

Miika Reinikainen

**INFORMAATION DIFFUUSION MERKITYS  
VAIKUTTAVUUDEN MUODOSTUMISEEN  
TWITTERISSÄ**



JYVÄSKYLÄN YLIOPISTO  
TIETOJENKÄSITTELYTIETEIDEN LAITOS  
2016

## TIIVISTELMÄ

Reinikainen, Miika

Informaation diffuusion merkitys vaikuttavuuden muodostumiseen Twitterissä

Jyväskylä: Jyväskylän yliopisto, 2016, 61 s.

Tietojärjestelmätiede, pro gradu-tutkielma

Ohjaaja: Veijalainen, Jari

Sisällön jakamisesta ja käyttäjien välisestä vuorovaikutuksesta internetissä on tullut helppoa ja nopeaa sosiaaliselle medialle tyypillisten yhteisöpalveluiden avulla. Nämä yhteisöpalvelut pohjautuvat rakenteeltaan tietoverkkoihin heijastellen samalla arkielämän ihmissuhteita. Tietyillä yksilöillä on koettu olevan erityisen merkityksellinen rooli sosiaalisissa verkostoissa tiedon leviämisen eli informaation diffuusion kannalta. Tällaisilla yksilöillä on havaittu olevan sellaisia ominaisuuksia, jotka tekevät heistä erityisiä muihin sosiaalisen verkoston yksilöihin verrattuna. Näitä erityisiä yksilöitä kutsutaan yleisemmin vaikuttajiksi.

Tässä tutkimuksessa käsitellään vaikuttavuuden muodostumista yhteisöpalvelu Twitterissä. Tarkoituksena on selvittää miten Twitterissä yksittäisten käyttäjien vaikuttavuus muodostuu, kun tarkastellaan informaation diffuusiota sekä sen taustalla olevaa sosiaalista verkostoa. Tutkimus toteutettiin kirjallisuuskatsauksen sekä tapaustutkimuksen avulla. Tutkimusta varten Twitteristä kerättiin huhtikuussa 2013 Bostonin maratonin pommi-iskuihin liittyvä aineisto. Aineisto sisälsi yli kahdeksan miljoonaa tviittiä ja retviittiä, jotka oli lähetetty yli neljän miljoonan Twitter-käyttäjän toimesta tapahtumaan liittyen.

Tutkimuksen tuloksena selvisi, että Twitterissä vaikuttavuuden muodostumisen kannalta informaation diffuusiolla on merkityksellinen rooli. Seuraajat luovat yleisön käyttäjän lähettämille tviiteille eli tilapäivityksille. Twitterissä informaation diffuusio syntyy, kun käyttäjän seuraajat retviittaavat eli lähettävät käyttäjän lähettämiä tviittejä eteenpäin omille seuraajilleen. Mikäli seuraajia on paljon tai käyttäjän seuraajat ovat itse vaikuttajia, voidaan myös laajamittaisen informaation diffuusion syntyminen nähdä ennustettavana tapahtumana. Mikäli yksi käyttäjä saa tällaisia ketjureaktioita aikaiseksi, voidaan hänellä todeta olevan enemmän vaikuttavuutta Twitterissä kuin tämän vertaiskäyttäjillä.

Asiasanat: Twitter, sosiaalinen verkosto, vaikuttavuus, informaation diffuusio

## ABSTRACT

Reinikainen, Miika

The effect of information diffusion in the formation of the user influence in Twitter

Jyväskylä: University of Jyväskylä, 2016, 61 p.

Information Systems, Master's Thesis

Supervisor: Veijalainen, Jari

Sharing content and communicating in the Internet with other users has become easy and fast because of social media networking services. The structures of these social networking services are based on the computer networks and cloud computing, at the same time reflecting daily life relationships. Certain individuals seems to play a crucial role in the information diffusion in social networks. Such individuals seems to have some special characteristics that make them special in comparison to other individuals in social networks in terms of information diffusion. These special individuals are being generally called influentials.

This study addresses the emergence of the user influence in the Twitter social network service. The purpose of this study is to find out how the influence of individual user is being formed when observing the information diffusion as well as the social network on Twitter. This study was carried out by using a literature review and a case study. The data set for this study was collected from Twitter in April 2013 after the Boston Marathon bombing and related with it. The data consists over eight million tweets that were created or retweeted by over four million different users.

As a result of this research it became clear that the information diffusion has its significant role to play in terms of user influence formation in Twitter. The followers are the main audience for the tweets sent by the user. When the followers are retweeting the tweets to their own followers this activity generates a phenomenon called diffusion of information. If the user has a lot of followers or the followers are influentials themselves, then the diffusion of information more widely can be seen as a predictable event. If one single Twitter user can cause such chain reactions, then the user has more influence than his peers.

Keywords: Twitter, social network, influence, information diffusion

## KUVIOT

KUVIO 1 Esimerkki #- ja @-merkintöjen hyödyntämisestä Twitterissä .....	16
KUVIO 2 Sosiaalisten verkostojen tyypit.....	18
KUVIO 3 Verkoston solmujen egotyypit .....	20
KUVIO 4 Sosiaaliset verkostot, jotka yhdistyvät yhteisen tuttavien kautta .....	22
KUVIO 5 Retviittipuun muodostuminen .....	43
KUVIO 6 Havaittujen retviittipuiden syvyyksien osuus aineistossa .....	46
KUVIO 7 Vaikuttavuuden muodostumiseen osatekijät .....	53

## TAULUKOT

TAULUKKO 1 Vaikuttavuus-käsitteen operationalisointi vaikuttavuutta mittaavaksi indikaattoriksi.....	34
TAULUKKO 2 Käyttäjien vaikuttavuus seuraajien perusteella .....	48
TAULUKKO 3 Käyttäjien vaikuttavuus retviittausten perusteella.....	49
TAULUKKO 4 Käyttäjien vaikuttavuus retviittipuiden arvon mukaan.....	50

# SISÄLLYS

TIIVISTELMÄ.....	2
ABSTRACT .....	3
KUVIOT .....	4
TAULUKOT .....	4
SISÄLLYS.....	5
1 JOHDANTO .....	7
1.1 Tausta.....	7
1.2 Tutkimuksen motivaatio .....	9
1.3 Tutkimuksen tavoitteet ja rajaus .....	10
1.3.1 Tutkimusongelma.....	10
1.3.2 Tutkimusmenetelmä.....	11
1.3.3 Tutkimuksen rajaus .....	11
1.3.4 Tutkielman rakenne.....	11
2 INFORMAATION DIFFUUSIO SOSIAALISISSA VERKOSTOISSA .....	13
2.1 Twitter.....	13
2.1.1 Twitterin perustoiminnallisuudet .....	14
2.1.2 Tviittaaminen ja retviittaaminen .....	15
2.2 Informaatio .....	16
2.2.1 Informaation diffuusio .....	17
2.3 Sosiaaliset verkostot.....	17
2.3.1 Käyttäjien väliset sosiaaliset suhteet .....	19
2.3.2 Käyttäjien väliset seuraajasuhteet verkostossa .....	20
2.3.3 Heikot ja vahvat siteet .....	21
2.3.4 Sosiaalisen verkoston vaikutus informaation diffuusion.....	22
2.4 Informaation diffuusio Twitterissä .....	25
3 VAIKUTTAJAT SOSIAALISISSA VERKOSTOISSA.....	28
3.1 Vaikuttavuuden muodostuminen .....	28
3.1.1 Vaikuttajat Twitterissä .....	29
3.2 Vaikuttavuuden mittaaminen.....	31
3.2.1 Vaikuttavuuden operationalisointi .....	32
3.3 Yhteenveto.....	34
4 EMPIIRINEN TUTKIMUS .....	37
4.1 Tutkimuksen tavoite.....	37
4.2 Tutkimustapa .....	37
4.3 Tutkimusprosessi.....	38

4.4	Tutkimusaineiston kerääminen.....	39
5	TUTKIMUKSEN TULOKSET .....	41
5.1	Taustatietoa tutkimusaineistosta.....	41
5.2	Käyttäjän vaikuttavuuden mittaaminen.....	42
5.3	Tviittianalyysi.....	44
5.4	Vaikuttavuusanalyysi.....	47
5.5	Johtopäätökset.....	50
6	YHTEENVETO.....	54
	LÄHTEET .....	56

# 1 JOHDANTO

## 1.1 Tausta

Joukkoviestintä on kokenut eräänlaista murrosvaihetta uusien viestintäkanavien myötä. Erilaisten medioiden käyttäjiä houkuttelevat yhä enenevässä määrin erilaiset sosiaalisen median palvelut, joiden rakenne pohjautuu tietoverkkoihin, niiden samalla heijastellessa internet-yhteydettömiä sosiaalisia rakenteita, kuten esimerkiksi perhe- ja ystävyys-suhteita. (Romero & Kleinberg, 2010). Näin ollen käyttäjien keskinäinen vuorovaikutus on vahvistunut erilaisten sosiaalisen median palveluiden vaikutuksesta tarjoten yksilöille mahdollisuuden olla helpommin vuorovaikutuksessa muihin käyttäjiin. Tällainen kehitys on hämärtänyt myös joukkoviestintän merkitystä tiedon jakamisen näkökulmasta. (Wu, Hoffman, Mason & Watts, 2011.)

Sosiaalinen media voidaan määritellä joukoksi internet-pohjaisia sovelluksia, joissa Web 2.0 toimii ideologisenä ja teknologisenä perustana ja jotka mahdollistavat käyttäjälähtöisen sisältöjen tuottamisen sekä jakamisen (Kaplan & Haenlein, 2010). Sosiaaliselle medialle tyypillisiä verkkopalveluita ovat esimerkiksi erilaiset sisällönjako- ja yhteisöpalvelut, joita erilaiset sosiaalisen verkoston sivustot (social networking sites) edustavat. Tällaisten palveluiden kautta sisällön tuottamisesta sekä jakamisesta on tullut helppoa ja nopeaa, jolloin myös yksilön näkökulmasta kynnys olla tässä prosessissa mukana on madaltunut.

Yhdysvalloissa toteutetun tutkimuksen mukaan mukaan 65% kaikista täysi-ikäisistä yhdysvaltalaisista käyttää yhteisöpalveluita, kuten Facebookia, Twitteriä ja YouTubea. Kasvu on ollut merkittävää, kun vuonna 2005 vastaava luku oli vain 7%. (Perrin, 2015.) Sosiaalisesta mediasta onkin tullut ihmisille osa jokapäiväistä elämää. Vuorovaikuttamisen tarve muiden käyttäjien kanssa voidaan katsoa yhdeksi merkittävimmistä syistä käyttää yhteisöpalveluita. Samalla näitä palveluita hyödynnetään sisältöjen tuottamiseen, jakamiseen, etsimiseen sekä kuluttamiseen. (Bolton ym, 2013.)

Tuotettujen sisältöjen määrä yhteisöpalveluissa on valtava, kun esimerkiksi Facebookissa käyttäjien on todettu jakavan yli 2,5 miljardia sisältöä (esimerkiksi

kuvia ja tilapäivityksiä) muiden käyttäjien nähtävillä päivittäin (Feitelson, Frachtenberg & Beck, 2013). Vastaavasti Twitterissä on ilmoitettu olevan kuukaudessa 310 miljoonaa aktiivista käyttäjää, jotka jakavat ja lukevat muiden lähettämiä tviittejä (tweets) eli käyttäjän palveluun lähettämiä tilapäivityksiä. (Twitter, 2016). Videopalvelu YouTubeen puolestaan ladataan ajallisesti yli 400 tuntia sisältöä minuutissa palvelun käyttäjien toimesta (Statista, 2016).

Käyttäjien jakaessa yhteisöpalveluissa erilaisia sisältöjä he hyödyntävät teknologian kautta sosiaalista verkostoaan (social network). Sosiaalista verkostoa voidaan tarkastella 3-tasoisien mallinnushierarkian avulla, joka kuvaa yhteisöpalveluissa olevien virtuaaliyhteisöjen (virtual communities) keskinäistä suhdetta, niiden kehitystä ja ominaisuuksia, sekä niiden esitysmuotoa tietojärjestelmissä. Mallinnushierarkian ensimmäisellä tasolla kuvataan ihmisille näkyvää sosiaalista todellisuutta, jossa ihmiset muodostavat virtuaaliyhteisöjä eli eräänlaisia sosiaalisia ryhmiä, jotka hyödyntävät informaatioteknologiaa keskinäiseen vuorovaikutukseen. Toinen taso kuvaa ensimmäisen tason käsitteitä sekä vuorovaikutussuhteita tyypillisesti tietynlaisen verkoston eli graafin avulla, joka koostuu joukosta solmuja (nodes) ja niitä yhdistäviä kaaria (edges). Kolmannella tasolla kuvataan niitä pysyviä tietovarastoja, jotka tyypillisesti ovat tietokantoja, joihin toisen tason tiedot tallentuvat. Jokainen sosiaalisen verkoston sivusto kuvaa virtuaaliyhteisöt eri tavoin kiinnittäen niiden sosiaalisen maailman digitaaliseen esitysmuotoon. Semenov ja Veijalainen kutsuvat tällaista mallia sivusto-ontologiaksi (site ontology), jossa tyypillisesti esiintyvät erityiset käsitteet kuten profiili (profile), ystävä (friend) ja seuraa (follows). (Semenov & Veijalainen, 2013)

Diffuusioprosessin kautta tapahtuva tiedon leviäminen sosiaalisessa verkostossa voidaan nähdä innovaation diffuusion kaltaisena tapahtumana, joka on sosiaalisessa järjestelmässä sen jäsenten kesken tapahtuva viestintäprosessi. Tällöin tieto jostain uudeksi koetusta ideasta, käytännöstä tai esineestä leviää verkostossa siihen kuuluvien jäsenten välillä tietyn ajan kuluessa tiettyjen kanavien kautta. Verkostoon kuuluvien yksilöiden keskinäinen vuorovaikutus voidaan nähdä diffuusioprosessin kannalta avainmekanismina. (Rogers, 2003, s. 5-6, 12; Kwon, 2013.) Twitterissä diffuusioprosessi ilmenee käyttäjän lähettäessä retviitin (retweet) eli jakamalla jonkun toisen käyttäjän lähettämän tviitin eteenpäin omille seuraajilleen. Jatkossa tässä tutkimuksessa puhuttaessa informaation diffuusiosta, tarkoitetaan Twitterin sosiaalisessa verkostossa käyttäjien kesken tapahtuvaa tiedon leviämistä käyttäjien lähettämien tviittien ja edelleen retviittien kautta.

Informaation diffuusion eli jonkin uudeksi koetun tiedon leviämisen kannalta yksilöiden päätöksentekoprosessiin voidaan nähdä vaikuttavan muiden yksilöiden päätösten. Brown ja Feng (2011) toteavat, että tällainen sosiaalinen vaikuttaminen voidaan nähdä eräänlaisena henkilön auktoriteetin omaisena kykynä, jonka avulla henkilö voi vaikuttaa muiden mielipiteisiin ja toimintoihin. Rogers puolestaan toteaa, että sosiaalisessa verkostossa on aina sellaisia erityisiä yksilöitä, joilla on muita yksilöitä enemmän vaikutusta koko verkoston populaatioon. Heidän avullaan on mahdollista suostutella muita sosiaalisessa järjestelmässä olevia, aikaansaaden näin laajamittaisia vaikuttavuuden ketjureaktioita



vähäisin kustannuksin. (Rogers, 2003, s. 26-28.) Esimerkiksi viraalimarkkinoinnin yhteydessä näillä yksilöillä on koettu olevan erityinen rooli suosittelijoina (Leskovec, Adamic & Huberman, 2007). Heillä on myös kyky aiheuttaa myös "sosiaalista epidemiaa" (social epidemic), joka viittaa siihen, että tiettyjen yksilöiden toimesta esimerkiksi informaation diffuusio voimistuu populaation keskuudessa.

Keller ja Berry (2003, s. 2) toteavat tällaisia yksilöitä, jotka vaikuttavat erityisellä tavalla muun väestön mielipiteiden ja trendien muodostumiseen, olevan 10 % väestöstä. Gladwell (2002, s. 33) viittaa myös tähän harvalukuisten yksilöiden joukkoon kutsuen vaikuttajien ilmiötä harvojen laiksi (the law of few). Tällaisia yksilöitä on kutsuttu kaksivaiheisen viestinnän mallissa mielipidejohtajiksi (opinion leaders) (Katz & Lazarsfeld, 1955, s. 32), perinteisen viestintäteorian yhteydessä vaikuttajiksi (influentials) (Rogers, 2003, s. 27, 282), sekä muissa yhteyksissä experteiksi (mavens), yhdistäjiksi (connectors) ja myyjiksi (salesman) (Gladwell, 2002, s. 24). Tässä tutkimuksessa käytetään termiä vaikuttaja tällaisista erityisistä yksilöistä puhuttaessa, koska kirjoittajan mielestä se kuvaa parhaiten näiden erityisten henkilöiden suhdetta muihin sosiaalisen vaikuttavuuden näkökulmasta.

## 1.2 Tutkimuksen motivaatio

Yhteisöpalveluiden käytön myötä periaatteessa jokainen käyttäjä voi olla tasavertainen sisällön tuottamisen sekä jakamisen näkökulmasta. Näin ollen myös vaikuttavuudesta on tullut sattumankauppaa, kun informaation diffuusion kanalta vallitsevilla olosuhteilla koetaan olevan suurempi merkitys kuin yksittäisen yksilön vaikuttavuudella. Tästä johtuen myös vaikuttajien merkitystä on kyseenalaistettu. Vaikuttamisesta ja sen roolista onkin tullut eräs kuumimmista puheenaiheista sosiaalisen median osalta.

