Google, ვებგვერდების რანჟირების ალგორითმი და მათემატიკა

1. საძიებო სისტემები

ციფრულ სამყაროში ახალი პრობლემები, როგორც წესი, მარტივად გადაიჭრება ახალი ალგორითმების ან ახალი მოწყობილობების საშუალებით. ვისაც ინტერნეტში მუშაობის თხუთმეტწლიანი გამოცდილება მაინც აქვს, ალბათ კარგად ახსოვს ძველი საძიებო სისტემები, მაგალითად, Alta Vista და Yahoo. დაბეჯითებით შეიძლება ითქვას, რომ თითქმის ყველა, ვინც იმ დროს ინტენსიურად იყენებდა ამ საძიებო სისტემებს, ამჟამად ინტერნეტში ინფორმაციის ძიებისას Google-ის საძიებო სისტემას იყენებს. გასაოცარია, მაგრამ Google-მა პოპულარობა სულ რამდენიმე თვეში მოიპოვა. ამაში მას დაეხმარა ინოვაცია, რომელსაც ძიების შედეგების რანჟირების ალგორითმი (PageRank Algorithm) ეწოდება. სტატიის მიზანია, მკითხველს, რომელიც ფლობს მცირე მათემა¬ტიკურ აპარატს, ამ ალგორითმის შესახებ წარმოდგენა შევუქმნათ.

ალგორითმი ეფუძნება თეორიას, რომელიც მათემატიკაში Google-ის მიერ მის გამოყენებამდე დიდი ხნით ადრე იყო ცნობილი. ამ თეორიას მარკოვის ჯაჭვების თეორია ეწოდება.

საძიებო სისტემის გამოყენება მარტივია. ინტერნეტში ჩართულ კომპიუტერთან მჯდარი მომხმარებელ საძიებო სისტემის ფორმის შესაბამის ველში წერს ფრაზას და ფორმის ღილაკზე დაჭერით ამ ფრაზას უგზავნის საძიებო სისტემის სერვერს. საძიებო სისტემა პასუხად აბრუნებს იმ მისამართების სიას, რომლებზე განთავსებული ინფორმაციაც შესაძლოა დაკავშირებული იყოს მომხმარებლის მიერ მიწოდებულ ფრაზასთან.

Google-ის გამოყენებას ის უპირატესობა აქვს, რომ მაღალია ალბათობა, მომხმარებლისთვის საინტერესო მისამართები ძიების შედეგების დასაწყისში მოხვდეს. ამიტომ მომხმარებელს არ უწევს ძიების შედეგების უზარმაზარი სიის გადარჩევა. საინტერესო ის არის, როგორ ახერხებს Google-ის საძიებო სისტემა მომხმარებლის სურვილის გამოცნობას მაღალი ალბათობით.

ავტომატიზებული საძიებო საშუალებები მრავალი ათწლეულის განმავლობაში ბევრ სხვადასხვა ადგილას გამოიყენებოდა. მაგალითად, საბიბლიოთეკო მონაცემთა ბაზები, საჯარო რეესტრის მონაცემთა ბაზები, სხვადასხვა პროფესიო¬ნალურ მონაცემთა ბაზები, რომლებიც სამედიცინო, იურიდიულ და სხვა სახის ინფორმაციას მოიცავს. ამ სახის მონაცემთა საცავებში საძიებო სისტემის მუშაობა საკმაოდ მარტივია იმის გამო, რომ ყველა ამ სახის ინფორმაციის საცავებს აქვს ერთი საერთო თვისება: თითოეული მათგანი შეიცავს წინასწარ განსაზღვრული სტრუქტურის მქონე ინფორმაციის ერთეულებს. მაგალითად, საბიბლიოთეკო მონაცემთა ბაზის თითოეული ჩანაწერი წიგნთან არის დაკავშირებული. აქედან გამომდინარე, ეს ჩანაწერი უნდა შეიცავდეს ინფორმაციას ავტორის, სათაურის, გამოცემის თარიღის და სხვა ისეთი მონაცემების შესახებ, რომლებიც დაკავშირებულია წიგნთან. იგივე შეიძლება ითქვას სადაზღვევო კომპანიის მონაცემთა ბაზის შესახებ, რომელიც შეიცავს ინფორმაციას ამ კომპანიის მომხმარებლების შესახებ. გარდა იმისა, რომ ეს ინფორმაცია სტრუქტურირებულია, მას აქვს კიდევ ერთი მნიშვნელოვანი თვისება - ერთგვაროვნება. მაგალითად, საბიბლიოთეკო ჩანაწერებში არასოდეს მოხვდება ჩანაწერი რომელიმე საკვები პროდუქტის შესახებ, რადგან ეს მონაცემთა ბაზა მხოლოდ წიგნებზეა სპეციალიზებული. სხვა სიტყვებით რომ ვთქვათ, ახალი მონაცემების შეტანისას ყოველთვის ითვალისწინებენ მონაცემთა სტრუქტურას და სპეციფიკურობას. ეს ყოველივე მთლიანად უგულებელყოფილია ინტერნეტში. აქ ინფორმაცია განთავ¬სებულია ყოველგვარი წინასწარი სტრუქტურის გაუთვალისწინებლად.

