Satunnaiskävelyt ja niiden kasvattamat puut

Joonas Laukka

Perustieteiden korkeakoulu

Kandidaatintyö
Espoo 1.5.2018

Vastuuopettaja

TkT Riikka Korte

Työn ohjaaja

Prof. Lasse Leskelä

Aalto-yliopisto
Perustieteiden korkeakoulu

Aalto-yliopisto
Perustieteiden korkeakoulu

Aalto-yliopisto, PL 11000, 00076 AALTO
www.aalto.fi

Tekniikan kandidaatintyön tiivistelmä

Tekijä Joonas Laukka	
Työn nimi Satunnaiskävelyt ja niiden kasvattamat puut	
Koulutusohjelma Teknillinen fysiikka ja matematiikka	
Pääaine Matematiikka ja systeemitieteet	**Pääaineen koodi** SCI3029
Vastuuopettaja TkT Riikka Korte	
Työn ohjaaja Prof. Lasse Leskelä	
Päivämäärä 1.5.2018	**Sivumäärä** 18+2 **Kieli** Suomi

Tiivistelmä
Nykymaailma on täynnä erilaisia verkkorakenteita Internetistä sosiaalisiin piireihin. Verkkojen kasvun mallintaminen onkin kiinnittänyt tutkijoiden huomion kuluneen vuosikymmenen aikana. Yksi tällainen kasvumalli on *No Restart Random Walk* (NRRW), jossa satunnaiskävelijä kasvattaa suuntaamatonta verkkoa kulkiessaan siinä. Kävelijä lähtee alkusolmusta, kulkee s:n kaaren läpi satunnaiseen solmuun, jonka naapuriksi lisätään uusi solmu. Uuden solmun asteluku on aina yksi. Tätä prosessia jatketaan satunnaiskävelijän viimeisimmästä tilasta, ja prosessin toistaminen kasvattaa satunnaisen puurakenteen. Mallin ominaisuudet määrää siis askelparametri, s.
NRRW-malli mukailee reaalimaailman verkoille tyypillisiä ominaisuuksia, kuten suosivaa kiinnittymistä (*eng. preferential attachment*). Mallissa satunnaiskävelyprosessi ja verkkoa kuvaavan satunnaisprosessi ovat voimakkaasti riippuvaisia. Prosessin käsittelyä voidaan helpottaa osoittamalla, että se on yhdenmukainen yksinkertaisemman stokastisen prosessin kanssa. Näytän, että NRRW-malli on stokastisesti yhdenmukainen rakentamani riippumattomien, tasajakautuneiden satunnaislukujen määräämän mallin kanssa. Lopuksi käytän tätä vaihtoehtoista esitystapaa puun solmujen palautuvuutta koskevan lauseen todistukseen.

Avainsanat Satunnaiskävely, puu, verkko, NRRW

Aalto-yliopisto
Perustieteiden korkeakoulu

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi
Abstract of the bachelor's thesis

Author Joonas Laukka

Title Random Walks and the Trees they Grow

Degree programme Engineering Physics and Mathematics

Major Mathematics and System Analysis **Code of major** SCI3029

Supervisor DSc (Tech.) Riikka Korte

Advisor Prof. Lasse Leskelä

Date 1.5.2018 **Number of pages** 18+2 **Language** Finnish

Abstract

The modern world is full of networks from the Internet to social circles. Modeling the growth of graphs has caught the attention of many researchers during the last decade. One such growth model is a *No Restart Random Walk* (NRRW). In an NRRW, a random walker grows an undirected graph while it is traversing it. The walker starts from the root vertex and crosses s edges to end up in a random destination vertex. Then, a new vertex with degree 1 is added as its neighbour. This process is repeated starting from the destination vertex and consequently, a random tree is generated. The properties of the model are unambiguously determined by the step parameters, s.

The NRRW model conforms to many properties that are common in real life graphs such as preferential attachment. The random walk process and the random graph process are mutually dependent. Examining these dependent processes can be simplified by proving that they are equivalent with a simpler stochastic process. I will show that the NRRW model is stochastically equivalent to a model seeded by a sequence of independent, uniformly distributed random numbers. Finally, I will use this alternative model to prove a theorem about the recurrence of the vertices in an NRRW generated graph.

Keywords Random walk, graph, tree, NRRW

Esipuhe

Otaniemi, 15.9.2017

Joonas Laukka

Sisällysluettelo

Tiivistelmä 3

Tiivistelmä (englanniksi) 4

Esipuhe 5

Sisällysluettelo 6

Symbolit ja lyhenteet 7

1 Johdanto 8

2 Malli 9
2.1 Yleiskatsaus 9
2.2 Satunnaisverkko ja -kävely 10
2.2.1 Satunnaisverkko 10
2.2.2 Satunnaiskävely 11
2.3 Kokonaisuus 11
2.3.1 Tilajoukko 11
2.3.2 Siirtymätodennäköisyydet 12

3 Stokastinen esitys 13
3.1 Teoria 13
3.2 Propositio 14
3.3 Todistus 14

4 Parillinen prosessi 16
4.1 Astelukujen jakauma 16
4.2 Palautuvuus 17

Viitteet 18

A Esimerkki liitteestä 19

B Toinen esimerkki liitteestä 20

Symbolit ja lyhenteet

Symbolit

$\mathbb{N}_0$	ei-negatiiviset kokonaisluvut $(0, 1, 2, \ldots)$

Operaattorit

$\mathbb{1}\{X\}$	indikaattorifunktio ehdolla X
$\mathbb{P}[X]$	tapahtuman X todennäköisyys

Lyhenteet

NRRW	No Restart Random Walk
i.i.d.	riippumaton ja identtisesti jakautunut

1 Johdanto

TODO: Esimerkkejä reaalimaailman verkoista!

