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** 25+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.
Tällainen NRRW-malli mukailee reaalimaailman verkoille tyypillisiä ominaisuuksia, kuten *preferential and local 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** 25+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 and local 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

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 todellisia verkostoja. Niille on tyypillistä preferential ja local attachment sekä se, että verkon solmujen astejakauma ei ole eksponentiaalisesti rajoitettu (*heavy-tailed distribution*). 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.

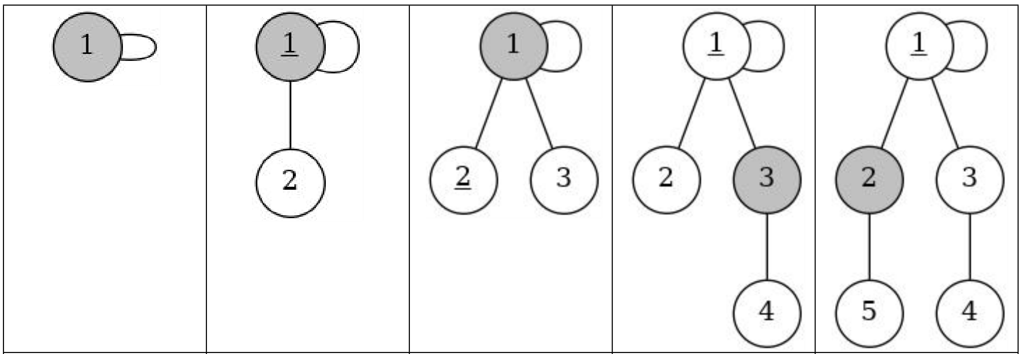
NRRW-prosessi voidaan nähdä joko satunnaisena verkkona tai satunnaiskävelynä. Näkemyksestä riippumatta satunnaiskävelyn palautuvuus ja sen kasvattaman verkon luonne ovat kietoutuneet toisiinsa. Näiden toisistaan riippuvien prosessien käsittelyä voidaan helpottaa osoittamalla, että koko prosessi on yhdenmukainen 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, $(G_t^s, W_t)_{t\geq 0}$. 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.



