Primjene permutacija i kombinacija

Špikić, Matija

SVEUČILIŠTE U ZAGREBU

FAKULTET ORGANIZACIJE I INFORMATIKE

V A R A Ž D I N

Matija Špikić

Matični broj: 43233/17–R

Studij: Primjena informacijske tehnologije u poslovanju

PRIMJENA PERMUTACIJA I KOMBINACIJA

ZAVRŠNI RAD

Mentor:

Doc.dr.sc. Bojan Žugec

Varaždin, srpanj 2019.

Matija Špikić

Izjava o izvornosti

Izjavljujem da je moj završni/diplomski rad izvorni rezultat mojeg rada te da se u izradi istoga nisam koristio drugim izvorima osim onima koji su u njemu navedeni. Za izradu rada su korištene etički prikladne i prihvatljive metode i tehnike rada.

Autor/ica potvrdio/la prihvaćanjem odredbi u sustavu FOI-radovi

---

Sažetak

Cilj ovog rada je prezentirati neke realne primjene permutacija i kombinacija. Permutacije i kombinacije će se matematički definirati i na jednostavnijim primjerima pojasniti njihova definicija. Kroz pojedine primjere će se vidjeti kako je široka primjena permutacija i kombinacija u komunikacijskim mrežama i arhitekturi računala s naglaskom na interkonekcijskim mrežama. Uz pojašnjenu primjenu, implementirana su i dva algoritma koja će nam dati konkretnije rezultate i za veće ulazne podatke. U radu će se vidjeti u kako širokom spektru područja primjenu nalaze permutacije i kombinacije što se konstatira, između ostalog, i u zaključku rada.

Ključne riječi: permutacije, kombinacije, vjerojatnost, komunikacijske mreže, arhitektura računala, rudarenje podataka

Sadržaj

1. Uvod

U ovom završnom radu, čiji je naslov Primjena permutacija i kombinacija prvo ću objasniti pojmove permutacije i kombinacije, a nakon toga i njihovu primjenu na nekim primjerima iz svakodnevnog života. Permutacije i kombinacije su dio matematičke discipline koja se zove kombinatorika. Svi smo se mi često u životu susreli sa kombinatorikom, odnosno s permutacijama i kombinacijama, a da toga nismo bili svjesni. Neki takvi tipični, svakodnevni problemi su npr. koliko kombinacija brojeva postoji na lotu ili na koliko načina možemo razmjestiti petero ljudi za okrugli stol. Naravno, to su neki jednostavniji problemi, a postoje i mnogo kompliciraniji i važniji problemi kao što su šifriranje i kriptografija, optimizacija raznih poslova u raznim procesima, te još izvlačenje bitnih informacija iz mnoštva njih u puno različitih područja. Odgovore na takva i slična pitanja daje nam kombinatorika. Baš iz tog razloga što kombinatorika ima jako puno primjena u životu koje mi niti ne primjećujemo, odlučio sam pomnije istražiti koje su to praktične primjene permutacija i kombinacija. Također ću i neke od algoritama za rješavanje zadataka iz kombinatorike implementirati na računalu kroz programiranje tih algoritama. Opisana i primjena permutacija i kombinacija u vjerojatnosti, te su prikazani neki primjeri gdje se to koristi.

2. Permutacije

Permutacije imaju jako bogatu kombinatornu strukturu. Dio razloga toj bogatoj strukturi jest to što permutacija konačnog seta može biti zastupljena na puno ekvivalentnih načina. Neki od njih su u obliku riječi ili sekvence, kao funkcija, kao kolekcija ili skup disjunktnih ciklusa, kao matrica itd. Svaka od tih reprezentacija sugerira mnoštvo prirodnih invarijanti ili "statistika", operacija, transformacija, struktura itd. [1]

2.1. Definicija permutacije

Neka je $S = \{a_1, a_2, \ldots, a_n\}$ neki n-člani skup. Permutacija od S je bijekcija $f : S \to S$, pa odabirom nekog uređaja elemenata od S, npr. $a_1 < a_2 < ... < a_n$, permutacijom f ti elementi prelaze redom u $f(a_1), f(a_2), ..., f(a_n)$, dakle u jedan novi niz (točnije uređenu n-torku) različitih elemenata iz S, npr. $a_{i_1}, a_{i_2}, ..., a_{i_n}$. Time se dobiva jedan novi (linearni) uređaj na $S : a_{i_k} < a_{i_l} \Leftrightarrow k < l$. Tako dobivamo da na S ima $n!$ različitih linearnih uređaja. Skup svih bijekcija (tj. permutacija) $S \to S$ s kompozicijom kao grupovnom operacijom očito čini grupu. Ako je $S = [n] = 1, 2, ..., n$, onda se ta grupa označava s S_n i zove simetrična grupa. Red te grupe je $|S_n| = n!$ [9].

Drugim riječima, permutacija je niz elemenata koja sadrži točno jednom svaki element iz nekog konačnog skupa (ne smije se taj element ponavljati) i ti elementi su poredani na određen način, a broj permutacija skupa S s n članova je $n!$.

2.2. Vrste permutacija

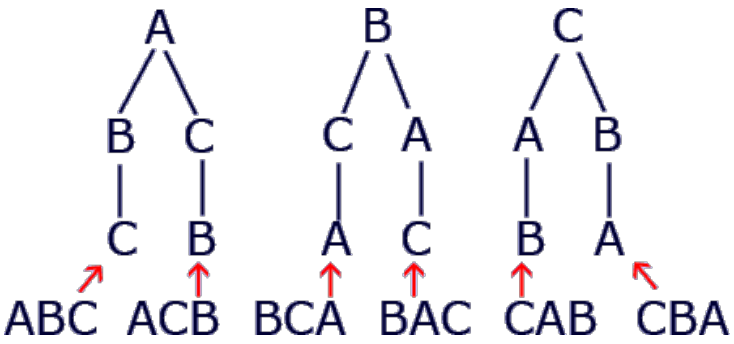
Ako su svi elementi međusobno različiti onda se takve permutacije nazivaju permutacije bez ponavljanja, a ako su neki elementi u permutaciji isti tada se takve permutacije nazivaju permutacije s ponavljanjem. [6]

2.2.1. Permutacije bez ponavljanja

Kao primjer permutacija bez ponavljanja uzmimo za primjer zadatak gdje su nam zadana 3 slova – A, B, C i treba izvesti sve moguće razmještaje tih elemenata. Kada bismo sami išli tražiti sva rješenja našli bismo ih 6. Ta rješenja su $ABC, ACB, BAC, BCA, CAB, CBA$. Znamo da formula za računanje permutacija bez ponavljanja glasi: $P_n = n!$, tako bi u našem primjeru bilo $P_3 = 3! = 3 \cdot 2 \cdot 1 = 6$.

No, kako se dođe do te formule. Na slici ispod prikazano je 3 stupca sa 3 retka. Svaki redak predstavlja jedno od 3 mjesta za slova A, B ili C. Na prvo mjesto možemo staviti bilo koje slovo: A, B ili C. Recimo da stavimo slovo A, za drugo mjesto nam preostaju dvije opcije: B ili C. Ako uzmemo B, na zadnje mjesto možemo još samo staviti C. Isto tako napravimo i kada je B na prvom mjestu i kada je C na prvom mjestu. Kako bismo dobili ukupan broj mogućnosti

(permutacija) za ta tri mjesta treba pomnožiti brojeve mogućnosti svih mjesta, u ovom slučaju $3 \cdot 2 \cdot 1 = 3! = 6$.

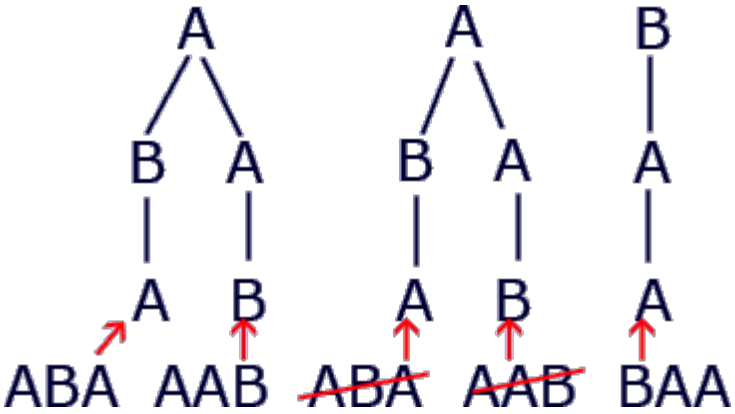


Slika 1: Permutacije bez ponavljanja

2.2.2. Permutacije s ponavljanjem

Da bismo prikazali permutacije s ponavljanjem, uzet ćemo sličan primjer kao i kod permutacija bez ponavljanja, samo ćemo ovaj puta imati slova A, A i B. Ovaj puta imamo sljedeća rješenja: AAB, ABA, BAA. To su dakle 3 rješenja s obzirom da imamo 2 ista slova. Nazovimo prvo slovo $A - A1$, a drugo $A2$. Rješenja $A1A2B$ i $A2A1B$ su jednaka stoga imamo upola manje konačnih rješenja nego kada bi sva slova bila različita. Formula za računanje permutacija s ponavljanjem glasi $P = \frac{n!}{(s! \cdot r! \cdot k!)}$, gdje je n broj elemenata u permutaciji, a s, r, k su brojevi istih elemenata u permutaciji. Uvjet u formuli jest $s + r + \leq n$. [6].