Mallin rakenne pyritään valitsemaan siten, että se mukailee parhaalla mahdollisella tavalla joitain todellisia verkostoja. Monille tosielämän verkoille on tyypillistä suosiva kiinnittyminen (*eng. preferential attachment*) sekä se, että verkon solmujen astejakauma ei ole eksponentiaalisesti rajoitettu. Suosivalla kiinnittymisellä tarkoitetaan satunnaisverkkojen yhteydessä sitä, että uusia solmuja liittyy todennäköisimmin sellaisiin solmuihin, joilla on jo ennestään korkea asteluku. Yksi lupaava kasvumalli on *No Restart Random Walk* (NRRW), jossa satunnaiskävelijä kasvattaa suuntaamatonta verkkoa kulkiessaan siinä. Nimen mukaisesti satunnaiskävely ei ala koskaan alusta, vaan jatkuu katkeamatta verkon kasvaessa. Prosessin ominaisuudet määrää askelparametri, s, ja se kehittyy seuraavanlaisen algoritmin mukaisesti:

0. Aluksi verkossa on yksi solmu, juuri, josta on kaari itseensä. Satunnaiskävelyn alkutila on juurisolmu.
1. Satunnaiskävely ottaa s askelta verkossa.
2. Verkkoon lisätään uusi solmu ja se liitetään satunnaiskävelyn senhetkiseen sijaintiin.
3. Palataan kohtaan 1.

Prosessin simulointi ja matemaattinen analyysi osoittaa, että parametrin s parillisuus vaikuttaa voimakkaasti satunnaiskävelyn ja satunnaisverkon luonteisiin:

Askelparametri	**Satunnaiskävely**	**Satunnaisverkko**
s = 1	Väistyvä	Solmujen asteluku on geometrisen jakauman stokastisesti dominoima.
s = 2n	Palautuva	Sisäsolmujen asteluku ei ole eksponentiaalisesti rajoitettu.

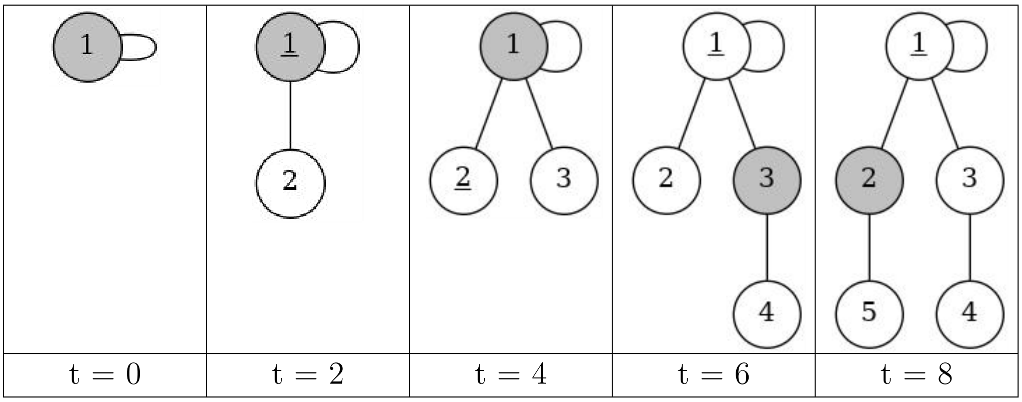
NRRW-prosessi voidaan nähdä joko satunnaisena verkkona tai satunnaiskävelynä. Näkemyksestä riippumatta satunnaiskävelyn palautuvuus ja sen kasvattaman verkon luonne ovat kietoutuneet toisiinsa. Näiden toisistaan riippuvien prosessien käsittelyä voidaan helpottaa osoittamalla, että koko prosessi on yhdenmukainen yksinkertaisemman stokastisen prosessin kanssa. Tässä kandidaatin työssä näytän, että NRRW-prosessi on stokastisesti yhdenmukainen riippumattomien, tasajakautuneiden satunnaislukujen määräämän prosessin kanssa. Lopuksi käytän tätä vaihtoehtoista esitystapaa puun solmujen palautuvuutta koskevan lauseen todistukseen.

2 Malli

2.1 Yleiskatsaus

NRRW-prosessi on kahden toisistaan riippuvan stokastisen prosessin pari, $(\mathcal{G}_t^s, \mathcal{W}_t)_{t\geq 0}$. Molemmat prosessit ovat kytkeytyneet vahvasti toisiinsa, mutta niitä voidaan jaksoittain tarkastella myös erillisinä. Taulukko 1 esittää yhden mahdollisen NRRW-prosessin realisaation askelparametrilla $s = 2$ ja parillisilla ajanhetkillä.

Taulukko 1: Yksi mahdollinen NRRW-prosessin realisaatio askelparametrilla $s = 2$. Satunnaiskävely on harmaan värin osoittamassa tilassa ja uusi solmu on juuri lisätty sen naapuriksi. Satunnaiskävely siirtyi nykyiseen tilaansa alleviivatun solmun kautta.