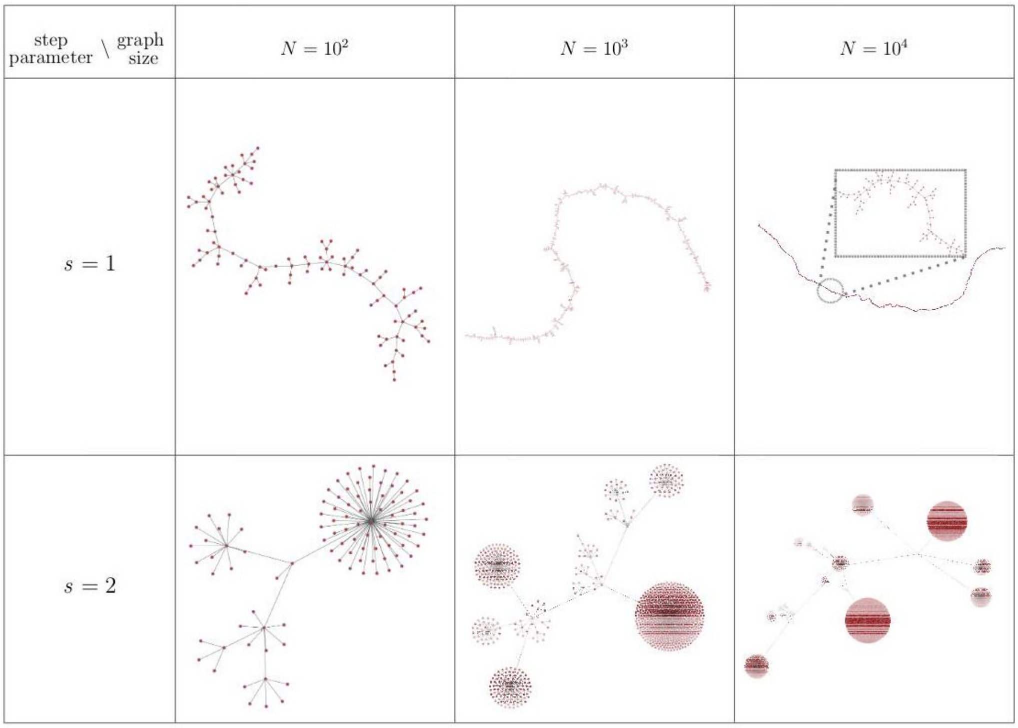
t = 0	t = 2	t = 4	t = 6	t = 8

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 *prefential attachment* -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. [1]

2.1.1 Satunnaisverkko

G_t^s, on satunnaisverkko, johon lisätään uusi solmu s aikayksikön välein. Satunnaisprosessin tilajoukko on kaikkien suuntaamattomien, yhtenäisten ja syklittömien verkkojen eli puiden joukko.

Määritelmä: 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 G_t^s solmut indeksoidaan kokonaisluvuilla $(0, 1, 2, \ldots)$ lisäysjärjestyksen mukaan. Juurisolmu on siis v_0 ja hetkelä $t = sn,\ n \in \mathbb{Z}$ lisätään solmu v_n. Tarkastellaan satunnaisverkkoa $(G_t^s)_{t\geq 0}$ itsenäisenä prosessina. Sen siirtymätodennäköisyydet pois nykytilasta ovat nollasta poikkeavia vain ajanhetkillä $t = sn,\ n \in \mathbb{Z}$. Ne määrää satunnaiskävelyn $(W_t)_{t\geq 0}$ siirtymämatriisi hetkellä $t = sn$, P_t:

$$\begin{aligned} &\mathbb{P}\left[G_{t+s}^s = (V_t \cup \{v_{n+1}\}, E_t \cup \{(v_i, v_{n+1})\}) \mid G_t = (V_t, E_t)\right] \\ &= \mathbb{P}\left[W_{t+s} = v_i \mid W_t = v_n\right] = (P_t^s)_{n,i} \end{aligned} \tag{1}$$

2.1.2 Satunnaiskävely

Vastaavasti satunnaiskävelyprosessin $(W_t)_{t\geq 0}$ tilajoukon ja siirtymätodennäköisyydet määrää sen kasvattama satunnaisverkko. Molemmat näistä 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

$$\psi^{(t)}(v_i, v_j) = \begin{cases} \mathbb{1}\{\{v_i, v_j\} \in E_t\} & \text{jos } v_i \neq v_j \\ 2 \cdot \mathbb{1}\{v_i = v_0\} & \text{muulloin} \end{cases}$$

missä E_t on verkon G_t kaarien joukko. Nyt todennäköisyys siirtyä solmusta v_i solmuun v_j on

$$\mathbb{P}\left[W_{t+1} = v_j \mid W_t = v_i\right] = \frac{\psi^{(t)}(v_i, v_j)}{d_t(v_i)} \tag{2}$$

missä $d_t(v_i)$ on solmun v_i asteluku hetkellä t. Toisin sanoen

$$d_t(v_i) = \sum_{w \in V} \psi^{(t)}(v_i, w) \tag{3}$$

eli kaikkien solmuun v_i yhdistyvien kaarien lukumäärä.

2.2 Kokonaisuus

Satunnaisverkon ja satunnaiskävelyn erillinen tarkastelu paljastaa, että näistä jälkimmäinen on NRRW-prosessin keskeisin satunnaisuuden lähde. Satunnaiskävely muuttaa liikkeillään sen omaa tilajoukkoa eli verkkoa G_t^s. Kun verkko muuttuu, myös satunnaiskävelyn siirtymätodennäköisyydet muuttuvat. Selvästi satunnaiskävelyprosessi on ajassa epähomogeeninen stokastinen prosessi. Satunnaisverkko sen sijaan kasvaa satunnaiskävelyn realisaation ohjaamana. Kokonaisuutena NRRW-prosessi on siis myös epähomogeeninen ajassa.