Yksi suosituimmista sosiaalisen median palveluista on mikroblogipalvelu Twitter, joka tarjoaa käyttäjälleen mahdollisuuden olla yhteydessä maailmanlaajuisesti muihin palvelun käyttäjiin. Twitterissä käyttäjät voivat luoda enintään 140 merkin mittaisia tilapäivityksiä (status update). Käyttäjä voi halutessaan sisällyttää tviitteihinsä myös videoita, kuvia sekä ulkopuolisiin internet-sivustoihin osoittavia URL-osoitteita, sekä etsiä hakutoiminnon kautta muiden käyttäjien lähettämien tviittien sisältöjä. Tämän tutkimuksen tarkoituksena on selvittää käyttäjien henkilökohtaisen vaikuttavuuden muodostumista Twitterissä informaation diffuusion määrällä mitaten.

## 1.3 Tutkimuksen tavoitteet ja raja

Tässä pro gradu-tutkielmassa on pyritty selvittämään kuinka tieto siirtyy informaation diffuusioprosessin mukaisesti käyttäjältä toiselle yhteisöpalvelu Twitterissä. Tämän lisäksi selvitetään mitä on vaikuttavuus ja miten se muotoutuu informaation diffuusion vaikutuksesta käyttäjäkohtaisesti.

### 1.3.1 Tutkimusongelma

Tässä tutkimuksessa on pyritty selvittämään sitä, mikä merkitys informaation diffuusiolla on vaikuttavuuden syntymisessä Twitter-käyttäjien keskuudessa. Tutkimusongelma on mielenkiintoinen, sillä aikaisempi tutkimus vaikuttamisesta on pohjautunut pitkälti sellaisiin sosiaalisiin verkostoihin, joissa käyttäjien suhde toisiinsa on vastavuoroista ja sen mukaisesti taustalla oleva sosiaalinen verkko on täysin symmetrinen. Esimerkiksi Facebookissa sosiaalinen verkko muodostuu nk. kavereista (friends), jolloin pääsy käyttäjän sosiaaliseen verkostoon vaatii toiselta käyttäjältä kaverikutsun (friendship request) lähettämisen ja kutsun vastaanottajalta sen hyväksymisen. Tällä tavalla muodostuu symmetrinen linkki eli kaksisuuntainen yhteys kahden käyttäjän välille. Twitter poikkeaa monista muista yhteisöpalveluista sen taustalla olevan käyttäjien muodostaman keskinäisen verkoston ollessa asymmetrinen eli epäsymmetrinen. Tämä tarkoittaa sitä, että Twitterissä käyttäjien vastavuoroinen seuraussuhde ei ole oletusarvoista ja kuka tahansa käyttäjä voi seurata toista käyttäjää ilman erillistä kutsua ja hyväksymistä, jos tämän yksityisyysasetukset eivät sitä estä. Mikäli tviitit ovat suojattuja, käyttäjän on hyväksyttävä jokainen seuraajapyyntö erikseen.

Sosiaalisessa verkostoissa on havaittu olevan aina sellaisia erityisiä yksilöitä, jotka pystyvät merkittävästi vaikuttamaan muiden yksilöiden päätöksentekoprosesseihin. Informaation diffuusion kannalta näillä yksilöillä on koettu olevan merkityksellinen rooli, sillä heidän voidaan nähdä olevan keskiössä omassa sosiaalisessa verkostossaan, jolloin heidän kauttaan voidaan aikaansaada laajamittaisia ketjureaktioita. Tällaisilla yksilöillä, joita kutsutaan vaikuttajiksi, on havaittu olevan sellaisia ominaisuuksia, joiden kautta he poikkeavat merkittävästi muista yksilöistä. (Goldenberg, Han, Lehmann & Hong, 2009; Rogers, 2003, s. 26-28.) Vaikuttavuuden muodostumisesta ja sen merkityksestä on erilaisia näkemyksiä, joten tämän tutkimuksen tarkoituksena on tarjota lisää vastauksia vaikuttavuudesta ilmiönä.

Tämän tutkimuksen keskeisin tutkimuskysymys on seuraava:

- *Mikä on informaation diffuusion merkitys käyttäjän vaikuttavuuden muodostumiselle Twitterissä?*

Jotta päätutkimuskysymystä voidaan tarkastella kokonaisuutena, on sen rinnalle asetettu seuraavat avustavat tutkimuskysymykset:

- *Miten sosiaalinen vaikuttavuus syntyy sosiaalisissa verkoissa?*
- *Voidaanko vaikuttavuus Twitterissä yksilöidä käyttäjäkohtaisesti siten, että joku käyttäjä on vaikuttaja ja joku ei?*

### 1.3.2 Tutkimusmenetelmä

Tutkimuksen teoreettisen osuuden tutkimusmenetelmäksi on valittu informaation diffuusioon, vaikuttavuuteen, sosiaalisen median palveluihin ja erityisesti Twitteriin liittyvän tieteellisen kirjallisuuden tutkiminen. Nämä tieteelliset lähteet kirjallisuuskatsausta varten on haettu pääasiassa eri informaatioteknologian, taloustieteiden sekä viestinnän tietokannoista, jossa apuna ovat olleet muun muassa ACM Portal, Business Source Elite, Emerald journals, IEEE Xplore sekä SpringerLink. Lisäksi tietoa on haettu myös käyttäen apuna NELLI-portaalia sekä Google Scholar -hakukonetta.

Tutkimusmenetelmänä tässä tutkimuksessa käytetään kvantitatiivista eli määrällistä tutkimusta. Kvantitatiivisessa tutkimuksessa keskeisiä tekijöitä ovat aikaisemmista tutkimuksista tehtävät johtopäätökset, aiemmat teoriat sekä tutkimuksen kannalta oleellisten käsitteiden määrittely. Tämän lisäksi merkityksellistä on havainnoitavan aineiston keruu, sen määrällinen eli numeerinen mittaaminen, sekä päätelmien tekeminen havaintoaineistosta tilastollisesti analysoimalla. (Hirsjärvi, Remes & Sajavaara, 2007, s. 136.)

Tutkimuksen kannalta merkitykselliset käsitteet määritellään teoreettisesti aikaisempien tutkimuksien osalta. Tämän lisäksi vaikuttavuuden käsite operationalisoidaan, jolloin on mahdollista mitata sen ilmenemistä empiirisesti. Operationaalisten määritelmien tehtävä on metodologinen, jolloin se viittaa mittamiseen eikä termin semanttiseen merkitykseen. (Niiniluoto, 2002, s. 187) Vaikuttavuus operationalisoidaan tämän tutkimuksen kannalta mitattavaan muotoon myöhemmin kohdassa 3.2.1.

### 1.3.3 Tutkimuksen rajaus

Tapaustutkimusta varten kerättiin 15.4.2013 tapahtuneeseen Bostonin maratonin pommi-iskuihin liittyvä aineisto, joka käsittää Twitter-käyttäjien lähettämiä twiittejä ja retviittejä kyseisen aiheen tiimoilta. Tapaustutkimuksen avulla pyritään selvittämään miten käyttäjien aikaansaama informaation diffuusio vaikuttaa yksittäisen käyttäjän vaikuttavuuden muodostumiseen Twitterissä.

### 1.3.4 Tutkielman rakenne

Tutkielma etenee siten, että sen luvussa 2 käsitellään informaation diffuusiota sosiaalisissa verkostoissa sekä siihen liittyviä keskeisiä käsitteitä. Luvussa esitellään tutkimuksen keskiössä oleva Twitter yhteisöpalveluna; miten se näyttäytyy käyttäjälleen, minkälaisia ominaisuuksia kyseinen palvelu pitää sisällään sekä

mitä on tviittaaminen ja retviittaaminen. Samassa luvussa käsitellään myös informaation diffuusiota tarkastellen informaatiota käsitteenä sekä diffuusioprosessiin merkityksellisesti vaikuttavia sosiaalisia verkostoja ominaispiirteineen. Luvun lopuksi havainnollistetaan miten informaation diffuusio näyttäytyy Twitterin kaltaisessa sosiaalisessa verkostossa. Luvussa 3 käsitellään vaikuttavuutta käsitteenä ja ilmiönä, sekä tarkastellaan vaikuttajien roolia sosiaalisissa verkostoissa. Luvussa käydään kirjallisuuskatsaukseen pohjautuen läpi sitä, miten vaikuttavuutta on pyritty aikaisemmin mittaamaan. Luvun lopuksi esitetään yhteenveto kirjallisuuskatsauksen pohjalta tehdyistä havainnoista. Luvussa 4 tarkastellaan tutkimuksen empiirisen osuuden tutkimusmenetelmää sekä sen suunnittelua ja toteutusta. Samassa luvussa esitellään menetelmä, jolla varsinainen tutkimusaineisto kerättiin. Luvussa 5 analysoidaan tutkimusaineistosta havaittuja tutkimustuloksia sekä kerrotaan johtopäätökset näiden tulosten pohjalta. Tutkimuksen luku 6 käsittää yhteenvedon, jossa kerrotaan tutkimuksen pääpiirteet sekä merkityksellisimmät tulokset. Yhteenvedon lopuksi otetaan myös kantaa tutkimuksen rajoitteisiin ja esitellään ehdotukset jatkotutkimusaiheiksi. Tutkimuksen loppuosa kattaa lähdeluettelon.

## 2 INFORMAATION DIFFUUSIO SOSIAALISISSA VERKOSTOISSA

Tässä luvussa tarkastellaan informaation diffuusion ilmenemistä sosiaalisissa verkostoissa, joita erilaiset yhteisöpalvelut edustavat. Aluksi esitellään tämän tutkimuksen kohteena oleva yhteisöpalvelu Twitter yleisesti toiminnallisuuksi-  
neen, minkä jälkeen esitellään tarkemmin tviittaaminen ja retviittaaminen, jotka ovat Twitterissä keskeisiä vuorovaikutuksen muotoja. Tämän jälkeen luvussa käydään läpi tarkemmin mitä ovat informaatio ja informaation diffuusio. Luvun lopuksi esitellään mitä ovat sosiaaliset verkostot, tarkastellaan niiden ominaispiirteitä sekä sitä, miten informaation diffuusio näyttäytyy Twitterissä.

### 2.1 Twitter

Vuonna 2006 julkistettu mikroblogipalvelu Twitter on nykyään yksi suosituimmista yhteisöpalveluista. Twitter ei julkaise tarkkoja käyttäjämääriään, mutta se on ilmoittanut, että palvelulla on aktiivisia käyttäjiä kuukausittain yli 310 miljoonaa (Twitter, 2016.) Twitterissä käyttäjä voi lähettää lyhyitä, enintään 140 merkin mittaisia tekstipohjaisia tviittejä eli tilapäivityksiä muiden Twitter-käyttäjien nähtäville. Mikrobloggaus on viestintämuodoltaan perinteiseen bloggaamiseen verrattuna nopeatempoisempaa. Kannustamalla luomaan lyhyempiä viestisisältöjä alennetaan samalla käyttäjän vaatimuksia esimerkiksi ajallisten resurssien puitteissa sisällön luomiselle. Eräs merkittävä ero on myös viestien tiheydellä; siinä missä perinteisiä aktiivisen käyttäjän blogipäivityksiä saattaa tulla muutamman päivän välein, aktiivinen Twitter-käyttäjä voi lähettää useita lyhyitä tilapäivityksiä päivittäin. (Java, Song, Finin & Belle, 2007.)

Suurin osa Twitterin käyttäjistä voidaan nähdä ennemmin seuraajina kuin sisällöntuottajina eli tviittaajina. Tämä tarkoittaa sitä, että seuraajat pääasiassa vain kuluttavat muiden käyttäjien lähettämiä tilapäivityksiä eivätkä aktiivisesti itse tviittaa. Sisällöntuotannon voidaankin nähdä Twitterissä seuraavan potenssilakia (power-law): aktiivisista käyttäjistä noin 10% tuottaa 90% kaikesta Twitteriin tuotetusta sisällöstä, mikä taas poikkeaa merkittävästi muista sosiaalisen median palveluista, joissa vastaava luku ei ole näin suuri. (Kaplan & Haenlein, 2011.) Twitterin käyttäjille palvelun käyttö on kuitenkin moninaista. Se tarjoaa kanavan olla vuorovaikutuksessa globaalisti sellaisten henkilöiden, organisaatioiden yms. tahojen kanssa mikä muutoin ei välttämättä olisi mahdollista. Käyttäjät kokevat Twitterin kuitenkin ennen kaikkea merkityksellisenä tiedon lähteenä, eräänlaisena uutismedianana eikä niinkään sosiaalisena kohtaamispaikkana. (Kwak, Lee, Park, & Moon, 2010; Zhao & Rosson, 2009.)

### 2.1.1 Twitterin perustoiminnallisuudet

Käyttääkseen Twitteriä, käyttäjän on luotava ensin henkilökohtainen käyttäjäprofiili joko Twitterin internet-sivuston tai mobiilisovelluksen kautta. Profiilin luomiseksi käyttäjän on annettava henkilökohtaisia perustietoja (esimerkiksi nimitiedot sekä sähköpostiosoite), luotava tunnistautumiseen tarvittava salasana sekä käyttäjätunnus, joka toimii samalla myös yksilöivänä tunnisteena käyttäjälle (esimerkiksi @BarackObama).

Twitterissä käyttäjän profiili on oletusarvoisesti julkinen (public), jolloin käyttäjän lähettämät tviitit ovat kaikkien palvelua käyttävien nähtävillä. Käyttäjä voi muokata itse käyttäjäprofiiliaan palvelun profiiliasetuksista siten, että määrittelee käyttäjätilin suojatuksi (protected). Tällöin vain käyttäjän hyväksymät seuraajat (followers) voivat lukea tämän lähettämiä tviittejä. Banerjee (2009) havaitsi, että 90% Twitter-käyttäjien profiileista ja näin ollen myös heidän lähettämänsä tviitit ovat julkisija (Banerjee, 2009). Tässä tutkimuksessa hyödynnetyn aineiston kohdalla kuitenkin havaittiin, että noin 15% käyttäjätileistä oli suojattu (Veijalainen, Semenov & Reinikainen, 2015a). Käyttäjäprofiilin luomisen jälkeen käyttäjä voi aloittaa Twitterin käytön. Halutessaan lukea muiden käyttäjien lähettämiä tviittejä, käyttäjän tulee etsiä ja aloittaa seuraamaan mielenkiintoisiksi kokemiaan muita käyttäjiä. Tällöin seurattavien (followees) käyttäjien tviitit alkavat ilmestyä Twitterin pääsivun aikajanelle (public timeline) käyttäjän nähtäväksi tämän ollessa kirjautuneena palveluun. Käyttäjän on myös mahdollista luoda listoja (list), joihin hän voi ryhmitellä esimerkiksi ystäviään tai tietyn aihepiirin tiimoilta yleensä tviittaavia käyttäjiä. Lista suodattaa tällöin käyttäjän näkyville ainoastaan niiden seurattaviensa tviitit, jotka on määritelty listaan. Käyttäjä voi myös liittyä seuraajaksi jonkin toisen käyttäjän luomaan listaan, mikäli kyseinen lista on määritelty kaikille käyttäjille avoimeksi listaksi. (Twitter Blog, 2009.) Käyttäjä voi itse lähettää tviittejä joko suoraan Twitter-sivuston tai Twitterin mobiilisovelluksen kautta.

Twitterin ideana on, että käyttäjä jakaa sen hetkisiä kuulumisiaan pyrkien vastaamaan kysymyksiin "Mitä teet?" ("What are you doing?") tai "Mitä tapahtuu?" ("What's happening?"). Kysymysten pitäisi innoittaa käyttäjää luomaan sisältöä Twitteriin sekä vastaamaan yksinkertaisesti kysymykseen mitä käyttäjä tekee kyseisellä hetkellä. Twitterin ensivaiheissa kuitenkin havaittiin, että palvelun käyttö on paljon laajempaa ja monipuolisempaa kuin mitä em. yksinkertaisiin kysymyksiin vastaaminen antaa olettaa. Käyttäjien lähettämistä tviiteistä jopa 58,5% ei suoranaisesti vastannut palvelun käyttäjälle valmiiksi ehdottamiin kysymyksiin. (Mischaud, 2007). Twitteristä on myös väitetty, että se on ennemminkin eräänlainen uutismedia, jolloin käyttäjät viestivät senhetkisistä merkityksellisistä tosielämän tapahtumista muille Twitterin käyttäjille, eikä niinkään sosiaalinen kohtaamispaikka verkossa, jossa käyttäjien kahdenvälinen keskustelu ja yhteydenpito on pääosassa. (Kwak ym., 2010).

Rakenteellisesti Twitter eroaa muista sosiaalisen median palveluista siten, että sen muodostama sosiaalinen verkosto on epäsymmetrinen eli asymmetrinen. Tämä tarkoittaa sitä, että toisen käyttäjän tviittien seuraaminen tai itse seuratuksi

tuleminen ei vaadi vastavuoroisuutta. Käyttäjä voi seurata Twitterissä ketä tahansa muuta käyttäjää ja tulla itse seuratuksi muiden käyttäjien toimesta ilman velvoitetta seurata takaisin.

