OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG

INF FAKULTÄT FÜR INFORMATIK

Discrete–Time Markov Chains

Eating Habits, Gamblers and Google

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG INF FAKULTÄT FÜR INFORMATIK

Applied Discrete Modeling

Topics of last ItS lecture

- (Discrete Time) Markov Chains, DTMC (Graham and the CIA)
- Hidden Markov Models (What illness does the patient have?)
- Hidden **non**–Markovian Models (Quality tester at a factory)
- Virtual Stochastic Sensors

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG INF FAKULTÄT FÜR INFORMATIK

Advanced Discrete Modeling

Today's agenda (= the part of the lecture relevant to the exam)

- (Discrete Time) Markov Chains, DTMC (Graham and the CIA)
- Hidden Markov Models (What illness does the patient have?)
- Hidden **non**–Markovian Models (Quality tester at a factory)
- Virtual Stochastic Sensors

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG INF FAKULTÄT FÜR INFORMATIK

Chinese, Greek or Italian?

We know Graham ate at the Chinese restaurant today

- Where will he eat the day after tomorrow?

Graham is dining out every night, always in one of these three restaurants: Chinese, Greek or Italian

We observed Graham for the past days and recorded his restaurant choices:

- C, I, G, G, I, I, G, I, G, I, C, C, C, C, I, I …

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG INF FAKULTÄT FÜR INFORMATIK

Chinese, Greek or Italian?

We now assume, that Graham' s dinner choice only depends on where he ate the evening before

Information extracted from the recorded sequence:

Restaurant	Frequency
Chinese	250
Greek	410
Italian	530

Choice one evening	Next choice	Frequency
Chinese	Chinese	50
Chinese	Greek	75
Chinese	Italian	125
Greek	Chinese	41
Greek	Greek	123
Greek	Italian	246
Italian	Chinese	159
Italian	Greek	212
Italian	Italian	159

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG
INF FAKULTÄT FÜR INFORMATIK

Chinese, Greek or Italian?

We can now compute the probabilities for each choice **B**, depending on the choice **A** of the night before

- $P(B|A) = H(B \wedge A)/H(A)$

Choice A	H(A)
Chinese	250
Greek	410
Italian	530

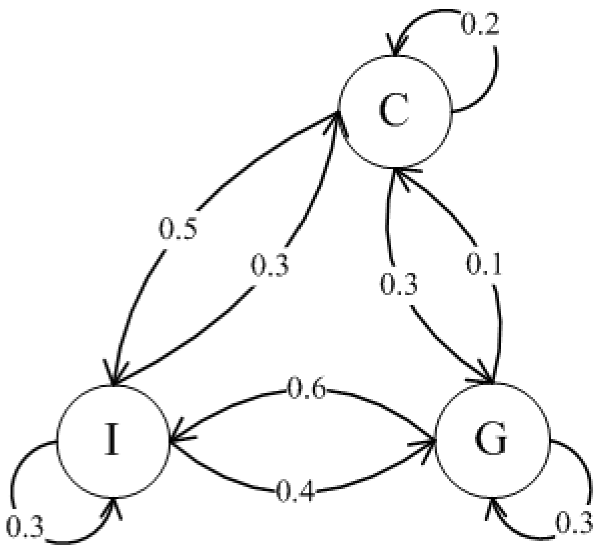
Choice A	Choice B	H(B∧A)	H(B∧A) / H(A)	P(B|A) = Probability
Chinese	Chinese	50	50 / 250	0.2
Chinese	Greek	75	75 / 250	0.3
Chinese	Italian	125	125 / 250	0.5
Greek	Chinese	41	41 / 410	0.1
Greek	Greek	123	123 / 410	0.3
Greek	Italian	246	246 / 410	0.6
Italian	Chinese	159	159 / 530	0.3
Italian	Greek	212	212 / 530	0.4
Italian	Italian	159	159 / 530	0.3

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG
INF FAKULTÄT FÜR INFORMATIK

Chinese, Greek or Italian?

We can now represent Graham's dining behavior as an annotated directed graph

Graphical representation



Tabular representation

A → B	Chinese	Greek	Italian
Chinese	0.2	0.3	0.5
Greek	0.1	0.3	0.6
Italian	0.3	0.4	0.3

$$P = \begin{bmatrix} 0.2 & 0.3 & 0.5 \\ 0.1 & 0.3 & 0.6 \\ 0.3 & 0.4 & 0.3 \end{bmatrix}$$

Chinese, Greek or Italian?