ვებგვერდები უსაზღვროდ მრავალფეროვანია და მოიცავს სხვადასხვა სფეროს, დაწყებული ტექნიკურითა და პროფესიონალურით და დამთავრებული გასართობი შინაარსის გვერდებით. ასევე მრავალფეროვანია მომხმარებელთა ინტერესები და ინფორმაციის მოძიების კომპეტენციები. ყოველივე ამასთან ერთად, ინტერნეტში განთავსებული ინფორმაცია სწრაფად იცვლება: ყოველ¬წამიერად ჩნდება და ქრება ინტერნეტრესურსები.

მიუხედავად ამისა, არსებობს რამდენიმე მნიშვნელოვანი რამ, რაც ინტერნეტში განთავსებულ ვებგვერდებს აერთიანებს: ისინი დაწერილია ერთსა და იმავე ენაზე ან მასთან დაკავშირებულ „დიალექტზე“ და ერთმანეთთან ჰიპერბმულების საშუალებით არის დაკავშირებული. ჰიპერბმულების ერთ-ერთი მნიშვნელოვანი ნაწილია მისამართი, რომლითაც იგი სხვა რესურსზე მიუთითებს. ეს მისამართები მომხმარებლისთვის უმეტესწილად უხილავია. მომხმარებელი ერთი რესურსიდან მეორეზე მხოლოდ ჰიპერბმულის შესაბამის ვიზუალურ კომპონენტზე (მაგ., ტექსტი, ღილაკი, ნახატი) დაჭერით გადადის.

1998 წლის იანვარში სტანფორდის უნივერსიტეტის ოთხმა მკვლევარმა (L. Page, S. Brin, R. Motwani, T. Winograd) შექმნა ალგორითმი, რომლის საშუალებითაც შესაძ¬ლებელი გახდა ინტერნეტგვერდების რანჟირება. ეს ალგორითმი იყენებს არა ვებგვერდის შინაარსს, მასზე განთავსებულ ტექსტურ, თუ ვიზუალურ მასალას, არამედ ინტერნეტში არსებულ ვებგვერდებს შორის ბმულების სტრუქტურას.

შევნიშნავთ, რომ რანჟირების ალგორითმის გამოქვეყნების დროს (1998 წელი), მისი ავტორების შეფასებით, ინტერნეტი შედგებოდა დაახლოებით 150 მილიონი გვერდისგან, ხოლო მათ შორის ბმულების რაოდენობა იყო დაახლოებით 1.7 მილიარდი. რამდენიმე წლის განმავლობაში გვერდების რაოდენობამ დაახლოებით 12 მილიარდს მიაღწია, ხოლო სადღეისოდ მათი რაოდენობა კიდევ რამდენიმეჯერ გაიზარდა.