Tako bi u našem slučaju formula glasila $P = \frac{3!}{2!} = \frac{(3 \cdot 2 \cdot 1)}{(2 \cdot 1)} = 3$. Na slici dolje se lijepo vidi kako prvi i drugi stupac sa slovima A na prvom mjestu daju jednaka rješenja.



Slika 2: Permutacije s ponavljanjem

2.3. Primjeri permutacija

Ovdje ću prikazati primjere nekoliko praktičnih zadataka za čije rješavanje koristimo permutacije.

Za prvi najjednostavniji primjer uzmimo da imamo 9 ljudi koji stoje u redu ispred nekog dućana i nas zanima na koliko načina se tih 9 ljudi može poredati. Prvo označimo svakog čovjeka brojevima od 1 do 9. Tu se pitamo koliko ustvari ima linearnih poredaka ovih 9 elemenata, odnosno na koliko načina možemo popuniti tablicu (, , , , , , , ,) a da iskoristimo sve elemente od 1 do 9. Na prvom mjestu može doći bilo tko, dakle imamo 9 mogućnosti, na drugom mjestu imamo jednu mogućnost manje jer smo jedno mjesto popunili. Isto tako vrijedi da za treće mjesto imamo 7 mogućnosti, četvrto 6 i sve tako do zadnjeg mjesta gdje je 1 mogućnost. Pošto moramo popuniti sva mjesta množimo broj svih mogućnosti na svakom mjestu i to je $9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 9! = 362880$. Dakle rješenje jest 362 880, što je jako velik broj načina na koje samo 9 ljudi može biti poredano u nizu.

Uzet ćemo još jedan sličan primjer, no ovog puta treba poredati 5 muškaraca i 4 žene u red na način da alterniraju u redu muškarac pa žena. Već znamo da bi 5 muškaraca mogli poredati na 5! načina. Ako se tih 5 muškaraca razmakne tako da stane žena između svakog od njih dobili bismo još $4!$ načina na koji bi se žene mogle poredati. Dakle u ovom slučaju dobili smo 2 zasebne permutacije, jednu od muškaraca i jednu od žena, a da bismo dobili jednu permutaciju treba te dvije vrijednosti pomnožiti. Prema tom pravilu množenja rezultat je: $5! \cdot 4! = 120 \cdot 24 = 2880$. Vidimo da iako i dalje imamo 9 ljudi, ali smo malo izmijenili uvjete, dobijemo znatno manji broj mogućih permutacije nego kod prvog primjera.

3. Kombinacije

3.1. Definicija kombinacije

Neka je S bilo koji skup, a $r \geq 0$ cijeli broj. Bilo koji r-člani podskup od S zove se r-kombinacija skupa S. Drugim riječima r-kombinacija od S je neuređeni izbor od r elemenata iz S. Skup svih r-kombinacija od S bilježimo simbolom $\binom{S}{r}$, ili katkada $Pr(S)$. Na primjer, $\binom{a,b,c,d}{3} = a,b,c,a,b,d,a,c,d,b,c,d$. S druge strane, broj svih r-kombinacija n-članog skupa bilježimo simbolom (čitaj n povrh r, ili n iznad r, engleski se čita „n choose r“): $\binom{n}{r}$, i zovemo ga binomni koeficijent. Dodatno definiramo $\binom{0}{0} = 1$. Ako je $r > n$, onda je jasno da je $\binom{n}{r} = 0$, a ako je $n = 0$, a $r \in \mathbb{N}$, onda je $\binom{0}{r} = 0$. Isto tako je jasno da je $\binom{n}{0} = 1, \binom{n}{1} = 1$. Pojam kombinacije toliko je fundamentalan da se čitava grana matematike prema tome pojmu naziva kombinatorika. Katkad se punim nazivom r-kombinacija skupa S zove i r-kombinacija od S bez ponavljanja. [9] Jednostavnije rečeno, kombinacija je način izbora nekog broja elemenata iz nekog skupa, gdje redoslijed (za razliku od permutacija) elemenata nije važan. Konačna formula za kombinacije jest $\binom{n}{r} = \frac{n!}{r!(n-r)!}$

3.2. Pascalov trokut

3.2.1. Povijest Pascalovog trokuta

Pascalov trokut se na primjer koristi u algebri, kombinatorici i vjerojatnosti, kod poligonalnih i Fibonaccijevih brojeva. Naziv je dobio po cijenjenom francuskom matematičaru i filozofu Blaise Pascalu (1623.-1662.), ali se je njegov eksplicitni prikaz javljao i mnogo ranije. Binomni koeficijenti javljaju se u 10. stoljeću u Chandashaastri, drevnoj indijskoj knjizi, koju je napisao Pingala negdje između 5. i 2. stoljeća prije Krista. Sljedeće javljanje bilo je u Kini, gdje je kineski matematičar Ji Xian u 11. stoljeću koristio trokut za izračunavanje drugog i trećeg korijena. Nakon toga u 13. stoljeću Yang Hui (1238.-1298.) prikazao je aritmetički trokut sa šest redaka koji je bio jako sličan Pascalovom trokutu kojeg mi danas znamo. Danas se u Kini Pascalov trokut naziva "Yang Hui-ov trokut". Oko 1303. godine još jedan kineski matematičar Zhu Shijiei proširuje trokut na osam redaka. Postoje i neki dokazi da je ovaj numerički trokut bio poznat arapskom astronomu, pjesniku i matematičaru Omaru Khayyamu iz 11. stoljeća. Zbog istovremenog pojavljivanja trokuta u Kini i Perziji ne zna se točno jesu li ga Arapi nezavisno otkrili ili su ga preuzeli od Kineza. Iako je ovaj numerički troku ime dobio po Blaise Pascalu, on nije bio jedini u Europi koji je proučavao binomne koeficijente u to vrijeme. Današnji se Pascalov trokut u Europi pojavio 1527. godine, a Blaise Pascal je sto godina poslije tog otkrića počeo proučavati aritmetički trokut. U njegovom djelu *'Traite du triangle arithmetique*, izdanom nakon njegove smrti, pojavljuje se aritmetički trokut, koji se od nama poznatog Pascalovog trokuta

razlikuje po tome što je zarotiran za 45 stupnjeva.

3.2.2. Formula

Pascal je do svoje formule došao tako da je binomne koeficijente $\binom{n}{r}$ za male vrijednosti n i r napisao u obliku tablice koja se zove Pascalov trokut. Pascalova formula je zapravo rekurzija za binomne koeficijente, jer se njome počevši s prva dva stupca može izračunati svaki daljnji član Pascalova trokuta. [9]

Za sve brojeve $r, n \in \mathbb{N}$, takve da je $r \leq n$ vrijedi:

$$\binom{n}{r} = \binom{n}{n-r}$$

To možemo dokazati koristeći definiciju binomnih koeficijenata.

$$\binom{n}{n-r} = \frac{n!}{(n-r)!(n-(n-r))!} = \frac{n!}{(n-r)!r!} = \binom{n}{r}$$

3.2.3. Primjer u kombinatorici

Uzmimo za primjer da imamo 3 olovke i želimo izračunati na koliko različitih načina možemo odabrati dvije od njih. Broj svih kombinacija od tri elementa drugog reda je $\binom{3}{2}$. Ovaj broj sadržan je u Pascalovom trokutu na drugom mjestu trećeg reda, stoga možemo vidjeti da imamo tri mogućnosti da odaberemo dvije različite olovke od njih 3.
Brojevi u Pascalovom trokutu lako se mogu generirati koristeći formulu za broj kombinacija, samo trebamo odrediti red i mjesto broja kojeg želimo.
Znamo da je formula za kombinacije $\binom{n}{r} = \frac{n!}{r!(n-r)!}$. Ako sada želimo pronaći 3. broj u 5. retku samo ubacimo u formulu $\binom{5}{3} = \frac{5!}{3!(5-3)!} = \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{3 \cdot 2 \cdot 1(2 \cdot 1)} = 10$

Ako pogledamo Pascalov trokut, vidimo da to uistinu i jest broj 10.