Back to our question:

- We know he ate in the Chinese restaurant two days ago (π_0)
- Where will Graham eat tonight?

Probability Vector $\pi_k = (\pi_C, \pi_G, \pi_I)$

- Each element represents the probability of the corresponding system state in step k
- Our initial probability vector:

$$\pi_0 = (1.00 \quad 0.00 \quad 0.00)$$

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG INF FAKULTÄT FÜR INFORMATIK

Chinese, Greek or Italian?

Back to our question:

- We know he ate in the Chinese restaurant two days ago (π_0)
- Where will Graham eat tonight? (π_2)

How to find the answer?

- Compute the probability of each restaurant choice two days (=steps) later → π_2

Iteratively compute π_2

- Use π_0 and P to obtain π_1; and then π_1 and P to obtain π_2
- In general: use current state probabilities (π_k) and the state transition probabilities (P) to gain the state probabilities (π_{k+1}) for the next step

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG INF FAKULTÄT FÜR INFORMATIK

Chinese, Greek or Italian?

Computing the next state probabilities

$$\pi_1 = (\pi_{1,C} \quad \pi_{1,G} \quad \pi_{1,I})$$

$$\pi_{1,C} = \pi_{0,C}\, p_{CC} + \pi_{0,G}\, p_{GC} + \pi_{0,I}\, p_{IC}$$

$$\pi_{1,G} = \pi_{0,C}\, p_{CG} + \pi_{0,G}\, p_{GG} + \pi_{0,I}\, p_{IG}$$

$$\pi_{1,I} = \pi_{0,C}\, p_{CI} + \pi_{0,G}\, p_{GI} + \pi_{0,I}\, p_{II}$$

$$\pi_{1,C} = 1.0*0.2 + 0*0.1 + 0*0.3 = 0.2$$

$$\pi_{1,G} = 1.0*0.3 + 0*0.3 + 0*0.4 = 0.3$$

$$\pi_{1,I} = 1.0*0.5 + 0*0.6 + 0*0.3 = 0.5$$

$$\pi_1 = (0.2 \quad 0.3 \quad 0.5)$$

… and now much easier!

$$\pi_0 P = (1.0 \quad 0.0 \quad 0.0)\begin{bmatrix} 0.2 & 0.3 & 0.5 \\ 0.1 & 0.3 & 0.6 \\ 0.3 & 0.4 & 0.3 \end{bmatrix} = (0.2 \quad 0.3 \quad 0.5) = \pi_1$$

Chinese, Greek or Italian?

Now the answer:

- Where will Graham eat tonight?

$$\pi_0 = (1.00 \quad 0.00 \quad 0.00)$$

$$\pi_1 = \pi_0 P = (0.20 \quad 0.30 \quad 0.50)$$

$$\pi_2 = \pi_1 P = (0.22 \quad 0.35 \quad 0.43)$$

- He will most likely (43% probability) eat at the Italian restaurant and least likely (22% probability) at the Chinese restaurant

DTMC Theory in a Nutshell

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG INF FAKULTÄT FÜR INFORMATIK

DTMC Basics

Properties of a stochastic matrix P

- The matrix is square $P = \{p_{ij}\}_{n \times n}$
- All elements are probabilities $0 \le p_{ij} \le 1$
- The sum of all outgoing transitions of each state is 1 $\sum_{j=1}^{n} p_{ij} = 1$

Properties of a stochastic vector π

- All elements are probabilities $0 \le \pi_i \le 1$
- The sum of all state probabilities is 1 $\sum_{i=1}^{n} \pi_i = 1$

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG INF FAKULTÄT FÜR INFORMATIK

Discrete–Time Markov Chains

Consider a system with n possible discrete states s_i

Transition probability matrix $P_{(nxn)}$

- Contains the probabilities to change from one state to the next in one step

$$p_{ij} = P(s_{k+1} = j \mid s_k = i)$$

Initial probability vector $\pi_{(n)}$

- Contains the initial probability of each system state

$$\pi_{0,i} = P(s_0 = i)$$

Graphical representation

- Interpret P as the adjacency matrix of an annotated directed graph

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG INF FAKULTÄT FÜR INFORMATIK

Solving DTMCs – The Power Method

An iterative method to compute the transient and steady state probabilities of a discrete time Markov chain

Transient solution:

- system state probability at a particular point in the future:

$$\pi_j^{k+1} = \sum_{i=1\ldots n} \pi_i^k p_{ij}$$

$$\pi_{k+1} = \pi_k P$$

Steady state solution π:

