Tuomo Pieniluoma

Salatut vertaisverkot

Tietotekniikan
kandidaatintutkielma
5. maaliskuuta 2009

Jyväskylän yliopisto

Tietotekniikan laitos

Jyväskylä

Tekijä: Tuomo Pieniluoma
Yhteystiedot: tuomo.pieniluoma@jyu.fi
Työn nimi: Salatut vertaisverkot
Title in English: Anonymous Peer-to-Peer Networks
Työ: Tietotekniikan kandidaatintutkielma
Sivumäärä: 27
Tiivistelmä: Internet-sensuuri ja kiristyvä tietoliikenteen valvonta ovat siivittäneet yksityisyyden suojaa ja informaation säilyvyyttä parantavien teknologioiden kehitystä. Salattuja vertaisverkkoja voidaan pitää yhtenä varteenotettavimpana ratkaisuna sensuurin ja valvonnan mitätöimisessä. Tässä tutkielmassa tarkastellaan yleisellä tasolla nykyisten salattujen vertaisverkkojen toimintaa ja vertaillaan niiden vahvuuksia ja heikkouksia.
English abstract: The gradually increasing Internet censorship and communications surveillance are stimulating the research in privacy and information persistence enhancing technologies. Anonymous peer-to-peer (P2P) networks are considered to provide an effective solution to these problems. This thesis provides an overview of the current privacy and anonymity providing mechanisms in anonymous P2P networks and a comparison of their strengths and weaknesses.
Avainsanat: anonymiteetti, Internet-sensuuri, luotettavuus, luottamus, mainejärjestelmät, P2P, pienen maailman verkostot, reititys, salatut vertaisverkot, sekoitusverkot, sipulireititys, tietoturva, vertaisverkot, yksityisyyden suoja
Keywords: anonymous peer-to-peer, communication security, darknet, Internet censorship, mix-net, onion routing, P2P, peer-to-peer, privacy, reliability, reputation systems, routing, small world networks, trust

Files available at <URL: `http://users.jyu.fi/~tujupien/bachelor/`>.

Sisältö

1 Johdanto

Kaupallisen Internetin aikakauden alkamisen seurauksena 1990-luvun lopulla Internetissä alkoi tapahtua monia rakenteellisia muutoksia. Internetin käyttäjämäärä kasvoi rajusti ja samalla käyttäjien vaatimien palveluiden luonne ja painopiste muuttuivat. Verkon alkuperäistä hajautettua rakennetta keskitettiin yhä enemmän erillisille palvelimille. Vuosituhannen vaihteessa tämä kehitys kuitenkin kääntyi, kun vertaisverkot alkoivat saada enemmän jalansijaa Internetin käyttäjien keskuudessa. Tiedostoja ja erityisesti musiikkia levitettiin suoraan käyttäjältä toiselle. [33]

Napster oli ensimmäinen suuremman yleisön tietoisuuteen tullut vertaisverkko. Nopean suosion kasvun mukana saapui myös pikainen kuolema – Napster suljettiin tekijänoikeudellisista syistä heinäkuussa 2001. Vertaisverkkona Napster ei ollut kuitenkaan täysin hajautettu, mutta se esitteli suuremmalle yleisölle vertaisverkkojen ja yleisemmin hajautettujen järjestelmien vahvuuksia. [33]

Vertaisverkkojen mukana onkin noussut esille monia moraalisia ja oikeudellisia kysymyksiä, joihin on vaikea löytää yksikäsitteisiä vastauksia jo pelkästään Internetin globaalisuuden vuoksi. Näiden kysymysten pohtiminen on johtanut moninaisiin toimenpiteisiin eri osapuolten keskuudessa. Jotkut osapuolet ovat yrittäneet rajoittaa Internetissä saatavilla olevaa aineistoa erilaisilla sensuuri- ja valvontamenetelmillä, kun taas toiset ovat kehitelleet keinoja säilyttää ja jakaa aineistoa turvallisesti sensuurista ja valvonnasta riippumatta. [11, 33]

Esimerkkeinä vertaisverkkojen aiheuttamista reaktioista eri puolilla maailmaa voidaan mainita muun muassa kiireelliset tekijänoikeuslakien päivittämishankkeet ja massiiviset tekijänoikeudelliset oikeudenkäynnit [33] sekä Kiinan palomuuri, yksityisyyden suojaa vahvistavien menetelmien kehittäminen ja Kyberavaruuden itsenäisyysjulistus, jossa Barlow onnistuneesti kiteyttää eri osapuolten välillä havaittavissa olevaa vastakkainasettelua:

> *[...] Olemme luomassa maailmaa, jossa kuka tahansa, missä tahansa voi ilman pelkoa vaiennetuksi tulemisesta vapaasti ilmaista omia ajatuksiaan ja uskomuksiaan, olivatpa ne kuinka tavattomia tahansa.*
>
> *Teidän oikeudelliset käsityksenne omaisuudesta, ilmaisusta, henkilöllisyydestä, liikkumisesta tai kontekstista eivät koske meitä. [...]*
>
> *— John Perry Barlow, Kyberavaruuden itsenäisyysjulistus [6]*

Salatut vertaisverkot mahdollistavat anonyymin tavan jakaa ja hakea informaatiota, koska käyttäjien yksityisyys on niissä turvattu erityisillä reititysalgoritmeilla ja verkon solmujen

välisen tietoliikenteen salauksella. Tämä vaikeuttaa merkittävästi käyttäjien henkilöllisyyksien ja roolien selvittämistä. Samoilla menetelmillä voidaan tarjota kiistanalaiselle informaatiolle turvapaikka, josta sen poistaminen on käytännössä erittäin vaikeaa tai jopa täysin mahdotonta. [12, 24]

Täydellistä salausta ja anonymiteettiä ei kuitenkaan ole vielä keksitty digitaaliseen ympäristöön vaan kaikilla salausmenetelmillä on jokin heikkous. Nykyisissä salattujen vertaisverkkojen toteutuksissa joudutaankin usein tasapainottelemaan yksityisyyden suojan ja suorituskyvyn välillä. Viime vuosina salatut vertaisverkot ja niiden ongelmat ovat saaneet paljon huomiota myös tiedeyhteisössä. Kattavaa yleisen tason tutkimusta aiheesta ei siitä huolimatta ole juurikaan saatavilla ja erityisesti suomenkielinen aineisto aiheesta on erittäin vähäistä.

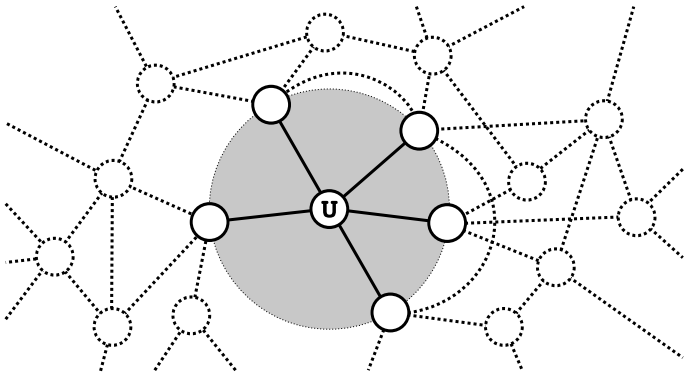
Tässä tutkielmassa käsitellään nykyisten salattujen vertaisverkkojen toimintaa erityisesti yksityisyyttä suojaavien ominaisuuksien kannalta. Luvussa 2 aihetta lähestytään hakualgoritmien ja reitityksen toiminnan kannalta. Luvussa 3 puolestaan tarkastellaan, kuinka anonyymissä ympäristössä käyttäjät voivat luottaa toisiinsa ja toistensa tarjoamiin resursseihin. Lisäksi luvussa 4 analysoidaan erilaisia hyökkäyksiä ja haavoittuvuuksia, joille salatut vertaisverkot ovat alttiita ja arvioidaan, kuinka hyvin nykyiset toteutukset niihin vastaavat. Lopuksi yhteenvedossa vertaillaan salatuista vertaisverkoista esiteltyjen lähestymistapojen keskeisiä ominaisuuksia sekä niiden vahvuuksia ja heikkouksia.

2 Reititys anonyymissä ympäristössä

Kaikille vertaisverkoille on yhteistä, että verkkoon voi muodostua tietoa hakiessa hyvinkin monimutkaisia reittejä, mutta salaamattomissa vertaisverkoissa voidaan lopullisessa tiedonsiirrossa hyödyntää alempien verkkokerrosten tarjoamaa reititystä. Tästä syystä hakualgoritmin tehokas toteutus muodostuu salaamattomien vertaisverkkojen suurimmaksi reitityksselliseksi haasteeksi [25, 33]. Salatuissa vertaisverkoissa reititys rakentuu salaamattomista vastineistaan poiketen omaksi kerroksekseen, sillä niissä on käytettävä itsenäistä reititysalgoritmia alempien verkkokerrosten päällä käyttäjien yksityisyyden suojaamiseksi. Ilman erillistä reititysalgoritmia hyökkääjän olisi vaivatonta selvittää IP-verkossa toimivan vertaisverkon käyttäjien henkilöllisyydet ja roolit kohteiden tietoliikennettä seuraamalla ja analysoimalla. Hyökkääjän ei tarvitsisi olla edes kytkeytyneenä vertaisverkkoon, sillä jokaiseen vertaisverkon pakettiin olisi valmiiksi merkitty alkuperäisen lähettäjän ja lopullisen vastaanottajan IP-osoitteet [38]. Salatuissa vertaisverkoissa käyttäjillä ei ole julkista IP-osoitetta vastaavaa tunnistetta. Joissakin toteutuksissa verkon käyttäjillä on väliaikainen ja salainen globaali tunniste, jota muut verkon käyttäjät eivät voi varmuudella yhdistää keneenkään [38], kun taas toisissa toteutuksissa käyttäjien sijaan tunnisteilla yksilöidään muodostetut yhteydet [12].

