Chapter 3

Markov Chains & PageRank

Systems are often modeled by automata, and discrete events are transitions from one state to another. In this part of the course we want to analyze such discrete events systems. We assume that events are stochastic, and we want to know how our system behaves on average.

If the events happen in discrete time (for example, there is an event every day), the tool to model the system is called Discrete Time Markov Chain (DTMC). The system transitions from one state to another, according to the probabilities of the transitions at the current state.

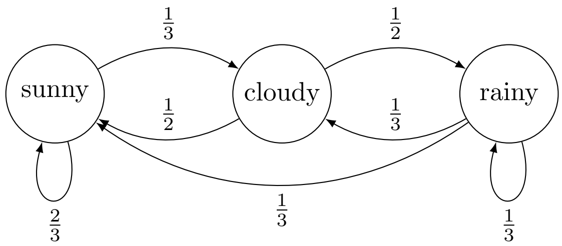

Figure 3.1: According to a self-proclaimed weather expert, the above graph models the weather in Zürich. On any given day, the weather is either *sunny*, *cloudy*, or *rainy*. The probability to have a cloudy day after a sunny day is $\frac{1}{3}$. In the context of Markov chains the nodes, in this case *sunny*, *rainy*, and *cloudy*, are called the *states* of the Markov chain.

Remarks:

- Figure 3.1 above is an example of a *Markov chain*—see the next section for a formal definition.

- If the weather is currently sunny, the predictions for the next few days according to the model from Figure 3.1 are:

Day	sunny	cloudy	rainy
0	1	0	0
1	$\frac{2}{3}$	$\frac{1}{3}$	0
2	0.611	0.222	0.167
3	0.574	0.259	0.167
4	0.568	0.247	0.185
$\vdots$	$\vdots$	$\vdots$	$\vdots$

3.1 Markov Chains

Markov chains are a tool for studying *stochastic processes* that evolve over time.

Definition 3.2 (Markov Chain)**.** *Let S be a finite or countably infinite set of* ***states****. A* ***(discrete time) Markov chain*** *is a sequence of random variables $X_0, X_1, X_2, \ldots \in S$ that satisfies the Markov property (see below).*

Definition 3.3 (Markov Property)**.** *A sequence (X_t) of random variables has the* ***Markov property*** *if for all t, the probability distribution for X_{t+1} depends only on X_t, but not on $X_{t-1}, \ldots, X_0$. More formally, for all $t \in \mathbb{N}_{>0}$ and $s_0, \ldots, s_{t+1} \in S$ it holds that $\Pr[X_{t+1} = s_{t+1} \mid X_0 = s_0, X_1 = s_1, \ldots, X_t = s_t] = \Pr[X_{t+1} = s_{t+1} \mid X_t = s_t]$.*

Remarks:

- A sequence of random variables is also called a *discrete time stochastic process*. Processes that satisfy the Markov property are also called *memoryless*.

- The probability distribution of X_0 does not depend on a previous state (since there is none). It is called the *initial distribution*, and we denote it by the vector $q_0 = (q_{0,s})_{s \in S}$ with the entries $\Pr[X_0 = s]$ for every state $s \in S$. If the first day is sunny, the initial distribution is $q_0 = (1, 0, 0)$.

Definition 3.4 (Time Homogeneous Markov Chains)**.** *A Markov chain is* ***time homogeneous*** *if $\Pr[X_{t+1} = s_{t+1} \mid X_t = s_t]$ is independent of t, and in that case $p_{i,j} = \Pr[X_{t+1} = j \mid X_t = i]$ is well defined.*

Remarks:

- We will only consider time homogeneous Markov chains.
- Markov chains are often modeled using *directed* graphs, as in Figure 3.1. The states are represented as nodes, and an edge from state i to state j is weighted with probability $p_{i,j}$.
- Just like directed graphs, Markov chains can be written in matrix form (using the adjacency matrix). In this context, the matrix is called the *transition matrix*, and we denote it by P. For the example from Figure 3.1, the transition matrix is:

		to		
		sunny	cloudy	rainy
from	sunny	2/3	1/3	0
	cloudy	1/2	0	1/2
	rainy	1/3	1/3	1/3

- Let $q_t = (q_{t,i})_{i \in S}$ be the probability distribution on S for time t, i.e., $q_{t,i} = \Pr[X_t = i]$. The probability to be in state j at time $t+1$ is $q_{t+1,j} = \sum_{i \in S} \Pr[X_t = i] \cdot \Pr[X_{t+1} = j \,|\, X_t = i] = \sum_{i \in S} q_{t,i} \cdot p_{i,j}$. This can be written as the vector-matrix-multiplication $q_{t+1} = q_t \cdot P$.
- The state distribution at time t is $q_t = q_0 \cdot P^t$. We denote by $p_{i,j}^{(t)}$ the entry at position i, j in P^t, i.e., the probability of reaching j from i in t steps.
- If we start a Markov chain in a single state (with probability 1), we do a random walk.

