4e. Łańcuchy Markowa i algorytm PageRank

Grzegorz Kosiorowski

Uniwersytet Ekonomiczny w Krakowie

Łańcuch Markowa

Łańcuch Markowa - definicja nieformalna

Łańcuchem Markowa lub *dyskretnym procesem Markowa* nazywamy zestaw złożony z:
1) V - zbioru możliwych stanów pewnego układu zmieniającego się w czasie. W tej prezentacji oznaczamy $|V| = n$.
2) p - pewnej zasady opisującej prawdopodobieństwa przejścia w pojedynczym kroku czasowym ze stanu do stanu.

Prawdziwą definicję musielibyśmy sformułować w języku procesów stochastycznych, dlatego nie będziemy się tu zagłębiać w formalizmy.

Łańcuch Markowa - interpretacja

Formalnie, łańcuchy Markowa ilustrują sytuacje, w których jakiś układ ewaluuje w ten sposób, że obiekty (ludzie i inne istoty, związki chemiczne, elementy sieci komputerów) zmieniają swój stan, a zmiana ta zależy tylko i wyłącznie od ich aktualnego stanu (tak zwana własność Markowa).
Oczywiście, najlepiej ten warunek spełniają systemy nieożywione, ale zastosowania łańcuchów Markowa nie ograniczają się do nich.
Najbardziej spektakularnym współczesnym zastosowaniem jest analiza sieci, takich jak Internet. Jeden z przykładów ich wykorzystania, słynny algorytm PageRank, omówimy w tej prezentacji.

Graf przejścia

Łańcuchy Markowa najlepiej są opisywane przez tak zwane grafy i macierze przejścia, znane z prezentacji o macierzach grafów.

Graf przejścia

Graf skierowany z wagami $G = (V, E)$ nazywamy *grafem przejścia łańcucha Markowa*, jeśli zbiór wierzchołków $V = \{v_1, \dots, v_n\}$ reprezentuje zbiór stanów łańcucha, krawędź (v_i, v_j) istnieje, gdy możliwe jest przejście ze stanu v_i do stanu v_j, a waga $w(v_i, v_j)$ przypisana takiej krawędzi oznacza prawdopodobieństwo przejścia ze stanu i do stanu j (lub pozostania w tym samym stanie, jeśli $i = j$).

Macierze przejścia

Macierz przejścia

Macierzą przejścia $P(G)$ grafu przejścia G nazywamy macierz kwadratową:

$$P(G) = \begin{bmatrix} p_{11} & \dots & p_{1n} \\ \dots & \dots & \dots \\ p_{n1} & \dots & p_{nn} \end{bmatrix},$$

gdzie p_{ij} to prawdopodobieństwo przejścia ze stanu j do stanu i (lub pozostania w tym samym stanie, jeśli $i = j$).

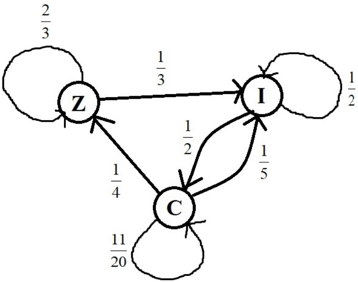
W większości zastosowań, $P(G)$ jest macierzą Markowa, czyli elementy każdej jej kolumny sumują się do 1.

Przykładowe zagadnienie modelowane łańcuchem Markowa

Przykład

W pewnym eksperymencie próbkę komórek poddano działaniu pewnego wirusa i pewnego leku. Co godzinę zapisywano liczbę komórek zdrowych, liczbę komórek zainfekowanych wirusem, ale nie przejawiających objawów choroby i komórek chorych. Średnio, co godzinę $1/3$ komórek zdrowych stawała się zainfekowana, $1/2$ komórek zainfekowanych stawała się chorymi, a z komórek chorych $1/4$ zdrowiała, a $1/5$ wracała do stanu „uśpionej infekcji”. Reszta pozostawała w takim stanie, w jakim była wcześniej. Zapisać graf i macierz przejścia tego układu.