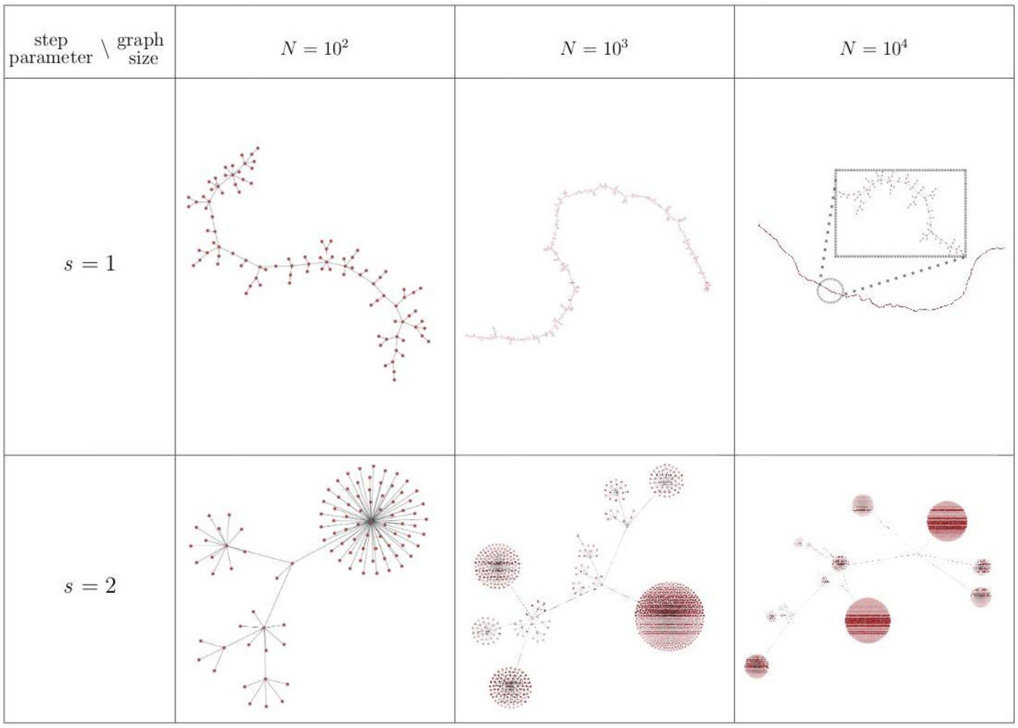


Askelparametri vaikuttaa huomattavasti siihen, millaisia verkkoja NRRW-prosessi kasvattaa. Tämä ilmenee hyvin tietokonesimuloiduista realisaatioista kuvassa 1. Parametrilla $s = 1$ jokaiseen satunnäiskävelyn vierailemaan tilaan lisätään uusi solmu. Tämä johtaa korkeisiin puihin, joissa solmut jakutuvat tasaisesti parillisille ja parittomille tasoille. Satunnaiskävelyn kaikki tilat ovat väistyviä.

Sen sijaan parillinen askelparametri ($s \in 2\mathbb{N}_0$) johtaa hyvin toisenlaisiin verkkoihin. Simulaatiotulokset viittaavat solmujen kasaantuvan suurissa määrin joidenkin yksittäisten solmujen jälkeläisiksi. Kyseessä on niin sanottu suosivan kiinnittymisen ilmiö, jossa jotain mitattavaa suuretta kertyy sellaisille joukon alkioille, joilla on jo ennestään paljon kyseistä suuretta. Satunnaisverkossa tämä ilmenee siten, että uusia solmuja liitetään todennäköisimmin sellaisiin verkon solmuihin, joilla on jo ennestään korkea asteluku.

Parillisen askelparametrin NRRW-prosesseihin liittyy myös toinen erikoisuus. Koska uusi solmuja lisätään vain parillisen askelmäärän välein, niitä kertyy vain verkon parillisille tasoille, ellei satunnaiskävely kulje juurisolmun silmukkakaaren läpi. Silloin satunnaiskävelyn pariteetti vaihtuu ja uusia solmuja lisätään vain parittomille tasoille.

Parillisen askelparametrin NRRW-prosessin ominaisuudet vastaavat paremmin todellisia verkostoja, kuten Internettiä ja sosiaalisia piirejä [2]. Rajoitun siis tarkastelemaan vain parillisen askelparametrin NRRW-prosesseja.



Kuva 1: NRRW-prosessin realisaatioita tietokonesimulaatioissa. Kuva on kopioitu aiemmasta tutkimuksesta (Iacobelli et al.) [1].

2.2 Satunnaisverkko ja -kävely

2.2.1 Satunnaisverkko

$(\mathcal{G}_t^s)_{t\in\mathbb{N}_0}$, on satunnaisverkko, johon lisätään uusi solmu s aikayksikön välein. Satunnaisprosessin tilajoukko on kaikkien suuntaamattomien, yhtenäisten ja syklittömien verkkojen eli puiden joukko.

Määritelmä 1. *Suuntaamaton verkko G on pari (V, E), missä joukko V on verkon solmujen joukko ($v_i \in V$) ja E on suuntaamattomien kaarien joukko. Kaarien joukon alkiot ovat muotoa $\{v_i, v_j\}$. Kaari $\{v_i, v_j\}$ yhdistää solmut v_i ja v_j toisiinsa.*

Kasvavan verkon käsittelyn helpottamiseksi satunnaisverkon $(\mathcal{G}_t^s)_{t\in\mathbb{N}_0}$ solmut indeksoidaan kokonaisluvuilla $(0, 1, 2, \ldots)$ lisäysjärjestyksen mukaan. Juurisolmu on siis v_0 ja hetkelä $t = sn,\ n \in \mathbb{N}_0$ lisätään solmu v_n. Jos satunnaisverkkoa $(\mathcal{G}_t^s)_{t\in\mathbb{N}_0}$ tarkastellaan itsenäisenä prosessina sen siirtymätodennäköisyydet pois nykytilasta ovat nollasta poikkeavia vain ajanhetkillä $t = sn,\ n \in \mathbb{N}_0$. Ne määrää satunnaiskävelyn $(\mathcal{W}_t)_{t\in\mathbb{N}_0}$ siirtymämatodennäköisyydet hetkellä $t = sn$:

$$\begin{aligned}&\mathbb{P}\left[\mathcal{G}_{t+s}^s = (V_t \cup \{v_{j+1}\}, E_t \cup \{(v_i, v_{j+1})\}) \mid \mathcal{G}_t^s = (V_j, E_j), \mathcal{W}_t = v_j\right] \\ &= \mathbb{P}\left[\mathcal{W}_{t+s} = v_i \mid \mathcal{W}_t = v_n, G_t = (V_t, E_t)\right]\end{aligned} \tag{1}$$

