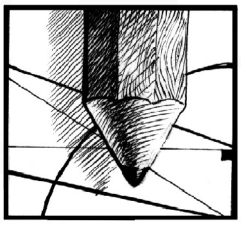


Verkkojen teoriaa Königsbergistä internetiin

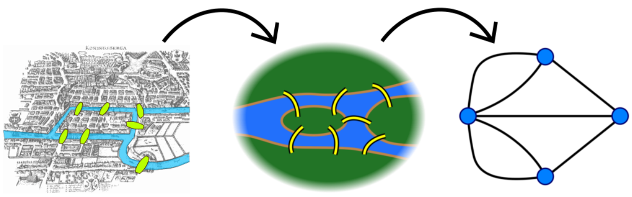
Ville Romanov
Matematiikan ja systeemianalyysin laitos, Aalto-yliopisto

Königsbergin sillat

1700-luvun Königsbergissä (nykyään Kaliningrad) saattoi satunnainen matkaaja päätyä pähkäilemään kinkkisen maantieteellisen pulman kanssa. Kaupungin läpi nimittäin virtaa Pregel-joki, jossa on kaupungin kohdalla kaksi suurikokoista saarta keskellä jokea. Tuohon aikaan joen eri puolia ja kahta saarta yhdisti seitsemän siltaa, joista yksi yhdisti saaret toisiinsa ja loput kuusi puolestaan kulkivat saarten ja mantereen välillä.

Hankala kysymys kuitenkin oli, olisiko mahdollista löytää sellainen kulkureitti, joka kulkee jokaisen sillan kautta tasan yhden kerran. Kartan kanssa pienen pohtimisen ja kenties yrityksen ja erehdyksen jälkeen voi huomata, ettei tuollaisen reitin löytäminen ole erityisen helppoa. Pitkällisemmän testaamisen jälkeen aloitteleva matemaatikko saattaa jopa hyväksyä, että ongelmaan ei näytä löytyvän ratkaisua.

Tämä empiirinen tosiseikka lienee ollut myös Königsbergin paikallisten asukkaiden tiedossa jo paljon aiemmin, mutta vasta paikallinen matemaatikko Leonhard Euler oli ensimmäinen, joka onnistui käsittelemään ongelmaa matemaattisesti mielekkäällä ja pätevällä tavalla. Vuonna 1736 julkaistussa artikkelissaan hän muotoili ongelman matemaattisesti ja ensimmäisenä myös todisti, ettei annettuun siltaongelmaan ole ratkaisua. Tätä Eulerin artikkelia voidaan pitää verkko- eli graafiteorian lähtölaukauksena, josta tähän päivään mennessä verkkoteoria on kehittynyt merkittäväksi diskreetin matematiikan osa-alueeksi. Suomessa puhutaan hieman asiayhteydestä riippuen sekä verkoista että graafeista (engl. graph), mutta nämä ovat käytännössä synonyymeja.



[1]

Verkkoteorian perusteita

Tarkastellaan vielä Königsbergin siltaongelmaa. Itse ongelman kannalta on selvästi täysin merkityksetöntä tarkastella siltoja ja saaria niiden aidossa maantieteellisessä yhteydessä esimerkiksi mittakaava tai maa-alueiden sisäiset reitit huomioiden. Sen sijaan siltojen muodostamat yhteydet maa-alueiden välillä ovat ainoita ongelman kannalta oleellisia seikkoja. Tällöin tilanne voidaan yksinkertaistaa piirtämällä varsin yksinkertainen verkko, jossa kutakin maa-aluetta merkitään pisteellä ja alueita yhdistäviä siltoja viivoilla; kaikki ongelman kannalta epäoleellinen on jätetty pois. Näin al-

[1]Kuva on mukailtu lähteistä `https://en.wikipedia.org/wiki/File:7_bridges.svg`, `https://en.wikipedia.org/wiki/File:Königsberg_graph.svg` ja `https://en.wikipedia.org/wiki/File:Konigsberg_bridges.png`. Lisenssi: CC BY-SA 3.0

kuperäisestä ongelmasta on saatu luotua yksinkertainen malli, johon voidaan soveltaa verkkoteorian tuloksia ja etsiä ratkaisua näiden valossa. Tarkastelun kohteena ovat tällöin erityisesti yhteyksien määrä ja niiden jakautuminen.

