Dinamika širenja zaraznih bolesti na kompleksnim mrežama

Sigmund, Margareta

Master's thesis / Diplomski rad

2021

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:217:782646

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-07-08**

Repository / Repozitorij:

Repository of the Faculty of Science - University of Zagreb

SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO-MATEMATIČKI FAKULTET
FIZIČKI ODSJEK

Margareta Sigmund

DINAMIKA ŠIRENJA ZARAZNIH BOLESTI NA KOMPLEKSNIM MREŽAMA

Diplomski rad

Zagreb, 2021.

SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO-MATEMATIČKI FAKULTET
FIZIČKI ODSJEK

INTEGRIRANI PREDDIPLOMSKI I DIPLOMSKI SVEUČILIŠNI STUDIJ
FIZIKA; SMJER ISTRAŽIVAČKI

Margareta Sigmund

Diplomski rad

Dinamika širenja zaraznih bolesti na kompleksnim mrežama

Voditelj diplomskog rada: prof. dr. sc., Hrvoje Štefančić

Ocjena diplomskog rada: _____

Povjerenstvo: 1. _____

2. _____

3. _____

Datum polaganja: _____

Zagreb, 2021.

Najprije, zahvalna sam dragome Bogu koji me je zaželio i stvorio i koji je u mene usadio ljubav i divljenje prema svemu stvorenom. Zahvalna sam mojim roditeljima što su me sa svom ljubavlju i brižnošću odgajali i poticali moju znatiželju. Zahvalna sam svojoj braći koji su sa mnom odrastali i sa mnom se veselili i glupirali. Od najranijeg djetinjstva dane su mi ispunjali brojni prijatelji te sam i njima zahvalna jer su me obogaćivali svojom prisutnošću i svojim primjerima, posebice oni od mene stariji. Zahvalna sam i svim učiteljima i nastavnicima koji su me pratili i poučavali i malo pomalo pomogli mi otkriti ljubav prema fizici. Svaki dan studiranja na ovome fakultetu bio je izazov i radost. Od srca zahvaljujem svim profesorima i asistentima što su se trudili, najbolje što znaju, prenijeti mi svoje znanje. Posebice sam zahvalna svim profesorima koji su od sebe davali uvijek ono nešto malo više kako bi nas poticali i savjetovali. Zahvalna sam svim svojim kolegama od kojih i s kojima sam učila. Na kraju, ali nipošto manje važno, od srca sam zahvalna mom mentoru prof.dr.sc. Hrvoju Štefančiću koji me je s velikom strpljivošću usmjeravao i vodio tijekom istraživanja i izrade ovog diplomskog rada.

Sažetak

Pojavom pandemije virusa COVID-19, tema širenja zaraze postala je vrlo aktualna te dio svakodnevice prosječnog laika. Modeliranje širenja zaraze već je i ranije izazivalo velik interes raznih znastvenih zajednica te je postalo područje multidisciplinarnog istraživanja budući da je zaintrigiralo matematičare, fizičare, biologe, sociologe i druge znastvenike. U ovome ćemo radu sažeti najvažnije doprinose epidemiološkom modeliranju, općenita svostva kompleksnih mreža te najznačajnije doprinose u problemu detekciji izvora zaraze na kompleksnim mreža. Također, razmotrit ćemo karakteristike prosječne udaljenosti promatranog vrha do svih zaraženih vrhova i njenog gradijenta na primjeru SIS modelu širenja zaraze na mreži malog svijeta. Dokazat ćemo da je gradijent prosječne udaljenosti promatranog vrha do svih zaraženih vrhova najveći za izvor zaraze, što ćemo provjeriti i na skupu podataka koji opisuju neku stvarnu mrežu. SIS model uvelike svojim karakteristikama podsjeća na dinamiku širenja koronavirusa (oporavljeni pacijent nakon nekog vremena ponovno postaje podložan novoj zarazi). Modeli homogenog miješanja tek aproksimativno opisuju širenje zaraze, stoga se u ovom radu promatra širenje zaraze SIS modelom na kompleksnim mrežama, konkretno na mrežama malog svijeta. Posebnost kompleksnih mreža jest vjeran prikaz raspodjele ljudi te njihovih interakcija, gdje su ljudi vrhovi grafa, a interakcije su veze među vrhovima. Ovaj je rad usmjeren na problem detekcije izvora zaraze pri širenju zaraze po SIS modelu na mreži malog svijeta.

Ključne riječi: epidemiološki modeli, SIS model širenja zaraze, kompleksne mreže, problem detekcije izvora širenja zaraze

Dynamics of contagious disease spreading on complex networks

Abstract

With the appearance of the COVID-19 virus pandemic, the topic of the spread of the infection became very usual and part of the everyday life of an average man. Modeling the spread of the infection has already aroused great interest in various scientific communities and became an area of multidisciplinary research since it has intrigued mathematicians, physicists, biologists, sociologists, and other scientists. In this thesis, we summarize the most important contributions to epidemiological modeling, the general properties of complex networks, and the most significant contributions to the problem of detecting sources of infection on complex networks. Also, we will consider the characteristics of the average distance of the observed node to all infected ones and its gradient on the example of the SIS model infection spreading on the small-world network. We will prove that the gradient of the average distance of the observed node to all infected ones is the largest for the source of infection, which we will also check on a set of data describing a real network. The SIS model is very similar in its characteristics to the dynamics of coronavirus spread (a recovered patient after some time becomes susceptible to a new infection again). Homogeneous mixing models only approximately describe the disease spread, so in this paper, we observe the SIS model spreading of infection on complex networks, specifically on small-world networks. The peculiarity of complex networks is a possibility of correct representation of the distribution of people and their interactions, where people are the vertices of the graph, and the interactions are the connections between the vertices. This paper deals with the source detection problem in SIS model disease spreading on small-world networks.

Keywords: edipemiological models, complex networks, SIS model of disease spreading, source detection problem

Sadržaj

1 Uvod **1**

2 Epidemiološko modeliranje **3**

2.1 Uvod 3

2.2 Modeli odjeljaka u epidemiologiji 4

2.2.1 SIS i SIR modeli 5

3 Kompleksne mreže **10**

3.1 Osnovni pojmovi 11

3.2 Mreža malog svijeta 15

3.3 Mreže bez skale 17

3.4 Primjena kompleksnih mreža u istraživanjima u fizici 19

4 Širenje zaraze na kompleksnim mrežama **21**

4.1 Širenje zaraze na mreži malog svijeta 21

4.2 Identifikacija izvora zaraze 23

5 Implementacija širenja zaraze na mreži malog svijeta **29**

5.1 Provjera valjanosti algoritma na lancu 30

5.1.1 Usmjereni lanac 30

5.1.2 Neusmjereni lanac 33

5.1.3 Apsolutna razlika analitičkih rješenja i rješenja dobivenih algoritmom 34

5.2 Kvalitativna svojstva veličine R stanju na mreži malog svijeta 34

5.2.1 Utjecaj promjene broja najbližih susjeda k na veličinu $R_i(t)$ . . 38

5.2.2 Utjecaj promjene vjerojatnosti premrežavanja p na veličinu $R_i(t)$ 41

5.2.3 Utjecaj promjene vjerojatnosti prijenosa zaraze između dva susjeda μ i vjerojatnosti oporavka ν zaraženog vrha na veličinu $R_i(t)$ 44

5.3 Kvalitativna svojstva ∇R na mreži malog svijeta 48

5.3.1 Utjecaj promjene broja najbližih susjeda k na $\nabla R_i(t)$ 48

5.3.2 Utjecaj promjene vjerojatnosti premrežavanja p mreže malog svijeta na $\nabla R_i(t)$ 50

5.3.3 Utjecaj parametara širenja zaraze λ na $\nabla R_i(t)$ 52
5.4 Ponašanje ∇R na živčanom sustavu oblića Caenorhabditis elegans . . . 56
5.5 Ponašanje ∇R na Zacharyjevoj mreži karate kluba 59
5.6 Zaključak rezultata 61

6 Zaključak **63**

Dodaci **65**

A Python kod algoritamske implementacije širenja zaraze na kompleksnim mrežama **65**
A.1 Paketi korišteni u simulacijama 65
A.2 Funkcija za generiranje mreže malog svijeta i simulaciju širenja zaraze na toj mreži 65
A.3 Simulacija SIS modela na mreži živčanog sustava oblića Caenorhabditis elegans 67
A.3.1 Generiranje mreže iz skupa podataka o mreži živčanog sustava oblića Caenorhabditis elegans 67
A.3.2 Simulacija širenja živčanih podražaja na živčanoj mreži C. elegans 68

Literatura **70**

1 Uvod

Zarazne bolesti obilježile su važne prekretnice u povijesti čovječanstva, pa je i njihovo proučavanje pristuno na razne načine kroz povijest. Budući da pojava epidemije neke zarazne bolesti snažno utječe na svakog čovjeka, neovisno o zanimanju, u istraživanje, modeliranje i opisivanje zaraznih bolesti uključeni su razni znanstvenici poput liječnika, biologa, kemičara, matematičara i sociologa. Epidemiologija je znanstvena metoda čiji je cij pronaći raspodjelu i uzorke stanja i događa povezanih sa zdravljem populacije (ne samo zaraznih bolesti) što omogućava predviđanje daljnjih ishoda. Razmatra se kolektivna dinamika zaraze za pojedince i zajednicu kao cjelinu. Značajna je primjena matematičkih modela i koncepata u epidemiologiji jer omogućuje predviđanje jasnih kvantitativnih ishoda širenja zaraze.

Početkom matematičkog modeliranja epidemiologije smatra se rad Daniela Bernoullija iz 18. stoljeća o utjecaju cijepljenja protiv boginja na duljinu životnog vijeka. Daljnje modeliranje širenja zaraznih bolesti razvija se primjenom modela koji su već ranije susretani u objašnjavanju nekih fizikalnih procesa poput zakona o djelovanju masa, raspodjela populacije u prepoznatljive odjeljke s obzirom na stanje zaraze (podložan zarazi S (eng. *susceptible*), zaražen I (eng. *infected*), oporavljen R (eng. *recovered*)i sl.), homogeno miješanje te modeliranje dinamike zaraze Markovljevim procesima. Prema nazivima odjeljaka u modelu razlikujemo SI, SIR, SIS, $SIRS$ i druge modele, ali u ovome radu posebno ćemo promotriti širenje zaraze prema SIS modelu. Primjena modela odjeljaka uz pretpostavku homogenog miješanja jedinki u populaciji daje aproksimativno točne rezultate te dosta dobar kvalitativan opis dinamike zaraze, no nemoguće je kvantitativno točno predvidjeti daljnji razvoj zaraze ili pak odrediti odakle se zaraza počela širiti. U posljednje vrijeme upotrebljavaju se modeli temeljeni na jedinkama (eng. *agent-based model*, ABM) koji su se pokazali vrlo točnima u kvantitativnom predviđanju izbijanja epidemije. Također, u zadnjim desetljećima razvila se primjena kompleksnih mreža u praćenju dinamike širenja zaraznih bolesti, čiji je napredak uvelike ubrzan razvitkom računala. Svaki vrh u grafu kompleksne mreže predstavalja jedinku unutar populacije koja može poprimiti obilježje stanja zaraze u kojem se nalazi. Cijela kompleksna mreža predstavlja sustav koji čini promatrana populacija. Veze među vrhovima predstavljaju intrakcije među pojedincima. Posebnu pozornost među kompleksnim mrežama zadobili su modeli

mreže malog svijeta i mreže bez skale koji poprilično dobro opisuju stvarne mreže poput mreže odnosa među glumcima, električne mreže, prometne mreže ili pak mreže živčanog sustava.

Budući da analitička rješenja za širenje zaraze na kompleksnoj mreži postaju iznimno teško rješiva ili gotovo nerješiva, u analizi širenja zaraze na kompleksnoj mreži koriste se Monte Carlo simulacije. Čak i Monte Carlo simulacije ubrzo postaju računalno vrlo zahtijevne, pa se koriste određene aproksimativne metode koje olakšavaju račun, ali daju kvantitativno točna rješenja. U primjeni kompleksnih mreža za razmatranje širenja zaraznih bolesti čest je problem detekcije izvora zaraze. Otkrivanjem izvorišta zaraze moguće je zaustaviti ili ograničiti širenje zaraze ili spriječiti širenje u nekim budućim sličnim situacijama. SIS model širenja zaraze nije primjenjiv isključivo za širenje zaraze već može biti primjenjen i na širenje informacija, glasina, podražaja ili signala u živčanoj ili električnoj mreži ili pak širenje računalnih virusa.

Cilj ovoga rada jest pomoću Monte Carlo simulacija širenja zaraze po SIS modelu na mreži malog svijeta pronaći karakteristiku po kojoj se ističe izvor zaraze od ostalih vrhova te provjeriti vrijedi li isti zaključak na nekoj stvarnoj mreži. U ovom radu su razmotrena dva primjera stvarnih mreža: mreža živčanog sustava oblića C. elegans te Zacharyjeva mreža karate kluba.

2 Epidemiološko modeliranje

2.1 Uvod

Prenosive bolesti oduvijek su bile važan dio ljudske povijesti poput kolere ili španjolske gripe. Također, postoje bolesti koje su postale endemske (stalno prisutne) te i danas uzrokuju veliku smrtnost u nekim krajevima svijeta. Cilj epidemiologije jest najprije razumijeti uzrok i mehanizam širenja zaraze, zatim predvidjeti njen razvitak te pronaći način da se širenje zaraze obuzda ili u potpunosti suzbije. Prvi korak u epidemiologiji jest prikupljanje i obrada podataka, što je razvitak računala uvelike olakšao.

Proučavanje zaraznih bolesti započeo je John Graunt objavivši 1662. godine svoju knjigu "Prirodna i politička promatranja prema Računima smrtnosti" (eng. *"Natural and Political Observations made upon the Bills of Mortality"*) [1]. Računi smrtnosti (eng. *The Bills of Mortality*) su bila tjedna izvješća o broju i uzroku smrti u londonskim župama koja su redovito bilježena od 1592. godine. Graunt je komparativno bilježio različite uzroke smrti i tako zadao osnove teorije kompetitivnih rizika (teorija da smrtnost kod ljudi uzrokuju različiti rizici koji se međusobno nadmeću).

Početak matematičkog modeliranja epidemiologije predstavlja rad Daniela Bernoullija o utjecaju cijepljenja protiv boginja na duljinu životnog vijeka ljudi. Naime, budući da su boginje u 18. stoljeću bile endemske, vodila se žestoka rasprava o korisnosti cijepljenja te je Bernoulli predvodio istraživanje o koristima cijepljenja. Računao je prosječni životni vijek sa i bez cijepiva te tako proučio utjecaj cjepiva na duljinu života. Računi su pokazali da cijepljenje povećava prosječan životni vijek s 26 godina i 7 mjeseci na 29 godina i 9 mjeseci. Svoj je rad opisao u kratkom pregledu 1760. godine, a detaljnije izložio 1766. godine.

Istraživanjem prostorno-vremenske raspodjele oboljelih od kolere u Londonu 1855. godine, John Snow dao je vrijedan doprinos razumijevanju zaraznih bolesti. Naime, uspio je pokazati da je crpka za vodu u Broad Streetu u Londonu izvorište zaraze.

Početkom 20. stoljeća, William Hamer i Ronald Ross primijenili su zakon o djelovanju masa[1] kako bi opisali širenje epidemije.

Dvadesetih godina 20. stoljeća uočila se potreba za modelom odjeljaka (eng. *com-*

[1]Brzina prirasta koncentracije produkta u kemijskoj reakciji proporcionalna je umnošku koncentracija tvari koje ulaze u reakciju

partmental model), gdje se odjeljci odnose na skupine populacije koje su nezaražene, zaražene ili oporavljene od zaraze. Kermack–McKendrickov epidemiološki model (1927. godine) i Reed–Frostov epidemiološki model (1928. godine) opisuju odnos između nezaraženih jedinki podložnih zarazi (eng. *susceptible*), zaraženih jedinki (eng. *infected*) te imunih/oporavljenih jedinki u populaciji te uspješno predviđa razvoj epidemija baš kao što je uočeno u mnogim zabilježenim slučajevima epidemija.

U posljednje vrijeme upotrebljavaju se modeli temeljeni na jedinkama (eng. *agent-based model*, ABM) koji su se pokazali vrlo točnima u kvantitativnom predviđanju izbijanja epidemije. Ovaj je model prostorno strukturiran te uzima u obzir diskretnu narav jedinki te njihovu mobilnost i stohastičnost interakcija s drugim jedinkama.

2.2 Modeli odjeljaka u epidemiologiji

Matematičko modeliranje u epidemiologiji razvilo se u značajno područje u kojem se isprepliću istraživanja iz raznih područja znanosti kao i razne znanstvene discipline. Svakako jedan od značajnih i jednostavnijih modela širenja zaraze jest model odjeljaka. Osnovni princip modela odjeljaka jest da se ukupna populacija podijeli u odjeljke koji predstavljaju skupine ljudi koje su u nekom stadiju razvitka zaraze, tako primjerice možemo razlikovati odjeljak pojedinaca podložnih zarazi (eng. *susceptible*, oznaka S), zaraženih pojedinaca (eng. *infected*, oznaka: I) te oporavljenih pojedinaca koji su razvili imunitet (eng. *recovered*, oznaka: R). Pojedinci prelaze iz odjeljka u odjeljak ovisno kako se zaraza razvija. Primjerice, nezaraženi pojedinac (S) nakon izloženosti zarazi postaje zaražen (I) te se spontano, nakon određenog vremena oporavi (R). Postoji više različitih modela odjeljaka koje razlikujemo po nazivima odjeljaka u koje se može razvrstati pojedince u populaciji. Tako primjerice razlikujemo SIS model gdje se pojedinac može oporaviti od zaraze, ali tako da ne razvije imunitet nego ponovno postaje podložan zarazi (takva je primjerice prehlada); SIR model gdje je pojedinca oporavljenog od zaraze nemoguće ponovno zaraziti jer je razvio imunitet (primjerice vodene kozice); $SEIR$ model koji je proširen odjeljkom pojedinaca izloženih zarazi (eng. *exposed*, oznaka: E) ili $SIRS$ model gdje je moguće jednom stečeni imunitet ponovno izgubiti i postati podložan zarazi te ostali slični modeli.

Razvitak takvih modela započinje u dvadesetom stoljeću radovima Rolanda Ro-

ssa i Hilde Hudson o primjeni teorije vjerojatnosti u epidemiologiji (*"An application of the theory of probabilities to the study of a priori pathometry*[2]*.—Part I,II, III"*) iz 1916. i 1917., radovima Kermacka i McKendricka o matematičkoj teoriji u epidemiji iz 1927. (*"A contribution to the mathematical theory of epidemics"*) te radom Kendalla o determinističkoj i stohastičkoj epidemiji u zatvorenim populacijama (*"Deterministic and stochastic epidemics in closed populations"*) 1956. godine.

Razvoj epidemije u ovim modelima opisan je diferencijalnim jednadžbama u koje može biti dodan i stohastički doprinos, što bolje opisuje realnu situaciju. Ovim je modelima moguće predvidjeti promjenu broja zaraženih kao i predvidjeti trajanje i razvoj epidemije u populaciji.

Slika 2.1: Shematski prikaz evolucije zaraze između odjeljaka za SIR, SIS, SIRS i SEIR modele.

2.2.1 SIS i SIR modeli

Pretpostavka za najjednostavniji princip evolucije zaraze u populaciji jest da je broj ljudi u populaciji stalan i jednak N. Zanemaruje se sve moguće demografske promjene (smrt, porod, migracije, itd.). Kako bi se u populaciji počela širiti zaraza, potrebno je da se nađe barem jedna zaražena jedinka. Prijenos zaraze događa se prilikom kontakta zaraženog (I) i nezaraženog (S) pojedinca. Budući da je interakcija zaraženog i nezaraženog pojedinca nužna za prijenos zaraze, prostorna raspodjela zaraženih i nezaraženih pojedinaca može imati značajnu ulogu u dinamici širenja

[2]Arhaičan naziv za epidemiologiju

zaraze, što se u ovom slučaju ne uzima u obzir. Vjerojatnost da se zaraza prenese sa zaraženog na nezaraženog pojedinca jednaka je μ. Radi jednostavnosti, pretpostavlja se da je μ konstantan (slika 2.1). Nakon što je pojedinac zaražen (I), spontano i neovisno o kontaktu s drugima, zaraženi se pojedinac u svakom vremenskom koraku može oporaviti s vjerojatnošću ν. I za tu se vjerojatnost, radi jednostavnosti, pretpostavlja da je konstantna. Nakon što se zaraženi pojedinac oporavi, on ponovno postaje podložan novoj zarazi i proces se nastavlja te je u nekim uvjetima moguće da nikad ne prestaje. Za razliku od SIS modela, u SIR modelu, nakon što se zaraženi pojedinac oporavio, nije ga moguće ponovno zaraziti (R). U SIR modelu se proces uvijek mora zaustaviti jer ponestane zaraženih jedinki. SIR modelu se mogu dodati koeficijenti umiranja i rađanja jedinki što omogućuje daljnju evoluciju zaraze. Proširenje SIR modela jest $SIRS$ model gdje oporavljeni pojedinac s imunitetom (R), spontano, u svakom vremenskom koraku može izgubiti imunitet s vjerojatnošću β te ponovno postati podložan novoj zarazi. Drugo moguće proširenje SIR modela jest da se između stanja nezaraženosti (S) i zaraženosti (I) dodaje stanje izloženosti (E) gdje je pojedinac zaražen, ali još ne može prenositi zarazu.