Przykład grafu i macierzy przejścia



Graf przejścia opisanego wcześniej układu (Z-zdrowe, I - zainfekowane, C - chore).

$$P(G) = \begin{bmatrix} \frac{2}{3} & 0 & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{5} \\ 0 & \frac{1}{2} & \frac{11}{20} \end{bmatrix}.$$

Macierz przejścia danego układu dla kolejności stanów (Z,I,C).

Interpretacja macierzy przejścia

Jeśli mamy dany wektor obecnego stanu układu (czyli wektor zapisujący, ilu uczestników układu jest w każdym ze stanów v_i), to mnożąc go przez macierz przejścia otrzymamy wektor kolejnego stanu układu.
Na przykład, dla rozważanego problemu, jeśli założymy, że w pewnym momencie było w układzie 60 komórek zdrowych, 40 komórek zainfekowanych i 100 komórek chorych, możemy obliczyć:

$$P(G) \cdot \begin{bmatrix} 60 \\ 40 \\ 100 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & 0 & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{5} \\ 0 & \frac{1}{2} & \frac{11}{20} \end{bmatrix} \cdot \begin{bmatrix} 60 \\ 40 \\ 100 \end{bmatrix} = \begin{bmatrix} 65 \\ 60 \\ 75 \end{bmatrix}$$

Zatem po godzinie możemy się spodziewać, że w układzie będzie 65 komórek zdrowych, 60 zainfekowanych i 75 chorych.

Stan graniczny

Generalnie, jeśli P jest macierzą przejścia modelującą pewien proces Markowa, a w_0 - wektorem początkowym, czyli takim, w którego i-ta współrzędna odpowiada liczbie badanych obiektów, które są w stanie i w momencie 0, to wtedy wektor stanu układu w momencie k (w_k) można wyrazić równością:

$$w_k = P^k w_0.$$

Czy opisywany układ się „na dłuższą metę" stabilizuje, czyli czy od pewnego momentu wektor stanu układu przestaje się zmieniać (lub zmienia się bardzo nieznacznie)? Jeśli tak jest, osiągnięty stan nazywamy *stanem granicznym* wektora w_0, zapisujemy jako $w_{+\infty}$ i możemy opisać równaniem:

$$\lim_{n\to\infty} Pw_0 = \lim_{n\to\infty} w_n = w_{+\infty}.$$

Stan stacjonarny

Okazuje się, że każdy stan graniczny musi być *stanem stacjonarnym (równowagi)*, czyli takim wektorem naszego układu, który nie zmienia się w kolejnych krokach czasowych łańcucha Markowa. Taki stan w musi spełniać równanie:

$$w = Pw,$$

które można przekształcić do wygodnej obliczeniowo (i mnemotechnicznie) postaci:

$$(P - I)w = 0,$$

gdzie I jest macierzą jednostkową odpowiedniego wymiaru, a 0 jest wektorem (ustawionym w jedną pionową kolumnę) złożonym z samych zer.

Twierdzenie Frobeniusa-Perrona (wniosek)

Twierdzenie znane z algebry dostosowane do potrzeb tego wykładu:

Twierdzenie Frobeniusa-Perrona (wniosek)

Niech P będzie macierzą przejścia pewnego łańcucha Markowa i P^n dla pewnego $n \geqslant 1$ ma wszystkie współczynniki dodatnie. Wtedy łańcuch Markowa posiada stan stacjonarny w o wszystkich współrzędnych dodatnich i wszystkie inne stany stacjonarne są wielokrotnościami wektora w.
Dodatkowo, każdy ze stanów stacjonarnych w o współrzędnych dodatnich jest stanem granicznym dla wszystkich stanów początkowych o tej samej sumie wszystkich współrzędnych co w.