Eräs luonnollinen sovelluskohde verkkoteorialle on siis sellaiset tosielämän systeemit, joissa tarkastelun kohteena ovat erityisesti yhteydet, niiden jakautuminen ja reitit systeemin eri osien tai olioiden välillä. Varsin ilmeisiä sovelluskohteita ovat esimerkiksi liikennejärjestelmät kuten katu- ja rataverkot, erilaiset kartat, sähkö- tai tietoverkot, WWW, sosiaaliset verkostot ja sukupuut.

Yksinkertaisimmillaan verkko tai graafi on **kokoelma**

- **pisteitä** (engl. node, vertex) ja
- **viivoja** (engl. edge) eli yhteyksiä näiden välillä.

Verkoissa yhteydet ovat aina kahden eri pisteen välillä, toisin sanottuna verkoissa ei ole silmukoita. Yhteydet ovat yleensä kaksisuuntaisia, mutta niin sanotussa suunnatussa verkossa voi olla myös yhdensuuntaisia yhteyksiä, joita voi edetä vain määrättyyn suuntaan. Pisteistä puhutaan yleisesti myös *solmuina.*

Tarkalleen ottaen verkossa on kahden pisteen välillä enintään yksi viiva. Mikäli viivoja on kaksi tai useampia, puhutaan yleensä **multigraafista**. Tämän perusteella Königsbergin siltoja kuvaava verkko on itse asiassa multigraafi, sillä joitakin pisteitä yhdistää kaksi viivaa. Siltaongelman tarkastelussa jaottelu tavallisen verkon ja multigraafin välillä ei kuitenkaan ole oleellinen seikka.

Lisäksi jos verkon voi piirtää tasoon siten, että viivat eivät leikkaa toisiaan, kutsutaan verkkoa **tasoverkoksi**.

Käytetään seuraavia merkintöjä verkoille:
Olkoon $G(P_G, V_G)$ jokin verkko. Tällä tarkoitetaan, että

- P_G on verkon G pisteiden joukko ja
- V_G on verkon G viivojen joukko,

jotka sitten muodostavat itse verkon G. Mikäli valitaan vain osa verkon G viivoista ja pisteistä ja muodostetaan niistä uusi verkko, on tämä uusi verkko G^* verkon G **aliverkko**.

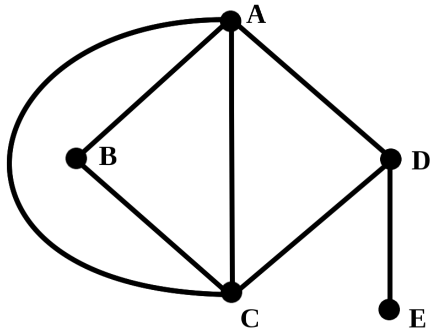
- Jos verkossa G on piste x, merkitään $x \in P_G$.
- Jos verkossa on viiva pisteiden x ja y välillä, merkitään yleensä $\overline{xy} \in V_G$.

Merkitään vielä:

- verkon viivojen määrää $v_G = |V_G|$ ja
- verkon pisteiden määrää $p_G = |P_G|$.

Määritellään lisäksi eristetty piste: $x \in P_G$ on **eristetty piste,** jos siihen ei liity yhtään viivaa.

Verkon jokaiseen pisteeseen liittyy myös **pisteen aste**, monien hyödyllisten tulosten kannalta oleellinen ominaisuus. Pisteen $x \in P_G$ asteeksi tai asteluvuksi $d(x)$ kutsutaan siihen liittyvien viivojen lukumäärää.