Na slici 3 dolje prikazan je Pascalov trokut. Po redcima su navedene vrijednosti koje n može poprimiti, a po stupcima su poredani binomni koeficijenti koji poprimaju različite vrijednosti, zavisno o poziciji u tablici. Pascalov trokut možemo koristiti za dobivanje broja r kombinacija n-članog skupa, a zove se trokut jer kao što vidimo svi rezultati čine oblik trokuta. Kao što smo već naveli, kod Pascalovog trokuta može se pomoću prva dva stupca izračunati svaki sljedeći član u trokutu. Na slici je grafički prikazano na koji način se dobivaju članovi trokuta, a to je da dva susjedna člana u stupcima zbrojimo i zapišemo zbroj ispod desnog člana i ako promotrimo to stvarno vrijedi za sve članove Pascalovog trokuta.

Tablica 1: Pascalov trokut

n	$\binom{n}{0}$	$\binom{n}{1}$	$\binom{n}{2}$	$\binom{n}{3}$	$\binom{n}{4}$	$\binom{n}{5}$	$\binom{n}{6}$	$\binom{n}{7}$	$\binom{n}{8}$	$\binom{n}{9}$
0	1									
1	1	1								
2	1	2	1							
3	1	3	3	1						
4	1	4	6	4	1					
5	1	5	10	10	5	1				
6	1	6	15	20	15	6	1			
7	1	7	21	35	35	21	7	1		
8	1	8	28	56	70	56	28	8	1	
9	1	9	36	84	126	126	84	36	9	1

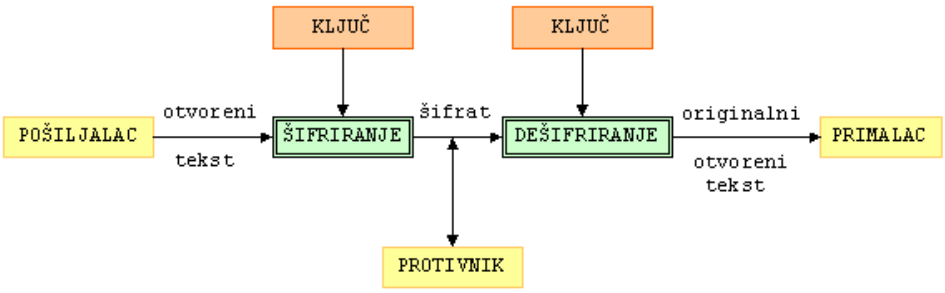
4. Primjene permutacija i kombinacija

U ovom odjeljku malo ćemo pomnije pogledati u kojim to sve područjima se primjenjuje kombinatorika, pa tako i permutacije i kombinacije prema T.Munakati. [5]

4.1. Komunikacijske mreže, kriptografija i mrežna sigurnost

Permutacije su često korištene u komunikacijskim mrežama i u paralelnim i distribuiranim sustavima. Usmjeravanje različitih permutacija na mreži za evaluaciju performansi je čest problem u tim poljima. Mnoge komunikacijske mreže zahtijevaju siguran prijenos informacija, što potiče razvoj kriptografije i mrežne sigurnosti. Ovo područje vrlo je bitno s obzirom da se danas jako puno informacija prenosi putem interneta. Problematika vezana uz to jest čuvanje privatnosti transakcija i drugih povjerljivih podataka, te uz to sačuvati mrežu od raznih napada virusa ili hakera. Proces enkripcije uključuje manipulaciju nizom kodova kao što su znamenke, znakovi ili riječi i zbog toga je taj proces usko vezan uz kombinatoriku.

Na slici 1 prikazan je proces enkripcije i dekripcije pomoću ključeva. Pošiljatelj šifrira poruku koju želi poslati javnim ključem, ta šifrirana poruka prolazi nesigurnim kanalom do primatelja koji ima privatni ključ koji dešifrira kriptiranu poruku.



Slika 3: Proces enkripcije i dekripcije ključevima [11]

4.1.1. Permutacijska šifra

Permutacijska šifra je vrsta šifriranja koja se koristi u kriptografiji. Ono po čemu se permutacijska šifra razlikuje od nekih drugih vrsta šifri jest to što je njezina ideja zadržati znakove otvorenog teksta onakvim kakvi jesu, bez zamjene s nekim znakovima šifrata. Dakle, permutacijska šifra radi, kao što i njeno ime govori, na način da svi znakovi teksta koji se šifrira ostanu, no promijeni se njihova pozicija koristeći permutacije.

Znamo da je permutacija način na koji se neki set elemenata može rasporediti, pa bi tako na primjer pravilo bilo da prvi element ide na treću poziciju, drugi element na prvu, te treći element na drugu poziciju u setu. To bi izgledalo ovako: $(1, 2, 3) \rightarrow (2, 3, 1)$. Ako je naš set $(plava, crvena, bijela)$, kada primjenimo permutaciju to bi se pretvorilo u $(crvena, bijela, plava)$.

Za setove koji imaju veći broj elemenata od skupa na kojem je zadana permutacija, u ovom slučaju veći od 3 elementa jer permutacija djeluje na 3 elementa, permutacija se primjenjuje na manje blokove elemenata. Ti blokovi su veliki koliko i permutacija, dakle ovdje bi imali blokove od 3 elementa. Uzmimo set od 6 elemenata: $(plava, crvena, bijela, zelena, siva, crna)$. Ovdje bismo primjenili permutaciju na način da uzmemo prva tri elementa kao jedan blok i druga tri elementa ko drugi blok i na kraju ih spojimo da dobijemo konačan rezultat. Na kraju bi dobili set $(crvena(2), bijela(3), plava(1), siva(2), crna(3), zelena(1))$.

Prilog radu je i algoritam za ovakav način enkripcije, implementiran pomoću JavaScript programskog jezika u obliku html dokumenta. Program funkcionira na način da riječ koju želimo enkriptirati grupira po broju slova koji odgovara broju slova ključa (ako ključ ima 3 slova tada riječ od 6 slova grupira u dvije grupe po 3 slova) i tada se ovisno o mjestu slova ključa u abecedi primjenjuje permutacija na zadanu riječ. Ako imamo slučaj da je zadana riječ od 8 slova, a ključ od 3, tada se prvih 6 slova grupira kao što je navedeno, a zadnja dva slova tvore zasebnu grupu. Ta zadnja grupa se permutira na isti način kao i ostale, samo što je njima pridruženo u treće "prazno" slovo koje se samo izostavi kada se ispisuje kriptirana zadana riječ.

4.2. Arhitektura računala

Kod dizajna računalnih čipova treba uzeti u obzir moguće permutacije ulaznih prema izlaznim konektorima. Programabilni interkonekcijski čipovi korisnicima omogućuju međusobno povezivanje za željenu permutaciju. Raspored logičkih sklopova je osnovni element u dizajniranju arhitekture računala. Programabilni interkonekcijski čipovi pružaju mogućnost ostvarivanja korisničkog programiranja interkonekcije za bilo koju željenu permutaciju. Takva interkonekcija je vrlo poželjna za podršku brzoj izradi prototipova hardverskih sustava i pružanju programabilnih komunikacijskih mreža za paralelno i distribuirano računalstvo. Takav čip bi trebao ostvariti svaku moguću permutaciju ulaznih i izlaznih pinova pomoću skupa programabilnih sklopki. [5]

4.2.1. Interkonekcijske mreže

U svakom sustavu koji uključuje više procesora potreban je interkonekcijski sustav da bi se procesorima pružio pristup memoriji ili da se napravi mehanizam za komunikaciju podacima između procesuiranih elemenata. U nekim sustavima da bi se osigurao protok podataka kroz interkonekcijsku strukturu, potrebno je definirati na koji način podatci protječu. Ako se definira da svaka izvorna adresa ima jedinstvenu adresu destinacije, tada govorimo o permutacijama na izvornim adresama i one se mogu matematički definirati. Ovdje ćemo za primjer uzeti da je broj ulaznih i izlaznih portova jednak, te ćemo prikazati neke od vrsta permutacija koje se koriste za spajanje svakog ulaznog s jednim od izlaznih portova. Navedeni primjeri bit će oni jednostavnijih permutacija, no oni su temelj za puno kompleksnije načine koji se danas koriste za pravljenje mehanizama za komunikaciju podacima. Svaka od ovih permutacija ima neke prednosti i mane (brzina protoka informacija, bolja protočnost, jednostavnost itd.), te su današnje mreže sastavljene od više vrsta kako bi se od svake dobilo najbolje svojstvo.