Reititys salatussa vertaisverkossa on peruspiirteiltään verrattavissa IP-verkon toimintaan, sillä verkon solmuilla on suora yhteys naapurisolmuihin ja reittiä muodostettaessa jokainen solmu on yksin vastuussa kauttansa kulkevan liikenteen ohjaamisesta oikeaan suuntaan. Tietoliikenne reititetään aina naapurilta toiselle, kunnes se saadaan lopulta toimitetuksi lopulliseen määränpäähänsä. Verkon solmut pitävät yllä mittavaa reititystaulua yhteyksistä ja niiden kulkusuunnista. Näiden taulujen avulla solmut osaavat jatkossakin reitittää vastaukset ja uudet yhteydet tietylle kohteelle samaa reittiä [12, 38]. Julkisten globaalien käyttäjätunnisteiden puuttuessa käyttäjät eivät pysty selvittämään toistensa sijaintia verkossa – tai välttämättä edes omaa sijaintiaan. Käyttäjät näkevät verkon eräänlaisena kuvan 2.1 esittämänä suljettuna pienen maailman verkostona *(small world network, darknet)*, jossa jokainen solmu näkee vain ja ainoastaan oman pienen maailmansa, ja kaikki sen ulkopuolinen sisältö projisoituu näkökentän reunalle [40]. Salattua vertaisverkkoa voitaisiin verrata myös sellaiseksi IP-verkoksi, jossa jokaisen verkon solmun välissä käytettäisiin osoitteenmuunnoksia siten, että jokainen verkon solmu peittäisi kaikki omat naapurinsa taakseen.

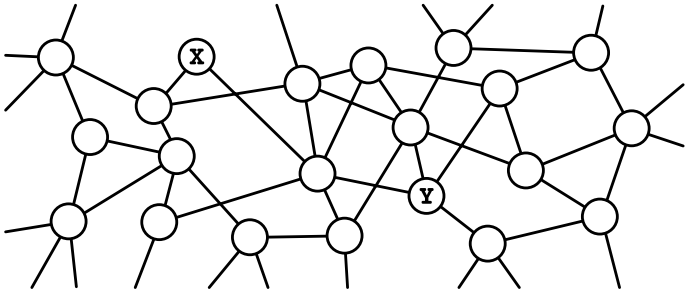
Salatun vertaisverkon luonne ja käyttötarkoitus määrittelevät, millainen reititysmenetelmä on yksityisyyden suojan ja verkon suorituskyvyn kannalta kannattava.



Kuva 2.1: Käyttäjä U ei voi tietää millaisen topologian sen naapurit peittävät taakseen.

2.1 Tulviminen

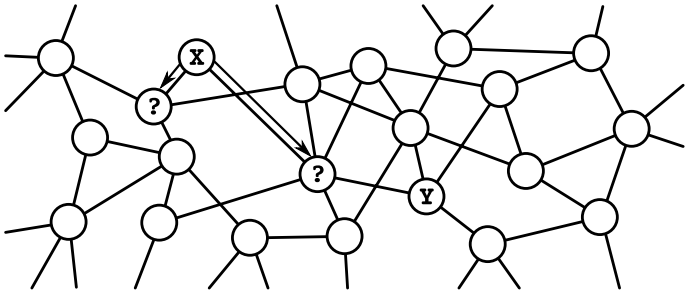
Erityisesti sellaisissa salatuissa vertaisverkoissa, joihin liittyessään käyttäjät tuovat mukanaan kokoelman omia tiedostojaan muiden käyttäjien saataville, käytetään reitinmuodostuksessa ja tiedonhaussa *tulvimismenetelmää* [28, 39]. Tulvimismenetelmässä viestit välitetään aina kaikille naapurisolmuille, jos viestin vastaanottajan suuntaan ei tiedetä reittiä jo entuudestaan. Tästä esimerkkinä alla oleva kuva 2.2. Jos solmu X haluaisi löytää reitin solmulle Y, se lähettäisi ensin viestin kaikille naapurisolmuilleen, jotka edelleen lähettäisivät viestin eteenpäin kaikkiin suuntiin, kunnes viesti lopulta päätyisi solmulle Y. Kuvissa 2.3 ja 2.4 on havainnollistettu tulvan etenemistä ja solmun Y löytymistä verkossa. [39]



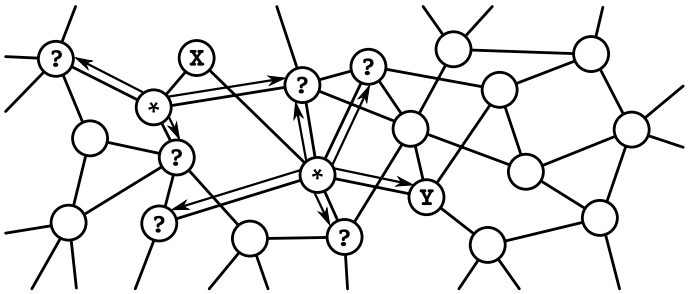
Kuva 2.2: Kuvitteellinen osa salattua vertaisverkkoa.

Tulviminen on tehokas tapa kattaa suuria alueita verkosta pienellä askelmäärällä, sillä tavoitettujen solmujen määrä kasvaa eksponentiaalisesti suhteessa kuljettujen askeleiden lukumäärään. On kuitenkin selvää, että väärällä tavalla sovellettuna tulviminen aiheuttaa verkolle valtavan kuormituksen. Rajoittamalla tulvintaa voidaan verkon kuormitusta vähentää huomattavasti. [39]

Jos tulvinta suoritettaisiin esimerkiksi verkossa, jossa jokaisella solmulla olisi viisi uutta naapurisolmua, kymmenennellä askeleella verkossa kulkisi jo noin 10 miljoonaa alkuperäi-



Kuva 2.3: X lähtee muodostamaan reittiä Y:lle tulvinnalla.



Kuva 2.4: Reitti X:ltä Y:lle löytyy, mutta verkko tulvii edelleen, koska kaikki verkon solmut eivät tiedä reitin jo löytyneen.

sen viestin kopiota. Seuraavalla askeleella viestit edelleen kertaantuisivat viidellä. Loputtoman monistuksen kierre voidaan katkaista yksinkertaisella TTL-laskurilla *(time-to-live counter)* samaan tapaan kuin IP-verkossa. Reitityksen kannalta TTL-laskuri tai jokin sitä vastaava rajoite tulvinnassa onkin välttämätön, jotta voidaan välttyä kuvassa 2.4 havainnollistetuilta ikuisesti monistuvilta viesteiltä ja vähentää myös muuta tulvinnan aiheuttamaa tarpeetonta verkon kuormitusta [21].

2.1.1 Muunneltu TTL-laskuri

Salatussa vertaisverkossa TTL-laskuria ei voida käyttäjien yksityisyyden suojaamisen vuoksi käyttää kiinteällä alkuarvolla. Mikäli TTL-laskuri lähtisi aina liikkeelle jostakin kiinteästä luvusta $k \in \mathbb{N}$, tällöin kaikki ne solmut, jotka saavat välitettäväkseen viestin TTL-arvolla k, tietäisivät viestin alkuperäisen lähettäjän olevan se naapuri, jolta viesti saatiin [39]. TTL-laskurin alkuarvon täytyy siis määräytyä satunnaisesti, mikä toisaalta aiheuttaa valtavaa vaihtelua verkon kuormituksessa.

Toisaalta tulvimisen aiheuttamaa yleistä verkon kuormitusta voidaan huomattavasti vähentää jos viestin haarautuminen verkossa asetetaan riippumaan käänteisesti TTL-laskurin arvosta [39]. Olkoon sääntönä esimerkiksi $\max\{1,\ B_{rajoite} - \mathrm{TTL}_n\}$, missä $B_{rajoite} \in \mathbb{N}$ on jokin vakio, joka määrittää kuinka monta kopiota viestistä saadaan tehdä viimeisellä askeleella.

Jos tämä sääntö määrittää kuinka moneen suuntaan viestin annetaan haarautua askeleella n, niin tällöin viesti voi tavoittaa maksimissaan $B_{rajoite}! + (\mathrm{TTL}_0 - B_{rajoite})$ verkon eri solmua. Tästä huomataan nopeasti, että suurin rajoittava tekijä tavoitettavien solmujen, ja siten myös verkon kuormituksen suhteen, on vakio $B_{rajoite}$, kun taas TTL-laskurin alkuarvo TTL_0 saa vaihdella vapaasti vaikuttamatta merkittävästi verkon kuormitukseen.

Erityisesti helposti ruuhkautuvissa ja asynkronisista yhteyksistä rakentuvissa verkoissa ongelmaksi voi muodostua kasvavat lähetysjonot. Esimerkiksi ADSL-yhteyksille tyypillinen lähetysnopeuteen verrattuna huomattavasti suurempi latausnopeus voi aiheuttaa joillekin verkon solmuille tilanteen, jossa solmulle saapuu enemmän viestejä välitettäväksi kuin niitä ehditään käsitellä. GNUnetissä TTL-laskuria käytetään hyödyksi tällaisten ruuhkatilanteiden selvittämisessä vähentämällä jonossa odottavien viestien TTL-laskurien arvoa tietyin väliajoin. Tällöin tulvimisen voidaan katsoa adaptoituvan automaattisesti verkon kuormitukseen [24].