Esimerkiksi yllä olevassa verkossa pisteen A aste $d(A) = 4$, tai pisteen E aste $d(E) = 1$. Eristetyn pisteen aste on aina 0. Pisteiden asteisiin liittyy myös ensimmäinen merkittävä tulos:

Lause:
Verkon viivojen lukumäärälle pätee:
$v_G = \frac{1}{2}\sum_{x \in P_G} d(x) =$ *verkon pisteiden asteiden summa jaettuna kahdella.*

Tästä tuloksesta puolestaan voidaan pian johtaa niin kutsuttu ”**kättelylemma**”:

Juhlissa, joissa on äärellinen määrä osallistujia, on aina ***parillinen*** *määrä osallistujia, jotka ovat kätelleet* ***paritonta*** *määrää muita osallistujia.*

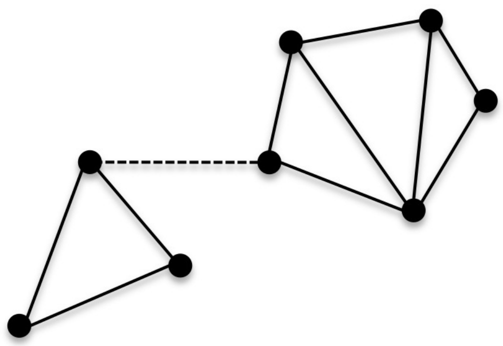
Tai toisin sanottuna: äärellisessä verkossa on aina parillinen määrä paritonasteisia pisteitä! Mistä tämän voisi päätellä? Osoita kättelylemman avulla, että talossa jossa on **tasan yksi** ulko-ovi, on ainakin yksi huone jossa on **pariton määrä** ovia.

Yhtenäisyys ja kulut verkossa

Kulku verkossa on reitti, jota pitkin verkossa voidaan viivoja pitkin edetä pisteestä toiseen. Muodollisesti kulku verkossa G on jono $\overline{x} = \{x_0, x_1, \ldots, x_n\}$ verkon pisteitä siten, että on olemassa viiva $\overline{x_{i-1}x_i} \in V_G$ kaikilla $i = 1, \ldots, n$. Yksinkertaiseksi kuluksi sanotaan sellaista kulkua, jossa kunkin viivan tai solmun kautta kuljetaan enintään kerran.

Verkko G on **yhtenäinen**, jos verkon minkä tahansa kahden pisteen välillä on kulku: mistä tahansa verkon pisteestä voidaan edetä viivoja pitkin mihin tahansa pisteeseen. Suunnistettujen verkkojen tapauksessa voidaan puhua erikseen yhtenäisyydestä ja vahvasta yhtenäisyydestä, mutta tavallisten verkkojen tapauksessa näiden kahden välillä ei ole eroa. Yhtenäisen verkon vastakohta on *epäyhtenäinen* verkko.

Täydelliseksi verkkoa kutsutaan, kun jokaisesta pisteestä on yhteys verkon **kaikkiin muihin** pisteisiin. Jos täydellisessä verkossa on k pistettä, voidaan viivojen lukumäärälle johtaa kaava $v_G = \frac{k(k-1)}{2}$.



[2]

Yhtenäinen verkko, josta saadaan epäyhtenäinen poistamalla katkoviivalla piirretty yhteys.

Eulerin kulut - takaisin Königsbergiin

Eulerin kuluksi verkossa G kutsutaan sellaista kulkua, jossa verkon G jokainen viiva esiintyy tasan yhden kerran. Itse asiassa Königsbergin siltaongelmassa on kysymys nimenomaan tällaisen kulun olemassaolosta, sillä ongelmassa halutaan löytää yksinkertainen reitti jokaisen sillan kautta. Jos tarkastellaan siltaongelmaa kuvaavaa verkkoa, voidaan havaita verkon kaikkien pisteiden olevan paritonasteisia: kaikkien pisteiden aste on pariton. Kysytään nyt kaksi oleellista kysymystä:

- Voiko Eulerin kulku päättyä paritonasteiseen pisteeseen, josta kulku myös alkaa?
- Voiko Eulerin kulku päättyä paritonasteiseen pisteeseen, josta kulku ei ole alkanut?