2.3 Parillinen prosessi

TODO: Erityishuomioita?

3 Stokastinen esitys

3.1 Teoria

Olkoon $(\omega_t)_{t \geq 0}$ jono riippumattomia välin $[0, 1]$ tasajakautuneita satunnaislukuja. Toisin sanoen $w_t \sim \text{Tas}(0, 1)$. 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

$$\begin{aligned} \mathbb{P}\left[\phi(x,\omega_t) = y\right] &= \int_0^1 \mathbb{1}\{\phi(x,w) = y\} \, \mathrm{d}w & (4) \\ &= \mathbb{P}\left[X_{t+1} = y \mid X_t = x\right] = P_{x,y} \quad \forall \, x, y \in S & (5) \end{aligned}$$

Yhtälön 5 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:

$$X_{t+1} = \phi(X_t, \omega_t) \tag{6}$$

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 5 vastaava ehto $\forall \, t \geq 0$:

$$\mathbb{P}\left[\phi^{(t)}(x,\omega_t) = y\right] = \mathbb{P}\left[X_{t+1} = y \mid X_t = x\right] = P^{(t)}_{x,y} \quad \forall \, x, y \in S \tag{7}$$

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

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

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 7. 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

$$\phi^{(t)}\left((v,G),w\right) = \left(\lambda^{(t)}\left((v,G),w\right), \mu^{(t)}\left(\lambda^{(t)}\left((v,G),w\right), G\right)\right) \tag{9}$$

missä $\lambda^{(t)}$ on satunnaiskävelyn $(W_t)_{t\geq 0}$ stokastisen esityksen rekursiofunktio ja μ päivittää verkkoa G satunnaiskävelyn realisaation mukaisesti. Satunnaiskävelyn rekursiofunktio

$$\lambda^{(t)}\left((v,G),w\right) = \begin{cases} v_1 & w \in [0, \rho(G)_{i,1}) \\ v_2 & w \in [\rho(G)_{i,1}, \rho(G)_{i,1} + \rho(G)_{i,2}) \\ & \quad \vdots \\ v_{\lfloor \frac{t}{s} \rfloor} & w \in [\rho(G)_{i,\lfloor \frac{t}{s} \rfloor - 1}, \rho(G)_{i,\lfloor \frac{t}{s} \rfloor - 1} + \rho(G)_{i,\lfloor \frac{t}{s} \rfloor}) \end{cases} \tag{10}$$

missä ρ on funktio verkosta sen siirtymämatriisiin. Hyödyntäen määritelmän 2.1.2 funktiota solmuja yhdistävien kaarien määrälle, saadaan kuvaukselle ρ alkioittainen määritelmä

$$\rho^{(t)}(G)_{(i,j)} = \rho^{(t)}\left((V,E)\right)_{(i,j)} = \frac{\psi^{(t)}(i,j)}{d_t(i)} = \frac{\psi^{(t)}(i,j)}{\sum_{w \in V} \psi^{(t)}(v_i, w)} \tag{11}$$

missä kaksi viimeistä yhtäsuuruutta seuraa kaavoista 2 ja 3.

Stokastisen esityksen verkkoa päivittävä funktio määritellään

$$\mu^{(t)}\left(v, (V,E)\right) = \begin{cases} (V \cup v_t, E \cup \{v, v_t\}) & \text{kun } t \bmod s = 0 \\ (V,E) & \text{muulloin} \end{cases} \tag{12}$$

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

3.3 Todistus

Esitän, että stokastisen esityksen $(\phi^{(t)}, \omega_t)_{t \geq 0}$ simuloima stokastinen prosessi on todellakin NRRW-prosessi askelparametrillä s.