Wyznaczanie stanu granicznego - przykład

Szukamy stanu granicznego dla naszego zadania z trzema stanami komórek, które realizowała poniższa macierz przejścia:

$$P = \begin{bmatrix} \frac{2}{3} & 0 & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{5} \\ 0 & \frac{1}{2} & \frac{11}{20} \end{bmatrix}.$$

P spełnia założenia wniosku z twierdzenia Frobeniusa-Perrona, bo

$$P^2 = \begin{bmatrix} \frac{4}{9} & \frac{1}{8} & \frac{73}{240} \\ \frac{7}{18} & \frac{7}{20} & \frac{22}{75} \\ \frac{1}{6} & \frac{21}{40} & \frac{161}{400} \end{bmatrix}.$$

Wyznaczanie stanu granicznego - przykład

Skoro tak, stany stacjonarne będą stanami granicznymi układu i można je obliczyć z równania $(P - I)w = 0$, które w tej sytuacji przyjmuje postać:

$$\begin{bmatrix} \frac{2}{3}-1 & 0 & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{2}-1 & \frac{1}{5} \\ 0 & \frac{1}{2} & \frac{11}{20}-1 \end{bmatrix} \begin{bmatrix} z \\ i \\ c \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} & 0 & \frac{1}{4} \\ \frac{1}{3} & -\frac{1}{2} & \frac{1}{5} \\ 0 & \frac{1}{2} & -\frac{9}{20} \end{bmatrix} \begin{bmatrix} z \\ i \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Wyznaczanie stanu granicznego - przykład

Można to przełożyć na język układów równań:

$$\begin{cases} -\frac{1}{3}z + \frac{1}{4}c = 0 \\ \frac{1}{3}z - \frac{1}{2}i + \frac{1}{5}c = 0 \\ \frac{1}{2}i - \frac{9}{20}c = 0 \end{cases} .$$

Rozwiązania tego układu, to wektory postaci $(\frac{3}{4}\alpha, \frac{9}{10}\alpha, \alpha)$, gdzie α jest dowolnie zadaną liczbą. Na przykład, jeśli komórek w badanej próbce jest w sumie 265, to możemy obliczyć z równania $\frac{3}{4}\alpha + \frac{9}{10}\alpha + \alpha = 265$ możemy obliczyć $\alpha = 100$ i wiemy, że stanem granicznym będzie $(\frac{3}{4}\alpha, \frac{9}{10}\alpha, \alpha) = (75, 90, 100)$. Dlatego możemy przewidzieć, że na dłuższą metę w próbce będzie około 75 komórek zdrowych, 90 zainfekowanych i 100 chorych: pojedyncze komórki mogą zmieniać stan, ale stan całego układu się już nie zmieni.

Łańcuchy Markowa i rankingi

Często nie jesteśmy w stanie ustalić, który ze stanów granicznych zostanie osiągnięty (np. ilość obiektów w układzie się zmienia lub po prostu jej nie znamy). Jednak wtedy można skorzystać z faktu, że możliwych stanów granicznych jest nieskończenie wiele, ale wszystkie są efektem przemnożenia jednego z nich przez liczbę. W szczególności, niezależnie jaki stan graniczny wybierzemy, uporządkowanie jego współrzędnych w kolejności od największej do najmniejszej wartości, oznaczające względną „popularność" każdego stanu w dalszej perspektywie czasowej. Dzięki temu, łańcuchy Markowa, grafy przejścia i macierze przejścia można wykorzystać do tworzenia rankingów: porządkowania pewnych opcji od najlepszej do najsłabszej, przy wiedzy, jak one wpływają na siebie nawzajem.

Łańcuchy Markowa i ich zastosowania

Przykładowe (z olbrzymiej liczby) zastosowania łańcuchów Markowa:

- Badania dynamik rozwoju populacji, w tym modelowanie epidemii.
- AHP (Analytic Hierarchy Process) - metoda hierarchicznej analizy procesów decyzyjnych.
- Modelowanie rynków np. giełdy.
- Pierwszy raz użycie łańcuchów Markowa do tworzenia rankingu było zaproponowane w XIX wieku (E.Landau) by rozstrzygnąć kolejność w turnieju szachowym, w którym nie udało się rozegrać wszystkich partii. Używane do tworzenia rankingów sportowych.
- **Ranking istotności stron internetowych na zadany temat.**