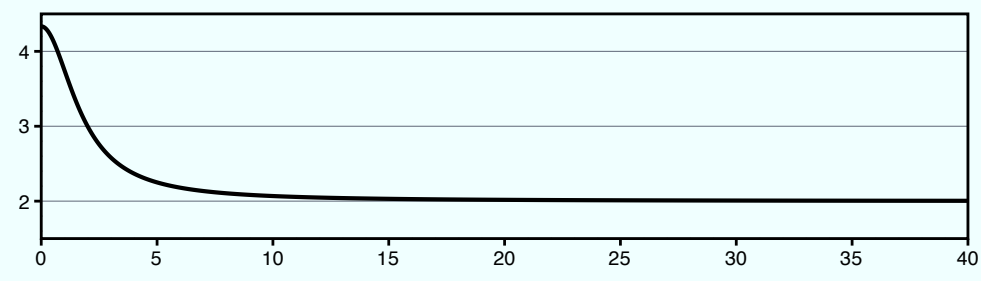
Lisäksi jos vastausviestiin merkitään tieto TTL-laskurin lopullisesta arvosta, niin silloin laskuria voidaan käyttää hyödyksi lyhimmän reitin selvittämisessä. Lyhimmän reitin etsimisessä voidaan myös käyttää Gnutellan tapaan askellaskuria TTL-laskurin rinnalla [21], mutta yksityisyyden suojaamiseksi myös askellaskurin alkuarvojen täytyy määräytyä satunnaisesti ainakin ylimmällä sovelluksen tasolla.

2.1.2 Hyötylaskuri

TTL- ja askellaskureiden puutteina voidaan pitää sitä, etteivät ne ota millään tavalla kantaa haarautumien lukumäärään tai korkeammalla tasolla löydettyihin hakutuloksiin. Sen sijaan hyötylaskurilla *(utility counter, UC)* voidaan yhtäaikaisesti kontrolloida verkon kuormitusta ja hakumenetelmän tehokkuutta, sillä siinä yhdellä muuttujalla seurataan reitin haarautumista, kuljettuja askeleita ja löydettyjä hakutuloksia. [39]

$$\mathrm{UC}_{n+1} := \mathrm{UC}_n + \alpha\, n\, R_{tulokset} + \beta\,(2 + \frac{7}{3n^2})\, B_{haarat} + \gamma \tag{2.1}$$

Hyötylaskurin seuraavan arvon laskeminen askeleella n voitaisiin parametrisoida esimerkiksi kaavan (2.1) mukaisesti, jossa painokertoimet α, β ja γ määrittelevät vastaavat painoarvot hakutuloksille, haarautumisille ja kuljetuille askelille. Lyhyemmillä etäisyyksillä n on pieni ja hakutulokset vaikuttavat hyötylaskurin kasvuun vähemmän; pidemmillä etäisyyksillä n:n kasvaessa samat hakutulokset kasvattavat hyötylaskurin arvoa enemmän. Haarautumisien kerrointermi $(2+\frac{7}{3n^2})$ varmistaa sen, että haarautumiset pienillä n:n arvoilla saavat enemmän painoarvoa. Kerrointermin käyttäytymistä on esitelty kuvassa 2.5. Näillä menetelmillä laskuri suosii lyhyemmältä etäisyydeltä saatuja vastauksia ja ottaa huomioon aikaisemmassa vaiheessa tehtyjen haarojen vaikutuksen verkon kuormitukselle. [39]



Kuva 2.5: $f(x) = 2 + \frac{7}{3x^2}$. Muuttujan x arvot kulkevat vaaka-akselilla ja pystyakselilla vastaavasti funktion $f(x)$ arvot.

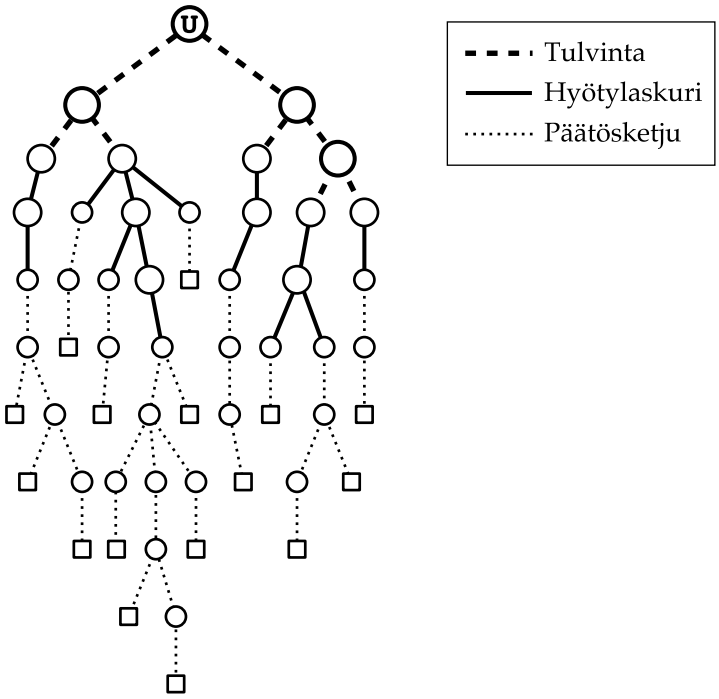
Sellaisenaan hyötylaskuria ei voida käyttää salatussa vertaisverkossa, sillä se uhkaa käyttäjien yksityisyyttä samaan tapaan kuin TTL-laskuri. Eräs keino salata viestin alkuperäisen lähettäjän henkilöllisyys on lähettää viesti aluksi satunnaisen pituiselle polulle verkkoon ilman hyöty- tai TTL-laskuria. Satunnaisuutta voidaan kontrolloida käyttämällä yhteisesti sovittua satunnaislukugeneraattoria ja välittää sen tilatunnisteet aina viestiketjussa seuraavalle. Lopulta satunnaislukugeneraattorin antaessa sopivan arvon voidaan viesti muuttaa satunnaiskulkijasta tai tiukasti rajoitetusta tulvinnasta laskuria noudattavaksi tulvaksi. Kontrolloimaton satunnaisuus johtaisi suuremmissa vertaisverkoissa jossakin vaiheessa ikuisesti kiertäviin viesteihin, jotka tuhlaisivat verkon resursseja. [39]

Vastaajan yksityisyyttä ei pystytä takaamaan hyöty- tai TTL-laskureilla, jos viestin alkuperäinen lähettäjä voi laskurin alkuarvoja muuntelemalla määrittää viestiketjun maksimipituudeksi yhden askeleen. Tällainen hyökkäys voidaan tehdä tehottomaksi määrittämällä jokaiselle verkon solmulle yksilölliset päätösketjut niitä kyselyitä varten, jotka päättyvät kyseiselle solmulle. Kuvassa 2.6 on esitetty puumuodossa, kuinka hyötylaskuria käyttävä haku voisi edetä esimerkiksi MUTEssa tai sen kaltaisessa vertaisverkossa alkuvaiheen rajoitetun tulvinnan ja päätösketjun kanssa. Hyökkääjän täytyisi pystyä seuraamaan kaikkea kohteensa liikennettä, jotta se voisi olla varma kohteensa roolista salatussa vertaisverkossa. [39]

2.2 Jyrkimmän nousun menetelmä ja ahne reititys

Salatuissa vertaisverkoissa, joissa tieto on mahdollista indeksoida ja lajitella, voidaan reitinmuodostuksessa käyttää tulvimisen sijaan erilaisia kulkijamenetelmiä. Kuva 2.7 esittelee erään kulkijamenetelmän kohteen löytämiseen.

Jyrkimmän nousun menetelmässä oletetaan, että verkon solmut ovat tietoisia – tai pystyvät ainakin tekemään hyviä arvauksia – naapurisolmujensa kautta tavoitettavissa olevista tiedostoista. Tätä tietoa verkkoon juuri kytkeytyneellä solmulla ei voi olla, ja sen täytyykin aluksi reitittää sen kautta kulkevat kyselyt täysin satunnaisesti, kunnes se saa reitittämiensä



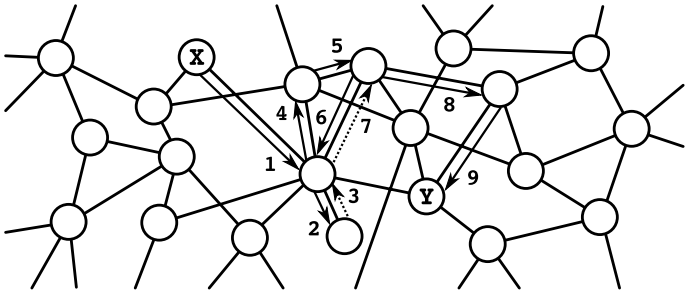
Kuva 2.6: Esimerkkipuuesitys hyötylaskurin käytöstä salatussa vertaisverkossa.

kyselyiden ja vastausten mukana tarpeeksi tietoa eri indeksien jakautumisesta verkkoon naapuriensa suhteen [12]. Jos verkon solmut pystyvät vielä reitittämään kyselyt aina kohdetta topologisesti lähimpänä olevalle naapurilleen, kutsutaan käytettyä reititystä *ahneeksi reitistykseksi* [40]. Uusien reittien muodostuksen vuoksi on tärkeää käyttää ajoittain satunnaista reititystä, vaikka oikeasta reitistä olisikin jo olemassa hyvä arvaus. [12]

Esimerkiksi jyrkimmän nousun menetelmä Freenetissä perustuu verkossa kulkevien tiedostojen indeksien ja tietoliikenteen suunnan seurantaan. Uudet kyselyt reititetään siihen suuntaan, josta samankaltaiset kyselyt ovat yleensä löytäneet myönteisen vastauksen. [12]

Tiedostojen indeksointi voidaan toteuttaa esimerkiksi hajautusalgoritmin avulla muodostamalla tiiviste jokaiselle tiedostolle. Freenetissä tiedostot, joilla on samankaltainen tiiviste, sijoitetaan verkossa topologisesti lähekkäin, ja tätä tietoa hyödyntämällä reitinmuodostus voidaan toteuttaa tehokkaasti ilman tulvintaa. Hyvällä hajautusalgoritmilla voidaan varmistaa sekä tiedostojen tasainen jakautuminen verkon eri osiin että turvata tiedon säilyvyys. Toimintaperiaatteeltaan Freenetin kaltaisia sovelluksia voidaankin pitää valtavina hajautettuina tiedostopankkeina. [12, 30]

Edellä kuvattu menetelmä ei ota mitään kantaa löydetyn reitin tehokkuuteen; menetelmällä löydetty reitti voi kulkea jonkin erittäin ruuhkautuneen solmun kautta tai voi olla olemassa jokin toinen muulla tavoin edullisempi reitti. Jos solmut taulukoivat pelkkien suuntien lisäksi myös kyselyiden vasteaikoja ja tekevät reitinvalintapäätöksensä näiden molem-



Kuva 2.7: Reitin muodostuminen X:ltä Y:lle eräällä kulkijamenetelmällä. Vaiheessa 2–3 reitti toipuu umpikujasta ja vaiheessa 6–7 vältytään silmukalta. Lopullinen reitti etenee siis askeleita 1, 4, 5, 8 ja 9.