Rakenne

Opinnäytteen rakenteen tulee olla hyvän tieteellisen kirjoittamisen käytännön mukainen ja sisältää vähintään seuraavat osat:

1. Nimiölehti
2. Tiivistelmä
3. Sisällysluettelo
4. Symboli- ja lyhenneluettelo
5. Johdanto
6. Aikaisempi tutkimus. Työn luonteen niin vaatiessa otsikko voi olla myös »Teoreettinen tausta» tai näiden otsikoiden yhdistelmä.
7. Tutkimusaineisto ja -menetelmät
8. Tulokset
9. Tarkastelu. Työn luonteen niin vaatiessa otsikko voi olla myös »Johtopäätökset» tai »Yhteenveto» tai edellä mainittujen otsikoiden yhdistelmä.
10. Lähteet
11. Liitteet.

Tiivistelmän ja symboli- sekä lyhenneluetteloiden väliin voi sijoittaa halutessaan esipuheen.

Työn osat 5–9 muodostavat *tekstiosan.* Työn yksittäisiä osia voidaan jakaa alaotsikoilla alaosiin, joita ei ole yllä esitetty. Alaotsikoiden käyttäminen selventää parhaimmillaan tekstiä, ja pahimmillaan sirpaloittaa sitä. Sirpaloitumista voi estää huolehtimalla siitä, että samalla sivulla ei esiinny useampaa alaotsikkoa. Tekstin jäsentelyssä on yleensä ongelmia, jos osassa on vain yksi alaosa, tai kirjoittaja joutuu käyttämään useampaa kuin kahta tasoa (osa ja alaosat): alaosien alaosat ovat harvoin tarpeen.

Sivut ja kirjaintyypit

Opinnäytteen tulee olla kirjoitettu koneella tai tekstinkäsittelyohjelmalla yksipuolisesti A4-kokoiselle paperille. Kandidaatintyön tekstiosan sopiva pituus on noin 15–20 sivua ja diplomityön noin 60 sivua. Työtä ei ole syytä tarpeettomasti pidentää.

Opinnäytteen tekstiosan kirjaintyypin tulee olla antiikva eli serif-tyyppinen ja lisäksi kursivoimaton, lihavoimaton sekä kooltaan 12 pistettä (kuten tässä esityksessä). Groteskeja eli Sans serif-tyyppisiä kirjaintyyppejä (kuten Helvetica tai Arial) ei saa käyttää varsinaisessa tekstissä, mutta otsikoissa näitä voidaan käyttää. Otsikoissa voidaan käyttää kooltaan edellä mainittua suurempaa kirjaintyyppiä sekä tyylikeinoja, kuten lihavointia tai kursivointia. Tekstissä samantasoisten otsikoiden on kuitenkin oltava tyyliltään ja kirjainlajeiltaan yhteneväisiä.

Taulukko 2: Taulukoissa ja kuvissa kirjaintyypin voi valita tarkoituksenmukaisesti, mutta kuva- ja taulukkoteksteissä tulee käyttää samaa kirjaintyyppiä kuin varsinaisessa tekstissä. Huomaa taulukon numeroinnin sijoittuminen taulukon yläpuolelle.

A	1	$e^{j\omega t}$
B	2	$\Re(c)$
C	3	$a \in \mathbb{A}$

Opinnäytteen vasen marginaali (sidonnan puoli) on 35 mm ja oikea 25 mm. Ylämarginaali on 25 mm. Leipätekstin korkeus on enimmillään 230mm. Tämän opinnäytepohjan marginaalien pitäisi olla paperille tulostettuna oikein, mutta tulostimesta ja paperista riippuen voi esiintyä yhden tai kahden millimetrin suuruisia eroja.

Asemointi

Tekstiosan tekstissä käytetään kappaleiden erottamiseen sisennystä, mutta ensimmäistä otsikon, väliotsikon tai muun katkon jälkeistä kappaletta ei sisennetä. Jos kuva tai muu katko tulee kappaleiden väliin, suositellaan katkon jälkeisen kappaleen sisentämistä.

Mikäli oikea reuna halutaan tasata, tulee käyttää tavutusta ja lisäksi tarkistaa, ettei tekstiin jää lukemista häiritseviä pitkiä sanavälejä. Jos käytät opinnäytteen tekemisessä LaTeX-järjestelmää, tämä asia hoituu automaattisest.