Algorytm PageRank

Rozważmy wyszukiwarkę internetową, która znalazła już strony zawierające zadane słowa kluczowe. W jakiej kolejności powinna je wyświetlić użytkownikowi? Oczywiście tak, by najużyteczniejsze wyniki wyszukiwania znalazły się jak najwyżej. Ale jak zdefiniować tę użyteczność?
W 1998 roku Sergey Brin i Larry Page, studenci uniwersytetu Stanford, opublikowali artykuł *The anatomy of a large-scale hypertextual Web search engine* w którym zaproponowali algorytm do szacowania istotności stron na takiej liście oparty na wyznaczniu stanów granicznych w łańcuchach Markowa. Algorytm ten nosi nazwę PageRank, a jego autorzy w tym samym czasie założyli znaną firmę Google, która ten algorytm długo wykorzystywała zgodnie z jego przeznaczeniem...

Idea PageRank

Główną ideą algorytmu było, by na podstawie sieci linków między stronami ustalić, z jakim prawdopodobieństwem się użytkownik, który przechodzi ze strony na stronę (głównie) po linkach znajdzie się na każdej z tych stron - i strony z większym prawdopodobieństwem znajdą się wyżej w naszym rankingu. Dzięki temu strony, na które wskazuje wiele linków są typowo ważniejsze od stron, na które wskazuje mało linków. Jednak liczy się też jakość linków: strona do której prowadzi jeden link z bardzo uczęszczanej strony witryny może być wyżej w rankingu niż strona z wieloma linkami z mało popularnych stron. W podstawowej wersji algorytmu, Brin i Page zakładali, że każdy link jest wybierany z tym samym prawdopodobieństwem.

Uproszczony PageRank

Przedstawię najpierw uproszczoną wersję algorytmu PageRank (bo łatwiej na niej pokazać podstawową zasadę algorytmu), a następnie pokażę, jak to uproszczenie było usunięte w oryginalnej pracy Brina i Page'a.

Uproszczenie polega na założeniu, że użytkownicy przemieszczają się między stronami **wyłącznie** za pomocą linków, chyba, że z jakiejś strony nie wychodzi żaden link. Prawdziwy algorytm uwzględnia (później zobaczymy, w jaki sposób), że użytkownicy mogą przemieszczać się między stronami też w inny sposób.

Dzięki temu, zagadnienie sprowadza się do wyznaczenia stanu granicznego łańcucha Markowa, którego stanami są strony, a przejścia między nimi symbolizują linki (z wagami/prawdopodobieństwami odwrotnie proporcjonalnymi do liczby linków na stronie wyjściowej).

Uproszczony PageRank

Uproszczony algorytm PageRank

Dane: Graf $G = (V(G), E(G))$ prosty skierowany, reprezentujący linki między stronami, $|V(G)| = n$.

Zmienne: $p(e)$ - funkcja wag zadana w - wektor rankingu.

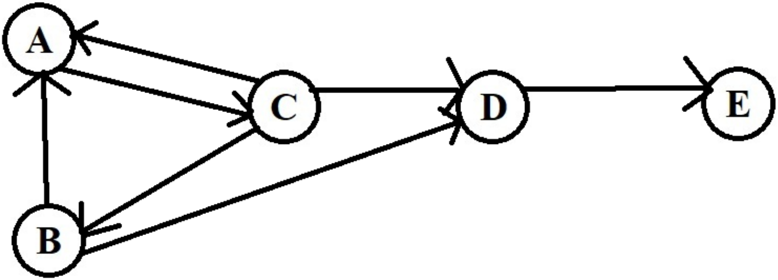
- I. Jeśli $(A, B) \in E(G)$ to $p(A, B) := \frac{1}{\deg^+ A}$.
- II. Jeśli dla jakiegoś $X \in V(G)$ zachodzi $\deg^+ X = 0$, to $E(G) := E(G) \cup \{(X, A) : A \in V(G)\}$ i $\forall_A \in V(G)\ p(X, A) = \frac{1}{n}$.
- III. Nowy graf $G(V(G), E(G))$ jest grafem przejścia pewnego łańcucha Markowa, którego dowolnym ze stanów granicznych jest wektor w.
- **Rezultat**: Lista wierzchołków uporządkowana od największej do najmniejszej wartości współrzędnej wektora w odpowiadającej temu wierzchołkowi.