### 2.1.2 Tviittaaminen ja retviittaaminen

Tviitti on käyttäjän luoma enintään 140 merkin mittainen tekstipohjainen tilapäisytytys, joka näkyy lähettämisen jälkeen käyttäjän profiilin aikajanalla. Tviitti voi olla myös käyttäjän esittämä toteamus tai vastaus jollekin toiselle käyttäjälle. Keskustelu Twitterissä tapahtuu käyttäjätunnusta eli @username-syntaksia hyödyntämällä, jolloin viesti osoitetaan suoraan erikseen valitulle käyttäjälle. On myös mahdollista jakaa jonkun toisen käyttäjän luoma tviitti eteenpäin. Tällaista toimintaa kutsutaan retviittaamiseksi. Retviittaaminen voidaan nähdä keinona, jonka avulla Twitteristä on tullut uudenlainen keino jakaa tietoa muiden käyttäjien ja yhä vain suuremman yleisön tietoisuuteen. Retviittaaminen toimii Twitterissä informaation diffuusion kannalta avainmekanismina. (Kwak ym., 2010.)

Käytännössä retviittaaminen tapahtuu lisäämällä käyttäjätunnuksen eteen jokin viestin eteenpäin lähettämistä määrittävä tarkenninlyhenne (esimerkiksi "RT", "via" tai "by"). Retviittaamisen yhteydessä retviittaajan on mahdollista lisätä eteenpäin lähetettävään tviittiin myös sisältöä, kuten omia kommentteja, tunteita tai mielipiteitä. Twitterissä on ominaisuus, niin sanottu yhden klikkauksen retviittaaminen (one-click retweet), joka nopeuttaa retviittamisprosessia ja jonka avulla käyttäjä pystyy retviittaamaan halutun viestin ilman tarkenninlyhenteitä. Näin retviitatessa tviitti näkyy alkuperäisessä muodossaan alkuperäistä lähettäjä myöten. Yhden klikkauksen retviittaaminen ei kuitenkaan tarjoa mahdollisuutta retviittaajan omien kommenttien lisäämiseen alkuperäisen tviitin yhteyteen. Retviittaaminen on eräänlainen sosiaalinen prosessi, jossa käyttäjät levittävät itsensä kannalta kiinnostavaa tietoa omasta tahdostaan. Tätä voidaan kutsua käyttäjälähtöiseksi diffuusiomenetelmäksi. (Kwon & Han, 2013).

Twitterissä on myös muotoutunut tapa käyttää aihetunnistetta eli hashtag-merkintää (#) halutun asianan yhteydessä. Näin kyseinen tviittaus saadaan liitettyä osaksi haluttua aihealuetta. Tämä helpottaa muiden käyttäjien mahdollisuutta seurata keskustelua juuri kyseisen aihealueen tiimoilta. Twitter listaa kymmenen suosituinta senhetkistä keskustelunaihetta Twitter.comin pääsivulle useimmin tviitattujen avainsanojen tai #-merkinnällä tarkennettujen aiheiden mukaan. Kuviossa 1. on havainnollistettu eri merkintätapojen hyödyntämistä Twitterissä.



KUVIO 1 Esimerkki #- ja @-merkintöjen hyödyntämisestä Twitterissä

Tviitteihin voidaan halutessaan lisätä sisältöjä myös muista sosiaalisen median palveluista, kuten esimerkiksi videopalvelu YouTube ja kuvapalvelu Instagramista ilman, että käyttäjän tarvitsee poistua Twitter-palvelusta. Twitterissä on myös mahdollista jakaa ulkopuolisiin sivustoihin osoittavia URL-osoitteita. Yleensä tällaiset osoitelinkit ovat merkkimäärältään pitkiä, ja koska yksittäisen tviittauksen merkkimäärä on rajallinen, hyödynnetäänkin Twitterissä lyhennettyjä URL-osoitteita (esimerkiksi <http://bit.ly>) ohjaamaan halutulle internetsivustolle. (Boyd, Golder & Lotan, 2010.) Esimerkiksi YouTube-sivusto tarjoaa käyttäjälle lyhennetyn URL-osoitteen helpottaakseen sisällönjakamista muissa sosiaalisen median palveluissa. Tämän lisäksi on myös olemassa erillisiä internet-sivustoja, kuten esimerkiksi Tinyurl.com, jossa voi tarvittaessa muuntaa halutun sivuston osoitteen lyhyempään muotoon.

## 2.2 Informaatio

Sosiaalisia verkostoja ja niissä käyttäjien keskinäisiä vuorovaikutuksia tarkasteltaessa huomio kiinnittyy usein käyttäjälähtöisesti luotuihin ja jaettuihin sisältöihin. Sisällöllä tarkoitetaan joukkoa koottua, muokattua tai tiettyyn käyttötarkoitukseen luotua tietoa. Termillä tieto viitataan sen käyttöyhteyden mukaan joko dataan koneen käsiteltävässä muodossa tai informaatioon ihmisen ymmärtämässä muodossa. Ihmisen ymmärtämässä muodossa olevana voidaan pitää ihmislähtöiselle ja -vastaanottajalle merkityksiä antavaa informaatiota. Näin ollen informaation voidaan nähdä tarkoittavan vastaanotetun datan käyttäjälle tuottamia mielteitä tai merkityksiä. (Atk-sanakirja, 2008, s. 42, 84, 326.)

Twitterissä informaatiota sisältävä data voidaan nähdä käyttäjien tuottamina tviitteinä, jotka ovat enimmillään 140 merkin mittaisia tilapäivyyksiä. Tviitit yleensä koostuvat käyttäjän henkilökohtaisesti tuottamasta tekstisisällöstä, kuvista, videoista sekä uutisiin ja artikkeleihin ohjaavista URL-linkeistä. (Romero, Galuba, Asur & Huberman, 2011.) Informaatiota voidaan siirtää esimerkiksi tietokonevälitteisesti ihmiseltä toiselle. Capurron ja Hjørland (2003, s. 369) mukaan Dretske (1986) esittää, ettei tietokoneiden kuitenkaan tarvitse ”ymmärtää” informaatiota samalla tavalla kuin ihminen, mutta ne voivat silti manipuloida sitä. Twitterissä käyttäjän luodessa sisältöjä tviittaamalla, hän luo samalla



merkityksiä antavaa informaatiota muiden käyttäjien, yleensä tämän seuraajien nähtäville.

### 2.2.1 Informaation diffuusio

Tässä tutkimuksessa diffuusio nähdään eräänlaisena viestintäprosessina, joka koostuu neljästä elementistä: innovaatiosta, innovaatioon vaikuttavasta sosiaalisesta järjestelmästä, sosiaalisen järjestelmän sisäisestä viestintäkanavasta, sekä ajasta, jonka aikana diffuusio tapahtuu. Diffuusio on siis erityistä viestintää, joka koskee nimenomaan uusia ideoita. Viestintä puolestaan on erillinen prosessi, jossa siihen osallistuvat yksilöt luovat ja jakavat informaatiota keskenään, päämääränään saavuttaa keskinäinen yhteisymmärrys. Diffuusion määritelmä viittaa siis erityisesti siihen, että viestintä on joko lähentävää tai erottavaa silloin, kun kaksi tai useampi henkilöä vaihtavat tietoa voidakseen lähentyä tai eriytyä antaessaan tapahtumille merkityksiä. Viestinnän voidaan katsoa olevan kaksisuuntainen prosessi, jossa yksilö haluaa siirtää haluamaansa viestiä eteenpäin saavuttaakseen tietynlaisia seurauksia. Näin ollen diffuusio voidaan nähdä sellaisena viestinnän tapana, jossa jaetut viestit koskevat jotain uutta ideaa. Tämän kaltainen idean "uutuus" viestisisällöissä antaa diffuusioprosessille sen erityisen luonteen, joka myös merkitsee sitä, että siihen sisältyy aina jonkinlaista epävarmuutta. (Rogers, 2003, s. 5-6.)

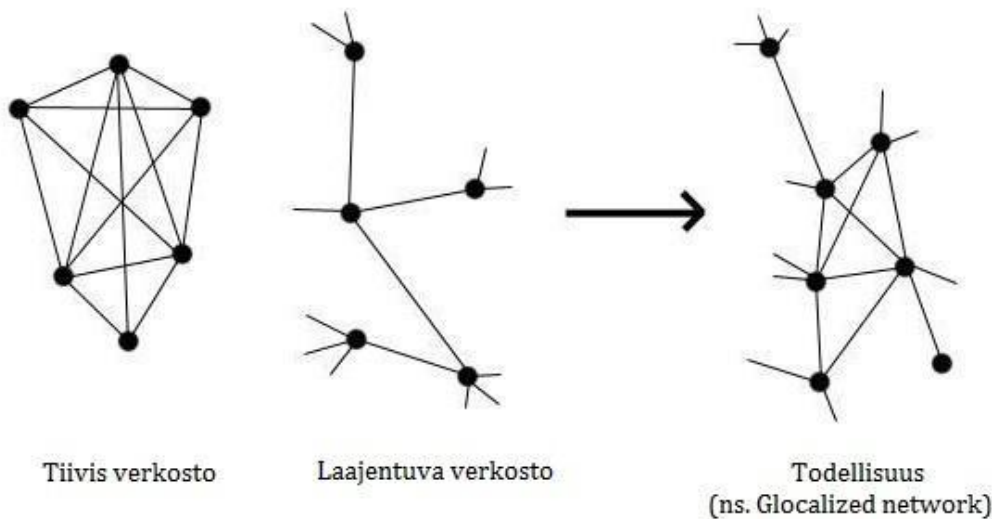
Diffuusioprosessia voidaan tarkastella myös eräänlaisena sosiaalisen muutoksen prosessina, jolloin varsinainen muutos tapahtuu sosiaalisen järjestelmän rakenteissa ja toiminnoissa. Uusien innovaatioiden tai informaatioiden syntymisen ja leviämisen seurauksena tapahtuu omaksumista tai hylkäämistä. Tämä puolestaan johtaa tiettyihin seurauksiin, jolloin myös tapahtuu sosiaalista muutosta järjestelmän sisällä. (Rogers, 2003, s. 6.)

Sosiaalinen järjestelmä, jota myös sosiaaliset verkostot edustavat, ovat joukkoja toisiinsa liitettyjä yksikköjä yhteisen tavoitteen saavuttamiseksi ongelmanratkaisussa. Näin ollen se viittaa diffuusioon sosiaalisen järjestelmän sisällä sen jäsenten kesken, jolloin käyttäjien keskinäisellä vuorovaikutuksella on tässä prosessissa merkityksellinen rooli. (Kwon, 2013.) Sosiaalisen verkoston vaikutus diffuusioprosessiin määritellään tarkemmin luvussa 2.3.4.

## 2.3 Sosiaaliset verkostot

Sosiaalinen verkosto voidaan mallintaa graafina, jossa on joukko yksilöitä kuvaavia solmuja sekä niitä yhdistäviä kaaria, jotka kuvaavat solmujen keskinäisiä vuorovaikutussuhteita. Monimutkaiset sosiaaliset verkostot muodostuvat tällaisten yksinkertaisten rakenteiden kautta muodostaen aina vain isompia ja monimuotoisempia kokonaisuuksia. (Scott, 1991, s. 13,67.) Verkosto voidaan mallintaa rakenteellisesti joko tiiviinä (densely knit group) tai laajenevana (ramified

network). Tiiviin verkoston jäsenet tuntevat toisensa ja ovat toistuvasti vuorovaikutuksessa keskenään, mutta samalla heillä on vain vähän yhteyksiä verkoston ulkopuolelle. Laajenevassa verkostossa osa verkoston jäsenistä on yhteydessä keskenään, mutta suurin osa vuorovaikutuksesta suuntautuu verkoston ulkopuolelle. Sosiaalisia verkostoja tarkasteltaessa verkostot ovat todellisuudessa yleensä näiden kahden sekoitus: sosiaalinen verkosto koostuu tiivistä ryhmistä, kuten perheestä ja työyhteisöstä, mutta samalla on myös paljon yhteyksiä tiiviiden verkostojen ulkopuolella. Tällaisesta verkostosta käytetään nimitystä "glocalized", joka on yhdistelmä sanoista "global" ja "local". (Boase, & Wellman, 2001.) Kuviossa 1 on havainnollistettu nämä eri verkostotyyppit.



KUVIO 2 Sosiaalisten verkostojen tyypit (Boase & Wellmann, 2001)

Verkostoa voidaan siis tarkastella kuin se olisi eräänlainen maailma pienoiskoossa. Milgram (1967) osoitti kuuden askeleen teoriassaan (six degrees of separation), että ainakin periaatteessa jokainen henkilö keskimäärin on kuuden hypäyksen päässä toisistaan. Hän kutsui tätä pienen maailman ilmiöksi (small-world phenomenon). Tällainen ilmiö näyttäytyy myös sosiaalisissa verkostoissa, jotka voivat olla kooltaan valtavia rakennelmia, sisältäen toisiinsa linkkien avulla yhdistettyjä solmujen joukkoja. (Watts & Storgatz, 1998).

Jokainen käyttäjäprofiili Twitterissä on verkostossa oleva yksittäinen solmu, sosiaalisen verkoston muodostuessa vertaiskäyttäjien profiilien seuraussuhteista. Käyttäjien välinen seuraussuhde voi olla vastavuoroista, mutta se ei kuitenkaan ole Twitterissä vaatimuksena. Näin ollen sosiaalisen verkoston voidaan sanoa Twitterissä olevan epäsymmetrinen eli asymmetrinen. Sosiaalisen verkoston osalta Twitter mallintaakin hyvin kapean osan reaalielämän vuorovaikutuksesta, sillä käyttäjä voi seurata sellaisia käyttäjiä, joita ei tosielämässä tunne lainkaan. Twitterin sosiaalisessa verkostossa seuraajina ja seurattavina olevat solmut ovat staattisia rakenteita. Niiden kautta muodostuu potentiaalinen vuorovaikutus,

joka tapahtuu tviittien lähettämisen, vastaanottamisen sekä retviittaamisen kautta.

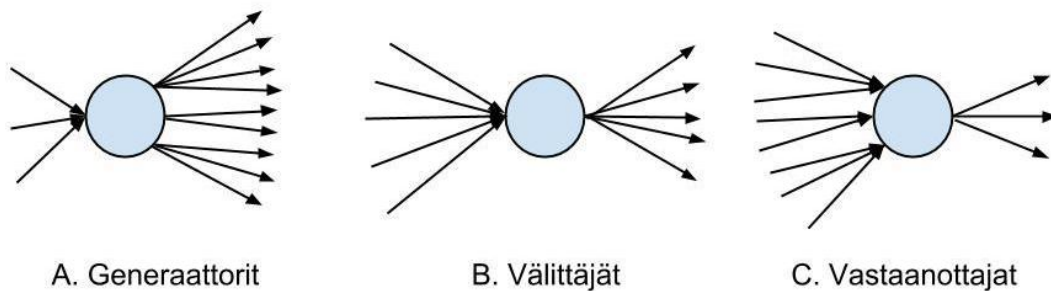
Twitterissä käyttäjät ovat keskimääräisesti vain hieman yli 4 hyppäyksen eli solmun päässä toisistaan, mikä poikkeaa merkityksellisesti Milgramin esittämään pienen maailman teoriaan verrattuna. (Kwak ym, 2010.) Vastaavanlaisia havaintoja on tehty myös Facebookin sosiaalisesta verkosta, jossa käyttäjien on myös havaittu olevan vain noin neljän hyppäyksen päässä toisistaan (Backstrom, Boldi, Rosa, Ugander & Vigna, 2012). Twitterissä käyttäjien keskinäinen seuraajasuhde ei myöskään noudata yleisesti sosiaalisissa verkostoissa toimivaa potenssilakia, koska vastavuoroisuus osapuolten välillä ei ole vaatimuksena. Voi siis olla, että polku, joka johtaa käyttäjästä toiseen, voi vaihdella hyppäyksistä riippuen tai jopa puuttua kokonaan päinvastaiseen suuntaan tarkasteltaessa. (Kwak ym, 2010.)

### 2.3.1 Käyttäjien väliset sosiaaliset suhteet

Keskinäiseen vuorovaikutukseen sosiaalisessa verkostossa vaikuttavat merkittävästi siihen osallistuvien yksilöiden samankaltaisuus sekä heidän keskinäiset kiinnostuksen kohteensa. Ilmiötä, jossa tällaisilla keskenään samankaltaisuuksia omaavilla henkilöillä on pyrkimys keskinäiseen vuorovaikutukseen ja ryhmytyymiseen, kutsutaan homofiliaksi. Mikäli verkoston yksilöiden suhde toisiinsa on pääasiassa homofilinen, kyseinen verkosto voidaan nähdä homogeenisenä. (McPhearson, Smith-Lovin & Cook, 2001.) Homofilia voidaan nähdä kahden keskenään kommunikoivan yksilön samankaltaisuusasteena. Tällainen samankaltaisuus syntyy, kun yksilöt jakavat keskenään tietynlaisia ominaisuuksia, kuten esimerkiksi uskomukset, koulutus, sosioekonominen status, maantieteellinen sijainti ym. Heterofilia, joka on homofilian vastakohta, tarkoittaa, että yksilöiden keskinäiset ominaisuudet eroavat toisistaan, jolloin samankaltaisuutta ei merkittävässä määrin pääse syntymään. Verkoston ollessa pääasiassa sekoittunut, sisältäen paljon heterofilisiä suhteita käyttäjien kesken, kutsutaan sitä tällöin heterogeeniseksi. (Rogers, 2003, s. 19-20.)