Definition 3.5 (Random Walk)**.** *Let $G = (V, E)$ be a directed graph, and let $\omega : E \to [0, 1]$ be a weight function so that $\sum_{v:(u,v) \in E} \omega(u, v) = 1$ for all nodes u. Let $u \in V$ be the* ***starting node****. A* ***weighted random walk on*** *G* ***starting at*** *u is the following discrete Markov chain in discrete time. Beginning with $X_0 = u$, in every step t, the node X_{t+1} is chosen according to the weights $\omega(X_t, v)$, where v are the neighbors of X_t. If G is undirected and unweighted, then X_{t+1} is chosen uniformly at random among X_t's neighbors and the random walk is called* ***simple****.*

Remarks:

- In Section 3.5 we will study some simple random walks.
- Now that we understand the definitions, we want to do some calculations. E.g., if it is sunny today, how long will it stay sunny?

3.2 Hitting Time & Arrival Probability

Definition 3.6 (Sojourn Time)**.** *The* ***sojourn time*** *T_i of state i is the time the process stays in state i.*

Remarks:

- It holds that $\Pr[T_i = k] = p_{i,i}^{k-1} \cdot (1 - p_{i,i})$, i.e., T_i is *geometrically distributed*. For example $\mathbb{E}[T_{\text{sunny}}] = 3$.
- The sojourn time T_i does not depend on the time the process has spent in state i already (memoryless property). The geometric distribution is the only discrete distribution that is memoryless.
- If it is currently sunny, how long does it take until we see the first rainy day?

Definition 3.7 (Hitting Time & Arrival Probability)**.** *Let i and j be two states. The* ***hitting time*** *$T_{i,j}$ is the random variable counting the number of steps until visiting j the first time when starting from state i, i.e., the value of $T_{i,j}$ is the smallest integer $t \geq 1$ for which $X_t = j$ under the condition that $X_0 = i$. The* ***expected hitting time*** *from i to j is the expected value $h_{i,j} = \mathbb{E}[T_{i,j}]$. The* ***arrival probability*** *from i to j is the probability $f_{i,j} = \Pr[T_{i,j} < \infty]$.*

Remarks:

- The time $c_{i,j} = h_{i,j} + h_{j,i}$ is referred to as the *commute time* between i and j.
- The following lemma states that the expected hitting time can be computed by solving a system of linear equations.

Lemma 3.8. *If $h_{i,j}$ exists for all $i, j \in S$, then the expected hitting times are*

$$h_{i,j} = 1 + \sum_{k \neq j} p_{i,k} h_{k,j} .$$

Proof. Plugging in the definition of $h_{i,j}$ and applying the law of total probability we get that

$$h_{i,j} = \mathbb{E}[T_{i,j}] = \sum_{k \in S} \mathbb{E}[T_{i,j} \mid X_1 = k] \cdot p_{i,k} .$$

Taking the j^{th} term out, we obtain

$$\begin{aligned} h_{i,j} &= \mathbb{E}[T_{i,j} \mid X_1 = j] \cdot p_{i,j} + \sum_{k \neq j} \mathbb{E}[T_{i,j} \mid X_1 = k] \cdot p_{i,k} \\ &= 1 \cdot p_{i,j} + \sum_{k \neq j} (1 + \mathbb{E}[T_{k,j}]) \cdot p_{i,k} . \end{aligned}$$

Since $p_{i,j}$ together with all the values $p_{i,k}$ sum up to 1, we can simplify to

$$h_{i,j} = 1 + \sum_{k \neq j} \mathbb{E}[T_{k,j}] \cdot p_{i,k} = 1 + \sum_{k \neq j} p_{i,k} h_{k,j} . \qquad \square$$

Remarks:

- On a sunny day it takes in expectation 8 days until it starts raining.
- Lemma 3.9 for the arrival probabilities can be established similarly to Lemma 3.8.