SIS i SIR [2] modeli predstavljaju osnovnu klasifikaciju epidemioloških modela s obzirom na evoluciju za vrlo velika vremenska razdoblja. Kao što se može uočiti na slici 2.2, SIS model je karakterističan po *endemskom*, stacionarnom stanju, nakon dovoljnog vremena trajanja zaraze. U stacionarnom je stanju broj zaraženih pojedinaca gotovo konstantan. U SIR modelu, broj zaraženih uvijek teži u nulu. I u SIS i u SIR modelu procesi zaraze i oporavka određuju evoluciju zaraze.

Vremena trajanja zaraze na pojedinacu raspodijeljena su oko srednje vrijednosti koja je dobro definirana vjerojatnošću oporavka ν. U slučaju da vrijeme zamišljamo kao diskretan niz vremenskih koraka, u svakom vremenskom koraku vjerojatnost da se pojedinac oporavi jest ν, tada će prosječno vrijeme za koje je pojedinac zaražen biti jednako ν^{-1}. U limesu kontinuiranog toka vremena, proces zaraze dobro je opisan Poissonovim procesom [3] gdje je ν stopa prijelaza (vjerojatnost po jedinici vremena) iz zaraženog u nezaraženo stanje, a vjerojatnost da pojedinac ostaje zaražen za vrijeme τ prati eksponencijalnu raspodjelu $P_{inf}(\tau) = \nu e^{-\nu\tau}$ s prosječnim trajanjem zaraze $\langle\tau\rangle = \nu^{-1}$.

Pretpostavka da vrijeme trajanja zaraze prati Poissonovu raspodjelu prirodno nas navodi na opis epidemije pomoću Markovljevih procesa [4].

Slika 2.2: Grafovi vremenske evolucije zaraze za SIS i SIR modele.

Proces prijenosa zaraze znatno je kompliciraniji od procesa oporavka budući da ovisi o interakcijama zaraženih i nezaraženih pojedinaca. U nedostatku detaljnih podataka o međuljuskim interakcijama, najčešća pretpostavka jest aproksimacija homogenog miješanja [5] koja pretpostavlja da sve jedinke u populaciji interagiraju nasumično jedna s drugom. Prema toj pretpostavci, što je veći broj zaraženih pojedinaca, veća je vjerojatnost prijenosa zaraze. Iz svega izrečenog, moguće je definrati *silu zaraze* (eng. *force of infection*) α koja označava vjerojatnost da će se nezaraženi pojedinac podložan zarazi zaraziti u jednom vremenskom koraku. U limesu kontinuiranog toka vremena α se definira na sljedeći način:

$$\alpha = \bar{\beta}\frac{N^I}{N} \tag{2.1}$$

gdje $\bar{\beta}$ ovisi o svojstvima zaraze kao i o interakcijama u populaciji, a N^I je broj zarzaženih jedinki u populaciji. α je proporcionalna udjelu zaraženih jedinki u populaciji $\rho^I = N^I/N$. Ponekad se $\bar{\beta}$ razdvaja na dva faktora βk gdje je β stopa zaraze

po kontaktu, a k je broj kontakata s drugim jedinkama u populaciji. Takav prikaz sile zaraze odgovara zakonu djelovanja mase te pronalazi široku upotrebu u opisima dinamičkih procesa s aproksimacijom srednjeg polja [6].

Prema prethodnim postavkama, epidemiju je moguće zamisliti kao stohastički *reakcijsko-difuzijski* proces. Jedinke se mogu zamisliti kao "čestice" koje pripadaju različitim odjeljcima te evoluiraju ovisno o međusobnim interakcijama uz različite stope prelaska iz odjeljka u odjeljak, stoga nije rijetkost da se ti prijelazi prikazuju stehiometrijskim jednadžbama. Primjerice, za SIS model vrijede sljedeće jednadžbe:

$$\begin{aligned} S + I &\xrightarrow{\mu} 2I \\ I &\xrightarrow{\nu} S \end{aligned} \tag{2.2}$$

gdje su μ i ν stope prijelaza za zarazu i oporavak. Slične jednadžbe vrijede za SIR model:

$$\begin{aligned} S + I &\xrightarrow{\mu} 2I \\ I &\xrightarrow{\nu} R \end{aligned} \tag{2.3}$$

U limesu kontinuiranog vremena i uzimajući u obzir pretpostavku o homogenom miješanju i zakon djelovanja masa, korisno je zapisati diferencijalne jednadžbe koje opisuju promjenu ukupnog broja zaraženih i broja nezaraženih u vremenu. Za SIS model vrijede sljedeće difenercijalne jednadžbe (po uzoru na diferencijalne jednadžbe za SIR model [6]):

$$\frac{dN^S}{dt} = -\mu \frac{N^S N^I}{N} + \nu N^I \tag{2.4}$$

$$\frac{dN^I}{dt} = \mu \frac{N^S N^I}{N} - \nu N^I \tag{2.5}$$

gdje su N^I i N^S broj zaraženih i nezaraženih pojedinaca u populaciji, a μ i ν su stope

zaraze i oporavka. Za N^I i N^S vrijedi i sljedeća jednadžba:

$$N^I + N^S = N \tag{2.6}$$

gdje je N ukupan broj jedinki u populaciji. Prethodna jednadžba vrijedi uz pretpostavku da se ukupan broj jedinki u populaciji ne mijenja.

Za SIS model, uvrštavajući jednadžbu (2.6) u diferencijalne jednadžbe, moguće ih je analitički riješiti, pa se za ukupan broj nezaraženih jedinki dobiva sljedeće rješenje:

$$N^S(t) = N\frac{\nu}{\mu} + N\frac{\nu-\mu}{\mu}\frac{1}{Ce^{(\mu-\nu)t}-1} \tag{2.7}$$

gdje je

$$C = \frac{\mu}{N(\nu-\mu)+\mu} \tag{2.8}$$

Budući da vrijedi jednadžba (2.6), za broj zaraženih u ovisnosti o vremenu dobivamo sljedeće rješenje:

$$N^I(t) = N\Big(1 - \frac{\nu}{\mu} - \frac{\nu-\mu}{\mu}\frac{N(\nu-\mu)+\mu}{\mu e^{(\mu-\nu)t} - N(\nu-\mu)-\mu}\Big) \tag{2.9}$$

Zanimljivo je proučiti limese kada vrijeme ide u beskonačnost $t \to \infty$. Za svaki $N^I(t=0) > 0$ vrijedi:

$$\frac{\mu}{\nu} \leq 1 \Rightarrow \lim_{t\to+\infty} N^I(t) = 0 \tag{2.10}$$

$$\frac{\mu}{\nu} > 1 \Rightarrow \lim_{t\to+\infty} N^I(t) = N\Big(1 - \frac{\nu}{\mu}\Big) \tag{2.11}$$

Uočavamo da zaraza iščezava za $\lambda \equiv \frac{\mu}{\nu} \leq 1$, a prelazi u endemsko stacionarno stanje za $\lambda \equiv \frac{\mu}{\nu} > 1$ za $t \to \infty$. Zbog takve važnosti veličine broja λ za evoluciju zaraze, taj se broj naziva *osnovni reprodukcijski broj* (eng. *basic reproduction number*). Također, prema analogiji s faznim prijelazima, λ bi se mogla nazvati koeficijentom uređenja. Za $\lambda > \lambda_{critical} = 1$ događa se epidemiološko aktivno stanje, tj. broj zaraženih ne trne u nulu.

Korisno je uvesti koncept [5] *epidemijskog praga* (eng. *epidemic threshold*), tj.

samo ako je $\lambda > 1$ (jedna zaražena osoba u prosjeku zarazi više od jedne nezaražene osobe) moguće je da prenosnik zaraze uzrokuje izboj epidemije konačnih veličina (za SIR model) ili da epidemija uđe u endemsko stanje (u SIS modelu). Ako je $\lambda < 1$ udio zaraženih postaje zanemarivo mali, te naposljetku i iščezava.

Za SIR model vrijede slične diferencijalne jednadžbe:

$$\frac{dN^S}{dt} = -\mu \frac{N^S N^I}{N} \tag{2.12}$$

$$\frac{dN^I}{dt} = \mu \frac{N^S N^I}{N} - \nu N^I \tag{2.13}$$

$$\frac{dN^R}{dt} = \nu N^I \tag{2.14}$$

no one nisu analitički riješive, već samo numerički.

Iako se za ovakve procese smatra da su difuzno-reakcijski sistemi, difuzija je gotovo u potpunosti zanemrena. Prostorni utjecaj i raspored odjeljaka moguće je uvesti uz koncepte kompleksnih mreža te na njima simulirati širenje zaraze. Epidemiološko širenje upravljano je nizom procesa koji se događaju s nekom vjerojatnošću. Za točnu analizu epidemije potrebno je eksplicitno uzeti u obzir stohastičnost procesa, a tada je uvjet $\lambda > 1$ nužan za širenje epidemije, no nije i dovoljan [7] za male populacije.

3 Kompleksne mreže

Svijet koji nas okružuje, moguće je zamisliti kao skup različitih mreža. Čovjek, kao pojedinac, predstavlja osnovnu jedinicu socijalne mreže dok su njegove interakcije s drugim ljudima veze među jedinicama. Mrežama je moguće opisati neke fizičke pojave poput odnosa u hranidbenom lancu, mreže zračnog prometa, vodovodne mreže naseljenog mjesta i drugih sličnih mreža koje su dobro definirane svojim položajem u euklidskom prostoru. Osim toga, mrežama je moguće opisati i apstraktne odnose poput mreže suradnji na znanstvenom članku ili glumačkih suradnji na filmovima.

Povijesno, proučavanje mreža započinje kroz teoriju grafova. Švicarski matematičar Leonard Euler bavio se problemom prelaska svih mostova u gradu Königsbergu točno jedanput. Rješenje tog problema objavio je 1736. godine što se smatra rođenjem ove grane matematike.

Posljednjih desetljeća razvija se proučavanje *kompleksnih mreža*, tj. mreža nepravilne, kompleksne strukture koja se može dinamično mijenjati u vremenu. Sustavi opisani kompleksnim mrežama nazivaju se kompleksnim jer nije moguće predvidjeti kolektivno ponašanje na temelju individualnih komponenti. Razumijevanje matematičkog opisa tih sustava omogućuje nam pretpostavku budućeg ponašanja sustava ili čak i kontrola istog. Ključni radovi za razvitak ovog područja istraživanja su rad Duncana J. Wattsa i Stevena H. Strogatza o mrežama malog svijeta [8] (eng. *small-world networks*) objavljenog 1998. u znanstvenom časopisu *Nature* te rad Alberta-Lászlóa Barabásija i Réke Albert [9] o mrežama bez skale (eng. *scale-free networks*) objavljen godinu dana kasnije u časopisu *Science*. Razvitak računala je ubrzao razvoj ove grane znanosti.

3.1 Osnovni pojmovi

Za razumijevanje i račun s kompleksnim mrežama nužno je uvesti osnovne pojmove iz teorije grafova [10].

Neusmjereni (usmjereni) graf $G = (\mathcal{N}, \mathcal{L})$ definiramo pomuću dva skupa $\mathcal{N}$ i $\mathcal{L}$ koja ga sačinjavaju. Za skupove vrijedi da $\mathcal{N} \neq \emptyset$ te da je skup $\mathcal{L}$ je skup neuređenih (uređenih) parova elemenata skupa $\mathcal{N}$. Elemente skupa $\mathcal{N} \equiv \{n_1, n_2, ..., n_N\}$ nazivamo čvorovima (vrhovima, točkama) grafa G dok elemente skupa $\mathcal{L} \equiv \{l_1, l_2, ..., l_K\}$ zovemo vezama (bridovima, crtama). U skupu $\mathcal{N}$ ima N elemenata, a u skupu $\mathcal{L}$ ima K elemenata, pa za graf G koristimo sljedeće oznake: $G(N, K) = (\mathcal{N}, \mathcal{L})$, $G(N, K)$ ili $G_{N,K}$. Vrhove grafa označavamo brojem *i* koji određuje njihov redni broj u skupu $\mathcal{N}$, dok oznaku brida među vrhovima određuju vrhovi koje taj brid povezuje, tj. označeni su kao par vrhova koje povezuju (i, j) ili l_{ij}.

Dva vrha povezana bridom nazivaju se *susjednim* vrhovima. U usmjerenom je grafu važan redoslijed vrhova i i j te vrijedi $l_{ij} \neq l_{ji}$. Uobičajeno, graf crtamo tako da svakom vrhu pridružimo jednu točku, a brid koji veže dva vrha prikazujemo crtom koja povezuje dvije odgovarajuće točke kao što možemo vidjeti na slici 3.1.

Za ovako definirane grafove nije dopušteno postojanje petlji (veze koje vežu vrh sam sa sobom, tj. veza izlazi iz istog vrha u koji ulazi, eng. *loop*) ili višestrukih veza (jedan par vrhova povezan s više različitih veza).

Za neusmjereni graf $G(N, K)$ najmanji mogući broj bridova među vrhovima je

Slika 3.1: Grafički prikaz neusmjerenog (a), usmjerenog (b) i težinskog usmjerenog (c) grafa sa N=7 čvorova i K=14 veza. U usmjerenom grafu veze u označene strelicama kako bi se pokazala njihova usmjerenost. U težinskom grafu debljina brida proporionalna je težini koju veza ima.

0, dok je najveći mogući broj bridova $N(N-1)/2$ te takav graf nazivamo *potpuno povezanim grafom* i označavamo s K_N. Usmjereni graf može imati duplo više bridova od neusmjerenog budući da vrijedi $l_{ij} \neq l_{ji}$, Za graf kažemo da je *rijedak* ako je $K \ll N^2$ te da je *gust* ako vrijedi $K = \mathcal{O}(N^2)$.

U raznim primjerima korisno je definirati pojam podgrafa. *Podgraf* $G' = (\mathcal{N}', \mathcal{L}')$ grafa $G = (\mathcal{N}, \mathcal{L})$ definiran je tako da vrijedi da je $\mathcal{N}' \subseteq \mathcal{N}$ i $\mathcal{L}' \subseteq \mathcal{L}$, pri čemu skup $\mathcal{L}'$ označava skup bridova kojima su povezani vrhovi u skupu $\mathcal{N}'$. Ako G' sadrži sve veze iz G kojima su povezani vrhovi u $\mathcal{N}'$, tada je G' *podgraf induciran* skupom $\mathcal{N}'$ što označavamo s $G' = G[\mathcal{N}']$.

Najvažnije svojstvo grafova jest povezanost vrhova unutar grafa. Iako nisu svi vrhovi susjedni, oni i dalje mogu biti povezani vezama preko drugih vrhova. *Šetnja* od vrha i do vrha j izmjeničan je niz vrhova i bridova među njima koji započinje u vrhu i i završava u vrhu j. Duljina šetnje određena je brojem veza među čvorovima. *Staza* jest hod u kojem se ni jedna veza ne ponavlja, dok je *put* hod u kojem ni jedan vrh nije posjećen više od točno jednog puta. Hod najmanje duljine (najmanjeg broja veza među vrhovima i i j) naziva se *najkraći put* ili *geodezik*. Zatvoreni hod (započinje završava u istom vrhu), u kojem se ne ponavalja ni jedan brid naziva se *ciklus* (eng. *cycle*) te se označava sa C_k, gdje je k duljina ciklusa (broj bridova od kojih se sastoji ciklus). Najmanji mogući ciklus jest trokut C_3. Petlja bi bila ciklus duljine 1, te ona nije dozvoljena ranijom definicijom grafa. Ako postoji barem jedan put između svakog čvora i i svakog čvora j tada je graf *povezan*. *Komponentom* grafa nazivamo maksimalno povezan podgraf grafa G (svi vrhovi u tom podgrafu povezani su svim

mogućim odgovarajućim bridovima kao i u grafu G). Za komponentu kažemo da je *velika komponenta* ako je reda veličine broja vrhova u grafu N.

Za potrebe računa, korisno je poznavati *matricu susjednosti* (eng. *adjacency matrix*) grafa $G = (\mathcal{N}, \mathcal{L})$. Matrica susjednosti $\mathcal{A}$, kvadratna je $N \times N$ matrica čiji su elementi a_{ij} $(i, j = 1, ..., N)$ jednaki 1 kada veza l_{ij} postoji ili 0 ako veza ne postoji. Budući da nisu dopuštene petlje, elementi na dijagonali jednaki su 0 što znači da je neusmjereni graf dobro opisan simetričnom matricom susjednosti jer vrijedi da je $l_{ij} = l_{ji}$.

Stupanj (povezanost) k_i pojedinog čvora i broj je veza koji upadaju u čvor te je definiran ovako:

$$k_i = \sum_{j \in \mathcal{N}} a_{ij} \tag{3.1}$$

Ako je graf usmjeren, stupanj vrha ima dvije komponente: k_i^{in} i k_i^{out}; čiji zbroj definira ukupan stupanj vrha.

Raspodjela stupnjeva $P(k)$ osnovna je topološka karakteristika grafova te je definirana kao raspodjela vjerojatnosti da pojedini vrh u grafu ima stupanj k ili kao udio vrhova pojedinog stupnja. N-ti moment raspodjele $P(k)$ definiran je kao:

$$\langle k^n \rangle = \sum_k k^n P(k) \tag{3.2}$$

Prvi je moment $\langle k \rangle$ srednja vrijednost stupnja vrhova u grafu G dok drugi moment $\langle k^2 \rangle$ mjeri fluktuacije u raspodjeli stupnjeva.

Duljina najkraćih puteva u grafu dobro opisuje globalnu povezanost grafa, tj. što je ona manja, to je veća povezanost grafa. Zbog takve važnosti najkraćih puteva, korisno je uvesti matricu najkraćih puteva $\mathcal{D}$ čiji su elementi d_{ij} duljina najkraćeg puta od vrha i do vrha j. Najveća vrijednost matričnog elementa d_{ij} matrice $\mathcal{D}$ naziva se *dijametar* grafa i piše se $Diam(G)$. Prosječnu udaljenost dva vrha mjeri se *prosječnom duljinom najkraćeg puta* što se također naziva i *karakteristična duljina puta*. Definirana je na sljedeći način:

$$L = \frac{1}{N(N-1)} \sum_{i,j \in \mathcal{N}, i \neq j} d_{ij} \tag{3.3}$$

Problem ovakve definicije jest u slučaju da postoje nepovezane komponente grafa

za koje ova definicija divergira. Alternativni pristup jest da se definira *učinkovitost* grafa:

$$E = \frac{1}{N(N-1)} \sum_{i,j \in \mathcal{N}, i \neq j} \frac{1}{d_{ij}} \tag{3.4}$$

Važnost pojedinog vrha za povezanost grafa moguće je odrediti mjerom *međupoloženosti vrha* i b_i (eng. *node betweenness*)*[3]

$$b_i = \sum_{j,k \in \mathcal{N}, j \neq k} \frac{n_{jk}(i)}{n_{jk}} \tag{3.5}$$

gdje je n_{jk} broj najkraćih puteva koji povezuju čvorove j i k, a $n_{jk}(i)$ je broj najkraćih putova koji prolaze kroz čvor i. Dakle, vrijednost međupoloženosti vrha i pojašnjava koliki udio svih puteva između bilo koja dva vrha u grafu prolaze točno kroz vrh i. Što je taj udio veći (veći *betweenness* vrha), veću ulogu taj vrh ima za povezanost promatranog grafa.

