



Kandidatutkielma

Tietojenkäsittelytieteen kandiohjelma

# **Vertaisverkkojen aiheuttamat ongelmat palveluntarjoajille**

Eero Ojala

16.3.2020

MATEMAATTIS-LUONNONTIETEELLINEN TIEDEKUNTA  
HELSINGIN YLIOPISTO

**Ohjaaja(t)**

S. Hätönen

**Tarkastaja(t)**

Prof. K. Heljanko

**Yhteystiedot**

PL 68 (Pietari Kalmin katu 5)  
00014 Helsingin yliopisto

Sähköpostiosoite: [info@cs.helsinki.fi](mailto:info@cs.helsinki.fi)

URL: <http://www.cs.helsinki.fi/>

|  |                               |  |  |
|--|-------------------------------|--|--|
| Tiedekunta — Fakultet — Faculty  |                               | Koulutusohjelma — Utbildningsprogram — Study programme |  |
| Matemaattis-luonnontieteellinen tiedekunta   |                               | Tietojenkäsittelytieteen kandiohjelma                  |  |
| Tekijä — Författare — Author   |                               |  |  |
| Eero Ojala   |                               |  |  |
| Työn nimi — Arbetets titel — Title   |                               |  |  |
| Vertaisverkkojen aiheuttamat ongelmat palveluntarjoajille  |                               |  |  |
| Ohjaajat — Handledare — Supervisors  |                               |  |  |
| S. Hätönen   |                               |  |  |
| Työn laji — Arbetets art — Level   | Aika — Datum — Month and year | Sivumäärä — Sidoantal — Number of pages                |  |
| Kandidutkielma   | 16.3.2020                     | 14 sivua   |  |
| Tiivistelmä — Referat — Abstract   |                               |  |  |
| <p>Vertaisverkot tarjoavat erilaisia hyötyjä sisällönjakajille ja käyttäjille. Nämä hyödyt kuitenkin tapahtuvat palveluntarjoajien kustannuksella, jotka kärsivät vertaisverkkojen aiheuttamasta liikenteenhallinnan hankaloittamisesta sekä liiallisesta ulkoliikenteestä. Näiden ongelmien välttämiseksi palveluntarjoajat ovat kehittäneet keinoja vertaisverkkoliikenteen estämiseksi, mutta ne ovat onnistuneet korkeintaan hidastamaan sitä.</p> <p>Ongelmien ratkaisuksi on ehdotettu saatavilla olevan sisällön paikallisuuden huomioon ottaminen, joko vertaisverkkosovellusten omasta toimesta tai sitten yhteistyönä palveluntarjoajien kanssa. Monet tutkimuspaperit ovat esittäneet lupaavilta vaikuttavia ratkaisuehdotuksia, mutta täydellistä ratkaisua ei ole vielä löytynyt. Jos halutaan paikallisuutta hyödyntäviä vertaisverkkosovelluksia otettavan käyttöön, tulee myös huolehtia siitä, että ulkoliikenteen vähentäminen ei kasvata käyttäjien latausaikoja.</p> |                               |  |  |
| <p><b>ACM Computing Classification System (CCS)</b><br/>         Networks → Network types → Overlay and other logical network structures → Peer-to-peer networks</p>   |                               |  |  |
| Avainsanat — Nyckelord — Keywords  |                               |  |  |
| vertaisverkot, palveluntarjoajat, verkkoliikenteen hallinta  |                               |  |  |
| Säilytyspaikka — Förvaringsställe — Where deposited  |                               |  |  |
| Helsingin yliopiston kirjasto  |                               |  |  |
| Muita tietoja — övriga uppgifter — Additional information  |                               |  |  |
| Ohjelmistojärjestelmien erikoistumislinja  |                               |  |  |



# Sisältö

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Johdanto</b>  | <b>1</b>  |
| <b>2</b> | <b>Vertaisverkot ja palveluntarjoajat</b>              | <b>2</b>  |
| 2.1      | Liikenteenhallinnan ongelmat . . . . .                 | 2         |
| 2.2      | Vertaisverkkoliikenteen määrä jaulkoliikenne . . . . . | 3         |
| 2.3      | Vertaisverkkoliikenteen hallitseminen . . . . .        | 4         |
| <b>3</b> | <b>Paikallisuuden hyödyntäminen</b>                    | <b>5</b>  |
| 3.1      | Paikallisuuden potentiaali ja hyödyt . . . . .         | 5         |
| 3.2      | Painotettu vertaisvalinta . . . . .                    | 7         |
| 3.3      | Vertaisverkkosovellusten omat ratkaisut . . . . .      | 8         |
| 3.4      | Yhteistyöratkaisut . . . . .                           | 9         |
| <b>4</b> | <b>Ratkaisuehdotusten ongelmat</b>                     | <b>10</b> |
| <b>5</b> | <b>Yhteenvedo</b>                                      | <b>11</b> |
|          | <b>Kirjallisuus</b>                                    | <b>13</b> |



# 1 Johdanto

Vertaisverkko (*Peer-To-Peer network, P2P*) on hajautettu arkkitehtuuriratkaisu, jossa vertaiset (*peer*), eli käyttäjät, jakavat resurssejaan jonkin yhteisen päämäärän saavuttamiseksi. Vertaisverkkoja voidaan käyttää moneen eri käyttötarkoitukseen. Esimerkiksi suoratoistoon (*streaming*), Internet-puhelimiin, ja tiedostojen jakamiseen, johon tässä paperissa pääasiassa keskitytään. Vertaisverkot tarjoavat tiedostojen jakamisessa monenlaisia etuja verrattuna perinteiseen asiakas-palvelin-malliin. Sisällöntarjoajalle vertaisverkot eivät vaadi suuria ja kalliita infrastruktuurihankintoja kuten sisällönjakeluverkkoja (*Content Distribution Network, CDN*)[12], koska sisällön jakaminen hajautetaan vertaisten kesken. Kun alkuperäinen sisällöntarjoaja on kerran saanut sisällön datan kokonaan jaetua vertaisille, voivat nämä periaatteessa hoitaa kaiken lopun sisällön jakamisesta. Tämä tuo paljon säästöjä sisällöntarjoajalle, kun sen ei enää mahdollisesti tarvitse jatkuvasti lähettää sisältöä usalle samanaikaiselle lataajalle ympäri maailmaa.