= system state probability after infinite number of steps

= state probabilities that do not change anymore in one step

- π_k converges to π to satisfy the following equations:

$$\pi_{k+1} = \pi_k$$

$$\pi = \pi P$$

Example 2: Gamblers Ruin

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG

INF FAKULTÄT FÜR INFORMATIK

Gamblers Ruin – A Random Walk

Two gamblers own a combined fortune of 200€

- Initially each gambler owns 100€

They throw a coin, betting 50€ every time

- Heads – Gambler A wins 50€ from gambler B
- Tails – Gambler B wins 50€ from gambler A

The game ends when either of the gamblers is broke!

Possible questions

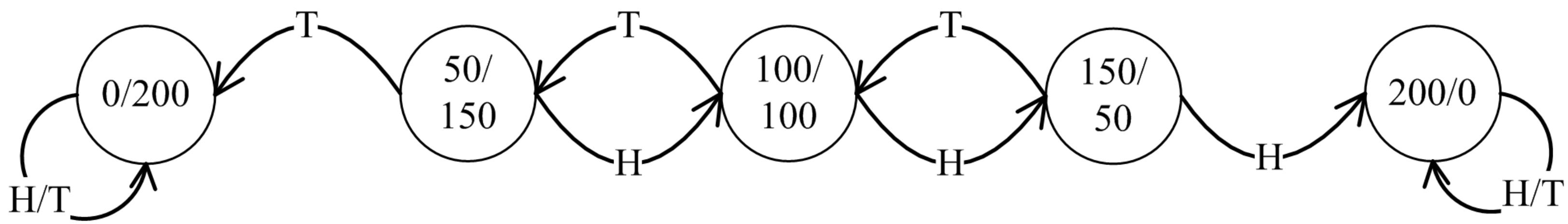
- Who will win the game most likely?
- What happens if the coin is not fair?

Gamblers Ruin – A Random Walk

Possible system states

State	Gambler A	Gambler B
0 / 200	0€	200€
50 / 150	50€	150€
100 / 100	100€	100€
150 / 50	150€	50€
200 / 0	200€	0€

Graphical representation of the system state space



OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG INF FAKULTÄT FÜR INFORMATIK

Gamblers Ruin – A Random Walk

Mathematical representation

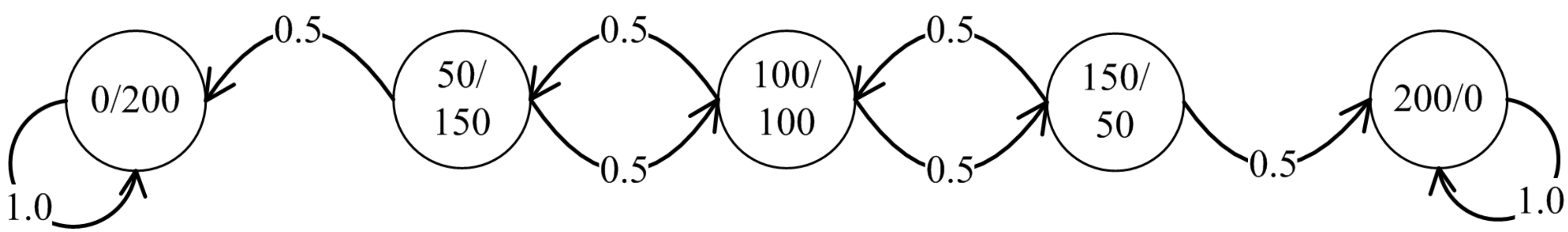
- Each coin throw corresponds to a state change
- Assuming we have a fair coin

$$P = \begin{bmatrix} 1.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.5 & 0.0 & 0.5 & 0.0 & 0.0 \\ 0.0 & 0.5 & 0.0 & 0.5 & 0.0 \\ 0.0 & 0.0 & 0.5 & 0.0 & 0.5 \\ 0.0 & 0.0 & 0.0 & 0.0 & 1.0 \end{bmatrix}$$

$$\pi = (\pi_{0/200} \quad \pi_{50/150} \quad \pi_{100/100} \quad \pi_{150/50} \quad \pi_{200/0})$$

$$\pi_0 = (0.0 \quad 0.0 \quad 1.0 \quad 0.0 \quad 0.0)$$

Graphical representation



OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG INF FAKULTÄT FÜR INFORMATIK

Gamblers Ruin – A Random Walk

What is the probability for each of the gamblers to win the game after a certain number of coin throws?

