Rani signali faznog prijelaza

Perković, Petar

Master's thesis / Diplomski rad

2018

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:217:947855

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-03-29**

Repository / Repozitorij:

Repository of the Faculty of Science - University of Zagreb

SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO-MATEMATIČKI FAKULTET
FIZIČKI ODSJEK

Petar Perković

Rani signali faznog prijelaza

Diplomski rad

Zagreb, 2018.

SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO-MATEMATIČKI FAKULTET
FIZIČKI ODSJEK

INTEGRIRANI PREDDIPLOMSKI I DIPLOMSKI SVEUČILIŠNI STUDIJ
FIZIKA; SMJER ISTRAŽIVAČKI

Petar Perković

Diplomski rad

Rani signali faznog prijelaza

Voditelj diplomskog rada: izv. prof. dr. sc. Davor Horvatić

Ocjena diplomskog rada: _____

Povjerenstvo: 1. _____

2. _____

3. _____

Datum polaganja: _____

Zagreb, 2018.

Sažetak

U ovom radu predstavljena je metodologija za dobivanje ranih signala faznog prijelaza u složenim sustavima uzimajući u obzir povezanost među članovima sustava. Metodologija je primijenjena na sustav od devet različitih kriptovaluta. Korištena je transferna entropija za dobivanje težinskih usmjerenih veza između različitih članova sustava. Pomoću dobivenih veza konstruirana je usmjerena mreža kriptovaluta u različitim vremenskim točkama. Iz promatranja ponašanja mreže kroz vrijeme uočeno je povećanje povezanosti u razdobljima naglih promjena na tržištu. Uveden je indeks povezanosti i uspoređen s kretanjem vrijednosti bitcoina. Dobiveni rezultati ukazuju na mogućnost primjene korištene metodologije za otkrivanje ranih signala u složenim sustavima.

Ključne riječi: rani signali, fazni prijelaz, transferna entropija, usmjerena mreža, kriptovalute

Early warning signals of phase transition

Abstract

In this thesis, we present the methodology for obtaining early warning signals of the phase transition for complex systems by taking into account the relationship between elements that build specific system. The methodology is applied to a system of nine different cryptocurrencies. Transfer entropy was used to obtain directed weighted links between different system members. By using obtained links a directed network of cryptocurrencies is constructed at different time points. From the observation of the network's behavior over time, we notice an increase in connectedness in times of severe market changes. The connectedness index is introduced and compared with the movement of bitcoin price. The obtained results indicate the possibility of using the presented methodology for detecting early warning signals in complex systems.

Keywords: early warning signals, phase transition, transfer entropy, directed network, cryptocurrencies

Sadržaj

1 Uvod **1**

1.1 Kritično usporavanje 2

1.2 Drugi rani signali 3

2 Entropija u statističkoj fizici i teoriji informacije **5**

2.1 Uvod u statističku fiziku 5

2.2 Entropija u statističkoj fizici 6

2.3 Gibbsova entropija 9

2.4 Entropija u teoriji informacija 10

2.5 Transferna entropija 12

2.6 Efektivna transferna entropija 13

3 Mreže **15**

3.1 Uvod u mreže 15

3.2 Stupanj čvora 16

3.3 Matrica susjedstva 17

3.4 Putevi i udaljenosti 18

3.5 Koeficijent grupiranja 19

3.6 Centralnost 19

4 Pregled modela **21**

4.1 Podaci 21

4.2 Procjena vjerojatnosti 22

4.3 Transferna entropija 23

4.4 Usmjerena mreža 25

5 Rezultati **27**

5.1 Entropija 27

5.2 Transferna entropija 28

5.3 Mreža kriptovaluta 31

6 Zaključak **37**

Dodaci **38**

A Kod **38**

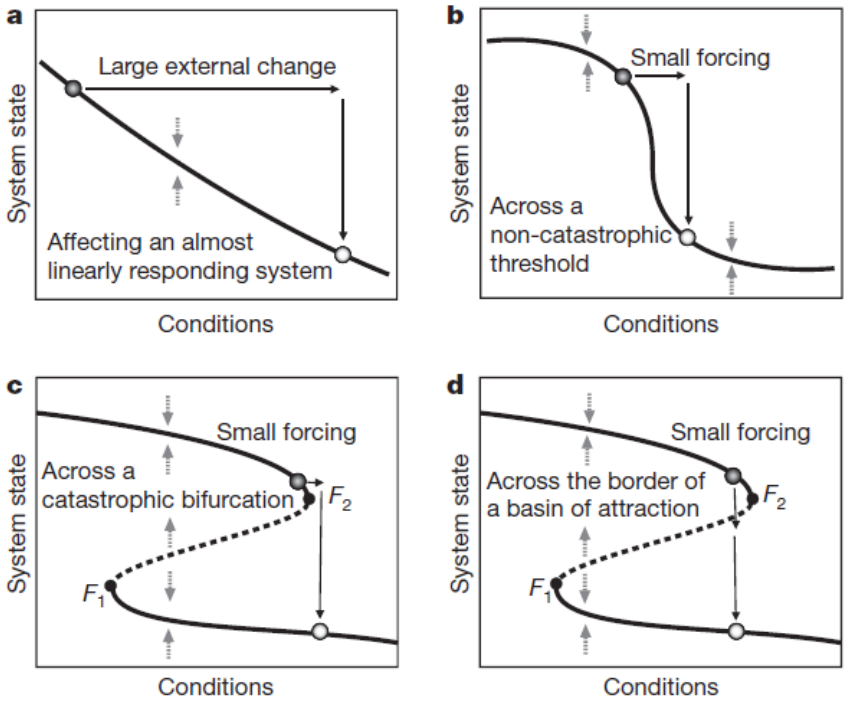
Literatura **49**

1 Uvod

Mnogi složeni sustavi imaju prijelomnu točku u kojoj sustav naglo prelazi iz jednog stanja u drugo. Primjeri se mogu naći u raznim granama znanosti, od naglih klimatskih promjena u klimatologiji do kraha burze u ekonomiji i epileptičnih napadaja u medicini.

Takvi prijelazi često su teško predvidljivi i sustavi prolaze kroz malo promjena prije faznog prijelaza. Razlog tome je što modeli složenih sustava ne opisuju sustav dovoljno precizno za predviđanje takvih situacija i sustavi uglavnom nemaju jasne signale koji upućuju na fazni prijelaz.

Unatoč tome, iz raznih istraživanja vidljivo je da postoje rani signali faznih prijelaza koji imaju slična svojstva za različite sustave. To govori da se mogu odrediti dinamike ponašanja sustava blizu kritične točke koje su općenite za širok opseg sustava. U takvim modelima kritične točke odgovaraju bifrukacijama. Posebno su bitne katastrofalne bifrukacije, gdje, kad se kritična točka prijeđe, sustav prolazi kroz naglu promjenu prema suprotnom stanju. Također su zanimljive bifrukacije gdje sustav prelazi iz stabilnog stanja u kaotično stanje.



Slika 1.1: Grafovi ovisnosti stanja o uvjetima za različite sustave. Pod a) i b) prikazani su sustavi s nekatastrofalnim bifrukacijama, a pod c) i d) prikazan je sustav s katastrofalnom bifrukacijom [1].

Sustavi mogu različito reagirati na promjenu uvjeta u kojima se nalaze. Neki sustavi reagirat će proporcionalno promjeni uvjeta i ne dolazi do velikih iznenađenja (Slika 1.1, a) dok će drugi sustavi reagirati burno na malu promjenu uvjeta (Slika 1.1,b). U oba ova slučaja efekt promjene kontinuiran je i reverzibilan te postoji samo jedno stanje za iste uvjete. Također postoje sustavi u kojima pri maloj promjeni uvjeta dolazi do kritičnog prijelaza (Slika 1.1, c i d). To je primjer katastrofalne bifrukacije. U takvim sustavima dolazi do nagle promjene stanja. Efekt takvih promjena nije kontinuiran i reverzibilan, a sustav se može nalaziti u više različitih stanja za iste uvjete.

1.1 Kritično usporavanje

Kritično je usporavanje fenomen koji je jedan od glavnih pokazatelja da se sustav približava kritičnoj granici [6]. Sustav se prilikom malih perturbacija to sporije vraća u osnovno stanje što je bliže točki prijelaza. Taj fenomen prisutan je u većini sustava. Analiza raznih modela pokazuje da usporavanje počinje daleko od točke bifrukacije i vrijeme povratka smanjuje se u nula kako se približava kritičnoj točki.

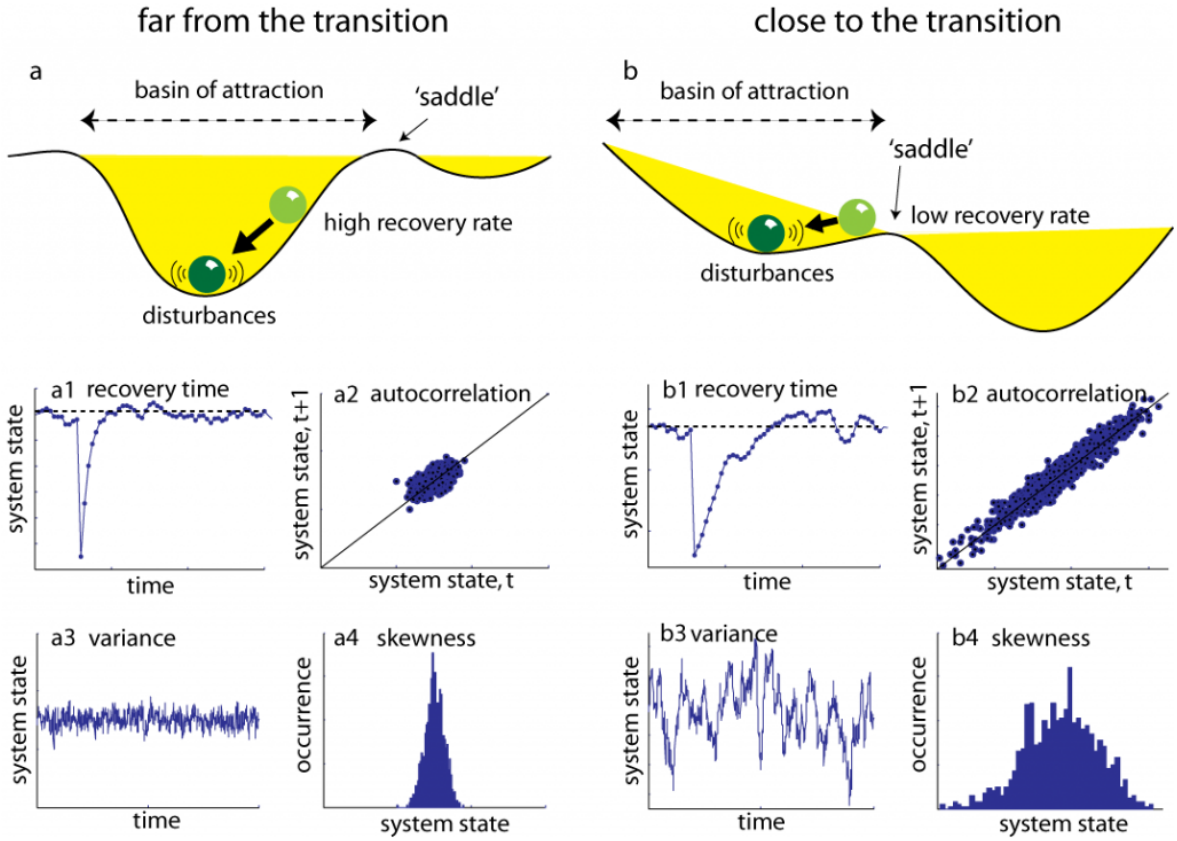
Stopa povratka u početno stanje jedna je od veličina koje se mijenjaju prilikom kritičnog usporavanja [7]. Nju se može mjeriti tako da se na sustav primijeni mala perturbacija koja neće uzrokovati prijelaz sustava preko granice. Pomoću toga može se odrediti blizina kritične točke. Primjer je toga uzimanje dijela ekosustava i micanje dijela biomase iz njega. Sustav koji je daleko od kritične točke puno će brže obnoviti tu biomasu nego sustav koji je blizu izumiranja [8]. Problem je ove metode što je često jako teško ili gotovo nemoguće napraviti perturbaciju u nekim složenim sustavima.

Porast autokorelacije druga je posljedica kritičnog usporavanja [9]. Kako se sustav sporije vraća u početno stanje, tako je i stanje u trenutku $t+1$ sličnije stanju u trenutku t. To govori da dolazi do povećanja autokorelacije u sustavu u blizini bifrukacije. Isto se može mjeriti računanjem autokorelacijskog koeficijenta među susjednim stanjima u vremenskom nizu (Slika 1.2, a2 i b2).

Porast varijance također se primjećuje prilikom kritičnog usporavanja [10]. Blizu kritične točke utjecaji šokova ne opadaju i akumulirano povećavaju varijancu u sustavu. Na Slici 1.2 vidimo primjer sustava daleko od kritične točke koji ima nisku

varijancu (a3) i primjer sustava blizu kritične točke s visokom varijancom (b3).

Ova tri rana signala posljedica su kritičnog usporavanja sustava u blizini bifrukacije. Uz njih postoje još neki rani signali koji ukazuju na fazne prijelaze.



Slika 1.2: Prikaz sustava u blizini (b) i daleko (a) od kritične točke. Prikazani su rani signali koji mogu upućivati na fazni prijelaz [1].

1.2 Drugi rani signali

Porast asimetrije fluktuacija javlja se kad je sustav blizu kritične točke [11] i nije posljedica kritičnog usporavanja. U smjeru prijelaza stopa promjene manja je nego na suprotnoj strani pa tu sustav provodi više vremena. To rezultira asimetričnom distribucijom oko točke stabilnosti.

Treperenje je još jedan primjer ranog signala. Događa se ako je stohastičko prisiljavanje dovoljno jako da pomiče sustav između dva područja privlačenja kad je sustav u bistabilnom području [12]. Može se uočiti u povećanju varijance i asimetrije, kao i prisutnosti bimodalnosti.

Uz rane signale kroz vremenske serije postoje i određeni prostorni uzorci koji se mogu pojavljivati prije faznog prijelaza. Mnogi se sustavi sastoje od manjih dijelova

koji međusobno interagiraju. Svaki od tih dijelova sustava teži prema stanju u kojem se nalazi njegov susjed. Primjer takvog ponašanja financijska su tržišta, gdje stanje jednog tržišta djeluje na stanje drugoga. U takvim sustavima može doći do promjene faze, kao kod feromagneta gdje pojedine čestice međusobno djeluju na spin. Porastom vanjskog utjecaja sustav dolazi blizu prijelomne točke, a distribucija stanja sustava mijenja se na karakterističan način. Pomoću promatranja promjene distribucije stanja može se prepoznati da je sustav blizu faznog prijelaza.

2 Entropija u statističkoj fizici i teoriji informacije

2.1 Uvod u statističku fiziku

Složeni sustav sastoji se od N čestica smještenih u prostor volumena V. U stvarnosti složeni sistemi sadrže velik broj čestica ($N > 10^{23}$) u velikom volumenu i na takve sisteme može se primijeniti termodinamički limit. Termodinamički limit definiran je kao limit sustava u kojem $N \to \infty$ i $V \to \infty$, a gustoća čestica n je konstantna ($n = N/V = konst.$). U takvom sustavu intenzivna svojstva su nepromjenjiva promjenom veličine sustava dok su ekstenzivna svojstva proporcionalna veličini sustava.

Ukupna energija sustava jednaka je zbroju energija svih čestica tog sustava. Ako je ϵ_i energija jedne čestice i n_i je broj čestica s tom energijom, onda pomoću toga možemo odrediti ukupnu energiju E i ukupni broj čestica N.

$$E = \sum_i n_i \epsilon_i, \tag{2.1}$$

$$N = \sum_i n_i. \tag{2.2}$$

Kvantna mehanika govori nam da pojedina čestica u volumenu V ima diskretne energije ϵ_i, pa i sustav takvih čestica ima diskretne energije E. Međutim, kako je sustav jako velik i razlike između njegovih diskretnih stanja energije jako su male s obzirom na njegovu ukupnu energiju, energija E može se smatrati kontinuiranom varijablom. Ovo vrijedi i za sustave u kojima čestice interagiraju, ali u njima ne vrijedi izraz za ukupnu energiju E (2.1).