Vertaisverkot skaalautuvat hyvin, koska jokainen uusi vertainen tuo verkkoon mukanaan lisää laskentatehoa, tallennustilaa, yms. Vertaisverkkojen kasvattaminen on myöskin helppoa, sillä sisällöntarjoajan ei tarvitse manuaalisesti lisätä uusia vertaisia verkkoon tai hankkia uusia palvelimia kuten asiakas-palvelin-mallissa. Sen sijaan seurantapalvelin tai jokin muu mekanismi pystyy automaattisesti mukauttamaan vertaisverkon kokoonpanoa vertaisten tullessa ja lähtiessä, mikä tekee vertaisverkkojen ylläpidosta hyvin vaivatonta.

Koska vertaisverkoissa sisällön jakamista ei olla keskitetty vain yhdelle palvelimelle, kestää vertaisverkko asiakas-palvelin-mallia paljon paremmin palvelunestohyökkäyksiä ja äkkinäisiä kysynnän nousuja (*flash crowd*). Tämä sietokyky sekä sisällön hajautuminen usean eri tarjoajan kesken parantaa sisällön saatavuutta käyttäjille. Lisäksi sisällön lataaminen samanaikaisesti usealta eri vertaiselta mahdollistaa nopeamman latausnopeuden kuin asiakas-palvelin-mallissa yhdeltä mahdollisesti ruuhkautuneelta palvelimelta lataaminen.

Vertaisverkkojen tuomat edut käyttäjille ja sisällöntarjoajille tapahtuvat kuitenkin palveluntarjoajien (*Internet Service Provider, ISP*) kustannuksella. Vertaisverkot aiheuttavat vakavia ongelmia palveluntarjoajille luomalla paljon verkkoliikennettä ja tekemällä mielivaltaista reititystä. Tämän aineen luvussa 2 tarkastellaan tarkemmin näitä ongelmia, ja luvussa 3 taas niihin esitettyjä ratkaisuja. Lopuksi luvussa 4 käsitellään vielä lyhyesti näiden ratkaisuehdotuksien omia ongelmia, ja luvussa 5 esitetään yhteenveto.

## 2 Vertaisverkot ja palveluntarjoajat

Vertaisverkot aiheuttavat palveluntarjoajille lukuisia ongelmia, mikäli niihin ei erikseen varauduta. Ongelmat johtuvat pääasiassa siitä, että vertaisverkot tekevät omaa mielivaltaista reititystä. Koska vertaisverkkojen reititys tapahtuu vain sovellustasolla, ne eivät pääse käsiksi verkkotason tarjoamaan tarkempaan tietoon verkon ominaisuuksista. Tästä johtuen vertaisverkot joutuvat tekemään omia estimointeja, ja niihin perustuvat reitityspäätökset aiheuttavat mahdollisesti ongelmia palveluntarjoajien verkoille.

Luvussa 2.1 tarkastellaan vertaisverkkojen aiheuttamia ongelmia palveluntarjoajien liikenteenhallinnalle. Luvussa 2.2 taas käsitellään vertaisverkkojen aiheuttamaa palveluntarjoajien verkkojen ulkopuolista liikennettä. Lopuksi luku 2.3 käsittelee palveluntarjoajien keinoja estää näitä ongelmia syntymästä, ja miten vertaisverkkosovellukset ovat näihin keinoihin reagoineet.

### 2.1 Liikenteenhallinnan ongelmat

Vertaisverkot hankaloittavat verkkoliikenteen ennustettavuutta. Perinteisessä asiakas-palvelin-mallissa tiedostot ladataan vain yhdestä lähteestä, jolloin palvelimelle vieville reitteihin kohdistuvaan ruuhkaan on helppo varautua. Vertaisverkoissa latauslähde voi olla kuka tahansa tavallinen asiakas erikseen sisällöntarjontaan dedikoitujen palvelimien sijaan. Lisäksi vertaisverkkojen myötä asiakkailta lähtee runsaasti enemmän lähetysliikennettä (*upload traffic*) verrattuna asiakas-palvelin-mallin mukaiseen sisällöntarjontaan.

Ennustettavuuden lisäksi vertaisverkot aiheuttavat myös monia muita konkreettisia ongelmia, joita käsitellään Ram Keralapura et.alin paperissa [8]. Ensimmäinen näistä ongelmista on kuormantasauskäytäntöjen rikkominen. On mahdollista, että palveluntarjoajat saattavat haluta dedikoida verkossa kahden solmun välisen linkin näille solmuille itselle kuuluvalla liikenteelle. Tällöin liikenteenhallinnassa pyritään ohjaamaan muiden solmujen paketteja kulkemaan jonkin muun reitin kautta. Vertaisverkkojen reitityksessä tätä ei kuitenkaan pystytäkään ottamaan huomioon. Palveluntarjoajan kuormantasausta voidaan sivuuttaa täysin, ja vertaisverkkoliikenne pystytään reitittämään näiden kahden solmun kautta. Tällöin palveluntarjoaja saa vääristyneen käsityksen solmujen välillä kulkevan liikenteen määrästä, jos se ei pysty erottelemaan niiden ja vertaisverkkojen liikenteen välillä.