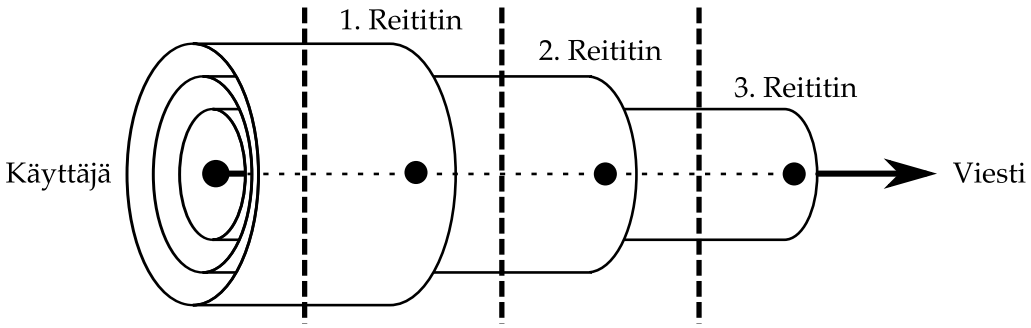
pien tietojen pohjalta, saadaan verkko tällöin adaptoitumaan automaattisesti sen rakenteellisiin muutoksiin sekä ruuhka- ja virhetilanteisiin. Uusien verkon solmujen sopeutumisaikaa voitaisiin lisäksi vähentää jakamalla niille satunnaisten verkon solmujen tuntemusta omista naapureistaan. [10]

Reititys voidaan tällaisessa ympäristössä toteuttaa yksinkertaisemmin nojautumalla enemmän todennäköisyyksiin. Esimerkiksi Freenetin reititys on nykyään toteutettu edellä kuvatun monimutkaisen verkon solmujen naapuruston kartoituksen sijaan Metropolis-Hastings -algoritmin mukaisella todennäköisyyslaskentaan perustuvalla ahneella reitityksellä [10, 41]. Metropolis-Hastings -algoritmia käytetään Freenetin tapauksessa Markovin ketjun Monte Carlo -menetelmän toteuttamiseen. Tällaisella satunnaiskulkijan ja ahneen reitityksen yhdistelmällä pyritään minimoimaan kyselyiden ajautumista umpikujiin ja reittien askelpituutta sekä yleistä verkon ruuhkautumista. [40, 41]

2.3 Sekoitusverkko

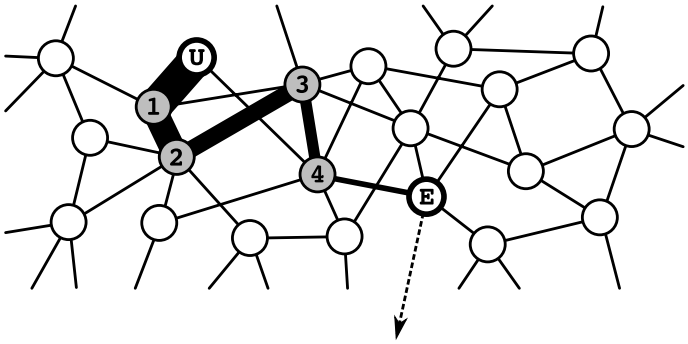
Sekoitusverkkojen *(mix-net)* perusperiaate on edelleen sama kuin Chaumin [8] 1980-luvun alussa esittelemä salausjärjestelmä. Toisin kuin luvuissa 2.1 ja 2.2 esitellyt reitit, sekoitusverkko ei tarjoa suoraa hakutoimintoa. Sekoitusverkkojen tarkoitus on muodostaa kerros, jota käyttäjät voivat hyödyntää oman yksityisyytensä suojaamisessa käyttäessään jotakin toista sovellusta, kuten lähettäessään sähköpostia [8, 31] tai selatessaan WWW-sivuja [44]. Alla oleva kuva 2.8 havainnollistaa kuinka satunnaisten osoitteenmuunnosten ja sisäkkäisten salattujen tunneleiden sarjalla saadaan reititettyä tietoliikenne kolmannelle osapuolelle ilman, että se tai kukaan viestiä välittänyt osapuoli voi olla varma, keneltä vertaisverkon käyttäjältä tietoliikenne oli alkuaan lähtöisin. [22, 44]

Käytännössä kaikissa sekoitusverkkojen lähestymistavoissa reititys toteutuu samalla tavalla. Tietoliikennettä ohjataan sopivan mittainen matka satunnaisesti valittua reittiä pitkin



Kuva 2.8: Kolmen askeleen tunnelointi sekoitusverkon satunnaisten solmujen eli reitittimien kautta.

ja lopulta se ohjataan sekoitusverkosta ulos kohti todellista määränpäätä [22, 31, 42], kuten kuva 2.9 selventää. Jos verrataan esimerkiksi I2P- [2] ja Tor-vertaisverkkoja [44] keskenään, huomataan niiden molempien reitityksen toimivan tällä samalla periaatteella, vaikka niiden kehityksessä onkin tehty monia vastakkaisia ratkaisuja.



Kuva 2.9: Käyttäjä U muodostaa tunnelin sekoitusverkkoon solmujen 1–4 kautta aina solmulle E asti, joka ohjaa U:lta lähtevän liikenteen ulos verkosta.

Reitityksen tehokkuutta sekoitusverkoissa on helpompi kontrolloida edellä esitettyihin menetelmiin verrattuna, sillä rakennettujen tunneleiden päätepisteillä ei ole mitään merkitystä; muodostetut tunnelit voidaan vaihtaa uusiin koska tahansa ja uudet tunnelit voidaan muodostaa täysin edellisistä riippumatta. Ylimääräiseltä kokeilevalta tunneleiden rakentelulta voidaan välttyä tarkkailemalla verkon osien ruuhkautumista ja välttämällä uusien tunneleiden muodostamista ruuhkautuneiden tai helposti ruuhkautuvien alueiden kautta. [16,17]

3 Mainejärjestelmät luotettavuuden takeena

Luotettavuus on yksi keskeinen salattujen vertaisverkkojen aihealue, sillä pelkästään vertaisverkon tietoliikenteen salaaminen ja käyttäjien yksityisyyden turvaaminen eivät tee salatusta vertaisverkosta käyttökelpoista. Anonymiteetin tarjoamin turvin käyttäjät voivat virheellisellä toiminnallaan helposti häiritä verkon muita käyttäjiä tai koko verkkoa. Vertaisverkon on tästä syystä pystyttävä jollakin tapaa takaamaan solmuille niiden luottamus toisiinsa, eli käytännössä varmistamaan verkon solmujen oikeellinen ja rehellinen toiminta. [15]

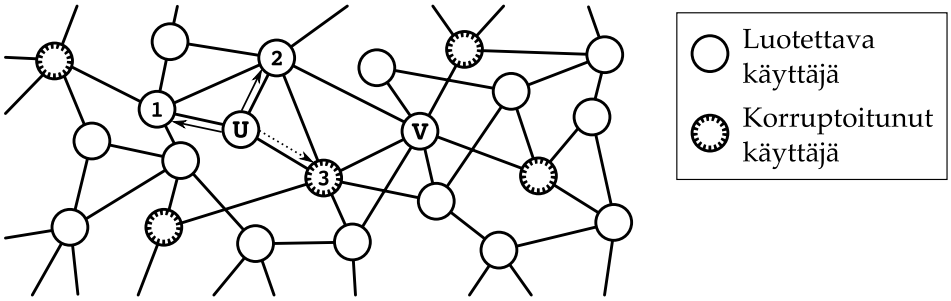
Salatuissa vertaisverkoissa erityisen haasteen luottamukselle asettaa juuri käyttäjien yksityisyyden suoja; on paljon helpompi arvioida jonkin tunnetun kuin täysin nimettömän tahon luotettavuutta. Täysin anonyymissä ympäristössä käyttäjä voi arvioida vain välittömien naapureidensa kautta muodostettavien ketjuuntuvien, transitiivisten yhteyksien luotettavuutta, mutta luotettavuustietoja ei voida yhdistää kehenkään tiettyyn käyttäjään. Lisäksi vertaisverkkojen hajautettu rakenne tuo luotettavuuden arviointiin ja luotettavuustietojen käsittelyyn oman haasteensa, sillä hajautetussa ympäristössä ei välttämättä ole mitään erillistä auktoriteettia, joka valvoisi yleistä järjestystä. [4, 15]

Luotettavuutta voidaan yrittää kontrolloida suoraan protokollan avulla tai kytkemällä käyttäjien oikeelliseen ja virheelliseen käyttäytymiseen toisistaan poikkeavia kustannuksia, jotka kannustavat käyttäjiä yleiseen luotettavuuteen [14]. Suosiota ja mainetta pidetään yleisesti ottaen luontevana suureena vertaisverkon käyttäjien luotettavuuden mittaamisessa. Useimmissa toteutuksissa järjestelmän eheys taataankin jonkinlaisella *mainejärjestelmällä*, joka helpottaa entuudestaan tuntemattomien verkon käyttäjien luotettavuuden arviointia. Tällaisen järjestelmän tarkoituksena ei yleensä ole taata täydellistä luotettavuutta vaan pikemminkin selvästi parantaa sitä verrattuna kontrolloimattomaan tapaukseen, ja sen kehittäminen onkin yleensä muihin vaihtoehtoihin nähden helpompaa ja kustannustehokkaampaa [14].