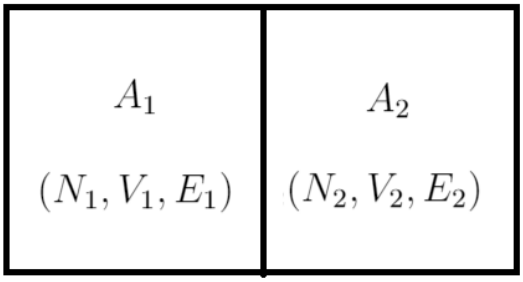
Vrijednosti parametara N, V i E definiraju makroskopska stanja danog sistema. Za određeno makroskopsko stanje postoji veliki broj mogućih stanja na molekularnoj razini. U neinteragirajućem sustavu s N čestica i volumena V, ukupna energija E može se raspodijeliti na mnogo načina među česticama od kojih se sastoji. Svaki od tih načina raspodjele jedno je mikroskopsko stanje sistema. Za svako od mikroskopskih stanja koja odgovaraju makroskopskom stanju može se pretpostaviti da ima jednaku vjerojatnost da će se u nekom trenutku t sustav nalaziti u jednome od njih ako nema drugih ograničenja. Ova pretpostavka još se naziva postulat jednake *a priori* vjerojatnosti svih mikroskopskih stanja za dano makroskopsko stanje.

Ukupni broj svih mogućih mikroskopskih stanja $\Omega(N, V, E)$ funkcija je volumena V, broja čestica N i energije sustava E. Iz vrijednosti $\Omega(N, V, E)$ i njegove ovisnosti o N, V i E moguće je izvesti cijelu termodinamiku danog sustava.

2.2 Entropija u statističkoj fizici

Dva odvojena sustava A_1 i A_2 nalaze se u ravnoteži. Makroskopsko stanje sustava A_1 opisano je s N_1, V_1 i E_1, a broj mikroskopskih stanja je $\Omega_1(N_1, V_1, E_1)$ dok je A_2 opisan s N_2, V_2 i E_2, i ima $\Omega_2(N_2, V_2, E_2)$ mikroskopskih stanja.

Kad se ta dva sustava stave u termalni kontakt, dolazi do izmjene energija među njima (Slika 2.1). Odvojeni su čvrstom pregradom koja ne dopušta izmjenu čestica tako da su N_1, N_2, V_1 i V_2 očuvani.



Slika 2.1: Dva sustava u termalnom kontaktu.

Energije E_1 i E_2 više nisu fiksne, ali zbroj energija oba sustava ostaje očuvan i jednak je $E^{(0)}$.

$$E_1 + E_2 = E^{(0)} = konst. \tag{2.3}$$

Sustav A_1 ima istu vjerojatnost da se nalazi u jednom od $\Omega_1(E_1)$ stanja, a sustav A_2 ima istu vjerojatnost da se nalazi u jednom od $\Omega_2(E_2)$ stanja. Sustavi A_1 i A_2 zajedno čine sustav $A^{(0)}$. Ukupan broj stanja $\Omega^{(0)}$ u kojem se može nalaziti sustav $A^{(0)}$ jednak je umnošku $\Omega_1(E_1)$ i $\Omega_2(E_2)$, i svako od tih stanja ima jednaku vjerojatnost.

$$\Omega_1(E_1)\Omega_2(E_2) = \Omega_1(E_1)\Omega_2(E^{(0)} - E_1) = \Omega^{(0)}(E^{(0)}, E_1). \tag{2.4}$$

Vidljivo je da se $\Omega^{(0)}$ mijenja u ovisnosti o E_1. Pitanje je za koju energiju E_1 će se sustavi naći u ravnoteži.

Kako je svako mikroskopsko stanje u kojem se sustav nalazi jednako vjerojatno, makroskopsko stanje s najviše mikroskopskih stanja je najvjerojatnije. Tvrdimo da će se sustav kretati prema takvom stanju u kojem postiže maksimalni broj mikrostanja . Za tipične sustave broj mikroskopskih stanja opada drastično pomakom iz makroskopskog stanja s najvećim brojem istih. Takvi sustavi veliku većinu vremena provode u tom makroskopskom stanju i možemo ga identificirati kao ravnotežno stanje. Sustav će se naći u ravnoteži za energiju E_1 za koju postiže maksimalan $\Omega^{(0)}$.

Pri energijama $\overline{E_1}$ i $\overline{E_2}$ dolazi do ravnotežnog stanja. Derivacija $\Omega^{(0)}$ po E_1 na tim energijama jednaka je 0, jer na tim energijama postiže maksimalnu vrijednost.

$$\frac{\partial \Omega^{(0)}}{\partial E_1} = \left(\frac{\partial \Omega_1(E_1)}{\partial E_1}\right)_{E_1=\overline{E_1}} \Omega_2(\overline{E_2}) + \Omega_1(\overline{E_1}) \left(\frac{\partial \Omega_2(E_2)}{\partial E_2}\right)_{E_2=\overline{E_2}} \frac{\partial E_2}{\partial E_1} = 0. \tag{2.5}$$

Iz jednadžbe (2.1) dobije se $\partial E_2/\partial E_1$=-1 pomoću čega se dobije izraz:

$$\left(\frac{\partial \ln \Omega_1(E_1)}{\partial E_1}\right)_{E_1=\overline{E_1}} = \left(\frac{\partial \ln \Omega_2(E_2)}{\partial E_2}\right)_{E_2=\overline{E_2}}. \tag{2.6}$$

Uvjet ravnoteže između sustava A_1 i A_2 na kraju se svodi na jednakost između parametara β_1 i β_2, gdje je parametar β definiran kao:

$$\beta = \left(\frac{\partial \ln \Omega(N, V, E)}{\partial E}\right)_{N,V,E=\overline{E}}. \tag{2.7}$$

Kada se dva fizikalna sustava dovedu u termalni kontakt, dolazi do razmjene energije među njima. Ta razmjena nastavlja se do trenutka kada se postigne ravnoteža, to jest kad energije E_1 i E_2 poprime vrijednosti $\overline{E_1}$ i $\overline{E_2}$. Nakon postizanja ravnoteže ne dolazi više do razmjene energije među sustavima i sustavi se nalaze u termodinamičkoj ravnoteži. To se dogodi kad parametri β_1 i β_2 poprime jednake vrijednosti. Razumno je očekivati da su ti parametri na neki način povezani s termodinamičkom temperaturom T danih sustava. Iz termodinamike se može preuzeti formula koja glasi:

$$\left(\frac{\partial S}{\partial E}\right)_{N,V} = \frac{1}{T}, \tag{2.8}$$

gdje je S entropija sustava.

Uspoređujući izraze (2.7) i (2.8) može se primijetiti da postoji veza između entropije S i broja stanja Ω. Odnos između te dvije vrijednosti možemo zapisati za bilo

koji fizikalni sustav u sljedećem obliku:

$$\frac{\Delta S}{\Delta \ln \Omega} = \frac{1}{\beta T} = konst. \tag{2.9}$$

Ludwig Boltzmann prvi je prepoznao ovaj odnos između termodinamičkog i statističkog pristupa. Također je pretpostavljao da je konstanta koja se pojavljuje u jednadžbi univerzalna. Max Planck izraz je zapisao u finalnom obliku:

$$S = k \ln \Omega. \tag{2.10}$$

Usporedbom (2.9) i (2.10) dobije se izraz za β.

$$\beta = \frac{1}{kT}. \tag{2.11}$$

Konstanta k se naziva Boltzmannovom konstantom.

Izraz (2.10) daje ukupnu entropiju danog sustava preko broja dostupnih mikroskopskih stanja za dano makroskopsko stanje. Ovaj izraz poprima minimalnu vrijednost 0 za stanje kada postoji samo jedno mikroskopsko stanje, što se slaže s trećim termodinamičkim zakonom. Ova formula je značajna jer predstavlja poveznicu između mikroskopskog i makroskopskog.

Entropija iz (2.10) također opisuje neuređenost sustava. Što sustav ima više mogućih mikroskopskih stanja to mu je neuređenost veća i teže je predvidljiv. Potpuna uređenost sustava nastupa kad postoji jedno moguće stanje i tad entropija iščezava.



Slika 2.2: Boltzmannova bista s formulom za entropiju.

2.3 Gibbsova entropija

Entropija koju mjerimo posljedica je različitih stanja u kojima se sustav može nalaziti. Svako od tih stanja može imati velik broj mikroskopskih stanja koji ne možemo mjeriti. Kako sustav može biti u svakom od tih stanja, ona pridonose ukupnoj entropiji prema formuli (2.10).

Pretpostavimo da se sustav može nalaziti u N različitih i jednako vjerojatnih mikroskopskih stanja. Ta mikroskopska stanja grupirana su u različita makroskopska stanja. Suma broja mikroskopskih stanja n_i u makroskopskom stanju jednaka je ukupnom broju mikroskopskih stanja N.

$$\sum_i n_i = N. \tag{2.12}$$

Vjerojatnost P_i da se sustav nađe u jednom od makroskopskih stanja je:

$$P_i = \frac{n_i}{N}. \tag{2.13}$$

Koristeći (2.12) pokazujemo da vrijedi:

$$\sum_i P_i = 1. \tag{2.14}$$

Ukupna entropija S_{tot} dobije se pomoću izraza (2.10) ($S_{tot} = k \ln N$). Kako se ne može direktno mjeriti broj mikroskopskih stanja, izmjerena entropija S jednaka je doprinosu od različitih makroskopskih stanja u kojima se sustav može nalaziti. Ukupnu entropiju S_{tot} dobijemo kad izmjerenoj entropiji S dodamo entropiju povezanu s mikroskopskim stanjima S_{micro}.

$$S_{tot} = S + S_{micro}. \tag{2.15}$$

Entropija povezana s mikroskopskim stanjima u kojima se sustav može nalaziti dana je izrazom:

$$S_{micro} = \langle S_i \rangle = \sum_i P_i S_i = \sum_i P_i k \ln n_i. \tag{2.16}$$

S_i je entropija mikroskopskih stanja povezanih s i-tim makroskopskim stanjem.

Za izmjerenu entropiju vrijedi:

$$S = S_{tot} - S_{micro} = k\left(\ln N - \sum_i P_i \ln n_i\right) = k\sum_i P_i(\ln N - \ln n_i). \tag{2.17}$$

Koristeći (2.13) dobijemo $\ln N - \ln n_i = -\ln P_i$ i uvrštavanjem u (2.17) dobijemo:

$$S = -k\sum_i P_i ln P_i. \tag{2.18}$$

Ovaj izraz definicija je **Gibbsove entropije**.

Gibbsova entropija smislena je i kad se sustav nalazi daleko od termodinamičke ravnoteže dok ostale definicije za entropiju pretpostavljaju da je sustav u ravnoteži. Gibbsova entropija minimalnu vrijednost 0 postiže kad postoji samo jedno stanje s vjerojatnošću 1. Također, što sustav ima više dostupnih mikroskopskih stanja vrijednost P_i opada i entropija sustava raste.

Može se pokazati da se (2.18) svodi na (2.10) kad imamo Ω različitih mikroskopskih stanja u sustavu, gdje svako stanje ima jednaku vjerojatnost. Vjerojatnost pojedinog stanja je $P_i = 1/\Omega$ i uvrštavanjem toga u (2.18) dobijemo :

$$S = -k\sum_i \frac{1}{\Omega}\ln\frac{1}{\Omega}. \tag{2.19}$$

Koristeći da je $\sum_i \frac{1}{\Omega} = 1$ dobijemo:

$$S = k\ln\Omega. \tag{2.20}$$

Vidljivo je da je entropija svedena na oblik iz formule (2.10).

2.4 Entropija u teoriji informacija

U prethodnom dijelu prošli smo kroz definicije entropije u statističkoj fizici, a sada ćemo uvesti entropiju u teoriji informacija. Dok u statističkoj fizici entropija opisuje neuređenost sustava, u teoriji informacija ona je mjera za nesigurnost događaja. Claude Elwood Shannon utemeljitelj je teorije informacija.

Shannonov informacijski sadržaj ishoda i definiran je kao:

$$h(P_i) = \log_2 \frac{1}{P_i}, \tag{2.21}$$

gdje je P_i vjerojatnost događaja i. Mjerna jedinica su bitovi koji poprimaju vrijednost između 0 i 1.

Entropija sustava X definirana je kao prosječni Shannonov informacijski sadržaj ishoda:

$$H(X) = \sum_i P_i h(P_i) = \sum_i P_i \log_2 \frac{1}{P_i}. \tag{2.22}$$

Drugo je ime za entropiju nesigurnost od X. Entropija se također mjeri u bitovima i možemo je kvantificirati kao prosječan broj bitova potrebnih za dešifrirati X. Samo kad je vjerojatnost događaja jednaka 1, entropija poprima minimalnu vrijednost jednaku 0. U slučaju kada postoji N stanja jednake vjerojatnosti, entropija logaritamski raste s brojem N. Entropija je maksimalna ako je raspodjela vjerojatnosti među događajima uniformna i poprima vrijednost $\log_2 N_X$, gdje je N_X broj mogućih događaja.

Redundancija sustava X definirana je kao:

$$R = 1 - \frac{H(X)}{\log_2 N_X}. \tag{2.23}$$

Zajednička entropija dvaju sustava X i Y dana je izrazom:

$$H(X,Y) = \sum_{i,j} P(i,j) \log_2 \frac{1}{P(i,j)}. \tag{2.24}$$

Ako su sustavi nezavisni, entropija je aditivna.

$$H(X,Y) = H(X) + H(Y) \text{ ako } P(i,j) = P_i P_j. \tag{2.25}$$

Entropija ima rekurzivno svojstvo koje je veoma korisno prilikom računanja iste. Rekurzivnost entropije omogućuje nam dekompoziciju entropije i njezino računanje u više manjih koraka. To svojstvo dano je izrazom:

$$H(P_1, P_2, ..., P_n) = H(P_1 + P_2, P_3, ..., P_n) + (P_1 + P_2) H\left(\frac{P_1}{P_1 + P_2}, \frac{P_2}{P_1 + P_2}\right). \tag{2.26}$$

2.5 Transferna entropija

Prethodna stanja nekog vremenskog niza mogu utjecati na njezina buduća stanja. Vremenski niz varijable X možemo promatrati kao Markovljev proces u kojem buduće stanje i_{n+1} ovisi o k prethodnih stanja. To znači da vrijedi:

$$P(i_{n+1}|i_n, i_{n-1}, ..., i_0) = P(i_{n+1}|i_n, i_{n-1}, ..., i_{n-k+1}) = P(i_{n+1}|i_n^{(k)}), \tag{2.27}$$

gdje je $P(A|B)$ uvjetna vjerojatnost za A ako se dogodilo B i definirana je kao:

$$P(A|B) = \frac{P(A,B)}{P(B)}. \tag{2.28}$$

Kada imamo više varijabli koje međudjeluju, onda vremenski niz varijable Y može utjecati na vremenski niz varijable X. U (2.24) definirali smo entropiju dvaju varijabli X i Y. Entropija u tom slučaju poprima minimalnu vrijednost ako su sva stanja sustava jednako vjerojatna.