Povezanost grafa na lokalnoj razini dobro je opisana *grupiranjem* (eng. *clustering*) grafa. Što je grupiranje grafa veće, veći broj susjednih vrhova nekom promatranom vrhu i, povezan je neovisno o vrhu i. Na primjeru mreže poznanstava veličina *grupiranja* određivala bi koliko je vjerojatno da se poznaju dvoje ljudi koji imaju zajedničkog prijatelja.

Veličina c_i, *lokalni koeficijent grupiranja za čvor* i, govori koliko je vjerojatno da je $a_{jm} = 1$ gdje su j i m susjedni čvorovi čvoru i. Vriijednost te veličine se dobiva prebrojavanjem veza e_i kojima su povezani susjedi čvora i (sve veze podskupa G_i susjednih čvorova čvoru i). Budući da je stupanj vrha i jednak k_i, najveći mogući broj veza jest $k_i(k_i - 1)/2$, pa je c_i definiran kao:

$$c_i = \frac{2e_i}{k_i(k_i - 1)} = \frac{\sum_{j,m} a_{ij} a_{jm} a_{mi}}{k_i(k_i - 1)} \tag{3.6}$$

Tada je *koeficijent grupiranja* C dan kao prosječan c_i za sve čvorove u G:

$$C = \langle c \rangle = \frac{1}{N} \sum_{i \in \mathcal{N}} c_i \tag{3.7}$$

te po definiciji vrijedi $0 \leq c_i \leq 1$ i $0 \leq C \leq 1$.

[3]U literaturi se najčešće engleski naziv ni ne prevodi nego se jednostavno govori o *betweenness*u nekog vrha.

Slika 3.2: Grafički prikaz procesa nastajanja mreže malog svijeta po metodi Watts-Strogatz metodi, preuzeto iz [8] uz dozvolu izdavača Springer Nature.

3.2 Mreža malog svijeta

Watts i Strogatz [8] su u svom 1998. radu pretpostavili model nastajanja (interpolacije) kompleksne mreže iz potpuno uređene (pravilne) u neuređenu (nasumičnu) mrežu.

Počevši od pravilne rešetke s n čvorova, od kojih je svaki povezan s prvih k susjeda (k stupnja), svaku vezu nasumično premještamo s vjerojatnošću p, kao na slici 3.2. Takva procedura omogućuje oblikovanje grafa od pravilne rešetke ($p = 0$) u potpuno slučajni graf $p = 1$. Područje vjerojatnosti premrežavanja (premještanja veza) $0 < p < 1$ nazivamo *mrežom malog svijeta* (eng. *small-world network*).

Procedura premrežavanja se izvršava na sljedeći način. Promotrimo sliku 3.2. Počnimo od pravilne prstenaste rešetke s n vrhova povezanih s k najbližih susjeda. Krećući se u smjeru kazaljke na satu, odaberemo vrh i brid kojim je povezan s najbližim susjedom (s desna). Odabranu vezu prespajamo s vjerojatnošću p na neki nasumično odabrani vrh ostatka mreže, uz uvjet da nisu dopuštene višestruke jednakovrijedne veze, ili vezu ostavljamo netaknutu s vjerojatnošću $(1 - p)$. Isti proces se ponavlja za svaki vrh i svaku vezu, tj. ukupno $nk/2$ puta, koliko i ima različitih veza.

Mreža malog svijeta specifična je po svojim strukturalnim osobinama koje mjerimo pomoću prosječne duljine puta $L(p)$ (globalno svojstvo mreže) i koeficijenta grupiranja $C(p)$ (lokalno svojstvo). Prosječna duljina puta $L(p)$ nam pojašnjava kolika je povezanost bilo koja dva čvora unutar mreže, dok koeficijent grupiranja govori kolika je povezanost bliskih susjeda. Premještanjem pojedinih veza, već za male vje-

Slika 3.3: Karakteristična duljina puta $L(p)$ i koeficijent grupiranja $C(p)$ za premrežene grafove s vjerojatnošću p ranije opisanom procedurom, preuzeto iz [8] uz dozvolu izdavača Springer Nature.

rojatnosti p, uvelike se povećava povezanost grafa, tj. poprilično se smanjuje duljina puta za vrlo udaljene vrhove.

Slika 3.3 prikazuje podatke, usrednjene vrijednosti 20 nasumičnih realizacija prethodno opisanog procesa premrežavanja, normaliziranih s vrijednostima prosječne duljine puta $L(0)$ i koeficijenta grupiranja $C(0)$ za pravilnu rešetku. Svih 20 grafova imaju $n = 1000$ čvorova prosječnog stupnja $k = 10$. Na x-osi korištena je logaritamska skala kako bi se jasnije predočilo naglo smanjenje $L(p)$ što predstavlja fenomen malog svijeta. Već za vrlo male vrijednosti p nastaju "prečaci" u mreži koji povezuju prethodno vrlo udaljenje vrhove te tako značajno smanjuju vrijednost $L(p)$, dok uklanjanje veze nema toliko velik utjecaj na promjenu $C(p)$. $C(p)$ ostaje gotovo konstantan te blizak onoj vrijednosti za pravilnu rešetku, dok se $L(p)$ značajno smanjio i približio vrijednosti koja odgovara potpuno nasumičnoj rešetci. To znači da je pojava svojstva malog svijeta gotovo neprimjetna na lokalnoj razini (vrijednost $C(p)$), dok na globalnoj ($L(p)$) postoji vrlo jasna razlika.

Watts i Strogatz posebno su razmatrali rijetko povezane mreže koje zadovoljavaju sljeeće uvjete: $n \gg k \gg ln(n) \gg 1$. Uvjet $k \gg ln(n)$ osigurava da će graf biti povezan [11]. U tom su režimu otkrili da je $L \sim n/2k \gg 1$ i $C \sim 3/4$ za $p \to 0$, dok je

$L \approx L_{random} \sim ln(n)/ln(k)$ i $C \approx C_{random} \sim k/n \ll 1$ za $p \to 1$. Za pravilnu rešetku, $p = 0$, mreža je jako grupirana te je L proporcionalan n, dok je potpuno nasumična mreža, $p = 1$, slabo grupirana, ali je L proporcionalan s $ln(n)$. Slika 3.3 nam otkriva da postoji širok spektar između graničnih vrijednosti p za koje je vrijednost $L(p)$ mala poput one za nasumičnu rešetku, dok je $C(p)$ puno veća od vrijednosti za nasumičnu rešetku.

Zbog takvog utjecaja postojanja prečaca u mreži, nastala je ideja da je fenomen malog svijeta moguće uočiti u rijetkim mrežama ako postoje "prečaci" koji povećavaju povezanost mreže. Tu su hipotezu testirali na mreži glumačkih suradnji na filmovima u to vrijeme (socijalna mreža), na električnoj mreži SAD-a te na živčanoj mreži oblića C. elegans (primjer potpuno mapirane živčane mreže). Pokazano je da sve tri mreže pokazuju svojstva mreže malog svijeta, tj. imaju vrlo visok koeficijent grupiranja $C(p)$ te vrlo malu vrijednost karakteristične duljine puta $L(p)$.

3.3 Mreže bez skale

Sustavi raznoliki poput genetskih mreža ili World Wide Web-a najbolje su opisane kao mreže kompleksne topologije. Zajedničko svojstvo mnogih velikih mreža jest da stupanj čvorova slijedi zakon raspodjele bez skale (eng. *scale-free law*) [9]. To je svojstvo posljedica dva mehanizma: mreža se širi kontinuiranim dodavanjem novih čvorova i novi čvorovi se radije vežu s već dobro povezanim čvorovima (čvorovima s velikim stupnjem). Model temeljen na tim dvjema pretpostavkama dobro opisuje promatrane stacionarne mreže bez skale.

Graf glumačkih suradnji na filmu dobro je dokumentiran primjer socijalne mreže. Svaki je glumac prikazan čvorom, a dva su čvora povezana ako su glumci zajedno glumili u istom filmu. Raspodjela vjerojatnosti broja suradnji za glumce slijedi sljedeći zakon: $P(k) \sim k^{-\gamma}$ gdje je $\gamma = 2.3 \pm 0.1$ što se vidi na slici 3.4 (A). Puno kompleksnija mreža s preko 800 milijuna čvorova je WWW (1999. godine), gdje je čvor dokument, a veze su poveznice s drugih dokumenta koje usmjeravaju na taj dokument. Topologija tog grafa određuje povezanost na WWW, a samim time i uspješnost pronalaska tražene informacije. Vjerojatnost da k dokumenata pokazuje na neku web stranicu jest dana na sljedeći način: $P(k) \sim k^{-\gamma_{WWW}}$, gdje je $\gamma_{WWW} = 2.1 \pm 0.1$ što se vidi na slici 3.4 (B). Topologija mreže napajanja u zapadnim državama SAD-a ocrtava

Slika 3.4: Prikaz raspodjele vjerojatnosti stupnja čvorova za različite mreže. (A) Mreža glumačkih suradnji s $N = 212,250$ čvorova sa prosječnim stupnjem $\langle k \rangle = 28.78$. (B) WWW, $N = 325,729$, $\langle k \rangle = 5.46$. (C) Mreža napajanja $N = 4941$ $\langle k \rangle = 2.67$. Crtkane linije imaju nagibe $\gamma = 2.3$, $\gamma_{WWW} = 2.1$ i $\gamma_{power} = 4$, preuzeto iz [9]. Ponovo upotrijebljeno uz dopuštenje izdavača American Association for the Advancement of Science.

povijesni urbani i industrijski razvoj, gdje su čvorovi generatori, transformatori i razdjelne stanice, a veze su visokonaponske prijenosne žice. Stupanj tih čvorova slijedi zakon s $\gamma_{power} \approx 4$ kao što se može vidjeti na slici 3.4 (C).

Napredak ovakvog modela nastanka realnih mreža u odnosu na prethodne modele slučajnih grafova i mreža malog svijeta jest taj što se ne pretpostavlja konačan broj čvorova, već čvorovi kontinuirano dodaju te njihov broj raste. Uz to, prethodni modeli pretpostavljaju da su veze među čvorovima nasumične i jednolike, dok u mnogo realnih mreža pokazuju preferirano povezivanje čvorova. Primjerice, novi glumac će najvjerojatnije dobiti sporednu ulogu s već dobro poznatim i aktivnim glumcima.

U svom radu, Albert-László Barabási i Réka Albert pokazali su da ta dva uvjeta doista tvore ranije spomenute distribucije stupnja. Započinju s malim brojem čvorova m_0. U svakom sljedećem vremenskom koraku dodaju novi čvor s m $(\leq m_0)$ veza s m postojećih čvorova u mreži. Pretpostavili su da vjerojatnost Π da je novi čvor povezan s nekim čvorom ovisi o stupnju tog čvora k, tako da je $\Pi(k_i) = \frac{k_i}{\sum_j k_j}$. Nakon t vremenskih koraka mreža se sastoji od $m_o + t$ čvorova i mt veza. Takvim principom nastaje mreža s $P(k) \sim k^{-\gamma_{model}}$ raspodjelom vjerojatnosti stupnja čvora gdje je $\gamma_{model} = 2.9 \pm 0.1$. Budući da raspodjela vjerojatnosti treba opisivati stanje mreže u različitim trenutcima razvoja, ona mora biti neovisna o vremenu što je i postignuto. Mreža se na ovakav način, neovisno o veličini, sama organizira u stacionarno stanje

bez skale.

Razmotrili su dvije verzije modela. Model A sadrži rastući karakter, no preferencijalni dodatak je uklonjen pretpostavljajući da se novi čvor veže s jednakom vjerojatnošću na bilo koji čvor, $\Pi(k) = konst. = 1/(m_0 + t - 1)$. Takav model uzrokuje raspodjelu vjerojatnosti $P(k) \sim exp(-\beta k)$, što znači da izostanak preferencijalnog vezanja čvorova ne vodi na raspodjelu vjerojatnosti bez skale. U modelu B, kreću od N nepovezanih čvorova. U svakom vremenskom koraku, nasumično se odabere čvor i poveže se s vjerojatnošću $\Pi(k_i) = \frac{k_i}{\sum_j k_j}$ s drugim čvorom u sustavu. U ranim trenutcima, model slijedi zakon distribucije vjerojatnosti bez skale, ali $P(k)$ nije stacionarna, tj. ovisi o vremenu jer broj čvorova ostaje stalan u vremenu, a broj veza se povećava. Nakon otprilike $T \approx N^2$, svi čvorovi su povezani. To pokazuje da su oba uvjeta nužna za nastanak raspodjele vjerojatnosti stupnja bez skale.

Rast i preferirano vezanje zajedničko su svojstvo mnogo kompleksnih mreža, stoga je očekivano da će se raspodjela vjerojatnosti povezanosti čvorava ponašati po zakonu bez skale u tim mrežama.

3.4 Primjena kompleksnih mreža u istraživanjima u fizici

Kompleksne mreže nalaze široku primjenu u raznim područjima prirodnih pojava i ljudskog života poput: modeliranja socijalnih mreža, mreža živčanih sustava u organizmima, kretanja prometa ili nekih drugih struktura poput apstraktnih mreža suradnji na znanstvenim člancima. Po definiciji fizika (grč. $\phi\upsilon\sigma\iota\kappa\eta$; prirodan, naravan) je "prirodna znanost koja se bavi materijom, gibanjem, energijom i međudjelovanjem"[4] u koju spadaju i gore navedeni primjeri. Ipak, u ovom ćemo poglavlju razmotriti primjenu kompleksnih mreža u područjima koja se nesumnjivo vežu uz studij fizike.

Čestice i sile među njima u zrnatim materijalima[5] imaju kompleksnu organizaciju na lokalnoj i globalnoj skali sustava. Raspored čestica u zrnatim materijalima značajno utječe na odgovor materijala na vanjske perturbacije. Prije uvođenja koncepta kompleksnih mreža, ovakvi su se materijali pučavali isključivo koristeći se aproksimacijom kontinuuma ili promatranjem djelovanja pojedinih čestica na lokalnoj razini, što onemogućuje sustavni opis organizacije čestica na svim relevantinim

[4]https://enciklopedija.hr/Natuknica.aspx?ID=19792

[5]Zrnati materijali su cjelina diskretnih, makroskopskih čestica koje međudjeluju dodirnim silama. Ovakvi sustavi su neravnotežni, tj. čestice su dovoljno velike da na njih ne utječu termalne fluktuacije te gube energiju trenjem i neelastičnim interakcijama sa susjednim česticama.

skalama. Lia Papadopoulos et al. [12] su 2018. sabrali dotadašnje primjene kompleksnih mreža u zrnatim materijalima te predložili nove metode za proučavanje i opis zrnatih materijala pomoću kompleksnih mreža.

Ističu se dvije metode uporabe komleksnih mreža za ops dinamike znatih materijala: mreža kontakata te težinska mreža sila. Pomoću mreže kontakata moguće je razumijeti kako različiti sustavi odgovaraju na vanjske perturbacije te ustvrditi koje svojstvo sustava poboljšava otpornost na perturbacije. Također, pomoću mreže kontakata proučava se utjecaj ciklusa te je pokazano da ciklusi s neparnim brojem vrhova (posebice s tri vrha) pružaju stabilnost strukturi materijala. Težinska mreža sila (težina veze proporcionalna je jakosti sile koja djeluje među vrhovima) motivirana je uočenom neravnomjernom raspodjelom opterećanja po česticama na fotoelastičnom disku. Eksperimentima i numeričkim simulacijama je pokazano da se sile prenose okosnicom lanaca sila. Težinskom mrežom sila se pokušava razumijeti nastanak lanaca sila te kako se njihova struktura mijenja u materijalu ili u različitim uvjetima što pruža ključne informacije za kontrolu modula elastičnosti i mehaničke stabilnosti te akustičku transmisiju u zrnatim materijalima.

Osim primjene kompleksnih mreža u fizici, zanimljiva je primjena fizičkih modela za predviđanje i kontrolu dinamike kompleksnih mreža. Tako su Krioukov et al. [13] u svom radu pokazali da je kauzalna mreža strukture prostor-vremena velikih skala akcelerirajućeg svemira zapravo mreža bez skale velikog koeficijenta grupiranja, slično kompleksnim mrežama Interneta, socijalnih i bioliških mreža. Također, pokazali su da je ta strukturalna sličnost uzrokovana asimptotskom jednakošću rasta kompleksnih mreža i kauzalnih mreža velikih skala, što upućuje na to da slični zakoni upravljaju dinamikom kompleksnih mreža i prostor-vremena u svemiru. Takva spoznaja omogućuje primjenu rezultata iz kozmologije na kompleksne mreže, kao i rezultata kompleksnih mreža na kozmologiju. Vrhovi u deSitterovom prostor-vremenu su zapravo kvanti prostor vremena koji su međusobno povezani ako ispunjaju uvjet kauzalnosti, odnosno ako se nalaze jedan drugome u svjetlosnom stošcu. Pokazano je da vrhovi u deSitterovom prostor-vremenu koji odgovara sadašnjosti slijede raspodjelu stupnjeva $P(k) \approx k^{-\gamma}$, gdje je $\gamma = 2$ uz velik koeficijent grupiranja što je slično velikom broju razmatranih kompleksnih mreža. Osim toga, pokazano je da se sličnim principom dodaju novi vrhovi kauzalnim i kompleksnim mrežama, što omogućuje buduće implikacije kozmologije u kompleksnim mrežama i kompleksnih

mreža u kozmologiji.

Osim ovdje istaknutih primjera, postoje i druge primjene kompleksnih mreža u fizici, poput razmatranja hijerarhije i uređenja kemijskih reakcija u planetarnim atmosferama i u međuzvjezdanom mediju [14] ili proučavanja kozmičke mreže, tj. raspodjele galaksija na velikim skalama [15]. Zanimljivo je također preplitanje fizike kvantnih informacija i kompleksnih mreža [16]. Kao što je iz danih primjera moguće vidjeti, kompleksne mreže su našle široku primjenu u fizici, a također rezultati iz fizike mogu pomoći u boljem razumijevanju dinamike kompleksnih mreža.

4 Širenje zaraze na kompleksnim mrežama

Primjena koncepta kompleksnih mreža u modeliranju dinamike širenja zaraznih bolesti omogućila je napredak te jasniji i točniji prikaz stvarnosti. Kompleksne mreže omogućuju vrlo jasno prikazati jedinke kao vrhove grafa te njihove interakcije kao veze među vrhovima, koji mogu imati i svoj smjer i svoju težinu, ovisno o važnosti pojedine interakcije. Takav koncept omogućio je vrlo točno modeliranje širenja zaraze unutar neke populacije ili nekog konceptualno sličnog problema poput širenja virusa na računalnim mrežama ili širenje glasina među populacijom.

Budući da mreže malog svijeta vrlo dobro opisuju stvarne mreže u svijetu, razmotrit ćemo karakteristike širenja zaraze na mreži malog svijeta. Primjenom kompleksnih mreža u epidemiološkom modeliranju, sve češće je u znanstvenim zajednicama zanimljiv problem pronalaska izvora zaraze, što omogućuje prevenciju budućeg širenja ili zaustavljanje trenutačnog širenja zaraze. Ovaj je problem vrlo čest i upotrebljiv u slučaju širenja zaraze na računalnim mrežama. U ovom je poglavlju predstavljen kratak prikaz najznačajnijih metoda prepoznavanja izvorišta zaraze.

4.1 Širenje zaraze na mreži malog svijeta

Utjecaj povezanosti mreže malog svijeta na širenje zaraze prema pojednostavljenom modelu, u svom su radu ispitali Watts i Strogatz [8]. Struktura populacije modelirana je mrežama malog svijeta nastalih kao što je opisano u potpoglavlju 3.2. U trenutku $t = 0$ u prethodno potpuno zdravu populaciju unese se jedna zaražena jedinka. Nakon proteka jedne vremenske jedinice, zaražena se jedinka trajno uklanja

Slika 4.1: Grafovi prikazuju utjecaj vjerojatnosti premrežavanja mreže na širenje zaraze. (a) Graf ovisnost kritične vrijednosti vjerojatnosti zaraze r_{half} o vjerojatnosti premrežavanja p. (b) Vrijeme $T(p)$ (skalirano s vrijednošću vremena $T(0)$ za pravilnu rešetku i uspoređeno s karakterističnom duljinom puta $L(p)$ skaliranu $L(0)$) potrebno za maksimalnu zaraznost ($r = 1$) da se proširi po cijeloj populaciji ima jednak oblik kao i karakteristična duljina puta $L(p)$. Čak i za vrlo male vrijednosti p (mali broj premještenih bridova), vrijeme globalne infekcije populacije postaje vrlo kratko poput vremena za slučajni graf ($p = 1$), preuzeto iz [8] uz dozvolu izdavača Springer Nature.