Uproszczony PageRank - komentarze

- Pierwszy etap ilustruje założenie o „równości” wszystkich linków: z danej strony A z jednakowym prawdopodobieństem ($\frac{1}{\deg^+ A}$) można wyjść każdym linkiem.
- W drugim etapie radzimy sobie ze stronami, z których nie prowadzą żadne linki. Zakładamy, że z takiej strony do każdej innej wychodzimy z równym prawdopodobieństwem.
- Nie możemy być pewni, czy macierz przejścia spełnia założenia twierdzenia Frobeniusa-Perrona (o dodatniości jakiejś potęgi), więc wektor w może nie być obliczalny. Jednak zazwyczaj jest.
- Tak naprawdę, otrzymujemy nieskończenie wiele wektorów granicznych, ale wszystkie zadają tę samą kolejność stron.

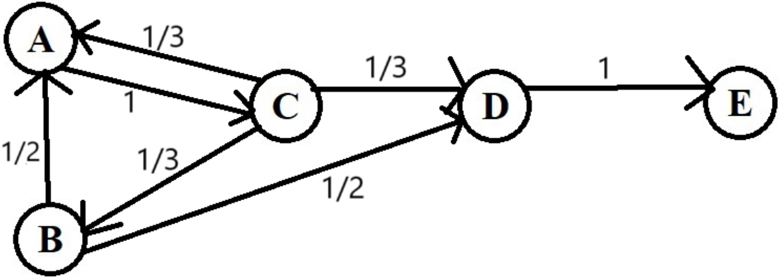
Uproszczony PageRank - przykład

Wykonajmy uproszczony algorytm PageRank dla 5 stron: A,B,C,D,E, połączonych linkami:
$A \to C, B \to A, B \to D, C \to A, C \to B, C \to D, D \to E$, co ilustruje poniższy graf:



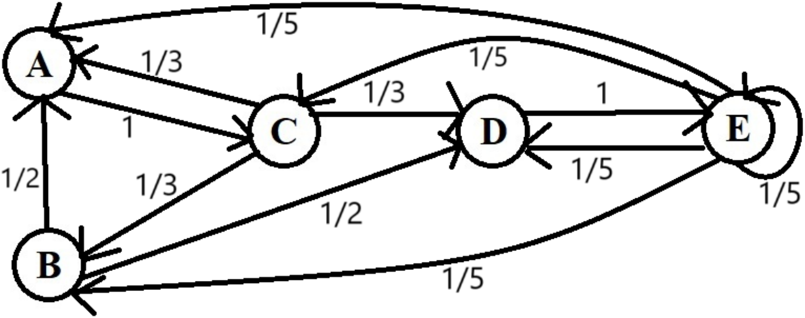
Uproszczony PageRank - przykład

Według pierwszego etapu algorytmu, każdej krawędzi w wyjściowym grafie przypisujemy wagę $\frac{1}{x}$, gdzie x to liczba krawędzi wychodzących z tego samego wierzchołka, co nasza krawędź (wszystkie linki, więc wszystkie te krawędzie, są równie prawdopodobne jako ścieżki wyjścia).



Uproszczony PageRank - przykład

W drugim etapie algorytmu radzimy sobie z wierzchołkiem E, z którego nie prowadzi żaden link: „dodajemy" do grafu krawędzie z E do każdego innego wierzchołka (również samego E) o tej samej wadze (w tym przypadku $\frac{1}{5}$).



To jest graf przejścia naszego łańcucha Markowa.