4.2.1.1. Permutacija savršenog miješanja

Ova permutacija naziva portova koristi se za mapiranje seta izvornih portova S do završnih portova D. Nazive portova podijelimo na dva podseta jednake veličine koji se zatim isprepletu. Ova permutacija može se dobiti jednostavnom manipulacijom binarnih reprezentacija portova. Za primjer uzmimo set binarnih reprezentacija ulaznih portova $S = \{000, 001, 010, 011, 100, 101, 110, 111\}$, te set završnih portova $D = S$. Označimo znamenke svakog elementa s brojevima $(1, 2, 3)$; permutacije dobijemo na način da promijenimo redoslijed znamenaka u $(2, 3, 1)$. Kada bi na našem primjeru spajali ulazne portove sa završnim, to bi izgledalo ovako: $000 \rightarrow 000$, $001 \rightarrow 010$, $010 \rightarrow 100$, $011 \rightarrow 110$, $100 \rightarrow 001$, $101 \rightarrow 011$, $110 \rightarrow 101$, $111 \rightarrow 111$

4.2.1.2. Permutacija zamjene

Ovo je još jedna korisna i jednostavna permutacija koja se može koristiti s permutacijom savršenog miješanja zbog toga što ona povezuje i portove s različitim brojem jedinica i nula, što prijašnjoj permutaciji fali.

Za primjer ćemo uzeti iste setove kao i u prethodnom primjeru;
$S = D = \{000, 001, 010, 011, 100, 101, 110, 111\}$. Sljedeće što napravimo jest da prvi port iz prve domene S spojimo s drugim iz domene D, te drugi iz prve s prvim iz druge domene. Tako napravimo i sa svim ostalim parovima u domeni i konačan rezultat je sljedeći. $000 \leftrightarrow 001$, $010 \leftrightarrow 011$, $100 \leftrightarrow 101$, $110 \leftrightarrow 111$. Može se primjetiti da se u spojenim parovima razlikuje samo treći bit binarnog broja.

4.2.1.3. Leptirasta permutacija

Kod ove permutacije najvažniji i najmanje važan bit (prvi i zadnji) se zamijene, tvoreći vezu. Za primjer ćemo uzeti setove adresa $S = D = \{000, 001, 010, 011, 100, 101, 110, 111\}$. Nakon primjene leptiraste permutacije na adrese dobiju se sljedeće veze: $000 \rightarrow 000$, $001 \rightarrow 100$, $010 \rightarrow 010$, $011 \rightarrow 110$, $100 \rightarrow 001$, $101 \rightarrow 101$, $110 \rightarrow 011$, $111 \rightarrow 111$
Na slici 4 ispod grafički je prikazano spajanje ulaznih i izlaznih portova na mreži pomoću leptiraste permutacije gdje je broj portova $N = 8$.

Slika 4: Leptirasta permutacija [11]

4.3. Jezici

Prirodni i računalni jezici su usko povezani s kombinatorikom jer su komponente tih jezika – rečenice, paragrafi, programi i blokovi složeni od manjih elemenata – riječi, znakova, naredbi. Na primjer, algoritam za pronalaženje stringova može se zasnivati na kombinatorici riječi i znakova. Jedan takav algoritam jest algoritam za traženje anagrama, odnosno traženje svih permutacija slova neke riječi. Direktna primjena tih algoritama može biti u smislu procesuiranja riječi i baza podataka.
Primjena permutacije i kombinacija kod jezika dešava se i u smislu zaštite poruka kroz neki nesiguran kanal. Da bi se poruka zaštitila koriste se permutacije i zamjene riječi u kombinaciji s tajnim ključevima koji osiguravaju da pravu poruku dobije samo onaj kome je ista i namijenjena.

4.4. Analiza uzoraka

U širem smislu, sva ova navedena područja primjene mogu se gledati kao specijalni slučajevi analiziranja ponavljajućih uzoraka. Molekularna biologija proučava uzorke atoma, molekula i DNA, dok jezici gledaju poredak rečenica, riječi i znakova. Uzorci mogu imati i mnoge druge oblike kao što su vizualne slike, akustični signali i ostale fizičke veličine – struja, tlak, temperatura i druge koje se pojavljuju kod inženjerskih problema. Uzorci mogu biti i apstraktni bez nadodanog fizičkog značenja, a takvi se pojavljuju kao digitalne, analogne i još neke druge jedinice. Elektronička glazba također može biti gledana kao specijalna domena kombinatorike kod akustičnih signala.

4.4.1. Kombinatorika u glazbi

Primjena kombinatorike na glazbu je mnogo - naglašavanje, mjerenje, kombinacije nekoliko glazbenih parametara, brojanje nota itd. Jedna poznata primjena kombinatorike u glazbi jesu "igre kockica", od kojih ćemo jednu takvu i ovdje opisati.
Mozart je razvio poznatu igru kockica. Sastoji se od tablice s 12 redaka i 8 stupaca. (i, j) -broj na tablici predstavlja glazbenu mjeru. Bacanjem dvije kockice prvi put i dodajući ih, dobivamo broj, vrijednost "i", odnosno broj retka. U stvari, broj redova je dan dodavanjem na dvije kockice. Dakle, prva mjera bit će prvi (stupac $j = 1$) element u i-bacanju. Bacanjem kockica ponovno dobivamo drugi broj i uzimamo element u i'-bacanju 2. stupca. Do sada smo dobili prve dvije mjere novog menueta. Nastavimo na isti način i dobivate niz od 8 mjera. Ako imamo 16 stupaca na stolu, možemo sastaviti minuet od 16 mjera. Sumirajući, za svaku mjeru, bacamo kockice i zapisujemo odgovarajuću glazbenu mjeru. Postoji 118 mogućih kombinacija od po osam mjera.[16]

4.5. Znanstvena otkrića

Za neke određene tipove znanstvenih otkrića potrebno je koristiti kombinatorne sekvence u procesu dobivanja mogućih rješenja za probleme. Za primjer uzmimo znanstvena otkrića u kemiji ili biologiji. Može nam biti potrebno da imamo sve moguće rezultate ili ishode kemijskih ili bioloških reakcija. U svakom koraku te reakcije moguće je generirati kombinacije radikala, baza i kemijskih spojeva kao moguća rješenja i od tih svih mogućih selektira se koje su najvjerojatnije kombinacije u danim uvjetima. Kao drugi primjer možemo uzeti matematiku gdje u područjima teorije grafova i teorije brojeva možemo generirati kombinatorne sekvence kao moguća rješenja.

4.5.1. Teorija grafova

Prvo trebamo definirati graf. Jednostavni graf G sastoji se od nepraznog konačnog skupa $V(G)$, čije elemente zovemo vrhovi (čvorovi) grafa G i konačnog skupa $E(G)$ različtih dvočlanih podskupova skupa $V(G)$ koje zovemo bridovi. Skup $V(G)$ zovemo skup vrhova i ako je jasno o kojem je grafu G riječ označavat ćemo ga kraće samo s V, a skup $E(G)$ zovemo skup bridova i označavat ćemo ga i samo s E. Formalno, ponekad ćemo pisati $G = (V(G), E(G))$ ili kraće još i $G = (V, E)$.[9]

Grafovi su jedna od osnovnih matematičkih struktura, stoga se oni i pojavljuju u raznim oblicima i u različitim situacijama. Mnoge se pojave modeliraju grafovima(dijagramima) koji se sastoje od točaka i njihovih spojnica. Na primjer, točke (vrhovi ili čvorovi) mogu predstavljati ljude iz nekakve skupine, a spojnice (bridovi) parove prijatelja, ili točke mogu predstavljati komunikacijske centre, a spojnice komunikacijske veze između komunikacijskih centara. Graf može predstavljati električnu mrežu, čiji su vrhovi električke komponente, a spojnice električne veze. Cestovne, željezničke, zrakoplovne i druge veze daljnji su primjeri primjene modela s

grafovima. U računarstvu se često dijagram toka nekog algoritma prikazuje grafom kojem su čvorovi naredbe (instrukcije), a lukovi iz jedne u drugu naredbu su bridovi. Isto se tako grafovima prezentiraju i razne kompjuterske strukture podataka, umrežavanje i paralelizam računala i njihov sekvencijalni rad, evolucijska ili porodična stabla u biologiji, kemijske veze među atomima ili molekulama, raspored poslova u velikim gospodarskim projektima itd. Dakle obično se kod primjene grafova u kombinatorici broje vrhovi i bridovi kojima pripišemo neko značenje ovisno o području gdje ih koristimo te se traže njihove permutacije i kombinacije da bi se dobio željeni rezultat (nekad je to optimalna ruta, nekad određeni kemijski spoj, poveznice u komunikaciji itd.) [9].