On helppoa päätellä, että jos Eulerin kulku alkaa paritonasteisesta pisteestä, ei se voi päättyä alkupisteeseensä, sillä kaikkien viivojen kautta kulkeminen tarkoittaisi, että viimeisen viivan kautta kulkeminen johtaisi pois pisteestä (ensimmäinen askel on pois aloituspisteestä, toinen takaisin, kolmas pois jne.). Ja kuitenkin huomataan, että Eulerin kulun on silti päätyttävä paritonasteiseen pisteeseen, mikäli verkossa on sellainen. Muuten tällaisesta paritonasteisesta pisteestä jäisi yksi viiva käymättä läpi. Koska Königsbergissä on neljä paritonasteista pistettä, ei ole mahdollista että kulku voisi samaan aikaan päättyä kolmeen näistä, joten Königsbergin verkossa ei voi olla Eulerin kulkua. Itse asiassa tämä tulos voidaan muotoilla yleisemminkin.

Lause:
Yhtenäisessä verkossa on Eulerin kulku, jos ja vain jos verkossa on tasan kaksi paritonasteista pistettä.

Suuret verkot maailmassamme - skaalautumattomat verkot

Arkielämässä ympärillämme on monia rakenteita, jotka ovat aivan ilmeisiä verkkoja tai joita ainakin voi mallintaa verkkoina. Esimerkiksi Internet ja muut tietoverkot ovat kirjaimellisesti verkkoja kuten myös sähkö- ja tieverkot. Konkreettisten verkkojen lisäksi ihmisten välisten suhteiden voidaan ajatella muodostavan verkon kaltaisen rakenteen, jossa ystävyyssuhteet muodostavat yhteyksiä yksilöiden eli verkon solmujen välille.

Modernin verkkoteorian kehittäjät Paul Erdős ja Alfréd Rényi olivat edelläkävijöitä suurten verkkojen tutkimuksessa. He kuitenkin tutkivat enimmäkseen satunnaisia verkkoja, joissa pisteiden välisten yhteyksien oletetaan olevan täysin sattumanvaraisesti muodostuneita. Kuitenkin monien tosielämän verkkojen tarkastelu osoittaa, ettei näiden verkkojen pisteiden astejakauma ole satunnainen. Satunnaisuuden sijaan monien näistä verkoista on havaittu noudattavan niin kutsuttua potenssilakia, jonka mukaan pisteiden astelukujen jakauma noudattelee potenssijakaumaa. Tällaista verkkoa kutsutaan **skaalautumattomaksi** tai **mittakaavattomaksi verkoksi** (*scale-free network*), sillä verkon pisteiden yhteydet muuttuvat verkon koon muuttuessa, eikä verkon koko rakennetta voida päätellä jostakin verkon osaverkosta.

Potenssilakia noudattavassa verkossa astetta d olevien pisteiden osuus on $P(d) \sim cd^{-\gamma}$, missä γ on jokin kyseiselle verkolle tyypillinen vakio. Tästä jakaumasta seuraa, että verkossa on useita matala-asteisia pisteitä, mutta myös huomattava määrä pisteitä, joiden asteluku on hyvin suuri. Tällaisia pisteitä kutsutaan usein **navoiksi** tai **hubeiksi,** ja ne ovat verkon solmukohtia. Nämä navat ovat yhteydessä poikkeuksellisen moneen pisteeseen verkossa, minkä takia verkossa kahden pisteen välinen keskimääräinen etäisyys on yleensä varsin pieni. Tätä ilmiötä kuvaa esimerkiksi *kuuden erottavan askeleen teoria*, jonka mukaan keiden tahansa kahden ihmisen välillä on enimmillään kuusi ihmistä, mikäli ystävyyssuhteet käsitetään yhteyksiksi. Esimerkiksi tietyssä elinpiirissä, kuten kouluissa, harrastusten piirissä tai saman kaupungin sisällä ihmiset tuntevat toisensa verrattain hyvin, jolloin näistä yhteyksistä muodostuu niin kutsuttu **pieni maailma** (*small world*). Tästä pienestä maailmasta on vähemmän yhteyksiä ulospäin, mutta todennäköisesti muutamat keskeisissä asemissa olevat yksilöt kuten esimerkiksi poliitikot tai julkisuuden henkilöt ovat erityisen verkottuneita ja yhdistävät useita pieniä maailmoja toimien näin verkon napoina.

