Contents

[1. Opis Markovljevog lanca 3](#_Toc168849341)

[2. Implementacija generatora 6](#_Toc168849342)

[3. Usporedba stvarnih i generiranih podataka 10](#_Toc168849343)

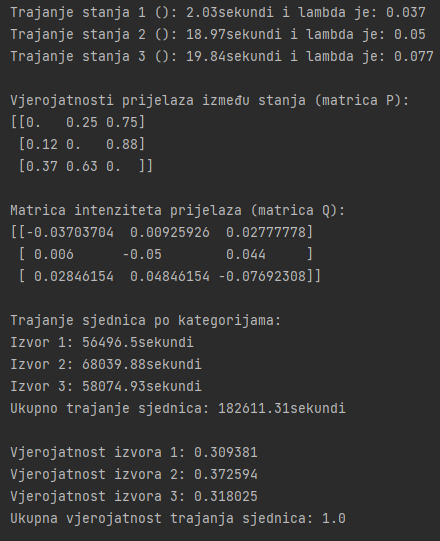
[3.1. Usporedba distribucija parametara 10](#_Toc168849344)

[3.2. Kolmogorov-Smirnov test 14](#_Toc168849345)

[3.3. Teoretska i empirijska vjerojatnost 15](#_Toc168849346)

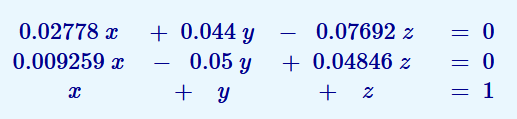
## Opis Markovljevog lanca

Pri kreiranju Markovljevog lanca koristile su se matrice vjerojatnosti prijelaza između stanja (matrica P) i matrica intenziteta prijelaza (matrica Q). One su proizvoljno popunjene vjerojatnostima i vrijednostima trajanja sesije kako je zadano u samom zadatku. Proizvoljno zadana prosječna vremena za naša 3 izvora su: online igranje igrice je u prosjeku 27 sekundi, strujanje zvuka je u prosjeku 20 sekundi i strujanje videa je u prosjeku trajalo 13 sekundi. Matrice P i Q prikazane su na slici (Slika 1).



Slika 1 - prikaz matrica prijelaznih vjerojatnosti i matrice gustoće prijelaza

Stacionarne vjerojatnosti dobivene su transponiranjem matrice Q i izmjenom 1 retka matrice QT sa retkom p1+p2+p3=1 (Slika 2) i dobivene stacionarne vjerojatnosti su: p1=0.3078, p2= 0.3697 i p3=0.3226 (Slika 3).