Palveluntarjoajat ylläpitävät verkoistaan liikennematriisia (*traffic matrix*), jonka alkiot sisältävät jokaisen mahdollisen solmuparin välisen liikenteen määrän. Vertaisverkoissa pakettien lopullinen päämäärä voidaan salata, jolloin julkiseksi päämääräksi määritellään reitin välisolmu. Vaikka palveluntarjoajan verkon sisällä kahden solmun välillä olisi suora linkki, on hyvinkin mahdollista että vertaisverkko päättää reitittää paketit kiertämään yhden tai useamman muun solmun kautta pienemmän viiveen tai jonkin muun syyn vuoksi. Koska palveluntarjoaja ei näe salattua lopullista päämäärää vaan jokaisen välietapin erikseen, liikennematriisin sisältämä tieto vääristyy. Liikenne jakautuu välietappisolmujen välille, ja alkuperäisen lähde- ja kohdesolmun alkioon ei tapahdu muutoksia. Tästä johtuen verkkomatriisia joudutaan usein korjailemaan tiedon oikeellisuuden ylläpitämiseksi.

Jos vertaisverkkoliikenne ruuhkauttaa jonkin yksittäisen solmujen välisen linkin erittäin pahasti, saattaa tilanne laukaista palveluntarjoajien verkon reittien kokoonpanoon muutoksia verkkotasolla. Vertaisverkot sitten huomaavat nämä muutokset, ja mukauttavat reititystään näiden mukaisesti. Tämä sitten saattaa johtaa uudelleen pahaan ruuhkautumistilanteeseen, joka taas laukaisee palveluntarjoajan tekemät muutokset, synnyttäen reaktiosyklin näiden välille. Toisenlainen syklinen tilanne voi syntyä, mikäli palveluntarjoajan verkkoa käyttää kaksi tai useampi vertaisverkko. Näiden vertaisverkkojen välille saattaa syntyä rinnakkaisjärjestelmille tyypillinen kilpailutilanne (*race condition*), jos ne reagoivat palveluntarjoajan verkoissa tapahtuviin muutoksiin samanaikaisesti. On mahdollista että vertaisverkot päätyvät samanaikaisesti käyttämään samaa parhaimmalta vaikuttavaa linkkiä ennen kuin ne ehtivät huomata toistensa tuottaman liikenteen vaikutuksia siihen. Tällöin valittu linkki ruuhkautuu, ja tilanne saattaa toistua niin kauan kunnes jokin verkoista onnistuu reagoimaan muita nopeammin.

## 2.2 Vertaisverkkoliikenteen määrä ja ulkoliikenne

Yllämainitut verkkoliikenneongelmat eivät koske pelkästään vertaisverkkoja, vaan kaikkia päällysverkkosovelluksia (*overlay network*). Nimensä mukaisesti päällysverkkosovellukset muodostavat oman päällystävän verkkonsa jonkin olemassaolevan alusverkon päälle. Esimerkiksi BitTorrent-protokollaa käyttävät vertaisverkkosovellukset muodostavat tiedostoa lataavien vertaisten välillä sovellustason verkon, joka toimii alusverkkona toimivan Internetin päällä verkkotasolla.

Muista päällysverkkosovelluksista vertaisverkot erottautuvat niiden tuottaman verkkoliikenteen määrän vuoksi. 2000-luvulla vertaisverkot olivat kuuma aihe, sillä ne aiheutti-

vat merkittävän osan kaikesta verkkoliikenteestä[6]. Vertaisverkkojen osuus liikenteestä on kuitenkin hiipunut 2010-luvulla esimerkiksi Netflixin ja Youtuben suosion nousun myötä. Tästä huolimatta BitTorrent on yhä dominoiva lähetyслиikenteen lähde, erityisesti Amerikan ulkopuolella[16].

Tämä vertaisverkkoliikenteen suuri määrä on palveluntarjoajien kannalta pelottavaa, sillä yllämanittujen liikenteenhallintaongelmien lisäksi vertaisverkot aiheuttavat paljon verkon ulkopuolelta tulevaa ja sinne menevää liikennettä. Tämä ulkoliikenne aiheuttaa palveluntarjoajille lisämaksuja, mikäli ne eivät ole erikseen sopineet vertaissopimuksia muiden palveluntarjoajien kanssa[11].

## 2.3 Vertaisverkkoliikenteen hallitseminen

Palveluntarjoajat hyötyvät vertaisverkoista siten, että ne motivoivat asiakkaita ostamaan parempia verkkoyhteysnopeampien latausaikojen vuoksi. Aiemmin mainitut ongelmat ovat kuitenkin niin merkittäviä että palveluntarjoajat eivät voi katsoa vertaisverkkoja läpi sormien. Ratkaisuna näihin ongelmiin palveluntarjoajat ovat yleensä päätyneet asettamaan lisämaksuja, ja rajoittamaan tai estämään vertaisverkkoliikennettä.

Perinteisesti palveluntarjoajat ovat myyneet Internet-yhteyksiä asiakkailleen kiinteillä kuukausihinnoilla. Tämä johtui siitä, koska 90-luvulla suurin osa latausliikenteestä kohdistui yksittäisiin palvelimiin asiakas-palvelin mallin mukaisesti. Näitä palvelimia käyttäviä sisällöntarjoajia maksettiin suhteellisesti niiden käyttämän kaistamäärän perusteella, mistä palveluntarjoajat saivat hyvin paljon tuloja. Täten henkilöasiakkaille pystyttiin tarjoamaan halvempia kiinteitä kuukausimaksuja, etenkin kun otaksuttiin että heiltä ei lähtesi juuri ollenkaan lähetyслиikennettä. Kun vertaisverkot alkoivat yleistymään, palveluntarjoajat eivät voineet vanhojen sopimusten perusteella siirtää niiden aiheuttamia ulkoliikennekuluja suoraan asiakkaille. Kiinteiden kuukausimaksujen sijaan jotkin palveluntarjoajat rupesivatkin tarjoamaan sopimuksia, joissa asiakkaat joutuvat maksamaan ylimääräistä mikäli he aiheuttavat ylimääräistä ulkoliikennettä.