Mainejärjestelmä on kuitenkin syytä jakaa sen toimintojen perusteella komponentteihin, jotta aihetta olisi helpompi tarkastella, ja kokonaisuuden suunnittelu ja kehittäminen olisi suoraviivaisempaa. Kattavan mainejärjestelmän voidaan katsoa koostuvan kolmesta eri komponentista: käyttäjistä ja aineistosta saatavilla olevien luotettavuustietojen keräämisestä; käyttäjien luotettavuuden arvioimisesta ja luotettavuuspisteiden laskennasta; sekä pisteiden vaikutuksesta muun järjestelmän toimintaan eli mainejärjestelmän vasteesta. [13, 15, 32]

3.1 Luotettavuustietojen selvittäminen

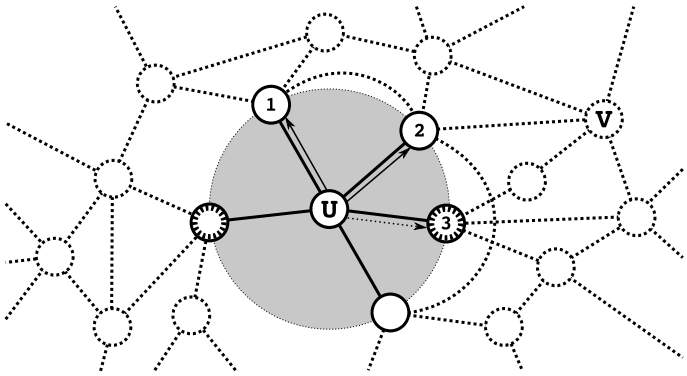
Käyttäjä voi lähteä kartoittamaan verkon luotettavuutta siitä oletuksesta, että kaikki verkon käyttäjät ovat oletusarvoisesti joko luotettavia tai epäluotettavia. Tällaisessa lähestymistavassa on kuitenkin se heikkous, että käyttäjä joko riskeeraa merkittävän määrän resursseja optimistisella asenteellaan tai joutuu kuluttamaan paljon resursseja luotettavien vertaisten etsimiseen. Käyttäjän on helppo arvioida sellaisten vertaistensa luotettavuutta, jotka ovat sille jo entuudestaan tuttuja, mutta ennalta tuntemattomien vertaisten luotettavuuden tarkka arviointi onkin jo huomattavasti vaikeampaa. Luotettavuustietoja täytyy tuolloin pyytää muilta käyttäjiltä, joiden luotettavuutta ei välttämättä myöskään tunneta. Jos kuitenkin voidaan olettaa, että enemmistö verkon käyttäjistä on luotettavia, niin satunnaiselta käyttäjäjoukolta kerätyt luotettavuustiedot ovat silloin keskimäärin luotettavia ja niistä voidaan edelleen tehdä johtopäätöksiä yksittäisten käyttäjien luotettavuudesta. Virheellisten luotettavuustietojen saantia voidaan vältellä myös pyytämällä luotettavuustietoja jo aiemmin luotettaviksi todetuilta käyttäjiltä, jotka edelleen pyytävät tietoja luottamiltaan käyttäjiltä. Näin muodostuvat transitiiviset luotettavuusketjut minimoivat epäluotettavien lähteiden vaikutuksen luotettavuuden arvioinnissa. [32, 35]



Kuva 3.1: Käyttäjä U pyytää välittömiltä naapureiltaan 1–3 käyttäjän V luotettavuustietoja.

Oletetaan, että kuvissa 3.1 ja 3.2 käyttäjällä U olisi jo hyvä käsitys välittömien naapureidensa 1–3 luotettavuudesta. Jos käyttäjä U haluaa selvittää käyttäjän V luotettavuuden, kannattaa käyttäjältä 3 saadut vastaukset jättää pienemmälle painoarvolle tai jopa kokonaan huomioimatta. Käyttäjällä 2 on luultavasti parempi käsitys V:n luotettavuudesta, mutta käyttäjä 1 joutuu todennäköisesti edelleen pyytämään tarkempia tietoja omilta välittömiltä naapureiltaan. Kuvassa 3.2 verkon anonymiteetti on oletettu niin vahvaksi, että U ei voi tietää V:n sijainnista enempää kuin sen, että se on helposti tavoitettavissa naapureiden 1–3 kautta.

Käyttäjä ei voi kuvan 3.2 kaltaisessa vahvasti salatussa verkossa varmasti tietää, onko sen välitön naapuri itse korruptoitunut vai muodostaako naapuri reittejä korruptoituneiden käyttäjien kautta. Tällaisessa tilanteessa naapuri voidaan luokitella epäluotettavaksi, koska se ei osaa muodostaa luotettavia reittejä.



Kuva 3.2: Käyttäjä U haluaa selvittää tuntemattoman käyttäjän tai resurssin V:n luotettavuuden ja yrittää selvittää sen naapureidensa 1–3 avustuksella.

Käyttäjä voi passiivisen luotettavuuden seuraamisen lisäksi itse aktiivisesti tutkia vertaistensa luotettavuutta [35]. Verkon solmut voivat erityisesti joutilaina ollessaan vaihtaa luotettavuustietoja toistensa kanssa ja päivittää esimerkiksi luotettavuutta kuvastavaa Bayes-verkkoaan näiden tietojen pohjalta [43]. Jos käytettävissä on jonkinlainen kuittaus- tai muu palautemekanismi, myös sitä voidaan hyödyntää yksittäisten käyttäjien luotettavuuden tutkimisessa; erityisesti niillä voidaan varmistaa, ettei yksittäistä käyttäjää voida leimata epäluotettavaksi ilman oikeaa syytä [14].

Uudet käyttäjät tuovat luotettavuuden arviointiin oman haasteensa. Koko verkolle tai sen paikalliselle osalle tuntemattoman käyttäjän luotettavuutta arvioitaessa voidaan käyttää aiemmin mainittuja optimistisia tai pessimistisiä lähestymistapoja, mutta niillä molemmilla on huomattavia heikkouksia. Korruptoituneet käyttäjät voivat käyttää hyödyksi optimistista lähestymistapaa valkaisussa *(whitewashing)* eli identiteettinsä toistuvassa muuttamisessa. Vastaavasti uusien käyttäjien on vaikea liittyä verkkoon jos ne aina oletetaan epäluotettaviksi. Uusien käyttäjien luotettavuutta voidaan arvioida adaptoituvalla muukalaispolitiikalla *(stranger policy)*, joka arvioi ennestään tuntemattomien käyttäjien luotettavuutta aikaisemmista tuntemattomista käyttäjistä kertyneiden kokemusten perusteella. Tällaisella politiikalla vältetään optimististen ja pessimististen lähestymistapojen ongelmat ja vähennetään samalla valkaisusta saavutettavia hyötyjä. [19, 29, 32]

3.2 Käyttäjien pisteytys ja rankkaus

Käyttäjistä kerätyistä luotettavuustiedoista voidaan lopulta laskea jokaiselle käyttäjälle jotakin sovittua metriikkaa noudattavat luotettavuuspisteet. Lopullisiin pisteisiin vaikuttavat yleensä sekä käyttäjistä saatu positiivinen että negatiivinen palaute. Positiivisessa palautteessa on kuitenkin usein se ongelma, että korruptoituneiden käyttäjien voi olla helppo manipu-

loida sitä [14], ja sen vuoksi joissakin toteutuksissa tarkastellaankin pelkästään epäluotettavuutta. Luotettavuuspisteet ohjeistavat käyttäjää valitsemaan verkosta luotettavia lähteitä. Todellisuudessa luotettavuuden lisäksi lopulliseen reitin ja kohteen valintaan saattaa vaikuttaa myös eri reittien tehokkuudet ja verkon eri osien ruuhkautuneisuus. [13, 26, 32]

Pisteytyksen kannalta on tärkeää tarkastella, mitkä asiat vaikuttavat käyttäjän luotettavuuspisteisiin. Käyttäjän korruptoituneisuus eli virheellinen toiminta vaikuttaa luotettavuuteen negatiivisesti. Korruption tiheys ja laatu yhdistettynä antaa paremman kuvan käyttäjän korruptiosta kuin kumpikaan niistä erikseen. Samaan tapaan käyttäjän tarjoaman hyvän palvelun laadun ja määrän yhdistelmä antaa kattavan kuvan käyttäjän luotettavuudesta. Jos vielä painotetaan viimeisimpiä käyttäjän toimia vanhoja enemmän, järjestelmä saadaan varautumaan mahdollisia äkillisiä korruptioita vastaan ja vastaavasti antamaan vanhat virheet nopeammin anteeksi. [32]

3.2.1 Vapaamatkustajat

Koko salatun vertaisverkon käytettävyys sen suunnitellussa käyttötarkoituksessa laskee sitä hyödyntävien vapaamatkustajien määrän kasvaessa; tilanne on verrattavissa yhteismaan ongelmaan *(tragedy of the commons)*, jossa käyttäjät ajavat itsekkäästi omaa hyötyään maksimoidessaan koko järjestelmän tilaan, joka on kaikkien edun vastainen. Mainejärjestelmän pisteytys ja sen vaste on juuri se työkalu, jolla vapaamatkustukseen voidaan tehokkaasti puuttua. [18, 32]

Kokonaisuutta hyödyttävä panostus verkon toimintaan voidaan palkita tai päinvastoin vapaamatkustuksesta voidaan rangaista. Antamalla positiivista palautetta kokonaisuutta hyödyttävää panostusta tarjoavista käyttäjistä, voidaan tehdä ero osallistujien ja vapaamatkustajien välille, ja tästä seuraavalla vasteella käyttäjiä voidaan kannustaa toimimaan vähemmän itsekkäästi ja enemmän kokonaisuuden hyväksi. Rehellisten ja luotettavien uusien käyttäjien kannalta on harmillista, että valkaisun hyötyä vapaamatkustuksessa voidaan vähentää vain rinnastamalla uudet käyttäjät vapaamatkustajiin; kokonaisuuden kannalta sillä on kuitenkin rakentava vaikutus. [18, 20]