Opinnäytteen riviväli on 1, mikä on myös tämän opinnäytepohjan käytäntö. Kappaleiden tulee yleensä olla ainakin kolmen rivin pituisia, mutta myös liian pitkiä kappaleita tulee välttää. Tässä opinnäytepohjassa ei tekstin luonteen vuoksi voida täysin toteuttaa kappaleen pituutta koskevia vaatimuksia.

Yksittäisiä, kappaleen päättäviä tai aloittavia rivejä sivun alussa tai lopussa on vältettävä koko työssä, myös luetteloissa ja liitteissä.

Numerointi

Opinnäytteen jokainen osa alkaa uudelta sivulta. Alaosa aloittaa uuden sivun vain edellisen sivun täytyttyä.

Työn osat numeroidaan siten, että johdanto on ensimmäinen numeroitava osa. Osien numeroinnissa käytetään arabialaisia numeroita.

Nimiölehti, tiivistelmä, esipuhe, sisällysluettelo ja symboli- ja lyhenneluettelo numeroidaan esipuheesta tai tämän puuttuessa ensimmäiseltä luettelosivulta alkaen roomalaisin numeroin.

Sivunumerointi alkaa toiselta varsinaiselta tekstisivulta, ja sivunumeroinnissa käytetään arabialaisia numeroita.

Lähdeluettelo alkaa uudelta sivulta. Lähdeluettelon sivunumerointi jatkuu viimeisestä tekstisivusta.

Jokainen liite alkaa uudelta sivulta. Liitteiden sivunumerointi jatkuu viimeisestä lähdeluettelon sivusta.

Sivunumero sijoitetaan sivun yläreunaan.

Matemaattiset kaavat numeroidaan arabialaisin numeroin. Kaavanumerointi ei saa katketa osien välissä (eikä niin tapahdukaan, jos käytät tätä opinnäytepohjaa). Kaikkia kaavoja ei tarvitse numeroida, vaan kirjoittaja voi käyttää harkintaa numeroinnin tarpeellisuudessa. Liitteissä olevat kaavat numeroidaan siten, että liitteen ajatellaan muodostavan numeroinnin kannalta itsenäisen ja yhtenäisen kokonaisuuden. Kaavan numero sijoitetaan oikealle puolelle alla olevan esimerkin mukaisesti

$$D(xy) = (Dx)y + x(Dy), \qquad x, y \in \mathbb{A}. \tag{13}$$

Kaikki kuvat ja taulukot numeroidaan erillisen juoksevan numeroinnin mukaisesti kuten taulukosta 2 ja kuvasta 2 käy ilmi. Liitteissä olevat kuvat ja taulukot numeroidaan siten, että liitteen ajatellaan muodostavan numeroinnin kannalta itsenäisen ja yhtenäisen kokonaisuuden. Liitteissä A ja B on esimerkkejä kaavojen (kaavat A1–A2 tai kaavat B1–B2), kuvien (kuva B1) ja taulukoiden (taulukko B1) numeroimisesta. Liitteet numeroidaan suuraakkosin (esimerkiksi Liite A, Liite B tai pelkästään A, B).

Kuva 2: Tämä on esimerkki numeroidusta kuvatekstistä.

Lähdeviittausten käyttö

Lähdeviittaukset tulee tehdä huolellisesti ja johdonmukaisesti numeroviitejärjestelmän mukaisesti. Numeroviitteet järjestetään lähdeluetteloon viittausjärjestykseen, mutta jos lähdeluettelo on hyvin laaja (useita sivuja), järjestetään viitteet pääsanan mukaiseen aakkosjärjestykseen. Alaviitejärjestelmää [1] ei käytetä.

Viitteen sijoittelussa noudatetaan seuraavia sääntöjä: Jos viite kohdistuu vain yhteen virkkeeseen tai virkkeen osaan, viite [1] sijoitetaan virkkeen sisään ennen virkettä päättävää pistettä. Jos taas viite koskee tekstin useampaa virkettä tai kokonaista kappaletta, sijoitetaan viite kappaleen loppuun pisteen jälkeen. [1]

Lähdeluettelo

Lähdeluettelossa esiintyy tavallisesti seuraavassa esitettäviä lähteitä, joista on numeroviitejärjestelmässä ilmoitettava asianomaisessa kohdassa vaaditut tiedot.