ინტერნეტი და მარკოვის ჯაჭვები

ინტერნეტი მილიარდობით გვერდისგან შედგება, ხოლო მათ შორის ბმულების რაოდენობა კიდევ უფრო მეტია. აქედან გამომდინარე, ინტერნეტის მოდელირებისთვის შეიძლება გამოვიყენოთ მიმართული გრაფი. ამ გრაფში წვეროები ვებგვერდებს შეესაბამება, ხოლო ჰიპერბმულები, რომლებიც გვერდებს ერთმანეთთან აკავშირებს, შეგვიძლია განვიხილოთ როგორც გრაფის წიბოები. ამ წერტილების შემაერთებელი ისრიანი წირები ნიშნავს, რომ:

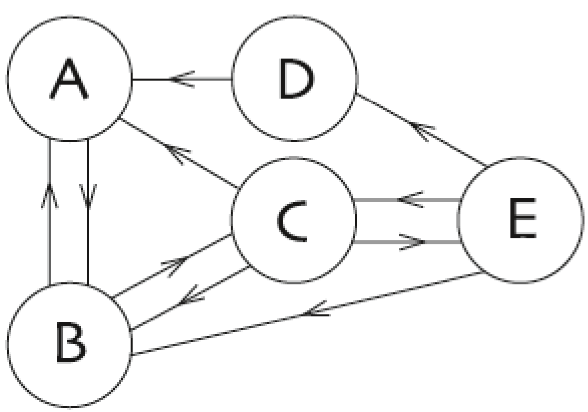
გვერდზე არსებობს ერთადერთი ჰიპერბმული გვერდისადმი;

გვერდებისადმი;

გვერდებისადმი;

გვერდისადმი;

გვერდებისადმი.



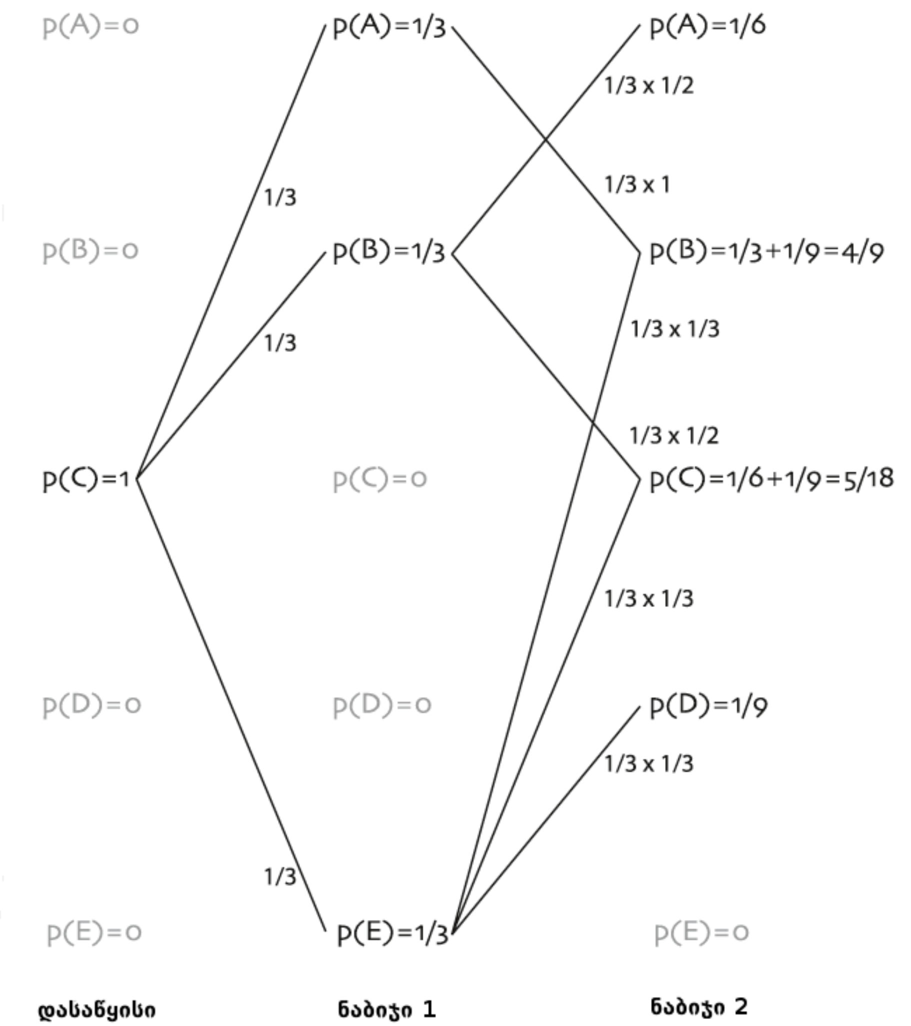
იმისათვის რომ მოვახდინოთ რანჟირების ალგორითმის დემონსტრირება ინტერ-ნენტის ამ წარმოსახვით, „მინი მოდელზე“, დავუშვათ რომ მომხმარებელი იწყებს ნავიგაციას შემთხვევით შერჩეული ჰიპერბმულების გასწვრივ. იმ შემთხვევაში, როდესაც მას აქვს მხოლოდ ერთი არჩევანი (მაგალითად, თუ იგი იწყებს D გვერდიდან), მაშინ მას შეუძლია გადაადგილდეს მხოლოდ A გვერდისაკენ. თუ მომხმარებელი იწყებს C გვერდიდან, მაშინ მას შეუძლია მიყვეს სამი ჰიპერბმულიდან ერთ-ერთს (ე.ი. რომელიმე მათგანზე გადაადგილების ალბათობა არის $\frac{1}{3}$).