(zbog stvaranja imuniteta ili smrti). Dok je pojedinac zaražen, može zaraziti sve svoje susjede s vjerojatnošću r. Takvim postupkom zaraza se širi bridovima grafa u narednim vremenskim tenutcima sve dok se ne zarazi cijela populacija ili dok zaraza ne iščezne, a da je pritom zarazila samo jedan udio populacije.

Razmatranja daju dva važna rezultata. Prvi je važan rezultat da kritična vrijednost zaraze r_{half}[6] opada za faktor $\cong 2$ već s vrlo malim vrijednostima vjerojatnosti premrežavanja p (slika 4.1(a)). Drugi je važan rezultat da vrijeme $T(p)$, potrebno da se zaraza proširi po cijeloj populaciji (za zarazu čije je vjerojatnost širenja zaraze do-

[6]Vjerojatnost širenja zaraze pri kojoj se zarazi polovica populacije na završetku širenja zaraze na toj mreži.

voljna da se zaraza može proširiti po cijeloj populaciji, neovisno o njenoj strukturi), nalikuje krivulji za karakteritičnu duljinu puta $L(p)$ (slika 4.1(b)).

Konačno, možemo zaključiti da i vrlo male vjerojatnosti premrežavanja značajno utječu na smanjenje kritične vjerojatnosti kao i na ukupno vrijeme potrebno zarazi da se proširi cijelom mrežom. Slično kao što smo zaključili u poglavlju 3.2, zaključujemo da već i vrlo mali broj "prečaca" značajno olakšava i ubrzava širenje zaraze mrežom.

4.2 Identifikacija izvora zaraze

Ideja problema identifikacije izvorišta zaraze jest na temelju dane slike o stanju epidemije u nekom trenutku t na dobro poznatoj mreži pokušati odrediti odakle se zaraza počela širiti. Rješenje takvog problema nalazi svoju primjenu u prevenciji daljnjeg širenja, kao i u potencijalnom konačnom zaustavljanju širenja zaraze. Pronalazak izvorišta zaraze može imati značajnu ulogu u zaustavljanju ili prevenciji širenja zaraze ili općenito razumijevanju širenja zaraze, kao što je to bio slučaj u radu Johna Snowa.

Prije razvitka epidemioloških modela na kompleksnim mrežama, epidemiolozi su pratili genetske promjene sojeva virusa kroz vrijeme kako bi identificirali izvorište zaraze što je zahtijevalo puno vremena. Drugi pristup je bio jednostavno praćenje niza kontakata unazad, što često nije bilo moguće jer su nedostajali neki ključni podatci. U posljednjem je desetljeću problem detekcije izvora zaraze privukao mnoga istraživanja definiran na sljedeći način: "Ako dobijemo trenutnu sliku stanja epidemije na nekoj mreži, može li se odrediti koji je vrh izvor zaraze?" Problem postaje vrlo zahtijevan već i za vrlo jednostavne modele poput SI ili SIR modela. Osnovna pretpostavka jest da je u $t = 0$ zaražen samo jedan vrh, a svi ostali su nezaraženi (S).

Za SIR model ističe se nekoliko različitih pristupa. Kai Zhu i Lei Ying u svom su radu [17] predložili metodu brojanja mogućih putanja uzorka (eng. *sample path counting*) pomoću koje su dokazali da izvorišni vrh ima najmanju udaljenost do najudaljenijih zaraženih vrhova za razliku od ostalih vrhova. Zaključuju da se problem može definirati kao problem procjene najveće vjerojatnosti (eng. *maximum likelihood estimation*, MLE) no za njegovo rješavanje bilo bi potrebno razmotriti sve moguće putove koji bi doveli do trenutnog stanja zaraze na mreži što je problem reda veličine t^N gdje je t vrijeme u kojem promatramo stanje zaraze, a N je broj čvorova u mreži.

Takav problem postaje već poprilično teško riješiv, čak i ako je t poznato, dok su se autori u radu odlučili razmatrati problem gdje t nije poznat, što čini problem puno težim. Definiranjem ekscentriciteta[7] vrha, uspjeli su dokazati da izvorište zaraze ima najmanji zarazni ekscentricitet. Vrh s najmanjim ekscentricitetom u grafu naziva se Jordanov centar, prema tome izvorište zaraze je Jordanov centar zaraze, tj. izvorište zaraze je onaj vrh (bilo da je on u stanju I ili R) koji ima najmanji ekscentricitet za zaražene vrhove. Odnosno, izvorište zaraze je najmanje udaljeno od svih zaraženih vrhova. U svom radu, razvili su algoritam reverzne zaraze kako bi odredili optimalan put širenja zaraze na principu da svaki zaraženi vrh šalje informaciju o zarazi po mreži, te vrh koji prvi prikupi sve informacije deklarira se izvorištem zaraze. Trajanje algoritma jednako je najmanjem ekscentricitetu zaraze, a broj poruka koje svaki vrh primi tijekom jedne iteracije jednak je stupnju tog vrha. Napravili su provjeru točnosti detekcije izvora zaraze na stvarnim mrežama poput IAS mreže (eng. *Internet Autonomous Systems*) i otprilike 80% procjenjenih izvora zaraze je udaljeno svega jedan ili dva vrha od točnog izvora zaraze.

Andrey Lokhov et al. 2014. su godine objavili rad [18] u kojem za otkrivanje izvorišta zaraze koriste algoritamsko zaključivanje temeljeno na jednadžbama dinamičkog prenošenja poruke (eng. *dynamic message-passing*, DMP) što je pokazalo napredak u odnosu na dotadašnje pristupe. Posebnost tog algoritma jest da je učinkovit čak i kada je poznato stanje samo dijela mreže. Za razliku od prethodnih radova u kojima se koristila metoda dinamičkog prenošenja zaraze, u svom su radu izveli DMP jednadžbe za raspodjelu vjerojatnosti $P_S^i(t)$, $P_I^i(t)$ i $P_R^i(t)$ nezaraženih, zaraženih i oporavljenih čvorova na temelju SIR modela na grafu u iterativnom obliku. Pretpostavka je bila da je u $t = 0$ samo jedan čvor u mreži zaražen (I), dok su svi ostali nezaraženi (S). DMP jednadžbe mogu podsjećati na diferencijalne jednadžbe srednjeg polja u kojima je korištena aproksimacija homogenog miješanja koje daju aproksimativno rješenje stvarnih vjerojatnosti, dok su DMP jednadžbe egzaktne u termodinamičkom limesu $N \to \infty$ sve dok za ansambl slučajnih grafova vrijedi da vjerojatnost da nasumično odabrani čvor pripada petlji konačne duljine ide u nula za grafove s vrlo velikim N. Koristeći Bayesov teorem, vjerojatnost da čvor i kao izvor zaraze stvara uočeno stanje $\mathcal{O}$ u trenutku t_0 proporcionalna je vjerojatnosti da se dobiva uočeno stanje $\mathcal{O}$ ako znamo da je i izvor zaraze, $P(i \mid \mathcal{O}) \sim P(\mathcal{O} \mid i)$. Također,

[7] Najveća udaljenost promatranog vrha do zaraženog vrha.

definirali su funkciju nalik energiji $E(i) \equiv -\log P(\mathcal{O} \mid i)$, tj. vrhovi s manjom energijom imaju veću vjerojatnost da su izvorište zaraze. Budući da nema sljedivog načina da se egzaktno izračuna $P(\mathcal{O} \mid i)$, aproksimirali su vjerojatnost s pristupom srednjeg polja kao produkt raspodjele vjerojatnosti dobivenih iz DMP:

$$P(\mathcal{O} \mid i) \simeq \prod_{\substack{k \in \mathcal{O} \\ q_k(t_0)=S}} P_S^k(t,i) \prod_{\substack{l \in \mathcal{O} \\ q_l(t_0)=I}} P_I^l(t,i) \prod_{\substack{m \in \mathcal{O} \\ q_m(t_0)=R}} P_R^m(t,i) \tag{4.1}$$

gdje su $q_{k,\ l,\ m}(t_0)$ oznake za vrhove $k,\ l,\ m$ u uočenom stanju $\mathcal{O}$ koji su redom u stanju S, I odnosno R. Kako bi se procijenio točan trenutak t_0 u kojem se promatra stanje mreže, računa se vrijednost $E(i,t)$ za različite moguće vrijednosti t i odabire se ona za koju je "particijska funkcija" $Z(t) \equiv \sum_i e^{-E(i,t)}$ maksimalna. Algoritamska složenost računanja $E(i)$ je red veličine $\mathcal{O}(t_0 N c)$ gdje je c prosječan stupanj čvorova u grafu. Takva metoda određivanja izvora zaraze značajno nadmašuje uspješnost ranijih metoda pomoću Jordanovih centara [17] i centra glasina [19] [20] (eng. *rumour centrality*).

Slična metodi DMP jest metoda širenja vjerovanja (eng. *belief propagation*, BP) koju su Altarelli et al. iskoristili u svom radu [21] kako bi izveli jednadžbe koje opisuju vremensku evoluciju raspodjele vjerojatnosti utemeljenu na nekom opaženom stanju epidemije na sustavu. Njihov princip računanja distribucija je vrlo učinkovit na širokom rasponu epidemioloških modela, u statičkom i dinamičkom slučaju.

Nino Anutulov-Fantulin et al. 2015. godine su u svom radu [22] predložili novi pristup identifikaciji izvorišta zaraze u statičnim i vremenski ovisnim mrežama te proučili njihovu robusnost i ograničenja. Razvili su egzaktna rješenja uz analitičku kombinatoriku [23] te egzaktne i aproksimativne raspodjele vjerojatnosti uz Monte Carlo simulacije (direktne i *Soft Margin*[8]) te su načinili referentna rješenja na kvadratnoj pravilnoj rešetci. Osim toga, mjerili su detektabilnost izvorišta zaraze koristeći normaliziranu Shannonovu[9] entropiju raspodjele vjerojatnosti da je promatrani vrh izvorište zaraze. Uočili su da postoje režimi velike detektabinosti kao i velike nedetektabilnosti za SIR model, koja prvotno ovisi o karakteristikama pro-

[8]Ne uzimaju se samo identično točna rješanja, nego se uzimaju i rješenja s određenim odstupanjem; često to znači da se razmatraju rješenja koja zadovoljavaju Gaussovu raspodjelu oko točnog rješenja.

[9]Mjera informacije koje je moguće dobiti iz pojedinih varijabli. Definirana je na sljedeći način: ako imamo diskretnu nasumičnu varijablu X, s mogućim ishodima $x_1, ..., x_n$, koji se događaju s vjerojatnošću $P(x_1), ..., P(x_n)$, tada je entropija X dana s $H(X) = -\sum_{i=1}^{n} P(x_i) \log P(x_i)$.

cesa širenja zaraze. Naposljetku, koristeći simulacije za seksualno prenosivu zarazu na vremenskom intervalu od 200 dana na empirijskoj mreži seksualnih kontakata provjerili su robustnost procjene izvorišta zaraze koristeći *Soft Margin* Monte Carlo simulacije. Slično radu Andreya Lokhova et al., cilj je izračunati vjerojatnost da je promatrani čvor θ_i izvor zaraze ako imamo dano stanje mreže $\vec{r}_*$ u nekom trenutku, tj. želimo izračunati $P(\Theta = \theta_i \mid \vec{R} = \vec{r}_*)$.

Za direktne Monte Carlo simulacije promatramo razvoj zaraze (p je vjerojatnost prijenosa zaraze sa zaraženog vrha (I) na nezaraženi vrh (S) u jednom diskretnom vremenskom koraku, a q je vjerojatnost da se zaraženi vrh I oporavi (R) u jednom diskretnom vremenskom koraku) od svakog od mogućih vrhova za T diskretnih vremenskih koraka, pri čemu se simulacija zaustavlja u trenutku $t < T$ ako je zaražen vrh koji je u danom stanju $\vec{r}_*$ nezaražen. Bilježi se n_i broj simulacija za koje se nakon vremena T postigne upravo željejno stanje mreže $\vec{r}_*$, tada je $P(\Theta = \theta_i \mid \vec{R} = \vec{r}_*) = n_i / \sum_j n_j$.

U analitičko-kominatoričkom pristupu svakom se vrhu stupnja n pridruži maksimalno $(n+1)$-dimenzionalna funkcija izvodnica koja bilježi događaje zaraze promatranog vrha te prijenosa zaraze na njegove susjedne vrhove. Množenjem funkcija izvodnica svih zaraženih vrhova moguće je dobiti raspodjelu vjerojatnosti da je pojedini vrh izvor zaraze. Nedostatak ovog pristupa jest što takav račun postaje iznimno zamršen za mreže koje ne nalikuju stablu.

Napravljen je niz referentnih Monte Carlo simulacija na kvadratnoj mreži s $N = 30 \times 30$ te su izračunate raspodjele vjerojatnosti te izračunata Shannonova entropija $H(\vec{r}_*)$ za različite vjerojatnosti širenja zaraze p i za tri različite vjerojatnosti oporavka q što je prikazano na slici 4.2. Detektabilnost izvora za uočeno stanje $\vec{r}_*$ definirana je kao $D(\vec{r}_*) = 1 - H(\vec{r}_*)$. Detektabilnost pokazuje kvalitativno jednako ponašanje za različite vrijednosti parametra q, što se uočava na slici 4.2. Važno je uočiti tri različita područja detektabilnosti. Područje male detektabinosti ($p < 0.2$), srednje detektabilno područje ($0.2 < p < 0.7$) i izrazito detektabilno područje ($p > 0.7$), što je uočeno i za različite epidemiološke modele poput SI modela.

Budući da direktne Monte Carlo simulacije kao i analitičko-kombinatorički pristup vrlo brzo postaju računalno vrlo zahtijevne, uvodi se *Soft-Margin* metoda. Sličnost vektora $\vec{r}_1$ i $\vec{r}_2$ mjeri se Jaccardovom funkcijom sličnosti $\phi : (R^N \times R^N) \to [0, 1]$ koja se računa kao omjer veličine interakcija zaraženih vrhova u oba vektora i veličine nji-

Slika 4.2: 2D histogrami (a), (b) i (c): Prikaz raspodjele vrijednosti entropije H raspodjele vjerojatnosti da je vrh izvor zaraze za generirane realizacije širenja zaraze na pravilnoj kvadratnoj mreži $N = 30 \times 30$ za $T = 5$, izračunato Monte Carlo simulacijom s $10^6 - 10^8$ simulacija po izvoru. Preuzeto iz [22] uz dozvolu izdavača American Physical Society.

hove unije. Empirička kumulativna funkcija raspodjele n uzoraka slučajne varijable $\phi(\vec{r_*}, \vec{R_\theta})$ označava se s $\hat{F}_\theta(x)$, gdje je x vrijednost varijable sličnosti. Derivacija prethodne funkcije daje funkciju gustoće vjerojatnosti (eng *probability density function*, PDF):

$$\hat{f}_\theta = \frac{d}{dx}\hat{F}_\theta(x) = \frac{1}{n}\sum_{i=1}^{n}\delta(x - \phi(\vec{r_*}, \vec{r}_{\theta,i})) \tag{4.2}$$

gdje je $\delta(x)$ Diracova delta funkcija. Definiramo Soft-Margin izraz:

$$\hat{P}(\vec{R} = \vec{r_*} \mid \Theta = \theta) = \int_0^1 w_a(x)\hat{f}_\theta(x)dx \tag{4.3}$$

gdje je $w_a(x)$ težinska funkcija. U radu je za težinsku funkciju korištena Gaussova težinska forma $w_a(x) = exp[-(x-1)^2/a^2]$. Za $a \to 0$ dobiva se direktna Monte Carlo raspodjela. Uvrštavajući definiciju $\hat{f}_\theta(x)$ u jednadžbu (4.3) te koristeći svojstva delta funkcije, dobiva se sljedeći izraz:

$$\hat{P}(\vec{R} = \vec{r_*} \mid \Theta = \theta) = \frac{1}{n}\sum_{i=1}^{n} exp\Big(-\frac{[\phi(\vec{r_*}, \vec{r}_{\theta,i}) - 1]^2}{a^2}\Big) \tag{4.4}$$

Parametar a su odabrali kao infimum skupa parametara za koje PDF konvergeriraju.

Slika 4.3: 2D histogrami (a), (b) i (c): Prikaz raspodjele vrijednosti entropije H za raspodjele vjerojatnosti izvora zaraze za realizacije širenja zaraze iz centra mreže na pravilnim četverokutnim mrežama (3×3, 5×5, 7×7) za $T = 5$, $q = 0.5$ i različite vrijednosti p, izračunato Soft Margin metodom s $10^6 - 10^8$ simulacija po izvoru. a je odabran prema uvjetu konvergecije. Preuzeto iz [22] uz dozvolu izdavača American Physical Society.

Na slici 4.3 prikazani su dijagrami entropije H (detektabilnosti (D)) za tri različite četverokutne mreže. Za mrežu $N = 3 \times 3$ entropija je gotovo 100% što znači da je gotovo nemoguće odrediti koji je vrh izvor zaraze, budući da za $T = 5$ na toliko maloj mreži gotovo svi vrhovi mogu postići uočeno stanje zaraze na mreži. Za ostala dva primjera mreže, raspodjela entropije podsjeća na raniju raspodjelu s direktnim Monte Carlo simulacijama, no rasap entropije je veći, što je očekivano budući da prihvaćamo rješenja s određenim odstupanjem od očekivanih.

U radu [22] su autori Soft Margin metodom pokušali odrediti izvorište zaraze na empirijskoj mreži seksualnih kontakata uz neodređenost početka širenja zaraze t_0 za $\varepsilon = 50$ dana, u ukupnom trajanju epidemije 300 dana te su uspješno odredili izvorište zaraze unutra prvih susjeda u 60% eksperimenata. Također su ispitali i utjecaj neodređenosti slijeda interakcija unutar vremenskog intervala Δ[10] te su zaključili da je utjecaj gotovo zanemariv, odnosno sustav je robustan na šum male veličine. Ta su mjerenja prikazana na slici 4.4. Osim toga, pokazali su da Soft Margin metoda određuje izvorište zaraze s puno većom preciznošću od Jordan i DMP metoda [23].

[10]Unutar vremena Δ nasumično je permutiran slijed interakcija.

Slika 4.4: Graf (a): Prikaz empirijske vremenske mreže seksualnih kontakata u Brazilu. Grafovi (b), (c) i (d) prikazuju izmjereni udio od 500 eksperimenata u odnosu na udaljenost izračunatog izvora zaraze od točnog izvora. Graf (b) prikazuje raspodjelu vjerojatnosti izvorišta zaraze za nasumičnu metodu (jednolika raspodjela vjerojatnosti potencijalnih izvora zaraze). Graf (c) prikazuje utjecaj vremenske neodređenosti pojavljivanja zaraze $[t_0 - \varepsilon, t_0 + \varepsilon]$. Graf (d) prikazuje utjecaj neodređenosti slijeda interakcija unutar vremenskog intervala Δ uz $\varepsilon = 0$. Preuzeto iz [22] uz dozvolu izdavača American Physical Society.

5 Implementacija širenja zaraze na mreži malog svijeta

Cilj ovog diplomskog rada jest odrediti po kojoj se karakteristici izvorište zaraze u SIS modelu razlikuje od ostalih vrhova na mreži malog svijeta. Posebnost SIS modela u odnosu na SIR model jest da se jednom zaraženi i oporavljeni vrh ponovno može zaraziti, što uzrokuje gubitak informacije, odnosno entropija sustava se povećava. Kako bismo uopće otkrili po kojoj se to karakteristici vrh koji je izvor zaraze ističe, potrebno je prvo napraviti algoritam širenja zaraze te provjeriti točnost u usporedbi s (postojećim) analitičkim rješenjima za širenje zaraze.

Stanje zaraze vrhova u mreži u nekom trenutku t pratimo pomoću vektora $X(t)$ koji ima onoliki broj komponenti koliko ima vrhova u sustvau koji se promatra. Svaka komponenta x_i može poprimiti vrijednost ili 1 ili 0 ovisno je li vrh i zaražen (I) ili nezaražen (S). Budući da pretpostavljamo da promatramo Markovljev proces, stanje mreže u trenutku t ovisit će samo o stanju mreže u trenutku $t - 1$.