2.2.2 Satunnaiskävely

Vastaavasti satunnaiskävelyprosessin $(\mathcal{W}_t)_{t\in\mathbb{N}_0}$ tilajoukon ja siirtymätodennäköisyydet määrää sen kasvattama satunnaisverkko. Sekä tilajoukko että siirtymätodennäköisyydet muuttuvat aina, kun verkkoon lisätään uusi solmu. Lisäyshetkien välillä tilajoukko on kuitenkin vakio ja satunnaiskävelyprosessi voidaankin silloin nähdä tavallisena äärellisen tilajoukon Markov-ketjuna. Jos satunnaiskävely on solmussa i hetkellä t, valitaan seuraava tila satunnaisesti valitsemalla yksi solmusta i lähtevä kaari ja siirtymällä sen osoittamaan solmuun. Todennäköisyysmassa on tasajakautunut kaikkien solmun i kaarien kesken.

Koska NRRW-prosessi rakentaa puita, ei verkossa ole muita syklejä kuin juuren kaari itseensä. Verkkoteoriassa kaaret solmusta itseensä lasketaan kahdesti. Merkitään solmun v_i solmuun v_j yhdistävien kaarien määrää verkossa G_i

$$\psi_{G_i}(v_i, v_j) = \begin{cases} \mathbb{1}\{\{v_i, v_j\} \in E_i\} & \text{jos } v_i \neq v_j \\ 2 \cdot \mathbb{1}\{v_i = v_0\} & \text{muulloin} \end{cases}$$

missä E_i on verkon G_i kaarien joukko. Nyt todennäköisyys siirtyä solmusta v_i solmuun v_j on yksinkertaisesti

$$\mathbb{P}\left[\mathcal{W}_{t+1} = v_j \mid \mathcal{W}_t = v_i, \mathcal{G}_t^s = G_i\right] = \frac{\psi_{G_i}(v_i, v_j)}{d_{G_i}(v_i)} \tag{2}$$

sillä oletuksella, että satunnaiskävely tapahtuu verkossa G_i. Kaavassa $d_{G_i}(v_i)$ on solmun v_i asteluku verkossa G_i. Toisin sanoen

$$d_{G_i}(v_i) = \sum_{w\in V} \psi_{G_i}(v_i, w) \tag{3}$$

eli kaikkien solmuun v_i yhdistyvien kaarien lukumäärä.

2.3 Kokonaisuus

Satunnaisverkon ja satunnaiskävelyn erillinen tarkastelu paljastaa, että näistä jälkimmäinen on NRRW-prosessin todellinen satunnaisuuden lähde. Satunnaiskävely muuttaa liikkeillään sen omaa tilajoukkoa eli satunnaisverkkoa $\mathcal{G}_t^s$. Kun verkko muuttuu, myös satunnaiskävelyn siirtymätodennäköisyydet muuttuvat. Satunnaisverkko sen sijaan kasvaa satunnaiskävelyn realisaation ohjaamana. Kokonaisuutena NRRW-prosessi on siis ajassa epähomogeeninen äärettömän tilajoukon Markov-prosessi. Prosessin epähomogeenisuus johtuu siitä, että satunnaisverkko voi muuttua vain s:llä jaollisilla ajanhetkillä.

2.3.1 Tilajoukko

Olkoon V kaikkien solmujen numeroituvasti ääretön joukko. Jos n:n solmun joukkoa merkitään $V_n = \{v_1, v_2, \ldots, v_n\} \subset V$ niin n:n solmun suuntaamattomien verkkojen joukko on

$$\mathbb{G}_n = \{(V_n, E) \mid E \subset V_n \times V_n\}$$

NRRW-prosessiin kuuluu myös satunnaiskävely, jonka tilajoukko n:n solmun verkkoa kulkiessa sen solmujen joukko V_n. Selvästi n:n solmun NRRW-prosessin tilajoukko on $S_n = V_n \times \mathbb{G}_n$. Kaikkien solmujen joukkoa merkitään numeroituvasti äärettömällä joukolla V. Vastaavasti kaikkien verkkojen joukkoa voidaan merkitä äärettömällä unionilla

$$\mathbb{G} = \bigcup_{n=1}^{\infty} \mathbb{G}_n \tag{4}$$

Koska satunnaisverkko kasvaa NRRW-prosessin edetessä, täytyy sen tilajoukon sisältää kaikki eri solmulukumäärien mahdollistamat tilat. NRRW-prosessin tilajoukko onkin siis

$$S = \bigcup_{n=1}^{\infty} S_n \tag{5}$$

2.3.2 Siirtymätodennäköisyydet

NRRW-prosessin siirtymätodennäköisyydet ovat ajassa epähomogeenisia, sillä satunnaiskävelyn realisaatio kasvattaa sen omaa tilajoukkoa, eli NRRW-prosessin verkkoa vain kun $t \bmod s = 0$. Siirtymätodennäköisyyksiä on siten syytä tarkastella kahdessa eri tapauksessa ajanhetken jaollisuuden mukaan.