4.6. Baze podataka i rudarenje podataka

Upiti u bazama podataka su operacije višestrukog pridruživanja koje su permutacije sastavnih operacija pridruživanja. Određivanje optimalne permutacije koja ima najmanje troškove jest čest i bitan problem. Rudarenje podataka ili otkrivanje znanja u bazama podataka jest sortiranje i grupiranje velikog broja podataka i ono cilja na izlučivanje korisnih informacija iz velikih baza podataka. Metode koje se koriste kod rudarenja su razni algoritmi ili neuronske mreže gdje je kombinatorika jako zastupljena i vrlo bitna.[10]

Važnost rudarenja podataka jest u tome što se danas dobiva sve veći broj raznovrsnih podataka u svim sferama života i poslovanja. Skup svih tih podataka jest kaotičan i nestrukturiran. Da bismo imali korist od tih podataka treba ih sortirati i izvući iz toga one informacije koje su nama bitne i relevantne da bismo mogli donijeti što bolje i informiranije odluke pomoću dobivenih podataka.[10]

Treba napomenuti da rudarenje podataka ne počinje s nekom pretpostavkom koja mora biti dokazana ili pobijena. Ono nije tradicionalni proces potvrđivanja znanja već je to istraživački proces koji cilja na otkrivanje novog znanja. Rudarenje podataka obično će otkriti neke neočekivane informacije, ponavljajuće uzorke ili prilike za poboljšanje područja u kojem se koristi. [13]

4.6.1. Područja primjene rudarenja podataka

Rudarenje podataka primjenjivo je na svim područjima u kojima se može dobiti velik broj podataka u neku bazu podataka i nalazi se u srcu analitike kroz razne industrijske grane.

4.6.1.1. Komunikacije

Na velikom tržištu gdje je jaka konkurencija potrebna je svaka korisna informacija da bi se postigla prednost pred drugima i to se često može dobiti ekstrakcijom informacija iz podataka o korisnicima. Multimedijske i telekomunikacijske kompanije tako mogu koristiti razne analize da bi iz brda podataka o svojim potrošaćima mogli predvidjeti njihove potrebe ili da bi im pružili ciljane i relevantne ponude.

Ključ uspjeha telekomunikacijskih tvrtki je segmentiranje tržišta i ciljanje sadržaja prema svakoj grupi. Ovo zlatno pravilo je relevantno za različita područja poslovanja. Kada govorimo o telekomunikacijama, postoje četiri sheme segmentacije od primarne važnosti: segmentacija vrijednosti kupaca, segmentacija ponašanja kupaca, segmentacija životnog ciklusa korisnika i segmentacija migracije korisnika. Napredno ciljanje omogućuje predviđanje potreba, preferencija i reakcija korisnika na telekomunikacijske usluge i proizvode u ponudi. Omogućuje poboljšano poslovno planiranje i ciljanje.

Telekomunikacijska industrija koja svakodnevno privlači jako velik broj korisnika ogromno je područje za prijevarne aktivnosti. Najrašireniji slučajevi prijevare u području telekomunikacija su ilegalni pristup, autorizacija, krađa ili lažni profili, kloniranje, prijevare u ponašanju, itd. Prijevara ima izravan utjecaj na odnos uspostavljen između tvrtke i korisnika. Stoga, sustavi za otkrivanje prijevara, alati i tehnike pronašli su široko korištenje. Primjenjujući algoritme strojnog učenja bez nadzora na ogromnu količinu podataka o klijentima i operatorima kako bi uočili karakteristike normalnog prometa, možete spriječiti prijevaru. Algoritmi definiraju anomalije i uz pomoć tehnika vizualizacije podataka predstavljaju ih kao upozorenja analitičarima u realnom vremenu. Primjena kombinatorike u ovom području jest u algoritmima koji su ranije navedeni. Ti algoritmi uzimaju razne podatke koji se smatraju važnima i pomoću matematičkih formula u koje su uključene i one iz kombinatorike dolaze do željenih rezultata. Učinkovitost ove tehnike je vrlo visoka jer omogućuje odgovor na sumnjivu aktivnost gotovo u realnom vremenu[15].

4.6.1.2. Edukacija

Kod studenata i učenika moguće je pratiti mnogo parametara vezanih uz njihovo znanje i napredovanje u edukaciji. Te sve podatke moguće je sortirati, analizirati, te naposlijetku iz njih vidjeti i predvidjeti kako pojedini studenti ili učenici napreduju, odnosno može se dobiti cjelovit uvid u njihove aktivnosti, uspjehe ili slabosti. Edukatori ili profesori pomoću tih podataka mogu napraviti i razviti strategije za pojedine studente koje će im olakšati savladavanje potrebnih znanja i vještina, te im olakšati da ostanu na pravom putu za izvršavanje svih obaveza i ispunjavanja njihovih ciljeva u školovanju ili edukaciji. Pomoću rudarenja tih podataka edukatori mogu i lakše predvidjeti kakav će uspjeh studenti postizati te identificirati gdje su čije slabosti ili jače strane i prema tome pridavati odgovarajuću pažnju.

Nekoliko specifičnih primjera upotrebe rudarenja podataka u edukaciji su[14]:

- primjena relacijskog rudarstva kako bi se pronašao odnos između uspjeha učenika u predmetima koji sadrže nekoliko poglavlja organiziranih u lekcije, pri čemu svaka lekcija uključuje više koncepata

-otkriće odnosa između studentskih ponašanja i karakteristika učenika ili kontekstualnih varijabli; Analiza istraživačkog pitanja u raznim kontekstima

-pronaći odnos između razine obrazovanja roditelja i studenata koji su se povukli iz škole, otkriće kurikularnih asocijacija u nizovima kolegija; otkrivanje koje pedagoške strategije dovode do učinkovitijeg / robusnijeg učenja

-identificirati rizične studente, razumjeti ishode učenika u obrazovanju (npr. predvidjeti

koji studenti će vjerojatno otpasti na prvoj godini fakulteta, te dokučiti kako čim više smanjiti taj rizik od otpadanja kod tih rizičnih skupina studenata) - ovaj primjer ćemo obraditi na temelju stvarnog istraživanja provedenog u Indiji od strane S.Pal [8]

Na slici 3 ispod vidimo upitnik koji je proveden u kojemu se žele dobiti informacije o studentima. Neke od tih informacija su kategorija studenata, spol, ocjene u srednjoj i osnovnoj školi, način upisa, jezik na kojem se školuju, veličina i status obitelji, mjesto stanovanja, obrazovanje roditelja, dohodak obitelji, zanimanje roditelja i na kraju je li student nastavio školovanje nakon prve godine. Ovo istraživanje provedeno je na 165 studenata jednog sveučilišta u Indiji.

Variables	Description	Possible Values
Branch	Students Branch	{CS, IT, ME)
Sex	Students Sex	{Male, Female}
Cat	Students category	{Unreserved, OBC, SC, ST}
HSG	Students grade in High School	{O – 90% -100%, A – 80% - 89%, B – 70% - 79%, C – 60% - 69%, D – 50% - 59%, E – 40% - 49%, F - < 40%}
SSG	Students grade in Senior Secondary	{O – 90% -100%, A – 80% - 89%, B – 70% - 79%, C – 60% - 69%, D – 50% - 59%, E – 40% - 49%, F - < 40% }
Atype	Admission Type	{UPSEE, Direct}
Med	Medium of Teaching	{Hindi, English}
LLoc	Living Location of Student	{Village, Town, Tahseel, District}
Hos	Student live in hostel or not	{Yes, No}
FSize	student's family size	{1, 2, 3, >3}
FStat	Students family status	{Joint, Individual}
FAIn	Family annual income status	{BPL, poor, medium, high}
FQual	Fathers qualification	{no-education, elementary, secondary, UG, PG, Ph.D., NA}
MQual	Mother's Qualification	{no-education, elementary, secondary, UG, PG, Ph.D., NA}
FOcc	Father's Occupation	{Service, Business, Agriculture, Retired, NA}
MOcc	Mother's Occupation	{House-wife (HW), Service, Retired, NA}
Dropout	Dropout: Continue to enroll or not after one year	{Yes, No}

Slika 5: Upitnik za istraživanje [8]

Dobiveni podaci iz upitnika ubačeni su u program za rudarenje podataka koji može korelirati sve kategorije s time je li student nastavio školovanje ili nije. Na slici 4 prikazani su dobiveni podaci od tog programa. Prikazane su one kategorije koje su dobile rezultat veći od 0.5 što znači da je ta kategorija značajno povezana s nastavkom školovanja studenata. Možemo vidjeti da je tipičan profil studenta koji ne nastavlja školovanje nakon prve godine muška osoba s niskim prosjekom ocjena u srednjoj školi sa direktnim upisom, pohađa nastavu na hindi jeziku, živi na selu, majka mu ima samo osnovnu školu i radi u nekoj službi

Variable	Values	Probability
Sex	Male	0.68
SSG	E	0.6623
Atype	Direct	0.6
Med	Hindi	0.76
Lloc	Village	0.55
Mqual	Elementry	0.50
Mocc	Service	.52

Slika 6: Relevantne varijable [8]

Tako dobiveni podaci tada su se primjenili na nove studente kako bi se predvidjelo hoće li oni nastaviti nakon prve godine, i na slici 5 prikazani su dobiveni rezultati. Od 165 studenata, predviđeni i stvarni rezultati podudaraju se u čak 144 slučaja, od toga 121 student za kojeg se predvidjelo da neće nastaviti školovanje uistinu nije nastavio.