სხვა სიტყვებით რომ ვთქვათ, იმ გვერდიდან, რომელზეც მოცემულ მომენტში იმყოფება მომხმარებელი, იგი შემთხვევით ირჩევს ამ გვერდზე არსებული ჰიპერბმულებიდან ერთ-ერთს, თანაბარი ალბათობით. დავუშვათ რომ ეს მომხმარებელი გადაადგილდება ერთი ჰიპერბმულის გასწვრივ ერთი წუთის განმავლობაში. ისმება შეკითხვა: რომელ გვერდზე აღმოჩნდება ეს მომხმარებელი ერთი საათის შემდეგ? ერთი დღის შემდეგ? ჰიპერბმულების გასწვრივ გადაად-გილებების ძალიან დიდი რაოდენობის შემდეგ?

უფრო ზუსტად, იმის დაშვებით რომ მისი გადაადგილების გზა ალბათურია, რისი ტოლი იქნება იმის ალბათობა იმისა, რომ დროის მოცემული მონაკვეთის შემდეგ იგი აღმოჩნდება მოცემულ გვერდზე?

აქ მოცემული ნახაზი დაგვეხმარება იმაში, რომ ამ კითხვას პასუხი გავცეთ იმ შედარებით მარტივ შემთხვევაში, როდესაც მომხმარებელი იწყებს ინტერნეტში გადაადგილებას C გვერდიდან და გადაადგილდება ორი ბმულით. როგორც ნახაზიდან ჩანს, ამ გვერდზე განთავსებულია სამი გარეთ გამავალი ბმული, რომლებიც დაკავშირებულია A, B და E გვერდებთან. ამგვარად, პირველი ნაბიჯის შემდეგ მომხმარებელი შეიძლება აღმოჩნდეს A გვერდზე ალბათობით $\frac{1}{3}$, B გვერ-დზე ალბათობით $\frac{1}{3}$ და E გვერდზე ასევე ალბათობით $\frac{1}{3}$. ეს მითითებულია ნახაზზე ასეთი ჩანაწერით: $p(A)=\frac{1}{3},\ p(B)=\frac{1}{3},\ p(E)=\frac{1}{3}$. ანალოგიური შინაარსისაა ჩანაწერები: $p(C)=0,\ p(D)=0$. ეს უკანასკნელი მიუთითებს იმას, რომ ერთი ნაბიჯის შემდეგ, მომხმარებელი ვერ მოხვდება C ან D გვერდებზე, რადგან წინა გვერდიდან არ არსებობს ჰიპერბმულები ამ გვერდებზე. მივაქციოთ ყურადღება იმას, რომ ყოველი ნაბიჯის შემდეგ, შესაბამისი ალბათობების ჯამი 1-ის ტოლია:

$$p(A+p(B)+p(C)+p(D)+p(E)=1.$$



მეორე ნაბიჯის შემდეგ, ალბათობა $p(B)$ (ორი გადაადგილების შემდეგ B გვერდზე მოხვედრის ალბათობა) უკვე აღარ არის $\frac{1}{3}$-ის ტოლი, რადგან გარდა $C \to B$ გზისა, არსებობს მეორე გზა, რომელსაც B წერტილში მივყავართ: $C \to E \to B$. თუ პირველი საფეხურის შემდეგ მომხმარებელი აღმოჩნდება E გვერდზე, მაშინ მეორე საფეხურზე მას შეუძლია თანაბარი ალბათობით აირჩიოს სამი ბმულიდან ერთ-ერთი: $E \to B$, $E \to C$, $E \to D$. წარმოდგენილ ნახაზზე ნათლად ჩანს თუ რისი ტოლი იქნება ალბათობები $p(A)$, $p(B)$, $p(C)$ და $p(D)$, ხდომილებების ნამრავლისა და ჯამის ალბათობების გამოთვლის წესის მიხედვით, მეორე საფეხურის შემდეგ:

$$p(A) = \frac{1}{6}, \quad p(B) = \frac{4}{9}, \quad p(C) = \frac{5}{18}, \quad p(D) = \frac{1}{9}, \quad p(E) = 0.$$

მივაქციოთ ყურადღება, რომ ამ ალბათობების ჯამი კვლავ 1-ის ტოლია, რაც შესაბამისობაშია ალბათობის თვისებასთან (ხდომილებების სრული სისტემის ალბათობების ჯამი 1-ის ტოლია).

ამის შემდეგ, უკვე შეგვიძლია ჩამოვაყალიბოთ მეთოდის ზოგადი აღწერა: გავაგრძელოთ სხვადასხვა გვერდებზე მოხვედრის ალბათობების გამოთვლა ნაბიჯების რაოდენობის გაზრდასთან ერთად და შევადაროთ ერთმანეთს ეს ალბათობები. კერძოდ, პირველი ორი ნაბიჯის შემდეგ, ყველაზე მაღალი ალბათობა აქვს იმას, რომ მოვხვდებით B გვერდზე, თუმცა ნაბიჯების რაოდენობის გაზრდის შემდეგ შესაძლოა სურათი შეიცვალოს და სხვა რომელიმე გვერდზე მოხვედრის ალბათობა გახდეს უფრო მეტი. საუკეთესო მათემატიკური აპარატი, რომლის გამოყენებითაც შესაძლებელია ამ პროცესის შესწავლა, არის მარკოვის ჯაჭვების თეორია, რომლის შესწავლის ძირითადი ობიექტი არის ე.წ. გადასვლის (ან ჰიპერბმულების) მატრიცა (ცხრილი):

$$P = \begin{bmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \vdots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{bmatrix}.$$

ამ მატრიცის ელემენტები განისაზღვრება ასე:

$$\boldsymbol{p_{ij}} = \begin{cases} \dfrac{\mathbf{1}}{\boldsymbol{k}}, & \textit{თუ } \boldsymbol{i} \textit{ გვერდიდან არსებობს გამომავალი ჰიპერბმული } \boldsymbol{j} \textit{ გვერდისაკენ} \\ \mathbf{0}, & \textit{თუ } \boldsymbol{i} \textit{ გვერდიდან არ არსებობს ჰიპერბმული } \boldsymbol{j} \textit{ გვერდისაკენ} \end{cases}$$