Merkitään NRRW-prosessia hetkellä t yksinkertaisesti $(\mathcal{W}_t, \mathcal{G}_t^s)$. Jos $v_i, v_j \in V$ ja $G_i, G_j \in \mathbb{G}$, niin NRRW-prosessin siirtymätodennäköisyydet saavat esityksen

$$\begin{aligned} P_{i,j} &= \mathbb{P}\left[(\mathcal{W}_{t+1}, \mathcal{G}_{t+1}^s) = (v_i, G_i) \mid (\mathcal{W}_t, \mathcal{G}_t^s) = (v_j, G_j)\right] \\ &= \mathbb{P}\left[\mathcal{W}_{t+1} = v_i, \mathcal{G}_{t+1}^s = G_i \mid \mathcal{W}_t = v_j, \mathcal{G}_t^s = G_j\right] \end{aligned}$$

Kun $t + 1 \bmod s \neq 0$, ei prosessin verkkoon lisätä uusia solmuja. Satunnaisverkko $\mathcal{G}_{t+1}^s$ on siis varmasti sama kuin sen edeltäjä $\mathcal{G}_t^s$. Siten

$$\mathbb{P}\left[\mathcal{G}_{t+1}^s = G_j \mid \mathcal{G}_t^s = G_j\right] = 1$$

riippumattomasti satunnaiskävelyn tilasta. Selvästi tapahtuman komplementti ($\mathcal{G}_{t+1}^s \neq G_j$) on mahdoton. Lausekkeen **??** mukaisesti

$$\begin{aligned} P_{i,j} &= \mathbb{P}\left[\mathcal{W}_{t+1} = v_i, \mathcal{G}_{t+1}^s = G_i \mid \mathcal{W}_t = v_j, \mathcal{G}_t^s = G_j\right] \\ &= \mathbb{P}\left[\mathcal{G}_{t+1}^s = G_i \mid \mathcal{G}_t^s = G_j, \mathcal{W}_t = v_j\right] \mathbb{P}\left[\mathcal{W}_{t+1} = v_i \mid \mathcal{W}_t = v_j, \mathcal{G}_{t+1}^s = G_i, \mathcal{G}_t^s = G_j\right] \\ &= \mathbb{P}\left[\mathcal{G}_{t+1}^s = G_i \mid \mathcal{G}_t^s = G_j, \mathcal{W}_t^s = v_i\right] \mathbb{P}\left[\mathcal{W}_{t+1} = v_i \mid \mathcal{W}_t = v_j, \mathcal{G}_{t+1}^s = G_i, \mathcal{G}_t^s = G_j\right] \\ &= \mathbb{1}\{G_i = G_j\}\, \mathbb{P}\left[\mathcal{W}_{t+1} = v_i \mid \mathcal{W}_t = v_j, \mathcal{G}_{t+1}^s = G_i, \mathcal{G}_t^s = G_j\right] \\ &= \mathbb{1}\{G_i = G_j\}\, \frac{\psi^{(t)}(v_i, v_j)}{d_t(v_i)} \end{aligned}$$

missä viimeinen yhtäsuuruus pohjautuu kaavan 2 lausekkeeseen.

Kun $t+1 \bmod s = 0$, satunnaisverkko $\mathcal{G}_{t+1}^s$ määräytyy satunnaiskävelyn $\mathcal{W}_{t+1}$ realisaatiosta.

$$\begin{aligned}
P_{i,j} &= \mathbb{P}\left[\mathcal{W}_{t+1} = v_i, \mathcal{G}_{t+1}^s = G_i \mid \mathcal{W}_t = v_j, \mathcal{G}_t^s = G_j\right] \\
&= \mathbb{P}\left[\mathcal{W}_{t+1} = v_i \mid \mathcal{W}_t = v_j, \mathcal{G}_t^s = G_j\right] \mathbb{P}\left[\mathcal{G}_{t+1}^s = G_i \mid \mathcal{W}_{t+1} = v_i, \mathcal{W}_t = v_j, \mathcal{G}_t^s = G_j\right]
\end{aligned}$$

Uusi solmu lisätään nyt siihen verkon G_j solmuun, johon satunnaiskävely siirtyy ajanhetkellä $t+1$. Selvästi mahdollisia satunnaisverkon $\mathcal{G}_{t+1}^s$ realisaatioita ovat vain sellaiset verkot, jotka on saatu liittämällä verkkoon G_j yksi uusi solmu, $v_{\lfloor \frac{t+1}{s} \rfloor}$. Kun verkkoja G_i ja G_j merkitään verkon määritelmän (1) mukaisina solmujen ja kaarien pareina (V_i, E_i) ja (V_j, E_j),

$$\begin{aligned}
P_{i,j} &= \mathbb{P}\left[\mathcal{G}_{t+1}^s = (V_i, E_i) \mid \mathcal{W}_{t+1} = v_i, \mathcal{W}_t = v_j, \mathcal{G}_t^s = (V_j, E_j)\right] \\
&\quad \cdot \mathbb{P}\left[\mathcal{W}_{t+1} = v_i \mid \mathcal{W}_t = v_j, \mathcal{G}_t^s = (V_j, E_j)\right] \\
&= \mathbb{1}\{(V_i, E_i) = (V_j \cup \{v_i\}, E_j \cup \{\{v_i, v_j\}\})\} \, \mathbb{P}\left[\mathcal{W}_{t+1} = v_i \mid \mathcal{W}_t = v_j, \mathcal{G}_t^s = (V_j, E_j)\right] \\
&= \mathbb{1}\{G_i = (V_j \cup \{v_i\}, E_j \cup \{\{v_i, v_j\}\})\} \, \frac{\psi^{(t)}(v_i, v_j)}{d_t(v_i)}
\end{aligned}$$