Henkilöitä, joiden kanssa on samanlaisia ominaisuuksia, arvioidaan yleensä myönteisesti, jolloin tällaisia yhteyksiä pyritään myös vahvistamaan. Homofilia voidaan nähdä psykologisena sekä rakenteellisena ilmiönä; ihmisellä on psykologinen taipumus luoda suhteita nimenomaan sellaisten henkilöiden kanssa, jotka ovat samankaltaisia hänen kanssaan, mutta samalla tällaisia henkilöitä on mahdollista kohdata aivan jokapäiväisessä elämässä, esimerkiksi perheessä, työpaikalla ja harrastusten parissa. Homofilian voidaankin nähdä muodostuvan yhteisten tuttavuuksien kautta. (Golder & Yardi, 2010; De Choudhury, Sundram, Seligmann & Kellihar, 2010.) Perusajatuksena kuitenkin on, että homofilia synnyttää uusia yhteyksiä, sillä sen voidaan nähdä lisäävän todennäköisyyttä kahden henkilön väliseen keskinäiseen kommunikointiin (McPhearson ym., 2001).

Homogeenisissä verkostoissa huomio kiinnittyy yksilöiden egotyyppeihin ja niiden vaikutukseen viestintäkäyttäytymisessä. Twitterissä voidaan nähdä olevan kolmenlaisia egotyyppisiä, jotka määräytyvät verkoston muista solmuista tiettyyn solmuun suuntautuvista ja tietyistä solmusta verkoston muihin solmuihin lähtevistä linkeistä. Näitä egotyyppisiä kutsutaan generaattoreiksi (generators), välittäjiksi (mediators) ja vastaanottajiksi (receptors). Generaattorit ovat käyttäjiä, jotka seuraavat useita käyttäjiä, mutta heillä itsellään on vähän seuraajia. Vastaanottajilla puolestaan on paljon seuraajia, jolloin he myös ovat tietyllä tavalla huomion keskipisteenä verkostossaan. Esimerkiksi Twitterissä vastaanottajiksi voidaan lukea julkisuuden henkilöt ja joukkoviestimien käyttäjäprofiilit, joilla on paljon seuraajia, mutta vastavuoroista seuraamista he eivät juuri toteuta. Näiden kahden em. egotyypin sekoitus ovat nk. välittäjät, jolloin käyttäjän seurattavien ja seuraajien määrä on suhteessa lähellä toisiaan. (De Choudhury, 2011; Yang & Leskovec, 2010.) Kuviossa 3. on havainnollistettu nämä kolme egotyyppiä, jossa pallokuviot esittävät käyttäjää ja nuolet seuraussuhteen suuntaa. Kun tätä tarkastellaan Twitteristä käsin, niin seuraajaa esittävä nuoli osoittaa kohti palloa ja seurattavaa esittävä nuoli pallosta poispäin.



KUVIO 3 Verkoston solmujen egotyypit (De Choudhury, 2011)

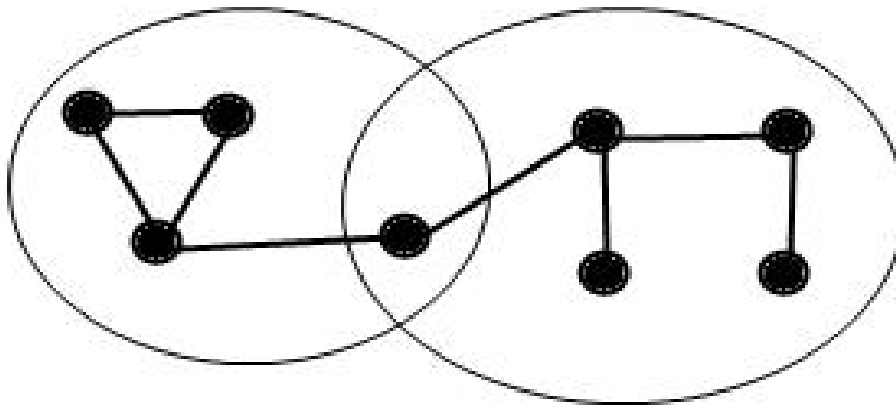
### 2.3.2 Käyttäjien väliset seuraajasuhteet verkostossa

Weng, Lim, Jiang ja He (2010) havaitsivat käyttäjien vastavuoroisen seuraamisen Twitterissä olevan varsin yleistä, kun jopa 72% käyttäjistä on todettu seuraavan yli 80% seuraajistaan. Heidän tutkimuksessa havainnot kuitenkin tehtiin ainoastaan yhden maan käyttäjien keskuudesta, jolloin voidaan päätellä, että esimerkiksi maantieteellinen sijainti ja lähetettyjen tviittien kieli vahvistaisivat tarkasteltavan verkoston homogeenisyyttä. Kwak ym. (2010) havainnot olivat puolestaan päinvastaisia todeten, että vastavuoroisen seuraamisen Twitterissä on verrattain alhaista: lähes 78 % käyttäjäparien seuraussuhteista nähtiin yksisuuntaisina, vain kolmasosa käyttäjistä seurasi ainakin yhtä seuraajaansa ja noin 22% käyttäjien välisistä yhteyksistä oli vastavuoroisia. Toisaalta niiden käyttäjien keskuudessa, joilla seuraussuhde oli vastavuoroista, oli myös havaittavissa jonkinasteista homofiliaa. Esimerkiksi lähetettyjen tviittien aiheilla voidaan nähdä ole-

van yhteys siihen, ketä käyttäjä alkaa seuraamaan, mutta sukupuoli tai esimerkiksi etnisellä taustalla ei vaikuttaisi olevan merkitystä. Käyttäjien keskinäisten seuraussuhteiden ei kuitenkaan voida suoranaisesti sanoa syntyvän homofilian seurauksena. Twitterissä alhainen vastavuoroinen seuraaminen selittäisi osaltaan sitä, että Twitter näyttäytyy käyttäjälleen useimmiten informaation lähteenä kuin sosiaalisena kohtaamispaikkana. (Kwak, yms., 2010; De Choudhury, 2011.) Jotkut käyttäjät myös seuraavat muita käyttäjiä pelkästään etiketin takia; on kohteliasta seurata sellaista käyttäjää, joka seuraa sinua, mikä ei kuitenkaan tarkoita sitä, että tämän tviitit tulisivat luetuiksi (Cha, Haddadi, Benevenuto & Gummadi, 2010).

### 2.3.3 Heikot ja vahvat siteet

Eräs keino tarkastella käyttäjien keskinäistä vuorovaikutusta sosiaalisessa verkostossa yksilötasolla on käyttäjien väliset keskinäiset suhteet. Kahdella toisensa hyvin tuntevilla yksilöllä voidaan katsoa olevan vahva keskinäinen suhde, eräänlainen side, joka ilmenee usein toistuvana, vakaana ja jopa intiiminä vuorovaikutuksena kahden yksilön välillä. Tällainen vahva side (strong tie) ei kuitenkaan välttämättä lisää vaikuttavuutta yksilöiden välillä, vaikka vaikuttavuus kahden toisilleen läheisen henkilön välillä olisikin merkittävää. Vaikutteet tulevat pikemminkin sellaisilta yksilöiltä, joihin on epävirallinen tai jopa satunnainen suhde eli heikko side (weak tie). (Goldenberg, Libai, & Muller, 2001.) Informaatio myös kulkee tehokkaammin sellaisten yksilöiden kautta, joilla on heikko keskinäinen side. Tämä johtuu siitä, että yleensä vahvojen siteiden yhdistämät yksilöt jakavat saman sosiaalisen verkoston eli tarkasteltava verkosto on tällöin homogeeninen. Tällöin verkoston sisällä olevilla yksilöillä tuskin on sellaista informaatiota, jota muilla samassa verkostossa olevilla ei jo valmiiksi ole. Sitä vastoin heikot siteet mahdollistavat yksilön pääsyn sellaisiin yhteyksiin, jotka ovat hänen suppean lähipiirinsä ulkopuolella ja jotka eivät ole sosiaalisten rakenteiden kautta suoraan saavutettavissa. Tällaiset heikkojen siteiden muodostamat sosiaaliset rakenteet syntyvät, kun yksilöiden sosiaaliset verkostot yhdistyvät esimerkiksi yhteisen tuttavien kautta. (Granovetter, 1973.) Kuviossa 4. on havainnollistettu sosiaalisten verkostojen yhdistyminen yhteisen tuttavien.



KUVIO 4 Sosiaaliset verkostot, jotka yhdistyvät yhteisen tuttavien kautta

Kun lähettäjän ja vastaanottajan välinen suhde on homofilinen, on viestintä tällöin myös tehokasta, jolloin homogeenisyyden voidaan nähdä kiihdyttävän informaation diffuusiota verkostossa. Samalla homofilia on myös este informaation laajamittaiselle leviämisen koko tarkasteltavan verkoston laajudelta. Tämä johtuu siitä, että tällöin diffuusio rajoittuu ainoastaan tiiviin käyttäjien keskinäisen verkoston sisällä tapahtuvaksi prosessiksi. Laajamittainen informaation diffuusio verkostossa vaatiikin aina jonkinlaisia heterofilisiä suhteita, jotka toimivat eräänlaisina "siltoina" tiiviisti muodostettujen homogeenisten verkostojen välillä. (Rogers, 2003, s. 19; Granovetter, 1973.) Heikot siteet tarjoavat kanavan, jonka kautta ideat, vaikutteet ja informaatio ovat yksilön saatavilla, joka muuten olisi sosiaalisen etäisyyden vuoksi yksilön ulottumattomissa. Tällöin ne yksilöt, jotka toimivat eräänlaisina siltoina verkostoryhmien välillä, saavat aikaisemmin, laajemmin ja monimuotoisemmin käyttöönsä informaatiota. Näin ollen heillä on mahdollisuus siirtää informaatiota niiden verkostojen kesken, joihin he vaikuttavat. (Burt, 2004.)

Twitterissä tviittaaminen jakaantuu verkoston rakenteen vuoksi, jolloin vuorovaikutus ei ole sidoksissa vain rajoitetusti tiettyjen käyttäjien kesken eikä vahvaa homogeenisyyttä verkostossa näin pääse muodostumaan. Twitterissä tehokas informaation diffuusio tapahtuu retviittaamisen kautta, joka puolestaan tuo uusia käyttäjiä tietoisiksi eri aiheista. Tämä kuitenkin vaatii sitä, että vuorovaikutus ei ole sidoksissa vain rajoitetusti tiettyjen käyttäjien kesken ja näiden käyttäjien seuraajat koostuvat muista Twitterin käyttäjistä. Näin ollen homogeenisyys ei pääse muodostumaan esteeksi informaation diffuusiolle. (Boyd ym., 2010.) Retviittaamista informaation diffuusion mekanismina tarkastellaan tarkemmin luvussa 2.4.

### 2.3.4 Sosiaalisen verkoston vaikutus informaation diffuusion

Verkostolla ja siihen kuuluvilla solmuilla voidaan nähdä olevan merkityksellinen rooli diffuusioprosessin kannalta. Tämän vuoksi informaation diffuusiota on

usein tutkittu tuotteiden omaksumisen ja innovaatioiden leviämisen lähtökohdista. Everett Rogersin esittämä teoria innovaation diffuusiosta määrittelee innovaation omaksumisprosessin vaiheet, johon kuuluvat tietämys (knowledge), suostuttelu (persuasion), päätös (decision), toteutus (implementation) ja vahvistus (confirmation). Populaation sisällä tapahtuva omaksuminen seuraa S-mallista käyrää, jossa yksilöt on jaettu viiteen eri kategoriaan sen perusteella milloin tämä omaksuu innovaation: innovaattorit (innovators), aikaiset omaksijat (early adopters), aikainen enemmistö (early majority), myöhäinen enemmistö (late majority) sekä hitaat omaksijat (laggards). (Rogers, 2003, s. 21, 280-281.)

Rogersin teoria pyrkii selittämään sitä ilmiötä, että sosiaalisessa verkostossa yksilöiden omaksumispäätöksiin vaikuttavat verkoston muiden yksilöiden päätökset. Yksilöllä, joka verkostoa mallinnettaessa usein kuvataan solmulla, on siis merkityksellinen rooli informaation difuusion kannalta. Uudenlainen käyttäytyminen tai informaatio saa alkunsa aina tietyistä yksilöstä tai muutaman yksilön muodostamasta joukosta, leviten vähitellen verkostossa ja lopulta saavuttaen huomattavasti laajemman käyttäjäkunnan. Verkostosta yleensä löytyy yksilöitä, jotka toimivat omaksujina joka tapauksessa. Toiset puolestaan tarvitsevat enemmän vakuuttelua sekä muiden yksilöiden liittymistä omaksujien joukkoon ennen kuin itse omaksuvat. Omaksujien määrän saavuttaessa kriittisen massan diffuusioprosessi ohjautuu tämän jälkeen omavaraisesti. Mikäli kriittistä massaa ei saavuteta, omaksumiskäyttäytyminen saattaa hiipua ja kuolla pois. (Rogers, 2003, s. 343-344; Schelling, 1978, s. 89; Granovetter, 1978.)

Verkostossa jokainen solmu voidaan nähdä joko aktiivisena (tartunnan saanut, vaikutukselle altistunut) tai ei-aktiivisena. Aktiivinen solmu on myös kykenevä levittämään "tartuntaa" (esimerkiksi informaatiota tai vaikutusta) edelleen verkon muihin ei-aktiivisiin solmuihin. (Yang & Leskovec, 2010.) Diffuusioprosessi on siis käytännössä tietynlainen verkostossa tapahtuva aktivoitumisten sarja. Esimerkiksi Twitterissä käyttäjän julkaistessa sisältöä, hän ottaa aktiivisesti osaa informaation levittämiseen ja tällöin käyttäjää kuvaavan solmun voidaan todeta aktivoituneen.

Yleensä tutkimuksissa, joissa on pyritty mallintamaan informaation diffuusiota sosiaalisissa verkostoissa vertaiskäyttäjien päätösten vaikutuksesta, on tällöin tarkasteltu diffuusioprosessia stokastisesta eli satunnaisuuden huomioon ottavasta näkökulmasta noudattaen joko kaskadimallia (Cascade Model) tai kynnysarvon mallia (Threshold Model). Perusmalleina näistä mainituista malleista voidaan nähdä olevan Mark Granovetterin (1978) esittämä lineaarisen kynnysarvon malli (Linear Threshold Model) sekä Goldenbergin ym. (2001) esittämä riippumaton kaskadimalli (Independent Cascade Model).

Lineaarisen kynnysarvon mallissa solmun aktivoituminen riippuu siitä, kuinka moni naapurisolmuista on aikaisemmin aktivoitunut. Solmujen välillä esiintyy tietty painotus, joka määrittää solmujen keskinäisen vaikutusasteen. Jokaisella solmulla on oma kynnysarvonsa, joka määrittyy satunnaisesti tietyistä todennäköisyysjakaumasta. Kynnysarvo puolestaan määrittää sen, kuinka monen naapurisolmun on aktivoitettava ennen kuin tietty solmu aktivoituu. Tarkasteltavan solmun aktivoituminen tapahtuu, mikäli naapurisolmujen painotuksen

summa ylittää tämän kynnysarvon. (Granovetter, 1978.) Informaation diffuusion tapauksessa kynnysarvo saavutetaan, kun yksilö on riittävän vakuuttunut omaksumaan informaation tiedostaessaan, että tietty joukko yksilöitä omasta henkilökohtaisesta verkostosta on omaksunut myös kyseisen informaation ja ovat tyytyväisiä siihen. Kynnysarvon mallin voidaan siis nähdä olevan vastaanottajakeskeinen (Guille & Hacid, 2012). On syytä kuitenkin huomata, että tällainen kynnysarvo esiintyy nimenomaan yksilötasolla, kun taas kriittinen massa toimii järjestelmän tasolla, kuten esimerkiksi yhteisöissä ja organisaatioissa. Yksittäiset kynnysarvot selittävät mikrotason prosessin, jossa yhteenlasketut yksittäiset päätökset muodostavat kriittisen massan järjestelmän. Kynnysarvon mallissa oletetaan, että yksittäinen päätös omaksua tietty informaatio riippuu lukuisista muista verkoston yksilöistä, jotka ovat muuttaneet käyttäytymistään. (Rogers, 2003, s. 355-356.)

Diffuusion dynamiikka määräytyy siis varhaisten omaksujien kynnysarvon myötä, josta Gladwell (2002, s. 12-13) käyttää nimitystä keikahduspiste (the tipping point). Kun tämä raja on ylittynyt, diffuusioprosessi omaksumiskäyttäytymisen myötä kasvaa räjähdysmäisesti. Mikäli varhaisten omaksujien kohdalla kynnysarvo ei ylity, myös omaksumiskäyttäytyminen tulee surkastumaan miniimiin ja lopulta kuolee pois. Mikäli omaksuminen saavuttaa keikahduspisteen, diffuusionprosessi etenee nopeasti hidastuen kuitenkin myöhemmin. (Jackson & Leeat, 2005.)