Uvjetna entropija $H(X|Y)$ govori koliko je informacije potrebno da bi se opisao ishod varijable X uzimajući vrijednost nasumične varijable Y. Pomoću nje se može odrediti neizvjesnost varijable X, ako se poznaje vrijednost varijable Y. Uvjetna entropija dana je izrazom:

$$H(X|Y) = -\sum P(x,y) \log_2 P(x|y). \tag{2.29}$$

Ako vremenske nizove X i Y promatramo kao Markovljeve procese gdje buduće stanje i_{n+1} od X ovisi o l prethodnih stanja j od varijable Y , uvjetna entropija jednaka je:

$$\begin{aligned} H(X_{t+1}|Y_{t:t-1-l}) &= -\sum P(i_{n+1}, j_n^{(l)}) \log_2 P(i_{n+1}|j_n^{(l)}) = \\ = -\sum P(i_{n+1}, j_n, j_{n-1}, ..., j_{n-l-1}) &\log_2 P(i_{n+1}|j_n, j_{n-1}, ..., j_{n-l+1}). \end{aligned} \tag{2.30}$$

Transferna entropija mjeri usmjereni transfer informacije između dva nasumična procesa. Ona je prosječna informacija iz vremenskog niza Y o sljedećem stanju vremenskog niza X koja već nije bila sadržana u vremenskom nizu X. Definirana je kao razlika uvjetne entropije gdje buduća vrijednost vremenskog niza X ovisi o svojoj prošlosti i uvjetne entropije gdje buduća vrijednost vremenskog niza X ovisi o svojoj

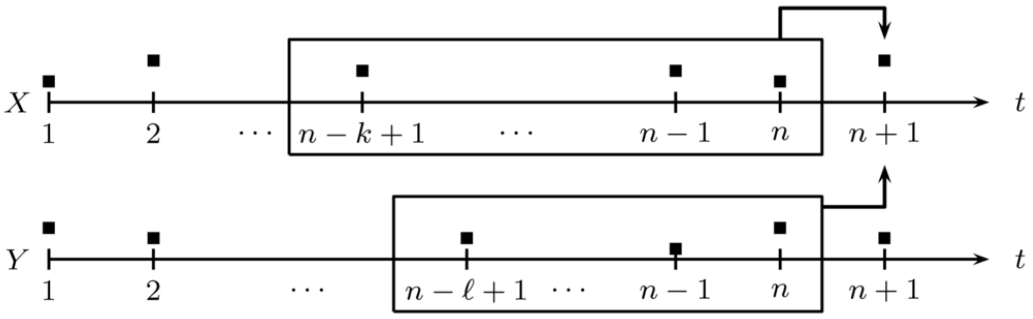
prošlosti i prošlosti Y:

$$T_{Y\to X} = H(X_{t+1}|X_{t:t-1-k}) - H(X_{t+1}|X_{t:t-1-k}, Y_{t:t-1-l}). \tag{2.31}$$

Uvrštavanjem izraza za uvjetnu entropiju (2.30) u (2.31) dobije se sljedeći izraz za transfernu entropiju:

$$\begin{aligned} T_{Y\to X} &= -\sum P(i_{n+1}, i_n^{(k)}) \log_2 P(i_{n+1}|i_n^{(k)}) + \\ &+ \sum P(i_{n+1}, i_n^{(k)}, j_n^{(l)}) \log_2 P(i_{n+1}|i_n^{(k)}, j_n^{(l)}) = \\ &= \sum P(i_{n+1}, i_n^{(k)}, j_n^{(l)}) \log_2 \frac{P(i_{n+1}|i_n^{(k)}, j_n^{(l)})}{P(i_{n+1}|i_n^{(k)})}. \end{aligned} \tag{2.32}$$

Gornji izraz jednak je Kullback-Leibler divergenciji. Ako se pomoću vremenskog niza Y može deterministički odrediti sljedeća vrijednost vremenskog niza X, tada uvjetna entropija $H(X|X,Y)$ poprima minimalnu vrijednost 0 i transferna entropija poprima maksimalnu vrijednost $H(X|X)$. U slučaju da buduća vrijednost vremenskog niza X ne ovisi o vremenskom nizu Y tada se uvjetna entropija $H(X|X,Y)$ svodi na $H(X|X)$ i transferna entropija poprima minimalnu vrijednost 0. Transferna entropija mjeri koliko korištenje povijesti vremenskog niza Y pomaže u predviđanju budućih vrijednosti vremenskog niza X.



Slika 2.3: Shematski prikaz transferne entropije.

2.6 *Efektivna transferna entropija*

Transferna entropija u sebi sadrži puno šuma jer su dostupni podatci o vremenskom nizu konačni i nestacionarni. Neke vremenske serije imaju veću volatilnost od drugih

pa je samim time i prijenos informacije u takvim sustavima veći. Također je vrijednost transferne entropije uvijek pozitivna.

Dio ovih efekata može se eliminirati korištenjem efektivne transferne entropije koju su predložili Marschinski i Kantz. Prvo je potrebno izračunati očekivanu vrijednost nasumične transferne entropije $T^{(r)}_{Y\to X}$ koristeći nasumične permutacije vremenske serije Y. Tu očekivanu vrijednost oduzmemo od transferne entropije vremenske serije da se dobije efektivna transferna entropija.

$$T^{(e)}_{Y\to X} = T_{Y\to X} - E[T^{(r)}_{Y\to X}]. \tag{2.33}$$

Da bi se efektivna entropija svela na udio maksimalne vrijednosti, dijeli se s maksimalnom vrijednosti $H(X|X)$. Tako se dobije normalizirana transferna entropija.

$$T^{(n)}_{Y\to X} = \frac{T^{(e)}_{Y\to X}}{H(X_{t+1}|X_{t:t-1-k})} = \frac{T_{Y\to X} - E[T^{(r)}_{Y\to X}]}{H(X_{t+1}|X_{t:t-1-k})}. \tag{2.34}$$

Normalizirana transferna entropija mjeri udio informacije u vremenskom nizu X koju daje vremenski niz Y, a da tu informaciju ne sadrži njegova vlastita prošlost.

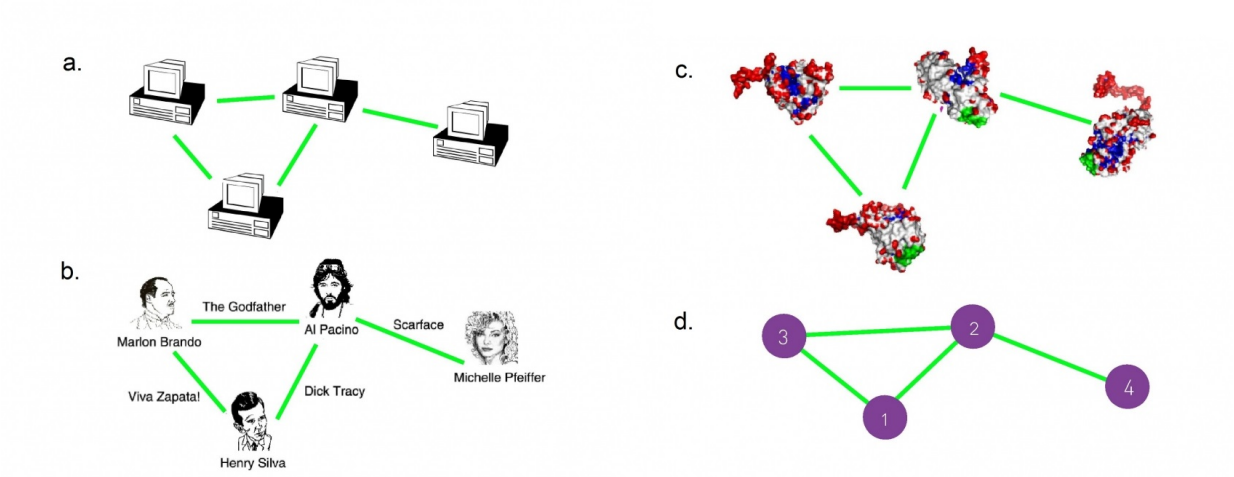
3 Mreže

3.1 Uvod u mreže

Znanost o mrežama proučava mrežni prikaz fizičkih, bioloških i društvenih fenomena dovodeći do prediktivnih modela tih fenomena. Da bismo razumjeli složene sustave, potrebno je poznavati kako njihovi dijelovi međusobno interagiraju. Te interakcije među dijelovima sustava mogu se prikazati preko mreže interakcija. Mreža je skup čvorova sustava i veza među njima. Pomoću mreža se mogu zajedničkim jezikom proučavati različiti sustavi. Iz Slike 3.1 vidljivo je da se tri potpuno različita sustava mogu svesti na istu mrežu.

Broj čvorova se označava s N i predstavlja broj dijelova sustava kojeg promatramo. N još se naziva i veličina mreže.

Broj veza označava se sa L i predstavlja broj interakcija među dijelovima sustava. Veze su uglavnom definirane čvorovima koje spajaju.



Slika 3.1: Pod a. je prikazana mreža među računalima, b. je mreža koja spaja glumce zajedničkim filmovima, c. je mreža interakcija među proteinima, a d. je grafički prikaz s 4 čvora i 4 veze na koji se svode sva tri prethodna primjera.

Veze među čvorovima mogu biti usmjerene i neusmjerene. Slanje poruke primjer je usmjerene veze jer jedna osoba šalje poruku drugoj, a obrnuta situacija ne mora biti slučaj. Primjer neusmjerene veze obiteljska je veza među osobama jer sve osobe vezane obiteljski jedna su drugoj obitelj. Neusmjerene mreže sadrže samo neusmjerene veze dok usmjerene mreže imaju samo usmjerene veze. Neke mreže mogu sadržavati i usmjerene i neusmjerene veze. Čvor također može imati vezu na samoga

sebe.

Kada sustav želimo prikazati u obliku mreže, bitno je kako definiramo čvorove i veze među čvorovima da bismo došli do željenih rezultata. Potrebno je paziti na njihov značaj pri razumijevanju problema koji želimo riješiti.

3.2 Stupanj čvora

Stupanj čvora jedno je od glavnih svojstava čvora. On predstavlja broj veza kojima je čvor vezan za druge čvorove. Stupanj čvora može predstavljati broj prijatelja na društvenoj mreži ili broj citata koje znanstveni rad ima.

S k_i označavamo stupanj od čvora i. U neusmjerenoj mreži ukupan broj veza L može biti prikazan kao:

$$L = \frac{1}{2}\sum_{i=1}^{N} k_i. \tag{3.1}$$

Imamo faktor $1/2$ ispred sume jer se svaka veza broji dva puta.

Jedno od bitnih svojstava mreže prosječan je stupanj čvora. On nam govori koliko je dobro mreža povezana. Za neusmjerenu mrežu on glasi:

$$\langle k \rangle = \frac{1}{N}\sum_{i=1}^{N} k_i = \frac{2L}{N}. \tag{3.2}$$

U usmjerenoj mreži razlikujemo dolazni k_i^{in} i odlazni k_i^{out} stupanj čvora. Dolazni je jednak broju veza usmjerenih u čvor dok odlazni odgovara broju veza koje su usmjerene iz čvora. Ukupan stupanj čvora je:

$$k_i = k_i^{in} + k_i^{out}. \tag{3.3}$$

Ukupan broj dolaznih i odlaznih veza međusobno je jednak jer je svaka odlazna veza jednog čvora dolazna veza drugoga i obrnuto. Ukupan broj veza u usmjerenoj mreži jednak je:

$$L = \sum_{i=1}^{N} k_i^{in} = \sum_{i=1}^{N} k_i^{out}. \tag{3.4}$$

Vidljivo je da nema više faktora $1/2$ jer se odvojeno broje odlazne i dolazne veze.

Prosječni stupanj mreže za nepovezanu mrežu je:

$$\langle k^{in} \rangle = \frac{1}{N} \sum_{i=1}^{N} k_i^{in} = \frac{L}{N}. \tag{3.5}$$

Distribucija stupnja čvora p_k vjerojatnost je da čvor ima određen broj stupnjeva k. Jednaka je omjeru broja čvorova N_k sa stupnjem k i ukupnog broja čvorova N.

$$p_k = \frac{N_k}{N}. \tag{3.6}$$

Distribucija stupnja čvora određuje mnoge fenomene mreža i posebno je bitna u razmatranju scale-free mreža.

3.3 Matrica susjedstva

Za potpuno opisivanje mreže potrebno je pratiti sve njezine veze. Veze se mogu pratiti preko liste koja sadrži sve moguće veze. Drugi način praćenja veza matrica je susjedstva. Ona je pogodnija za matematičko razmatranje. Matrica mreže s N čvorova ima N redaka i N stupaca. Za elemente matrice usmjerene mreže vrijedi sljedeće:

- $A_{ij} = 1$ ako postoji veza iz čvora i u čvor j.
- $A_{ij} = 0$ ako ne postoji veza iz čvora i u čvor j.

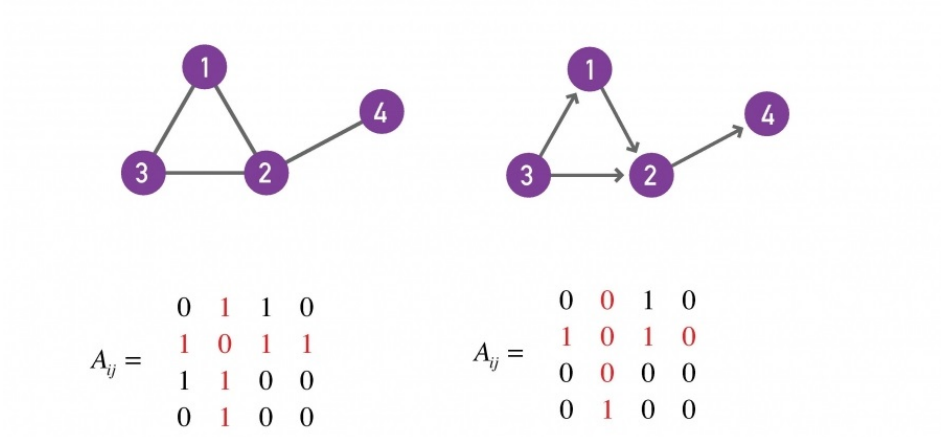
Za neusmjerene mreže, ako postoji veza između dva čvora i i j, vrijedi $A_{ij} = A_{ji} = 1$. Matrica za neusmjerene mreže je simetrična. Na Slici 3.2 su prikazani grafovi i pripadajuće matrice za usmjerene i neusmjerene mreže. Za usmjerene mreže vrijedi da su dolazni i odlazni stupnjevi čvora jednaki redom sumi po redovima i stupcima matrice.

$$k_i^{in} = \sum_{j=1}^{N} A_{ji}, \quad k_i^{out} = \sum_{j=1}^{N} A_{ij}. \tag{3.7}$$

Isto vrijedi i za neusmjerene mreže.

$$k_i = \sum_{j=1}^{N} A_{ij} = \sum_{j=1}^{N} A_{ji}. \tag{3.8}$$

Do sada smo samo razmatrali mreže kojima sve veze imaju identičnu težinu jednaku 1. Veze mogu imati i različite težine. Mreže s različitim težinama nazivamo težinske



Slika 3.2: Na slici je s lijeve strane prikazana neusmjerena mreža i njezina matrica susjedstva, dok je s desne strane prikazana usmjerena matrica i njezina matrica susjedstva.

mreže. U takvim mrežama svaka veza ima pripadajuću težinu w_{ij}. Matrični element za težišne mreže poprima vrijednost težine veze.

$$A_{ij} = w_{ij}. \tag{3.9}$$

Težinske mreže često su prisutne u znanosti. One se koriste jer bolje opisuju veze među čvorovima i pomoću njih je moguće bolje opisati željeni sustav.

3.4 Putevi i udaljenosti

Najkraći put između dva čvora onaj je put koji sadrži najmanje veza. Najkraći put često nije jedinstven, tj. može postojati više najkraćih puteva između dva čvora i ne može presijecati sam sebe. Udaljenost d_{ij} između čvorova jednaka je najkraćem putu.

U neusmjerenim mreža udaljenost među čvorovima i i j jednaka je udaljenosti između j i i ($d_{ij} = d_{ji}$) dok u usmjerenim mrežama to ne mora biti slučaj. Za netežinske mreže udaljenost je jednaka broju veza koje se prijeđu od cilja do odredišta dok je za težinske jednaka sumi inverza težina veza koje se prijeđu.