Lemma 3.9. *For all $i, j \in S$, the arrival probability is*

$$f_{i,j} = p_{i,j} + \sum_{k \neq j} p_{i,k} f_{k,j} \,.$$

3.3 Stationary Distribution & Ergodicity

What is the "climate" in Zürich? Often one is particularly interested in the long term behavior of Markov chains and random walks. The mathematical notion that captures a Markov chain's long term behavior is the *stationary distribution*, which we will introduce and study in the following.

Remarks:

- The entries in P^t contain the probability of entering a certain weather condition (state). What happens for large values of t? The matrix seems to converge!

$$P^3 \approx \begin{pmatrix} 0.574 & 0.259 & 0.167 \\ 0.556 & 0.222 & 0.222 \\ 0.537 & 0.259 & 0.204 \end{pmatrix} \quad P^{10} \approx \begin{pmatrix} 0.563 & 0.250 & 0.187 \\ 0.562 & 0.250 & 0.187 \\ 0.562 & 0.250 & 0.188 \end{pmatrix}$$

- No matter what the initial weather q_0 is, the product $q_0 \cdot P^t$ seems to approach $\tilde{q} \approx (0.563, 0.250, 0.188)$ as t grows. Moreover, if we multiply the vector $\tilde{q}$ with P we *almost* get $\tilde{q}$ again. In other words, $\tilde{q}$ is almost an eigenvector of P with eigenvalue 1.

Definition 3.10 (Stationary Distribution)**.** *A distribution π over the states is called* ***stationary distribution*** *of the Markov chain with transition matrix P if $\pi = \pi \cdot P$.*

Remarks:

- Our weather Markov chain converges towards $\pi = (9/16, 4/16, 3/16)$, which is an eigenvector of P with eigenvalue 1. We conclude that in the long run, 9 out of 16 days are sunny in Zürich. The weather model appears to be not as accurate as the weather expert led us to believe ...
- Consider the sequence $q_i = q_{i-1} \cdot P$, where q_0 is the initial distribution. In general, this sequence does not necessarily converge as t grows. However, if it does converge to some distribution π, then it must hold that $\pi = \pi \cdot P$.

Lemma 3.11. *The transition matrix of every Markov chain has a left eigenvector with eigenvalue* 1*.*

Proof. Let P be the transition matrix of a Markov chain, and denote by $e = (1, \ldots, 1)^\top$ the all-ones vector. Because in P the entries in each row sum up to 1 (P is row stochastic), it holds that $Pe = e$. Denoting by I the identity matrix, it follows that $(P - I)e = 0$. In other words, e is an eigenvector with eigenvalue 0 for $(P-I)$, which implies that $(P-I)$ is singular, i.e., not invertible. Thus, also $(P - I)^\top$ is singular, and it follows that there is a vector $\pi \neq 0$ so that $0 = (P - I)^\top \pi = P^\top \pi - I\pi$. Transposing and rearranging we obtain that $\pi^\top P = \pi^\top$, as desired. □

Remarks:

- Using Brouwer's fixed point theorem one can show that there is also a left eigenvector π that corresponds to a probability distribution.
- The stationary distribution is not necessarily unique, see Figure 3.12. The issue is that some states are not reachable from all other states.

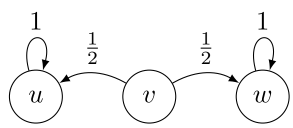

Figure 3.12: This Markov chain has infinitely many stationary distributions, for example $\pi_0 = (1, 0, 0)$, $\pi_1 = (0, 0, 1)$, and $\pi_{0.8} = (0.2, 0, 0.8)$. The states u and w are called *absorbing* states, since they are never left once they are entered.

Definition 3.13 (Irreducible Markov Chains)**.** *A Markov chain is* ***irreducible*** *if all states are reachable from all other states. That is, if for all $i, j \in S$ there is some $t \in \mathbb{N}$, such that $p_{i,j}^{(t)} > 0$.*

Lemma 3.14. *In an irreducible Markov chain it holds that $h_{i,j} < \infty$ for all states i, j.*

Proof. Fix some state j, and observe that due to Definition 3.13 for every $s \in S$, there is some t_s so that $p_{s,j}^{(t_s)} > 0$. Denote by $t = \max\{t_s \mid s \in S\}$ the largest such value. State j can be reached from every state in at most t steps. We partition the random walk into *trials* of t successive steps. Within each trial, state j is reached with probability at least $p = \min\{p_{s,j}^{t_s} \mid s \in S\}$. The number of trials until the random walk reaches j is thus upper bounded by a geometric distribution with parameter p. It follows that at most $1/p$ trials are necessary to reach j, and we conclude that $h_{i,j} \leq t/p$ for any i. □

Remarks:

- Similarly, it follows that $f_{i,j} = 1$ for all states i, j if the Markov chain is irreducible.