Vertaisverkkoliikennettä pystytään havaitsemaan erilaisilla pakettianalysaattoreilla, joita voidaan sijoittaa verkkojen pääsylinkkien kohdalle. Nämä analysaattorit erottelevat vertaisverkkojen lähettämät paketit muun liikenteen joukosta, jolloin vertaisverkkoliikenteeseen pystytään kohdistamaan rajotuksia. Vertaisverkkoliikennettä pystyttiin hallitsemaan myös esimerkiksi rajoittamalla tai estämällä tiettyjen porttien kautta kulkevaa liikennettä, joita vertaisverkkosovellusten tiedettiin suosivan.

## 3 Paikallisuuden hyödyntäminen

Edellisessä luvussa esitetyt ratkaisuehdotukset eivät kuitenkaan ole osoittautuneet kovinkaan toimiviksi. Tiettyjen porttien kautta kulkevan verkkoliikenteen estämisen seurauksena vertaisverkkosovellukset rupesivat käyttämään satunnaisia portteja, jolloin palveluntarjoajat eivät voisi estää vertaisverkkoliikennettä pelkkien porttien perusteella haittaamatta muita sovelluksia. Pakettianalysointoreiden kiertämiseksi vertaisverkkosovellukset taas rupesivat salaamaan pakettejaan. Nämä kierrot johtivat eräänlaiseen “asevarustelukilpaan” palveluntarjoajien ja vertaisverkkojen välille.

Estomenetelmät pystyivät lähinnä vain hidastamaan vertaisverkkojen aiheuttamaa liikennettä, ja ne eivät onnistuneet poistamaan pohjimmaista ongelmaa. Vertaisverkkojen tekemä satunnaisuuteen tai epätäydelliseen estimointiin perustuvat mielivaltainen reititys ja vertaisuussuhteiden asettaminen johtaa usein verkon käytön kannalta tehottomiin tilanteisiin. Näissä tilanteissa vertaisverkkojen käyttäjät lataavat sisällön usein palveluntarjoajien verkkojen ulkopuolelta, vaikka se olisikin saatavilla paikallisesti verkon sisältä muilta vertaisilta. Luvussa 3.1 tarkastellaan tämän paikallisuuden hyödyntämisen potentiaalia sekä hyötyjä, ja luvuissa 3.2-3.4 tarkemmin käsitellään ehdotettuja ratkaisumalleja.

### 3.1 Paikallisuuden potentiaali ja hyödyt

Thomas Kargiannis et. al.[7] tutkivat vertaisverkkojen aiheuttamaa liikennettä ja paikallisuuden hyödyntämisen mahdollistamia säästöjä. He tekivät paperiaan varten kaksi eri tutkimusta. Ensimmäisessä tutkimuksessa tilastoitiin yli 20 000 käyttäjän yliopistoverkon pääsylinkin läpi meneviä BitTorrent-paketteja noin kolmen päivän ajalta. Tilastojen perusteella tutkijat pystyivät laskemaan BitTorrentin aiheuttaman ulkoliikenteen määrän. Toisessa tutkimuksessa taas tarkasteltiin yhden BitTorrent parven (*swarm*, termi yksittäiselle BitTorrent-vertaisverkolle) seurantal palvelimen lokia viiden kuukauden ajalta. Lokista tutkijat saivat selville vertaisten IP-osoitteet sekä liikkuneen datan määrän. Näistä tiedoista he pystyivät määrittämään vertaisten sijainnit eri palveluntarjoajien verkoissa, sekä laskemaan niissä kulkeneen BitTorrent-datan määrän.

Tutkimusten perusteella tutkijat pystyivät laskemaan paikallisuutta puoltavia tuloksia. Päivittäin noin yksi kymmenesosa BitTorrentilla ladatuista tiedostoista ladataan koko-

naan uudelleen verkon ulkopuolelta, vaikka kaikki tarvittava data olisi jo saatavilla verkon sisällä. Paikallisuutta voitaisiin hyödyntää varsin usein, sillä noin 30% ajasta ladattavalle tiedostolle löytyisi vertainen verkon sisältä. Jos kyseisessä verkossa olisi ollut ideaalinen paikallisuusskenaario, jossa kaikki paikallisesti saatavilla oleva data ladattaisiin aktiivisilta käyttäjiltä, niin jokaisesta yksittäisestä tiedostosta olisi tarvinnut lähettää keskimäärin vain yksi kokonainen kopio verkon ulkopuolelle. Myös vertaisverkkojen käyttäjät voisivat mahdollisesti hyötyä tästä ideaaliskenaariosta. Tutkimusten mukaan se hyödyttäisi noin 70% käyttäjistä mahdollistamalla heille nopeammat latausajat vaikuttamatta kuitenkaan negatiivisesti loppuihin 30% käyttäjistä.

Paikallisuuden potentiaalille on sinänsä loogiset syyt. Sisällön kielestä ja maantieteellisistä kiinnostuksen kohteista johtuen ladattaville tiedostoille löytyy usein samoilta alueilta paljon vertaisia. Lisäksi kaikkein suosituimmalle sisällölle löytyy paljon maailmanlaajuista kiinnostusta, kuten esimerkiksi uusille Hollywood-elokuville. Näitä tiedostoja jakelevat vertaisverkot saattavat sisältää tuhansia vertaisia. Tällöin on hyvin todennäköistä, että saman palveluntarjoajan verkon sisältä löytyy useampikin vertainen. Eritoten jos kyseessä on suuremman palveluntarjoajan verkko.

Tutkimuksen esittämät nopeammat latausajat taas selittyvät sillä, että paikallisuuden suosiminen muodostaa vertaisten lähelle lyhyempiä reittejä. Kun reitit ovat lyhyitä, paketit kulkevat keskimäärin harvemmin pääsylinkkien lävitse. Nämä pääsylinkit ovat yleensä Internetin pahimpia pullonkauloja, joten ne hidastavat niiden läpi kulkevaa liikennettä. Pääsylinkkien kautta kulkeva liikenne aiheuttaa siis palveluntarjoajien ylimääräisten maksujen lisäksi myös käyttäjille haittaa hidastamalla heidän latausnopeuksiaan.