U prvom ćemo potpoglavlju dokazati da algoritam doista odgovara dostupnim analitičkim rješenjima za usmjerene i neusmjerene lance s 2 i 3 vrha. Zatim ćemo

u drugom potpoglavlju provjeriti kako se razvija veličina $R_I^i(t)$ koju ćemo definirati kao prosječnu udaljenost svih zaraženih vrhova od promatranog vrha i (dobro je definirana samo za povezane komponente grafa, dok za nepovezane komponente grafa ona divergira). Budući da je moguće uočiti da je gradijent $R_I^i(t)$ specifičan za vrhove, ovisno o tome koliko su udaljeni od izvorišta zaraze, u sljedećem ćemo potpoglavlju provjeriti kako se ta mjera mijenja sa svojstvima mreže te dokazati kako je upravo to karakteristika po kojoj se ističe izvor zaraze. U posljednjim potpoglavljima provjerit ćemo vrijedi li isti zaključak na nekim stvarnim mrežama.

5.1 Provjera valjanosti algoritma na lancu

Praćenje širenja zaraze na kompleksnim mrežama već za vrlo jednostavne sustave vrlo brzo postaje analitički prekompleksno i nerješivo te je nužno uvesti aproksimacije ili numerički riješavati problem. Primjena simulacija na problemu širenja zaraze pokazala se vrlo korisna i točna u opisivanju realnog razvoja događaja te je računski manje zahtijevna, a vremenski učinkovitija.

Algoritam koji oponaša širenje zaraze u ovom radu implementiran je u programskom jeziku *python*u uz pomoć programskog paketa *NetworkX* [24] specijaliziranog za izradu, prikaz te proučavnje strukture, dinamike i funkcija kompleksnih mreža u pythonu. Više detalja vezanih uz sam kod algoritma, priloženo je u dodatku. Kako bismo bili sigurni da algoritam dovoljno dobro oponaša realnu situaciju, algoritam je testiran na vrlo jednostavnim i analitički riješivim sustavima poput usmjerenog i neusmjerenog lanca s dva ili tri vrha, shematski prikaz je na slici 5.1.

5.1.1 Usmjereni lanac

Usmjereni lanac s dva vrha

Stanje	Prvi vrh - Drugi vrh
1	0 - 0
2	0 - 1
3	1 - 0
4	1 - 1

Tablica 5.1: Klasifikacija realiziranih stanja na vrhovima lanca s dva vrha (stanje 0 odgovara nezaraženom S stanju, a stanje 1 odgovara zaraženom I stanju.

Slika 5.1: Shematski prikaz usmjernog i neusmjerenog lanca s 2 ili 3 vrha.

Usmjereni lanac podskup je neusmjerenog lanca, te se na njemu zaraza može širiti samo u jednom smjeru. U konkretnom slučaju kojeg razmatramo, zaraza se može širiti samo sa vrha 1 na vrh 2, s vrha 2 na vrh 3 itd. Usmjereni lanac s dva vrha najjednostavnija je moguća verzija nekog sustava na kojem bismo mogli provjeriti ispravnost algoritma. Zaraza se može širiti samo s vrha 1 na vrh 2 s vjerojatnošću μ ako i samo ako je vrh 1 zaražen, a vrh 2 nezaražen. Oporavak zaraženih vrhova događa se spontano i neovisno o stanju drugih vrhova s vjerojatnošću ν. Vjerojatnost da se zaraza ne širi sa zaraženog na nezaraženi vrh je $(1-\mu)$, dok je vjerojatnost da se zaraženi vrh ne oporavi $(1-\nu)$. Budući da promatramo lanac s dva vrha od kojih svaki može biti u dva različita stanja, ukupan broj mogućih realizacija stanja zaraženosti lanca je $2^2 = 4$. Sve moguće realizacije stanja zaraze na (usmjerenom) lancu s dva vrha imenujemo brojevima od 1 do 4 kao u tablici 5.1, tada dobivamo matricu vjerojatnosti prelaska iz stanja ostvarenih u trenutku $t-1$ u

stanja u trenutku t danu sa:

$$\begin{bmatrix} P_1(t) \\ P_2(t) \\ P_3(t) \\ P_4(t) \end{bmatrix} = \begin{bmatrix} 1 & \nu & (1-\mu)\nu & \nu^2 \\ 0 & (1-\nu) & \mu\nu & (1-\nu)\nu \\ 0 & 0 & (1-\mu)(1-\nu) & (1-\nu)\nu \\ 0 & 0 & \mu(1-\nu) & (1-\nu)^2 \end{bmatrix} \begin{bmatrix} P_1(t-1) \\ P_2(t-1) \\ P_3(t-1) \\ P_4(t-1) \end{bmatrix}$$

Ako zadamo početni uvjet da je u $t = 0$ ostvareno stanje 3 ($P(0) = (0,\ 0,\ 1,\ 0)$, prvi vrh je zaražen, a drugi nezaražen) dobivamo analitički riješiv sustav s kojim je moguće usporediti vjerojatnosti ostvarivanja stanja u trenutku t koje daje algoritam.

Usmjereni lanac s tri vrha

Stanje	Prvi - Drugi - Treći vrh
1	0 - 0 - 0
2	0 - 0 - 1
3	0 - 1 - 0
4	0 - 1 - 1
5	1 - 0 - 0
6	1 - 0 - 1
7	1 - 1 - 0
8	1 - 1 - 1

Tablica 5.2: Klasifikacija realiziranh stanja na vrhovima lanca s tri vrha (stanje 0 odgovara nezaraženom S stanju, a stanje 1 odgovara zaraženom I stanju).

Analogno definiranju stanja realizacije zaraze na (usmjerenom) lancu s dva vrha, definiramo stanja zaraze na (usmjerenom) lancu s tri vrha. Budući da sad promatramo lanac s tri vrha koja mogu biti u dva različita stanja, ukupan broj mogućih realizacija zaraze na lancu je $2^3 = 8$, stoga ćemo ta stanja i imenovati brojevima od 1 do 8 kao u tablici 5.2.

Kao i ranije, zaraza može preći samo sa zaraženog vrha na nezaraženi vrh u svakom vremenskom koraku s vjerojatnošću μ, ali samo u jednom smjeru (usmjereni lanac), dok se oporavak zaraženih vrhova događa spontano i neovisno o stanjima drugih vrhova u svakom vremenskom koraku s vjerojatnošću ν. Stoga dobivamo matricu koja povezuje stanja realizirana na usmjrenom lancu u trenutku $t-1$ sa stanjima u trenutku t.

$$\begin{bmatrix} 1 & \nu & (1-\mu)\nu & \nu^2 & (1-\mu)\nu & (1-\mu)\nu^2 & (1-\mu)\nu^2 & \nu^3 \\ 0 & (1-\nu) & \mu\nu & (1-\nu)\nu & 0 & (1-\mu)\nu(1-\nu) & \mu\nu^2 & \nu^2(1-\nu) \\ 0 & 0 & (1-\mu)(1-\nu) & (1-\nu)\nu & \mu\nu & \mu\nu^2 & (1-\mu)\nu(1-\nu) & \nu^2(1-\nu) \\ 0 & 0 & \mu(1-\nu) & (1-\nu)^2 & 0 & \mu\nu(1-\nu) & \mu\nu(1-\nu) & \nu(1-\nu)^2 \\ 0 & 0 & 0 & 0 & (1-\mu)(1-\nu) & (1-\mu)(1-\nu)\nu & (1-\mu)\nu(1-\nu) & \nu^2(1-\nu) \\ 0 & 0 & 0 & 0 & 0 & (1-\mu)(1-\nu)^2 & \mu\nu(1-\nu) & \nu(1-\nu)^2 \\ 0 & 0 & 0 & 0 & \mu(1-\nu) & \mu(1-\nu)\nu & (1-\mu)(1-\nu)^2 & \nu(1-\nu)^2 \\ 0 & 0 & 0 & 0 & 0 & \mu(1-\nu)^2 & \mu(1-\nu)^2 & (1-\nu)^3 \end{bmatrix}$$

5.1.2 Neusmjereni lanac

Slično prethodno opisanom, ista stanja kao na usmjerenom lancu s dva ili tri vrha, moguće je realizirati i na neusmjerenom lancu s dva ili tri vrha. Na sličan način računaju se vjerojatnosti ostvarivanja mogućih stanja u trenutku t preko stanja ostvarenih u trenutku $t-1$, uz razliku da se zaraza sada može širiti u dva smjera. Za neusmjereni lanac s dva vrha dobiva se sljedeća matrica:

$$\begin{bmatrix} P_1(t) \\ P_2(t) \\ P_3(t) \\ P_4(t) \end{bmatrix} = \begin{bmatrix} 1 & (1-\mu)\nu & (1-\mu)\nu & \nu^2 \\ 0 & (1-\mu)(1-\nu) & \mu\nu & (1-\nu)\nu \\ 0 & \mu\nu & (1-\mu)(1-\nu) & (1-\nu)\nu \\ 0 & \mu(1-\nu) & \mu(1-\nu) & (1-\nu)^2 \end{bmatrix} \begin{bmatrix} P_1(t-1) \\ P_2(t-1) \\ P_3(t-1) \\ P_4(t-1) \end{bmatrix}$$

Na neusmjerenom lanacu s tri vrha postoji mogućnost da se zaraza prenese s dva zaražena vrha na jedan nezaraženi vrh istovremeno (moguće je samo u stanju 6). U tom slučaju dopuštamo mogućnost da se zaraza prenijela ili sa zaraženog vrha 1 na nezaraženi vrh 2 ili sa zaraženog vrha 3 na nezaraženi vrh 2 ili s vrha 1 i s vrha 3 istovremeno na vrh 2. Tom slučaju odgovara vjerojatnosti prijenosa zaraze $(1-(1-\mu)^2)$, što znači: "Neka se zaraza prenijela **barem** sa jednog zaraženog vrha na nezaraženi". Ostale vjerojatnosti ostaju jednake kao i ranije i stoga dobivamo sljedeću matricu

$$\begin{bmatrix} 1 & \nu(1-\mu) & (1-\mu)^2\nu & (1-\mu)\nu^2 & (1-\mu)\nu & (1-\mu)^2\nu^2 & (1-\mu)\nu^2 & \nu^3 \\ 0 & (1-\mu)(1-\nu) & (1-\mu)\mu\nu & (1-\mu)(1-\nu)\nu & 0 & (1-\mu)^2\nu(1-\nu) & \mu\nu^2 & \nu^2(1-\nu) \\ 0 & \mu\nu & (1-\mu)^2(1-\nu) & (1-\mu)(1-\nu)\nu & \mu\nu & (1-(1-\mu)^2)\nu^2 & (1-\mu)\nu(1-\nu) & \nu^2(1-\nu) \\ 0 & \mu(1-\nu) & \mu(1-\mu)(1-\nu) & (1-\mu)(1-\nu)^2 & 0 & (1-(1-\mu)^2)\nu(1-\nu) & \mu\nu(1-\nu) & \nu(1-\nu)^2 \\ 0 & 0 & \mu\nu(1-\mu) & \mu\nu^2 & (1-\mu)(1-\nu) & (1-\mu)^2(1-\nu)\nu & (1-\mu)\nu(1-\nu) & \nu^2(1-\nu) \\ 0 & 0 & \mu^2\nu & \mu\nu(1-\nu) & 0 & (1-\mu)^2(1-\nu)^2 & \mu\nu(1-\nu) & \nu(1-\nu)^2 \\ 0 & 0 & \mu(1-\mu)(1-\nu) & \mu\nu(1-\nu) & \mu(1-\nu) & (1-(1-\mu)^2)(1-\nu)\nu & (1-\mu)(1-\nu)^2 & \nu(1-\nu)^2 \\ 0 & 0 & \mu^2(1-\nu) & \mu(1-\nu)^2 & 0 & (1-(1-\mu)^2)(1-\nu)^2 & \mu(1-\nu)^2 & (1-\nu)^3 \end{bmatrix}$$

Kao što možemo uočiti, već za malo kompliciraniji problem s lancem od četiri vrha, da bismo mogli doći do analitičkog rješenja bila bi nam potrebna 16×16 matrica (budući da bismo imali $2^4 = 16$ mogućih realizacija stanja zaraze na lancu s četiri vrha) s kojom je daleko teže računati. Jasno je da svi realističniji problemi postaju

neusporedivo teži i gotovo nerješivi analitički zbog čega se koriste numeričke metode.

5.1.3 Apsolutna razlika analitičkih rješenja i rješenja dobivenih algoritmom

Na slikama 5.2 i 5.3 prikazana je konvergencija apsolutne razlike između analitičkih rješenja i rješanja dobivenih simulacijama širenja zaraze na usmjerenim i neusmjerenim lancima s 2 i 3 vrha nakon $T = 20$ diskretnih vremenskih koraka od početka zaraze na lancu. Vjerojatnost širenja zaraze sa zaraženog na nezaraženi vrh je $\mu = 0.3$, a vjerojatnost oporavka zaraženog vrha je $\nu = 0.075$. Uočavamo da se već za 10^7 algoritamskih simulacija dobivaju gotovo identična rješenja analitičkim. U ovom koraku bila nam je važna vrlo velika preciznost kako bismo dokazali točnost algoritma, tj. da algoritam doista reproducira širenje zaraze jednako onom koje očekujemo, što smo ovim mjerenjima i pokazali. U daljnjim koracima koristit će se manji broj ponavljanja jer više ne će biti važna preciznost koliko generalno ponašanje pojedinih promatranih veličina, što će smanjiti računalnu zahtijevnost.

5.2 Kvalitativna svojstva veličine R stanju na mreži malog svijeta

Monte Carlo simulacije na mreži malog svijeta započinjemo s pretpostavkom da je u $t = 0$ zaražen vrh 1. Budući da je u radu [17] pokazano da je izvorište zaraze Jordanov centar zaraze[11] u slučaju SIR modela, s pravom možemo očekivati da će slično definirana veličina (poput ekscentriciteta) pokazivati slična svojstva za vrh koji je izvorište zaraze. Veličinu R definiramo na sljedeći način:

$$R_i(t) \equiv \frac{1}{N^I(t)} \sum_{j \in I(t)} d_{ij} \tag{5.1}$$

gdje je vrh j element skupa $I(t)$ zaraženih vrhova u trenutku t, a d_{ij} je duljina najkraćeg puta između između promatranog vrha i i zaraženog vrha j i $N^I(t)$ je ukupan broj zaraženih vrhova u trenutku t. Postavljamo uvjet da ako je $I(t) = \emptyset$ tada je $R_i(t) \equiv 0$. Ako vrhovi i i j nisu unutar povezane komponente d_{ij} nije dobro definiran, tj. divergira u beskonačnost, no da bismo izbjegli tu divergenciju d_{ij} u obzir uzimamo samo d_{ij} za koje postoji put, tj. za vrhove i i j koji su unutar

[11] Izvorište zaraze je najmanje udaljeno od svih zaraženih vrhova.

(a) Usmjereni lanac s 2 čvora.

(b) Usmjereni lanac s 3 čvora.

Slika 5.2: Kovergencija apsolutne razlike vjerojatnosti ostvarivanja pojedinih stanja između analitičkih i algoritamskih rezultata za različite brojeve ponavljanja algoritma na usmjerenom lancu s 2 i 3 vrha, za širenje zaraze do $T = 20$ diskretnih vremenskih koraka nakon početne zaraze uz vjerojatnost prijenosa zaraze $\mu = 0.3$ i vjerojatnost oporavka zaraženog vrha $\nu = 0.075$.

(a) Neusmjereni lanac s 2 čvora.

(b) Neusmjereni lanac s 3 čvora.

Slika 5.3: Kovergencija apsolutne razlike vjerojatnosti ostvarivanja pojedinih stanja između analitičkih i algoritamskih rezultata za različite brojeve ponavljanja algoritma na neumjerenom lancu s 2 i 3 vrha, za širenje zaraze do $T = 20$ diskretnih vremenskih koraka nakon početne zaraze uz vjerojatnost prijenosa zaraze $\mu = 0.3$ i vjerojatnost oporavka zaraženog vrha $\nu = 0.075$.

povezane komponente. Takvo rješenje problema divergencije je opravdano jer se zaraza ni ne može širiti u nepovezani dio grafa, te nam to ne utječe na problem detekcije izvora zaraze koji nas zanima. Druga stvar koja se može učiniti problematična u ovoj definiciji jest da je $R = 0$ u dva slučaja. Vrijednost $R = 0$ najprije u slučaju kada je zaražen samo vrh i u odnosu na koji promatramo udaljenost ostalih zaraženih vrhova, a zatim i u slučaju kada nema zaraženih vrhova. Budući da se ta dvoznačnost $R = 0$ javlja u različitim stadijima epidemije; prvi se slučaj najčešće javlja na početku epidemije, a drugi na kraju epidemije ako $\lambda = \mu/\nu < 1$, dvoznačnost vrijednosti $R = 0$ ne predstavlja problem u razumijevanju razvoja epidemije na mreži.

Za SIS očekujemo da će se vrijednost $R_i(t)$ povećavati do neke vrijednosti za vrhove koji su bliže izvoru jer će se zaraza sve više širiti dalje od izvora, dok će se za vrhove koji su dalje od izvora vrijednost $R_i(t)$ smanjivati jer će im se zaraza sve više približavati, odnosno vrijednosti d_{ij} u sumi će biti manje.

Ako je $\lambda > 1$ očekujemo da će nakon nekog vremena $R_i(t)$ saturirati na neku vrijednost R_{sat} jer će se uspostaviti ravnoteža između broja novooporavljenih vrhova i novozaraženih vrhova. To će u prosjeku uzrokovati jednaku vrijednost R_{sat} za sve vrhove i. Kažemo da je tada postignut endemski režim širenja zaraze.

Za $\lambda < 1$ vrijednost $R_i(t)$ će za sve vrhove i s vremenom trnuti u 0, odnosno zaraza će iščeznuti iz sustava.

Najprije simuliramo širenje zaraze po SIS modelu na mreži malog svijeta. Širenje zaraze simuliramo prethodno testiranim algoritmom[12], dok mrežu malog svijeta generiramo naredbom $networkx.watts_strogatz_graph(n, k, p)$ [25]. Veličina n označava broj vrhova u mreži, k broj najbližih susjeda s kojima je svaki vrh povezan prije premrežavanja [8], a p je vjerojatnost da se svaka od mogućih veza među vrhovima odspoji od jednog vrha i spoji se s nekim nasumično odabrani vrhom (vjerojatnost premrežavanja).

Simuliramo širenje zaraze na mreži malog svijeta na sljedeći način: prvo će se svaki nezaraženi vrh zaraziti s vjerojatnošću $1 - (1 - \mu)^b$, gdje je b ukupan broj najbližih zaraženih susjeda (s kojima je promatrani nezaraženi vrh direktno povezan), zatim će se svaki zaraženi vrh oporaviti sa vjerojatnošću ν. Nakon simuliranja širenja zaraze, računamo ranije definiranu veličinu $R_i(t)$ usrednjenu na 1000 ili 10000

[12]Detalji algoritma su u Dodatku.

simulacija.
U promatranim slučajevima i sama mreža malog svijeta može biti izvor stohastičnosti (vjerojatnost premrežavanja p) i širenje zaraze je stohastično (vjerojatnosti širenja zaraze μ i oporavka ν), pa je u svakom ponavljanju simulacije iznova generirana mreža malog svijeta i na njoj puštena simulacija širenja zaraze. Iako se može činiti da je takvm principom izgubljena točna informacija o utjecaju pojedinog izvora stohastičnosti na širenje zaraze na mreži, zapravo je napravljen bolji prosjek za dane parametre mreže ($N,\ p,\ k$) i zaraze ($\mu,\ \nu$). Važno je uočiti da je broj simulacija puno manji nego broj simulacija s kojim smo provjeravali točnost algoritma, a tomu je tako najprije jer račun potaje računalno vrlo zahtijevan. Osim toga, više nije nužno imati onoliko velik broj simulacija jer nekoliko različitih konfiguracija zaraženih vrhova daju istu vrijednost $R_i(t)$ pa je ukupan broja različitih vrijednosti $R_i(t)$ reduciran.