Lemma 3.15. *Every finite irreducible Markov chain has a unique stationary distribution π. The distribution is $\pi_j = \frac{1}{h_{j,j}}$ for all $j \in S$.*

Proof. Denote by P the transition matrix of an irreducible Markov chain. Let $\pi \neq 0$ be a left eigenvector of P with eigenvalue 1 as promised by Lemma 3.11. Denote further by $h_{i,j}$ the expected hitting times guaranteed by Lemma 3.14.

We first consider the case that $\sum_i \pi_i \neq 0$ and w.l.o.g. assume that $\sum_i \pi_i = 1$. Due to Lemma 3.8 it holds that for any $j \in S$,

$$\pi_i h_{i,j} = \pi_i \left(1 + \sum_{k \neq j} p_{i,k} h_{k,j}\right) \text{ for all } i \in S\,.$$

Since $\sum_i \pi_i = 1$, summing up those equations over all i yields

$$\begin{aligned} \pi_j h_{j,j} + \sum_{i \neq j} \pi_i h_{i,j} &= 1 + \sum_i \pi_i \sum_{k \neq j} p_{i,k} h_{k,j} \\ &= 1 + \sum_{k \neq j} h_{k,j} \sum_i \pi_i p_{i,k}\,, \end{aligned}$$

by switching the summation on the right hand side. Since π is an eigenvector with eigenvalue 1, it holds that $\sum_i \pi_i p_{i,k} = \pi_k$, and thus the equation becomes

$$\pi_j h_{j,j} + \sum_{i \neq j} \pi_i h_{i,j} = 1 + \sum_{k \neq j} h_{k,j} \pi_k\,.$$

Noting that all $h_{j,j} > 1$ we conclude that $\pi_j = 1/h_{j,j}$, as desired. In the remaining case where $\sum_i \pi_i = 0$, the equation turns into

$$\pi_j h_{j,j} + \sum_{i \neq j} \pi_i h_{i,j} = \sum_{k \neq j} h_{k,j} \pi_k\,,$$

yielding that $\pi_j = 0$ for all j. This contradicts that π is an eigenvector. □

Remarks:

- Irreducible Markov chains with an infinite number of states do not necessarily have a stationary distribution.
- Depending on the choice of the initial distribution, even an irreducible Markov chain does not necessarily converge towards its stationary distribution, see Figure 3.16.

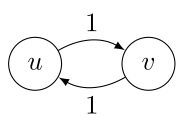

Figure 3.16: This Markov chain is irreducible, and has the unique stationary distribution $\pi = (0.5, 0.5)$. In this particular chain, each state can only be reached every other step, or in other words, both states have *period* 2. Therefore, the initial distribution is attained in every second step, and only $q_0 = \pi$ "converges" towards the stationary distribution.

Definition 3.17 (Aperiodic Markov Chains)**.** *The* ***period*** *of a state $j \in S$ is the largest $\xi \in \mathbb{N}$ such that*

$$\{n \in \mathbb{N} \mid p_{j,j}^{(n)} > 0\} \subseteq \{i \cdot \xi \mid i \in \mathbb{N}\}$$

A state with period $\xi = 1$ is called ***aperiodic****, and the Markov chain is* ***aperiodic*** *if all its states are.*

Remarks:

- One can show that if the Markov chain is irreducible, then all states have the same period.
- A state j with $p_{j,j} > 0$ is trivially aperiodic.
- If $p_{j,j} = 0$, then one can check whether state j is aperiodic by testing, as illustrated in Figure 3.18, if the following holds: Does j lie on two directed cycles of lengths k and l (counting the edges in the chain) so that k and l are relatively prime, i.e., have a greatest common divisor of 1? Or, using the k^{th} and l^{th} powers of P, are there relatively prime k and l such that both $p_{j,j}^{(k)}$ and $p_{j,j}^{(l)} > 0$?