Uproszczony PageRank - przykład

Dla naszego grafu przejścia, tworzymy macierz przejścia

$$P = \begin{bmatrix} 0 & 1/2 & 1/3 & 0 & 1/5 \\ 0 & 0 & 1/3 & 0 & 1/5 \\ 1 & 0 & 0 & 0 & 1/5 \\ 0 & 1/2 & 1/3 & 0 & 1/5 \\ 0 & 0 & 0 & 1 & 1/5 \end{bmatrix}$$

i rozwiązujemy równanie $(P - I)w = 0$, by otrzymać stan graniczny.

Uproszczony PageRank - przykład

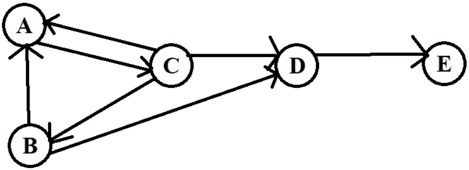
Zakładając, że $w = (w_A, w_B, w_C, w_D, w_E)$, równanie $(P - I)w = 0$ generuje następujący układ równań: przełożyć na język układów równań:

$$\begin{cases} -w_A + \frac{1}{2}w_B + \frac{1}{3}w_C + \frac{1}{5}w_E = 0 \\ -w_B + \frac{1}{3}w_C + \frac{1}{5}w_E = 0 \\ w_A - w_C + \frac{1}{5}w_E = 0 \\ \frac{1}{2}w_B + \frac{1}{3}w_C - w_D + \frac{1}{5}w_E = 0 \\ w_D - \frac{4}{5}w_E = 0. \end{cases}$$

Z ostatniego równania: $w_D = \frac{4}{5}w_E$. Odejmując stronami równanie pierwsze i czwarte otrzymamy $w_A = w_D$. Wstawiając te dwie informacje do równania trzeciego otrzymamy, że $w_E = w_C$. Po wstawieniu tej ostatniej równości do drugiego równania, otrzymamy $w_B = \frac{8}{15}w_E$. Dlatego wszystkie rozwiązania są postaci $w = (\frac{4}{5}w_E, \frac{8}{15}w_E, w_E, \frac{4}{5}w_E, w_E)$.

Uproszczony PageRank - przykład

Wszystkie rozwiązania są postaci $w = (\frac{4}{5}w_E, \frac{8}{15}w_E, w_E, \frac{4}{5}w_E, w_E)$. Wybieram przykładowe rozwiązanie, na przykład: $(12, 8, 15, 12, 15)$ - są to kolejno rankingi stron A,B,C,D,E (inny wybór dałby inne rankingi, ale tę samą kolejność). Stąd kolejność stron wyznaczona algorytmem PageRank to:
C=E, A=D, B (= oznacza, że kolejność tych dwóch stron na liście jest dowolna).



Wcale nie strony, do których prowadzi najwięcej linków, są najważniejsze!

Prawdziwy PageRank

Oryginalny program PageRank, przedstawiony w pracy Brina i Page'a i będący podstawą pierwszej wyszukiwarki Google różnił się od przedstawionego powyżej założeniem, że użytkownik przemieszcza się między stronami za pomocą linków tylko z pewnym prawdopodobieństwem d (znanym jako współczynnik tłumienia), a z prawdopodobieństwem $1 - d$ przeskakuje na losową stronę w sieci. Dlatego prawdziwa macierz przejścia w algorytmie PageRank (znana również jako macierz Google) ma postać:

$$P_R = dP + (1-d)\frac{1}{n}J,$$

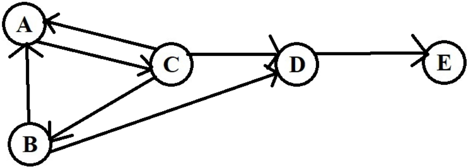
gdzie J jest macierzą złożoną z samych jedynek. W oryginalnym algorytmie PageRank $d = \frac{17}{20}$.