Kirjasta ilmoitetaan seuraavat tiedot:

- tekijät
- julkaisun nimi
- painos, jos useita
- kustannuspaikka
- julkaisija tai kustantaja
- julkaisuaika
- mahdollinen sarjamerkintö.

Viitteet [1]–[3] ovat esimerkkejä kirjan esittämisestä lähdeluettelossa. Viite [3, s. 83–124] on esimerkki lähdeluettelossa esiintyvän kirjan tiettyjen sivujen esittämisestä tekstissä.

Artikkelista kausijulkaisussa ilmoitetaan seuraavat tiedot:

- tekijät
- artikkelin nimi
- kausijulkaisun nimi
- julkaisuvuosi
- kausijulkaisun volyymi tai ilmestymisvuosi
- kausijulkaisun numero

[1] Myöskään alaviitteenä olevia kommentteja ei suositella käytettäviksi.

– sivut, joilla artikkeli on.

Viitteet [4]–[5] ovat esimerkkejä artikkelin esittämisestä lähdeluettelossa.
Kokoomateoksen luvusta tai osasta ilmoitetaan seuraavat tiedot:

– luvun tai osan tekijät

– luvun tai osan nimi

– maininta »Teoksessa»

– koko teoksen toimittajat sekä maininta »(toim.)»

– koko teoksen tai konferenssin nimi

– konferenssiesitelmän kyseessä ollessa sen pitopaikka ja -aika

– painos, jos useita

– kustannuspaikka

– julkaisija tai kustantaja, jos aihetta tämän ilmoittamiseen on

– julkaisuaika

– sivut, joilla luku tai osa on

– mahdollinen sarjamerkintä.

Viitteet [6]–[7] ovat esimerkkejä kokoomateoksen luvun tai osan esittämisestä lähdeluettelossa.
Opinnäytetyöstä ilmoitetaan seuraavat tiedot:

– tekijä

– työn nimi

– opinnäytetyön tyyppi

– oppilaitoksen nimi

– osaston, laitoksen tai ohjelman nimi

– oppilaitoksen sijaintipaikka

– vuosiluku.

Viitteet [8]–[10] ovat esimerkkejä opinnäytteen esittämisestä lähdeluettelossa.
Standardista ilmoitetaan seuraavat tiedot:

– standardin tunnus ja numero

– standardin nimi

– painos, mikäli ei ole ensimmäinen

– julkaisupaikka

– julkaisija

– julkaisuvuosi

– sivumäärä.

Viite [11] on esimerkki standardin esittämisestä opinnäytteen lähdeluettelossa.

Haastattelusta ilmoitetaan seuraavat tiedot:

– haastatellun henkilön nimi

– haastatellun henkilön arvo tai asema

– haastatellun henkilön edustama organisaatio

– organisaation osoite

– maininta siitä, että kyseessä on haastattelu ja haastattelun päivämäärä.

Viite [12] on esimerkki haastattelun esittämisestä lähdeluettelossa.

Osa sähköisessä muodossa olevista artikkeleista on saatavissa myös painettuina. *Vain verkosta saatavissa olevasta artikkelista* esitetään seuraavat tiedot:

– tekijät

– artikkelin nimi

– kausijulkaisun nimi

– viestintyyppi

– laitos tai volyymi

– kausijulkaisun yksittäistä osaa koskeva merkintä tai numero

– julkaisuvuosi tai maininta »Päivitetty» ja päivitysaika

– maininta »Viitattu» ja viittaamisen ajankohta

– maininta »Saatavissa» ja URL tai maininta »DOI» ja DOI-numero (DOI=Digital Object Identifier).