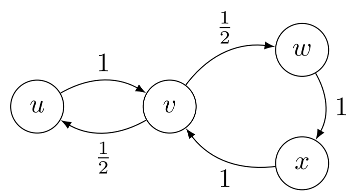

Figure 3.18: Starting at state v there is a cycle $v \to u \to v$ using 2 edges, and a cycle $v \to w \to x \to v$ using 3 edges. Because 2 and 3 are relatively prime, the state v is aperiodic.

Definition 3.19 (Ergodic Markov Chains)**.** *If a finite Markov chain is irreducible and aperiodic, then it is called* ***ergodic****.*

Theorem 3.20. *If a Markov chain is ergodic it holds that*

$$\lim_{t \to \infty} q_t = \pi,$$

where π is the unique stationary distribution of the chain.

Remarks:

- The theorem holds regardless of the initial distribution.
- The stationary distribution of ergodic Markov chains can thus be approximated efficiently, namely by successively multiplying a vector with a matrix instead of computing the powers of a matrix.

3.4 PageRank Algorithm

Google's PageRank algorithm is based on a Markov chain obtained from a variant of a random walk.

Remarks:

- Google provides search results that match the user's search terms. Under the hood Google maintains a ranking among websites to make sure "better" or "more important" websites appear early in the search results. Instead of solving the whole problem at once, this ranking is first established globally (independent of the search terms), and only later websites matching the search query are sorted according to some rank. In this section we focus on the ranking part.

- The first step to ranking websites is to crawl the web graph, i.e., a directed graph in which the nodes are websites, and an edge (u, v) indicates that website u contains a hyperlink to website v.

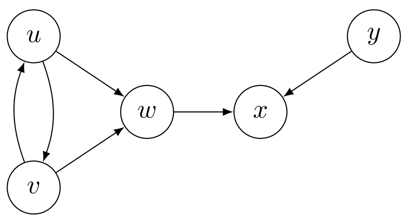

Figure 3.21: An example of a web graph with 5 websites. Website x does not link to any other website, i.e., x is a *sink*.

- A naïve approach is to rank the sites by the number of incoming hyperlinks. In the example from Figure 3.21 this yields the same rank for websites w and x. One could, however, argue that the link from w to x means that x is more important than w.

- Google's idea is to model a *random surfer* who follows hyperlinks in the web graph, i.e., performs a simple random walk. After sufficiently many steps, the websites can be ranked by how many times they were visited. The intuition is that websites are visited more often if they are linked by many other sites, which should be a good measure of how important a website is.

- Since the walk is directed, the random surfer can get stuck in sinks (nodes with no outgoing edges). To fix this issue, a random website is chosen for the next step whenever the random surfer reaches a sink.

- Let us denote the *random surfer matrix* describing this simple random walk by W.

- Simulating the simple random walk described by W to find a stationary distribution is not feasible: There are over 1 billion websites—meaning that a lot of steps have to be simulated to get a good estimation of the stationary distribution. Using our knowledge about

Markov chains we can simulate many random walks at once by repeatedly multiplying some initial distribution q_0 with W.

- There is no guarantee that this process converges to a stationary distribution. We know that this can be fixed by making the Markov chain ergodic.

- One way to make a Markov chain ergodic is to insert an edge between every two nodes.

Definition 3.22 (Google Matrix)**.** *Let W be a random surfer matrix, and let $\alpha \in (0, 1)$ be a constant. Denote further by R the matrix in which all entries are $1/n$. The following matrix M is called the **Google Matrix**:*

$$M = \alpha \cdot W + (1 - \alpha) \cdot R\,.$$

Remarks:

- The intuition behind R is that in every step, with probability $1 - \alpha$, the random surfer "gets bored" by the current website and surfs to a new random site.

- While the R-component in M ensures that the Markov chain converges, it also changes the stationary distribution. To ensure the impact is not too large, α should be chosen close to 1. A typical value for α is 0.85.

- The rate at which the process converges depends on the magnitude of M's second largest eigenvalue. One can show that for M the second largest eigenvalue is at most α, and that the error decreases by a factor of α in each step.

- In the example from Figure 3.21, the page ranks are

Website	Rank
x	0.384615
w	0.230769
u	0.153846
v	0.153846
y	0.0769231

- This initial version of the PageRank algorithm worked well at the time it was invented. However, it can be (and has been) fooled. Consider the following example.