- For a fair coin, each of the gamblers will have won with equal probability after every step

$$
\begin{aligned}
\pi_0 &= (0.0 \quad 0.0 \quad 1.0 \quad 0.0 \quad 0.0) \\
\pi_1 &= (0.0 \quad 0.5 \quad 0.0 \quad 0.5 \quad 0.0) \\
\pi_2 &= (0.25 \quad 0.0 \quad 0.5 \quad 0.0 \quad 0.25) \\
\pi_3 &= (0.25 \quad 0.25 \quad 0.0 \quad 0.25 \quad 0.25) \\
\pi_4 &= (0.375 \quad 0.0 \quad 0.25 \quad 0.0 \quad 0.375) \\
&\vdots \\
\pi_k &= (0.5 \quad 0.0 \quad 0.0 \quad 0.0 \quad 0.5)
\end{aligned}
$$

- What will happen if the coin is not fair? $P(Heads) = 0.6$
- How about a different initial distribution of the money?

Example 3: Google Page Rank

Google's PageRank

The initial question

- How to rank the results of a query with a particular key phrase?

… can be "reduced" to the question

- How to rank all the websites in the WWW?

Assumptions:

- A user on the internet surfs by clicking links most of the time
- Sometimes he directly enters a URL

Implications:

- A website is more important, the more links point to it
- A link from an important web site (e.g. Wikipedia) is more important than a link from a less important web site (e.g. a student's hompage)

This random walk can be represented by a discrete–time Markov chain (… and it really is done that way!)

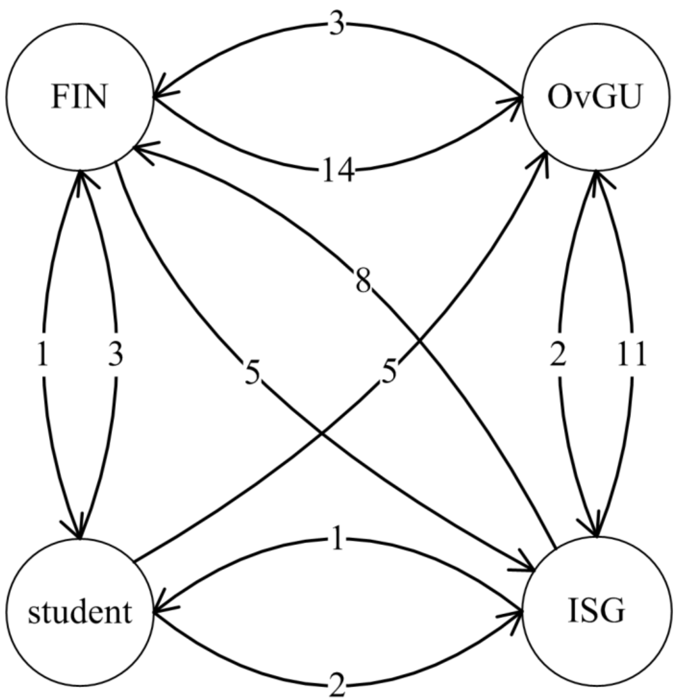
OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG INF FAKULTÄT FÜR INFORMATIK

Google's PageRank

Consider the following system of four websites

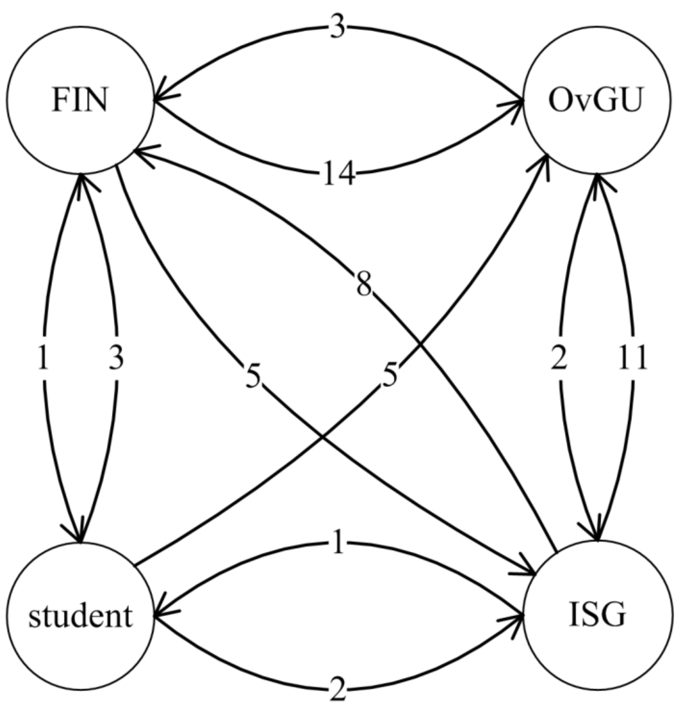
- The websites of OvGU, FIN, and ISG and …
- … a particular student's personal home page