Kaskadimallissa omaksuminen tietyn ajan sisällä riippuu pelkästään tarkasteltavan yksilön lähellä olevista muista yksilöistä, jotka ovat viimeisimpänä omaksuneet uuden informaation, eikä näin ollen ota huomioon tarkasteltavan yksilön koko verkostoa, kuten kynnysarvon malli tekee (Acemoglu, Ozdaglar & Yildiz, 2011). Toisin kuin kynnysarvon malli, kaskadimalli voidaan nähdä lähettäjakeskeisenä: prosessin aikana jokaisella yksilön aktiivisella naapurisolmulla on yksi tilaisuus vaikuttaa ja onnistua aktivoimaan ei-aktiivinen solmu. Toisin sanoen, joka kerta kun solmun naapuri omaksuu tietyn informaation, on olemassa mahdollisuus, että solmun vieressä oleva ei-aktiivinen solmu myös omaksuu kyseisen informaation. (Leskovec ym., 2007; Guille & Hacid, 2012.) Varhaisessa vaiheessa jotkin ei-aktiiviset solmut aktivoituvat, jolloin tapahtuu informaation omaksuminen. Jokaisen onnistuneen vaiheen aikana tietty määrä solmuja aktivoituu. Tapahtumasarjan voidaan nähdä päättyvän siinä vaiheessa, kun uusia solmuja ei enää aktivoitu. (Gruhl, Guha, Liben-Nowell & Tomkins, 2004.)

Myös tartuntatautien leviämisenä on yhteisiä piirteitä diffuusioprosessin näkökulmasta tarkasteltuna. Tämän vuoksi lääketieteessä esiintyvä tautien leviäminen toimii analogiana sille, miten informaatio leviää verkossa tartunnan saaneista solmuista tartunnalle alttiisiin solmuihin. Perinteiset taudin etenemistä kuvaavat nk. SIR-mallit (Susceptible, Infected, Recovered) perustuvat taudin sykliin kantajassaan. SIR-mallissa henkilö ensin altistuu (susceptible) taudille. Jos kyseinen henkilö on kontaktissa mahdolliseen taudinkantajaan, hän saa itse tartunnan (infected). Tämän jälkeen tauti etenee kantajassaan siten, että taudin kantaja lopulta paranee (recovered) tai mahdollisesti kuolee (death). (Newman, 2001.)



Taudin leviämisen malliin on tutkimuksissa lisätty myös muita rooleja yksilöille, kuten esimerkiksi muiden käyttäjien altistaminen (exposed) (Bhattacharya & Ram, 2012) ja uudelleen altistuminen (susceptible) taudille, jolloin parantunut yksilö voi saada tartunnan myöhemmin uudestaan. (Leskovec ym., 2007).

Sosiaalisessa verkostossa tapahtuvaan informaation diffuusioon taudin leviämisen malli kuitenkin soveltuu huonosti, sillä kyseinen malli ei ota diffuusioprosessin taustalla olevan verkoston vaikutusta huomioon ja olettaa samalla yksilöiden käyttäytymisen olevan yhdenmukaista. (Iribarren & Moro, 2011; Bhattacharya & Ram, 2012.) SIR-mallissa tartunnalle altistuneella yksilöllä ei myöskään ole valinnan mahdollisuutta haluaako tulla tartutetuksi vai ei. Tämä johtuu siitä, että yleensä pelkästään havainnoimalla ei tiedetä kuka on taudinkantaja jolloin myös yksilön valinnan mahdollisuus kaventuu minimiin. Harvoin on myöskään yksilöstä itsestään riippuvainen missä muodossa hän levittää mahdollista tautia eteenpäin. (Kimura, Saito & Nakano, 2007; Magnani, Montesi & Rossi, 2013.)

Diffuusioprosessissa tietynlaisen informaation leviäminen ja omaksuminen voidaan yleensä nähdä lähettäjän ja omaksujan välillä tapahtuvana harkittuna tekona. Informaation omaksuminen voi vaatia kypsyykseen myös aikaa. Tämä saattaa edellyttää myös omaksujan näkökulmasta lisäinformaation aktiivista hankkimista. Yleensä uuden informaation omaksumista voidaan pitää hyödyllisenä tapahtumana, mikä taas ei yleensä päde tautitartuntojen tapauksissa. (Betencourt, Cintron-Arias, Kaiser & Castillo-Chavez, 2006)

Vaikka nämä eri diffuusiomallit pyrkivät vastaamaan kysymykseen miten informaatio leviää verkostossa ja mikä merkitys vertaiskäyttäjien vaikutuksella on diffuusioprosessiin, ne perustuvat usein kuitenkin tiettyihin olettamuksiin, kuin varsinaisesti vaikuttavuuden mitattuihin vaikutuksiin. Nämä mallit olettavat verkoston, jossa diffuusioprosessi tapahtuu, olevan kokonaan verkoston ulkopuolelta havaittavissa. Tällöin ne myös olettavat, että informaation on mahdollista levitä ainoastaan kyseisen verkoston sisällä, eikä se näin ollen voi tulla sen ulkopuolelta. Tällaiset olettamukset ovat useimmiten kuitenkin epärealistisia, sillä tosiasiallisesti verkosto voi olla täysin tuntematon tai vain osittain tunnettu. Lisäksi verkoston todellista kokoa on useimmiten mahdoton määrittellä sen valtaavan koon vuoksi. (Yang & Leskovec, 2010.) Sosiaalisissa verkostoissa uudeksi koettu informaatio harvoin myöskään tulee yksilön lähellä olevista solmuista, vaan useimmiten nimenomaan tiiviin käyttäjäryhmän eli klusterin ulkopuolelta (Rogers, 2003, s. 340-341). Sama informaatio voi myös esiintyä eri aikaan useassa paikassa ilman, että sen olisi tarvinnut levitä koko verkon laajuudelta (Najar, Denoyer & Gallinari, 2012).

## 2.4 Informaation diffuusio Twitterissä

Usein sisältö, jonka käyttäjä Twitterissä jakaa seuraajilleen tulee luetuksi ja saman tien unohdetuksi. Joissain tapauksissa lähetetty tviitti ei tule luetuksi lain-

kaan sen vastaanottajan toimesta. Joissain tapauksissa tviitin sisältö voidaan kokea niin mielenkiintoiseksi, että seuraajat päättävät jakaa sen eteenpäin retviittaamalla omille seuraajilleen. (Kaplan & Haenlein, 2011.)

Twitterissä informaation diffuusio tapahtuu nopealla aikavälillä ja tviitin hyppäysten määrä, edeten käyttäjältä toiselle, saavuttaa nopeasti lakipisteensä. Retviittaamisen voidaan nähdä noudattavan potenssilakia: siinä missä esimerkiksi kuvanjakopalvelu Flickrissä suosituimmat kuvat leviävät hitaasti usean päivän jaksoissa, Twitterissä tviittit leviävät välittömästi siitä hetkestä, kun tviitti tulee retviitatuksi ensimmäisen kerran. Retviitattu tviitti voi saavuttaa hetkessä jopa tuhat käyttäjää riippumatta siitä mikä alkuperäisen tviitin lähettäjän todellinen seuraajamäärä on. Retviitattu tviitti tulee usein myös uudelleen retviitattua välittömästi eteenpäin. (Kwak ym., 2010.) Tiedon jakamisen välineenä Twitterin puolesta puhuukin nimenomaan tiedon nopea leviäminen käyttäjien muodostamassa sosiaalisessa verkostossa.

Retviittaaminen voidaan nähdä tarkoituksenmukaisena, mutta ei itseisarvoisena toimintana. Tviittejä ei lueta retviittaamisen takia, vaan siksi että tviitti ylipäättään ilmestyy käyttäjän nähtäväksi. Mikäli tviitti koetaan jollain tavalla merkitykselliseksi, halutaan se jakaa eteenpäin muille käyttäjille. (Kawamoto, 2013.) Tviittien potentiaalisten lukijoiden määrän voidaan nähdä kasvavan nopeasti jokaisella retviitin seurauksena tapahtuvalla hyppäyksellä käyttäjältä toiselle, mutta samalla lukijoiden määrän kasvu hidastuu ajan myötä. Tällainen käyttäytyminen osoittaa, että varsinainen diffuusio tapahtuu useimmiten aikaisessa vaiheessa tviitin alkaessa levitä sosiaalisessa verkostossa käyttäjältä toiselle retviittausten seurauksena. (Lee, Kwak, Park & Moon, 2010.)

Retviittaamisen teho perustuu nimenomaan käyttäjien kollektiiviseen toimintaan. Kun tviitti tulee retviitatuksi ja alkaa näin saada useampia potentiaalisia lukijoita, leviää se nopeasti verkossa saavuttaen tietynlaisen yleisön joka tapauksessa. Jokaisella käyttäjällä on tässä kohden päätäntävalta, mitkä tviittit hän haluaa jakaa eteenpäin muiden käyttäjien nähtäväksi. Tällä tavoin voidaan määrittää tietyn tviitin merkitystä verkoston käyttäjien näkökulmasta. (Kwak ym., 2010.) Voidaan ajatella, että retviittausten määrällä on käyttäjän vaikuttavuuden kannalta suurempi merkitys, kuin esimerkiksi käyttäjän seuraajien lukumäärällä. Suurin osa tviiteistä ei kuitenkaan leviä verkossa lainkaan tai sitten vain yhden hyppäyksen verran. Tämä kertoo siitä, että myös tviittien leviäminen verkostossa noudattaa potenssilakia; vain harvat tviittit, jotka tulevat retviitatuksi, leviävät laajasti käyttäjältä toiselle useamman hyppäyksen verran. (Bakshy, Hofman, Mason & Watts, 2011.)

Retviittaaminen on tehokasta silloin, kun käyttäjällä on laaja verkosto ja hän täyttää verkossa olevia rakenteellisia aukkoja eri yhteisöjen välillä hyödyntämällä heikkoja siteitä. Jotta informaatiota voi levittää uusille käyttäjille, käyttäjän täytyykin olla yhdistettynä sellaisiin käyttäjiin, joihin alkuperäinen informaation lähde ei ole vielä yhdistettynä. (Boyd ym., 2010.)

Myös kollektiivinen identiteetti muodostuu yhteisöllisen retviittaamisen seurauksena. Käyttäjät voivat esimerkiksi eräällä tavalla liittoutua tietyn aihepii-

rin taakse ja näin ollen manipuloida järjestelmää aiheuttaakseen sosiaalista toimintaa. Tästä esimerkkinä ovat eräänlaiset käyttäjien aiheuttamat ”kampanjat”, joita pyritään levittämään koko verkon laajuudelta kehottamalla muita käyttäjiä retviittaamaan kyseinen tviitti esimerkiksi mainitsemalla tviitissään ”RETWEET THIS UNTIL IT TRENDS” tai vain yksinkertaisesti ”RTs”. Tällä tavoin kyseisen tviitin aiheesta pyritään tekemään niin suosittu, että aiheen taustalla oleva ilmiö nousee trendiksi Twitterin käyttäjien kesken. (Boyd ym., 2010.) On kuitenkin havaittu, että käyttäjän lähettämien tviittien määrä ei kuitenkaan suoraan vaikuta retviittausten määrään. Näin ollen sisällön laadulla voidaan nähdä olevan vaikutusta siihen, millaisia tviittejä retviitataan eteenpäin. (Suh, Hong, Piroli & Chi, 2010.)

Diffuusion kannalta on myös merkitystä ketä käyttäjä itse Twitterissä seuraa, sillä tviitti tulee retviitatuksi todennäköisemmin mitä enemmän henkilö seuraa muita käyttäjiä. Tämä viittaa siihen, että mitä useampia lähteitä käyttäjä seuraa, sitä mielenkiintoisempia hänen itsensä tuottamat tviitit myös ovat. Tosin seurattavien vaikutus diffuusion kannalta ei ole yhtä voimakas kuin käyttäjän omien seuraajien määrä. (Suh ym., 2010.)

Seuraajien määrä ei kuitenkaan kerro sitä, mikä on tviitin todellinen yleisö. Kun viesti tulee ensimmäisen kerran retviitatuksi, sillä on mahdollista levitä laajasti verkossa, vaikka alkuperäinen yleisö olisikin vain pieni määrä käyttäjän omia seuraajia. Tässä piilee retviittauksen teho informaation jakamisen näkökulmasta. Ensimmäisen retviittauksen jälkeen tviitillä on mahdollisuus levitä nopeasti laajalle koko verkon laajuudelta. Jokaisella käyttäjällä on siis periaatteessa yhtäläinen mahdollisuus aiheuttaa ”sosiaalista epidemiaa” käyttäjien keskuudessa, mutta se vaatii muiden käyttäjien aktiivista osallistumista prosessiin. (Kwak ym., 2010.)

Informaation diffuusiota Twitterissä on tarkasteltu diffuusiomallin avulla, jossa jokainen tviitin eteenpäin retviitannut seuraaja on luokiteltu sukupolviin sen perusteella seuraako tämä retviittaajaa. Näin diffuusioketju etenee aina niin pitkälle, kunnes retviittaamista ei enää tapahdu tai viimeisen sukupolven seuraajat ovat jo saaneet kyseisen tviitin aikaisemmin jostain muualta. Retviittausten aktiivisuus voidaan näin mallintaa monilukuseksi satunnaisprosesseiksi, jossa jokaisella seuraajasukupolvella on merkityksellinen rooli diffuusion kannalta. Informaation diffuusio tapahtuu seuraajien kannalta yhteistyössä toistettujen toimintojen kautta. (Kawamoto, 2012.)

Kawamoton mallissa kuitenkin oletetaan, että diffuusio tapahtuu seuraaja-perusteisesti. Näin ollen siinä ei huomioida, että Twitter edustaa monille käyttäjille informaation lähdettä eikä sosiaalista kohtaamispaikkaa. Tviitti voi tulla käyttäjän havainnoimaksi ja täten retviitatuksi myös sosiaalisen verkoston ulkopuolelta, eli käyttäjältä johon ei ole suoraa seuraajasuhdetta. Näin voi tapahtua esimerkiksi Twitterin hakutoiminnon kautta. Esimerkiksi Tviittien sisältämällä aihetunnisteilla (#) on havaittu olevan merkittävä yhteys siihen, tuleeko tviitti retviitatuksi. Näiden aihetunnisteiden avulla pyritään sitomaan tviitti tietyn aiheeseen, jolloin se on myös helpommin muiden käyttäjien havaittavissa joko hakutoiminnon tai suosituimpien keskustelunaiheiden kautta. (Suh ym., 2010)

### 3 Vaikuttajat sosiaalisissa verkostoissa

Tässä luvussa tarkastellaan vaikuttamista käsitteenä, sen merkitystä sosiaalisessa verkostossa sekä muodostumista yksilön tasolla. Luvun aluksi määritellään vaikuttavuuden käsite, jonka jälkeen selvitetään kuinka vaikuttavuus muodostuu yksilön tasolla ja miten vaikuttavuus näyttäytyy Twitterissä. Luvussa tarkastellaan myös miten vaikuttavuutta on mahdollista mitata yhteisöpalveluissa käyttäjäkohtaisesti ja kuinka sitä on pyritty mittaamaan informaation diffuusion näkökulmasta. Tässä luvussa vaikuttavuuden käsite operationalisoidaan empiiristä tutkimusosiota varten mitattavaan muotoon. Luvun lopuksi esitetään yhteenvedo kirjallisuuskatsauksen pohjalta.

#### 3.1 Vaikuttavuuden muodostuminen

Gabriel Weimann (1991) määrittelee yksilön vaikuttavuuden muodostuvan kolmen henkilökohtaisen ja sosiaalisen asian yhdistelmästä: henkilöityminen arvojen mukaan (”kuka olet?”), yksilön kompetenssi (”mitä tiedät?”) sekä yksilön strateginen sosiaalinen sijainti omassa sosiaalisessa verkostossaan (”kenet tunnet?”). Gladwellin määritelmä vaikuttajan ominaisuuksista mukailee Weimannin määritelmää tarkentaen sitä kuitenkin joiltain osin. Hänen mukaansa vaikuttajilla on poikkeuksellisen paljon yhteyksiä ja ennen kaikkea merkityksellisiä yhteyksiä verkostossaan. Tällaisten yhteyksien voidaan katsoa olevan nimenomaan oman sosiaalisen verkoston ulkopuolelta tulevat yhteydet, merkitykselliset heikot siteet, jolloin vaikuttajalle tarjoutuu mahdollisuus päästä käsiksi laajaan tarjontaan uutta informaatiota. Vaikuttajat ovat myös sosiaalisesti motivoituneita, jonka mukaan heillä on halu jakaa informaatiota muille yksilöille. He myös omaavat erityistä tietämystä sekä sosiaalisia taitoja käynnistää ”sosiaalisia epidemioita”; pelkästään informaation määrä tai laatu ei siis ole merkityksellinen vaan se miten informaatio siirretään eteenpäin. (Gladwell, 2002, s. 38, 46, 62, 70.)

Tällaisten vaikuttajayksilöiden voidaan nähdä olevan keskiössä sosiaalisessa verkossaan, sillä lukuisten yhteyksien vuoksi heillä on mahdollisuus vaikuttaa ja tartuttaa ”sosiaalista epidemiaa” myös muihin verkon yksilöihin ilman, että heillä olisi edes suoraa yhteyttä heihin (Goldenberg ym., 2009). Koska vaikuttajat tuntevat paljon ihmisiä, toimivat he samalla myös yhdistäjinä muiden verkostossa olevien yksilöiden välillä (Gladwell, 2002, s.38). Milgram (1967) ei myöskään oletta pienen maailman ilmiössään henkilöiden olevan tasaveroisia verkostossaan, vaan he linkittyvät toisiinsa juuri tiettyjen erityisten yksilöiden kautta. Käytännössä tällä tarkoitetaan sitä, että pieni määrä erityisiä ihmisiä on yhdistettynä toisiinsa pienillä hyppäyksillä ja loput ihmisistä ovat linkitettyinä toisiinsa näiden erityisten ihmisten kautta (Gladwell, 2002, s. 36-37).