[2] `http://commons.wikimedia.org/wiki/File:Sample-graph.jpg` Lisenssi: CC BY 3.0

Tällä ajatuksella leikkivät muun muassa *Kevin Bacon number* ja *Erdősin luku,* jotka kertovat henkilön etäisyyden Kevin Baconiin sen perusteella, keiden kanssa kyseinen henkilö on näytellyt elokuvissa, tai etäisyyden matemaatikko Paul Erdősiin yhteisten tieteellisten julkaisujen perusteella.

Tätä pienen maailman ilmiötä tutki ensimmäisen kerran Stanley Milgram jo vuonna 1967 lähettämällä kirjeitä Yhdysvalloissa satunnaisille vastaanottajille pyytäen heitä lähettämään kirjeen sellaiselle tuttavalleen, jonka ajattelee olevan lähimpänä kirjeen lopullista vastaanottajaa. Tällä tavoin kokeilemalla Milgram sai kulun keskimääräiseksi pituudeksi kuusi ja näin pani alulle kuuden erottavan askeleen teorian. Nykyään kun tietoverkot ovat mahdollistaneet nopean sähköisen viestinnän ja sosiaalinen media yhdistää yhä useamman, on kulun keskimääräinen pituus luultavasti vain lyhentynyt. Teorian voi ajatella pätevän myös Internetin tapauksessa, sillä muun muassa hakukonesivustot yhdistävät lukemattomia vähemmän linkittyneitä sivustoja toisiinsa. Esimerkiksi vuosituhannen vaihteessa kahden sivun väliseksi keskimääräiseksi etäisyydeksi WWW:ssä mitattiin noin 19 erottavaa askelta, missä laskettiin sivulta toiselle siirtymiseen tarvittavien linkkien määrää.[3]

Skaalautumattomien verkkojen erityisominaisuus on kulkujen keskimääräisesti lyhyt pituus, koska verkon navat tai hubit yhdistävät suuren määrän pienempiasteisia pisteitä. Tämä rakenne on kuitenkin varsin haavoittuvainen sellaisille tapauksille, että verkosta poistuu yksi tai useampi hubi, jolloin äärimmäisimmillään verkko muuttuu jopa epäyhtenäiseksi. Sen sijaan jos verkosta poistetaan sattumanvaraisesti yksittäisiä pisteitä, voidaan niitä poistaa hyvin suuri määrä hajottamatta verkkoa, sillä skaalautumattomassa verkossa on suhteessa erittäin paljon merkityksettömiä pisteitä, joiden asteluku on pieni. Täysin satunnaisilla verkoilla sen sijaan on kriittinen määrä poistettavia pisteitä, jonka jälkeen verkko hajoaa nopeasti epäyhtenäisiin osiin.