5.2.1 Utjecaj promjene broja najbližih susjeda k na veličinu $R_i(t)$

Na slici 5.4 prikazan je niz grafova razvitka $R_i(t)$ u vremenu na mreži malog svijeta s $N = 35$ vrhova i s vjerojatnošću premrežavanja $p = 0.01$ uz promjenu parametra k. Širenje zaraze je simulirano uz parametre: $\mu = 0.3,\ \nu = 0.075$. Zaraza se širi iz vrha 1 (izvorište zaraze), a vrijednosti svih $R_i(t)$ usrednjene su na rezultatima 1000 simulacija.
Najprije ćemo razmotriti općenita svojstva i ponašanje $R_i(t)$ za $i \in \{1, ..., N\}$ u ovisnosti o diskretnim vremenskim koracima t. Uočavamo da sve vrijednosti $R_i(t)$ nakon nekog vremena saturiraju na vrijednost R_{sat} u vremenu τ_{sat}. Budući da se u ovom konkretnom slučaju radi o parametrima epidemije za koje je $\lambda > 1$, epidemija ulazi u endemsko stanje u kojem se izmijenjuje otprilike podjednak broj zaraženih I i nezaraženih S vrhova među odjeljcima. Da su odabrani parametri za koje vrijedi da je $\lambda < 1$, prema definiciji $R_i(t)$ sve bi vrijednosti trnule u 0, budući da bi zaraza iščeznula iz mreže.
Drugo svojstvo koje možemo uočiti na grafovima jest da je većina vrijednosti $R_i(t)$ udvostručena. Uzrok udvostručenim vrijednostima $R_i(t)$ jest postojanje dvaju vrhova jednake udaljenosti od izvorišta zaraze, po jedan sa svake strane mreže gledano u odnosu na izvorište. Posebice za vjerojatnosti premrežavanja mreža malog svijeta $p \leq 0.1$, gotovo da je nemoguće razlikovati ”lijevu” ili ”desnu” stranu

(a) $k = 2$

(b) $k = 4$

(c) $k = 6$

(d) $k = 8$

Slika 5.4: Grafovi udaljenosti $R_i(t)$ u ovisnosti o t za različite brojeve najbližih susjeda k na mreži malog svijeta s $N = 35$ vrhova i sa vjerojatnosti premrežavanja $p = 0.01$.

mreže u odnosu na izvorište zaraze. Stoga na grafovima vidimo potpuno poklopljene ili vrlo slične vrijednosti za $R_{35}(t)$ i $R_2(t)$ ili za $R_{34}(t)$ i $R_3(t)$ i slično. Sljedeće svojstvo uočeno na grafovima, koje je među najvažnijima za ovaj rad, jest svojstvo da se $R_i(t)$ za polovicu mreže bližu izvoru zaraza povećava sve do R_{sat}, dok se $R_j(t)$ za polovicu mreže udaljeniju od izvorišta smanjuje sve do R_{sat}. Takvo je ponašanje u skladu s očekivanjem budući da se za vrhove koji su blizu izvorištu zaraza širi na vrhove sve udaljenije od njih te se $R_i(t)$ povećava, dok se za vrhove na drugom kraju mreže, zaraza njima sve više približava, u odnosu na početni trenutak kada im je bila maksimalno udaljena, pa se i $R_j(t)$ smanjuje prema R_{sat}. Možemo zaključiti da će vrhovi bliže izvorištu zaraze imati najveći pozitivni gradijent od $R_i(t)$, tj. $\nabla R_i(t)$. Gradijent će se sve više smanjivati i prelaziti u negativnu vrijednost što je promatrani vrh udaljeniji od izvora zaraze.

Na slici 5.4 prikazana su četiri grafa s vrijednostima $k \in \{2, 4, 6, 8\}$. Najlakše se uočava značajno smanjenje vremena τ_{sat} kao i promjena vrijednosti R_{sat} s promjenom broja najbližih susjeda k.[13]

Očekujemo da povećanje broja najbližih susjeda uzrokuje bolju povezanost mreže, a time i lakši prijenos zaraze na veći broj vrhova u jednom vremenskom koraka. Zbog toga se vrijeme τ_{sat} potrebno za postizanje R_{sat} smanjuje povećanjem broja najbližih susjeda k, a budući da je karakteristična duljina puta među vrhovima manja, jer ima više veza, i sama vrijednost R_{sat} se smanjuje. Za promjenu vrijednosti R_{sat} s promjenom broja najbližih susjeda k, zabilježene su sljedeće vrijednosti: za $k = 2$ vrijednost $R_{sat} \approx 8$, za $k = 4$ $R_{sat} \approx 4.5$ $k = 6$ je vrijednost $R_{sat} \approx 3$, dok je za $k = 8$ vrijednost $R_{sat} \approx 2.5$. Skala za vrijeme nije dovoljno precizna, pa je i teže odrediti vrijeme potrebno da mreža dođe u endemsko stanje, no bilježimo sljedeće vrijednosti: za $k = 2$ je otprilike $\tau_{sat} \approx 110$ diskretnih vremenskih koraka, za $k = 4$ potrbno je otprilike $\tau_{sat} \approx 20$ vremenskih koraka, za $k = 6$ svega $\tau_{sat} \approx 10$ vremenskih koraka, dok je za $k = 8$ nešto manje od 10, no nemoguće je reći točnu vrijednost jer skala nema potrebnu preciznost.

[13]Naredba $networkx.watts_strogatz_graph(n, k, p)$ generira mrežu malog svijeta samo za parne brojeve k ili za $k - 1$ ako je k neparan, budući da mogućnost crtanja mreže s neparnim brojem k u ovom konkretnom slučaju nije od posebnog značaja, tj. ponašanje takvih mreža se može pretpostaviti interpolacijom rješenja za mreže s parnim brojem k, primjere s neparnim brojem k se nije ni razmatralo.

5.2.2 Utjecaj promjene vjerojatnosti premrežavanja p na veličinu $R_i(t)$

Mreža malog svijeta s $N = 20$ vrhova

Na slici 5.5 prikazan je niz grafova razvoja $R_i(t)$ kroz vrijeme za različite vrijednosti vjerojatnosti premrežavanja p na mreži malog svijeta s $N = 20$ vrhova s $k = 4$ najbližih susjeda. Parametri širenja zaraze jednaki su za svih šest grafova i iznose: $\mu = 0.3$ i $\nu = 0.075$, tj. $\lambda = 4$, odnosno zaraza nakon nekog vremena ulazi u endemsko stanje. Sve vrijednosti $R_i(t)$ usrednjene su na 10000 simulacija.
Prema razmatranju u radu Wattsa i Strogatza o mrežama malog svijeta [8] očekujemo da će se povezanost grafa povećavati s povećanjem vjerojatnosti premrežavanja p, tj. karakteristična duljina puta $L(p)$ između bilo koja dva vrha u mreži će se smanjiti, a samim time i udaljenost bilo kojeg promatranog vrha od zaraženih vrhova će se smanjiti. Budući da je povezanost bolja i karakteristična duljina puta među vrhovima $L(p)$ manja za veće p, očekujemo da će se R_{sat} i τ_{sat} smanjiti, odnosno zaraza će se biti bliža svim vrhovima te će se efikasnije širiti.
Kako bismo što više usmjerili pozornost na utjecaj vjerojatnosti p na vrijednosti $R_i(t)$ na povezanoj mreži, potrebno je da broj najbližih susjeda bude barem jednak $k = 4$ kako se s povećanjem p mreža ne bi podijelila na nepovezane komponente.
Prema radu Wattsa i Strogatza [8] očekujemo da će za vjerojatnost premrežavanja $p = 0.01$ udaljenost $L(p)$ odstupati od vrijednosti $L(p = 1)$ za slučajni graf za maksimalno 25%. Od $p = 0.001$ počinje se uočavati smanjena početna udaljenost najudaljenijih vrhova od izvorišta zaraze, no ne značajna. Za $p = 0.1$ uočava se prvo značajno smanjenje vrijednosti R_{sat} uzrokovan smanjenjem prosječne udaljenosti među vrhovima, značajnija promjena bi se ranije primjetila da je mreža veća, odnosno da je utjecaj postojanja "prečaca" veći. Također, moguće je uočiti i smanjenje vremena potrebnog da sustav dođe u ravnotežu, odnosno τ_{sat}, ali i to smanjenje nije izričito značajno iz istog razloga kao i za R_{sat}.
Za $p = 0.01$ očitavamo vrijednost $R_{sat} \approx 3.5$ i $\tau_{sat} \approx 15$, za $p = 0.1$ vrijednost R_{sat} se značajnije smanjuje te ona iznosi $R_{sat} \approx 2.25$ dok je vrijeme potrebno da proces zaraze i oporavka dođu u ravnotežu $\tau_{sat} \approx 15$. Za slučajni graf $p = 1$ $R_{sat} \approx 2$ dok je $\tau_{sat} \approx 10$.

Mreža malog svijeta s $N = 35$ vrhova

Slično kao ranije opisano ponašanje mreže s $N = 20$ vrhova s obzirom na promjenu

(a) $p = 0$

(b) $p = 0.0001$

(c) $p = 0.001$

(d) $p = 0.01$

(e) $p = 0.1$

(f) $p = 1$

Slika 5.5: Grafovi udaljenosti $R_i(t)$ u ovisnosti o t za različite vrijednosti vjerojatnosti premrežavanja p za mrežu malog svijeta s $N = 20$ vrhova i sa $k = 4$ najbližih susjeda.

(a) $p = 0$

(b) $p = 0.0001$

(c) $p = 0.001$

(d) $p = 0.01$

(e) $p = 0.1$

(f) $p = 1$

Slika 5.6: Grafovi udaljenosti $R_i(t)$ u ovisnosti o t za različite vrijednosti vjerojatnosti premrežavanja p za mrežu malog svijeta s $N = 35$ vrhova i sa $k = 4$ najbližih susjeda.

vjerojatnosti premrežavanja p za $R_i(t)$ događa se i na mreži s $N = 35$ vrhova.
Na slici 5.6 prikazan je niz grafova razvoja $R_i(t)$ u vremenu na mreži s $N = 35$ vrhova i s $k = 4$ najbližih susjeda uz iste parametre širenja zaraze $\mu = 0.3$ i $\nu = 0.075$, ali se vjerojatnost premrežavanja p mijenja za svaki od grafova. Sve vrijednosti $R_i(t)$ usrednjene su na 1000 simulacija.
Budući da na ovim mrežama imamo ukupno više vrhova nego u prethodnom primjeru i sve su vrijednosti R_{sat} veće. Kao i u prethdnom primjeru jasnije promjene počinju se primjećivati tek iza $p = 0.01$. Za $p = 0.01$ vrijednost $R_{sat} \approx 4$ dok je vrijeme potrebno za postizanje ravnoteže $\tau_{sat} \approx 25$. Prvi sljedeći graf prikazuje ravoj $R_i(t)$ za $p = 0.1$ te je vrijednost $R_{sat} \approx 3$, a vrijeme $\tau_{sat} \approx 20$. Za slučajni graf $p = 1$ imamo $R_{sat} \approx 2.5$ dok je $\tau_{sat} \approx 12.5$ diskretnih vremenskih trenutaka.
Kao što smo ranije pretpostavili vidimo da je već za mrežu s $N = 35$ uočljiviji utjecaj vjerojatnosti p na vrijednosti R_{sat} i τ_{sat}.

5.2.3 Utjecaj promjene vjerojatnosti prijenosa zaraze između dva susjeda μ i vjerojatnosti oporavka ν zaraženog vrha na veličinu $R_i(t)$

Promjena vjerojatnosti oporavka zaraženog vrha ν

Na slici 5.7 prikazan je niz grafova razvoja $R_i(t)$ u vremenu za vjerojatnost prijenosa zaraze sa zaraženog na nezaraženi vrh $\mu = 0.3$ i za različite vrijednosti vjerojatnosti oporavka zaraženog vrha ν kako bismo ispitali utjecaj ν na vrijednosti $R_i(t)$. Svi grafovi napravljeni su simulacijama širenja zaraze na mreži malog svijeta s $N = 35$ vrhova, vjerojatnosti premrežavanja $p = 0.01$ i s $k = 2$ najbližih susjeda te su sve vrijednosti usrednjene na 1000 simulacija.
Očekujemo da će najveću ulogu u razvoju $R_i(t)$ u vremenu imati omjer $\mu/\nu = \lambda$, tj. ako je $\lambda < 1$ sve vrijednosti $R_i(t)$ će trnuti u 0, dok će za $\lambda > 1$ sve vrijednosti $R_i(t)$ saturirati na neku vrijednost R_{sat}. Vrijednost $\lambda = 1$ granica je između tih dvaju ponašanja vrijednosti $R_i(t)$, odnosno u ovom konkretnom slučaju ta se granica postiže za $\nu = 0.3$.
U nizu grafova uočavamo da povećanjem vrijednosti $\nu \in [0,\ 0.3\rangle$ da se τ_{sat} povećava, odnosno potrebno je dulje vrijeme da sustav uđe u enedemsko stanje zaraze budući da se oporavak puno brže događa, odnosno dulje se natječu procesi zaraze i oporavka prije ulaska u ravnotežu. Za granično stanje $\lambda = 1$, odnosno $\nu = 0.3$ vrijeme potrebno za saturacije divergira $\tau_{sat} \to \infty$. Također, uočavamo da se R_{sat}

(a) $\mu = 0.3$ *i* $\nu = 0$

(b) $\mu = 0.3$ *i* $\nu = 0.0375$

(c) $\mu = 0.3$ *i* $\nu = 0.075$

(d) $\mu = 0.3$ *i* $\nu = 0.15$

(e) $\mu = 0.3$ *i* $\nu = 0.4$

(f) $\mu = 0.3$ *i* $\nu = 0.5$

Slika 5.7: Grafovi razvoja $R_i(t)$ u vremenu za sve vrhove i i za različite vrijednosti vjerojatnosti oporavka ν, dok je vjerojatnost prenošenja zaraze ista $\mu = 0.3$ za sve grafove te su i parametri mreže isti na svim grafovima $N = 35$, $p = 0.01$ i $k = 2$.

snižava od vrijednosti $R_{sat} \approx 9$ do $R_{sat} \approx 7$ zato što se vrhovi brže oporavljaju, odnosno manje je vrhova zaraženo pa ih manje sudjeluje u ukupnom zbroju u R_{sat}. Za $\lambda < 1$, odnosno za $\nu > 0.3$ (grafovi (e) i (f) na slici 5.7), vrijednost $R_{sat} \to 0$ kao što i očekujemo prema definiciji $R_i(t)$. Povećanjem $\nu \in \langle 0.3,\ 1]$ vidimo da se τ_{sat} smanjuje jer proces oporavka zaraženih vrhova prije nadvladava proces zaraze, pa i zaraza brže iščezava iz sustava.

Promjena vjerojatnosti prijenosa zaraze između I i S vrha μ

(a) $\mu = 0.075$ *i* $\nu = 0.3$

(b) $\mu = 0.15$ *i* $\nu = 0.3$

(c) $\mu = 0.4$ *i* $\nu = 0.3$

(d) $\mu = 0.75$ *i* $\nu = 0.3$

Slika 5.8: Grafovi udaljenosti $R_i(t)$ u ovisnosti o t za različite vrijednosti vjerojatnosti prenošenja zaraze μ za vjerojatnost oporavka jednaku $\nu = 0.3$ za sve grafove, na mrežama malog svijeta s istim parametrima $N = 35$, $k = 2$ i $p = 0.01$.

Na slici 5.8 prikazan je niz grafova razvoja $R_i(t)$ u vremenu uz stalnu vrijednost vjerojatnosti oporavka $\nu = 0.3$ zaraženog vrha, a vrijednost vjerojatnosti prijenosa zaraze μ sa zaraženog na nezaraženi vrh se mijenja. Simulacije su rađene na mreži malog svijeta s $N = 35$ vrhova uz vjerojatnost premrežavanja $p = 0.01$ i s $k = 2$

najbližih susjeda, a sve vrijednosti $R_i(t)$ usrednjene su na 1000 simulacija.
Za utjecaj promjene vjerojatnosti μ ili ν na razvoj epidemije na mreži malog svijeta, najvažniji je omjer tih dvaju veličina, odnosno λ. Vjerojatnost $\mu \in [0,\ 0.3\rangle$ u ovom konkretnom slučaju daje vrijednost $\lambda < 1$, odnosno zaraza s vremenom iščezava. Povećavanjem vjerojatnosti μ od 0 do 0.3 produljuje se vrijeme potrebno da zaraza iščezne τ_{sat} jer je potrebno dulje vrijeme da proces oporavka nadvlada proces zaraze budući da se zaraza događa sa sve većom vjerojatnošću. Nakon što vrijednost vjerojatnosti μ preraste 0.3, zaraza zadovoljava uvjet $\lambda > 1$ potreban za ulazak u endemsko stanje zaraze. Što je veća vrijednost $\mu \in \langle 0.3,\ 1]$ vrijeme τ_{sat} je kraće, odnosno proces zaraze prije uđe u ravnotežu s procesom oporavka. Osim što se τ_{sat} smanji, vrijednost R_{sat} se poveća jer se više vrhova uspije zaraziti prije ulaska u ravnotežu dvaju procesa, pa više vrhova pridonosi vrijednosti $R_i(t)$.
Sveukupno, možemo zaključiti da je najvažniji odnos vjerojatnosti prijenosa zaraze μ i vjerojatnosti oporavka ν u omjeru λ. Taj omjer određuje hoće li proces zaraze i oporavka ući u ravnotežu za $\lambda > 1$ i vrijednost $R_i(t)$ saturirati na neku vrijednost R_{sat} za sve $i \in \{1, .., N\}$ ili će zaraza iščeznuti iz mreže i $R_i(t) \to 0$ za $\lambda < 1$.
Osim toga, za $\lambda > 1$, što je vrijednost λ veća (ili se vrijednost μ povećava, ili se vrijednost ν smanjuje) vrijednost R_{sat} je veća. R_{sat} raste povećanjem λ jer se više vrhova zarazi prije postizanja ravnoteže. Također, vrijeme τ_{sat} potrebno za postizanje ravnoteže između dva procesa kraće je jer se ravnoteža postiže brže budući da se zaraza brže proširi cijelom mrežom.
Ako je $\lambda < 1$, tada će se smanjivanjem vrijednosti $\lambda \to 0$ i vrijeme potrebno da zaraza potpuno iščezne iz sustava smanjivati jer će proces oporavka puno brže nadvladati proces prijenosa zaraze. Vrijednost λ se smanjuje ili smanjivanjem vjerojatnosti prijenosa zaraze μ ili povećavanjem vjerojatnosti oporavka ν. Budući da vrijednost λ sadrži najvažnije informacije za razvoj zaraze na mreži, dalje ćemo razmatrati samo utjecaj promjene vrijednosti λ na $\nabla R_i(t)$.
Za kraj, možemo zaključiti da je vrijednost $R_i(t)$ najmanja za izvorište zaraze i raste prema R_{sat} za sve kombinacije vjeojatnosti p, broja najbližih susjeda k i omjera epidemijskih parametara λ. Svaki od navedenih parametara uzrokuje drugačiju promjenu na $R_i(t)$ koji se mogu sažeti u promjenu povezanosti mreže (p i k), ukupnog broja dostupnih veza (k) te promjenu značaja procesa širenja zaraze ili oporavka za stanje epidemije na mreži malog svijeta. Za $\nabla R_i(t)$ možemo očekivati

da će biti najveći za izvorište zaraze te da će navedeni parametri na sličan način utjecati na razvoj $\nabla R_i(t)$ u vremenu.

5.3 Kvalitativna svojstva ∇R na mreži malog svijeta

U prethodnom smo potpoglavlju vidjeli da $R_i(t)$ za izvorišni vrh, kao i za vrhove koji su mu najbliži, raste iz najmanjih vrijednosti prema R_{sat}, odnosno imaju pozitivan gradijent. Još preciznije, očekujemo da će gradijent $\nabla R_i(t)$ biti najveći za izvorište zaraze, što ćemo i pokazati u ovom potpoglavlju. Također ćemo u ovom potpoglavlju i razmotriti koja su svojstva $\nabla R_i(t)$ s obzirom na promjenu svojstva širenja zaraze λ te svojstava mreže k i p.