Dropout		Predicted	
		Yes	No
Actual	Yes	121	10
	No	11	23

Slika 7: Točnost predviđanja[8]

4.6.1.3. Bankarstvo

Bankarski svijet danas je jako kompetitivan i mnogo banaka se natječe za iste klijente. Automatizirani algoritmi koji se koriste u rudarenju podataka pomažu bankama da bolje razumiju svoje klijente i samim time da ponude što bolju uslugu istima što im u konačnici donosi i

više zadovoljnih i novih klijenata. Algoritmi također služe za razumijevanja miljuna transakcija koje se svakodnevno dešavaju u bankama što opet rezultira optimizacijom i poboljšanjem bankovnog sustava. Rudarenje bankovnih podataka još omogućuje bolji uvid u tržišne rizike, brže prepoznavanje i detekciju prevara i dobivanje optimalnih povrata na marketinška ulaganja.

Prije nego što se rudarenje podataka može početi provoditi, banke moraju napraviti skladište podataka. Skladištenje podataka jest proces ekstrahiranja, čišćenja, transformiranja i standardiziranja nekompatibilnih podataka iz bankovnog sustava kako se ti podaci mogu koristiti za analiziranje i otkrivanje korisnih uzoraka, veza i poveznica. [13]

U bankarstvu, neka konkretna pitanja na koja bi se rudarenjem podataka moglo odgovoriti su:

1. Koje transakcije klijent napravi prije prebacivanja u konkurentsku banku? (Zbog sprečavanja osipanja klijenata)
2. Kakav slijed kreditnih transakcija dovodi do prevara ili pronevjera. (da bi se spriječile i odbile prevare)
3. Kakav je profil dužnika s visokim rizikom? (za sprječavanje zadanih, loših kredita i za poboljšanje provjeravanja)
4. Koje bankovne proizvode često koriste zajedno koje skupine kupaca? (za prodaju i ciljani marketing)
5. Koje usluge i pogodnosti bi sadašnji kupci vjerojatno željeli? (da bi se povećala lojalnost i zadržavanje korisnika)

4.6.1.4. Osiguranje

Osiguravajuće kuće također imaju jako velik broj klijenata te stalno traže nove načine da bi preciznije i bolje sastavili ponude koje bi bolje odgovarale klijentima, a istovremeno donijele njima veću zaradu. Ovdje rudarenje podataka može biti od iznimno velike pomoći gdje tvrtke koriste razne tehnike da bi efektivnije odredile cijenu njihovih osiguravajućih polica, a isto tako i da pronađu nove načine i oblike usluga i polica koje mogu ponuditi svojim klijentima tako da se bolje zadovolje njihove potrebe. Tvrtke pomoću raznih analitičkih alata mogu i bolje rješavati kompleksne probleme koji se tiču raznih prevara, zakonske probleme, upravljanje rizikom te moguće smanjenje broja klijenata.

Osiguravatelji sada grade svoje personalizirane ponude svojim klijentima na temelju svojih preferencija i podataka o ponašanju,a nude im i inovativne usluge koje pojednostavljuju proces kupnje. Zdravstvena osiguravajuća društva koriste podatke o aplikacijama i nosivim podacima, što im omogućuje da proaktivno prate svoje klijente, pomažući klijentima u upravljanju njihovim zdravstvenim uvjetima i kroničnim bolestima. Neke tvrtke koriste narukvice za praćenje stanja klijenata (npr. prevencija i upravljanje dijabetesom) i to integriraju s korisnikovim planom osiguranja.

Još jedan primjer je iz sektora životnih osiguranja gdje je korisnicima omogućeno da brzo donesu odluku o svojoj osiguravajućoj polici putem online upitnika, državnih evidencija o motornim vozilima i drugih izvora podataka, koristeći tehnologije velikih podataka.

Neki osiguravatelji također poboljšavaju svoje iskustvo kupaca pomažući im poboljšati sigurnost. Postoji aplikacija koja procjenjuje ponašanje vozača u vožnji i dijeli savjete za poboljšanje navika vožnje. Također koriste i informacije o ponašanju vozača da bi kreirali odgovarajuću policu (sigurniji i bolji vozači dobivaju bolje uvjete osiguranja).

4.7. Operativna optimizacija

Mnogi problemi optimizacije u operativnom aspektu raznih poslova rješavaju se kombinatorikom. Problem raspoređivanja poslova je u suštini problem određivanja kojim redoslijedom bi se poslovi trebali raditi da im se što više skrati vrijeme i trošak. Ti poslovi mogu biti unutar računalnog sustava, mreže ili neke tvornice. Mnogi problemi vezani uz grafove i mreže uključuju slaganje rasporeda vrhova i rubova. Prodavači koji putuju moraju odrediti kojim redoslijedom koje gradove posjetiti da bi smanjili ukupan put.
Jedan dobar primjer optimizacije korištenjem raznih algoritama, permutacija i kombinacija jesu Googleove karte. U tom programu gdje se nalaze sve karte i putevi svijeta, pronalaze se najbolje rute, uzimajući u obzir brzinu putovanja, dakle ceste kojima se ide (autocesta, brza ili obična cesta), moguće gužve na pojedinim putevima i dužina puta. Taj program koristi sve mu dostupne informacije kako bi odredio put koji odgovara našim zahtjevima, koji su nekad što kraće vrijeme, a nekad izbjegavanje autocesta ili što kraći put.
Još jedan dobar primjer kojeg Veljan [9] navodi u svojoj knjizi jest problem kineskog poštara, koji je prvi razmatrao kineski matematičar Kuan 1962. godine. Kod tog problema poštar u poštanskom uredu pokupi poštu, odlazi je razdijeliti i potom se vraća u ured. On mora barem jedanput proći svakom ulicom svoje četvrti, a želi rutu kojom će on najmanje hodati. Taj problem povezuje se i s teorijom grafova, što je već ranije navedeno da se u tom području koristi kombinatorika. Prevedeno na jezik teorije grafova problem je ekvivalentan s time da se u povezanom težinskom grafu s težinama koje predstavljaju duljine ulica nađe optimalna tura. Da ovdje sad ne ulazimo dublje u detalje, suština jest da se problem kineskog poštara rješava jednim algoritmom koji konstruira optimalnu turu na način da počne u nekom vrhu i u svakom koraku bira rezni brid (brid čijim se izostavljanjem povećava broj komponenti povezanosti) neprijeđenog podgrafa samo ako nema druge alternative.
Sličan problem jest i problem trgovačkog putnika koji se razlikuje od problema kineskog poštara po tome što trgovački putnik mora obići neke gradove i vratiti se na mjesto polaska tako da svaki grad posjeti točno jedanput i svakom cestom prođe najviše jedanput. Dakle matematička razlika jest u tome što se kod ovog problema bavimo obilaskom svih vrhova grafa, a ne svih bridova. Taj problem razradio je William R.Hamilton na grafu dodekaedra (pravilnog poliedra u čijem se vrhu sastaju tri pravilna peterokuta, te ima 20 vrhova, 30 bridova i 12 peterokuta)

4.8. Računalna sigurnost

Ovo je domena kojoj se u zadnjem desetljeću pridaje sve više pažnje zbog učestalih i sve većih prijetnji nacionalnoj sigurnosti. Da bi se suočili s izazovima sigurnosti, razvijene su mnoge inteligentne računalne tehnike koje uključuju i inteligentnu analizu uzorka ljudskog lica, rendgenskih fotografija, kemijskih supstanci, podatke od mreže bežičnih senzora itd. Razvijene su tehnike procesuiranja podataka koje iz tih svih informacija probiru i pregledavaju najbitnije, odnosno koje pronalaze potencijalne prijetnje.