Udaljenost među čvorovima lako je odrediti u malim mrežama dok je u većima to zahtjevnije. Za računanje udaljenosti u većim mrežama koristi se BFS (Breadth-First Search) algoritam.

Mreža je povezana ako iz svakog čvora mreže možemo doći do bilo kojeg drugog čvora mreže. U slučaju da ne možemo povezati barem dva čvora mreža je nepove-

zana. To znači da je udaljenost među njima beskonačna $d_{ij} = \inf$. Nepovezane mreže možemo razdvojiti na podskupove od kojih je svaki povezan, ali ti podskupovi nisu međusobno povezani.

3.5 Koeficijent grupiranja

Lokalni koeficijent grupiranja C_i govori koliki je stupanj povezanosti između čvorova s kojima je povezan čvor i. Definiran je kao:

$$C_i = \frac{2L_i}{k_i(k_i - 1)}, \tag{3.10}$$

gdje je k_i stupanj čvora i, a L_i je broj veza između k_i susjeda čvora i. Koeficijent C_i poprima vrijednost između nula i jedan. Minimalnu vrijednost postiže kada nema veza među susjedima, a maksimalnu kad su svi susjedi povezani. Ova vrijednost zapravo predstavlja šansu da su dva čvora u susjedstvu povezana. Što je susjedstvo povezanije, koeficijent je veći.

Prosječni koeficijent grupiranja mreže $\langle C \rangle$ dan je kao prosjek lokalnih koeficijenata grupiranja:

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^{N} C_i. \tag{3.11}$$

On govori kolika je vjerojatnost da su bilo koja dva čvora koja su u susjedstvu jednog čvora međusobno povezana.

3.6 Centralnost

U teoriji mreža pomoću indikatora centralnosti određujemo najbitnije čvorove u mreži. Postoji nekoliko mjera centralnosti koje se koriste, a neke su od najbitnijih stupanj čvora, bliskost i međupovezanost.

Stupanj čvora jedna je od mjera centralnosti. On govori s koliko čvorova je povezan dani čvor. Čvorovi s velikim stupnjem direktno su povezani s najviše čvorova. Možemo ga koristiti za identifikaciju najbitnijih čvorova.

Bliskost je mjera centralnosti koja govori koliko je čvor blizu svih ostalih čvorova u mreži. Jednaka je inverzu prosječne udaljenosti čvora od svih ostalih čvorova. Iz čvorova s velikom bliskosti se jako brzo dođe do svih ostalih čvorova u mreži. Bliskost

c_i čvora i računamo kao:

$$c_i = \frac{N-1}{\sum_j d_{ij}}, \tag{3.12}$$

gdje je N broj čvorova u mreži, a d_{ij} udaljenost između čvorova.

Međupovezanost je mjera centralnosti koja govori koliko najkraćih puteva u mreži prolazi preko određenog čvora. Ona se koristi za određivanje najvažnijih čvorova za povezivanje mreže. Međupovezanost je dana sljedećim izrazom:

$$b_i = \sum_{j,k} \frac{n^i_{j,k}}{n_{j,k}}, \tag{3.13}$$

gdje je $n_{j,k}$ broj najkraćih puteva između j i k, a $n^i_{j,k}$ broj najkraćih puteva između j i k koji prolaze preko i.

4 Pregled modela

4.1 Podaci

Analizirane su vremenske serije cijena raznih kriptovaluta. Na tržištu kriptovaluta su jake fluktuacije cijena. Kroz povijest kriptovaluta je došlo do nekoliko naglih skokova i padova cijene koji su pogodili cijelo tržište. Teško je predvidjeti kretanje cijene samih kriptovaluta, jer su sustavi u pitanju jako složeni. Također je prisutna velika korelacija među različitim kriptovalutama i specifična događanja na jednoj kriptovaluti se vrlo lako mogu odraziti na cijelo tržište. Posljedica toga može biti simultani krah tržišta svih kriptovaluta.

Analizirane kriptovalute su: bitcoin, ethereum, litecoin, monero, ripple, bitcoin cash, dash, zcash i starcoin. Promatrano je razdoblje od dvije i pol godine zaključeno s 15. studenog 2018. godine. Dostupni su dnevni podaci o cijeni i volumenu kriptovaluta s kojima se trgovalo. Podaci su direktno preuzeti s mjenjačnice za kriptovalute `https://poloniex.com`. Promatrano je kretanje dnevne prosječne cijene kriptovalute. Cijene su izražene u američkim dolarima.

Za analiziranje kriptovaluta na određeni datum korišten je vremenski prozor od prethodnih 30 dana. Same cijene kriptovaluta generalno imaju eksponencijalni trend rasta cijena. Potrebno je eliminirati taj efekt i svesti vremenski niz koji je promatran na isti red veličine kroz cijelu povijest. Da bi se to postiglo, promatrana je razlika logaritma cijene $S(t)$ između dva susjedna dana.

$$R(t) = \log_2 S(t) - \log_2 S(t-1). \tag{4.1}$$

Dobiveni vremenski nizovi korištene su za daljnju analizu korelacija među kriptovalutama. Računate su transferne entropije među svim kombinacijama kriptovaluta te je promatrana njihova međupovezanost. Rezultati su izračunati za cijelu povijest i analizirano je njihovo kretanje kroz vrijeme. Konstruirane su usmjerene mreže pomoću dobivenih veza i promatran je utjecaj pojedinih kriptovaluta unutar dobivene mreže. Pokazano je kako dobiveni rezultati mogu biti korišteni kao rani signali za predviđanje faznih prijelaza.

4.2 Procjena vjerojatnosti

Cijene kriptovaluta nisu diskretne vrijednosti i povijest koja je dostupna za promatranje ograničena je. To nam predstavlja problem pri izračunu vjerojatnosti potrebnih za računanje transferne entropije. Da bi se taj problem riješio, potrebno je vrijednosti smjestiti u kategorije prema intervalu unutar kojeg pripadaju. Zbog dinamike kretanja cijena kripotvaluta korišten je vremenski prozor za promatranje vremenskog niza. To čini promatrani vremenski niz relativno kratkim i javljala bi se velika greška kada bismo cijene smjestili u više kategorija. Zbog toga je vremenski niz smješten u dvije kategorije, jednu kad je došlo do porasta cijene i drugu kad je došlo do pada cijene.

Kako bi se ipak sačuvala informacija o jačini promjene cijene, korišteni su koncepti iz teorije fuzzy skupova. Definirana je sigmoid funkcija pripadnosti svakoj od dvije kategorije na temelju veličine koju promatramo. Pripadnost pozitivnoj kategoriji μ_+ definirana je kao:

$$\mu_+ = \frac{1}{1 + e^{-\alpha R(t)}}. \tag{4.2}$$

Pripadnost negativnoj kategoriji definirana je kao:

$$\mu_- = \frac{e^{-\alpha R(t)}}{1 + e^{-\alpha R(t)}}. \tag{4.3}$$

α je koeficijent strmosti simgoid funkcije. Vidljivo je da za sve vrijednosti $R(t)$ vrijedi:

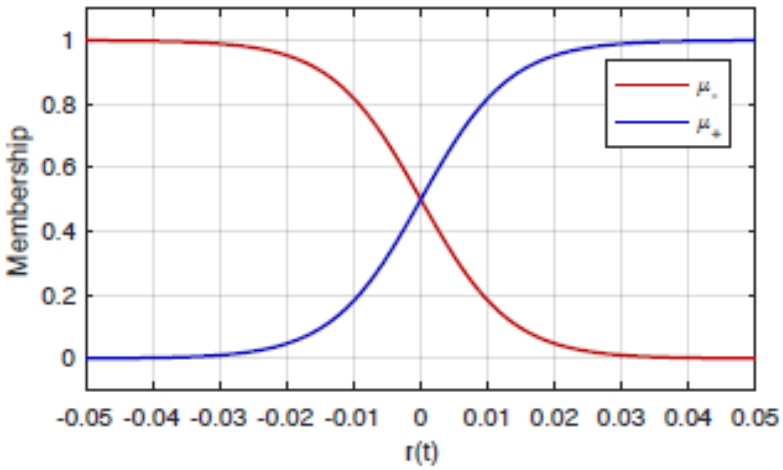
$$\mu_+ + \mu_- = 1. \tag{4.4}$$

Vidljivo je da na temelju veličine $R(t)$ sigmoid funkcija odredi kolika je pripadnost određenoj kategoriji. Za jako negativne vrijednosti $R(t)$ pripadnost negativnoj kategoriji μ_- poprima vrijednost blizu 1 i pripadnost pozitivnoj kategoriji μ_+ je blizu 0 dok je za jako pozitivne vrijednosti obrnuta situacija. U slučaju da $R(t)$ poprimi vrijednost 0, pripadnost za obje kategorije jednaka je i iznosi 0.5. Vjerojatnost p_+ da će vrijednost $R(t)$ biti pozitivna jednaka je srednjoj vrijednosti μ_+.

$$p_+ = \frac{1}{n} \sum \mu_+, \tag{4.5}$$

gdje je n broj članova vremenskog niza. Isto vrijedi i za p_- samo s μ_-.

$$p_- = \frac{1}{n}\sum \mu_-. \tag{4.6}$$



Slika 4.1: Sigmoid funkcija pripadnosti za koeficijent strmosti $\alpha = 150$.

Ovim pristupom, za razliku od smještanja u kategorije po intervalu, uzima se u obzir magnituda promjene. Pri računanju vjerojatnosti veću težinu od malih imaju velike promjene te više utječu na krajnji rezultat. Na primjer, u situaciji u kojoj imamo podjednak broj negativnih i pozitivnih vrijednosti $R(t)$, ali je apsolutni iznos pozitivnih vrijednosti u prosjeku dosta veći od negativnih, vjerojatnost p_+ će biti veća od p_-. Ovom metodom održan je nizak broj kategorija i šum je sveden na minimum dok je ipak sačuvana informacija o jačini promjene cijene.

4.3 Transferna entropija

Normalizirana transferna entropija korištena je za određivanje usmjerenih veza između različitih kriptovaluta. Ona nam govori koliko informacije o kriptovaluti X sadrži kriptovaluta Y, a da ta informacija nije sadržana u X. Što je transferna entropija veća, to je veća povezanost među kriptovalutama. Vremenski nizovi kriptovaluta promatrani su kao Markovljevi procesi gdje buduća vrijednost ovisi samo o jednom prethodnom stanju kriptovaluta ($k = 1$ i $l = 1$). Vrijednosti su računate u vremenskom prozoru od 30 dana.

Uvjetna entropija za kriptovalutu X o svojoj prošlosti kad uvrstimo $k = 1$ u (2.30)

jest:

$$H(X_{t+1}|X_t) = -\sum P(i_{n+1}, i_n) \log_2 P(i_{n+1}|i_n). \tag{4.7}$$

Transferna entropija iz jednadžbe (2.32), kada se uvrsti $k = 1$ i $l = 1$, svodi se na sljedeći izraz:

$$T_{Y\to X} = \sum P(i_{n+1}, i_n, j_n) \log_2 \frac{P(i_{n+1}|i_n, j_n)}{P(i_{n+1}|i_n)}. \tag{4.8}$$

Efektivna transferna entropija je računata koristeći 100 različitih permutacija za računanje srednje vrijednosti nasumične transferne entropije $T^{(r)}_{Y\to X}$. Normalizirana transferna entropija izračunata je po formuli (3.34).

$$T^{(n)}_{Y\to X} = \frac{T^{(e)}_{Y\to X}}{H(X_{t+1}|X_t)} = \frac{T_{Y\to X} - E[T^{(r)}_{Y\to X}]}{H(X_{t+1}|X_t)}. \tag{4.9}$$

Da bi se mogla promatrati veza kriptovalute o samoj sebi, korištena je međusobna informacija I_X. Jednaka je razlici entropije kriptovalute i uvjetne entropije kriptovalute o svojoj prošlosti.

$$I_X = H(X_{t+1}) - H(X_{t+1}|X_t) = \sum P(i_{n+1}, i_n) \log_2 \frac{P(i_{n+1}|i_n)}{P(i_{n+1})P(i_n)}. \tag{4.10}$$

Međusobna informacija I_X poprima minimalnu vrijednost 0 kada buduća vrijednost od X uopće ne ovisi o svojoj prošlosti dok maksimalnu vrijednost od $H(X_{t+1})$ poprima kada prošlost u potpunosti opisuje budućnost.

Efektivna međusobna informacija računa se na sličan način kao efektivna transferna entropija. Prvo se nasumično ispremješa vremenski niz te se računa srednja vrijednost za 100 različitih permutacija. Ta se vrijednost oduzima od međusobne informacije da se dobije efektivna međusobna informacija.

$$I^{(e)}_X = I_X - E[I^{(r)}_X]. \tag{4.11}$$

Normalizirana međusobna informacija dobije se kada se efektivna međusobna informacija podijeli s maksimalnom vrijednosti $H(X_{t+1})$.

$$I^{(n)}_X = \frac{I^{(e)}_X}{H(X_{t+1})} = \frac{I_X - E[I^{(r)}_X]}{H(X_{t+1})}. \tag{4.12}$$

Normalizirana međusobna informacija korištena je za određivanje veze kriptovalute

na sebe samu.

Normalizirana transferna entropija i međusobna informacija su izračunate za sve moguće kombinacije i dostupne datume za dani vremenski prozor. Pomoću njih je promatrana povezanost među različitim kriptovalutama kroz vrijeme i pokušalo se odrediti ukazuju li na nagle promjene na tržištu kriptovaluta.

4.4 Usmjerena mreža

Da bismo odredili usmjerenu težinsku mrežu za izabrane kriptovalute, potrebno je odrediti veze među njima. Za jačinu veze uzimana je vrijednost normalizirane transferne entropije za sve kombinacije kriptovaluta i normalizirana međusobna informacija za vezu svake kriptovalute sa samom sobom. Veza postoji samo ako je jačina veze veća od 0. Vezu možemo definirati na sljedeći način:

$$w_{Y,X} = \begin{cases} \max(T^{(n)}_{Y\to X}, 0), & \text{ako } Y \neq X \\ \max(I^{(n)}_{X}, 0), & \text{ako } Y = X. \end{cases} \tag{4.13}$$

Očigledno će veza postojati ako je efektivna transferna entropija veća od 0, to jest kada je transferna entropija veća od srednje vrijednosti transferne entropije za nasumičan redoslijed vremenske serije Y ($T_{Y\to X} > E[T^{(r)}_{Y\to X}]$). Isto vrijedi za međusobnu informaciju ($I_X > E[I^{(r)}_X]$). Dobivena matrica susjedstva koristila se za konstruiranje usmjerene težinske mreže.

Dobivene su usmjerene težinske mreže za različite točke u vremenu. Mreža je povezanija što ima više veza i što su te veze jače. Promatran je omjer ostvarenih i mogućih mreža kroz vrijeme, kao i prosječna jačina veze u mreži. Postotak ostvarenih veza je korišten umjesto broja veza zbog toga što nisu sve kriptovalute postojale kroz cijelo razdoblje.

Izračunat je i koeficijent grupiranja za sve kriptovalute u mreži (3.10). Koristeći (3.11) izračunat je prosječni koeficijent grupiranja. On govori kolika je vjerojatnost da su dvije kriptovalute u susjedstvu druge kriptovalute međusobno povezane.

Povezanost među kriptovalutama se može prikazati preko bliskosti. To je mjera centralnosti koja govori koliko je blizu jedna kriptovaluta ostalim kriptovalutama u mreži i računa se prema formuli (3.12). Izračunata je za sve kriptovalute u mreži i promatran je njezin prosjek kroz vrijeme za cijelu mrežu kriptovaluta.