Viitteet [13]–[15] ovat esimerkkejä sähköisessä muodossa olevan artikkelin esittämisestä opinnäytteen lähdeluettelossa. Viitteet [13] ja [14] ovat saatavissa sekä painettuna että verkosta, joten viitteiden esitystapa mukailee painetun artikkelin viitteen esitystapaa, mutta sen lisäksi kerrotaan julkaisun olevan verkkolehti ja lehden olevan saatavissa myös painettuna. Viite [15] on saatavissa vain verkosta ja siitä esitetään yllä vaaditut tiedot.

Valitettavasti sähköisessä muodosssa olevasta artikkelista ei ole aina saatavissa laitos-, volyymi- tai numerotietoja.

Sähköisessä muodossa olevasta opinnäytetyöstä ilmoitetaan seuraavat tiedot:

– tekijä

– työn nimi

– viestintyyppi

– opinnäytetyön tyyppi

– oppilaitoksen nimi

– osaston, laitoksen tai ohjelman nimi

– oppilaitoksen sijaintipaikka

– vuosiluku

– viittamisen ajankohta

– maininta "Saatavissa"ja URL tai maininta "DOI"ja DOI-numero.

Viite [16] on esimerkki sähköisessä muodossa olevan opinnäytteen esittämisestä lähdeluettelossa.

Viite [17] on esimerkki itsenäisen kirjoituksen sisältävästä verkkosivusta. Tällainen lähde on rinnastettavissa erillisteokseen. *Verkkosivusta* esitetään tiedot:

– tekijät

– otsikko

– maininta "Päivitetty"ja päivitysaika

– maininta "Viitattu"ja viittaamisen ajankohta

– Maininta "Saatavissa"ja URL.

Joskus verkkosivun kirjoitus on jaettu useammalle sivulle, jolloin lähdeluetteloon kirjataan vain sellainen verkko-osoite, joka koskee koko kirjoitusta tai sen etusivua, ellei sitten todella tarkoiteta kirjoituksen yksittäistä sivua.

Muuta huomioitavaa lähdeluettelossa

Lähdeluettelossa työn ja julkaisun nimi kirjoitetaan alkuperäisessä muodossaan. Julkaisijan kotipaikka kirjoitetaan alkukielisessä muodossaan.

Viittamista koskevassa suomalaisessa standardissa SFS 5342 [11] vaaditaan julkaisuista ilmoitettavaksi myös ISBN- tai ISSN-numerot, mutta näissä opinnäyteohjeissa ei ISBN- ja ISSN-numeroita vaadita.

4 Tutkimusaineisto ja -menetelmät

Tässä osassa kuvataan käytetty tutkimusaineisto ja tutkimuksen metodologiset valinnat, sekä kerrotaan tutkimuksen toteutustapa ja käytetyt menetelmät.

5 Tulokset

Tässä osassa esitetään tulokset ja vastataan tutkielman alussa esitettyihin tutkimuskysymyksiin. Tieteellisen kirjoitelman arvo mitataan tässä osassa esitettyjen tulosten perusteella.

Tutkimustuloksien merkitystä on aina syytä arvioida ja tarkastella kriittisesti. Joskus tarkastelu voi olla tässä osassa, mutta se voidaan myös jättää viimeiseen osaan, jolloin viimeisen osan nimeksi tulee »Tarkastelu». Tutkimustulosten merkitystä voi arvioida myös »Johtopäätökset»-otsikon alla viimeisessä osassa.

Tässä osassa on syytä myös arvioida tutkimustulosten luotettavuutta. Jos tutkimustulosten merkitystä arvioidaan »Tarkastelu»-osassa, voi luotettavuuden arviointi olla myös siellä.

6 Yhteenveto

Opinnäytteen tekijä vastaa siitä, että opinnäyte on tässä dokumentissa ja opinnäytteen tekemistä käsittelevillä luennoilla sekä harjoituksissa annettujen ohjeiden mukainen muotoseikoiltaan, rakenteeltaan ja ulkoasultaan.

Viitteet

[1] Kauranen, I., Mustakallio, M. ja Palmgren, V. *Tutkimusraportin kirjoittamisen opas opinnäytetyön tekijöille.* Espoo, Teknillinen korkeakoulu, 2006.