Danas postoje i razvijaju se razni programi koji uzimaju velik uzorak podataka s računala koji uključuju sve prošle događaje i sigurnosne prijetnje, te različite metapodatke. Metapodaci se koriste za organizaciju, pronalaženje, preuzimanje, arhiviranje i očuvanje informacija elektroničkog materijala te za cijeli niz drugih specijaliziranih stručnih potreba i potreba iz svakodnevnog života. Oni pomažu ljudima da pronađu podatke koje trebaju i da osmisle kako ih najbolje mogu koristiti. To su strukturirane informacije koje opisuju i lociraju, odnosno olakšavaju pristup i primjenu izvora informacija. [7] Uzevši sve te opsežne podatke u obzir, ti programi mogu kvalitetno iskoristiti te podatke, pronaći uzorak u njima koji daje indikaciju za neki problem i tako detektirati i otklanjati različite prijetnje računalnoj sigurnosti kojih je danas izrazito mnogo.

5. Vjerojatnost

5.1. Statistička definicija vjerojatnosti

Za vjerojatnost nekog događaja u eksperimentu pretpostavljamo da "mjeri" učestalost pojavljivanja promatranog događaja kod velikog broja ponavljanja tog eksperimenta. Na primjer, kod bacanja "poštenog" novčića očekujemo da će se jednaki broj puta pojaviti broj odnosno grb, dakle vjerojatnost obaju događaja će iznositi 0.5. Međutim, jasno nam je da kod malog broja ponavljanja pokusa ne možemo očekivati da će zastupljenost obaju događaja biti jednaka, može biti čak i prilično različita, ali da će se s druge strane kod velikog broja ponavljanja eksperimenta zastupljenost promatranih događaja približiti vrijednosti od 50 posto. [2]
Ako označimo s n_A broj pojavljivanja događaja A u n ponavljanja eksperimenta, relativna frekvencija događaja A jednaka je

$$f_n(A) = \frac{n_A}{n}$$

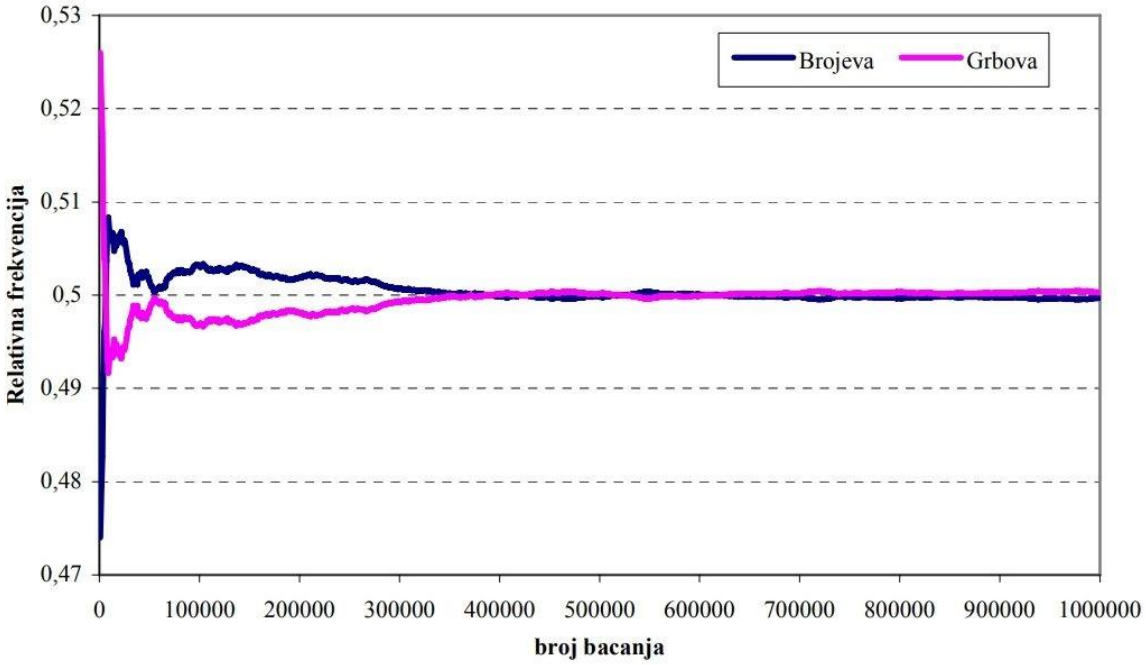
Očekujemo da će se s povećanjem n, vrijednost $f_n(A)$ stabilizirati, tj. da postoji

$$\lim_{n\to\infty} f_n(A)$$

Tu ćemo vrijednost nazvati **vjerojatnost** događaja A i označiti

$$P(A) = \lim_{n\to\infty} f_n(A)$$

Na grafu ispod su prikazane relativne frekvencije pojavljivanja brojeva odnosno grbova koje su dobivene korištenjem generatora slučajnih brojeva (nula i jedinica) u programskom paketu *Mathematica*. Izvršeno je 1 000 000 slučajnih odabira (svaki odabir roja predstavlja jedno bacanje novčića), te su nakon svakog izračunate relativne frekvencije. Vidimo da se s povećanjem broja bacanja obje vrijednosti približavaju vrijednosti 0.5. Na opisani se način izvršila simulacija slučajnog procesa. Simulacije u današnje vrijeme postaju sve ozbiljniji alat kod predviđanja slučajnih procesa. [2]



Slika 8: Vjerojatnost kod bacanja novčića

Naravno, ovako dobivena vjerojatnost nas može u nekim situacijama dovesti do pogrešnih zaključaka i zato da bi se smanjila ta mogućnost potrebno je napraviti velik broj eksperimenata i simulacija za dobru procjenu vjerojatnosti.

5.2. Klasična definicija vjerojatnosti

Ako promatramo eksperiment koji ima n mogućih ishoda, od kojih su svi jednako vjerojatni i disjunktni. Neka je za odabrani događaj A, m ishoda povoljnih. Onda je vjerojatnost događaja A jednaka kvocijentu broja povoljnih ishoda za događaj A i broja svih mogućih ishoda, tj. $P(A) = \frac{m}{n}$

5.3. Primjer permutacija u vjerojatnosti

Da bismo prikazali primjenu permutacija u izračunavanju vjerojatnosti uzet ćemo 2 relevantna primjera.

Prvi primjer glasi ovako: Od dijelova koji su označeni brojevima $1, 2..., 8$ može se sastaviti neki sklop proizvoljnim rasporedom elemenata. Treba odrediti vjerojatnost da se dobije sklop u kojem su elementi 2567 jedan do drugoga: a) u prikazanom rasporedu; b)u proizvoljnom rasporedu.

Rješenje
Broj svih mogućih rasporeda od 8 elemenata jednak je broju permutacija tj. $8!$.
a) Broj povoljnih rasporeda možemo dobiti na sljedeći način: ako elemente 2567 promatramo kao jednu cjelinu (1 element) uz preostala 4 elementa je moguće dobiti $5!$ različitih rasporeda. Prema tome je tražena vjerojatnost jednaka $\frac{5!}{8!} = \frac{1}{8 \cdot 7 \cdot 6} = 0.003$.

b) Za svaki mogući raspored na način kako je objašnjeno pod točkom a), elemente $2, 5, 6, 7$ možemo među sobom permutirati, što nam daje $4!$ mogućnosti. Stoga je ukupno povoljnih rasporeda $5! \cdot 4!$, odnosno tražena vjerojatnost je jednaka $\frac{5! \cdot 4!}{8!} = 0.071$.

Uzet ćemo još jedan primjer koji se odnosi na bacanje kocke, a glasi ovako: da li je vjerojatnije da u 4 bacanja kocke barem jednom padne šestica ili da u 24 bacanja dviju kocki padne barem jednom par šestica?
Rješenje
Kod 4 bacanja kocke je broj mogućih ishoda jednak 6^4. Prebrojimo broj ishoda koji za promatrani događaj nisu povoljni (događaj da nije pala niti jedna 6). Znači u četiri bacanja su pali brojevi od 1 do 5, što nam daje 5^4 mogućih ishoda. To znači da je povoljnih događaja $6^4 - 5^4$. Prema tome je vjerojatnost traženog događaja jednaka $\frac{6^4 - 5^4}{6^4} = 0.51775$.
Kod bacanja dviju kocki je broj mogućih ishoda jednak 36, pa je broj ishoda za 24 ponavljanja bacanja jednak 36^{24}. Kao i prije prebrojimo broj ishoda koji za promatrani događaj nisu povoljni (događaj da niti jednom nije pao par šestica). Znači u svakom se bacanju dogodio bilo koji od ishoda osim para $(6, 6)$. Kako je broj takvih ishoda 35, u sva 24 bacanja imamo ukupno 35^{24} mogućih ishoda. To znači da je povoljnih događaja $36^{24} - 35^{24}$. Prema tome je vjerojatnost traženog događaja jednaka $\frac{36^{24} - 35^{24}}{36^{24}} = 0.4914$.