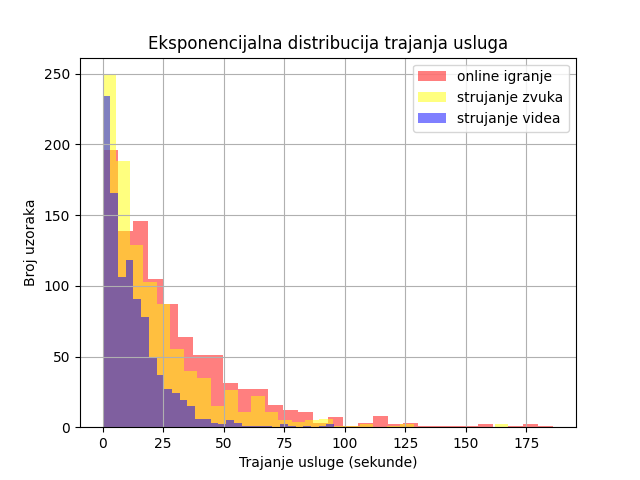


Slika 2 - sustav za računanje stacionarnih vjerojatnosti

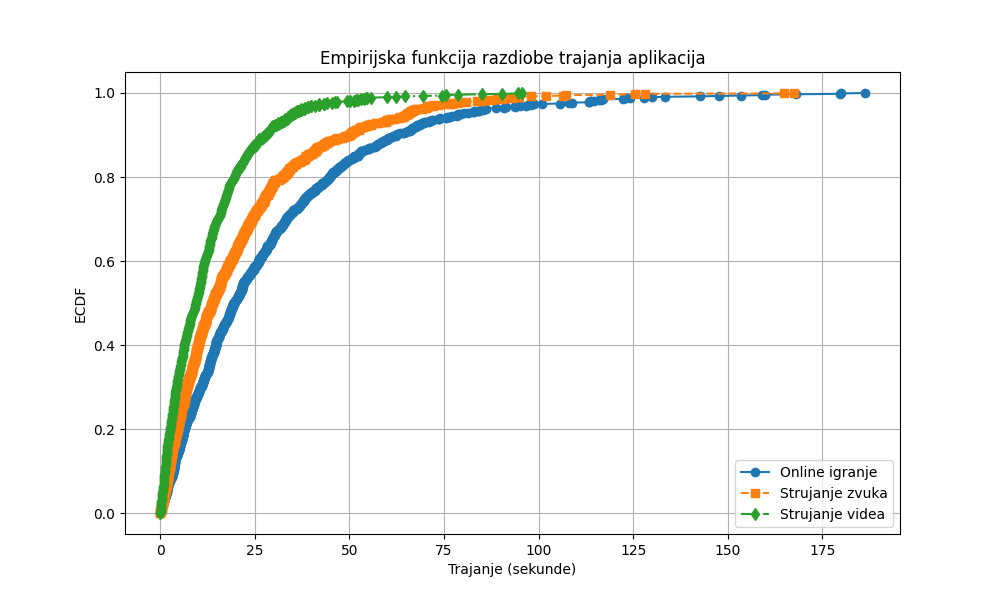


Slika 3 - rezultat stacionarnih vjerojatnosti

Vizualizirani sakupljeni podaci su prikazani na slici 4 (Slika 4) i 5 (Slika 5).

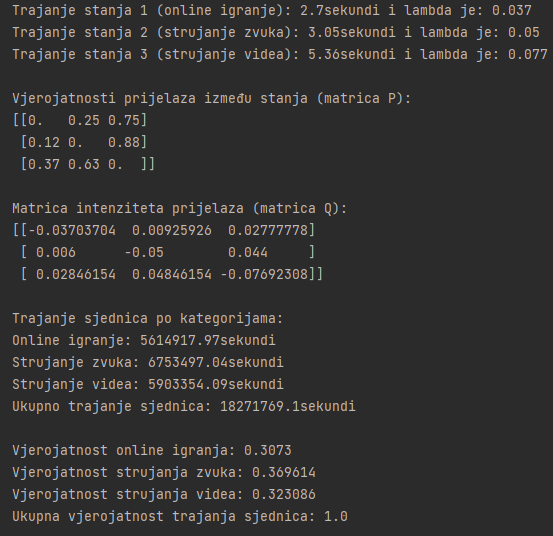


Slika 5 – vizualizacija distribucije trajanja usluga

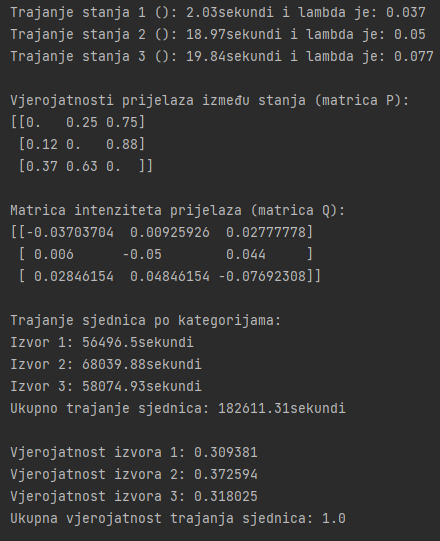


Slika 4 - vizualizirani podaci trajanja korištenjem ECDF

Simulirane vjerojatnosti za 10000 koraka prikazane su na slici (Slika 6). Simulirane vjerojatnosti poklapaju se sa izračunatim stacionarnim vjerojatnostima do druge decimale tj. tek se na trećoj decimali razlikuju vjerojatnosti (osim u slučaju zadnjeg izvora tj. strujanja videa). Vjerojatnosti su puno točnije ako se simulacija odradi za milijun koraka te su rezultati onda jako slični izračunatim matematičkim rezultatima. Rezultati simulacije za milijun koraka su prikazani na slici (Slika 7). Za deset tisuća koraka, kako je i zadano za labos, je preciznost dobivenih simulacijskih vjerojatnosti manja od ove prikazane za milijun koraka na slici (Slika 6). Tako velika preciznost nakon milijun koraka pokazuje podudaranje simulacijskih rezultata sa očekivanim matematičkim rezultatima te samo potvrđuje matematičku točnost izračuna. To samo potvrđuje da se nakon jako puno koraka dolazi do stacionarnih vjerojatnosti pojavljivanja tj. provođenja vremena u određenim stanjima u Markovljevim lancima kao što i matematika „predviđa“.