Očekivano je da će doći do porasta povezanosti između kriptovaluta u situacijama kada je prisutna nagla promjena na cijelom tržištu kriptovaluta. Kako bi se testirala ta hipoteza, uveden je indeks povezanosti. Indeks povezanosti definiran je kao suma težina svih veza u mreži podijeljena s maksimalnim mogućim brojem veza. Za usmjerenu težinsku mrežu gdje čvorovi imaju vezu na samog sebe koeficijent povezanosti je:

$$S = \frac{1}{N^2} \sum_{Y \in C} \sum_{X \in C} w_{Y,X}. \tag{4.14}$$

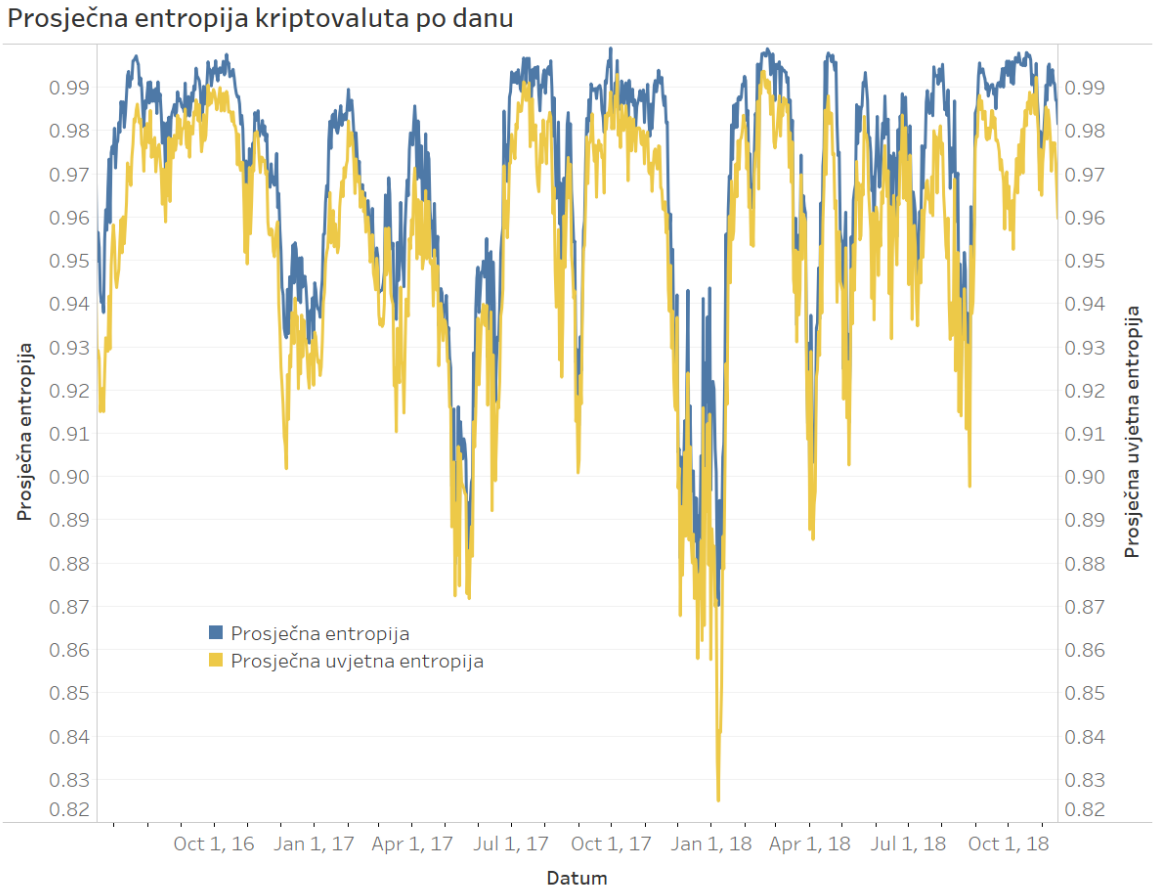
Promatrano je ukazuje li indeks povezanosti u situacijama faznog prijelaza sustava na samu promjenu. Da bi indeks povezanosti bio kvalitetan rani signal, potrebno je da na situaciju ukaže prije samog faznog prijelaza i da signal bude intenzivniji od signala za ostatak promatranog razdoblja. Izračunat je indeks povezanosti za svaki korak koji je promatran i uspoređen s kretanjem cijene bitcoina. Bitcoin je izabran jer čini većinu tržišne kapitalizacije kriptovaluta i glavni je pokazatelj kretanja tržišta.

5 Rezultati

5.1 Entropija

Analizirani su povijesni podaci o cijeni 9 kriptovaluta: bitcoin, ethereum, litecoin, monero, ripple, bitcoincash, dash, zcash i starcoin. Promatrano je razdoblje od protekle dvije i pol godine. Za računanje vrijednosti korišten je vremenski prozor od 30 dana. Izračunate su razlike logaritma cijene susjednih dana $R(t)$ te pripadnosti pozitivnoj μ_+ i negativnoj μ_- kategoriji. Koeficijent strmosti je postavljen na 150. Pomoću dobivenih pripadnosti dobivene su vjerojatnosti za računanje potrebnih entropija.

Prije računanja transferne entropije prvo je bilo potrebno izračunati vrijednosti entropije $H(X_t)$ i uvjetne entropije $H(X_{t+1}|X_t)$ za sve kriptovalute. Maksimalna vrijednost koju entropija može poprimiti je 1, a minimalna 0. Vrijednosti prosječne entropije i uvjetne entropije kroz vrijeme mogu se vidjeti na Slici 5.1.



Slika 5.1: Graf prosječne entropije $H(X_t)$ i uvjetne entropije $H(X_{t+1}|X_t)$ za sve kriptovalute kroz vrijeme.

Iz grafa je vidljivo da obje entropije poprimaju dosta visoke vrijednosti. To znači

da je u sustavu prisutna velika neuređenost i teško je predvidljiv. Također, uvjetna entropija poprima manje vrijednosti od entropije jer je u njoj sadržano više informacija o sustavu.

5.2 Transferna entropija

U sljedećem koraku izračunate su vrijednosti transferne entropije $T_{Y\to X}$ svih kombinacija parova kriptovaluta prema formuli (4.8). Na Slici 5.2 može se vidjeti prosječna vrijednost transferne entropije za sve kriptovalute kroz vrijeme.

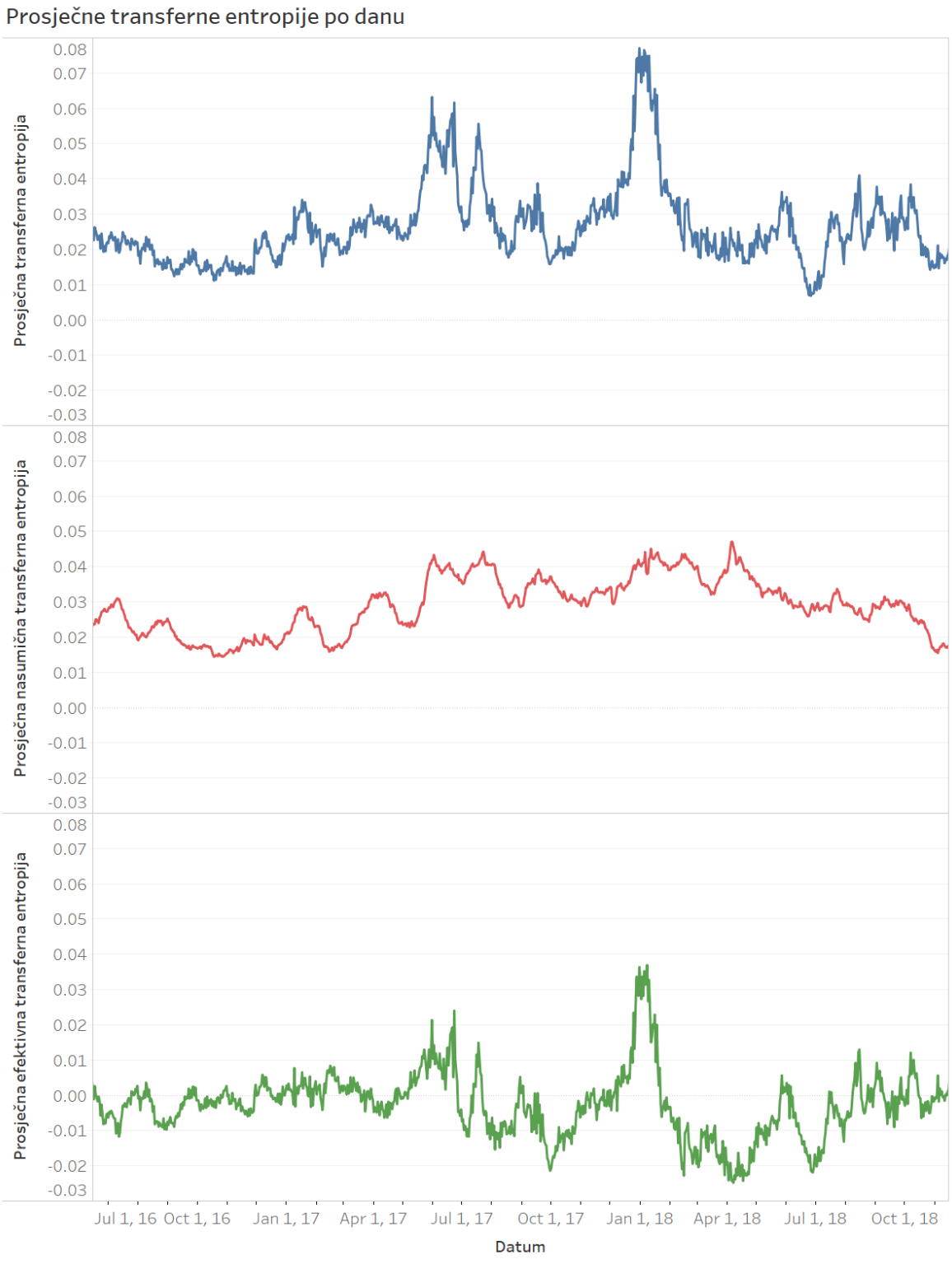
Vidljivo je nekoliko izraženijih vrhova vrijednosti prosječne transferne entropije kroz vrijeme, a najizraženiji se nalazi oko kraja 2017. i početka 2018. godine. Prosječna transferna entropija poprima maksimalnu vrijednost na dan 31. prosinca 2017. godine.

Transferna entropija govori koliko nove informacije o budućnosti cijene kriptovalute X sadrži povijest kriptovalute Y. Zbog toga vrhovi vrijednosti prosječne transferne entropije ukazuju da je u tom trenutku prisutna jaka povezanost cijena među kriptovalutama i da prošlost pojedinih kriptovaluuta može pomoći u predviđanju budućih vrijednosti kriptovaluta.

Kako bi se iz transferne entropije eliminirao šum zbog ograničenosti broja podataka i velike volatilnosti cijena, bilo je potrebno izračunati transfernu entropiju. Prvi korak bio je računanje transferne entropije $T^{(r)}_{Y\to X}$ za nasumičan redoslijed vremenskog niza unutar vremenskog prozora. Postupak je ponovljen 200 puta kako bi se dobila srednja vrijednost $T^{(r)}_{Y\to X}$. Dobivene srednje vrijednosti $E[T^{(r)}_{Y\to X}]$ oduzete su od transfernih entropija za sve parove kriptovaluta i dobivena je efektivna transferna entropija.

Prosječne vrijednosti za sve transferne entropije prikazane su na Slici 5.2. Transferna entropija za nasumičan redoslijed vremenskog niza sličnog je reda veličine kao i transferna entropija. Kod nje jedino nisu prisutni izraženiji vrhovi jer su oni posljedica specifičnog redoslijeda članova vremenske serije. Njezinim oduzimanjem od transferne entropije dobivena je efektivna transferna entropija.

Vidljivo je da efektivna transferna entropija može poprimiti negativne vrijednosti za razliku od drugih entropija. Oblik krivulje sličan je kao i za transfernu entropiju, ali vrijednosti se kreću oko nule.

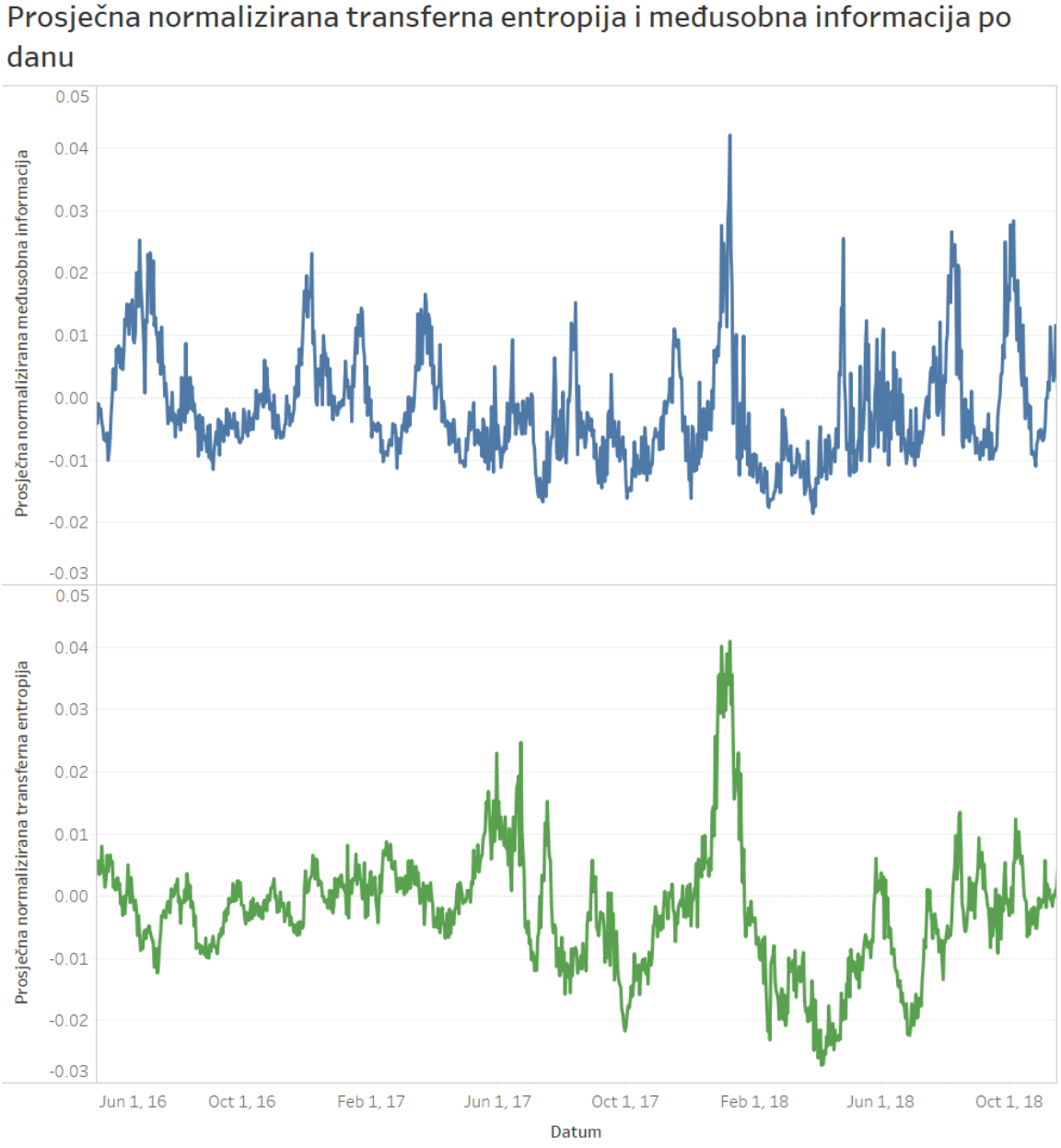


Slika 5.2: Graf prosječne transferne entropije $T_{Y\to X}$, nasumične transferne entropije $T^{(r)}_{Y\to X}$ i efektivne transferne entropije $T^{(e)}_{Y\to X}$ po danu.

Isti postupak je napravljen i za međusobnu informaciju I_X. Izračunata je međusobna informacija I_X kroz vrijeme, kao i međusobna informacija za nasumični vremenski niz $I^{(r)}_X$. Računanjem međusobne informacije za nasumični vremenski $I^{(r)}_X$ na

200 primjera dobije se srednja vrijednost $E(I_X^{(r)})$, pomoću koje se dobije efektivna međusobna informacija $I_X^{(e)}$.

