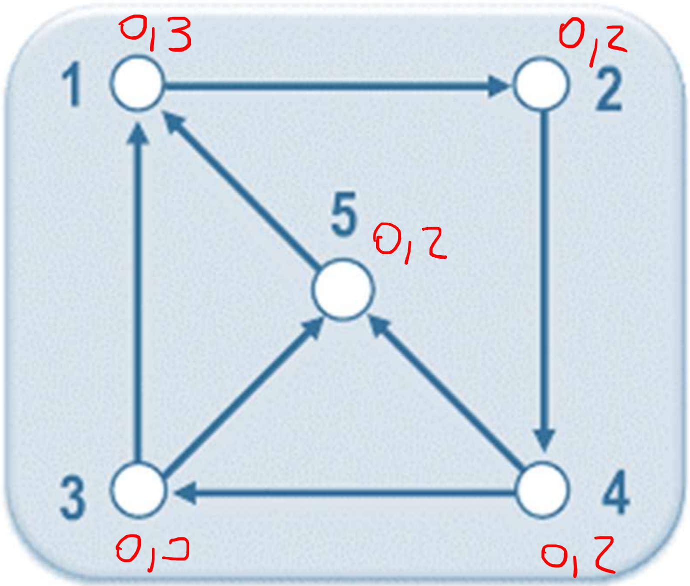
Professor
Douglas Maioli

$t = 0$

Nó	Page Rank
1	0,2
2	0,2
3	0,2
4	0,2
5	0,2

$n = 5 \qquad \frac{1}{5} = 0,2$





$$PR(p,0) = \frac{1}{n}$$

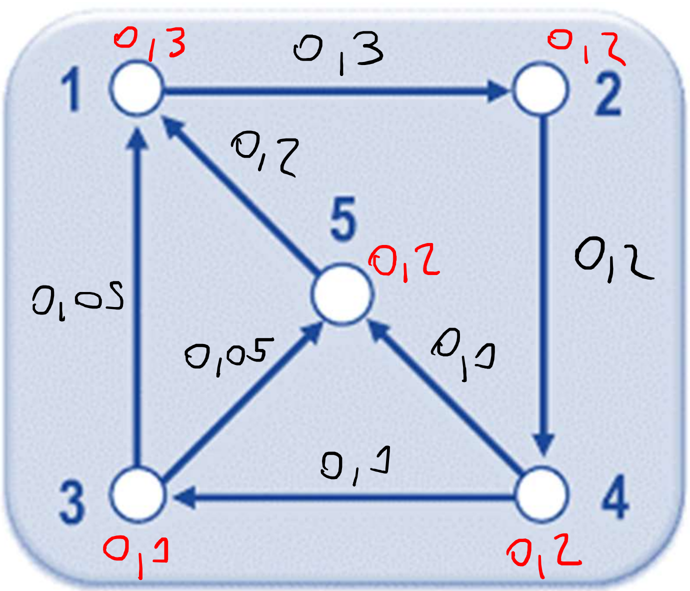
$$PR(p,t) = \sum_{p' \in C(p)} \frac{PR(p', t-1)}{NumLinks(p')}$$