5.3.1 Utjecaj promjene broja najbližih susjeda k na $\nabla R_i(t)$

Na slici 5.9 prikazan je utjecaj promjene broja najbližih susjeda k na ovisnost $\nabla R_i(t)$ o vremenu. Napravljeno je 1000 simulacija na mrežama malog svijeta s $N = 35$ vrhova s vjerojatnošću premrežavanja $p = 0.01$ Isto kao u prethodnom potpoglavlju simuliramo širenje zaraze iz vrha 1 za $T = 80$ diskretnih vremenskih koraka uz vjerojatnost prijenosa zaraze $\mu = 0.3$ i vjerojatnost oporavka zaraženog vrha $\nu = 0.075$. U svim grafovima odrezana su prva dva trenutka budući da oni daju puno veće vrijednosti $\nabla R_i(t)$ što je izričito uzrokovano definicijom $R_i(t)$ i $\nabla R_i(t)$, tj. ne nosi nikakvu važnu informaciju.
Najprije i najvažnije što uočavamo jest da je gradijent $R_{source}(t)$ (za izvor zaraze, na grafovima jedina puna linije) zaista najveći. Za razliku od $R_i(t)$, vrijednost $\nabla R_i(t)$ usrednjena na 1000 pokazuje veće oscilacije u vrijednostima $\nabla R_i(t)$ za susjedne trenutke, no prepoznatljiva su općenita svojstva funkcije.
Vrijednost gradijenta je do saturacije najveća za izvorište zaraze (vrh 1), a zatim i za njegove najbliže susjedne vrhove (2 i 35). Što je veći broj najbližih susjeda k, povećava se i broj vrhova čiji je gradijent iznosom vrlo sličan $\nabla R_{source}(t)$, no budući da ipak postoji određena vjerojatnost premrežavanja $p = 0.01$ te vrijednosti nisu identične za prvih $2,\ 4,\ 6$ ili 8 susjednih vrhova. Očekujemo da će za vrijednost R_{sat} sve vrijednosti $\nabla R_i(t)$ biti jednake nuli, što vidimo da se događa na grafovima. Kao i u prethodnom potpoglavlju, vrijeme potrebno za postizanje saturacije τ_{sat}[14] se

[14]Ovaj τ_{sat} za $\nabla R_i(t)$ nije identičan τ_{sat} za $R_i(t)$, no ima istu funkciju, odnosno označava trenutak u kojem su sve vrijednosti $\nabla R_i(t)$ približno jednake.

(a) $k = 2$

(b) $k = 4$

(c) $k = 6$

(d) $k = 8$

Slika 5.9: Grafovi $\nabla R_i(t)$ u ovisnosti o t za različite brojeve najbližih susjeda k na mreži malog svijeta s $N = 35$ vrhova i sa vjerojatnosti premrežavanja $p = 0.01$; parametri zaraze su $\mu = 0.3$ i $\nu = 0.075$.

smanjuje povećavanjem broja najbližih susjeda k. Za $k = 2$ vrijeme potrebno za saturaciju τ_{sat} premašuje 80 diskretnih vremenskih koraka. Za $k = 4$ vrijeme potrebno za saturaciju se značajno smanjuje te je ono $\tau_{sat} \approx 25$ diskretnih vremenskih koraka, dok je za $k = 6$ vrijeme τ_{sat} nešto dulje od 10 vremenskih koraka, a za $k = 8$ ono je nešto kraće od 10 diskretnih vremenskih koraka. Funkcije $\nabla R_i(t)$ za veće vrijednosti k pokazuju manje oscilacije u vrijednosti $\nabla R_i(t)$ za susjedne trenutke, a uzrok tome jest što postoji veći broj veza u mreži, pa u ukupnom rezultatu vrijednosti $R_i(t)$ ima više različitih konfiguracija zaraze na mreži koji daju jednaku vrijednost $R_i(t)$ koja se manje razlikuje u trenutcima t i $t - 1$. Možemo zaključiti da se gradijent funkcije $R_i(t)$ ponaša kako bismo očekivali, odnosno iznos gradijenta najveći je za izvor zaraze, kao i za njegove najbliže susjede čiji se broj povećava s brojem k. Slično kao i za $R_i(t)$, povećanje broja k utječe na smanjenje τ_{sat} te su manje oscilacije između vrijednosti $\nabla R_i(t)$ za susjedne trenutke.

5.3.2 Utjecaj promjene vjerojatnosti premrežavanja p mreže malog svijeta na $\nabla R_i(t)$

Na primjeru utjecaja promjene vjerojatnosti premrežavanja p na vrijednost $R_i(t)$ vidjeli smo da povećanje p uvjetuje smanjenje vremena potrebnog za postizanje ravnoteže τ_{sat} kao i na smanjenje same vrijednosti R_{sat} jer se karakteristična udaljenost među vrhovima $L(p)$ smanjuje. Slično ponašanje očekujemo i od $\nabla R_i(t)$. Na slici 5.10 prikazani su grafovi ovisnosti $\nabla R_i(t)$ o vremenu za vrijednosti vjerojatnosti premrežavanja $p \in \{0,\ 0.0001,\ 0.001,\ 0.01,\ 0.1,\ 1\}$. Broj najbližih susjeda $k = 4$, broj vrhova u mreži $N = 35$ i parametar širenja zaraze $\lambda = 4.0$ jednaki su za sve grafove. Sve vrijednosti $\nabla R_i(t)$ usrednjene su na 1000 simulacija na mrežama s danim svojstvima. Kao i u slučaju $R_i(t)$ nije vidljiv značajniji utjecaj vjerojatnosti p prije $p = 0.01$. Najviše se razlike uočava u povećanju početnih vrijednosti $\nabla R_i(t)$ za najudaljenije vrhova (negativne vrijednosti $\nabla R_i(t)$). Vrijednost $R_i(t)$ za udaljenije vrhove od izvorišta u početku je to bliža konačnoj vrijednosti R_{sat} što je veća vjerojatnost p, pa je i $|\nabla R_i(t)|$ manja. Vrijeme potrebno za postizanje ravonteže τ_{sat} za $p = 0.01$ je $\tau_{sat} \approx 25$, dok je za vjeorjatnost $p = 0.1$ $\tau_{sat} \approx 20$, dok je za slučajni graf $p = 1$ potrebno $\tau_{sat} \approx 15$ diskretnih vremenskih koraka. Zaključujemo da za $p < 0.01$ vjerojatnost premrežavanja nema toliko značajnu ulogu u razvoju $\nabla R_i(t)$ na mreži malog svijeta s $N = 35$ vrhova.

(a) $p = 0$

(b) $p = 0.0001$

(c) $p = 0.001$

(d) $p = 0.01$

(e) $p = 0.1$

(f) $p = 1$

Slika 5.10: Grafovi $\nabla R_i(t)$ u ovisnosti o t za različite vrijednosti vjerojatnosti premrežavanja p za mrežu malog svijeta s $N = 35$ vrhova i sa $k = 4$ najbližih susjeda; parametri zaraze su $\mu = 0.3$ i $\nu = 0.075$.

5.3.3 Utjecaj parametara širenja zaraze λ na $\nabla R_i(t)$

Omjer parametara širenja zaraze $\lambda > 1$

(a) $\lambda = 40$

(b) $\lambda = 4$

(c) $\lambda = 3$

(d) $\lambda = 1.5$

Slika 5.11: Grafovi $\nabla R_i(t)$ u ovisnosti o t za različite omjere vjerojatnosti prijenosa zaraze μ i vjerojatnosti oporavka zaraženog vrha ν; $\lambda = \mu/\nu > 1$ u endemskom režimu. Sve simulacije rađene su na mreži malog svijeta s $N = 35$ vrhova, $p = 0.01$ i $k = 4$.

Na slici 5.11 prikazana su četiri grafa ovisnosti $\nabla R_i(t)$ o vremenu na mrežama s $N = 35$ vrhova, $k = 4$ najbliža susjeda i vjerojatnošću premrežavanja $p = 0.01$. Sve vrijednosti $\nabla R_i(t)$ usrednjene su na 1000 simulacija na mrežama s danim parametrima za četiri različita parametra λ (četiri grafa). Budući da smo na primjeru utjecaja vjerojatnosti širenja zaraze μ i vjerojatnosti oporavka ν na

vrijednosti $R_i(t)$ zaključili da je najvažniji omjer dvije vjerojatnosti $\mu/\nu = \lambda$, a ne toliko konkretne vrijednosti μ ili ν, u ovim se grafovima prati samo promjena parametra λ. U sva četiri grafa $\lambda > 1$, što znači da se epidemija širi u endemskom režimu, tj. nakon nekog vremena τ_{sat} sve vrijednosti $R_i(t)$ postižu vrijednost R_{sat}. Prvo što uočavamo jest da su funkcije oscilacije između vrijednosti $\nabla R_i(t)$ za susjedne trenutke to manje što je λ veća. Slično kao u grafovima s promjenom broja najbližih susjeda k, uzrok manjih oscilacija između vrijednosti $\nabla R_i(t)$ i $\nabla R_i(t-1)$ jest što za veće vrijednosti λ proces širenja zaraze nadmašuje proces oporavka pa se više vrhova stigne zaraziti prije τ_{sat}, tj. više različitih konfiguracija zaraze na mreži doprinosi istom stanju $R_i(t)$ što smanjuje oscilacije između $\nabla R_i(t)$ i $\nabla R_i(t-1)$ za sve $t > 0$. Što je manja i bliža vrijednost parametra λ kritičnoj vrijednosti $\lambda_{crit} = 1$, to je teže odrediti vrijeme potrebno za saturaciju svih vrijednosti $\nabla R_i(t)$ na nulu, što je u skladu s razmatranjem o utjecaju λ na $R_i(t)$. Budući da se za manje vrijednosti λ dulje natječu procesi širenja zaraze i oporavka više različitih vrhova na različitim udaljenostima od izvora zaraze se oporavlja i zaražava što uzrokuje da je više vremena potrebno da se $\nabla R_i(t)$ stabilizira na nuli.

Omjer parametara širenja zaraze $\lambda < 1$

Na slici 5.12 prikazana su četiri grafa ovisnosti $\nabla R_i(t)$ o vremenu za četiri različite vrijednosti $\lambda \leq 1$. Sva četiri grafa napravljena su na mrežama malog svijeta s $N = 35$ vrhova, vjerojatnošću $p = 0.01$ i s $k = 4$ najbliža susjeda. Vrijednosti $\nabla R_i(t)$ usrednjene su na 1000 simulacija. Kao i za $\lambda > 1$ uočavamo da postoje veće oscilacije u vrijednostima $\nabla R_i(t)$ za susjedne trenutke što je manja vrijednost λ. Za razliku od $R_i(t)$, režim $\lambda < 1$ u kojem $R_i(t)$ s vremenom trne u nulu nije prepoznatljiv prema značajnom promjenu ponašanja grafa. Ovaj režim pokazuje puno veće razlike u vrijednosti $\nabla R_i(t)$ za susjedne trenutke što znači da 1000 simulacija ne pokriva savršeno sve moguće konfiguracije $\nabla R_i(t)$ što smo objasnili i u ranijim primjerima. Naposljetku, možemo zaključiti da bez obzira na vrijednot vjerojatnosti p, parametra λ ili broja najbližih sujseda k, $\nabla R_i(t)$ zaista je najveći za izvorište zaraze. Sada još preostaje provjeriti vrijedi li isto svojstvo i na stvarim mrežama.

(a) $\lambda = 1$

(b) $\lambda = 0.75$

(c) $\lambda = 0.6$

(d) $\lambda = 0.4$

Slika 5.12: Grafovi $\nabla R_i(t)$ u ovisnosti o t za različite omjere vjerojatnosti prijenosa zaraze μ i vjerojatnosti oporavka zaraženog vrha ν; $\lambda = \mu/\nu \leq 1$ što nije endemski režim. Sve simulacije rađene su na mreži malog svijeta s $N = 35$ vrhova, $p = 0.01$ i $k = 4$.

Slika 5.13: Shematski prikaz mreže s $N = 279$ vrhova i $E = 2287$ veza među vrhovima što odgovara mreži živčanog sustava oblića Caenorhabditis elegans; na slici je prikazano stanje širenja živčanog impulsa u trenutku $T = 8\ ms$ gdje su crvenom bojom označeni živci koji se još oporavljaju od podražaja (I), a sivo označeni su živci oporavljeni od podražaja i podložni novom podražaju S.

5.4 *Ponašanje ∇R na živčanom sustavu oblića Caenorhabditis elegans*

U radu Wattsa i Strogatza [8] pokazano je da je živčani sustav oblića C. elegans dobro opisan mrežom malog svijeta, što nam je već dovoljno dobra motivacija za odabir baš te realne mreže. SIS model širenja zaraze u ovom bi slučaju opisivao širenje živčanih signala na živčanom sustavu oblića, gdje bi stanje S bio živac prije primanja signala ili nakon otpuštanja signala, dok bi stanje I bio živac na kojem se u tom trenutku živčani signal nalazi. Živčani sustav C. elegansa sastoji se od $N = 279$ živčanih stanica, tj. vrhova u mreži, te $E = 2287$ veza među stanicama.

Slika 5.14: Graf ovisnosti $\nabla R_i(t)$ o vremenu pri širenju živčanih podražaja na mreži živčanog sustava oblića C. elegans prema SIS modelu uz $\mu = 0.3$ što je vjerojatost prijenosa signala s aktiviranog živca (I) na neaktivnog S u $1\ ms$ i uz $\nu = 0.01$ vjerojatnost da se živac oslobodi signala u $1\ ms$; na grafu su prikazani samo $\nabla R_i(t)$ za 29 vrhova radi preglednosti grafa.

Ova mreža nije usmjerena niti veze imaju težinu, što je čini jednostavnijom za razmatranje. Mreža živčanog sustava oblića C. elegans potpuno je opisana i mapirana mreža živčanog sustava [26] te su je u obliku kompleksne mreže u svom radu generirali Watts i Strogatz. Podatci o mreži strukturirani su u obliku tekstualne

datoteke u dva stupca u koje su zapisani povezani vrhovi (tj. zapisane su sve veze u mreži) te je iz tih podataka moguće generirati mrežu, a preuzeti su iz dodatnih materijala knjige "Complex Networks: Principles, Methods and Applications" [27].
Na slici 5.13 prikazana je shema mreže živčanog sustava oblića C. elegans u trenutku $T = 8\ ms$ nakon početnog pojavljivanja signala. Crveni vrhovi označuju živce (I) na kojima se trenutno nalazi živčani signal, a sivi vrhovi označuju vrhove koji su spremni primiti novi signal (S). Shema ne prikazuje stvaran položaj živaca u obliću već prikazuje odnose između živaca u obliću.
Na slici 5.14 prikazana je ovisnost $\nabla R_i(t)$ o vremenu. Vrijednosti $\nabla R_i(t)$ usrednjene su na 100 simulacija širenja zaraze na identičnoj mreži i s jednakim početnim uvjetima (vrh 1 je zaražen u trenutku $t = 0\ ms$). Širenje zaraze simulirano je uz vjerojatnost $\mu = 0.3\ ms^{-1}$ prenošenja podražaja u jednoj milisekundi [15] te vjerojatnost $\nu = 0.01\ ms^{-1}$ otpuštanja podražaja [28].
Prvo što uočavamo na grafu je $\tau_{sat} \approx 6\ ms$, što znači da je već za $6\ ms$ živčani podražaj ravnomjerno raspodijeljen po cijelom živčanom sustavu oblića C. elegans. Lako uočavamo da vrijednost $R_{source}(t)$ nije najveća, ali je među najvećima. U tablici

Udaljenost od izvora	vrhovi
1	44, 160, 5, 13, 73, 11
2	7, 2, 22, 48, 115, 75, 21, 114, 50, 6, 19, 17, 43, 77, 46, 47, 42, 12, 78, 40, 10, 74, 9, 18, 41
3	49, 20

Tablica 5.3: Popis vrhova čija je vrijednost $\nabla R_i(2\ ms)$ veća od vrijednosti $\nabla R_1(2\ ms)$ za izvor, te njihova udaljenost od izvora zaraze.

5.3 zapisani su vrhovi čija je vrijednost $\nabla R_i(2\ ms)$ u trenutku $t = 2\ ms$ veća od $\nabla R_1(2\ ms)$, tj. vrijednosti gradijenta za izvor zaraze. U tablici su vrhovi podijeljeni po udaljenosti od izvora zaraze. Najviše vrhova s vrijednosti $\nabla R_i(2\ ms)$ većom od vrijednosti za izvor zaraze je na udaljenosti 2 od izvora zaraze, odnosno oni su drugi susjedi izvoru zaraze.
Na slici 5.15 je prikazana ovisnost $\nabla R_i(t)$ samo za vrhove i čija je vrijednost gradijenta u trenutku $t = 2\ ms$ veća od vrijednosti gradijenta za izvor zaraze kako bi se jasnije vidio odnos tih gradijenta tijekom vremena. Uočavamo da gradijent za

[15] Zbog nedostatka traženih podataka za oblića C. elegans, radi ilustracije korišten je podatak preuzet sa stranice https://aiimpacts.org/rate-of-neuron-firing/ dan za čovjeka. Kako bi živac bio spreman primiti novi podražaj potrebno mu je $3 - 4\ ms$ (eng. *Absolute Refractory Period*, ARP).

Slika 5.15: Graf ovisnosti $\nabla R_i(t)$ o vremenu pri širenju živčanih podražaja na mreži živčanog sustava oblića C. elegans prema SIS modelu uz $\mu = 0.3$ što je vjerojatost prijenosa signala s aktiviranog živca (I) na neaktivnog S u $1\ ms$ i uz $\nu = 0.01$ vjerojatnost da se živac oslobodi signala u $1\ ms$; na grafu su prikazani samo $R_i(t)$ za vrhove i čija je vrijednost $\nabla R_i(t = 2\ ms)$ u trenutku $t = 2\ ms$ veća od vrijednosti $\nabla R_1(t)$.

izvor zaraze nije u svim trenutcima t manji od svih promatranih vrhova, nego ih u nekim trenutcima premašuje, što nam ukazuje da ipak gradijent izvora zaraze ima značajnu uogu među najvećim vrijednostima gradijenta. To potvrđuje našu pretpostavku da izvor zaraze ima jednu od najvećih vrijednosti gradijenta $R_i(t)$.

5.5 Ponašanje ∇R na Zacharyjevoj mreži karate kluba

Slika 5.16: Shematski prikaz mreže s $N = 34$ vrhova i $E = 78$ veza među vrhovima koji označavaju članove sveučilišnog karate kluba te njihova međusobna prijateljstva; na slici je prikazano stanje širenja zaraze u trenutku $T = 8\ dana$ gdje su crvenom bojom označeni zaraženi članovi (I), a sivo označeni su članovi podložni (ponovnoj) zarazi S.

Drugi primjer stvarne mreže koju ćemo promotriti je Zacharyjeva mreža sveučilišnog karate kluba. To je socijalna mreža koju je proučavao Wayne W. Zachary u razdoblju od 1970. do 1972. [29] Mreža sadrži $N = 34$ člana karate kluba s ukupno $E = 78$ veza među njima koje označuju prijateljstvo među članovima. Mreža je neusmjerena, a veze nemaju težinu što čini razmatranje jednostavnijim. Ova se mreža često koristi u proučavanju dinamike kompleksne mreže na primjeru problema da dođe do svađe dvaju trenera i klub se razdvoji na dva. Ovaj skup podataka zapisan je u tekstualnu datoteku u obliku dva stupca koja

Slika 5.17: Graf ovisnosti $\nabla R_i(t)$ o vremenu pri širenju zaraze na Zacharyjevoj mreži sveučilišnog karate kluba prema SIS modelu uz $\mu = 0.81$ vjerojatnost prijenosa zaraze u jednom danu i $\nu = 1/14$ vjerojatnost oporavka jednom danu; na grafu je označen redoslijed $\nabla R_i(2\ dana)$.

označuju sve veze unutar mreže. Podatci su preuzeti iz dodatnih materijala knjige "Complex Networks: Principles, Methods and Applications". [27]

Na mreži ćemo proučiti primjer širenja zaraze koronavirusa. Shema kompleksne mreže dobivene iz podataka prikazana je na slici 5.16 uz stanje zaraze u trenutku $T = 8\ dana$ nakon pojave zaraze. Crveni vrhovi označuju zaražene (I) članove kluba, a sivi članove kluba prodložne (ponovnoj) zarazi (S). Važno je uočiti da shema nema za cilj prikazati stvarnu prostornu raspodjelu članova karate kluba već samo prijateljstva među članovima.

Na mreži je simulirano širenje zaraze koronavirusa uz $\mu = 0.18$ novozaraženih u danu[16] te $\nu = 1/14$[17] vjerojatnost oporavka zaraženog vrha od trenutka kontakta prilikom kojeg je prenesana zaraza u jednom danu. Vrijednosti $\nabla R_i(t)$ dobivene su usrednjavanjem 1000 simulacija na identičnoj mreži i s jednakim početnim uvjetima.