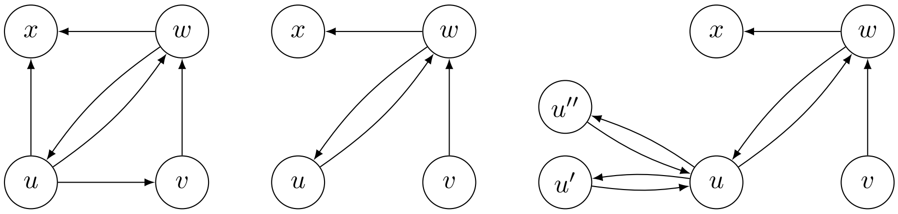

Figure 3.23: Website u wants to improve its PageRank, which is ≈ 0.23 in the initial setting on the left. First, all outgoing links to websites that do not link back are removed. The PageRank improves to ≈ 0.27. In a *Sybil attack* (right) the owner of u creates fake websites u' and u'' whose purpose is to exchange links with u. Moreover, the new websites increase the probability to visit u after a sink. Now, website u is the highest ranked site in the network with a rank of ≈ 0.41.

- Attacks where a single party pretends to be more than one individual are called *Sybil Attacks*.
- It is unknown how exactly Google ranks websites today, and specifically how the engineers at Google mitigate the effects of attacks.
- A different kind of attack on Google is *Google bombing*. This attack relies on the fact that the search terms for which a website v is considered relevant also take the *anchor text* of hyperlinks to website v into account. If, for instance, many websites link to `http://www.ethz.ch` using the anchor text "Smartest People Alive", then a search query for smart people might end up presenting ETH's website.

3.5 Simple Random Walks

In this section, all random walks are considered to be simple. This means that the edges are undirected, and the node for the next step is chosen uniformly at random among the current node's neighbors.

Lemma 3.24. *Let G be a graph with m edges. The stationary distribution π of any simple random walk on G is*

$$\pi_u = \frac{\delta(u)}{2m}\,.$$

Proof. The Markov chain underlying the random walk is irreducible, with Lemma 3.15 it has a unique stationary distribution. We first verify that π from above satisfies the equation $\pi = \pi \cdot P$ from Definition 3.10:

$$\pi_u = \sum_{v\,\in\,N(u)} \pi_v \cdot p_{v,u} = \sum_{v\,\in\,N(u)} \frac{\delta(v)}{2m} \cdot \frac{1}{\delta(v)} = \frac{\delta(u)}{2m}\,,$$

for some arbitrary node $u \in V$. Since $\sum_{v\in V} \delta(v) = 2m$, π is a distribution. This proves that π is the unique stationary distribution. $\square$

Remarks:

- It follows from Lemma 3.15 that for a simple random walk, $h_{u,u}$ is $2m/\delta(u)$.
- The *cover time* $\mathrm{cov}(v)$ is the expected number of steps until all nodes in G were visited at least once, starting at v.
- One could use the following Markov chain to compute the cover time of a simple random walk on the graph $G = (V, E)$. The set of states is $\{(v, I) \mid v \in V \text{ and } I \in 2^V\} \cup \{t\}$, where v denotes the current state, I denotes the visited states, and t is an additional sink state. The probabilities $p_{(v,I),(w,I\cup\{w\})}$ are 0 if either $I = V$ or $\{v, w\} \notin E$, and $1/\delta(v)$ else (if $\{v, w\} \in E$). Additionally, each state (v, V) has an edge to the sink t with $p_{(v,V),\,t} = 1$. Then, the cover time is $\mathrm{cov}(v) = h_{(v,\{v\}),\,t} - 1$.

Lemma 3.25. *Let $G = (V, E)$ be a graph with n nodes and m edges. It holds that* $\mathrm{cov}(s) < 4m(n - 1)$ *for any starting node $s \in V$.*

Proof. Let $\{u, v\} \in E$ be an edge. It holds that

$$\frac{2m}{\delta(u)} = h_{u,u} = \frac{1}{\delta(u)} \sum_{w \in N(u)} (h_{w,u} + 1)\,,$$

and thus it must be true that $h_{u,v} < 2m$. Next, observe that it is possible to traverse all nodes in G by using no more than $2n-2$ edges, e.g., by traversing a spanning tree rooted at s. Since $h_{u,v} < 2m$ holds for every edge $\{u, v\}$ used in the traversal, it follows that $\mathrm{cov}(s) < (2n-2) \cdot 2m = 4m(n-1)$, as desired. □