Myös vertaisverkkosovellukset voisivat hyötyä paikallisuudesta. Jos vertaisverkkosovellukset hyödyntäisivät paikallisuutta, voisivat ne pitää verkkoliikenteensä mahdollisimman paljon palveluntarjoajien verkkojen sisällä. Tällöin vertaisverkot eivät aiheuttaisi ylimääräistä ulkoliikennettä, jolloin palveluntarjoajilla olisi vähemmän syytä rajoittaa tai estää vertaisverkkoliikennettä. Jos vertaisverkkosovelluksen liikennettä ei rajoiteta, sen käyttäjät eivät myöskään välttämättä vaihda toiseen sovellukseen. Lisäksi ylläolevien tutkimustulosten mukaan paikallisuuden hyödyntäminen mahdollisesti parantaa latausnopeuksia, joka voisi olla sovellukselle hyvä myyntivaltti.

## 3.2 Painotettu vertaisvalinta

Ruchir Bindal et.al. esittävät paperissaan[2] tekniikan nimeltä painotettu vertaisvalinta (*Biased Peer Selection, BPS*), minkä avulla BitTorrent-sovellukset voisivat hyödyntää paikallisuutta. Kyseisestä ratkaisusta on hyvä lähteä liikkeelle, sillä se on melko yksinkertainen ja esittelee hyvin paikallisuuden hyödyntämisen perusideoita. Se on myös yksi ensimmäisistä paikallisuuteen perustuvista ratkaisuista, ja monet myöhemmin ilmestyneet aiheetta käsittelevät paperit ovat viitanneet tähän paperiin.

Painotetussa vertaisvallinnassa BitTorrent-vertaisille pyritään antamaan suurin osa muista vertaisista saman palveluntarjoajan verkon sisältä. Vertaisille kuitenkin annetaan aina pieni määrä vertaisia verkon ulkopuolelta, jotta sisältöä olisi mahdollista ladata mikäli sitä ei löydy verkon sisäpuolelta. Ulkopuoliset vertaiset myös edesauttavat vertaisverkon eheyttä. Jos palveluntarjoajan verkon sisäiset vertaiset olisivat yhteydessä ulkopuolelle vain yhden tai muutaman vertaisen kautta, olisi vertaisverkko hyvin riskialtis jakautumiselle näiden vertaisten poistuttua vertaisverkosta.

BitTorrentissa parven vertaiset pitävät listaa muiden vertaisten omistamista datan osista, joita BitTorrentin sanastossa kutsutaan palasiksi (*piece*). BitTorrent priorisoi sellaisten palasten lataamista, joita on vertaisten keskuudessa kaikista vähiten. Tätä kutsutaan ”harvinaisin pala ensin”(*rarest piece first*)-periaatteeksi[10], ja se soveltuu hyvin painotetun vertaisvallinnan kanssa. Kun palaset saadaan ensimmäistä kertaa verkon sisälle, voidaan ne sitten tehokkaasti levittää muille paikallisille vertaisille.

Tutkijat testasivat painotettua vertaisvalintaa simuloiduilla verkoilla, ja testien tulokset osoittautuivat positiivisiksi. Ratkaisulla pystyntämään hyödyntämään paikallisuutta hyvin, ja palveluntarjoajien verkon ulkoliikenne ei kasva yksittäisen vertaisverkon vertaisten määrän lisääntyessä. Myös käyttäjien latausnopeudet pysyvät ennallaan, kunhan sisällön alkuperäisellä jakajalla on tarpeeksi nopea lähetysyhteys.

Tutkijat ehdottavat kaksi eri implementointitapaa painotetulle vertaisvallinnalle. Vertaisverkkosovelluksia ja/tai niiden käyttämiä seurantapalvelimia voitaisiin muokata, jolloin vertaisverkkosovellukset voisivat huolehtia paikallisuuden hyödyntämisestä itsenäisesti. Toisena vaihtoehtona painotettu vertaisvalinta voitaisiin tehdä käyttämällä hyväksi palveluntarjoajien infrastruktuuria, jolloin palveluntarjoajat ja vertaisverkot voisivat tehdä yhteistyötä ulkoliikenteen välttämiseksi. Näitä eri implementaatiotapoja käsitellään tarkemmin seuraavissa aliluvuissa.

### 3.3 Vertaisverkkosovellusten omat ratkaisut

Vertaisverkkosovellukset voisivat itsenäisesti hyödyntää paikallisuutta, jolloin palveluntarjoajien ei tarvitsisi hankkia uutta kallista infrastruktuuria. Lisäksi vertaisverkkosovellusten käyttäjät eivät välttämättä luottaisi palveluntarjoajien tarjoamaan yhteistyöhön, koska nämä voisivat salaa ilmiantaa tekijänoikeuksia rikkovaa materiaalia lataavia vertaisia. Täten tekemällä paikallisen vertaisvalinnan täysin itse vertaisverkkosovellukset eivät menettäisi mahdollisesti käyttäjiään toisille sovelluksille.

Yksi tällainen sovellusratkaisu on David Choffnessin ja Fabián Bustamanten kehittämä Ono-lisäosa[3] Azureus BitTorrent-sovellukselle. Ono käyttää hyväkseen sisällönjakeluverkkojen tekemiä käyttäjien uudelleenlähetyksiä paikallisille palvelimille. Tutkijoiden mukaan samoille palvelimille lähetettävät vertaiset sijaitsevat todennäköisesti lähemmäs, mahdollisesti jopa saman palveluntarjoajan verkon sisällä. Näiden uudelleenlähetysten pohjalta Ono sitten pystyy tekemään paikallisuutta hyödyntävän vertaisvalinnan käyttäjilleen.