3.3 Luotettavuuden ja epäluotettavuuden seuraukset

Käyttäjille voidaan tarjota erilaisia etuja jos he osoittautuvat luotettaviksi käyttäjiksi. Edut toimivat samalla kannusteina koko järjestelmää hyödyttävästä ja rakentavasta yhteistyöstä. Yleisimmin eri toteutuksissa on käytetty etuina suurempaa kaistanleveyttä, parempaa palvelun laatua ja sen suurempaa määrää. Esimerkiksi latausnopeus voidaan sitoa käyttäjän asettamiin lähetysnopeuden rajoituksiin ja jaettuun aineistoon tiedostojen levitykseen tarkoite-

tuissa toteutuksissa. Yhteyksiä voidaan lisäksi priorisoida käyttäjien oman osallistumisen perusteella, jolloin luotettaville käyttäjille voidaan taata paremmat yhteydet. [32]

Mainejärjestelmät osoittavat tehokkuutensa erityisesti korruptoituneiden käyttäjien hallinnassa. Tahallisesti korruptoituneet ja jatkuvasti epäluotettaviksi osoittautuvat verkon käyttäjät voidaan parhaassa tapauksessa sulkea kokonaan verkon ulkopuolelle. [32]

Mikä tahansa mainejärjestelmä kuitenkin aiheuttaa vertaisverkolle lisää yleiskustannuksia *(overhead)*, kun normaalien hakujen lisäksi joudutaan tekemään luotettavuuskyselyitä. Lisäksi joudutaan reititystaulujen ohella tallentamaan luotettavuustietoja ja tekemään niihin liittyvää laskentaa. Optimoinnilla voidaan vähentää kustannuksia huomattavasti ja toisaalta mainejärjestelmällä saavutettavat hyödyt ovat yleensä niin suuret, että niiden katsotaan kattavan järjestelmän mukanaan tuomat kustannukset. [4, 26]

4 Turvallisuus ja yksityisyyden suoja

Salattujen vertaisverkkojen turvallisuuden suurimmat haasteet ovat yksityisyyden suojassa. Tietoliikenteen sisältö voidaan salata tehokkaasti kryptografisilla menetelmillä ilman, että käytetty menetelmä vaikuttaisi suuresti verkon käytettävyyteen sen suunnitellussa käyttötarkoituksessa. Käyttäjien yksityisyyttä on kuitenkin vaikeampi suojata tehokkaasti kuluttamatta siihen suhteessa huomattavasti enemmän resursseja. Käytännössä yksityisyys turvataan salatuissa vertaisverkoissa aina epävarmuuden periaatteella – kuka tahansa verkon käyttäjistä voi olla minkä tahansa viestin alkuperäinen lähettäjä tai lopullinen vastaanottaja. Epävarmuuden periaate on sitä tehokkaampi mitä enemmän käyttäjiä vertaisverkolla on, ja sen tarjoamaa yksityisyyden suojaa voitaisiinkin parantaa, jos kaikki vertaisverkkoon liittyvä tietoliikenne pystyttäisiin naamioimaan siten, että sen erottaminen muusta Internetin tietoliikenteestä olisi mahdotonta; silloin kaikki Internetin käyttäjät olisivat potentiaalisia vertaisverkon käyttäjiä. [37, 40]

4.1 Hyökkäykset, haavoittuvuudet ja puolustuskeinot

Salattujen vertaisverkkojen haavoittuvuuksia tarkasteltaessa on tärkeää analysoida eri hyökkäysmenetelmien tehokkuutta vertaisverkon omia ominaisuuksia vastaan. Seuraavassa hyökkäysmenetelmien tarkastelussa oletetaan, että vertaisverkon alla oleva verkko- ja järjestelmäarkkitehtuuri ei kuulu vertaisverkon turvallisuusmääritykseen ja, että hyökkääjä voi käyttää niitä hyväkseen murtaessaan käyttäjän yksityisyyden suojaa. Lisäksi on oletettu, että hyökkääjän käytössä on sellaisia välineitä ja resursseja, joilla salattujen vertaisverkkojen käyttämät kryptografiset suojausmenetelmät voidaan murtaa riittävän lyhyessä ajassa, jotta hyökkäystä voidaan ja sitä on kannattavaa jatkaa. Näiden oletusten turvin voidaan tarkastella hyökkäyksen viimeistä vaihetta, jossa hyökkääjä on jo onnistuneesti raivannut tieltään kaikki muut esteet.

Käytännössä hyökkääjän voi olla erittäin vaikeaa – ellei jopa mahdotonta – toteuttaa edellä esitettyjä oletuksia, koska tarpeeksi vahvaa kryptografista salausta käyttämällä hyökkäystä voidaan hidastaa merkittävästi; avaimet ja informaatio ovat jo vanhentuneet, kun hyökkääjä lopulta saa ne murrettua [24, 38]. Salatun vertaisverkon turvallisuus ja yksityisyyden suojan vahvuus voidaan kuitenkin määritellä hyökkäyksen viimeisen vaiheen tehokkuuden ja seurausten avulla. Heikkoa suojaa tarjoavissa toteutuksissa hyökkäyksen viimeinen vaihe on helppo suorittaa ja sillä voidaan saada aikaiseksi mittavaa vahinkoa, kun taas vahvaa suo-

jaa tarjoavissa toteutuksissa viimeinen vaihe voi olla mahdoton tai saavutettaviin hyötyihin nähden kannattamaton toteuttaa.

4.1.1 Soluttautuminen

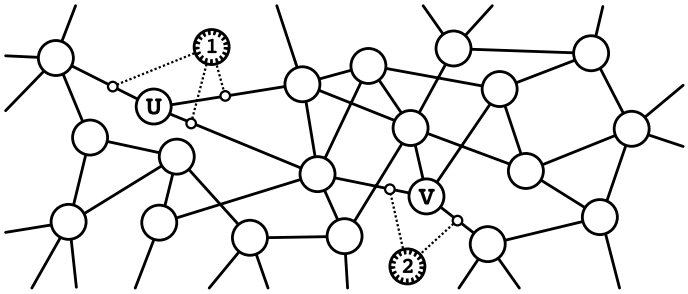
Soluttautuminen vertaisverkkoon on muihin hyökkäysmenetelmiin verrattuna helppo tapa päästä käsiksi vertaisverkon tietoliikenteeseen. Soluttautumalla yhdeksi vertaisverkon käyttäjistä hyökkääjä pääsee heti käsittelemään reitittämäänsä tietoliikennettä ja samalla hän välttää vertaisverkon ulkopuolisia hyökkäyksiä vastaan suunnitellut kryptografiset suojausmenetelmät, kuten vertaisverkon solmujen väliset salatut yhteydet *(link encryption)*. Pelkästään soluttautumalla kerätyistä tiedoista ei välttämättä ole mitään hyötyä yksittäisiä käyttäjiä vastaan kohdistetussa hyökkäyksessä, sillä hyökkääjä ei voi olla varma yksittäisen käyttäjän osallisuudesta tarkkailemaansa tietoliikenteeseen, ellei hän kontrolloi kaikkia kohteensa naapurisolmuja [9, 38]. Jos hyökkääjä kontrolloisi esimerkiksi kaikkia kuvissa 3.1 ja 3.2 esitettyjä korruptoituneita verkon solmuja, hän ei silti voisi olla täysin varma käyttäjän U osallisuudesta kaappaamaansa liikenteeseen. Tämän lisäksi hyökkäys salatun vertaisverkon sisältä koko järjestelmää vastaan voidaan estää tehokkaasti esimerkiksi luvussa 3 kuvatulla mainejärjestelmällä.

Joissain tapauksissa soluttautuminen voi kuitenkin olla merkittävä turvallisuusriski salatulle vertaisverkolle. Esimerkiksi sekoitusverkossa toimiva soluttautuja voi verkostapoistumissolmuna *(exit node)* toimiessaan saada melko vaivattomasti käsiinsä arkaluonteista tietoa, kuten toisten käyttäjien käyttäjätunnuksia, salasanoja ja luottokorttitietoja sekä muita henkilökohtaisia tai yrityksen tietoja. On kuitenkin huomattava, että käyttäjä voi altistua aivan samoille turvallisuusriskeille ilman sekoitusverkkoakin. [44]

4.1.2 Välistäveto

Hyökkääjä voi olla asettautunut vertaisverkon ulkopuolelle salakuuntelemaan *(eavesdrop)* jotakin linjaa tai yhdyskäytävää. Välistävetomenetelmällä *(man-in-the-middle)* hyökkääjä voi seurata kaikkea tietoliikennettä, joka kulkee salakuunneltua linjaa pitkin. Esimerkiksi kuvassa 4.1 hyökkääjä 1 olisi voinut asettautua käyttäjän U Internet-palveluntarjoajan reitittimelle kuuntelemaan tietoliikennettä, jossa hän kontrolloisi kaikkea U:n kautta kulkevaa liikennettä. Vastaavasti hyökkääjällä 2 olisi kontrolli vain osaan käyttäjän V liikenteestä, jos V:llä olisi useita fyysisesti erillisiä yhteyksiä vertaisverkkoon tai hyökkääjä olisi asettautunut topologisesti liian kauaksi V:stä.

Käytännössä hyökkääjän voi olla kryptografisten suojausten vuoksi erittäin vaikeaa toteuttaa reaaliaikaista tai riittävän pienen viiveen välistävetohyökkäystä [38]. Suuremmalla viiveellä välistävetohyökkäys on kuitenkin todellinen uhka salattujen vertaisverkkojen



Kuva 4.1: Hyökkääjät 1 ja 2 salakuuntelemassa käyttäjien U ja V tietoliikennettä.

käyttäjien yksityisyyttä kohtaan. Yksi keino suojautua myös täydelliseltä salakuuntelulta on generoida koneellisesti ylimääräistä ja turhaa häiriöliikennettä todellisen tietoliikenteen sekaan [9, 27]. Jos voidaan osoittaa, että osa tietoliikenteestä on pelkkää häiriötä, minkä tahansa tietoliikenteen osan osoittaminen käyttäjältä lähtöisin olevaksi tai käyttäjälle tarkoitetuksi voi parhaassa tapauksessa olla mahdotonta.