Možemo zaključiti da je vjerojatnije da u 4 bacanja kocke barem jednom padne šestica nego da u 24 bacanja dviju kocki padne barem jednom par šestica.

5.4. Primjer kombinacija u vjerojatnosti

Primjer na kojem ćemo vidjeti primjenu kombinacija u izračunavanju vjerojatnosti je sljedeći: računalo ispisuje brojeve s 8 znamenaka. Uz pretpostavku da je pojava svake znamenke jednako vjerojatna, izračunat ćemo vjerojatnosti sljedećih događaja.

1. sve znamenke su različite

2. četiri posljednje znamenke su jednake

3. broj počinje s 5

4. broj sadrži tri znamenke jednake 5

Rješenje

S obzirom da svaku znameknu možemo birati iz skupa $0, 1, ..., 9$, imamo sveukupno $\binom{8}{10} = 10^8$ mogućnosti.
a)Broj, za ovaj slučaj povoljnih ishoda, jednak je broju permutacija osmog razreda od 10 elemenata, tj $P\binom{8}{10} = \frac{10!}{(10-8)!} = \frac{10!}{2!}$. Dakle vjerojatnost promatranog događaja je jednaka $\frac{10!}{2! \cdot 10^8} = 0,01814$.
b)Na prva četiri mjesta možemo proizvoljno odabrati znamenke na 10^4 načina, a na posljednja četiri mjesta se odabrana znamenka ponavlja tj. imamo 10 načina za odabir. Sveukupno je 10^5 povoljnih mogućnosti, pa je pripadajuća vjerojatnost jednaka $\frac{10^5}{10^8} = 0.001$.
c)Samo prva znamenka je fiksirana, a ostalih 7 biramo proizvoljno, što daje 10^7 mogućnosti, odnosno vjerojatnost $\frac{10^7}{10^8} = 0.1$.
d)Mjesta na kojima će se pojaviti znamenka 5 možemo odabrati na $\binom{8}{3}$ načina. Na preostalih 5 mjesta znamenke biramo proizvoljno iz skupa $0, 1, 2, 3, 4, 6, 7, 8, 9$ (bez broja 5). To možemo učiniti na 9^5 načina. Prema pravilu produkta možemo izračunati broj povoljnih ishoda, pa je vjerojatnost promatranog događaja jednaka $\frac{\binom{8}{3} \cdot 9^5}{10^8} = 0.0331$.

5.4.1. Isprogramiran primjer lutrije

Lutrija ili "lotto" je dobar primjer primjene vjerojatnosti i kombinacija, stoga sam odlučio napraviti jedan program koji izračunava koje su šanse da se od zadane količine brojeva pogodi određen broj brojeva ili "kuglica". Program prvo računa koje su šanse da se pogodi određen broj brojeva iz zadane količine brojeva. Nakon toga, podijeli se dobiven broj sa brojem kombinacija koje se mogu napraviti od pogođenih brojeva jer redoslijed nije bitan i tako se na kraju dobije koja je šansa da se pogode određeni brojevi, bez obzira na njihovu kombinaciju.

Uzmimo za primjer lutriju gdje treba pogoditi 6 brojeva od njih 45. Program radi na sljedećem principu: da bi se osvojio zgoditak potrebno je pogoditi prvi i drugi i treći i četvrti i peti i šesti broj. U statistici "i" obično znači da te sve šanse treba pomnožiti, pa tako u ovom slučaju šansa za pogađanje prvog broja je 1 u 45, matematički to je $\frac{1}{45}$. Tako je za drugi broj šansa 1 od 44, odnosno $\frac{1}{44}$. Napravimo to i za sve ostale brojeve i dobijemo račun: $\frac{1}{45} \cdot \frac{1}{44} \cdot \frac{1}{43} \cdot \frac{1}{42} \cdot \frac{1}{41} \cdot \frac{1}{40} = \frac{1}{5864443200}$

Dobivanje ovog broja d napravljeno je pomoću sljedeće petlje:

```
a = koliko brojeva treba pogoditi
b = koliko je ukupno brojeva
d=1
for(j=1;j<=a;j++){
d=d*;
b--;
}
```

No, moramo uzeti u obzir da se dobitna kombinacija može uzeti u bilo kojem redoslijedu, znači da se šanse poboljšavaju za broj različitih kombinacija u koje se u ovom primjeru 6 brojeva može napisati. Dakle to je $6! = 720$. Sad podijelimo prvotno dobiveni broj sa brojem kombinacija $\frac{5864443200}{720} = 8145060$. Možemo vidjeti da su nam šanse da pogodimo 6 brojeva od 45 jednake 1 u 8145060

Izračun faktorijele f također je napravljen pomoću petlje i kasnije izračunata konačna vjerojatnost r

```
 f = 1;
  for(i=1;i<=a;i++){
    f = f*i;
  }

r=d/f;
  document.getElementById("rezultat").value= r;
```

6. Zaključak

U ovom radu dane su definicije i opisane vrste permutacija i kombinacija koje postoje, te su dani razni jednostavniji i svakodnevni primjeri kako se koriste permutacijske i kombinacijske formule. Nadalje, navedena su različita područja primjene permutacija i kombinacija, a to su konkretno područja komunikacijskih mreža, kriptografije, mrežne sigurnosti, arhitekture računala, jezika, analize uzoraka, znanstvenih otkrića, baza podataka i rudarenja podataka, operativne optimizacije i računalne sigurnosti. Neka područja istražena su dublje i detaljnije, a neka malo manje s obzirom na to koliko dostupnih informacija o tim područjima je dostupno, te s obizirom na to koliko je široka i zastupljena upotreba permutacija i kombinacije u nekim područjima. Neka područja bilo je moguće dublje i pomnije istražiti, no ona su pak toliko široka da se može napraviti zaseban rad o njima, pa sam prikazao dijelove za koje sam smatrao da su najbitniji i najviše ovise o permutacijama i kombinacijama. Također je opisana i primjena permutacija i kombinacija u vjerojatnosti, te su prikazani neki primjeri gdje se to koristi. Isprogramiran je i jedan primjer korištenja kombinacija u vjerojatnosti. Na kraju se potvrdila konstatacija iz uvoda da su permutacije i kombinacije jako široko rasprostranjene i da se uistinu pojavljuju u svim sferama naših života. One se koriste od najjednostavnijih problema pa sve do kompleksnih sigurnosnih, znanstvenih i ekonomskih zahtjeva.

Popis literature

[1] M.Bona, Combinatorics of permutations. Chapman and Hall/CRC

[2] Črnjarić-Žic N."Numerička i stohastička matematika", nastavni materijali na kolegiju Numerička i stohastička matematika, Tehnički fakultet, Rijeka

[3] I.Hartmann, Primjena Pascalovog trokuta u nastavi matematike, Sveučilište J.J.Strossmayera u osijeku, odjel za matematiku, Osijek, 2010.

[4] Z.Hećimović, Metapodaci, Fakultet građevinarstva, arhitekture i geodezije, katedra za geodeziju i geoinformatiku, Split, 2016

[5] T.Munakata, Područja primjene kombinatorike, posebice permutacija i kombinacija (eng. Application Areas of Combinatorics, Especially Permutations and Combinations), 2005.

[6] B. Nikolić, Statistika, nastavni materijali na kolegiju Statistika, Sveučilište u Zagrebu, Edukacijsko rehabilitacijski fakultet, Zagreb, 2015.

[7] NISO (2004): Understanding Metadata. Bethesda, NISO Press

[8] S. Pal, “Mining educational data using classification to decrease dropoutrate of students,”arXiv preprint arXiv:1206.3078, 2012

[9] D. Veljan, Kombinatorna i diskretna matematika, Algoritam, Zagreb, 2001.

[10] `https://www.sas.com/en_us/insights/analytics/data-mining.html#dmhistory` <3.6.2019.>

[11] `https://www.fidelissecurity.com/threatgeek/threat-intelligence/whole-brain-technology/`, <18.5.2019.>

[12] `http://www.expertsmind.com/questions/butterfly-permutation-30138485.aspx`, <24.6.2019.>

[13] `https://www.rightpoint.com/thought/2011/11/08/data-mining-in-banks-and-financial-institutions`, <25.6.2019.>

[14] `https://www.researchgate.net/publication/304808426_Data_Mining_in_Education`, <26.6.2019.>

[15] `https://www.kdnuggets.com/2019/02/top-10-data-science-use-cases-telecom.html`, <26.6.2019.>

[16] https://www.researchgate.net/publication/325675331_Musical_Combinatorics_Tonnetz_and_the_CubeHarmonic, <7.7.2019.>

Popis slika

Popis tablica