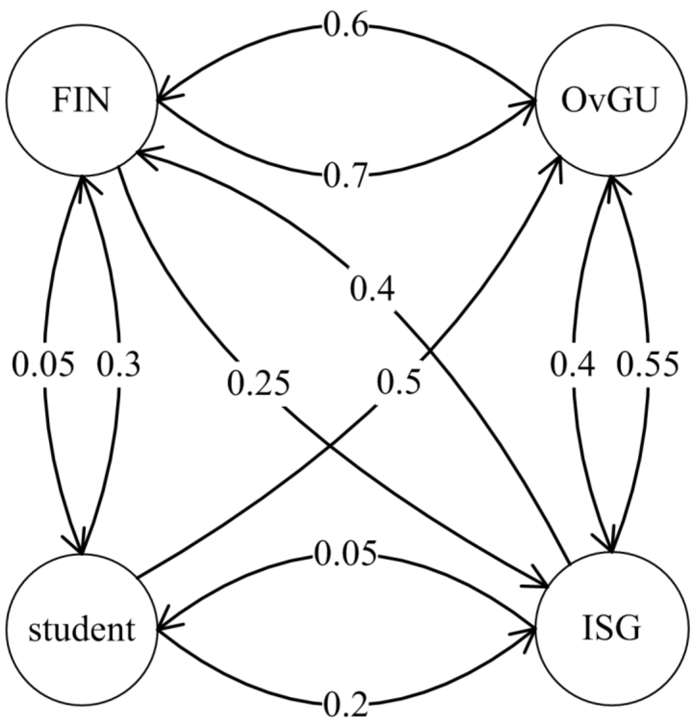
Assume the following link structure in this system



Google's PageRank

Normalize the edge weights by dividing each one by the total number of outgoing links of the web-site



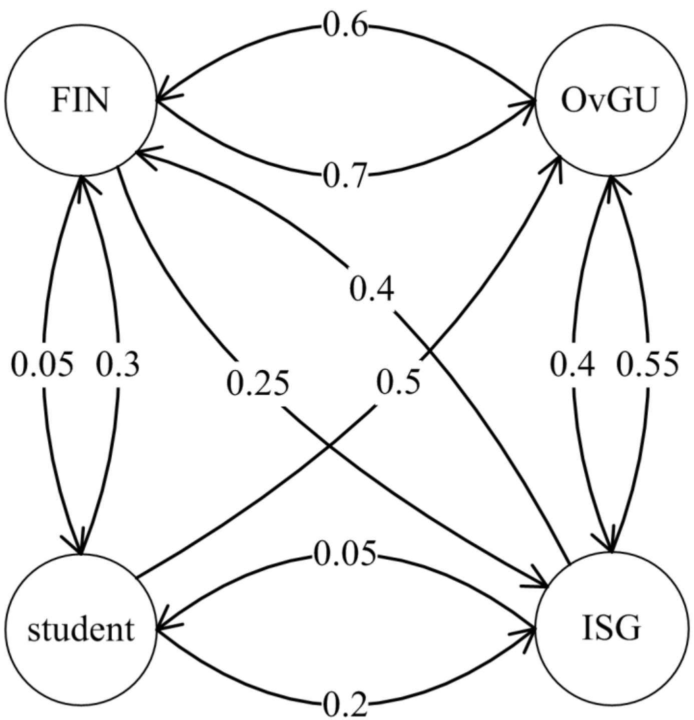


This is already a DTMC!