[2] Itkonen, M. *Typografian käsikirja.* 3. painos. Helsinki, RPS-yhtiöt, 2007.

[3] Koblitz, N. *A Course in Number Theory and Cryptography. Graduate Texts in Mathematics 114.* 2. painos. New York, Springer, 1994.

[4] Bardeen, J., Cooper, L. N. ja Schrieffer, J. R. Theory of Superconductivity. *Physical Review,* 1957, vol. 108, nro 5, s. 1175–1204.

[5] Deschamps, G. A. Electromagnetics and Differential Forms. *Proceedings of the IEEE,* 1981, vol. 69, nro 6, s. 676–696.

[6] Sihvola, A. et al. Interpretation of measurements of helix and bihelix superchiral structures. Teoksessa: Jacob, A. F. ja Reinert, J. (toim.) *Bianisotropics '98 7th International Conference on Complex Media.* Braunschweig, 3.–6.6.1998. Braunscweig, Technische Universität Braunschweig, 1998, s. 317–320.

[7] Lindblom-Ylänne, S. ja Wager, M. Tieteellisten opinnäytetöiden ohjaaminen. Teoksessa: Lindblom-Ylänne, S. ja Nevgi, A. (toim.) *Yliopisto- ja korkeakouluopettajan käsikirja.* Helsinki, WSOY, 2004, s. 314–325.

[8] Miinusmaa, H. Neliskulmaisen reiän poraamisesta kolmikulmaisella poralla. Diplomityö, Teknillinen korkeakoulu, konetekniikan osasto, Espoo, 1977.

[9] Loh, N. C. High-Resolution Micromachined Interferometric Accelerometer. Master's Thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1992.

[10] Lönnqvist, A. Applications of hologram-based compact range: antenna radiation pattern, radar cross section, and absorber reflectivity measurements. Väitöskirja, Teknillinen korkeakoulu, sähkö- ja tietoliikennetekniikan osasto, 2006.

[11] SFS 5342. Kirjallisuusviitteiden laatiminen. 2. painos. Helsinki, Suomen standardisoimisliitto, 2004. 20 s.

[12] Palmgren, V. Suunnittelija. Teknillinen korkeakoulu, kirjasto. Otaniementie 9, 02150 Espoo. Haastattelu 15.1.2007.

[13] Ribeiro, C. B., Ollila, E. ja Koivunen, V. Stochastic Maximum-Likelihood Method for MIMO Propagation Parameter Estimation. *IEEE Transactions on Signal Processing,* verkkolehti, vol. 55, nro 1, s. 46–55. Viitattu 19.1.2007. Lehti ilmestyy myös painettuna. DOI: 10.1109/TSP.2006.882057.

[14] Stieber, T. GnuPG Hacks. *Linux Journal,* verkkolehti, 2006, maaliskuu, nro 143. Viitattu 19.1.2007. Lehti ilmestyy myös painettuna. Saatavissa: `http://www.linuxjournal.com/article/8732`.

[15] Pohjois-Koivisto, T. Voiko kone tulevaisuudessa arvata tahtosi? *Apropos,* verkkolehti, helmikuu, nro 1, 2005. Viitattu 19.1.2007. Saatavissa: `http://www.apropos.fi/1-2005/prima.php`.

[16] Adida, B. Advances in Cryptographic Voting Systems. Verkkodokumentti. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts, 2006. Viitattu 19.1.2007. Saatavissa: `http://crypto.csail.mit.edu/~cis/theses/adida-phd.pdf`.

[17] Kilpeläinen, P. WWW-lähteisiin viittaaminen tutkielmatekstissä. Verkkodokumentti. Päivitetty 26.11.2001. Viitattu 19.1.2007. Saatavissa: `http://www.cs.uku.fi/~kilpelai/wwwlahteet.html`.

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:

$$T_{ik} = -pg_{ik} + wu_i u_k + \tau_{ik}, \quad \text{(B1)}$$

$$n_i = nu_i + v_i. \quad \text{(B2)}$$