[16]Podatak je temeljen na omjeru novozaraženih pacijenata te ukupnog broja zaraženih 10. 09. 2021. iz podataka dostupnih na stranici https://www.koronavirus.hr/koronavirus-statisticki-pokazatelji-za-hrvatsku-i-eu/901.

[17]Podatak temeljen na informacijama na stranici https://www.webmd.com/lung/covid-recovery-overview#2.

Ovisnost $\nabla R_i(t)$ o vremenu prikazana je na slici 5.17. Radi lakšeg raspoznavanja, jedino je $\nabla R_1(t)$ prikazan punom linijom. Na grafu su prikazane vrijednosti $\nabla R_i(t)$ za sve vrhove $i \in \{1, \ ..., \ 34\}$.
Na slici 5.17 moguće je uočiti da izvor zaraze (vrh 1) nema najveći gradijet u svim trenutcima, ali uvijek je gradijent izvora zaraze među najvećim vrijednostima. U trenutku $t = 6 \ dana$ vrhovi $i \in \{2, \ 12, \ 18, \ 13, \ 11, \ 5, \ 22\}$ imaju vrijednost $\nabla R_i(t)$ veću od vrijednosti $R_1(t)$. Svi ti vrhovi su prvi susjedi vrhu 1, što je moguće ustvrditi uvidom u skup podatak iz kojeg je generirana mreža. Taj rezultat podsjeća na rezultat uspješnosti detekcije izvora zaraze u radu s Jordanovim centrom zaraze [17]. Zaključujemo da će na stvarnoj mreži poput Zacharyheve mreže karate kluba iznos $\nabla R_i(t)$ biti najveći za vrhove koji su najbliži susjedi izvoru zaraze, što znači da je raspodjela procjenjenih rješenja centrirana oko izvora zaraze sa širinom udaljenosti 1. Za razliku od prethodnog primjera, ovdje se radi o osam puta manjoj mreži, pa je i raspodjela vrhova s najvećom vrijednosti $\nabla R_i(t)$ uža i bliža stvarnom izvoru zaraze.

5.6 Zaključak rezultata

U prvom potpoglavlju pokazana je točnost algoritamskih rješenja usporedbom s analitičkim rješenjima za širenje zaraze na usmjerenim i neusmjerenim lancima s dva ili tri vrha. U drugom potpoglavlju proučene su karakteristike prosječne udaljenosti promatranog vrha od svih zaraženih vrhova $R_i(t)$ s obzirom na promjene parametara mreže malog svijeta ili širenja zaraze. Može se zaključiti da najznačajniju promjenu u $R_i(t)$ unosi broj prvih susjeda k s kojim je povezan svaki vrh, a za krajnji ishod zaraze najznačajniji omjer vjerojatnosti širenja zaraze i oporvaka od zaraze $\lambda = \mu/\nu$. U skladu s očekivanjem prema radu [17], pokazano je da je najmanja vrijednost $R_i(t)$ za izvor zaraze na mreži malog svijeta te je već napravljena pretpostavka da za $\nabla R_i(t)$ očekujemo najveću vrijednost za izvor zaraze. U trećem potpoglavlju, dokazana je hipoteza da je vrijednost $\nabla R_i(t)$ najveća na mreži malog svijeta za izvor zaraze neovisno o parametrima mreže malog svijeta i širenja zaraze. Budući da smo dokazali da je vrijednost $\nabla R_i(t)$ najveća za izvor zaraze potrebno je provjeriti vrijedi li isti zaključak na stvarnim mrežama, poput mreže živčanog sustava oblića C. elegans ili mreže prijateljstva među članovima

karate kluba. Na tim stvarnim mrežama pokazano je da najveću vrijednost $\nabla R_i(t)$ imaju vrhovi koji su najbliži susjedi izvoru zaraze, no nije nužno vrijednost najveća za izvor zaraze.

6 Zaključak

U ovome radu sažeti su i obrađeni najvažniji doprinosi epidemiološkom modeliranju. Predstavljen je model odjeljaka, posebice SIS model širenja zaraze koji je specifičan budući da je nemoguće razlikovati jedinke koje nikad nisu bile zaražene od onih koje su se oporavile od zaraze te su ponovno podložne zarazi. Osim što se u SIS modelu gubi informacija o prethodnom stanju zaraze na pojedinim jedinkama, SIS model je prepoznatljiv po svom ponašanju u limesu $t \to \infty$. Ovisno o omjeru vjerojatnosti širenja zaraze između zaražene i nezaražene jedinke μ i vjerojatnosti oporavka zaražene jedinke ν u jedinici vremena, zaraza na razini sustava za $\lambda = \mu/\nu > 1$ ulazi u endemsko stanje ili iščezava iz sustava za $\lambda = \mu/\nu < 1$. Endemsko stanje zaraze znači da u nekom trenutku postaje podjednak broj jedinki koje se u jedinici vremena zaraze kao i onih koje se u jedinici vremena oporave od zaraze.

U epidemiloškom modeliranju, koncept kompleksnih mreža značajno je povećao kvantitativnu točnost predviđanja daljnjeg razvoja zaraze te je i olakšao praćenje zaraze na lokalnoj razini koja ima odjek na globalno stanje zaraze. Zbog takve važnosti kompleksnih mreža u epidemiološkom modeliranju, u ovom su radu sažeti najvažniji pojmovi i svojstva raznih mreža, te su detaljnije razmotrene mreže malog svijeta i mreže bez skale koje su pokazala vrlo veliku točnost u opisivanju stvarnih mreža.

Problem detekcije izvora zaraze privukao je značajnu pozornost u području istraživanja širenja zaraze na kompleksnim mrežama. U ovome radu prezentirani su najvažniji rezultati i algoritmi korišteni za rješavanje problema detekcije izvora zaraze. Posebno se ističu algoritmi detekcije izvora zaraze pomoću Jordanovih centara, zaključivanje temeljeno na jednadžbama dinamičkog prenošenja poruke te jednostavno određivanje izvora zaraze provođenjem Monte Carlo i Soft Margin simulacija.

Pomoću *Netwokx* paketa specijaliziranog za izradu, prikaz i proučavanje kompleknih mreža unutar pythona napravljen je algoritam širenja zaraze prema SIS modelu iz jednog izvora. Napravljena je provjera točnosti algoritma s obzirom na analitička rješenja za usmjerene i neusmjerene lance s dva i tri vrha te je pokazano odlično slaganje. To slaganje daje nam pouzdanost da možemo algoritam koristiti za

simulaciju širenja zaraze na mreži malog svijeta. Motivirani radom [17] u ovom se diplomskom radu promatra prosječna udaljenost svih zaraženih vrhova od nekog vrha i $R_i(t)$ na mreži malog svijeta. Također se promatra ponašanje $R_i(t)$ s obzirom na promjenu parametara mreže malog svijeta (broj najbližih susjeda k, vjerojatnost premrežavanja p) te promjenu parametara širenja zaraze (vjerojatnost prijenosa zaraze μ i vjerojatnost oporavka zaraženog vrha ν). Na ponašanje $R_i(t)$ najznačajnije utječe broj najbližih susjeda k, čijim se povećanjem poboljšava povezanost mreže i olakšava širenje zaraze. Asimptotsko ponašanje $R_i(t)$ određuje omjer $\mu/\nu = \lambda$, tj. određuje hoće li zaraza postati endemska ili će potpuno iščeznuti. Vrijednost $R_i(t)$ je najmanja za izvor zaraze te sve vrijednosti $R_i(t)$ teže nekoj asimptotskoj vrijednosti R_{sat}. Dakle, nameće se hipoteza da će vrijednost $\nabla R_i(t)$ biti najveća za izvor zaraze na mreži malog svijeta. Proučavanjem ponašanja $\nabla R_i(t)$ u vremenu s obzirom na promjenu parametara mreže i omjera λ, uočava se očekivano ponašanje, tj. $\nabla R_i(t)$ jest najveći za izvor zaraze.
Budući da je dokazano da $\nabla R_i(t)$ ima najveću vrijednost na mreži malog svijeta, bilo je potrebno provjeriti vrijedi li isti zaključak na nekim stvarnim mrežama. Takva provjera napravljena je na mreži živčanog sustava oblića C. elegans, gdje je SIS model opisivao širenje živčanog podražaja, te na Zacharyjevoj mreži prijateljstava među članovima karate kluba, gdje je SIS model opisivao širenje koronavirusa prema aktualnim parametrima. Na Zacharyjevoj mreži karakte kluba $\nabla R_i(t)$ je najveći za vrhove koji su prvi susjedi izvoru zaraze, što približno odgovara pretpostavci da će vrijednost gradijenta biti najveća za izvor zaraze. Na mreži živčanog sustava oblića C. elegans, budući da je ona veća, gradijent ima najveću vrijednost većinom za druge susjede izvoru zaraze, odnosno točnost hipoteze se smanjuje povećanjem mreže.
Prednost ovakve metode detekcije izvora zaraze jest da je dovoljno imati dvije slike trenutnog stanja zaraze na mreži kako bi se mogao odrediti gradijent te iz njega s dosta dobrom točnošću odrediti koji bi vrh, tj. skup vrhova, mogao biti izvor zaraze. Nedostatak ovakve metode jest da se točnost smanjuje povećanjem mreže. Osim toga, nakon vremena τ_{sat}, potrebnog za postizanje ravnoteže procesa zaraze i procesa oporavka, teško je razlučiti vrijednosti $\nabla R_i(t > \tau_{sat})$ za različite vrhove.

Dodaci

Dodatak A Python kod algoritamske implementacije širenja zaraze na kompleksnim mrežama

A.1 Paketi korišteni u simulacijama

```
import matplotlib.pyplot as plt
import networkx as nx
from random import random
import numpy as np
```

A.2 Funkcija za generiranje mreže malog svijeta i simulaciju širenja zaraze na toj mreži

```
def SIS_SmallWorld(T, N, pInfect, pRecover, k, p):
#T - ukupan broj vremenskih koraka simulacije
#N - broj vrhova mreže
#pInfect - vjerojatnost prijenosa zaraze
#pRecover - vjerojatnost oporavka zaraženog vrha
#k - broj najbližih susjeda (parametar mreže malog svijeta)
#p - vjerojatnost premrežavanja mreže malog svijeta

    #generiranje mreže malog svijeta
    G=nx.watts_strogatz_graph(N,k,p)

    #2D polje - spremanje R za svaki vrh u t<=T
    R=[[0 for i in range(N)] for dt in range(T)]

    #početni uvjeti
    for i in nx.nodes(G):
        G.nodes[i]['infection']=0
    G.nodes[0]['infection']=1
```

```
for dt in range(1,T):
    #pomoćna lista - stanja zaraze u trenutku t-1
    R_probe = [G.nodes[i]['infection'] for i in nx.nodes(G)]

    #širenje zaraze u jednom dt
    for i in nx.nodes(G):
        if R_probe[i] == 0:

            brzs=0 #broj zaraženih susjeda vrhu 'i'
            for j in nx.neighbors(G,i):
                if R_probe[j] == 1:
                    brzs+=1
            #prijenos zaraze na vrh 'i'
            if random()<(1-(1-pInfect)**brzs):
                G.nodes[i]['infection']=1

        #oporavak zaraženog vrha
        if R_probe[i]==1:
            if random()<pRecover:
                G.nodes[i]['infection']=0

    #bilježenje podataka R u trenutku t
    for i in nx.nodes(G):
        #stanje zaraze na mreži G u trenutku t
        zaraza = [G.nodes[i]['infection'] for i in nx.nodes(G)]
        for j in nx.nodes(G):

            #provjera povezanosti dvaju vrhova
            if nx.has_path(G, i, j):
                bz=0 #ukupan broj zaraženih vrhova
                for c in zaraza:
                    if c==1:
```

```
                        bz+=1
                    if bz!=0:
                        #udaljenost vrha 'i' i vrha 'j'
                        dij=len(nx.shortest_path(G, i, j))-1
                        R[dt][i]+=zaraza[j]*dij/bz
    return R
```

Funkcija kao rezultat vraća $2D$ polje sa zapisanim vrijednostima $R_i(t)$ za svaki vrh $i \in \{1, .., N\}$ u svakom trenutku $t \leq T$. Ovu funkciju se u ostatku koda ponovi željeni broj puta te se sve odgovarajuće vrijednosti $R_i(t)$ (u odgovarajućem trenutku i za odgovarajući vrh) zbroje te podijele s brojem ponavljanja. Iz toga se crta graf usrednjenih vrijednosti $R_i(t)$ u ovisnosti o vremenu. Za $\nabla R_i(t)$ jednostavno se usrednjene vrijednosti u trenutku t oduzimaju od usrednjenih vrijednosti u trenutku $t+1$ te se iz dobivenih rezultata dobiva ovisnost gradijenta o vremenu.

A.3 Simulacija SIS modela na mreži živčanog sustava oblića Caenorhabditis elegans

A.3.1 Generiranje mreže iz skupa podataka o mreži živčanog sustava oblića Caenorhabditis elegans

```
f = open("C_elegans_edges.txt", "r")
lines=f.readlines()
source=[]
target=[]
for line in lines:
    pomocna=line.split(' ')
    source.append(int(pomocna[0]))
    target.append(int(pomocna[1]))
edges=list(zip(source,target))

#generiranje mreže živčanog sustava
G = nx.Graph()
G.add_edges_from(edges)
```

Podatci o mreži živčanog sustava zapisani su u tekstualnu datoteku u dva stupca, gdje stupci odgovaraju vrhovima koji su povezani odgovarajućom vezom. Očitavanjem redova u datoteci, zapravo očitamo dva vrha povazana vezom. Iz očitanih veza generira se mreža.

A.3.2 Simulacija širenja živčanih podražaja na živčanoj mreži C. elegans

Za razliku od ranijih SIS simulacija, mreža se generira samo jedanput te se iz te iste mreže simulira širenje zaraze iz zadanih početnih uvjeta. Generirana mreža tada ulazi kao argument funkcije za simulaciju širenja podražaja prema SIS modelu.

```
def SIS_graph(T, G, pInfect, pRecover):
#T - ukupan broj diskretnih vremenskih koraka
#G - mreža na kojoj se simulira širenje po SIS modelu
#pInfect, pRecover - parametri SIS modela

    #2D polje - spremanje R za svaki vrh u t<=T
    R=[[0 for i in range(N)] for dt in range(T)]

    #početni uvjeti
    for i in nx.nodes(G):
        G.nodes[i]['infection']=0
    G.nodes[0]['infection']=1

    for dt in range(1,T):
        #pomoćno polje
        R_probe = [G.nodes[i]['infection'] for i in nx.nodes(G)]

        #širenje podražaja u jednom dt
        for i in nx.nodes(G):
            if R_probe[i] == 0:
                brzs=0
                for j in nx.neighbors(G,i):
                    if R_probe[j] == 1:
```

```
                brzs+=1
            if random()<(1-(1-pInfect)**brzs):
                G.nodes[i]['infection']=1
        if R_probe[i]==1:
            if random()<pRecover:
                G.nodes[i]['infection']=0
    for i in nx.nodes(G):
        zaraza = [G.nodes[i]['infection'] for i in nx.nodes(G)]
        for j in nx.nodes(G):
            if nx.has_path(G, i, j):
                bz=0
                for c in zaraza:
                    if c==1:
                        bz+=1
                if bz!=0:
                    dij=len(nx.shortest_path(G, i, j))-1
                    R[dt][i]+=zaraza[j]*dij/bz
return R
```

Literatura

[1] Brauer, F. Mathematical epidemiology: Past, present, and future. // Infectious Disease Modelling. Vol. 2, 2 (2017), str. 113–127.

[2] Pastor-Satorras, R.; Castellano, C.; Mieghem, P. V. et al. Epidemic processes in complex networks. // Rev. Mod. Phys. Vol. 87, 3 (2015.), str. 925–979.

[3] Cox, D. R. Renewal Theory. Methuen Publishing Ltd, 1967. London.

[4] Ross, S. M. Stochastic Processes. John Wiley and Sons. Inc., 1996. New York.

[5] Anderson, R. M.; May, R. M. Infectious Diseases of Humans. Oxford University Press, 1992. Oxford.

[6] Hethcote, H. W. The Mathematics of Infectious Diseases. // SIAM Review. Vol. 42, 4 (2000.), str. 599–653.

[7] Andersson, H.; Britton, T. Stochastic Epidemic Models and Their Statistical Analysis. Springer, New York, NY, 2000.

[8] Watts, D. J.; Strogatz, S. H. Collective dynamics of 'small-world' networks. // Nature. Vol. 393 (1998.), str. 440–442.

[9] Barabási, A.-L.; Albert, R. Emergence of Scaling in Random Networks. // Science. Vol. 286 (1999.), str. 509–512.

[10] Boccaletti, S.; Latora, V.; Moreno, Y. et al. Complex networks: Structure and dynamics. // Physics Reports. Vol. 424, 4 (2006.), str. 175–308.

[11] Bollobás, B. Random graphs. Academic Press, London, 1985. London Mathematical Society Monographs.

[12] Papadopoulos, L.; Porter, M. A.; Daniels, K. E. et al. Network analysis of particles and grains. // Journal of Complex Networks. Vol. 6, 4 (2018.), str. 485.–565.

[13] Krioukov, D.; Kitsak, M.; Sinkovits, R. S. et al. Network Cosmology. // Sci Rep. Vol. 2, 793.

[14] Solé, R. V.; Munteanu, A. The large-scale organization of chemical reaction networks in astrophysics. // EPL. Vol. 68, 2 (2004.), str. 170.

[15] Hong1, S.; Dey, A. Network analysis of cosmic structures: network centrality and topological environment. // Monthly Notices of the Royal Astronomical Society. Vol. 450, 2 (2015.), str. 1999.–2015.

[16] Biamonte, J.; Faccin, M.; Domenico, M. D. Complex networks from classical to quantum. // Communications Physics. Vol. 2, 53.

[17] Zhu, K.; Ying, L. Information Source Detection in the SIR Model: A Sample-Path-Based Approach. // IEEE/ACM Transactions on Networking. Vol. 24, 1.

[18] Lokhov, A. Y.; Mézard, M.; Ohta, H. et al. Inferring the origin of an epidemic with a dynamic message-passing algorithm. // Phys. Rev. E. Vol. 90, 1 (2014.), str. 012801.

[19] Shah, D.; Zaman, T. Detecting sources of computer viruses in networks: theory and experiment. // ACM SIGMETRICS Perform. Eval. Rev. Vol. 38, 1 (2010.), str. 203–214.

[20] Shah, D.; Zaman, T. Rumors in a Network: Who's the Culprit?. // IEEE Transactions on Information Theory. Vol. 57, 8 (2011.), str. 5163–5181.

[21] Altarelli, F.; Braunstein, A.; Dall'Asta, L. et al. Bayesian Inference of Epidemics on Networks via Belief Propagration. // Phys. Rev. Lett. Vol. 112, 11 (2014.), str. 118701.

[22] Antulov-Fantulin, N.; Lančić, A.; Šmuc, T. et al. Identification of Patient Zero in Static and Temporal Networks: Robustness and Limitations. // Phys. Rev. Lett. Vol. 114, 24 (2015.), str. 248701.

[23] Antulov-Fantulin, N.; Lančić, A.; Šmuc, T. et al. Supplementary information: Identification of patient zero in static and temporal networks - robustness and limitations. // .

[24] Hagberg, A. A.; Schult, D. A.; Swart, P. J. Exploring Network Structure, Dynamics, and Function using NetworkX. // Proceedings of the 7th Python in

Science Conference. str. 11. – 15. Edited by G. Varoquaux, T. Vaught and J. Millman.

[25] watts_strogatz_graph, https://networkx.org/documentation/networkx-1.9/reference/generated/networkx.generators.random_graphs.watts_strogatz_graph.html, .

[26] Achacoso, T. B.; Yamamoto, W. S. Ay's Neuroanatomy of C. Elegans for Computation. CRC Press, 1992. , Boca Raton, FL.

[27] Latora, V.; Nicosia, V.; Russo, G. Complex Networks: Principles, Methods and Applications. Cambridge University Press, 2017.

[28] Lockery, S. R.; Goodman, M. B. The Quest for Action Potentials in C. elegans Neurons Hits a Plateau. // Nature neuroscience. Vol. 12, 4 (2009.), str. 377. – 378.

[29] Zachary, W. W. An Information Flow Model for Conflict and Fission in Small Groups. // Journal of Anthropological Research. Vol. 33, 4 (1977.), str. 452–473.