Professor
Douglas Maioli

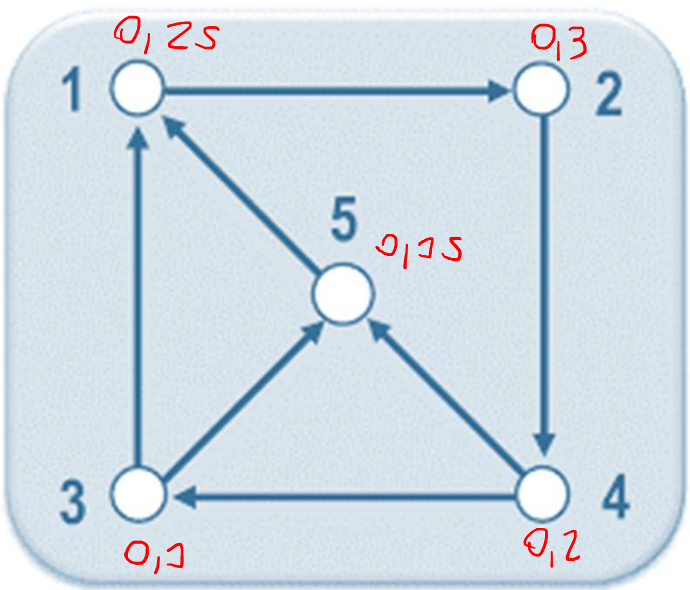
$t = 1$

Nó	Page Rank
1	0,3
2	0,2
3	0,1
4	0,2
5	0,2

t = 1



t = 2



$$PR(p,0) = \frac{1}{n}$$

$$PR(p,t) = \sum_{p' \in C(p)} \frac{PR(p', t-1)}{NumLinks(p')}$$

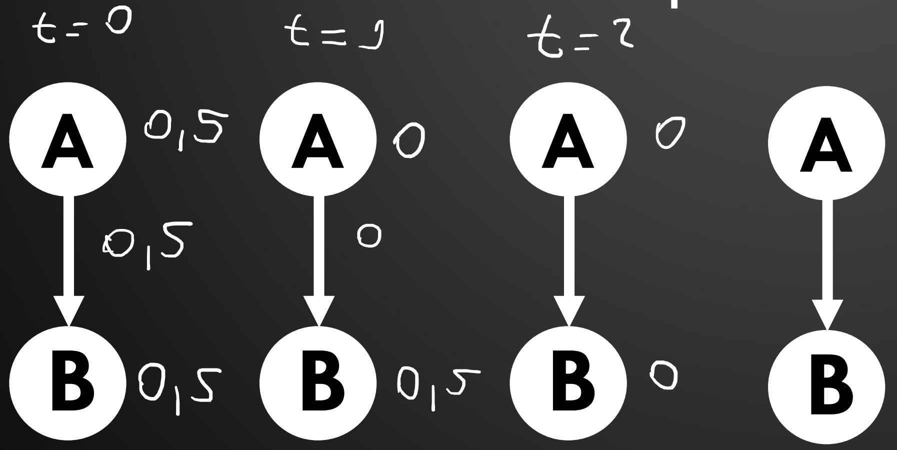
Professor
Douglas Maioli

PAGE RANK
PROBLEMA DO MODELO SIMPLIFICADO

Páginas sem links de saída podem drenar os Page Ranks do sistema:

$n=2$

$\frac{1}{2}=0{,}5$



Vértices/Iteração	t=0	t=1	t=2	t=3
A	0,5	0	0	0
B	0,5	0,5	0	0

Professor
Douglas Maioli

PAGE RANK
FÓRMULA ORIGINAL

Em um dado um grafo G com n vértices, temos que o Page Rank do vértice p na iteração t ($PR(p,t)$) é dado por:

$$\boldsymbol{PR(p,0) = \frac{1}{n}}$$

$$\boldsymbol{PR(p,t) = (1-d) + d.\sum_{p' \in C(p)} \frac{PR(p',t-1)}{NumLinks(p')}}$$

Em que:

$C(p)$ é o conjunto de página que tem links para p

$NumLinks(p')$ é a quantidade de links na página p'

d é o fator de amortecimento (usualmente $0{,}85$)

Professor Douglas Maioli

PAGE RANK
COM FATOR DE AMORTECIMENTO

Em um dado um grafo G com n vértices, temos que o Page Rank do vértice p na iteração t ($PR(p,t)$) é dado por:

$$\boldsymbol{PR(p,0) = \frac{1}{n}}$$

$$\boldsymbol{PR(p,t) = \frac{(1-d)}{n} + d. \sum_{p' \in C(p)} \frac{PR(p',t-1)}{NumLinks(p')}}$$

Em que:

$C(p)$ é o conjunto de página que tem links para p

$NumLinks(p')$ é a quantidade de links na página p'

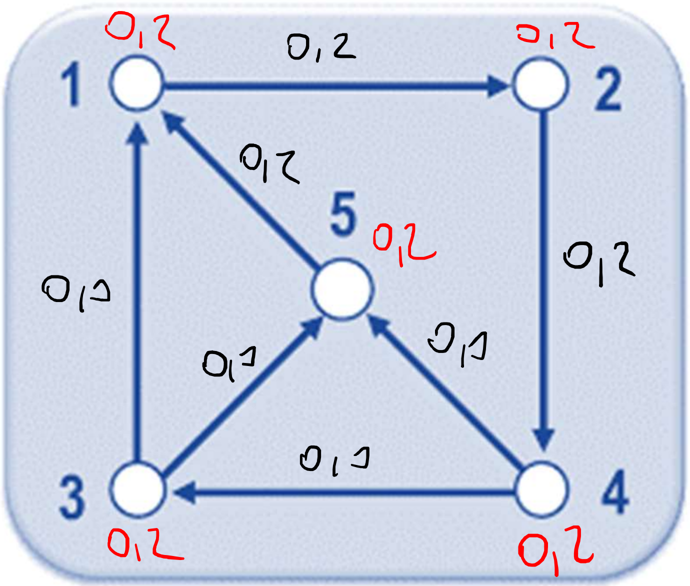
d é o fator de amortecimento (usualmente $0{,}85$)