Idući korak je bio normaliziranje dobivenih vrijednosti transferne entropije i međusobne informacije za sve kriptovalute. Normalizirane vrijednosti su izračunate dijeljenjem efektivnih vrijednosti s maksimalnim vrijednostima za transfernu entropiju i međusobnu informaciju. Normalizirana transferna entropija i međusobna informacija su prikazane na Slici 5.3.



Slika 5.3: Graf prosječne entropije $H(X_t)$ i uvjetne entropije $H(X_{t+1}|X_t)$ za sve kriptovalute kroz vrijeme.

Vidljivo je da normalizirana transferna entropija i međusobna informacija imaju

isti red veličine i obje vrijednosti imaju najizraženije vrhove oko početka 2018. godine. Normalizirana međusobna povezanost govori koliki je udio informacije o kretanju cijene kriptovalute sadržan u njezinoj prošlosti dok normalizirana transferna entropija govori koliki je udio informacije o kretanju cijene kriptovalute sadržan u prošlosti druge kriptovalute. Dobivene vrijednosti mogu se koristiti za dobivanje veza između kriptovaluta. Pomoću veza između kriptovaluta konstruirana je usmjerena težinska mreža.

5.3 *Mreža kriptovaluta*

Normalizirana transferna entropija korištena je za dobivanje veza između različitih kriptovaluta dok je međusobna informacija korištena za dobivanje veze kriptovaluta na sebe samu. Korištena je formula (4.13) za dobivanje veza $w_{Y,X}$. Ako je vrijednost normalizirane transferne entropije i međusobne informacije manja ili jednaka nuli, težina veze je nula i veza nije uspostavljena. Ako su vrijednosti veće od nula, veza postoji i težina je jednaka samim vrijednostima. Na Slici 5.4 prikazana je dobivena matrica susjedstva na dan 1.1.2018.

Matrica susjedstva za kriptovalute na 1.1.2018.

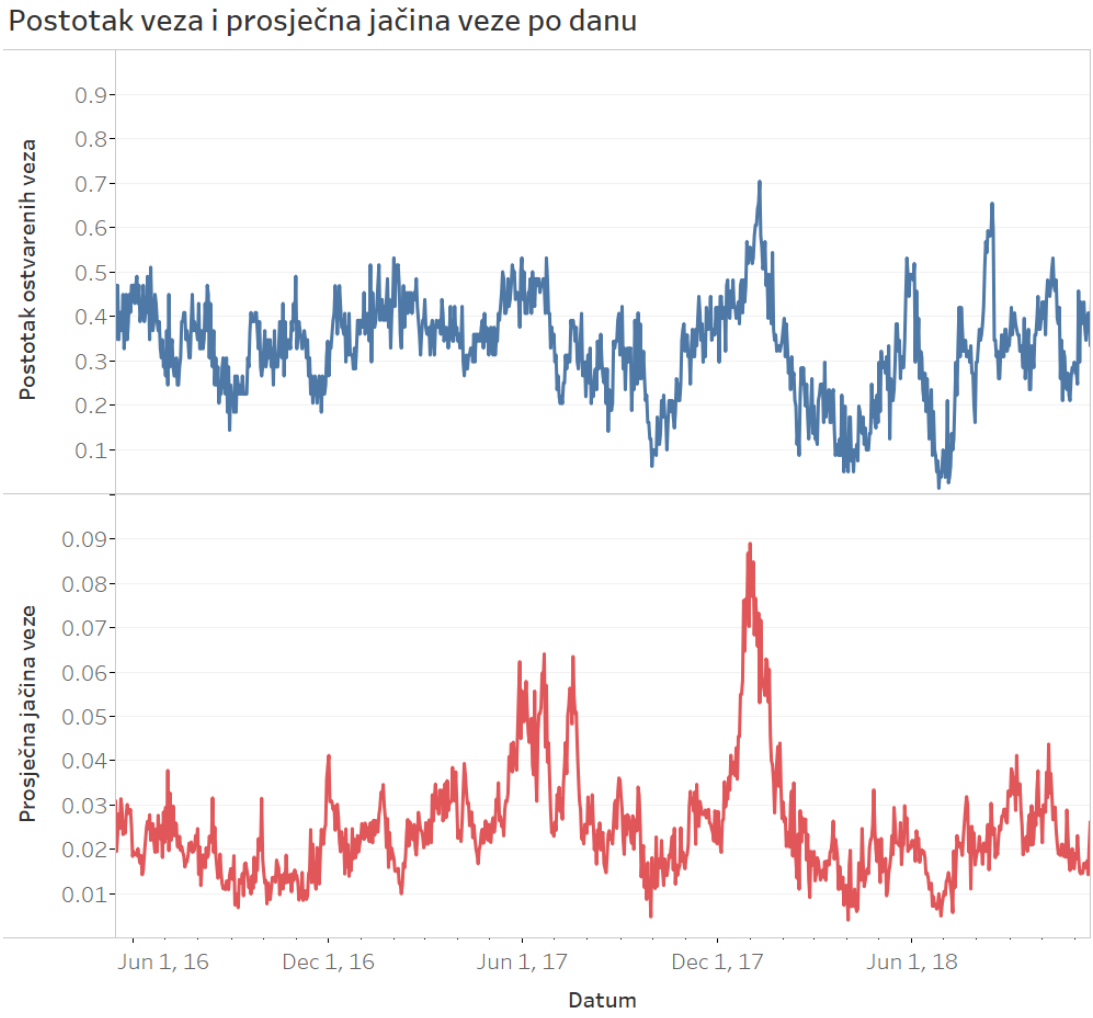
Kriptovaluta Y \ Kriptovaluta X	BCH	BTC	DASH	ETH	LTC	STR	XMR	XRP	ZEC
BCH	0.0000	0.0716	0.0795	0.0449	0.0000	0.0000	0.1269	0.0149	0.0734
BTC	0.0000	0.0000	0.0112	0.0410	0.0000	0.0164	0.0000	0.0392	0.0000
DASH	0.0891	0.2469	0.0000	0.0000	0.0514	0.0000	0.0000	0.0000	0.0638
ETH	0.1368	0.1316	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0192
LTC	0.0501	0.0131	0.0186	0.1016	0.0000	0.0000	0.1390	0.0310	0.0394
STR	0.1071	0.0142	0.0441	0.0430	0.0000	0.0995	0.0000	0.0401	0.0000
XMR	0.1993	0.1921	0.0778	0.0509	0.1930	0.0000	0.0933	0.0102	0.0120
XRP	0.0000	0.0000	0.0231	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ZEC	0.1543	0.1793	0.0883	0.0550	0.0730	0.0000	0.0000	0.0000	0.0000

Jačina veze
0.0000 — 0.2469

Slika 5.4: Matrica susjedstva za kriptovalute na 1.1.2018. dobivena pomoću normalizirane transferne entropije $T^{(n)}_{Y\to X}$ i međusobne informacije $I^{(n)}_X$.

Dobivene veze su korištene za konstruiranje usmjerene težinske mreže u različitim točkama u vremenu. Preko mreže je promatrana povezanost kriptovaluta kroz vrijeme. Parametri koji mogu pokazati povezanost mreže broj su veza u mreži i jačina veze u mreži. Kako nisu sve kriptovalute postojale kroz period koji je promatran, umjesto broja veza korišten je omjer broja veza i maksimalnog mogućeg broja veza. Također su izračunate prosječne jačine veza u mreži. Rezultati su prikazani na Slici 5.5.

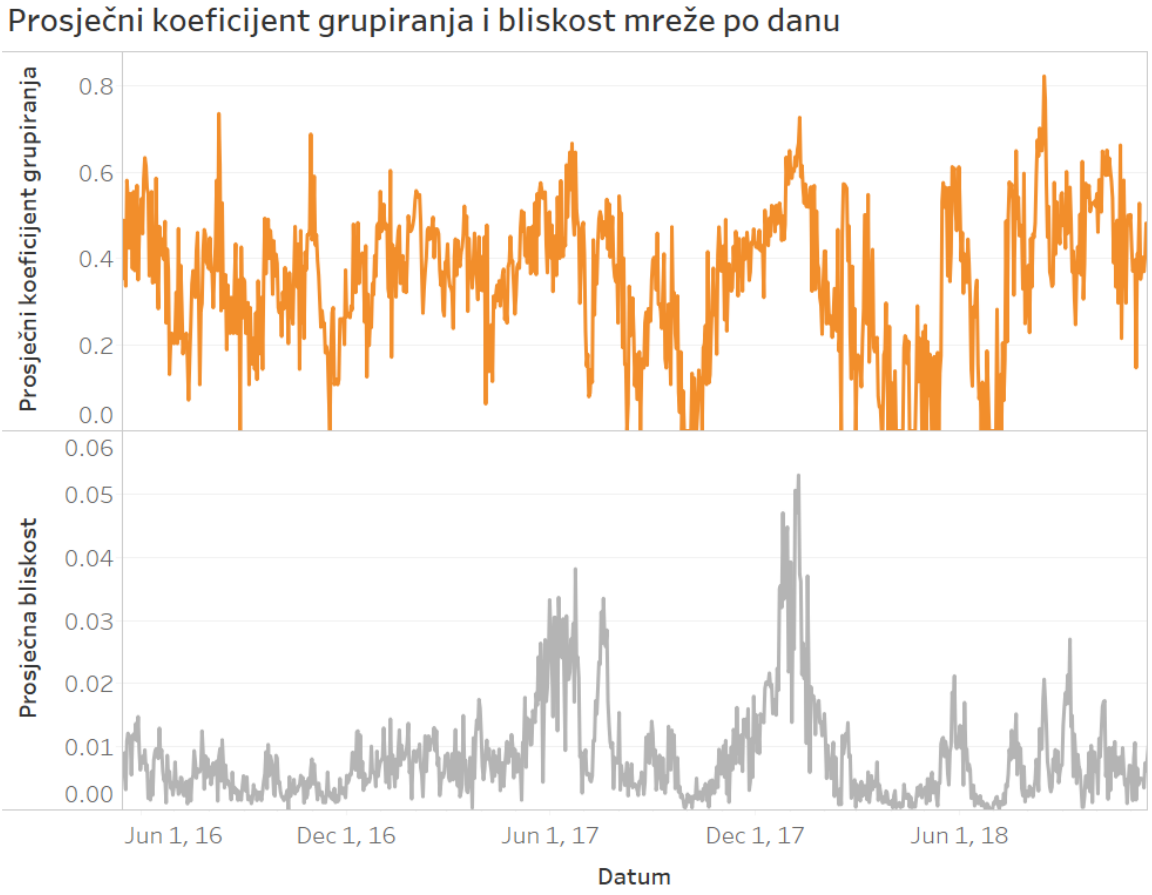
Iz grafa je vidljivo da postotak ostvarenih veza i prosječna jačina veze imaju najizraženiji vrh oko kraja 2017. i početka 2018. godine. Veliki postotak ostvarenih veza prisutan je i kroz 8. mjesec 2018. godine dok je drugi najveći vrh prosječne jačine veze kroz 6. mjesec 2017. godine.



Slika 5.5: Graf postotka ostvarenih veza (plavo) i prosječne jačine veza (crveno) kroz vrijeme.

Nadalje, promatran je prosječni koeficijent grupiranja u mreži. On govori kolika

je vjerojatnost da su prvi susjedi neke kriptovalute u mreži povezani. Za njegovo računanje korištene su formule (3.10) i (3.11). Vrijednosti prosječnog koeficijenta grupiranja mreže kroz vrijeme prikazani su na Slici 5.6. Vidljivo je da vrhovi prosječnog koeficijenta grupiranja nisu toliko izraženi, kao u drugih vrijednosti koje su promatrane. Vrhovi se pojavljuju u istim razdobljima kao i za vrijednosti postotka ostvarenih veza i prosječne jačine veza dok prosječni koeficijent centralnosti ima dva vrha kroz 2016. godinu.



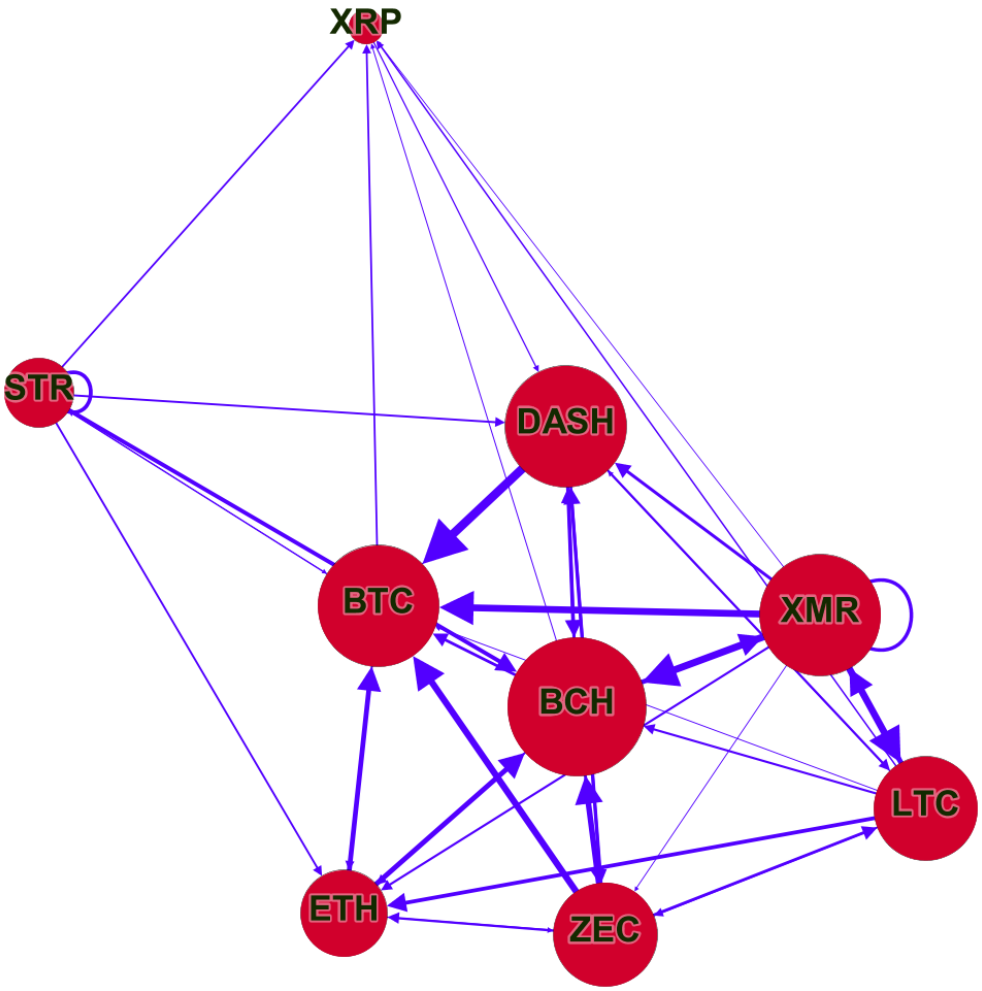
Slika 5.6: Graf prosječnog koeficijenta grupiranja (narančasto) i prosječne bliskosti (sivo) kroz vrijeme.

Izračunata je i bliskost kriptovaluta kroz vrijeme. Ta mjera centralnosti govori koliko je kriptovaluta blizu ostalih kriptovaluta u mreži. Prosječna vrijednost za sve kriptovalute izračunata je u svim vremenskim koracima i prikazana na Slici 5.6. Bliskost, kao i većina promatranih vrijednosti, ima najsnažniji vrh početkom 2018. godine. U tom razdoblju došlo je do naglog pada cijena svih kriptovaluta.