Vaikuttajista ja heidän merkityksestään on olemassa kaksi vastakkaista näkemystä. Ensimmäisen näkemyksen mukaan vaikuttajat ovat yksilöitä, joilla on

sellaisia erityislaatuisia ominaisuuksia, että esimerkiksi huhun leviäminen voidaan ennustaa pelkästään sen alullepanijoista. Tämä väittäminen perustuu kaksivaiheisen viestinnän teoriaan, jonka mukaan joukkoviestimien vaikutus yleisöön on epäsuoraa, eli informaatio virtaa joukkoviestimistä ensin potentiaalisille vaikuttajille ja heiltä eteenpäin muille verkostossa olijoille (Katz & Lazarsfeld, 1955, s. 32). Tätä tukee myös Gladwellin (2002, s. 33) esittämä ajatus siitä, että ”sosiaaliset epidemiat” lähtevät liikkeelle yksilöistä, jotka omaavat erityislaatuista ja harvinaista sosiaalista lahjakkuutta.

Toinen näkemys sosiaalisesta vaikuttamisesta on, että periaatteessa kuka tahansa voi olla vaikuttaja, jolloin vallitsevilla olosuhteilla voidaan nähdä olevan enemmän merkitystä kuin tietyllä yksilöllä. Duncan Watts (2007) toteaa, että esimerkiksi huhulla on mahdollisuus levitä laajasti vain siinä tapauksessa, jos olosuhteet ovat riittävän otolliset ja omaksumiseen löytyy sosiaalista alttiutta verkostossa. Näin ollen ”sosiaalisen epidemian” leviämistä ei voitaisi ennustaa pelkästään tarkastelemalla sen alullepanijaa, jonka ilmiö olisikin enemmän sattumanvarainen, ajasta ja paikasta riippuva tapahtuma. Tästä syystä Watts käyttääkin termiä ”sattumavaikuttajat” (accidental influentials). Hänen mukaansa ”sosiaaliset epidemiat” syntyvät pikemminkin helposti vaikutteille alttiiden yksilöiden saavuttaessa kriittisen massan eikä niinkään yksittäisten vaikuttajien vaikutuksesta. (Watts, 2007).

Vaikuttavuuden voidaan kuitenkin nähdä toimivan potenssilain mukaisesti sosiaalisissa verkostoissa. Tämä tarkoittaa sitä, että on olemassa pieni joukko yksilöitä, jotka vaikuttavat laajaan kuulijakuntaan, mutta samalla on iso joukko yksilöitä, joiden vaikutus on erittäin vähäistä tarkasteltaessa koko verkostoa. Vaikuttavuuden osalta potenssilaki toimii myös toiseen suuntaan, eli vaikutuksille alttiita ovat suurin osa verkossa olijoista ja on vain pieni joukko, jotka ovat vain vähän alttiita muiden vaikutuksille. (Junquero-Trabado, Trench-Ribes, Aguila-Lorente & Dominguez-Sal, 2011.) Usein vaikuttajista puhuttaessa ei kuitenkaan tarkoiteta esimerkiksi julkisuuden henkilöitä, vaan nimenomaan ”tavallisia” yksilöitä, joilla on erityistä vaikuttavuutta omassa verkostossaan. Kuitenkin esimerkiksi monilla julkisuuden henkilöillä ja asiantuntijoilla on erityistä vaikuttavuutta suureen yleisöön, sillä nämä pystyvät vaikuttamaan usein myös muiden kanavien, kuten esimerkiksi median kautta, jolloin vaikuttamisen mahdollisuus tarjoutuu laajemmassa mittakaavassa. (Bakshy ym., 2011.)

### 3.1.1 Vaikuttajat Twitterissä

Vaikuttaja on omassa sosiaalisessa verkostossaan eräänlainen solmukohta, jota muut käyttäjät ympäröivät. Vaikuttaja on myös vahvojen sekä heikkojen siteiden kautta yhdistyneenä määrällisesti isoon joukkoon muita solmuja (Gladwell, 2002, s. 38,46). Twitterissä seuraajat luovat yleisön käyttäjälle: mitä enemmän seuraajia sitä laajempi yleisö on käyttäjän tuottamalla sisällöllä. On myös mahdollista, että kahdella käyttäjällä voi olla lukumäärällisesti yhtä paljon seuraajia, mutta heidän vaikuttavuutensa eivät ole keskenään verrannollisia. Näin ollen pelkästään seu-

raajien lukumäärää tarkastelemalla ei voida arvioida yksittäisten käyttäjien vaikuttavuutta omassa verkossaan, jolloin käyttäjiä on vertailtava myös heidän vuorovaikutuksensa kautta. Käyttäjien keskinäisestä vaikuttavuudesta kertoo pikemminkin se, että heidän seuraajansa ovat aktiivisia ja retviittaavat tämän lähettämiä viestejä. (Cha ym., 2010.) Twitterissä voidaan tarkastella potenssilakia vaikuttavuuden osalta siten, että mitä enemmän käyttäjällä on seuraajia sitä enemmän, ainakin teoriassa, kyseinen käyttäjä on mainittuna muiden tviitteissä ja tulee muiden käyttäjien osalta huomioiduksi (Junquero-Trabado ym., 2011).

Myös sisällön laadulla on ympäröivän sosiaalisen verkon lisäksi merkitystä. Laadullisia tekijöitä voidaan katsoa olevan viestisisältöjen ajankohtaisuus sekä kiinnostavuus, jolloin sisältö on kohdistettu tiettyyn seuraajakuntaan. Yleisö, jonka voidaan katsoa olevan aihepiirin tiimoilta homogeeninen, on yleensä kiinnostunut juuri niistä asioista, jotka ovat heille merkityksellisiä. (Romero ym., 2011.) Yksittäisen tviittauksen semanttista merkityksellisyyttä voidaan tarkastella suoraan retviittausten kautta: mitä enemmän tviitti tulee retviitatuksi, sitä merkityksellisempi sen voidaan nähdä olevan (Raux, Grünwald & Prieur, 2011). Chan ym. (2010) mukaan vaikuttajat ovat Twitterissä yksilöitä, joiden käyttäytyminen osoittaa erityistä sitoutuneisuutta ja jotka rajoittavat tviittinsä tietyn aihealueen ympärille.

Vaikuttajat ovat kykeneviä myös luomaan uutta ja mielenkiintoista sisältöä riittävän tiheällä aikavälillä, mikä puolestaan edesauttaa sitä, että muut käyttäjät pysyvät tämän seuraajina. Näin ollen sosiaalisen verkostonsa keskiössä olevilla vaikuttajilla on selkeä halu jakaa informaatiota eteenpäin ja he usein pystyvät luomaan sellaista sisältöä, jollaista tämän seuraajat arvostavat. Samalla heidän jakamansa sisällön kautta nousee esiin kyseisen käyttäjän tietynlainen karismaattisuus, joka erottaa heidät vaikuttavuuden osalta muista käyttäjistä. He ovat usein erityisen aktiivisia käyttäjiä ja näin ollen heidän oma toimintansa heijastuu vaikuttavuuteen (Romero ym., 2011; Gladwell, 2002, s. 67-68.). Myös vaikuttavuuden ylläpitäminen vaatii panostusta; vaikka vaikuttavuus ei yksilön kohdalla olisikaan tarkoitushakuista, silti se pohjautuu yksilön omaan motivaatioon toimia verkostossa tietyllä tavalla (Cha ym., 2010).

Twitterin kontekstissa informaatio ei myöskään siirry suoraan medialta suurelle yleisölle, vaikka joukkoviestinten käyttäjäprofiilien onkin todettu olevan aktiivisimpia sisällöntuottajia Twitterissä. Sitä vastoin informaatio siirtyy tiettyjen välikäsien kautta, jotka ovat tavallisia käyttäjiä, mutta jotka samalla toimivat omassa sosiaalisessa verkostossaan keskiössä ja ovat näin ollen alttiimpia median suuntaan kuin heidän seuraajansa. (Wu ym., 2011.) Tämä havainto tukee Katz ja Lazarsfeldin (1955) kaksivaiheisen viestinnän teoriaa sen osalta, että myös Twitterissä informaatio siirtyy tiettyjen välikäsien kautta massoille.

Twitterissä vaikuttajat voidaan nähdä ennalta-arvattavampina kuin mitä Watts (2007) yleisesti esittää, eikä vaikuttavuus näin ollen synny vahingossa ja sattumanvaraisesti. Vaikuttavuuden muodostuminen kuitenkin muodostuu osittain vertaiskäyttäjien vaikutuksesta, sillä he joutuvat antamaan panoksensa esimerkiksi retviittaamiseen liittyen, joka näyttäytyy merkityksellisenä itse diffuusion kannalta Twitterissä. (Cha ym., 2010.)

### 3.2 Vaikuttavuuden mittaaminen

Vaikuttavuutta yhteisöpalveluissa, Twitter mukaan lukien, on tutkittu verrattain paljon, ilmiön ollessa tällä hetkellä vahvasti pinnalla. Weng ym. (2010) ovat tutkineet Twitter-käyttäjien vaikuttavuutta erillisellä TwitterRank-algoritmilla, jossa käyttäjän vaikuttavuus määräytyy paitsi tämän seuraajien vaikuttavuuden mukaan, myös sen perusteella, kuinka paljon käyttäjä on jakanut omille seuraajilleen tviittejä. Näin ollen käyttäjä, joka tuottaa enemmän tviittejä seuraajiensa nähtäväksi olisi suhteessa vähemmän tviittaavia käyttäjiä vaikuttavampi. Vaikuttavuuden mittaamisessa on pyritty huomioimaan myös käyttäjien keskinäisen homofilian merkitys, joka syntyy, mikäli käyttäjät ovat kiinnostuneita samasta aihepiiristä. Tutkimuksessa kuitenkin jätetään huomioimatta esimerkiksi käyttäjän seuraajien lukumäärä, sillä sitä ei nähdä perusteltuna mittarina vaikuttavuuden syntymisessä. Periaatteessa tälle tavalle mitata yksittäisen käyttäjän vaikuttavuutta voidaan esittää kritiikkiä, sillä useimmat tutkimukset vaikuttamisesta pitävät nimenomaan käyttäjän henkilökohtaista yleisöä yhtenä merkityksellisenä osana vaikuttavuuden syntymisessä. Twitterissä käyttäjän seuraajat ovat kuitenkin ensisijainen kohde käyttäjän lähettämille tviiteille, mikä tarjoaa myös yhteyden merkityksellisiin heikkoihin siteisiin.

Cha ym. (2010) puolestaan ovat ottaneet huomioon myös käyttäjien keskinäisen vuorovaikutuksen havainnoidessaan yksilön vaikuttavuutta. He käyttävät vaikuttavuuden mittareina käyttäjän seuraajien lukumäärän lisäksi käyttäjän retviitattujen tviittien sekä käyttäjän @username-syntaksin mainintojen määrää. Esimerkiksi lähetettyjen tviittien määrä on sitä vastoin koettu heikoksi vaikuttavuuden mittariksi, sillä sen avulla tilastoissa nousevat esille ”spämmääjät” sekä mahdolliset ei-ihmiskäyttäjät eli robotit.

Yang ja Lescovec (2010) ovat pyrkineet ennustamaan lineaarisen vaikuttavuuden mallin (Linear Influence Model) avulla solmujen vaikuttavuutta koko verkoston laajuudelta ja mikä suhde sillä on informaation diffuusion kannalta. Kyseisessä mallissa oletetaan, että uusien tartunnan saaneiden solmujen lukumäärä on riippuvainen aikaisemmin tartunnan saaneista solmuista. Näin aikaisemmin tartunnan saaneilla solmuilla on vaikutusta niihin solmuihin, jotka myöhemmin saavat tartunnan. Solmujen vaikuttavuutta on tarkasteltu määrittämällä, kuinka monta myöhempää tartuntaa voidaan katsoa johtuvan tietyn solmun vaikutuksesta ajan myötä. Vaikuttavuuden mittareina on käytetty seuraajien määrää ja käyttäjän omaa aktiivisuutta, joka tässä kohden tarkoittaa ainoastaan yksilön käyttämien #-merkkien kokonaismäärä tämän lähettämässä tviiteissä. Mallissa ei kuitenkaan huomioida verkoston rakennetta, eikä sitä mikä yksittäinen solmu aiheuttaa vaikuttavuudellaan muiden solmujen kohdalla informaation diffuusiota.

Erilaisia internet-pohjaisia palveluita henkilökohtaisen vaikuttavuuden mittaamiseen on tarjolla sisällönjako- ja verkkoyhteisöpalveluiden käyttäjille. Näiden palveluiden avulla käyttäjä on mahdollisuus tarkastella omaa henkilökohtaista vaikuttavuuttaan käyttämässään yhteisöpalveluissa, kuten esimerkiksi

Twitterissä ja Facebookissa. Eräs tällainen palvelu on Klout.com-sivusto, joka tarjoaa käyttäjälle vaikuttavuutta kuvaavan Klout-arvon (Klout Score) välille 1-100 mittaamalla sitä erilaisten mittareiden avulla. Käyttäjällä on sitä enemmän vaikuttavuutta sosiaalisessa verkostossaan mitä lähempänä hänen Klout-arvonsa on lukua 100. Twitterin osalta Klout.com käyttää merkityksellisimpinä vaikuttavuuden mittareina muiden käyttäjien retviittaamia käyttäjän tviittejä sekä käyttäjän @username-mainintoja, jolloin muut käyttäjien tviiteissä käyttäjä on mainittuna em. syntaksin kautta. Klout.com mittaa käyttäjän vaikuttavuutta viimeisen 90 päivän ajalta kertyneen informaation avulla. Käytännössä Klout-arvo osoittaa sen, kuinka hyvin käyttäjä osallistuu vuorovaikutukseen muiden käyttäjien kanssa sosiaalisessa verkostossa, mutta samalla on kritisoitu sitä, että ilmentääkö tämä nimenomaan käyttäjän henkilökohtaista vaikuttavuutta. (Bacile, 2013.)

Twitterissä käyttäjien lähettämiin tviitteihin pohjautuvaa vaikuttavuusanalyysia voidaan siis tehdä monella tapaa. Yleisesti voidaan kuitenkin todeta, että käyttäjien välinen vaikuttavuus syntyy ennen kaikkea vuorovaikutuksen kautta. Mikäli käyttäjä ei ikinä itse tviittaa tai retviittaa muiden käyttäjien tviittejä, ei hänellä tällöin myöskään ole vaikuttavuutta suhteessa vertaiskäyttäjiin. Vaikuttavuuden voidaan tässä tutkimuksessa todeta olevan mitattu arvo, joka perustuu käyttäjien tviiteistä ja retviiteistä muodostuvaan tviittivirtaan tietyn havainnointijakson aikana, jolloin tarkasteluväli voi olla aina tunneista jopa kuukausiin. Se käyttäjä, kenellä on eniten vaikuttavuutta kyseisellä aikavälillä, voi vaihdella paljon tarkastelujaksosta riippuen, eikä vaikuttavuus näin ollen ole välttämättä yleistettävissä tarkastelujakson ulkopuolelle. (Veijalainen ym., 2015a.)

### 3.2.1 Vaikuttavuuden operationalisointi

Jotta Twitterin käyttäjien yksilöllistä vaikuttavuutta voitaisiin arvioida, tulee vaikuttavuuden käsite operationalisoida. Operationaalisen määritelmän tehtävä ei ole niinkään kyseisen termin määrittelemineen, vaan sen muuttaminen mitattavaan muotoon, jolloin määritelmä on pikemminkin metodologinen kuin semanttinen. Määritelmän muuttamista mitattavaan muotoon käytetään sellaisissa tilanteissa, joissa on tarkoitus luoda indikaattori tietyn tutkimuksen tarkoituksia varten, jolloin prosessia kutsutaan tarkasteltavan termin operationalisoinniksi. Operationalisoinnissa termin määritelmä muunnetaan mittariksi koostuvista pienistä osatekijöistä, joista syntyvä kokonaisuus on tärkeämpi kuin sen yksittäinen osio. (Metsämuuronen, 2002, s. 25; Niiniluoto, 1999, s. 187.) Kun tarkasteltavan käsitteen teoreettinen määritelmä paikantaa ilmiön teorian maailmaan, niin operationaalinen määritelmä puolestaan kytkee teorian ja reaali maailman toisiinsa. Tavoitteena tällöin on, että teoreettinen ja operationaalinen määritelmä peittäisivät täsmälleen saman alan. (Uusitalo, s. 39.)

Tässä tutkimuksessa on tieteellisen kirjallisuuden kautta nostettu esiin niitä erityisiä piirteitä, joita vaikuttaja-yksilöllä on havaittu olevan. Näiden piirteiden kautta voidaan tarkastella miten vaikuttavuus käyttäjäkohtaisesti muodostuu.



Kirjallisuuden ja aikaisempien tutkimusten perusteella vaikuttajilla on poikkeuksellisen paljon yhteyksiä verkostossaan. Twitterissä tätä voidaan tarkastella seuraajien lukumäärän kautta. Mitä enemmän käyttäjällä on seuraajia, sitä todennäköisemmin tviitit tulevat myös retviitatuksi. Koska seuraajien lukumäärä ei vaikuttavuuden osalta välttämättä tee merkityksellistä eroa käyttäjien välillä, on tällöin tarkasteltava myös käyttäjän seuraajien kautta ilmenevää vuorovaikutusta.