Vertaisvalinnan lisäksi Ono keräsi tietoa käyttäjien välisestä verkonkäytöstä ja tiedonsiirrosta tutkimusta varten. Tammikuuhun 2008 mennessä Onoa oli ladattu 120 000 kertaa 108 eri maasta, ja se keräsi tietoa yli 100 000 000 vertaisesta jotka olivat levittäytyneet yli 10 000 eri autonomisen järjestelmän[5] välille. Kerätyistä tiedoista tutkijat vertailivat ranskalaisen sanomalehden Le Monden sisällönjakoverkkoa käyttäneen Onon sekä BitTorrentin oman metodin valitsemia vertaisia, ja näiden mahdollistamia säästöjä ulkoliikenteeseen sekä vaikutuksia latausnopeuksiin.

Tietojen pohjalta tutkijat saivat selville, että heidän ratkaisunsa skaalautuu hyvin yli 100 000 käyttäjälle, ja vähentää merkittävästi BitTorrentin aiheuttamaa ulkoliikennettä. Noin kolmasosa Onon suosittelemista vertaisista sijaitsee yhden autonomisen järjestelmän sisällä, ja yleisemmin Onon ehdottamien vertaisten välinen mediaani AJ-etäisyys (AJ: Autonominen Järjestelmä) on tasan yksi. Lisäksi ehdotettujen vertaisten välisen latenssin suuruusluokka on kaksi kertaa pienempi ja häviö-aste on kolmasosan pienempi kuin BitTorrentin satunnaisesti valitsemilla vertaisilla, mikä mahdollistaa paremmat latausajat.

Tutkijat kuitenkin huomauttavat, että käytetyllä sisällönjakeluverkolla on mahdollisesti suuri vaikutus tuloksiin. Paperin loppupuolella tutkijat vertailivat kuuden eri sisällönjakeluverkon käyttämää Onon vertaisvalintaa BitTorrentin omaan vertaisvalintaan. Tuloksista näkyi että vaikka pienempiä sisällönjakeluverkkoja käyttämällä saadaan heikompiä tuloksia, mahdollistavat ne silti paremmat latausnopeudet ja ulkoliikennesäästöt kuin BitTorrentin oma vertaisvalinta.

### 3.4 Yhteistyöratkaisut

Yhteistyö palveluntarjoajien ja vertaisverkkosovellusten välillä vaikuttaisi parhaalta ratkaisulta, sillä vertaisverkkosovellusten tekemillä verkon rakenteen estimoinneilla on rajansa. Ne ovat hyvin virhealttiita, ja usein johtavat epäoptimaalisiin ratkaisuihin. Myöskään pelkästään sovellustasolla ei voida estimoida tietoa kuten esimerkiksi palveluntarjoajien liikenteenhallinnan käytäntöjä, joiden tietäminen olisi hyödyllistä reititystä varten.

Yhteistyö antaisi vertaisverkoille pääsyn palveluntarjoajien tietämään tarkempaan tietoon verkon ominaisuuksista. Näiden tietojen avulla vertaisverkkosovellukset voisivat tehdä optimaalisempaa reititystä ja löytää käyttäjille paikallisia vertaisia, mikä voisi mahdollisesti nopeuttaa tiedostojen latausaikoja. Tarkemmat tiedot myös poistaisivat tarpeen estimoinnille ja sen aiheuttamalle ylimääräiselle luotainliikenteelle, mikä hyödyttäisi palveluntarjoajia. Yhteistyön avulla palveluntarjoajat voisivat pitää vertaisverkkojen liikenteen mahdollisimman paljon verkkonsa sisällä esimerkiksi tarjoamalla paikallisia vertaisia, jolloin kallista ulkoliikennettä voitaisiin karsia. Yhteistyö mahdollistaisi myös palveluntarjoajille keinon vaikuttaa vertaisverkkojen liikenteeseen, jolloin niiden aiheuttamia liikenteenhallinnan ongelmia voitaisiin myös välttää ylimääräisen ulkoliikenteen lisäksi.

Tällainen yhteistyö voitaisiin toteuttaa esimerkiksi Vinay Aggarwalin et.alin ehdottaman “oraakkelin”[1] avulla. Oraakkeli on päälysverkkosovelluksille tarkoitettu rajapinta, jolla on pääsy verkkokerroksen tarkkoihin tietoihin. Vertaisverkkosovellukset voisivat käyttää tätä oraakkelia siten, että vertaisverkkoihin yhdyistyvät uudet käyttäjät lähettäisivät sille listan potentiaalisista vertaisehdokkaista. Oraakkeli voisi sitten tietojensa pohjalta asettaa nämä vertaiset paremmuusjärjestykseen, ja palauttaa ne takaisin käyttäjille. Tämän listan perusteella sovellukset voivat siten valita paikallisempia vertaisia, jotka voisivat mahdollistaa nopeammat latausajat ja vähentää ulkoliikenteen määrää.

Tutkijat testasivat oraakkelin vaikutusta vertaisverkkoihin erilaisilla simulaatioilla sekä käyttämällä pientä rajattua testiympäristöä. Yksinkertaistetun päälysverkon lisäksi tutkijat käyttivät Gnutellaa[14] esimerkkinä vertaisverkoista. Hyödyntämällä oraakkelin vertaisehdokkaita Gnutellaan syntyi paikallisia ryppäitä kuitenkin haittaamatta vertaisverkon kokonaisuheyttä. Vertaisten välinen keskimääräinen AJ-etäisyys oli noin yhden luokkaa. Luku usein pieneni alle yhteen, mikäli vertaisillaan oli käytössä tuhannen vertaisen välimuisti. Paikallisryppäiden ansiosta myös vertaisten keskinäisestä kommunikoinnista aiheutuvat yleiskustannusviestit (*overhead*) vähenivät roimasti.