3 Stokastinen esitys

3.1 Teoria

Olkoon $(\omega_t)_{t \geq 0}$ jono riippumattomia välin $[0,1]$ tasajakautuneita satunnaislukuja. Toisin sanoen $w_t \sim \mathrm{Tas}(0,1)$. Tällaista satunnaislukujen sarjaa voidaan käyttää monimutkaisemman stokastisen prosessin algoritmiseen simulointiin [3]. Simulointi rakennetaan etsimällä deterministinen funktio $\phi : S \times [0,1] \to S$, joka liittää stokastisen prosessin tilan ja välin $[0,1]$ reaaliluvun prosessin seuraavaan tilaan. Funktion ϕ tulee toteuttaa yhtälö

$$\mathbb{P}\left[\phi(x, \omega_t) = y\right] = \int_0^1 \mathbb{1}\{\phi(x, w) = y\} \,\mathrm{d}w \tag{6}$$

$$= \mathbb{P}\left[X_{t+1} = y \mid X_t = x\right] = P_{x,y} \quad \forall\, x, y \in S \tag{7}$$

Yhtälön 7 toteuttavaa funktiota ϕ ja satunnaislukujonoa $(\omega_t)_{t \geq 0}$ kutsutaan siirtymämatriisin P *stokastiseksi esitykseksi.*

Jos stokastisen prosessin alkutila on $X_0 = x_0$, sitä voidaan simuloida sen stokastisen esityksen avulla rekursiivisesti:

$$X_{t+1} = \phi(X_t, \omega_t) \tag{8}$$

Mikäli stokastinen prosessi on epähomogeeninen ajassa, vaaditaan stokastiseen esitykseen jono deterministisiä funktioita $(\phi^{(t)})_{t \geq 0}$, jotka ovat muotoa $\phi^{(t)} : S \times [0,1] \to S \;\forall\, t \geq 0$. Funktioiden $\phi^{(t)}$ tulee toteuttaa yhtälöä 7 vastaava ehto $\forall\, t \geq 0$:

$$\mathbb{P}\left[\phi^{(t)}(x, \omega_t) = y\right] = \mathbb{P}\left[X_{t+1} = y \mid X_t = x\right] = P_{x,y}^{(t)} \quad \forall\, x, y \in S \tag{9}$$

Prosessin simulointi tehdään, kuten homogeenisessa tilanteessa, mutta rekursiofunktio valitaan funktiojonosta $(\phi^{(t)})_{t\geq 0}$ järjestyksessä:

$$X_{t+1} = \phi^{(t)}(X_t, \omega_t) \tag{10}$$

3.2 Propositio

Edellä havaittiin, että NRRW-prosessi on epähomogeeninen ajassa. Sen stokastinen esitys on siis muotoa $(\phi_t, \omega_t)_{t\geq 0}$, missä funktiot ϕ_t toteuttavat ehdon 9. Muodostetaan sopiva jono deterministisiä funktioita NRRW-prosessille. Prosessin tilat ovat muotoa (v, G), missä $G = (V, E)$ on verkko ja $v \in V$ on verkon G solmu. Jakamalla rekursiofunktiot satunnaiskävely- ja satunnaisverkko-osiin saadaan niille korkeatasoinen muoto

$$\phi^{(t)}\left((v, G), w\right) = \Big(\lambda\left((v, G), w\right), \mu^{(t)}\left(\lambda\left((v, G), w\right), G\right)\Big) \tag{11}$$

missä λ on satunnaiskävelyn $(\mathcal{W}_t)_{t\geq 0}$ stokastisen esityksen rekursiofunktio parametriksi annetussa verkossa G ja μ päivittää verkkoa G satunnaiskävelyn realisaation mukaisesti. Satunnaiskävelyn rekursiofunktio

$$\lambda\left((v_i, G), w\right) = \begin{cases} v_1 & w \in [0, \rho(G)_{i,1}) \\ v_2 & w \in [\rho(G)_{i,1}, \rho(G)_{i,1} + \rho(G)_{i,2}) \\ & \quad \vdots \\ v_{|V|} & w \in [\rho(G)_{i,|V|-1}, \rho(G)_{i,|V|-1} + \rho(G)_{i,|V|}) \end{cases} \tag{12}$$

missä ρ on funktio verkosta sen siirtymämatriisiin ja $|V|$ on verkon G solmujen lukumäärä. Hyödyntäen määritelmän 2.2.2 funktiota solmuja yhdistävien kaarien määrälle, saadaan kuvaukselle ρ alkioittainen määritelmä

$$\rho(G)_{i,j} = \rho\left((V, E)\right)_{i,j} = \frac{\psi^{(t)}(i, j)}{d_G(i)} = \frac{\psi^{(t)}(i, j)}{\sum_{w\in V} \psi^{(t)}(v_i, w)} \tag{13}$$

missä kaksi viimeistä yhtäsuuruutta seuraa kaavoista 2 ja 3.

Stokastisen esityksen verkkoa päivittävä funktio määritellään

$$\mu^{(t)}\left(v, (V, E)\right) = \begin{cases} (V \cup \{v_t\}, E \cup \{\{v, v_t\}\}) & \text{kun } t \bmod s = 0 \\ (V, E) & \text{muulloin} \end{cases} \tag{14}$$

Se siis lisää verkon solmuun v uuden solmun v_t aina kun aika on jaollinen askelparametrillä s. Muulloin verkkoa ei muuteta.

3.3 Todistus

Todistus. Esitän, että stokastisen esityksen $(\phi^{(t)}, \omega_t)_{t\geq 0}$ simuloima stokastinen prosessi on todellakin NRRW-prosessi askelparametrillä s. NRRW-prosessin siirtymätodennäköisyydet on laskettu kappaleessa 2.3.1. Prosessin käyttäytyminen ajanhetkellä

t riippuu siitä, onko t jaollinen askelparametrilla s. Osoitetaan stokastisen esityk-seen yhdenmukaisuus NRRW-prosessin kanssa molemmissa tapauksissa. Olkoon $X_t = (\mathcal{W}_t, \mathcal{G}_t^s)$ NRRW-prosessin tila hetkellä t.