Slika 7 - vjerojatnosti pojavljivanja pojedine aplikacije iz simulacijskih rezultata za milijun koraka



Slika 6 - vjerojatnosti pojavljivanja pojedine aplikacije iz simulacijskih rezultata za 10000 koraka

## Implementacija generatora

Implementacija funkcionalnosti generatora paketa navedenog u uputama napisana je u tri skripte:

* Receiver.py
* Generator.py
* MarkovChain.py

Zbog ograničenja mreže u kojoj smo obavljali implementaciju (Eduroam), nismo bili u mogućnosti slati pakete preko lokalne mreže, već smo pakete slali na loopback adresu (127.0.0.1).

Skripta Receiver.py implementira jednostavnog primatelja UDP paketa. Sluša na vratima 8080, prima UDP pakete te ispisuje njihovu veličinu na konzolu. Izvorni kod je na slici 1. Također prikazana je obrada najčešćih pogrešaka (opis u komentarima koda).

A screen shot of a computer program

Description automatically generated

Slika 1. Receiver.py skripta

Skripta Generator.py implementira generator prometa. U kratko ova skripta generira pakete različitih veličina i među dolaznih vremena, a sve na osnovu nasumičnih vrijednosti (dobivenih iz odgovarajućih, analizom dobivenih, distribucija).

A screen shot of a computer

Description automatically generated

Slika 2. Generator.py Formiranje i slanje paketa određene veličine

A screen shot of a computer program

Description automatically generated

Slika 3. Generator.py Generiranje paketa za odabranu vrstu usluge

Kao što je prikazano na slici 3, ovisno o vrsti usluge generira se slučajna vrijednost veličine paketa te među dolazno vrijeme (delay) na osnovu odgovarajućih kontinuiranih distribucija (dobivene analizom).  
Nakon što se odrede vrijednosti veličine paketa i vremena uspavljuje se dretva na broj milisekundi definirano u varijabli delay, te se poziva funkcija za formiranje i slanje paketa s dobivenom vrijednosti.

Unutar skripte MarkovChain.py implementiran je automat stanja (Markovljev lanac). Sastoji se od 3 stanja: Agario, Radio i Video. Na slici 4. su prikazane i dodatne pomoćne funkcije.

A screenshot of a computer program

Description automatically generated

Slika 4. MarkovChain.py Definicija stanja, vjerojatnosti prijelaza te pomoćne funkcije

Na slici 5. prikazana je implementacija automata. Počinje u nasumičnom stanju. Zatim za trenutno stanje se izvode sljedeće akcije:

* Računa se vrijeme ostanka u stanju na osnovu eksponencijalne razdiobe
* Za dobiveni broj sekundi se izvodi funkcija do\_action\_for\_seconds koja za određeno stanje izvodi odgovarajuće generiranje paketa
* Na osnovu matrice prijelaza se računa sljedeće stanje

Akcije se izvode u beskonačnoj petlji (do prekida programa).

A screen shot of a computer program

Description automatically generated

Slika 5. Implementacija stanja i akcija

## Usporedba stvarnih i generiranih podataka

### Usporedba distribucija parametara

U zagradi su ispisane vrijednosti koje su dobivene iz mjerenja prometa preko interneta te su uspoređene parametri distribucije na sljedeći način: (Najbolja generirana – Najbolja snimljena s interneta koja su korištena za generator). Sve distribucije od generatora koje su jednake onima s interneta budu vrlo dobro opisane. Može se vidjeti razlika na Agar.io međudolazna vremena distribucijama gdje je od generatora dobivena dweibull, a snimljena normalna s interneta da nisu svi podaci pokriveni krivuljom. Dodatno smo više puta generirali podatke generatorom za jednu vrstu paketa te su najbolje distribucije bile različite u recimo dvije različite Radio generacije podataka generatorom. Također Kolmogorov-Smirnov test je još uvijek izbacivao p-value jednak nula.

**Agar.io**:

* Parametri najbolje distribucije za veličinu paketa (gamma - gamma)
  + Alpha: 0.71 (0.65)
  + Loc: 92 (60)
  + Scale: 61.56 (63.3)

Slika na kojoj se prikazuje tekst, snimka zaslona, dijagram, radnja

Opis je automatski generiran

Sl. 1. Agar.io distribucija za veličinu paketa

Grafovi snimljenog s interneta i generiranog prometa za veličinu paketa su najsličniji kod Agar.io igre. Internet promet generira veće i raspršenije pakete koji idu i iznad krivulje dok su podaci iz generatora djelomično uz krivulju.