Na Slici 5.7 prikazana je mreža na 1. siječnja 2018. godine. U tom razdoblju kod

većine prethodno promatranih veličina prisutni su najintenzivniji vrhovi vrijednosti. U tom razdoblju ima jako puno veza u mreži i same veze su dosta jake. Debljina veze na slici proporcionalna je samoj jačini veze dok je veličina čvora proporcionalna samom stupnju čvora. Strelicom je označen smjer veze.

Kriptovalute na slici jako su gusto povezane. Jedino su izdvojene starcoin (STR) i ripple (XRP). To je i očekivano jer su starcoin i ripple korištene od strane banaka za transakcije među državama i ponašaju se malo drugačije od ostalih kriptovaluta. Bitcoin (BTC) i bitcoin cash (BCH) su kriptovalute s najvećim stupnjem u mreži. Obje kriptovalute smještene su u sredini mreže i dobro su povezane sa svim ostalim kriptovalutama. U mreži se od jedne do druge kriptovalute dođe u najviše dva koraka i putevi u mreži dosta su kratki.



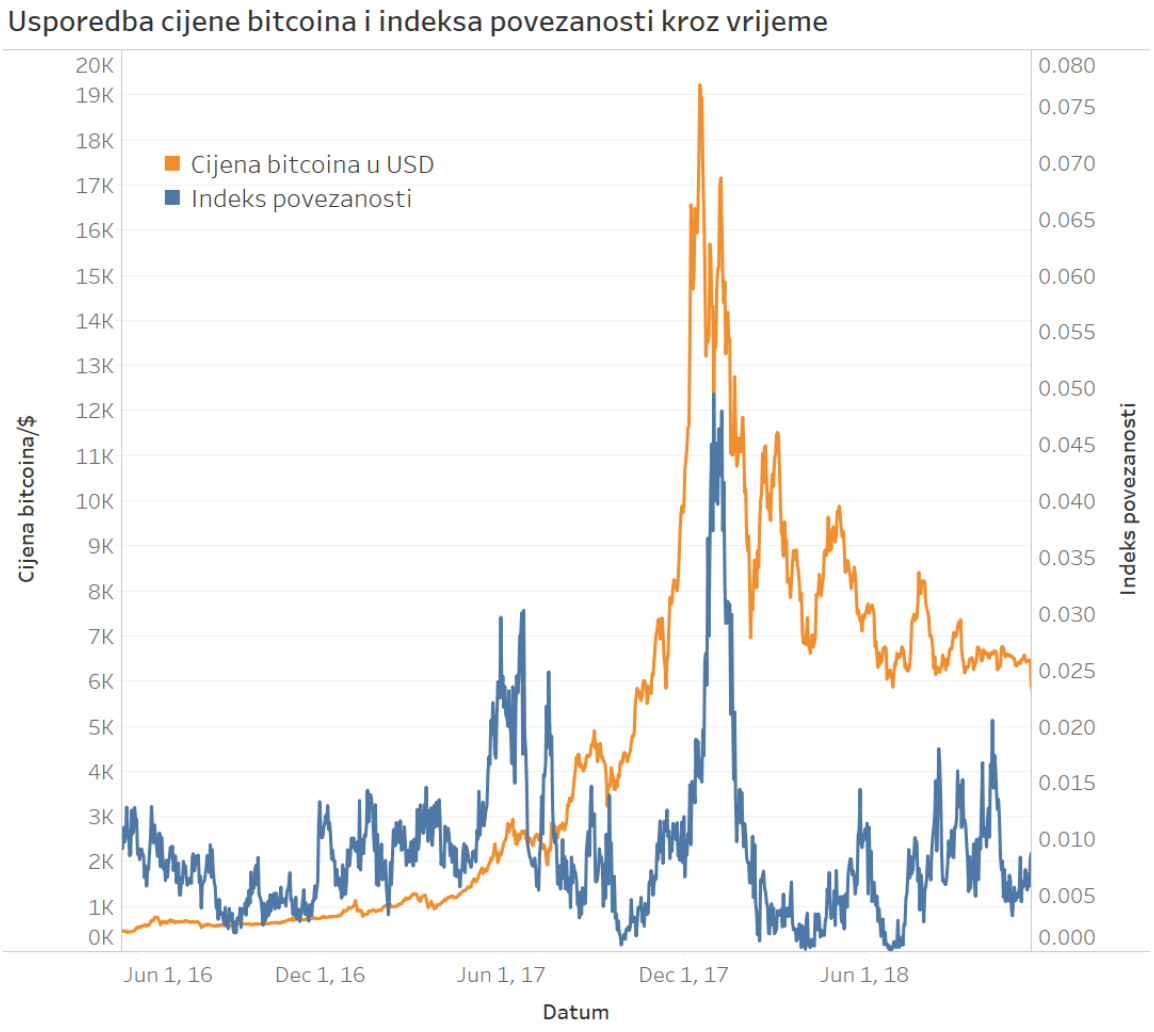
Slika 5.7: Usmjerena težinska mreža kriptovaluta na 1. siječnja 2018. godine.

Izračunat je indeks povezanosti sustava za svaku točku u vremenu. Računat je

preko formule (4.14). Indeks povezanosti poprima veću vrijednost što ima više veza i što je jačina veza veća. Njegova ovisnost o vremenu prikazana je na Slici 5.8.

On govori koliki je udio informacije o kretanju cijena na tržištu kriptovaluta sadržan u samoj povijesti tržišta. Očekivano je da se, kada on postigne velike vrijednosti, trenutno kretanje cijena na tržištu odražava na samu budućnost tržišta. U tim situacijama sve su kriptovalute jako povezane. Također se efekti koji zahvate jedne kriptovalute odražavaju na buduće kretanje cijena drugih kriptovaluta. U tim je situacijama prisutan utjecaj koji mijenja sustav iz jednog stanja u drugo.

Da bi se dana hipoteza potvrdila, potrebno je usporediti kretanje indeksa povezanosti i kretanje cijena bitcoina. Bitcoin je izabran jer čini preko pola kapitalizacije tržišta kriptovaluta. Indeks povezanosti i cijena bitcoina uspoređeni su na Slici 5.8. Vidljivo je da se najveći vrh vrijednosti indeksa povezanosti poklapa s trenutkom kada



Slika 5.8: Usmjerena težinska mreža kriptovaluta na 1. lipnja 2018. godine.

je bitcoin poprimio najveću vrijednost u svojoj povijesti. Nakon tog trenutka dolazi do naglog pada cijene same kriptovalute, koja je u mjesec dana izgubila preko 50% svoje vrijednosti. U tom trenutku bitcoin prelazi iz dugoročnog trenda rasta cijena koji je bio prisutan u 2017. godini u dugoročni trend pada cijena u 2018. godini. Isti je efekt primjetan na svim drugim kriptovalutama.

Drugi porast vrijednosti indeks povezanosti ima krajem svibnja i početkom lipnja 2017. godine. U tom periodu je počeo pad cijene bitcoina koji iznosio 30% tadašnje vrijednosti i trajao je do sredine srpnja.

Treći vrh vrijednosti događa se sredinom srpnja. U tom trenutku nastupa nagli rast cijene bitcoina, gdje mu se vrijednost udvostručila.

Iz ovih je rezultata empirijski pokazano da indeks povezanosti može ukazivati na promjene stanja sustava. Signali koji su prisutni događaju se ranije nego što je došlo do faznog prijelaza sustava, što omogućuje njegovo predviđanje. Također su vrijednosti koje postiže indeks povezanosti u tim razdobljima veće nego u ostatku povijesti pa je sam signal lako razaznati. Ovi rezultati potvrđuju danu hipotezu da se pomoću informacijske povezanosti kriptovaluta mogu dobiti rani signali koji ukazuju na naglu promjenu faze sustava.

6 Zaključak

U ovom radu predstavljena je metoda analiziranja kompleksnog sustava radi prepoznavanja ranih signala prije faznog prijelaza. Promatrane su statističke ovisnosti među članovima sustava koje ukazuju na promjenu na makroskopskom nivou.

Metoda je primjenjena na sustavu kriptovaluta. Promatrane su cijene kriptovaluta u posljednje dvije i pol godine. Veze između samih kriptovaluta su dobivene koristeći koncepte transferne entropije i međusobne informacije. Kako bi se eliminirala njihova pozitivna pristranost i šum, korišteni su njihovi normalizirani oblici. Pomoću tih vrijednosti dobivene su usmjerene težinske veze između kritpovaluta kroz vrijeme.

Promatrane su dobivene mreže kriptovaluta u različitim vremenskim točkama i analizirano je kretanje postotka ostvarenih veza i njihove prosječne jačine. Izračunati su prosječni koeficijenti grupiranja i bliskost unutar mreže. Vidljivo je da gotovo sve promatrane vrijednosti imaju najintenzivniji vrh vrijednosti u razdoblju između samog kraja 2017. godine i početka 2018. godine.

Iz prethodnih rezultata je uočeno da dolazi do porasta vrijednosti promatranih parametara u trenutcima turbulencije na tržištu. Očekivano je da će mreže biti gušće prilikom događaja koji djeluju na cijelo tržište nego u mirnom razdoblju. Kako bi se testirala ta hipoteza, uveden je indeks povezanosti koji govori kolika je jako vezan sustav. Izračunati indeks povezanosti uspoređen je s kretanjem cijene bitcoina kroz vrijeme. Vidljivo je da se dobiveni signali preko indeksa povezanosti događaju prije naglih promjena cijene bitcoina. Također, u tim razdobljima indeks povezanosti poprima značajno veće vrijednosti nego u preostalom razdoblju.

Navedeno ukazuje da predložena metodologija daje rane signale preko kojih se mogu predvidjeti fazni prijelazi na tržištu kriptovaluta i otvara mogućnost primjene na drugim složenim sustavima.