On näytettävä kaavan 9 mukaisesti, että

$$\mathbb{P}\left[\phi^{(t)}(x, \omega_t) = y\right] = \mathbb{P}\left[X_{t+1} = y \mid X_t = x\right] \tag{15}$$

kaikilla $x, y \in S$ ja $t \geq 0$, kuten kaavassa 9. Sijoittamalla aiemmin määritelty stokastinen esitys yhtälön vasempaan puoleen saadaan

$$\begin{aligned}
\mathbb{P}\left[\phi^{(t)}(x, \omega_t) = y\right] &= \int_0^1 \mathbb{1}\left\{\phi^{(t)}(x, w) = y\right\} \mathrm{d}w \\
&= \int_0^1 \mathbb{1}\left\{\left(\lambda\left((v_x, G_x), w\right), \mu^{(t)}\left(\lambda\left((v_x, G_x), w\right), G_x\right)\right) = (v_y, G_y)\right\} \mathrm{d}w \\
&= \int_0^1 \mathbb{1}\{\lambda\left((v_x, G_x), w\right) = v_y\} \mathbb{1}\left\{\mu^{(t)}\left(\lambda\left((v_x, G_x), w\right), G_x\right) = G_y\right\} \mathrm{d}w
\end{aligned}$$

Kun $t + 1 \bmod s \neq 0$

$$\begin{aligned}
\mathbb{P}\left[\phi^{(t)}(x, \omega_t) = y\right] &= \int_0^1 \mathbb{1}\{\lambda\left((v_x, G_x), w\right) = v_y\} \mathbb{1}\left\{\mu^{(t)}\left(\lambda\left((v_x, G_x), w\right), G_x\right) = G_y\right\} \mathrm{d}w \\
&= \mathbb{1}\{G_x = G_y\} \int_0^1 \mathbb{1}\{\lambda\left((v_x, G_x), w\right) = v_y\} \mathrm{d}w \\
&= \mathbb{1}\{G_x = G_y\}\, \rho(G_x)_{x,y} = \mathbb{1}\{G_x = G_y\} \frac{\psi_{G_x}(v_x, v_y)}{d_{G_x}(v_x)} \\
&= \mathbb{P}\left[X_{t+1} = y \mid X_t = x\right]
\end{aligned}$$

Kun $t + 1 \bmod s = 0$

$$\begin{aligned}
\mathbb{P}\left[\phi^{(t)}(x, \omega_t) = y\right] &= \int_0^1 \mathbb{1}\{\lambda\left((v_x, G_x), w\right) = v_y\} \mathbb{1}\left\{\mu^{(t)}\left(\lambda\left((v_x, G_x), w\right), G_x\right) = G_y\right\} \mathrm{d}w \\
&= \mathbb{1}\{G_y = (V_x \cup \{v_y\}, E_x \cup \{\{v_x, v_y\}\})\} \int_0^1 \mathbb{1}\{\lambda\left((v_x, G_x), w\right) = v_y\} \mathrm{d}w \\
&= \mathbb{1}\{G_y = (V_x \cup \{v_y\}, E_x \cup \{\{v_x, v_y\}\})\}\, \rho(G_x)_{x,y} \\
&= \mathbb{1}\{G_y = (V_x \cup \{v_y\}, E_x \cup \{\{v_x, v_y\}\})\} \frac{\psi^{(t)}(v_i, v_j)}{d_t(v_i)} \\
&= \mathbb{P}\left[X_{t+1} = y \mid X_t = x\right]
\end{aligned}$$

Selvästi rakentamani stokastinen esitys on ekvivalentti NRRW-prosessin kanssa.

$\square$

4 Parillinen prosessi

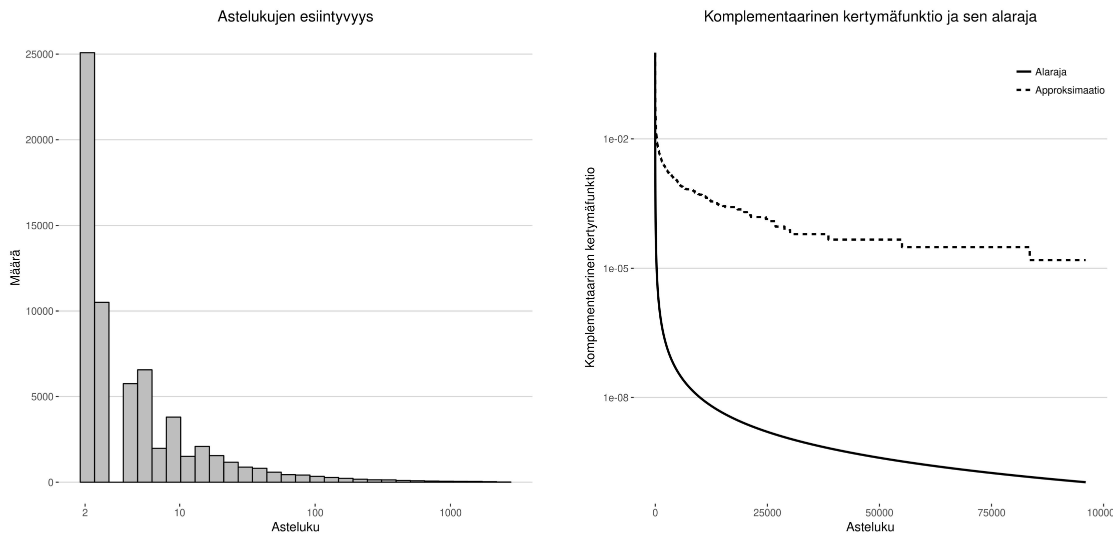
NRRW-prosessin käyttäytymiseen vaikuttaa voimakkaasti askelparametrin s parillisuus. Parillisen parametrin prosessissa satunnaiskävely $(\mathcal{W}_t)_{t\in\mathbb{N}}$ on palautuva eli se vierailee kaikissa sen tilajoukon solmuissa äärettömän monta kertaa lähes varmasti. Lisäksi satunnaisverkon $(\mathcal{G}_t^s)_{t\in\mathbb{N}}$ sellaisten solmujen, joiden asteluku on yli kaksi, astelukujen jakauma on potenssilain alhaalta rajoittama.