Korzyści z modyfikacji

$$P_R = dP + (1-d)\frac{1}{n}J.$$

- Łańcuch Markowa, którego macierzą przejścia jest P_R modeluje zachowania użytkowników realistyczniej niż łańcuch Markowa z macierzą przejścia P.
- Macierz P_R ma wszystkie wyrazy dodatnie, więc automatycznie spełnia założenia twierdzenia Frobeniusa-Perrona, dzięki czemu stan graniczny układu zawsze może być wyznaczony jednoznacznie z dokładnością do przeskalowania.
- Oczywiście, rozwiązanie powstającego, najczęściej gigantycznego układu równań wymaga dużej mocy obliczeniowej i efektywnych algorytmów rozwiązywania równań liniowych (więcej na ten temat na przedmiocie Metody Numeryczne).

Macierz PageRank



Dla naszego przykładu:

$$P_R = \frac{17}{20}\begin{bmatrix} 0 & 1/2 & 1/3 & 0 & 1/5 \\ 0 & 0 & 1/3 & 0 & 1/5 \\ 1 & 0 & 0 & 0 & 1/5 \\ 0 & 1/2 & 1/3 & 0 & 1/5 \\ 0 & 0 & 0 & 1 & 1/5 \end{bmatrix} + \frac{3}{20}\cdot\frac{1}{5}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 3/100 & 91/200 & 47/150 & 3/100 & 1/5 \\ 3/100 & 3/100 & 47/150 & 3/100 & 1/5 \\ 22/25 & 3/100 & 3/100 & 3/100 & 1/5 \\ 3/100 & 91/200 & 47/150 & 3/100 & 1/5 \\ 3/100 & 3/100 & 3/100 & 22/25 & 1/5 \end{bmatrix}$$

Zmodyfikowane rozwiązanie

$$P_R = \begin{bmatrix} 3/100 & 91/200 & 47/150 & 3/100 & 1/5 \\ 3/100 & 3/100 & 47/150 & 3/100 & 1/5 \\ 22/25 & 3/100 & 3/100 & 3/100 & 1/5 \\ 3/100 & 91/200 & 47/150 & 3/100 & 1/5 \\ 3/100 & 3/100 & 3/100 & 22/25 & 1/5 \end{bmatrix}$$

W tym przypadku mogą Państwo sprawdzić, że modyfikacja o współczynnik tłumienia zupełnie nie wpływa na uporządkowanie stron w porównaniu z uproszczonym algorytmem PageRank - nadal otrzymujemy ranking (C=E, A=D, B). W ogólnym przypadku,przejście z uproszczonego PageRanka na oryginalny PageRank może to uporządkowanie zmienić).

Historia i problemy z Page Rank

- Ranking oparty o PageRank jest dość prosty do manipulowania, zwłaszcza w sytuacji, gdy opis algorytmu jest powszechnie znany. wystarczy do sieci dokładać odpowiednio dużo linków do strony, którą chcemy promować i jej ranking wzrośnie.
- Dlatego niemal natychmiast w zastosowaniach rankingi wyszukiwarek opartych na PageRank zaczęły być modyfikowane. PageRank stał się tylko jednym z kryteriów oceny strony, a ponadto sam był modyfikowany, czy to w celu urealnienia oceny wartości linków (które nie są równie istotne), czy też w celu ochrony przed wcześniej wspomnianymi manipulacjami. Dokładny opis rankingu pozostaje tajny.
- Według informacji podawanych przez przedstawicieli Google, około 2006 roku jej wyszukiwarka kompletnie odeszła od wykorzystywania algorytmu PageRank (ale niekoniecznie od podejścia opartego o łańcuchy Markowa). Jednak wyciek informacji z tej firmy, do którego doszło w roku 2024 wskazuje, że rankingi PageRank są nadal tam używane, choć dokładne zastosowania są nieznane.

Page Rank - współczesne zastosowania

Nawet zakładając, że PageRank nie jest już bezpośrednio wykorzystywany do tworzenia rankingów stron w wyszukiwarkach, nadal jest stosowany w co najmniej kilku zagadnieniach:

- Ocena wartości prac i czasopism naukowych na podstawie cytowań.
- Analiza sieci reakcji białkowych w biochemii.
- Porównywanie istotności różnych gatunków w ekosystemach.
- W neuronauce, przewidywanie aktywności konkretnych neuronów w zależności od ich lokalizacji w sieci neuronowej.