Professor
Douglas Maioli

$t = 0$

Nó	Page Rank
1	0,2
2	0,2
3	0,2
4	0,2
5	0,2

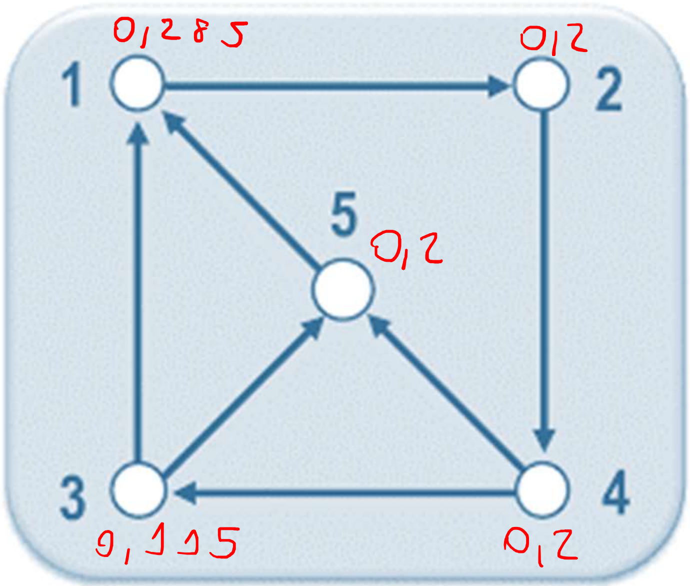
$n = 5 \quad t = 0 \quad \frac{1}{5} = 0,2$



$5 \to 0,03 + 0,85 . 0,2 = 0,2$

$2 \to 0,03 + 0,85 . 0,2 = 0,2$

$3 \to 0,03 + 0,85 . 0,1 = 0,115$



$1 \to 0,03 + 0,85 . 0,3 =$

$$d = 0,85$$

$$\frac{1-d}{n} = \frac{0,15}{5} = 0,03$$

$$PR(p,0) = \frac{1}{n}$$

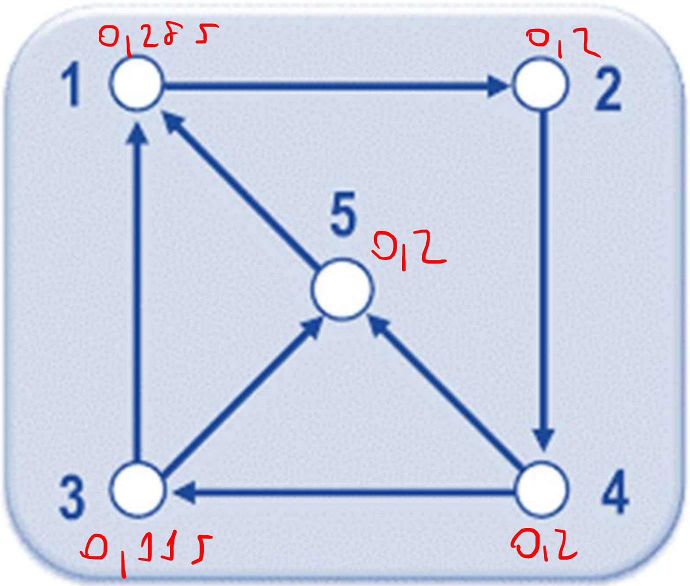
$$PR(p,t) = \frac{(1-d)}{n} + d.\sum_{p' \in C(p)} \frac{PR(p', t-1)}{NumLinks(p')}$$

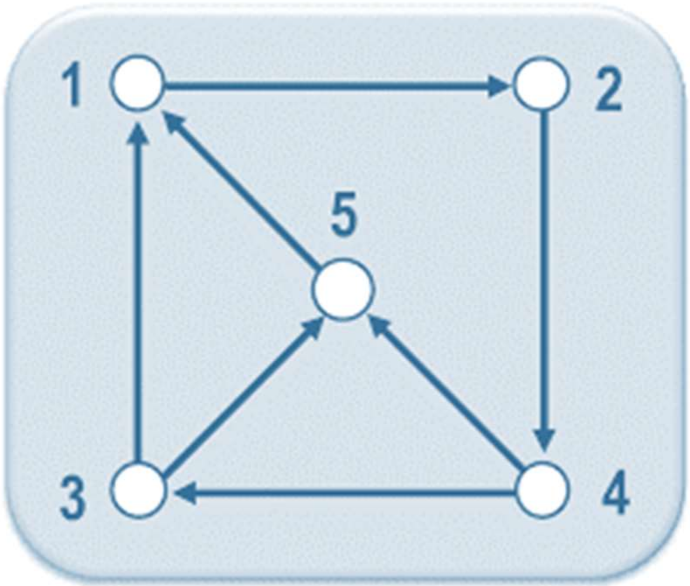
Professor
Douglas Maioli

$$t = 1$$

Nó	Page Rank
1	0,285
2	0,2
3	0,115
4	0,2
5	0,2

$t = 1$





$$d = 0,85$$

$$\frac{1-d}{n} = \frac{0,15}{5} = 0,03$$

$$PR(p,0) = \frac{1}{n}$$

$$PR(p,t) = \frac{(1-d)}{n} + d.\sum_{p' \in C(p)} \frac{PR(p', t-1)}{NumLinks(p')}$$

Professor
Douglas Maioli