სადაც k რიცხვი არის i გვერდიდან გამომავალი ჰიპერბმულების მთლიანი რაოდენობა. მაგალითად, აქ განხილული მინი ინტერნეტის შემთხვევაში (5 გვერდი), ჰიპერბმულების მატრიცას აქვს ასეთი სახე

$$\begin{array}{cc} & \begin{array}{ccccc} \;A\; & \;B\; & \;C\; & \;D\; & \;E\; \end{array} \\ P = & \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{3} & 1 & 0 \\ 1 & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} & 0 & 0 \end{pmatrix} \end{array} \quad \begin{array}{c} A \\ B \\ C \\ D \\ E \end{array}$$

ამ მატრიცის სვეტებში წერია შესაძლო მდგომარეობებში გადასვლის ალბათობები. მაგალითად, A გვერდიდან შესაძლებელია მხოლოდ B გვერდზე მოხვედრა, ამიტომ იმ სვეტში, რომელსაც თავზე აწერია A, მხოლოდ ერთი რიცხვი განსხვავდება 0-საგან

და იგი 1-ის ტოლია. B გვერდიდან შესაძლებელია მხოლოდ A და C გვერდებზე მოხვედრა (ერთნაირი ალბათობით), ამიტომ იმ სვეტში, რომელსაც აწერია B, ამ ორი გვერდის შესაბამის ადგილებზე წერია $\frac{1}{2}$. იგივე შეიძლება ითქვას სხვა სვეტებისთვისაც. ინტერნეტში ნავიგაციის დროს, რომელიმე ნაბიჯის შესაბამისი ალბათობის ვექტორი არის ვექტორი, რომლის კომპონენტები მიუთითებს იმ ალბათობას, რომლითაც მომხმარებელი ამ ნაბიჯის შემდეგ მოხვდება ამ ვებ-გვერდზე. მაგალითად, თუ მომხმარებელი ნავიგაციას იწყებს C გვერდიდან, მაშინ დასაწყისში ალბათობის ვექტორი იქნება

$$p^0 = \begin{pmatrix} p(X_0 = A) \\ p(X_0 = B) \\ p(X_0 = C) \\ p(X_0 = D) \\ p(X_0 = E) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

ალბათობის თვისებების გამოყენებით ადვილი სანახავია, რომ პირველი ნაბიჯის შემ-დეგ ალბათობის ვექტორი მიიღება P მატრიცის გამრავლებით საწყის ვექტორზე: $p^1 = P \cdot p^0$, და ასე შემდეგ:

$$p^1 = \begin{pmatrix} p(X_1 = A) \\ p(X_1 = B) \\ p(X_1 = C) \\ p(X_1 = D) \\ p(X_1 = E) \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{3} & 1 & 0 \\ 1 & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ 0 \\ 0 \\ \frac{1}{3} \end{pmatrix},$$

$p^2 = P \cdot p_1$:

$$p^2 = \begin{pmatrix} p(X_2 = A) \\ p(X_2 = B) \\ p(X_2 = C) \\ p(X_2 = D) \\ p(X_2 = E) \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{3} & 1 & 0 \\ 1 & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ 0 \\ 0 \\ \frac{1}{3} \end{pmatrix} = \begin{pmatrix} \frac{1}{6} \\ \frac{4}{9} \\ \frac{5}{18} \\ \frac{1}{9} \\ 0 \end{pmatrix}.$$

საზოგადოდ: $p^n = P \cdot p^{n-1}$.

ამ პროცედურის შინაარსიდან ადვილად ჩანს, რომ ვებ-გვერდების რანჟირება ნიშნავს მათ დალაგებას იმის მიხედვით, თუ რომელ მათგანში მოხვედრის ალბათობა არის უფრო მეტი ინეტრნეტში ნავიგაციის პროცესის უსასრულოდ გაგრძელების შედეგად. ე.ი. თუ ვექტორების მიმდევრობისათვის $p^n = P \cdot p^{n-1},\ n = 1,\ 2, \dots$, ვიპოვით „ზღვრულ" ვექტორს და შემდეგ გვერდებს დავალაგებთ ამ ვექტორის შესაბამისი კომპონენტების ზრდის მიხედვით, ეს მოგვცემს გვერდების დალაგებას მათი რანგების ზრდის მიხედვით. როოგორც მარკოვის ჯაჭვების თეორია ამბობს, ეს „ზღვრული" ვექტორი კი, უმეტეს შემთხვევებში, არის ისეთი ვექტორი, რომელიც არის გადასვლის მატრიცის საკუთარი ვექტორი, რომლის შესაბამისი საკუთარი მნიშვნელობა 1-ის ტოლია („უძრავი წერტილი") და რომლის კომპონენტების ჯამი 1-ის ტოლია (ალბათობების ჯამი).