Dodaci

Dodatak A Kod

```
#UČITAVANJE SVIH POTREBNIH PAKETA
import networkx as nx
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import warnings
from tqdm import tqdm
warnings.filterwarnings("ignore")
#UČITAVANJE PODATAKA
import json
from urllib.request import urlopen

#DEFINIRANJE FUNKCIJE ZA POVLAČENJE PODATAKA O KRIPTOVALUTAMA S POLONEIXA
def GetData(name='BTC'):
   #UČITAVANJE JSONA S POLONIEXA, PERIOD ZA CIJENE JEDAN DAN
   url = 'https://poloniex.com/public?command=returnChartData&currencyPair=USDT_'
   +name+'&start=0&end=15230592000&period=86400'
   openUrl = urlopen(url)
   # PARSIRANJE JSONA U PANDAS DATAFRAME
   r = openUrl.read()
   openUrl.close()
   d = json.loads(r.decode())
   df = pd.DataFrame(d)
   # returning origninal meaning to data
   original_columns=[u'close', u'date', u'high', u'low', u'open',
   u'volume', u'quoteVolume']
   new_columns = ['Close','Timestamp','High','Low','Open','Volume_USD','Volume']
   df = df.loc[:,original_columns]
   df.columns = new_columns
   return(df)
```

```
#DEFINIRANJE FUNKCIJA ZA POTREBNIH ZA RACUNANJE
def sigmoid(x):
  return 1 / (1 + np.exp(-x))

def alfa(x):
   return 150*x

#DEFINIRANJE FUNKCIJE ZA RAČUNANJE PROSJEČNE VRIJEDNOSTI U NEKOM TIME WINDOWU
def time_window(df, old_column,new_column,window_size):
   temp=df[old_column]
   df[new_column]=temp
   for i in range(1,window_size):
      df[new_column]=df[new_column]+temp.shift(i+1)
   df[new_column]=df[new_column]/window_size
   return df

#POSTAVLJANJE ŽELJENOG TIME WINDOWA ZA KOJI ĆEMO RAČUNAT ENTROPIJU
window_size=30
#KRIPTOVALUTE KOJE PROMATRAMO
crypto=["ETH","BTC","XMR","DASH","LTC","XRP","BCH","ZEC",'STR']

#UČITAVANJE PODATAKA ZA KRIPTOVALUTE I RAČUNANJE ENTROPIJA
final=pd.DataFrame()
for c in crypto:
#bit=pd.read_csv("bitcoin_price.csv")
   bit=GetData(c)
   bit["Crypto"]=c
   bit=bit[bit['Open']!=0]
   bit=bit[bit['Volume']!=0]
   bit=bit.sort_values(["Timestamp"])
   bit["Date"]=pd.to_datetime(bit["Timestamp"], unit='s')

   bit["Average"]=bit["Volume_USD"]/bit["Volume"]
   bit["Average"]=np.log2(bit["Average"]/bit["Average"].shift(1))
   bit=bit[-bit["Average"].isnull()]
```

```
    bit["u+"]=sigmoid(alfa(np.array(bit["Average"])))
    bit["u-"]=1-bit["u+"]
    bit["u++"]=bit["u+"]*bit["u+"].shift(1)
    bit["u+-"]=bit["u+"]*bit["u-"].shift(1)
    bit["u-+"]=bit["u-"]*bit["u+"].shift(1)
    bit["u--"]=bit["u-"]*bit["u-"].shift(1)
    bit=time_window(bit, "u++","avg_u++",window_size)
    bit=time_window(bit, "u+-","avg_u+-",window_size)
    bit=time_window(bit, "u-+","avg_u-+",window_size)
    bit=time_window(bit, "u--","avg_u--",window_size)
    bit=time_window(bit, "u+","avg_u+",window_size)
    bit=time_window(bit, "u-","avg_u-",window_size)



    p=bit["avg_u--"]/bit["avg_u-"].shift(1)
    bit["CONDITIONAL_ENTROPY"]=-bit["avg_u--"]*np.log2(p)
    p=bit["avg_u+-"]/bit["avg_u-"].shift(1)
    bit["CONDITIONAL_ENTROPY"]=bit["CONDITIONAL_ENTROPY"]-bit["avg_u+-"]*np.log2(p)
    p=bit["avg_u-+"]/bit["avg_u+"].shift(1)
    bit["CONDITIONAL_ENTROPY"]=bit["CONDITIONAL_ENTROPY"]-bit["avg_u-+"]*np.log2(p)
    p=bit["avg_u++"]/bit["avg_u+"].shift(1)
    bit["CONDITIONAL_ENTROPY"]=bit["CONDITIONAL_ENTROPY"]-bit["avg_u++"]*np.log2(p)
    p=bit["avg_u+"]
    bit["ENTROPY"]=-p*np.log2(p)
    p=bit["avg_u-"]
    bit["ENTROPY"]=bit["ENTROPY"]-p*np.log2(p)
    final=final.append(bit)
#FUNKCIJA ZA RAČUNANJE TRANSFER ENTROPIJE
def transfer_entropy(df,window_size):
    df["TRANSFER_ENTROPY"]=df["CONDITIONAL_ENTROPY_x"]
    state=["-","+"]
    for k in state:
        for j in state:
            df["p"+j+k]=df["u"+j+"_x"]*df["u"+k+"_y"]
            df=time_window(df, "p"+j+k,"avg_p"+j+k,window_size)
```

```
        for i in state:
            df["u"+i+j+k]=df["u"+i+"_x"]*df["u"+j+"_x"].shift(1)*
            *df["u"+k+"_y"].shift(1)
            df=time_window(df, "u"+i+j+k,"avg_u"+i+j+k,window_size)
            p=df["avg_u"+i+j+k]/df["avg_p"+j+k].shift(1)
            df["TRANSFER_ENTROPY"]=df["TRANSFER_ENTROPY"]+
            +df["avg_u"+i+j+k]*np.log2(p)
    return df

#FUNKCIJA ZA RAČUNANJE TRANSFERNE ENTROPIJE ZA
#RANDOMIZIRAN SLJED VREMENSKE SERIJE
def randomize_transfer_entropy(df,window_size,random_samples):
    random_entropy = np.zeros(shape= (window_size+1))
    tab_a=np.array(df[["u+_x","u-_x"]])
    tab_b=np.array(df[["u+_y","u-_y"]])
    for i in range((window_size+1),df.shape[0]):
        a=tab_a[i-window_size-1:i,:]
        b=tab_b[i-window_size-1:i,:]
        t=np.array(range(0,window_size+1))
        entropy=0
        #start=time.time()
        for r in range(random_samples):
            #np.random.shuffle(t)
            #a=a[t,:]
            np.random.shuffle(t)
            b=b[t,:]
            for k in range(2):
                p2=np.mean(a[:window_size,k])
                for l in range(2):
                    p23=np.mean(a[:window_size,k]*b[:window_size,l])
                    for j in range(2):
                        p123=np.mean(a[1:,j]*a[:window_size,k]*b[:window_size,l])
                        p12=np.mean(a[1:,j]*a[:window_size,k])
                        entropy=entropy+p123*np.log2((p123*p2)/(p23*p12))
        #print(time.time()-start)
```

```
      random_entropy=np.append(random_entropy,entropy)
   df["RANDOM_ENTROPY"]=random_entropy/random_samples
   return df

#FUNKCIJA ZA RAČUNANJE KONDICIJSKE ENTROPIJE ZA
#RANDOMIZIRAN SLJED VREMENSKE SERIJE
def randomize_entropy(df,window_size,random_samples):
   random_entropy = np.zeros(shape= (window_size+1))
   tab_a=np.array(df[["u+","u-"]])
   for i in range((window_size+1),df.shape[0]):
      a=tab_a[i-window_size-1:i,:]
      t=np.array(range(0,window_size+1))
      entropy=0
      #start=time.time()
      for r in range(random_samples):
         np.random.shuffle(t)
         a=a[t,:]
         for k in range(2):
            p2=np.mean(a[:window_size,k])
            for j in range(2):
               p12=np.mean(a[1:,j]*a[:window_size,k])
               p1=np.mean(a[1:,j])
               entropy=entropy+p12*np.log2((p12)/(p2*p1))
      #print(time.time()-start)
      random_entropy=np.append(random_entropy,entropy)
   df["RANDOM_ENTROPY"]=random_entropy/random_samples
   return df

#POSTAVLJAMO NA KOLIKO ĆE SE RANDOM PRIMJERA VRTITI RAČUNANJE EFEKTIVNE ENTROPIJE
#!!!OVAJ BROJ SMANJITI AKO SE SPORO VRTI!!!
random_samples=200

#PROLAZAK KROZ SVE MOGUĆE KOMBINACIJE KRIPTOVALUTA I
#RAČUNAJE NORMALIZIRANE TRANSFER ENTROPIJE
transfer=pd.DataFrame()
```

```
for i in tqdm(crypto):
   for j in crypto:
      if i!=j:
         temp1=final[final["Crypto"]==i]
         temp2=final[final["Crypto"]==j]
         temp=temp1.merge(temp2,how='inner',left_on="Date", right_on="Date")
         temp=transfer_entropy(temp,window_size)
         temp=randomize_transfer_entropy(temp,window_size,random_samples)
         temp["EFFECTIVE_TRANSFER_ENTROPY"]=
         =(temp["TRANSFER_ENTROPY"]-temp["RANDOM_ENTROPY"])
         temp["NORMALIZED_TRANSFER_ENTROPY"]=
         =(temp["TRANSFER_ENTROPY"]-temp["RANDOM_ENTROPY"])/
         /(temp["CONDITIONAL_ENTROPY_x"])
         transfer=transfer.append(temp[["Crypto_x",
         "Crypto_y",'CONDITIONAL_ENTROPY_x',
         'ENTROPY_x',"Date","TRANSFER_ENTROPY","RANDOM_ENTROPY",
         "EFFECTIVE_TRANSFER_ENTROPY","NORMALIZED_TRANSFER_ENTROPY"]])
      else:
         temp=final[final["Crypto"]==i]
         temp["Crypto_x"]=i
         temp["Crypto_y"]=i
         temp['ENTROPY_x']=temp['ENTROPY']
         temp['CONDITIONAL_ENTROPY_x']=temp['CONDITIONAL_ENTROPY']
         temp["TRANSFER_ENTROPY"]=-temp["CONDITIONAL_ENTROPY"]+temp["ENTROPY"]
         temp=randomize_entropy(temp,window_size,random_samples)
         temp["EFFECTIVE_TRANSFER_ENTROPY"]=
         =(temp["TRANSFER_ENTROPY"]-temp["RANDOM_ENTROPY"])
         temp["NORMALIZED_TRANSFER_ENTROPY"]=
         =(temp["TRANSFER_ENTROPY"]-temp["RANDOM_ENTROPY"])/
         /(temp["ENTROPY"])
         transfer=transfer.append(temp[["Crypto_x",
         "Crypto_y",'CONDITIONAL_ENTROPY_x',
         'ENTROPY_x',"Date","TRANSFER_ENTROPY","RANDOM_ENTROPY",
         "EFFECTIVE_TRANSFER_ENTROPY","NORMALIZED_TRANSFER_ENTROPY"]])
```

```
transfer=transfer[-transfer["NORMALIZED_TRANSFER_ENTROPY"].isna()]

avg_transfer=transfer.copy()

avg_transfer=avg_transfer.groupby(["Date"]).agg({"TRANSFER_ENTROPY":"mean",
"NORMALIZED_TRANSFER_ENTROPY":"mean"}).reset_index()

plt.xlabel("Datum")
plt.ylabel("Prosječna  transfer entropija")
plt.plot(avg_transfer["Date"].iloc[300:],
avg_transfer["TRANSFER_ENTROPY"].iloc[300:],color='green')

plt.xlabel("Datum")
plt.ylabel("Prosječna normalizirana transfer entropija")
plt.plot(avg_transfer["Date"].iloc[300:],
avg_transfer["NORMALIZED_TRANSFER_ENTROPY"].iloc[300:])

transfer["NORMALIZED_TRANSFER_ENTROPY"]
[transfer["NORMALIZED_TRANSFER_ENTROPY"]<0]=0
transfer["CONNECTION"]=0
transfer["CONNECTION"][transfer["NORMALIZED_TRANSFER_ENTROPY"]>0]=1

avg_transfer=transfer.copy()
avg_transfer=avg_transfer.groupby(["Date"]).agg({"TRANSFER_ENTROPY":"mean",
"NORMALIZED_TRANSFER_ENTROPY":"mean","CONNECTION":"sum"}).reset_index()

plt.xlabel("Datum")
plt.ylabel("Broj veza")
plt.plot(avg_transfer["Date"].iloc[300:],
avg_transfer["CONNECTION"].iloc[300:],color='red')

avg_con=transfer[(-transfer["NORMALIZED_TRANSFER_ENTROPY"].isna())
& (transfer["CONNECTION"]==1)]
avg_con=avg_con.groupby(["Date"]).agg({"TRANSFER_ENTROPY":"mean",
"NORMALIZED_TRANSFER_ENTROPY":"mean","CONNECTION":"sum"}).reset_index()
```

```
plt.xlabel("Datum")
plt.ylabel("Prosječna jačina veze")
plt.plot(avg_con["Date"].iloc[300:],
avg_con["NORMALIZED_TRANSFER_ENTROPY"].iloc[300:],color='red')

#ODABIR DATUMA KOJI ŽELIMO PROMATRATI
day=pd.to_datetime("20180201")
#ODABIR VEZA ZA TAJ DATUM
sna=transfer[transfer["Date"]==day]
sna=sna[sna["CONNECTION"]==1]
sna=sna.rename(columns={"NORMALIZED_TRANSFER_ENTROPY":"weight"})
#DEFINIRANJE DVOSMJERNOG GRAFA
G = nx.DiGraph()
#DODAVANJE VEZA ZA KRITPOVALUTE
for _,row in sna.iterrows():
   G.add_edge(row["Crypto_y"], row["Crypto_x"], weight=row["weight"])
#ODREĐIVANJE POZICIJA ZA ČVOROVE
pos = nx.spring_layout(G,weight='weight')
#ODREĐIVANJE STUPNJEVA ZA SVAKI ČVOR I PRIPISIVANJE VELIČINE
#ČVORU PROPORCIONALNE STUPNJU
degree = nx.degree(G)
degree=[20*degree[node] for node in G.nodes()]

# CRTANJE ČVOROVA
nx.draw_networkx_nodes(G, pos,node_size=degree, node_color='red')
#PRIPISIVANJE DEBLJINE VEZE PROPORCIONALNO JAČINI VEZE
edges = G.edges()
weights = [G[u][v]['weight']*10 for u,v in edges]
#CRTANJE VEZA
nx.draw_networkx_edges(G, pos, arrows='True',edge_cmap=plt.cm.Blues,
width=weights,arrow_size=10)
# CRTANJE LABELA
nx.draw_networkx_labels(G, pos, font_size=10, font_family='sans-serif')
#PLOTANJE
plt.axis('off')
```

```
plt.show()
edges = G.edges()
for u,v in edges:
   G[u][v]['distance']=1/G[u][v]['weight']

#FUNKCIJA ZA RAČUNANJE KOEFICIJENTA GRUPIRANJA
def clustering_coefficient(crypto,sna):
   f={}
   for c in crypto:
      temp=sna[sna["Crypto_x"]!=sna["Crypto_y"]]
      temp=temp[(temp["Crypto_x"]==c) | (temp["Crypto_y"]==c)]
      un=pd.concat([temp["Crypto_x"],temp["Crypto_y"]]).unique()
      if len(un)<3:
         f[c]=0
      else:
         temp=sna[sna["Crypto_x"]!=sna["Crypto_y"]]
         temp=temp[(temp["Crypto_x"].isin(un)) & (temp["Crypto_y"].isin(un))]
         temp=temp[(temp["Crypto_x"]!=c) & (temp["Crypto_y"]!=c)]
         f[c]=temp.shape[0]/((len(un)-2)*(len(un)-1))
   return f

#POMOĆNE FUNKCIJE
def fill_data(c,final,day,cnt_type):
   temp=pd.DataFrame.from_dict(c,orient='index')
   temp["Crypto"]=temp.index
   temp["Centrality"]=temp[0]
   temp["Date"]=day
   temp["Centrality_type"]=cnt_type
   del temp[0]
   final=final.append(temp)
   return final

def add_zero(c,crypto):
   l=list(c.keys())
   for i in crypto:
```

```
      if i not in l:
         c[i]=0
   return c

#RAČUNANJE RAZLIČITIH CENTRAALNOSTI ZA SVE DATUME
dates=transfer["Date"].unique()
centrality=pd.DataFrame()

for day in tqdm(dates):
   sna=transfer[transfer["Date"]==day]
   crypto=pd.concat([sna["Crypto_x"],sna["Crypto_x"]]).unique()
   sna=sna[sna["CONNECTION"]==1]
   sna=sna.rename(columns={"NORMALIZED_TRANSFER_ENTROPY":"weight"})
   if sna.shape[0]>1:
      G = nx.DiGraph()
      #DODAVANJE VEZA ZA KRITPOVALUTE
      for _,row in sna.iterrows():
         G.add_edge(row["Crypto_y"], row["Crypto_x"], weight=row["weight"])
      edges = G.edges()
      for u,v in edges:
         G[u][v]['distance']=1/G[u][v]['weight']
      #STUPANJ ČVORA
      c=nx.degree_centrality(G)
      c=add_zero(c,crypto)
      centrality=fill_data(c,centrality,day,"degree_centrality")
      #ULAZNI STUPANJ ČVORA
      c=nx.in_degree_centrality(G)
      c=add_zero(c,crypto)
      centrality=fill_data(c,centrality,day,"in_degree_centrality")
      #IZLAZNI STUPANJ ČVORA
      c=nx.out_degree_centrality(G)
      c=add_zero(c,crypto)
      centrality=fill_data(c,centrality,day,"out_degree_centrality")
      #closeness
      c=nx.closeness_centrality(G,distance='distance')
```

```
        c=add_zero(c,crypto)
        centrality=fill_data(c,centrality,day,"closeness_centrality")
        #BETWEENNESS
        c=nx.betweenness_centrality(G,weight='weight')
        c=add_zero(c,crypto)
        centrality=fill_data(c,centrality,day,"betweenness_centrality")
        #KOEFICIJENT GRUPIRANJA
        c=clustering_coefficient(crypto,sna)
        c=add_zero(c,crypto)
        centrality=fill_data(c,centrality,day,"clustering_coefficient")

#SPREMANJE REZULTATA
centrality.to_excel("centrality.xlsx")
transfer.to_excel("transfer.xlsx")
crypto_data=final[['Close', 'Timestamp', 'High', 'Low', 'Open', 'Volume_USD',
'Volume','Crypto', 'Date', 'Average','CONDITIONAL_ENTROPY', 'ENTROPY']]
crypto_data.to_excel("crypto_data.xlsx")
```

Bibliography

[1] Marten Scheffer, Jordi Bascompte, William A. Brock, Victor Brovkin, Stephen R. Carpenter, Vasilis Dakos, Hermann Held, Egbert H. van Nes, Max Rietkerk & George Sugihara: Early-warning signals for critical transitions, Nature 461, 53-59 (03 September 2009)

[2] Lenton, T. M. et al. Tipping elements in the Earth's climate system. Proc. Natl Acad. Sci. USA 105, 1786-1793 (2008).

[3] Kambhu, J., Weidman, S. & Krishnan, N. New Directions for Understanding Systemic Risk: A Report on a Conference Cosponsored by the Federal Reserve Bank of New York and the National Academy of Sciences (The National Academies Press, 2007).

[4] Litt, B. et al. Epileptic seizures may begin hours in advance of clinical onset: a report of five patients. Neuron 30, 51-64 (2001).

[5] Schroeder, M. Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise (Freeman, 1991).

[6] Wissel, C. A universal law of the characteristic return time near thresholds. Oecologia 65, 101-107 (1984).

[7] Van Nes, E. H. & Scheffer, M. Slow recovery from perturbations as a generic indicator of a nearby catastrophic shift. Am. Nat. 169, 738-747 (2007).

[8] Ovaskainen, O. & Hanski, I. Transient dynamics in metapopulation response to perturbation. Theor. Popul. Biol. 61, 285-295 (2002).

[9] Ives, A. R. Measuring resilience in stochasticsystems. Ecol. Monogr. 65, 217-233 (1995).

[10] Biggs, R., Carpenter, S. R. & Brock, W. A. Turning back from the brink: detecting an impending regime shift in time to avert it. Proc. Natl Acad. Sci. USA 106, 826-831 (2009).

[11] Guttal, V. & Jayaprakash, C. Changing skewness: an early warning signal of regime shifts in ecosystems. Ecol. Lett. 11, 450-460 (2008).

[12] Berglund, N. & Gentz, B. Metastability in simple climate models: pathwise analysis of slowly driven Langevin equations. Stoch. Dyn. 2, 327-356 (2002).

[13] Schroeder, M. Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise (Freeman, 1991).

[14] R. K. Pathria, Paul D. Beale: Statistical Mechanics, Third Edition, Elsevier, 2011.

[15] Stephen J. Blundell, Katherine M. Blundell: Concepts in Thermal Physics, Second Edition, Oxford 2009.

[16] David J.C. MacKay: Information Theory, Inference, and Learning Algorithms, Version 7.2 (fourth printing) March 28, 2005

[17] T. Schreiber, Measuring information transfer, Physical Review Letters, vol. 85, pp. 461-464, Jul 2000.

[18] R. Marschinski and H. Kantz, Analysing the information flow between financial timeseries, The European Physical Journal B, vol. 30, pp. 275-281, Nov 2002.

[19] Shannon, C.E. A Mathematical Theory of Communication. Bell Syst. Tech. J 1948, 27.

[20] Schreiber, Thomas: Measuring Information Transfer, Physical Review Letters. 85 (2): 461–464.

[21] S. Begušić, Z. Kostanjčar, D. Kovač, H. E. Stanley, B. Podobnik Information Feedback in Temporal Networks as a Predictor of Market Crashes // Complexity, 2018 (2018), 1-13

[22] Albert-László Barabási: Network Science, Cambridge University Press, 2015

[23] Newman, M.E.J. 2010. Networks: An Introduction. Oxford, UK: Oxford University Press.

[24] `https://poloniex.com/`, Izvor svih podataka korištenih u obradi kriptovaluta.