Samaan tapaan lyhyen askelpituuden ansiosta yksittäinen häiriö voi edetä verkossa nopeasti hyvin eri puolille verkkoa nimenomaan verkon hubien kautta. Tosielämän tapausesimerkki tällaisesta *kaskadihäiriöstä* nähtiin Yhdysvaltain itärannikolla vuonna 2003, jolloin yksinkertainen paikallinen vika sähköverkossa levisi nopeasti laajoille alueille jättäen jopa 45 miljoonaa ihmistä ilman sähköä vaihteleviksi ajoiksi. Toinen mahdollinen tosielämän tapaus olisi yhteyksien katkeaminen kriittisten Internet-reitittimien välillä. Molemmissa tapauksissa kriittisen yhteyden katkeaminen tai hubin poistuminen ohjaa tieto- tai sähköverkon kuormituksen hubien sijaan pienemmän kapasiteetin solmukohtiin. Ylikuormittuessaan nekin poistuvat verkosta, jolloin kuorma ohjautuu jälleen pienemmille solmuille, ja näin kaskadihäiriö onkin valmis leviämään salamannopeasti.

Pulmia pohdittaviksi

Älä katkaise lankaa!

Tarvikkeet: narukerä. Osallistujien tehtävänä on järjestäytyä rinkiin ja muodostaa annetulla narulla sitä katkaisematta täydellinen verkko, jossa naru kahden osallistujan välillä tarkoittaa siis viivaa verkossa. Missä tapauksissa verkon muodostaminen on mahdollista?

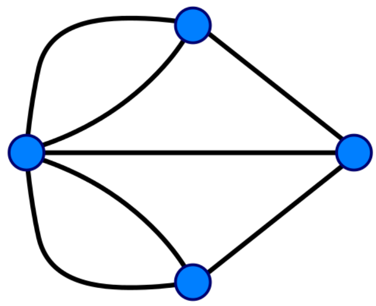
No more tears, Königsberg!

Eulerin kulku esiteltiin jo aiemmin artikkelissa: kyseessä on siis kulku, joka kulkee verkon **kaikkien viivojen** kautta tasan yhden kerran. Eulerin kulun olemassaololle on löydetty varsin käyttökelpoisia riittäviä ehtoja. Tässä niistä yksi:

Yhtenäisessä verkossa on Eulerin kulku, jos ja vain jos siinä on korkeintaan kaksi paritonasteista pistettä.

Mieti, miksi asia on näin!

Hamiltonin kuluksi taas kutsutaan sellaista kulkua, joka kulkee verkon **kaikkien pisteiden** kautta tasan yhden kerran. Hamiltonin kulun olemassaololle ei ole löydetty yleisiä riittäviä ehtoja, ja kulun etsiminen verkosta tai sen olemassaolon näyttäminen on ns. NP-täydellinen ongelma. Eulerin tai Hamiltonin kierrokseksi kutsutaan sellaista kulkua, joka on Eulerin tai Hamilton kulku ja palaa lähtöpisteeseensä.



Ohessa Königsbergin siltojen muodostama verkko.

a) Lisää verkkoon yksi silta (=viiva) siten, että verkossa on **Eulerin kulku**.
b) Lisää verkkoon vielä yksi silta (=viiva) siten, että verkosta löytyy **Eulerin kierros**, eli sellainen Eulerin kulku, joka palaa lähtöpisteeseensä.

[3] R. Albert, J. Hawoong ja A.-L. Barabási. "Internet: Diameter of the world-wide web." Nature 401 (1999): 130–131.

Cocktail party graph

”Cocktailkutsu-verkoksi” kutsutaan verkkoa, joka kuvaa cocktailkutsujen kättelemisiä. Ajatuksena on, että kutsuilla on n pariskuntaa (yhteensä $2n$ osallistujaa), ja kaikki kutsuvieraat kättelevät kaikkia muita vieraita paitsi omaa pariaan. Verkossa tämä tarkoittaa sitä, että kaikista pisteistä on viiva verkon muihin pisteisiin paitsi kyseisen pisteen omaan pariin.