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG INF FAKULTÄT FÜR INFORMATIK

Google's PageRank

The matrix P_{link} contains the probabilities to go from one website to the next by following an existing link



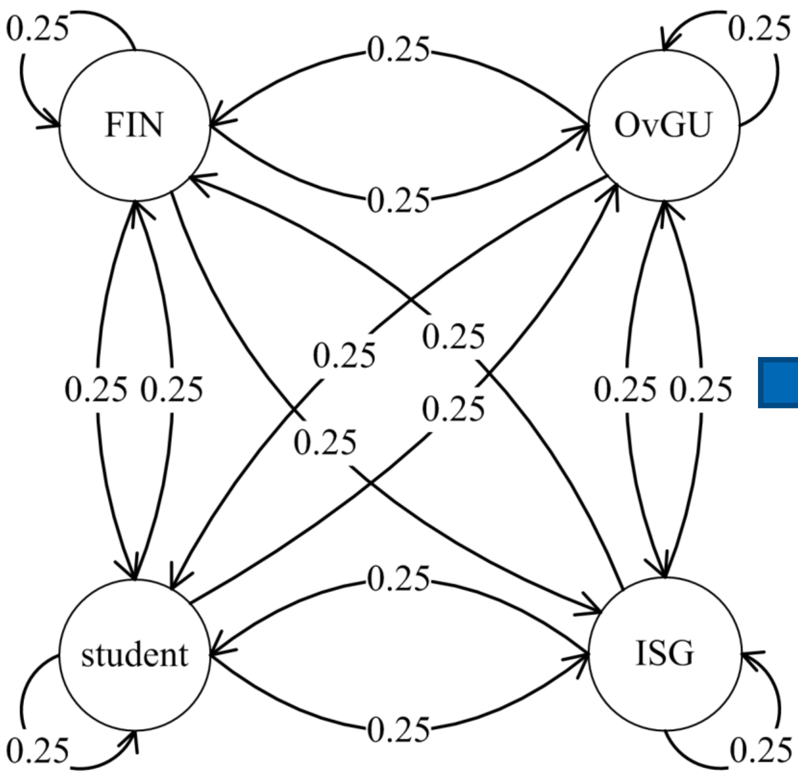
$$\pi = \left(\pi_{FIN} \quad \pi_{OvGU} \quad \pi_{ISG} \quad \pi_{student}\right)$$

$$P_{link} = \begin{bmatrix} 0.0 & 0.7 & 0.25 & 0.05 \\ 0.6 & 0.0 & 0.4 & 0.0 \\ 0.4 & 0.55 & 0.0 & 0.05 \\ 0.3 & 0.5 & 0.2 & 0.0 \end{bmatrix}$$

Mathematical representation

Google's PageRank

The fully connected matrix P_{full} contains the probabilities to go from one website to the next by directly entering a URL



$$P_{full} = \begin{bmatrix} 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \end{bmatrix}$$

Mathematical representation

Google's PageRank

Quantifying the assumptions

- Clicking links most of the time → 85%
- Sometimes directly entering a URL → 15%

$$P = 0.85 * P_{link} + 0.15 * P_{full}$$

… leads to a combined matrix P that contains the probabilities to go from one website to the next by entering a URL or following a link

$$P = \begin{bmatrix} 0.0375 & 0.6325 & 0.25 & 0.08 \\ 0.5475 & 0.0375 & 0.3775 & 0.0375 \\ 0.3775 & 0.505 & 0.0375 & 0.08 \\ 0.2925 & 0.4625 & 0.2075 & 0.0375 \end{bmatrix}$$

Google's PageRank

The steady state solution of this discrete-time Markov chain represents the probability for a random surfer to be on each website

$$\pi = (\pi_{FIN} \quad \pi_{OvGU} \quad \pi_{ISG} \quad \pi_{student})$$

$$\pi_0 = (1.0 \quad 0.0 \quad 0.0 \quad 0.0)$$

$$\pi_1 = (0.0375 \quad 0.6325 \quad 0.25 \quad 0.08)$$

$$\pi_2 = (0.4655 \quad 0.2107 \quad 0.2741 \quad 0.0497)$$

$$\vdots$$

$$\pi_k = (0.3248 \quad 0.3706 \quad 0.243 \quad 0.0616)$$

The larger this probability, the higher the PageRank!

- The OvGU website has the highest rank
- The student's website has the lowest rank

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG INF FAKULTÄT FÜR INFORMATIK

Google's PageRank

Some remarks

- The Google matrix had an n=2.7 Billion in October 2002
- We presented you with the original algorithm, which is susceptible to manipulation
- The currently running algorithm contains some modifications to counteract manipulations and generate more useful results
- There are numerous resources on the web regarding PageRank

To learn more about DTMCs, HMMs and HnMMs

… visit our Master module "Applied Discrete Modeling"

Learning Goals

You are now able to solve question 8 of the ItS Exam