4.1.3 Tietoliikenneanalyysi

Tietoliikenneanalyysiä voidaan käyttää yhdessä soluttautumisen tai välistävetomenetelmän kanssa. Siinä yritetään yleensä tilastollisesti ratkaista jonkin käyttäjän tai käyttäjäryhmän osallisuus tutkittavaan tietoliikennekaappaukseen.

Hyökkääjän kontrolloidessa kohteensa välitöntä naapuria ja päästessä käsiksi kohdetta palvelevan osapuolen lokitietoihin, hyökkääjä voi käyttää esimerkiksi ajoitusanalyysiä *(timing analysis)* murtamaan kohteensa yksityisyyden suojan [22]. Sekoitusverkossa hyökkääjä voi muutamalla verkkoon solutetulla solmulla saada aikaan kuvan 2.9 tapaisia sekoitusreittejä, joissa hänellä on sekä ensimmäisen että viimeisen reitittimen kontrolli. Ajoitusanalyysillä eli kohteen tietoliikenteen aikaleimojen ja viimeisen reitittimen lokitietojen vertailulla hyökkääjä voi tilastollisesti murtaa kohteensa yksityisyyden. Esimerkiksi Torvertaisverkossa hyökkääjä voi valehdella omien solmujensa ominaisuuksia ja kasvattaa todennäköisyyttä sellaisten reittien muodostumisille, joissa hän kontrolloi sekä ensimmäistä että viimeistä reititintä [7]. Vastaavasti tiedostojen levitykseen tarkoitetuissa salatuissa vertaisverkoissa hyökkääjä voisi ajoitusanalyysillä selvittää välittömien naapureidensa osallisuuden jakamiensa tiedostojen lataamisessa.

Kuten välistävetomenetelmää myös tietoliikenneanalyysiä vastaan toimii häiriöliikenteen lisääminen. Ajoitushyökkäyksien tehokkuutta voidaan lisäksi huomattavasti vähentää asettamalla reititykseen erilaisia viiveitä. Viivästyttämällä ja sekoittamalla reititettävien pakettien järjestystä verkon solmut voivat merkittävästi vaikeuttaa ajoitusanalyysin tekemistä. [9, 22, 27]

5 Yhteenveto

Salatut vertaisverkot turvaavat käyttäjiensä yksityisyyttä erityisesti siihen tarkoitukseen kehitetyillä reititysalgoritmeilla ja vahvoilla kryptografisilla menetelmillä. Samat yksityisyyden suojaamismenetelmät eivät kuitenkaan sovellu kaikkiin käyttötarkoituksiin, ja tämän vuoksi eri toteutusten välillä saattaa olla suuriakin eroavaisuuksia. Tässä tutkielmassa esiteltiin salattujen vertaisverkkojen kolme keskeisintä toisistaan eroavaa käyttötarkoitusta: tiedostojen anonyymi levittäminen, tiedon sensuurinkestävä säilyttäminen ja tietoliikenteen jäljittämätön reititys. Esittelyssä tarkasteltiin erityisesti reitityksen toteuttamisen sekä järjestelmän luotettavuuden piirteitä nojautumalla uusimpiin aihetta koskeviin teorioihin ja toteutuksiin.

Taulukkoon 5.1 on listattu tässä tutkielmassa käsiteltyjen lähestymistapojen keskeisimmät ominaisuudet. Kaikkia lähestymistapoja yhdistää se, että vertaisverkon suorituskyky laskee helposti, kun käyttäjien yksityisyyttä ja anonymiteettiä yritetään varmistaa entistä voimakkaammin. Vertaisverkolle tulee helposti lisää yleiskustannuksia, kun yritetään vähentää yhä monimutkaisempien hyökkäysten vaikutus käyttäjien yksityisyyden suojaan ja järjestelmään kokonaisuutena; toisaalta tehokkuutta parantavat uudet menetelmät saattavat avata hyökkääjille uusia tietoturva-aukkoja.

Nykyisillä toteutuksilla voidaan varmistaa melko voimakas käyttäjien yksityisyyden suoja. Käyttäjämäärien kasvu salatuissa vertaisverkoissa tuo käyttäjilleen yhä enemmän turvaa, sillä mitä enemmän vertaisverkossa on käyttäjiä, sitä suurempaan joukkoon yksittäiset käyttäjät voivat sulautua. On kuitenkin suuri arvoitus miksi salattujen vertaisverkkojen käyttäjämäärät ovat suhteellisen pieniä verrattuna salaamattomiin vastineisiinsa. Sananvapautta rajoittava ennakkosensuuri sekä erinäisten tahojen suorittama Internetin käyttäjien yksityisyyttä loukkaava valvonta lisääntyy eri puolilla maailmaa [1, 3, 23, 34], mutta sekään ei ole vielä nähtävästi vaikuttanut salattujen vertaisverkkojen suosioon [44]. Ohjelmistojen ja tiedostojen laittoman levityksen voidaan olettaa vaikuttavan suurelta osin salattujen ja salaamattomien vertaisverkkojen keskinäiseen suhteeseen. Jos piratismin vastaista taistelua entisestään tehostettaisiin, voitaisiin olettaa, että suuri osa salaamattomien vertaisverkkojen käyttäjistä siirtyisi salattuihin vastineisiin.

Nykyisillä toteutuksilla voidaan nähdä olevan myös muita heikkouksia. Useimmat niistä tukeutuvat suorituskyvyllisistä syistä pääosin kryptografisiin suojausmenetelmiin, joka on osaltaan huolestuttavaa, sillä useissa nykyisissä toteutuksissa viimeiset voimakkaat varokeinot, kuten häiriöliikenne, on jätetty kokonaan pois. Luvussa 4 esiteltyjen turvallisuusriskien perusteella voidaankin sanoa, että jos hyökkääjällä on riittävän suuret resurssit käytettävis-

Taulukko 5.1: Reititysmenetelmien ominaisuuksia salattujen vertaisverkkojen toteutuksissa

Menetelmä	**Keskeiset ominaisuudet**
Tulvinta	Verkon kuormitus on rajoittamattomana suuri → rajoitettava
	Käyttäjät näkevät vain pienen osan suuresta verkosta
	Massiiviset reititystaulut
	Soveltuvuus: Tiedonhaku ja -siirto järjestäytymättömässä ympäristössä
Ahne reititys	Verkon kuormitus skaalautuu tulvintaa paremmin
	Tehoton jos ympäristön järjestäytyneisyyttä ei tunneta
	Massiiviset reititystaulut
	Tehokas reititys ei välttämättä ole tarkka
	Soveltuvuus: Tiedonhaku ja -siirto järjestäytyneessä ympäristössä
Sekoitusverkko	Suoraviivainen reititys
	Vanhoja yhteyksiä ei tarvitse muistaa
	Voidaan tarjota pienempi viive
	Ei tarjoa hakutoimintoa
	Soveltuvuus: Anonyymi reititys kolmannen osapuolen järjestelmään

sään, voidaan joidenkin toteutusten osalta käyttäjien yksityisyys murtaa triviaalisti. [9, 22]

Lisäksi useimmissa tapauksissa soluttautumisen kautta tapahtuva salakuuntelu sekä järjestelmän muu väärinkäyttö ovat täysin tiedossa olevia ongelmia. Salattuja vertaisverkkoja voidaan helposti käyttää esimerkiksi haittaohjelmien levitykseen tai jäljittämättömään roskapostin lähettämiseen. On kuitenkin perusteltua olettaa, että ajamalla salatut vertaisverkot alas, ei näiltä ongelmilta voitaisi välttyä, vaan niille löydettäisiin välittömästi muita levityskanavia, koska nämä ongelmat ovat iältään salattuja vertaisverkkoja vanhempia ja siten salatuista vertaisverkoista riippumattomia [44]. Pikemminkin voitaisiin kysyä kuinka paljon nämä ongelmat itsessään ovat inspiroineet ja osaltaan siivittäneet salattujen vertaisverkkojen kehitystä.

Viimeisimpänä, mutta ei niinkään merkityksettömimpänä, salattujen vertaisverkkojen heikkoutena voitaisiin mainita niiden käytettävyys ja erityisesti käyttöön ottaminen. Nykyiset toteutukset eivät vielä tarjoa samanlaista vaivatonta käyttöä ja käyttöön ottamista kuin salaamattomat vastineensa ja tulevaisuudessa kannattaa mielestäni keskittyä näiden osa-alueiden kehittämiseen. Vähintään yhtä tärkeää on selvittää voidaanko salattujen vertaisverkkojen tarjoamaa yksityisyyden suojaa entisestään parantaa kasvattamatta merkittävästi yleiskustannuksia ja mitkä ovat salattujen vertaisverkkojen merkittävimmät turvallisuusuhat.

Mielestäni tutkielma antaa vielä aiheen esittää yhden yhteiskunnallisesti merkittävän kysymyksen: onko yksityisyys menettänyt arvonsa digitaalisella aikakaudella?

Lähteet

[1] *HE 48/2008 vp, Hallituksen esitys eduskunnalle sähköisen viestinnän tietosuojalain ja eräiden siihen liittyvien lakien muuttamisesta*, saatavilla Internetistä: <URL: http://www.eduskunta.fi/valtiopaivaasiat/he+48/2008>.

[2] *Introducing I2P*, saatavilla Internetistä: <URL: http://www.i2p2.de/techintro.html>, viitattu 10.1.2009.

[3] *L 1.12.2006/1068, Laki lapsipornografian levittämisen estotoimista*, saatavilla Internetistä: <URL: http://www.finlex.fi/fi/laki/ajantasa/2006/20061068>.