* Parametri najbolje distribucije za međudolazna vremena (dweibull - normalna)
  + Alpha: 0.69
  + Loc: 0.03 (18.79)
  + Scale: 0.01 (9.70)

Slika na kojoj se prikazuje tekst, snimka zaslona, dijagram, radnja

Opis je automatski generiran

Sl. 2. Agar.io distribucija za međudolazna vremena

Grafovi se vrlo razlikuju te generator slabo predstavlja normalnu distribuciju. Moguće je da bi uz više generiranih podataka (više od 5000) mogao pratiti normalnu distribuciju.

**Video:**

* Parametri najbolje distribucije za veličinu paketa (gamma - dweibull)
  + Alpha: 0.71 (0.70)
  + Loc: 92 (1434)
  + Scale: 61.56 (488.39)

Slika na kojoj se prikazuje tekst, snimka zaslona, dijagram, radnja

Opis je automatski generiran

Sl. 3. Video distribucija za veličinu paketa

Parametri dweibull distribucije u našem generatoru uzrokuju nastanak gamma distribucije za veličine. Dweibull distribucija za pakete s interneta jako loše pokušava namjestiti se podacima te zato dobivamo drugačiji izgled distribucije generatora.

* Parametri najbolje distribucije za međudolazna vremena (dweibull – expon)
  + Alpha: 0.69
  + Loc: 0.03 (0)
  + Scale: 0.01 (9.74)

Slika na kojoj se prikazuje tekst, snimka zaslona, dijagram, radnja

Opis je automatski generiran

Sl. 4. Video distribucija za međudolazna vremena

Isto tako vrijedi i za međudolazna vremena kod videa.

**Radio**:

* Parametri najbolje distribucije za veličinu paketa (dweibull - dweibull)
  + Alpha: 0.59 (0.48)
  + Loc: 1519 (1488)
  + Scale: 432.21 (150.84)

Slika na kojoj se prikazuje tekst, snimka zaslona, dijagram, radnja

Opis je automatski generiran

Sl. 5. Radio distribucija za veličinu paketa

* Parametri najbolje distribucije za međudolazna vremena (dweibull -dweibull)
  + Alpha: 0.54 (0.49)
  + Loc: 0.22 (204.79)
  + Scale: 0.05 (49.59)

Slika na kojoj se prikazuje tekst, snimka zaslona, dijagram, radnja

Opis je automatski generiran

Sl. 6. Radio distribucija za međudolazna vremena

Ovdje su grafovi distribucije vrlo slični, ali imaju vrlo različite parametre. Vremena paketa na Internet prometu su puno više rašireniji i izvan krivulje.

### Kolmogorov-Smirnov test

Pošto je izrađeno točno 5000 generiranih stanja, također se reducirao broj međudolaznih paketa i veličine paketa nastalim direktnim snimanjem na oko 5000 tisuća podataka. Rezultati su prikazani na sljedećoj slici.

Slika na kojoj se prikazuje tekst, snimka zaslona, Font

Opis je automatski generiran

*Sl. 7.* Kolmogorov-Smirnov test

Razlike su puno veće tijekom u međudolaznim vremenima nego u veličini paketa.KS statistics vrijednost govori koja je razlika između distribucija te što je veća vrijednost (prema 1) to je veća razlika. *p* vrijednost će nam reći jesu li distribucije izvučene iz istog skupa distribucije, manja vrijednost *p* nam daje do znanja da nisu. Pošto je većina *p* vrijednosti 0, to može označavati da je došlo do velike pogreške prilikom traženja najbolje distribucije koja opisuje podatke koji su nastale snimanjem s interneta i još kumulativno se dodaju pogreške nastale pri određivanju distribucija generatora. Postoje također odstupanja u varijablama loc, scale i alpha kod generiranih distribucija što može ukazati na potreban puno veći skup podataka. Najbližu distribuciju u našem primjeru imaju veličine paketa tijekom snimanja i generiranja igre Agar.io.

### Teoretska i empirijska vjerojatnost

Slika na kojoj se prikazuje tekst, snimka zaslona, Font

Opis je automatski generiran

Sl. 8. Empirijska vjerojatnost za internet i generirani promet

Teoretske vjerojatnosti su u oba slučaja jednake pošto se koristi ista P matrica koja se množi s lambda. Empirijske vjerojatnosti su imaju približno jednake vrijednosti. Stanje 0 predstavlja Agar.io (lambda = 1/27), stanje 1 radio (lambda = 1/20) i stanje 2 (lambda = 1/13), video. U oba slučaja najkraće će se ostati u stanju gledanja videa, dok će se igranje igrice i slušanja radija duže odvijati. Korištenjem većeg broja stanja, empirijske vjerojatnosti bi se trebale približavati teoretskim vrijednostima.