Remarks:

- Consider the resistor network obtained from G by replacing every edge with a 1Ω resistor. Let u and v be two nodes in the resistor network. It can be shown that the commute time $c_{u,v} = 2m \cdot R(u, v)$, where $R(u, v)$ denotes the effective resistance between u and v.
- Foster's Theorem states that for every connected graph $G = (V, E)$ with n nodes,

  $$\sum_{(u,v)\,\in\, E} R(u, v) = n - 1\,,$$

  i.e., that adding/removing an edge in G reduces/increases the effective resistance, respectively.

Chapter Notes

Historic background on the development of Markov chains can be found in [1]. The short version is that in a 1902 paper [9], the theologist Pavel Alekseevich Nekrasov, in his effort to establish free will on a mathematical basis, (falsely) postulated that independence of events is necessary for the law of large numbers. Markov, being an atheist and considering Nekrasov's reasoning an "abuse of mathematics", set out to prove him wrong.

In 1906, Markov published his first findings on chains of pairwise dependent random variables [7]. This work already includes a variant of Theorem 3.20, thus disproving Nekrasov's claim. Markov also studied the notion of irreducibility [8], proving that for irreducible Markov chains 1 is a single eigenvalue and the largest by magnitude. Today, Markov's ideas are widely applied in, e.g., physics, chemistry, and economics.

Markov chains are the basis for queueing theory, an important transport layer concept. Another application in computer science is the PageRank algorithm [10]. The bound on the Google matrix' second eigenvalue is from [5]. Sybil attacks were originally studied in the context of peer to peer systems [3], and PageRank's sensitivity to such attacks was investigated in [2].

The connection from random walks to resistor networks is investigated in depth in [4]. By associating a word with each state, random walks can be used to generate random text [11]. More than 120 "scientific" papers were generated using such methods [6] and later withdrawn by the publishers.

This chapter was written in collaboration with Jochen Seidel.

Bibliography

[1] Gely P. Basharin, Amy N. Langville, and Valeriy A. Naumov. The life and work of a.a. markov. *Linear Algebra and its Applications*, 386:3 – 26, 2004.

[2] Alice Cheng and Eric Friedman. Manipulability of pagerank under sybil strategies, 2006.

[3] John R. Douceur. The sybil attack. In *Peer-to-Peer Systems*, volume 2429 of *LNCS*, pages 251–260. Springer Berlin Heidelberg, 2002.

[4] Peter G. Doyle and J. Laurie Snell. Random walks and electric networks. `http://math.dartmouth.edu/~doyle/docs/walks/walks.pdf`, July 2006. Originally published 1984. Website accessed Sep. 23, 2015.

[5] Taher Haveliwala and Sepandar Kamvar. The second eigenvalue of the google matrix. Technical Report 2003-20, Stanford InfoLab, 2003.

[6] Cyril Labbé and Dominique Labbé. Duplicate and fake publications in the scientific literature: How many SCIgen papers in computer science? *Scientometrics*, 94(1):379–396, 2013.

[7] Andrey Andreyevich Markov. Rasprostranenie zakona bol'shih chisel na velichiny, zavisyaschie drug ot druga. *Izvestiya Fiziko-matematicheskogo obschestva pri Kazanskom universitete*, 2-ya seriya 15 (94):135–156, 1906. (Extension of the law of large numbers to random variables dependent on each other).

[8] Andrey Andreyevich Markov. Rasprostranenie predel'nyh teorem ischisleniya veroyatnostej na summu velichin svyazannyh v cep'. *Zapiski Akademii Nauk po Fiziko-matematicheskomu otdeleniyu*, VIII seriya 25 (3), 1908. (Extension of the limit theorems of probability theory to a sum of variables connected in a chain).

[9] Pavel Alekseevich Nekrasov. The philosophy and logic of science of mass phenomena in human activity, Moscow 1902. In Russian.

[10] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. Technical Report 1999-66, Stanford InfoLab, November 1999.

[11] Jeremy Stribling, Max Krohn, and Dan Aguayo. SCIgen - an automatic CS paper generator. `https://pdos.csail.mit.edu/archive/scigen/`, 2005. Website accessed Sep. 23, 2015.