ჩვენი „მინი ინტერნეტის" შემთხვევაში ეს ვექტორი ადვილი მოსაძებნია წრფივი ალგებრის ელემენტარული აპარატის გამოყენებით და იგი არის:

$$\pi = \frac{1}{41}\begin{pmatrix} 12 \\ 16 \\ 9 \\ 1 \\ 3 \end{pmatrix}.$$

მიღებული შედეგის მიხედვით შეიძლება ითქვას, რომ საკმაოდ ხანგრძლივი ნავიგაციის შემდეგ მომხმარებელი ყველაზე ხშირად მოხვდება B გვერდზე (ალბათობა ტოლია $\frac{16}{41}$-ის). ხოლო გვერდების რანჟირება კი ასეთია: B გვერდის რანგი იქნება 1-ის ტოლი, რადგან მასში მოხვერდის ალბათობა ყველაზე მაღალია; A გვერდის რანგი იქნება 2-ის ტოლი, რადგან მასში მოხვედრის ალბათობა B-ს მერე ყველაზე მაღალია; მას მოსდევს C გვერდი, ხოლო ბოლო ადგილზეა D გვერდი, რადგან მასში მოხვედრის ალბათობა ყველაზე მცირეა.

შევნიშნავთ, რომ მართალია აქ აღწერილი პროცედურა წარმოადგენს რეალობაში გამოყენებული პროცედურის საფუძველს, Google იყენებს ამ პროცედურის გაუმჯობესების სხვადასხვა ხერხებს და ინოვაციებს, რომლებიც კომერციულ საიდუმლოებას წარმოადგენს.

გარდა ამისა, ჩვენ განვიხილეთ წარმოსახვითი შემთხვევა, როდესაც „მინი ინტერნეტი" შედგებოდა მხოლოდ 5 გვერდისაგან. ამ შემთხვევაში, ალბათობების ვექტორის მოძებნა საკმაოდ მარტივია და შესაძლებელია მფოლოდ ფურცლისა და კალმის

გამოყენებითაც. რეალური ინტერნეტი კი მოიცავს მილიარდობით გვერდს. ე.ი. რანჟირებისათვის საჭიროა ისეთი მატრიცის საკუთარი ვექტორის მოძებნა, რომელიც შეიცავს რამოდენიმე ასეულ მილიარდ ელემენტს. როგორც მკვლევარები აღნიშნავენ, ეს არის მატრიცებთან დაკავშირებული გამოთვლების ამოცანებს შორის ერთ-ერთი ყველაზე რთული ამოცანა. თუ რომელ ხერხს იყენებს Google ამ ამოცანის ამოსახსნელად, ასევე კომერციული საიდუმლოა.

ანალოგიური მიდგომა გამოიყენება სამეცნიერო ჟურნალების რანჟირებისას ციტირების ინდექსის (impact factor) მიხედვით. ამ შემთხვევაში, ვებ-გვერდის მაგივრად შეიძლება განვიხილოთ სამეცნიერო ჟურნალი, ხოლო ჰიპერბმულების მაგივრად შეგვიძლია განვიხილოთ ამ ჟურნალში არსებული ციტირებები, ხოლო გადასვლის მატრიცა აიგება იგივე წესით, როგორც ეს ხდება ინტერნეტ-გვერდების შემთხვევაში. თუმცა უნდა აღინიშნოს, რომ ამ რანჟირებაში მონაწილე სამეცნიერო ჟურნალების რაოდენობა გაცილებით მცირეა ვიდრე ვებ-გვერდების რაოდენობა და მათი რანჟირების ალგორითმი ნაკლებ რესურსებს მოითხოვს.