## 4 Ratkaisuehdotusten ongelmat

Paikallisuuden hyödyntäminen ei ole kuitenkaan maaginen kaiken ratkaiseva keino, vaan sillä on myös omat ongelmansa. Monet ratkaisuehdotukset saattavat kuulostaa teoriassa hyvältä, mutta käytännössä osoittautua tehottomaksi. Menetelmien testaaminen Internetissä suuremmalla skaalalla on kieltämättä hankalaa, joten usein joudutaan turvautumaan simulaatioihin tai rajattuihin testiympäristöihin. Michael Piatek et.al vertasivat paperissaan[13] Onon suorituskykyä ja sen aiheuttamaa ulkoliikennettä muokkaamattomaan BitTorrenttiin. Tutkimustaan varten tutkijat liittyivät 32:een suosittuun ja tuoreeseen BitTorrent-parveen. Näissä parvissa jaetuista sisällöistä ladattiin ensimmäiset 30 megatavua Azureus-sovelluksella käyttäen eri vertaisvalintakeinoja, mukaan lukien Ono-lisäosan sisällönjakeluverkkoihin perustuvaa ja BitTorrentin omaa satunnaista vertaisvalintaa.

Tutkimusten tulosten mukaan Onoa käyttävässä ja käyttämättömässä sovelluksissa ei ollut juurikaan eroja. Onoa käytettäessä latausajat olivat noin 1,02 kertaa hitaampia kuin normaalin BitTorrentin. Myös Onon ja BitTorrentin valitsemien vertaisten välinen mediaani AJ-etäisyydet olivat vastaavasti 3,99 ja 4,02. Aiemmin viitatus Onon esittelevän tutkimuspaperin esittämät väitteet valittujen vertaisten nopeammista latausajoista ja lyhyimmistä etäisyyksistä pitivät sinänsä paikkansa tämänkin paperin mukaan. Ongelmana kuitenkin oli, että Ono pystyi suosittulemaan vain hyvin pienen määrän paikallisista vertaisista, vaikka kyseiset parvet olivat tuoreita ja suosittuja.

Tutkijat esittävät, että Onon tyylliset täysin käyttäjäsovellus-pohjaiset ratkaisut eivät välttämättä ole riittäviä ulkoliikenteen vähentämiseksi, sillä niiltä puuttuu globaali kokonaiskuva samanaikaisista lataajista. Parempi ratkaisu paikallisuuden määrittämiseksi olisi globaalimmalla tasolla, esimerkiksi seurantapalvelimilla. Lisäksi BitTorrent itsessään ei välttämättä ole ylipäättänsä soveltuva paikallisuuden hyödyntämiseen, mikä johtuu pääasiallisesti kahdesta syystä.

BitTorrentille tyypillisessä sisällönjakelun elinkaareissa tiedostoja jaetaan alussa ahkeasti, mutta kannustinmekanismien puutteesta johtuen vertaiset eivät yleensä jää jakamaan sisältöä muille vertaisille saatuaan omat latauksensa valmiiksi. Pikkuhiljaa tapahtuvassa BitTorrent-parvien pienenemisen yhteydessä samalla myös paikallisuuden hyödyntämisen potentiaali laskee. Toinen syy on, että verkkojen siirtonopeuksien eriävyyksien vuoksi BitTorrenttien käyttäjien kokonaiskapasiteetti on jakautunut hyvin epätasaisesti. Suu-



rin osa kapasiteetista tulee pieneltä määrältä käyttäjiltä, jotka eivät ole tasaisesti jakaantuneet ympäri maailmaa, vaan pääasiassa sijaitsevat tietyillä alueilla. Paikallisuuden hyödyntämisessä nämä korkean kapasiteetin käyttäjät rypästyvät omille alueilleen, jolloin muiden aluiden latausnopeudet kärsivät.

Myös palveluntarjoajien rajapinnoilla yms. yhteistyöratkaisuilla on potentiaalisesti omat ongelmansa. Ne eivät saa kertoa verkon ominaisuuksista tai vertaisista liian riskialtista tietoa, jota pahantahtoiset tekijät voisivat hyväksikäyttää. Rajapintapalvelut vaatisivat myös verkkojen infrastruktuurin muutoksia, joten rajapintapalveluiden toteuttaminen voisi olla hyvin hankalaa ja kallista. Palveluntarjoaja voisi myös näitä menetelmiä käyttämällä manipuloida vertaisverkkoliikennettä omien etujensa mukaisesti. Esimerkiksi jos kahden vertaisverkon välillä olisi kaupallinen sopimus, voisivat nämä ohjata vertaisverkkoliikennettä toisilleen rahallisen hyödyn vuoksi nopeampien vaihtoehtojen sijasta.

## 5 Yhteenveto

Vertaisverkot tuovat paljon hyötyjä sisällönjakajille ja käyttäjille, kuten esimerkiksi halvempi ylläpito, parempi saatavuus sekä nopeammat latausajat. Nämä edut kuitenkin tapahtuvat palveluntarjoajien kustannuksella. Vertaisverkot aiheuttavat paljon vaikeasti enustettavaa verkkoliikennettä, jota palveluntarjoajat eivät pysty hallitsemaan, haittaen näiden liikenteenhallintaa. Lisäksi vertaisverkkojen tekemän vertaisvalinnan vuoksi ne saavat aikaan paljon ulkoliikennettä, tuoden palveluntarjoajille ylimääräisiä kuluja.

Näiden ongelmien torjumiseksi palveluntarjoajat ovat ottaneet käyttöön eri keinoja vertaisverkkoliikenteen torjumiseksi tai rajoittamiseksi. Vertaisverkkosovellukset taas vastavasti ovat kehittäneet omat keinonsa näiden estojen välttämiseksi. Tämä on johtanut eräänlaiseen asevarustelukilpaan palveluntarjoajien ja vertaisverkkosovellusten välille.