[4] Roberto Aringhieri, Ernesto Damiani, Sabine De Capitani Di Vimercati, Stefano Paraboschi, ja Pierangelo Samarati. *Fuzzy Techniques for Trust and Reputation Management in Anonymous Peer-to-Peer Systems*. Journal of the American Society for Information Science and Technology, 57(4):528–537, tammikuu 2006.

[5] Vijay Atluri, toim. *CCS '02: Proceedings of the 9th ACM Conference on Computer and Communications Security*, New York, NY, USA, 2002. Association for Computing Machinery.

[6] John Perry Barlow. *A Declaration of the Independence of Cyberspace*, saatavilla Internetistä: <URL: http://homes.eff.org/~barlow/Declaration-Final.html>, 8.2.1996.

[7] Kevin Bauer, Damon McCoy, Dirk Grunwald, Tadayoshi Kohno, ja Douglas Sicker. *Low-Resource Routing Attacks Against Anonymous Systems*. Tekninen raportti, University of Colorado, Boulder, CO, USA, 2007.

[8] David L. Chaum. *Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms*. Communications of the ACM, 24(2):84–90, helmikuu 1981.

[9] Tom Chothia ja Konstantinos Chatzikokolakis. *A Survey of Anonymous Peer-to-Peer File-Sharing*. Kirjassa *Proceedings of the IFIP International Symposium on Network-Centric Ubiquitous Systems (NCUS 2005)*, sarjan *Lecture Notes on Computer Science* osa 3823, s. 744–755, Heidelberg, Germany, 2005. Springer-Verlag.

[10] Ian Clarke. *Freenet's Next Generation Routing Protocol*, saatavilla Internetistä: <URL: http://freenetproject.org/ngrouting.html>, 20.7.2003.

[11] Ian Clarke. *A Distributed Decentralised Information Storage and Retrieval System*. Tekninen raportti, University of Edinburgh, 1999, saatavilla Internetistä: <URL: http://freenetproject.org/papers/ddisrs.pdf>.

[12] Ian Clarke, Scott G. Miller, Theodore W. Hong, Oskar Sandberg, ja Brandon Wiley. *Protecting Free Expression Online with Freenet*. IEEE Internet Computing, 6(1):40–49, 2002.

[13] Ernesto Damiani, De Capitani di Vimercati, Stefano Paraboschi, Pierangela Samarati, ja Fabio Violante. *A Reputation-Based Approach for Choosing Reliable Resources in Peer-to-Peer Networks*. Ks. Atluri [5], s. 207–216.

[14] Roger Dingledine, Michael J. Freedman, David Hopwood, ja David Molnar. *A Reputation System to Increase MIX-net Reliability*. Kirjassa Ira S. Moskowitz, toim., *Proceedings of the 2001 Information Hiding Workshop*, sarjassa *Lecture Notes in Computer Science*, Lecture Notes in Computer Science, s. 126–141, Heidelberg, Germany, huhtikuu 2001. Springer Verlag.

[15] Roger Dingledine, Michael J. Freedman, ja David Molnar. *Accountability*. Ks. Oram [36], 16. luku, s. 271–340.

[16] Roger Dingledine ja Nick Mathewson. *Tor Protocol Specification*, saatavilla Internetistä: <URL: https://svn.torproject.org/svn/tor/trunk/doc/spec/tor-spec.txt>, viitattu 24.12.2008.

[17] Roger Dingledine, Nick Mathewson, ja Paul Syverson. *Tor: The Second-Generation Onion Router*. Kirjassa *Proceedings of the 13th USENIX Security Symposium*, s. 303–320, Berkeley, CA, USA, elokuu 2004. The USENIX Association.

[18] Michal Feldman ja John Chuang. *Overcoming Free-Riding Behavior in Peer-to-Peer Systems*. ACM SIGecom Exchanges, 5(4):41–50, heinäkuu 2005.

[19] Michal Feldman, Kevin Lai, Ion Stoica, ja John Chuang. *Robust Incentive Techniques for Peer-to-Peer Networks*. Kirjassa *EC'04: Proceedings of the 5th ACM conference on Electronic commerce*, s. 102–111, New York, NY, USA, toukokuu 2004. Association for Computing Machinery.

[20] Michal Feldman, Christos Papadimitriou, John Chuang, ja Ion Stoica. *Free-Riding and Whitewashing in Peer-to-Peer Systems*. IEEE Journal on Selected Areas in Communications, 24(5):1010–1019, toukokuu 2006.

[21] The Gnutella Developer Forum. *The Annotated Gnutella Protocol Specification v0.4*, saatavilla Internetistä: <URL: http://rfc-gnutella.sourceforge.net/developer/stable/index.html>, viitattu 26.12.2008.

[22] Michael J. Freedman ja Robert Morris. *Tarzan: a Peer-to-Peer Anonymizing Network Layer*. Ks. Atluri [5], s. 193–206.

[23] Reporters Sans Frontières. *China*, saatavilla Internetistä: <URL: http://www.rsf.org/article.php3?id_article=26134>, viitattu 4.2.2009.

[24] Christian Grothoff, Ioana Patrascu, Krista Bennett, Tiberiu Stef, ja Tzvetan Horozov. *GNET*, saatavilla Internetistä: <URL: http://www.gnunet.org/download/main.pdf>, viitattu 24.12.2008.

[25] Vikas Gupta, Avnish Dass, Harpreet Singh Matharu, Ankur Verma, ja Yashraj Chauhan. *Peer-to-Peer Application Development: Cracking the Code*. Hungry Minds, Inc., New York, NY, USA, 2002.

[26] Theodore Hong. *Performance*. Ks. Oram [36], 14. luku, s. 203–241.

[27] Yoosuk Jung, Juyung Seo, Kyungsuk Lhee, ja Manpyo Hong. *Unobservable Mix: Hiding Communication with Uniform Shape of Network Traffic*. Kirjassa *Proceedings of the International Conference on Convergence Information Technology*, s. 2386–2393, marraskuu 2007.

[28] Gene Kan. *Gnutella*. Ks. Oram [36], 8. luku, s. 94–122.

[29] Kevin Lai, Michal Feldman, Ion Stoica, ja John Chuang. *Incentives for Cooperation in Peer-to-Peer Networks*. Kirjassa *Proceedings (online) of the First Workshop on Economics of Peer-to-Peer Systems*, UC Berkeley, Berkeley, CA, USA, kesäkuu 2003. saatavilla Internetistä: <URL: http://www.sims.berkeley.edu/p2pecon/papers/s1-lai.pdf>.

[30] Adam Langley. *Freenet*. Ks. Oram [36], 9. luku, s. 123–132.

[31] Adam Langley. *Mixmaster Remailers*. Ks. Oram [36], 7. luku, s. 89–93.

[32] Sergio Marti ja Hector Garcia-Molina. *Taxonomy of Trust: Categorizing P2P Reputation Systems*. Computer Networks, 50(4):472–484, joulukuu 2005.

[33] Nelson Minar ja Marc Hedlund. *A Network of Peers: Peer-to-Peer Models Through the History of the Internet*. Ks. Oram [36], 1. luku, s. 3–20.

[34] Matti Nikki. *Suodatuslista*, saatavilla Internetistä: <URL: http://www.lapsiporno.info/suodatuslista/>, viitattu 4.2.2009.

[35] Kieron O'Hara, Harith Alani, Yannis Kalfoglou, ja Nigel Shadbolt. *Trust Strategies for the Semantic Web*. Kirjassa *ISWC'04: Proceedings of the 3rd International Workshop on Trust, Security, and Reputation on the Semantic Web*, marraskuu 2004.

[36] Andy Oram, toim. *Peer-to-Peer: Harnessing the Benefits of a Disruptive Technology*. O'Reilly & Associates, Inc., Sebastopol, CA, USA, ensimmäinen laitos, maaliskuu 2001.

[37] Jean-François Raymond. *Traffic Analysis: Protocols, Attacks, Design Issues and Open Problems*. Kirjassa *Proceedings of International Workshop on Design Issues in Anonymity and Unobservability*, sarjan *Lecture Notes on Computer Science* osa 2009, s. 10–29, Heidelberg, Germany, 2001. Springer-Verlag.

[38] Jason Rohrer. *How File Sharing Reveals Your Identity*, saatavilla Internetistä: <URL: http://mute-net.sourceforge.net/howPrivacy.shtml>, viitattu 23.12.2008.

[39] Jason Rohrer. *Scalable Real-Time Search in P2P Networks*, saatavilla Internetistä: <URL: http://mute-net.sourceforge.net/utilityCounters.shtml>, viitattu 26.12.2008.

[40] Oskar Sandberg. *Distributed Routing in Small-World Networks*. Kirjassa Rajeev Raman, Robert Sedgewick, ja Matthias F. Stallmann, toim., *2006 Proceedings of the Ninth Workshop on Algorithm Engineering and Experiments (ALENEX)*, s. 144–155, Philadelphia, PA, USA, 2006. Society for Industrial and Applied Mathematics.

[41] Oskar Sandberg. *Decentralized Search with Random Costs*. ArXiv e-prints, huhtikuu 2008, saatavilla Internetistä: <URL: http://arxiv.org/pdf/0804.0577v1>.

[42] Paul F. Syverson, David M. Goldschlag, ja Michael G. Reed. *Anonymous Connections and Onion Routing*. Kirjassa *Proceedings of the 1997 IEEE Symposium on Security and Privacy*, s. 44–54, Los Alamitos, CA, USA, toukokuu 1997. IEEE Computer Society.

[43] Yao Wang ja Julita Vassileva. *Bayesian Network-Based Trust Model*. Kirjassa *Proceedings of the IEEE/WIC International Conference on Web Intelligence 2003*, s. 372–378, Los Alamitos, CA, USA, lokakuu 2003. IEEE Computer Society.

[44] The Tor Project Webmasters. *Tor: Overview*, saatavilla Internetistä: <URL: https://www.torproject.org/overview.html.en>, 13.6.2008.