Miltä cocktailkutsu-verkko näyttää, jos juhliin osallistuu

a) $n = 3$ pariskuntaa

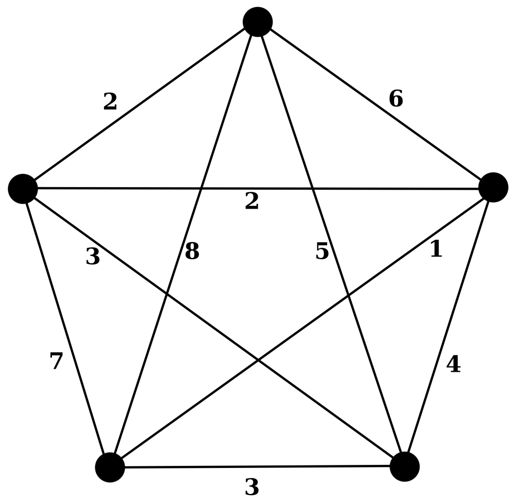
b) $n = 5$ pariskuntaa?

Saatko piirrettyä verkot siten, etteivät viivat leikkaa toisiaan? Ovatko verkot siis tasoverkkoja?

Reittikartan optimointi

Oheinen verkko kuvaa reittejä, joilla julkinen liikenne voidaan järjestää Yli-Hilseen kaupungissa. Verkko on painotettu yhteysvälien kustannusten mukaan, eli jokaisella viivalla on tietty painokerroin. Julkisten menojen leikkausten vuoksi liikenneverkko on toteutettava

- siten, että kaikista pisteistä voi kulkea minne tahansa (ts. verkko on yhtenäinen),
- mahdollisimman halvalla.



Halvin mahdollinen reittikartta on niin kutsuttu verkon virittävä puu. **Puu** on yhtenäinen verkko, jossa kahden pisteen välillä on olemassa tasan yksi (yksinkertainen) kulku: kahden eri pisteen välillä on siis vain yksi mahdollinen reitti. Verkon G **virittäväksi puuksi** sanotaan jotakin G:n puumuotoista aliverkkoa, joka koostuu joistakin verkon G viivoista, jotka yhdistävät kaikki pisteet P_G.

Tehtävänäsi on etsiä halvin mahdollinen reittikartta käyttäen kahta eri menetelmää:

a) ”Ahne algoritmi”:
Aloita puun rakentaminen ”halvimmista viivoista” eli viivoista, joiden painokerroin on mahdollisimman pieni, ja lisää niitä puuhun yksi kerrallaan, kunnes kaikki pisteet on saavutettu.

b) Renkaiden poistaminen:
Puun määritelmän mukaan puussa ei ole yhtään rengasta, siis yksinkertaista kulkua jonka alku- ja päätepisteet ovat samat. Lähde liikkeelle valmiista verkosta ja ala sitten poistaa sieltä viivoja, jotka kuuluvat johonkin renkaaseen.

Saitko molemmista kohdista samanlaisen puun? Jos et, onko niillä kuitenkin samat kokonaiskustannukset?

Kirjallisuutta ja muita mielenkiintoisia lähteitä

A. Barabási, *Linkit: verkostojen uusi teoria.* Terra Cognita, 2002.
Unkarilais-amerikkalainen fyysikko Albert-László Barabási on yksi keskeisiä verkostojen uuden teorian kehittäjiä. Muun muassa mittakaavattoman verkon käsite on hänen keksintönsä ja hän on tutkinut suurikokoisten sekä luonnossa esiintyvien että keinotekoisten verkkojen ominaisuuksia, kuten esimerkiksi häiriöalttiutta.

R. Diestel, *Graph theory.* Graduate texts in mathematics. Springer-Verlag Berlin and Heidelberg, 2000.

The Opte Project on hanke, jonka tarkoituksena on luoda taiteellinen näkemys Internetistä verkkomuodossa. Internetin kartasta voi helposti hahmottaa WWW:n ”pieni maailma” -rakenteen.
`http://www.opte.org/`

The Oracle of Bacon on verkkosivusto, joka laskee lähes kenen tahansa jossakin elokuvassa esiintyneen henkilön Kevin Baconin luvun.
`http://oracleofbacon.org/`