Vertaisverkkojen aiheuttama runsas ulkoliikenne olisi kuitenkin mahdollisesti vältettävissä, sillä tutkimusten mukaan vertaisverkoissa jaettava sisältö olisi usein saatavilla paikallisesti saman verkon sisällä. Palveluntarjoajien lisäksi myös vertaisverkot itse voisivat hyötyä paikallisuuden hyödyntämisestä, sillä ulkoliikenteen vähentäminen poistaisi ison syyn palveluntarjoajien tekemeille estoille. Käyttäjien olisi myös mahdollisesti nopeampi ladata sisältö paikallisemmilta vertaisilta kuin palveluntarjoajan verkon ulkopuolelta.

Paikallisuuden hyödyntämiselle on esitetty erilaisia ratkaisuehdotuksia. Pääasiassa ne jakautuvat kahteen eri kastiin: vertaisverkkosovellusten tekemiin estimointeihin, ja palveluntarjoajien tarjoamiin verkon informaation rajapintoihin. Molemmilla tavoilla on omat hyötynsä, ja näitä ratkaisuja ehdottavat paperit kertovat lupaavia testituloksia.

Täydellistä ratkaisua ongelmiin ei ole kuitenkaan vielä löytynyt. Mitään ratkaisuehdotusta ei olla otettu laajamittaisesti käyttöön, ja monet simulaatioilla tai rajatuilla testiympäristöillä testatut ratkaisuehdotukset ovat osoittautuneet käytännössä tehottomiksi. Vertaisverkot kuten BitTorrent eivät välttämättä edes ylipäättänsä sovellu paikallisuuden hyödyntämiseen, johtuen mm. vertaisverkkojen vertaisten ajan mittaisesta pienemisestä ja eri alueiden tiedonsiirtonopeuksien välisistä eroista. Vaikka paikallisuutta hyödyntämällä onnistuttaisiin vähentämään ulkoliikennettä, ei tällä olisi kuitenkaan paljoakaan merkitystä mikäli vertaisverkkosovellusten latausnopeukset kärsisivät. Ratkaisujen tulisi siis myös taata vähintään vertaisverkkosovellusten latausnopeuksien ennallaan pysymisen, mikäli käyttäjiä halutaan rohkaista siirtymään paikallisuutta hyödyntävien sovellusten pariin. Muuten palveluntarjoajien ja vertaisverkkosovellusten välinen kilpailu kiihtyy entisestään, ja ongelmat eivät ole lähdössä minnekään.

# Kirjallisuus

- [1] V. Aggarwal, A. Feldmann ja C. Scheideler. "Can ISPs and P2P users cooperate for improved performance?" *ACM SIGCOMM Computer Communication Review* 37.3 (2007), s. 29–40.
- [2] R. Bindal, P. Cao, W. Chan, J. Medved, G. Suwala, T. Bates ja A. Zhang. "Improving traffic locality in BitTorrent via biased neighbor selection". Teoksessa: *26th IEEE International Conference on Distributed Computing Systems (ICDCS'06)*. IEEE. 2006.
- [3] D. R. Choffnes ja F. E. Bustamante. "Taming the torrent: a practical approach to reducing cross-isp traffic in peer-to-peer systems". *ACM SIGCOMM Computer Communication Review* 38.4 (2008), s. 363–374.
- [4] V. K. Gurbani, V. Hilt, I. Rimac, M. Tomsu ja E. Marocco. "A survey of research on the application-layer traffic optimization problem and the need for layer cooperation". *IEEE Communications Magazine* 47.8 (2009), s. 107–112.
- [5] J. Hawkinson ja T. Bates. *Guidelines for creation, selection, and registration of an Autonomous System (AS)*. 1996.
- [6] T. Karagiannis, A. Broido, N. Brownlee, K. C. Claffy ja M. Faloutsos. "Is P2P dying or just hiding?" Teoksessa: *IEEE Global Telecommunications Conference, 2004. GLOBECOM'04*. Vol. 3. IEEE. 2004, s. 1532–1538.
- [7] T. Karagiannis, P. Rodriguez ja K. Papagiannaki. "Should Internet Service Providers Fear Peer-Assisted Content Distribution?" Teoksessa: *Proceedings of the 5th ACM SIGCOMM Conference on Internet Measurement*. IMC '05. Berkeley, CA: USENIX Association, 2005, s. 63–76.
- [8] R. Keralapura, N. Taft, C.-N. Chuah ja G. Iannaccone. "Can ISPs take the heat from overlay networks". Teoksessa: *ACM SIGCOMM Workshop on Hot Topics in Networks (HotNets)*. Citeseer. 2004.
- [9] S. Le Blond, A. Legout ja W. Dabbous. "Pushing bittorrent locality to the limit". *Computer Networks* 55.3 (2011), s. 541–557.
- [10] A. Legout, G. Urvoy-Keller ja P. Michiardi. "Rarest first and choke algorithms are enough". Teoksessa: *Proceedings of the 6th ACM SIGCOMM conference on Internet measurement*. 2006, s. 203–216.

- [11] W. B. Norton. "The evolution of the US Internet peering ecosystem". *Equinix white papers* (2004).
- [12] G. Peng. "CDN: Content Distribution Network". *ArXiv* cs.NI/0411069 (2004).
- [13] M. Piatek, H. V. Madhyastha, J. P. John, A. Krishnamurthy ja T. E. Anderson. "Pitfalls for ISP-friendly P2P design." Teoksessa: *HotNets*. 2009.
- [14] M. Ripeanu. "Peer-to-peer architecture case study: Gnutella network". Teoksessa: *Proceedings first international conference on peer-to-peer computing*. IEEE. 2001.
- [15] R. Rodrigues ja P. Druschel. "Peer-to-peer systems". *Communications of the ACM* 53.10 (2010), s. 72–82.
- [16] Sandvine. *Global Internet Phenomena Report - October 2018*. Tekninen raportti. 2018.
- [17] H. Xie, Y. R. Yang, A. Krishnamurthy, Y. G. Liu ja A. Silberschatz. "P4P: Provider portal for applications". *ACM SIGCOMM Computer Communication Review* 38.4 (2008), s. 351–362.