Seuraajilla on vertaiskäyttäjien roolissa myös oma merkityksellinen roolinsa vaikuttavuuden muodostumisessa, sillä seuraajien kautta informaation diffuusio voimistuu heidän retviitatessaan käyttäjän lähettämiä tviittejä. Jotta voidaan tarkastella sitä, mikä merkitys seuraajilla varsinaisesti on vaikuttavuuden syntymisessä informaation diffuusion kannalta, on tarkasteltava myös retviittaavien käyttäjien merkitystä heidän omien seuraajiensa kautta. Kun käyttäjällä on omassa sosiaalisessa verkostossaan sellaisia seuraajia, joilla itsellään on paljon seuraajia, luo tämä pohjan Granovetterin mainitsemille merkityksellisille heikkoille siteille. Näiden heikkojen siteiden kautta informaatio välittyy tehokkaasti ympäri sosiaalista verkostoa. Tämä merkitsee sitä, että vaikka alkuperäisen tviitin lähettäjällä ei olisi paljon omia seuraajia, mutta niiden vähäisten seuraajien joukosta löytyy käyttäjiä, joilla on paljon seuraajia, voi tviitti saada yhdenkin retviittauksen kautta laajan yleisön osakseen.

Kun käyttäjän tviitti tulee retviitatuksi useammin kuin kerran, muodostuu tällöin retviittipuu, jonka kautta voidaan tarkastella miten retviitti etenee verkostossa käyttäjältä toiselle. Retviittipuun muodostuminen on selitetty tarkemmin luvussa 5.2. Käyttäjän vaikuttavuuden muodostumiseen vaikuttaa se, kuinka pitkä retviittipuusta lopulta muodostuu. Tällä tavalla voidaan tarkastella miten käyttäjä on levittänyt ”sosiaalista epidemiaa” sosiaalisessa verkostossa ilman, että käyttäjällä olisi varsinaisesti suoraa yhteyttä retviittaaviin käyttäjiin.

Tässä tutkimuksessa oli tarkoitus selvittää vaikuttavuuden merkitystä informaation diffuusion kannalta, jolloin merkityksellisiksi mittareiksi on pyritty ottamaan ne tekijät, jotka on katsottu kirjallisuuskatsauksen pohjalta oleellisimmiksi vaikuttavuuden indikaattoreiksi. Näitä tekijöitä on yhteensä kolme. Ensimmäinen merkityksellinen mittari on käyttäjän sosiaalinen verkosto, joka Twitterissä koostuu käyttäjän seuraajista. Seuraajien lukumäärän avulla pystytään tarkastelemaan käyttäjän potentiaalista yleisöä. Toinen mittari kuvaa sosiaalisessa verkostossa nk. heikkojen siteiden vaikutusta informaation diffuusion ja näin myös vaikuttavuuden syntyyn. Kolmas mittari kuvaa retviitattujen tviittien muodostamaa retviittipuuta, joka kuvaa käyttäjän tviitin aikaansaamaa ”sosiaalista epidemiaa” verkostossa. Taulukossa 1 on lukijalle havainnollistettu nämä edellä mainitut vaikuttavuuden mittarit, joiden avulla käyttäjän vaikuttavuutta on tämän tutkimuksen empiirisessä tutkimusosuudessa mitattu.

TAULUKKO 1 Vaikuttavuus-käsitteen operationalisointi vaikuttavuutta mittaavaksi indikaattoriksi

Teoreettinen käsite	Teoreettisen käsitteen määritelmä	Operationaalinen määritelmä
	Käyttäjän ominaisuudet:	Käyttäjän vaikuttavuus:
Vaikuttavuus	A) Sosiaalisen verkoston koko	1) Seuraajien määrä
	B) Yhteydet heikkoihin siteisiin	2) Käyttäjän merkitykselliset seuraajat, joiden kautta yhteys oman verkoston ulkopuolelle
	C) "Sosiaalisten epidemioiden" käynnistäjä	3) Retviittpuun koko

### 3.3 Yhteenveto

Tässä luvussa on tarkasteltu informaation diffuusiota ja siihen vaikuttavia merkityksellisiä tekijöitä sosiaalisessa verkostossa. Twitterissä käyttäjien keskinäinen vuorovaikuttaminen tapahtuu tviittaamalla eli lähettämällä viestisisältöjä muiden käyttäjien nähtäville. Näin toimiessaan he luovat ja jakavat ihmiskäyttäjälle merkityksiä antavaa informaatiota. Käyttäjien retviitatessa muiden käyttäjien tviittejä informaatio leviää käyttäjien muodostaman sosiaalisessa verkostossa, joka Twitterissä muodostuu verkoston solmuja kuvaavista käyttäjäprofiileista ja heidän keskinäisistä vuorovaikutussuhteistaan. Moniin muihin suosittuihin yhteisöpalveluihin verrattuna Twitter poikkeaa verkoston rakenteen osalta siten, että se on epäsymmetrinen eli asymmetrinen. Tällöin käyttäjien vastavuoroinen seuraussuhde ei ole vaatimuksena. Yleensä käyttäjät seuraavat Twitterissä toisiaan sen takia, että ovat kiinnostuneita tuotetuista sisällöistä, joka ilmentää mielenkiinnon osalta käyttäjien välistä samankaltaisuutta. Tällaista samankaltaisuutta kutsutaan homofiliaksi, joka saattaa ilmentyä esimerkiksi statuksen ja maantieteellisen sijainnin kautta (Rogers, 2003, s. 19-20). Homofilia synnyttää uusia yhteyksiä, sillä se kasvattaa mahdollisuuksia kahden henkilön väliseen vuorovaikutukseen ja näin ollen sen voidaan nähdä kiihdyttävän käyttäjien keskinäistä informaation diffuusiota (McPhearson ym., 2001). Homofilian vastakohtana oleva heterofilia kuvaa käyttäjien eriarvoisuusastetta. Heterofiliset suhteet verkoston sisällä puolestaan tuovat henkilöitä tietoisiksi uudenlaisesta informaatiosta. Tämä johtuu siitä, että mikäli verkosto on muodostunut pääasiassa homofilisista suhteista, uuden informaation pääsyyn verkoston sisälle vaaditaan yleensä heterofilinen kontakti eli heikon siteen muodostama yhteys toiseen verkostoon. Nämä heikot siteet mahdollistavat kontaktin sellaisiin yhteyksiin, jotka ovat suppean verkoston ulkopuolella ja näin ollen estynyt sosiaalisten

rakenteiden vuoksi. (Rogers, 2003, s. 19; Granovetter, 1973.) Twitterissä tämä ilmenee siten, että käyttäjä itse seuraa paljon muita käyttäjiä ja hänellä on paljon omia seuraajia eikä viestintä tällöin ole keskittynyt vain tiettyjen käyttäjien väliin vuorovaikutukseen.

Sosiaalisessa verkostoissa tapahtuvaa diffuusioprosessia on pyritty tutkimaan usein satunnaisuutta sisältävien mallien mukaan, jossa sekä käyttäjien keskinäisellä vuorovaikutuksella että heidän muodostamallaan sosiaalisella verkostolla on merkityksellinen rooli diffuusioprosessin kannalta. Kirjallisuudessa ja tutkimuksissa esitetyt diffuusiomallit kuitenkin tekevät usein tiettyjä oletuksia, jotka ovat informaation diffuusion kannalta epärealistisia. Todellisuudessa esimerkiksi verkostoa ei yleensä pystytä havaitsemaan ollenkaan tai siitä havaitaan vain suppea osa. Tällöin informaation leviämistä verkoston ulkopuolelta heikkojen siteiden kautta ei pystytä havaitsemaan.

Twitterin kontekstissa retviittaaminen on diffuusion kannalta avainmekanismi, jonka teho perustuu ennen kaikkea käyttäjän laajaan sosiaaliseen verkostoon, jossa pyrkimyksenä on täyttää verkossa olevia rakenteellisia aukkoja eri yhteisöjen välillä hyödyntämällä heikkoja siteitä. (Boyd ym., 2010.) Näin informaatio voi levitä sellaisiin solmuihin, johon informaation lähettäjä ei ole itse vielä yhdistettynä. Retviitattu tviitti leviää nopeasti verkossa, saaden näin joka tapauksessa tietynlaisen yleisön. Retviittaaminen on kuitenkin ennen kaikkea käyttäjälähtöistä toimintaa. Jokaisella käyttäjällä on kuitenkin yksilönä päätäntävalta siitä, haluaako ylipäänsä retviitata saamiaan tviittejä eteenpäin muiden käyttäjien nähtäväksi.

Informaation diffuusion kannalta vaikuttaja-yksilöillä on koettu olevan erityisen merkityksellinen rooli. Vaikuttajia on pidetty merkityksellisenä "sosiaalisten epidemioiden" syntymisessä, kuten esimerkiksi trendien ja huhujen levittämisessä laajalle kuulijakunnalle. Kaksivaiheisen viestintäteorian mukaan nämä kyseiset yksilöt ovat niitä, jotka toimivat informaation välittäjinä joukkoviestimistä laajemmalle yleisölle (Katz & Lazarsfeld, 1955, s. 32). Vaikuttajille on kirjallisuuden perusteella esitetty tiettyjä ominaisuuksia, jotka tekevät heistä erityislaatuista muiden yksilöiden keskuudessa. He ovat karismaattisia, omaavat laajan sosiaalisen verkoston ja osaavat hyödyntää käsillä olevaa informaatiota esimerkiksi suositteluun sekä sen eteenpäin siirtämiseen. (Gladwell, 2002, s. 67-68.)

Samalla kun näitä vaikuttajia on pyritty hyödyntämään esimerkiksi viraalimarkkinoinnin yhteydessä, on esitetty myös vastakkainen näkökulma vaikuttajien merkityksellisyydestä. Vallitsevilla olosuhteilla väitetään olevan enemmän merkitystä "sosiaalisten epidemioiden" syntymisessä kuin erityisellä yksilöllä tapahtuman taustalla. Näin ollen vaikuttaminen olisi ennemmin sattumankauppaa, jossa vallitsevien olosuhteiden merkitys korostuu yksilöä enemmän. (Watts, 2007)

Twitterissä vaikuttajilla on kuitenkin havaittu olevan merkityksellinen rooli informaation diffuusion eli tviittien leviämisen näkökulmasta. Yksittäisen käyttäjän osalta vaikuttavuuden voidaan nähdä muodostuvan monesta eri tekijästä. Yksi merkityksellisimmistä tekijöistä on vertaiskäyttäjät, jotka seuraamalla käyttäjää luovat tämän lähettämille tviiteille yleisön. Koska vertaiskäyttäjillä on oma merkityksensä vaikuttavuuden muodostumisessa, ei vaikuttavuuden voida

nähdä olevan tarkoitushakuista. Twitterissä vaikuttajien voidaan nähdä olevan erityisen aktiivisia sisällöntuottajia ja ennen kaikkea motivoituneita jakamaan informaatiota muille käyttäjille. Näin ollen vertaiskäyttäjillä on halu liittyä heidän sosiaalisen verkkonsa piiriin seuraamalla näitä erityisiä yksilöitä saadakseen mielenkiintoista informaatiota. Pelkästään seuraajien lukumäärän ei kuitenkaan voida nähdä olevan vaikuttavuuden kannalta ainoa merkitsevä mittari, sillä seuraajien täytyy toimia myös itse aktiivisesti retviitaten käyttäjän lähettämiä tviittejä, jotta voidaan havaita ”sosiaalista epidemiaa” eli informaation diffuusiota. Vaikuttajien täytyykin olla omassa sosiaalisessa verkostossaan muiden aktiivisten käyttäjien ympäröimä. Näin ollen vaikuttaminen voidaan nähdä syntyvän vastavuoroisen prosessin kautta.

Havainnot aikaisemmista tutkimuksista osoittavat, että vaikuttavuuden määritelmä ei ole kovin yksiselitteinen, eikä sen mittaamiseen ole myöskään mitään yhteneväistä tai selkeää tapaa. Tämän vuoksi tässä tutkimuksessa vaikuttavuutta on pyritty tarkastelemaan aikaisemman tutkimuksen pohjalta parhaimmiksi havaittujen mittaustapojen pohjalta, jonka vuoksi vaikuttavuuden määritelmä operationalisoitiin tämän tutkimuksen kannalta mitattavaan muotoon. Vaikuttavuuden muodostumisen kannalta tärkeimmiksi indikaattoreiksi Twitterissä todettiin olevan yhteensä kolme tekijää, joiden kautta tämän tutkimuksen empiirisessä osuudessa vaikuttavuutta tullaan mittaamaan. Ensimmäinen merkityksellinen tekijä on käyttäjän omien seuraajien määrä, joka määrittää minkä kokoiselle yleisölle käyttäjän lähettämät tviitit tulevat ensisijaisesti luettaviksi. Mitä enemmän käyttäjällä on seuraajia, sitä potentiaalisemmin käyttäjän tviitit tulevat myös retviitatuksi eteenpäin. Toisena tekijänä todettiin olevan käyttäjän merkitykselliset heikot siteet omassa sosiaalisessa verkostossa. Mikäli käyttäjän seuraajilla on paljon omia seuraajia, on mahdollista, että retviitatussa tviitti tulee saamaan laajan yleisön ja tällöin myös ”sosiaalinen epidemia” leviää verkostossa laajasti. Tällaista tviitin leviämistä retviittien kautta käyttäjältä toiselle voidaan tarkastella retviittipuiden syvyyttä havainnoimalla, joka toimii kolmantena tekijänä vaikuttavuutta tarkasteltaessa. Mitä syvempi retviittipuusta muodostuu, sitä laajemmalle tviitti leviää Twitterissä eri käyttäjien nähtäville.

## 4 EMPIIRINEN TUTKIMUS

Tässä luvussa tarkastellaan tutkimuksen empiiristä menetelmää sekä sen toteuttamista. Luvun aluksi esitellään tutkimuksen tavoite eli tutkimusongelma, sekä siihen liittyvät tarkentavat tutkimuskysymykset. Seuraavaksi luvussa tarkastellaan tutkimuksen toteuttamisessa käytettyä tutkimustapaa sekä -prosessia. Luvun lopuksi esitellään, kuinka tutkimusaineiston kerääminen Twitteristä toteutettiin.

### 4.1 Tutkimuksen tavoite

Tämän tutkimuksessa tarkastellaan informaation diffuusion vaikutusta Twitterin sosiaalisessa verkostossa käyttäjän vaikuttavuuden muodostumisen näkökulmasta. Twitteristä kerätyn aineiston perusteella pyritään havaitsemaan miten informaatio siirtyy käyttäjältä toiselle ja miten vaikuttavuus informaation diffuusion näkökulmasta muotoutuu käyttäjäkohtaisesti. Teoria- ja empiriaosuiden avulla pyritään saamaan vastauksia siihen, miten informaation diffuusio ilmenee Twitterin kaltaisessa sosiaalisissa verkostoissa ja miten vaikuttavuus muodostuu informaation diffuusion vaikutuksesta.

Yksilön vaikuttavuuden muodostumiseen liittyvä tutkimusongelma on mielenkiintoinen ja se on herättänyt aikaisempien tutkimusten pohjalta vastakkainasettelua. Samalla tutkimuksen taustalla on kirjoittajan henkilökohtainen halu tarkastella näitä jollain tavalla jopa mysteerisiksi luokiteltuja vaikuttajakäyttäjiä.

### 4.2 Tutkimustapa

Tutkimuksessa hyödynnetään case- eli tapaustutkimusta, jolla on kvantitatiivinen kehys. Kvantitatiivisuudella viitataan siihen, että tutkimustavan paradigmassa korostetaan yleispätevän syyn ja seurauksen lakeja. Tutkimustavan taustalla on realistinen ontologia, jonka mukaan todellisuus rakentuu objektiivisesti todettavista tosiasioista. Ajatuksena on, että kaikki tieto on peräisin suorasta aistihavainnosta ja loogisesta päättelystä, joka perustuu näihin havaintoihin. (Hirsjärvi ym., 2007, 135.)

Kvantitatiivisessa tutkimustavassa keskeisiä asioita ovat mm. johtopäätökset aiemmista tutkimuksista, aiemmat teoriat sekä käsitteiden määrittely. Näitä asioita on pyritty kattavasti huomioimaan jo aikaisemmin tieteellisen kirjallisuuden pohjalta teoriaosuudessa. Lisäksi kvantitatiivisessa tutkimustavassa havaintoaineisto perustuu määrälliseen ja numeeriseen mittaamiseen, jossa tutkittava joukko määritellään perusjoukosta otannan avulla, muuttujista muodostetaan

taulukko ja aineisto saatetaan tilastollisesti käsiteltävään muotoon, joka käsitellään tilastolliseen analyysiin perustuen. (Hirsjärvi ym., 2007, s. 136.)

Case- eli tapaustutkimuksessa puolestaan pyritään tarkastelemaan yhtä tai useampaa toisiinsa suhteessa olevaa tapausta, jolloin tarkoituksena on tuottaa yksityiskohtaista ja intensiivistä tietoa. Järvisen ja Järvisen (2004, s. 79) mukaan Yin (1989) määrittelee tapaustutkimuksen siten, että kyseessä on empiirinen tutkimusote, jossa tarkastellaan tämän päivän ilmiötä sen todellisessa kontekstissa käyttäen monia evidenssin lähteitä tiedostaen samalla, että ilmiön ja kontekstin rajapinta ei ole selkeä. Tiedonhankintatapoina tapaustutkimuksessa käytetään tavallisesti kyselyitä, haastatteluita, havainnointia sekä arkistomateriaalia, jolloin kerättävä tieto voi yhtä hyvin olla sekä kvantitatiivista että kvalitatiivista. Tapaustutkimus voi olla luonteeltaan kuvailevaa, teoriaa testaavaa tai uutta teoriaa luovaa. Vaikkei teoreettisesti mitään uutta löydykään, voi tapauksen kuvaus kuitenkin tarjota uutta näkemystä siitä, millainen maailma on. (Järvinen & Järvinen, 2004, s. 75, 79.)