Erityisesti vain parillisen askelparametrin NRRW-prosessissa kaksi peräkkäistä uutta solmua voidaan lisätä samaan verkon solmuun. Parittoman askelparametrin prosessissa tämä ei ole mahdollista muualla kuin puun juuressa, sillä satunnaiskävely siirtyy s:ssä askeleessa puun parittomalta tasolta parilliselle ja parilliselta tasolta parittomalle. Parillisen askelparametrin prosessissa satunnaiskävelyn pariteetti muuttuu vain, jos satunnaiskävely kulkee puun juuren silmukan läpi.

Uuden, peräkkäisen solmun lisääminen samaan solmuun parillisessa NRRW-verkossa on paitsi mahdollista niin myös sitä todennäköisempää mitä suurempi kyseisen solmun asteluku on. Koska uusia solmuja lisätään ennen kaikkea saman pariteetin tasoille puussa, jäävät uudet solmut lehdiksi (*eng. leaf*), kunnes prosessin pariteetti vaihtuu. Lehtisolmuista voi siirtyä vain takaisin lähtöpisteeseen, joten todennäköisyys palata johonkin satunnaisverkon solmuun suurenee, kun solmussa vieraillaan ja sen asteluku kasvaa.

Tämä parillisen prosessin ominaisuus liittyy vahvasti suosivaan kiinnittymiseen (*eng. preferential attachment*), sillä uusia solmuja kertyy erityisen todennäköisesti sellaisille verkon solmuille, joilla on jo ennestään paljon naapureita. Suosiva kiinnittyminen on tosielämän verkkorakenteiden kannalta kiinnostava ominaisuus. Keskitynkin tässä kandidaatintyössä parillisen NRRW-prosessin ominaisuuksia käsittelemiseen.

4.1 Astelukujen jakauma



(a) Jotain. (b) Jotain.

Teoreema 1. *Olkoon NRRW-prosessin askelparametri s parillinen. Oletetaan, että T_j on se ajanhetki, jolloin solmuun j lisätään sen ensimmäinen naapuri. Tällöin kaikilla $t \geq T_j$ ja $k \in \{1, \dots, \lfloor \frac{t-T_j}{s} \rfloor + 1\}$ pätee, että*

$$\mathbb{P}\left[d_t(j) \geq k+1 \mid T_j < \infty\right] \geq k^{-\frac{s}{2}}, \quad \textit{jos } j \neq \textit{root}$$

$$\mathbb{P}\left[d_t(j) \geq k+2\right] \geq \left(\frac{k(k+1)}{2}\right)^{-\frac{s}{2}}, \qquad \textit{jos } j = \textit{root}$$

4.2 Palautuvuus

Teoreema 2. *Olkoon i solmu ja $\{i, j\}$ kaari NRRW-prosessin verkossa. Jos i on palautuva, niin satunnaiskävely $(\mathcal{W}_t)_{t \in \mathbb{N}}$ kulkee kaaren $\{i, j\}$ läpi äärettömän monta kertaa lähes varmasti.*

Todistus. TODO: INSERT PROOF HERE! $\square$

5 Yhteenveto

Viitteet

A Esimerkki liitteestä

Liitteet eivät ole opinnäytteen kannalta välttämättömiä ja opinnäytteen tekijän on kirjoittamaan ryhtyessään hyvä ajatella pärjäävänsä ilman liitteitä. Kokemattomat kirjoittajat, jotka ovat huolissaan tekstiosan pituudesta, paisuttavat turhan helposti liitteitä pitääkseen tekstiosan pituuden annetuissa rajoissa. Tällä tavalla ei synny hyvää opinnäytettä.

Liite on itsenäinen kokonaisuus, vaikka se täydentääkin tekstiosaa. Liite ei siten ole pelkkä listaus, kuva tai taulukko, vaan liitteessä selitetään aina sisällön laatu ja tarkoitus.

Liitteeseen voi laittaa esimerkiksi listauksia. Alla on listausesimerkki tämän liitteen luomisesta.

```
\clearpage
\appendix
\addcontentsline{toc}{section}{Liite A}
\section*{Liite A}
...
\thispagestyle{empty}
...
tekstiä
...
\clearpage
```

Kaavojen numerointi muodostaa liitteissä oman kokonaisuutensa:

$$d \wedge A = F, \tag{A1}$$

$$d \wedge F = 0. \tag{A2}$$

B Toinen esimerkki liitteestä

Liitteissä voi myös olla kuvia, jotka eivät sovi leipätekstin joukkoon: Liitteiden taulu-

Kuva B1: Kuvateksti, jossa on liitteen numerointi

koiden numerointi on kuvien ja kaavojen kaltainen: Kaavojen numerointi muodostaa

Taulukko B1: Taulukon kuvateksti.

9.00–9.55	Käytettävyystestauksen tiedotustilaisuus (osanottajat ovat saaneet sähköpostitse valmistautumistehtävät, joten tiedotustilaisuus voidaan pitää lyhyenä).
9.55–10.00	Testausalueelle siirtyminen

liitteissä oman kokonaisuutensa:

$$T_{ik} = -pg_{ik} + wu_i u_k + \tau_{ik}, \tag{B1}$$

$$n_i = nu_i + v_i. \tag{B2}$$