### 4.3 Tutkimusprosessi

Tutkimusprosessi alkoi keväällä 2012 tutkimusaiheen suunnittelulla. Alkuperäinen suunnitelma oli tarkastella sitä, mitkä eri osa-alueet vaikuttavat yksittäisen käyttäjän vaikuttavuuden syntymiseen sosiaalisessa mediassa. Tutkimuksen ensimmäinen vaihe sisälsi kirjallisuuteen ja teoreettiseen aineistoon perehtymistä aiheen tiimoilta. Tutkimuksen aihealuetta rajattiin siten, että varsinaiseksi tutkimuskohteeksi muodostui yksittäiseen tapahtumaan liittyvä vaikuttavuuden muodostumisen havainnointi Twitterissä informaation diffuusion näkökulmasta. Kun tieto Bostonin pommi-iskusta tuli julki, aloitettiin heti empiirisen aineiston kerääminen, jolloin myös varsinainen syklinen tutkimusprosessi muotoutui. Tutkimuksen osalta varsinainen kirjoitusprosessi tapahtui samanaikaisesti tiedonkeruun kanssa sekä varsinaisen tutkimusaineiston analyysin jälkeen.

Teoreettisen aineiston osalta primäärisinä lähteinä hyödynnettiin erilaisia elektronisia aineistokantoja, kuten ACM Portal, Business Source Elite, Emerald journals, IEEE Xplore sekä SpringerLink. Lisäksi kirjallista aineistoa haettiin käyttäen apuna NELLI-portaalia sekä tieteellisten julkaisujen Google Scholar -hakupalvelua. Kirjallisen aineiston etsintä tapahtui tekemällä erilaisia sanahakuja mahdollisimman monipuolisesti elektronisten ja painettujen lähteiden tietokantoihin.

Tutkimusongelma haluttiin rajata jo varhaisessa vaiheessa, jotta tutkimuksesta muotoutuisi selkeä kokonaisuus. Tutkimusprosessin ohessa kirjoitettiin keväällä 2015 yhdessä tämän tutkimuksen ohjaaja Jari Veijalaisen sekä Jyväskylän yliopiston tutkijatohtori Alexander Semenovin kanssa kaksi tutkimusartikkelia liittyen tässä tutkimuksessa tutkittavaan aihealueeseen. Tässä tutkimuksessa hyödynnettyä tutkimusaineistoa hyödynnettiin myös näiden tutkimusartikkelien yhteydessä.

## 4.4 Tutkimusaineiston kerääminen

Tässä tutkimuksessa tarkasteltavan aineiston kerääminen aloitettiin 15. huhtikuuta 2013, noin yksi tunti sen jälkeen kun ensimmäisen pommi räjähti Bostonissa. Aineiston kerääminen jatkui noin viiden päivän ajan. Keräämisprosessissa hyödynnettiin Twitterin sovellusrajapintaa (API 1.1) määrittelemällä avainsanat, jotka toimivat hauille parametreinä. Hakusanoina käytettiin seuraavia sanoja ilman #-merkkiä: "Boston" ja "bombing".

Kerätystä aineistosta tutkittiin vuonna 2014 huhti-kesäkuussa käyttäjien seuraajarelaatiot, jotta voitiin tarkastella Twitter-käyttäjien sosiaalista verkostoa sekä tviittien leviämistä seuraajien kesken. Seuraajarelaatioiden tutkiminen jatkui saman vuoden marras-joulukuussa. Apuna seuraajarelaatioiden keräämisessä käytettiin Tweepy-sovellusta (<https://pypi.python.org/pypi/tweepy>). Koska Twitter rajoittaa API-rajapinnan hyödyntämistä seuraajainformaation keräämiseen (15 käyttäjää per 15 minuuttia), luotiin tätä varten noin 100 erillistä Twitter-sovellusta suorittamaan rinnakkaista keräämistä.

Varsinaisen aineiston keräämiseen hyödynnettiin Python-pohjaista sovellusta, joka valitsi 15 ennestään käsittelemätöntä käyttäjätiliä käyttäjätunnuksen (`user_id`) perusteella käyttäjätunnustaulusta (`input table twitter_user_ids`) sekä haki jokaiselle tunnukselle seuraaja-id:t (`follower user_ids`) Twitterin API:sta. Seuraajaparit (`user_id1`, `user_id2`) tallennettiin seuraajatauluun (`follower_user_ids`) suorituksen toteutumista kuvaavan aikaleiman kanssa. Jotta välttyttiin API:n rajoittimen ylitykseltä, loppui suoritus 15 minuutiksi siinä vaiheessa, kun Tweepyn toiminnallisuus `api.rate_limit_status()` palautti arvon nolla (0).

Kokonaisuudessaan tutkittava aineisto käsitti 8090803 tviittia. Näistä 3743696 olivat alkuperäisiä ja 4347107 retviitattuja, eli muiden käyttäjien toimesta uudelleenlähetettyjä tviittejä. Tviittien sekä niiden sisältämän metadatan avulla kerättiin myös aineistoon liittyvät käyttäjätiedot. Aineistosta pystyttiin erittelemään kokonaisuudessaan 4150569 yksittäistä käyttäjätiliä, joista tallennettiin seuraavat tiedot:

- `screen_name`: tviittauksen lähettäjän käyttäjänimi Twitterissä
- `user_id` (id): käyttäjän yksilöllinen käyttäjänimen tunnus Twitterissä
- `created_at` (`created_at`): milloin käyttäjätili on perustettu
- `followee_count` (`friends_count`): niiden käyttäjien lukumäärä, joita käyttäjä itse seuraa
- `followers_count` (`followers_count`): niiden käyttäjien lukumäärä, jotka seuraavat kyseistä käyttäjää
- `number_of_tweets` (`statuses_count`): käyttäjän lähettämien tviittien lukumäärä

Prosessina aineiston kerääminen oli hyvin altis virheille useista eri syistä. Useimmiten aineiston keräämistä varten luotujen sovellusten API-kutsu kesti

liian kauan, jolloin hakutoiminta kaatui. Toinen yleinen syy oli vähäinen päämuisti, jonka vuoksi tietokantaan tallentaminen epäonnistui.

Aineiston keräämiseen liittyvät ongelmat aiheuttivat myös sen ongelman, että kerätyt seuraajatiedot saattoivat olla vaillinaisia. Tämä tapahtui eritoten silloin, kun käyttäjällä oli enemmän kuin 5000 seuraajaa, koska tietojen kerääminen tapahtui maksimissaan 5000 seuraajan erissä ja heti haun jälkeen seuraajat lisättiin seuraajarelaatioon. Syynä tähän oli Twitterin API:n rajoitus, joka rajoitti yksittäisen haun 5000 seuraajaan. Näin ollen yksi sovellus pystyi hakemaan enimmillään 300000 seuraajaa tunnissa ja vain 7,2 miljoonaa seuraajaa päivässä, eli 5000 seuraajaa 15 minuutin välein yhden käyttäjän osalta. Mikäli tietojen kerääminen kaatui kesken haun, kerääminen jouduttiin aloittamaan alusta. Twitterin API ei myöskään tarjonnut tältä osin jo kerätyn aineiston palauttamista. Useilla käyttäjillä saattoi olla satojatuhansia tai jopa miljoonia seuraajia, joten tietojen keräämiseen meni useita tunteja. Enimmillään seuraajia yhdellä käyttäjällä oli yli 60 miljoonaa. Käytännössä tämä tarkoitti sitä, että yhden tällaisen todella suosittun käyttäjän seuraajien kerääminen kesti 8-9 päivää jolloin keräimen vikaantumiseen oli suuri todennäköisyys näin pitkässä keräysjaksossa.



## 5 TUTKIMUKSEN TULOKSET

Tässä luvussa käydään läpi tutkimuksen kannalta merkityksellisimmät tutkimustulokset ja niistä tehdyt havainnot. Aluksi lukijalle taustoitetaan tutkimusaineisto, jonka jälkeen selitetään ne mittaukset, joiden kautta vaikuttavuutta tutkimuksessa on mitattu. Tämän jälkeen luvussa suoritetaan tutkimusaineistoon pohjautuva tviitti- ja vaikuttavuusanalyysi. Nämä raportoidut tulokset on esitetty Veijalaisen, Semenovin ja Reinikaisen (2015b) tutkimusartikkelissa, jossa on tarkasteltu vaikuttavuuden leviämistä alkuperäisen tviitin tai retviitin lähettäjän tunnettuuden perusteella. Luvun lopuksi annetaan vastaukset tutkimuskysymyksiin kirjallisuuskatsauksen ja empiirisen tutkimusosuuden perusteella, sekä esitetään tutkimuksen pohjalta tehdyt johtopäätökset.

### 5.1 Taustatietoa tutkimusaineistosta

Kuten aikaisemmin on mainittu, seuraajarelaatioiden tutkiminen suoritettiin yli vuosi sen jälkeen, kun varsinainen aineisto kerättiin. Tämä tarkoittaa sitä, että todennäköisesti ne käyttäjien seuraajat, jotka aineistossa olivat aikavälillä huhtikuu-kesäkuu 2013, eivät olleet täysin identtisiä aikavälillä huhtikuu-joulukuu 2014. Yksi syy tähän on se, että monien käyttäjätilien seuraajien määrä oli muuttunut tällä aikavälillä. Esimerkiksi Justin Bieberillä (@justinbieber) oli vuonna 2014 noin 58,5 miljoonaa seuraajaa, kun taas huhtikuussa 2016 seuraajien määrä oli kasvanut yli 81 miljoonaan seuraajaan.

On myös huomioitava, että kaikkien käyttäjien seuraajia ei saatu kerättyä. Yleisin syy tähän oli se, että käyttäjä oli määritellyt käyttäjätilinsä suojaetuksi. Toinen yleinen syy oli, että käyttäjätili oli suljettu. Kesällä 2014 kerätystä aineistosta 10-12% osalta ei pystytty keräämään seuraajatietoja. Tämä luku kasvoi marras-joulukuun tarkastelussa, jolloin noin 15% seuraajatiedoista ei ollut saatavilla. Näin ollen kerätyssä aineistossa voidaan nähdä olleen tuntuva aukko.

Vuonna 2013 kerätystä aineistosta kaikkien seuraajien lukumääräksi saatiin noin 6,7 miljardia seuraajaa ja keskiarvoksi 1615 seuraajaa. Vuonna 2014 kaikkien seuraajien lukumäärä oli yli 10 miljardia, seuraajien keskiarvon ollessa 2052 seuraajaa. Huhtikuussa 2013 seuraajien mediaani oli 176 seuraajaa, kun taas vuoden 2014 tarkastelujakson aikana vastaava luku oli 280.

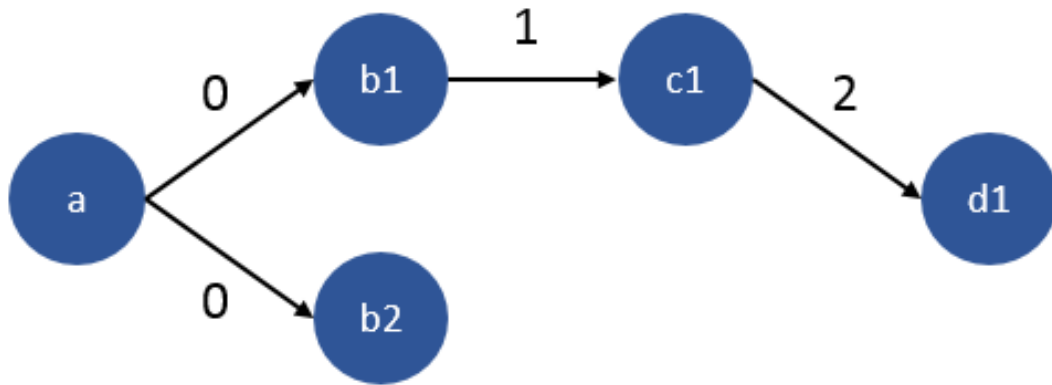
Aineistosta kävi ilmi, että vuonna 2013 kerätystä aineistosta hieman yli 2,6 miljoonalla käyttäjällä oli enemmän seurattavia kuin seuraajia. Noin 1,5 miljoonalla käyttäjällä tulos oli päinvastainen ja noin 42000 käyttäjällä oli yhtä paljon seuraajia kuin seurattavia. Niillä käyttäjillä, joilla oli enemmän seurattavia kuin seuraajia, seuraajien mediaaniarvo oli 109. Näin ollen mediaaniarvon yläpuolella olevilla käyttäjillä oli enemmän seurattavia kuin seuraajia ja mediaanin alapuolella päinvastoin. Vain 94 käyttäjällä, joilla oli yli 100000 seuraajaa, oli enemmän seurattavia kuin seuraajia. Huhtikuussa 2013 noin 100000 käyttäjällä, joilla oli

enemmän kuin 1000 seuraajaa, oli enemmän seurattavia kuin seuraajia. Näin ollen 2,5 miljoonalla käyttäjällä, joilla seurattavasuhde dominoi, oli vähemmän kuin 1000 seuraajaa. Kun tarkastellaan niitä noin 1,5 miljoonaa käyttäjää, joilla seuraajia oli vuonna 2013 huhtikuussa enemmän kuin seurattavia, seuraajien mediaaniarvo 365 oli ymmärrettävästi korkeampi, koska tähän joukkoon lukeutui ne käyttäjät, joilla oli yli 100000 seuraajaa.

Aineistosta tehtyjen havaintojen mukaan noin 2,13 miljoonaa erillistä käyttäjää lähetti yhden tai useamman alkuperäisen tviitin, kokonaismäärän oltua 3,74 miljoonaa alkuperäistä tviittiä tarkastelujakson aikana. Noin 2,02 miljoonaa käyttäjää ainoastaan retviittasivat muiden käyttäjien lähettämiä tviittejä. Tässä joukossa oli noin 0,52 miljoonaa käyttäjätiliä, joiden yksi tai useampi tviittaus retviitattiin muiden käyttäjien toimesta vähintään kerran. Näin ollen hieman yli 3 miljoonaa tviittiä, jotka tulivat tviitatuksi noin 1,57 miljoonan käyttäjän toimesta, ei retviitattu tarkastelujakson aikana kertaakaan. Tätä havaintoa tukevat aikaisemmissa tutkimuksissa esitetyt havainnot siitä, että retviittaaminen voidaan nähdä tarkoituksenmukaisena, mutta ei suinkaan itseisarvoisena toimintana. Tviittejä ei siis lueta retviittaamisen takia, vaan siksi että tviitti yleensäkin ilmestyy käyttäjän nähtäväksi. Ja mikäli käyttäjä kokee tviitin merkitykselliseksi, jaetaan se eteenpäin omille seuraajille retviittaamalla. (Boyd ym. 2010; Kawamoton 2013.)

## 5.2 Käyttäjän vaikuttavuuden mittaaminen

Tässä tutkimuksessa vaikuttavuuden mittaaminen perustui käyttäjän seuraajien lukumäärään, retviitattujen tviittien lukumäärän suhteesta seuraajien lukumäärään sekä retviitatuista tviiteistä muodostuvan retviittipuun kokoon. Retviittipuun muodostuu kun alkuperäinen tviitti retviitataan yhden tai useamman käyttäjän toimesta eteenpäin. Retviittien aikaleimoista ja käyttäjien seuraajista voidaan tällöin muodostaa nk. ketju, joka on syvyydeltään niin pitkä kuin retviittien määrä on seuraajaperusteisesti. Retviittipuun muodostuminen on havainnollistettu kuviossa 5. Kuviossa nuolet kuvaavat tviitin suuntaa lähettäjältä katsoen. Alkuperäinen lähettäjä a lähettää tviitin, jolla on yhteensä kaksi vastaanottajaa b1 ja b2. Vastaanottajista b1 retviittaa alkuperäisen tviitin eteenpäin, jonka jälkeen myös retviittin vastaanottaja c1 retviittaa sen eteenpäin. Kun vastaanottaja d1 ei enää päättää retviitata tviittiä eteenpäin, voidaan retviittipuun syvyydeksi laskea kaksi sen perusteella, kuinka monta retviittausta ensimmäisen retviittajan ja viimeisen vastaanottajan välille muodostui. Kuviossa b1:n kautta muodostuva retviittipuun saa arvokseen kaksi. Kuviossa alkuperäisen tviitin vastaanottaja b2 ei retviittaa tviittiä eteenpäin, joten retviittipuun saa arvokseen nolla.



KUVIO 5 Retviittipuun muodostuminen

Mitä useammin käyttäjät saivat suoraan alkuperäiseltä lähettäjältä tviitin tai tuo tviitti tuli retviitatuksi muiden käyttäjien toimesta, sitä enemmän tällä lähettäjällä voidaan todeta olevan vaikuttavuutta. Tällöin ennen kaikkea seuraajien lukumäärällä on merkitystä, koska potentiaalisten tviitin saajien lukumäärä on riippuvainen seuraajien lukumäärästä. Mikäli käyttäjällä on paljon seuraajia, myös potentiaalisten retviittausten määrä on suuri. Potentiaalisten tviitin lukijoiden määrä voidaan laskea suoraan tviittivirrassa olevista tviiteistä, koska Twitterin API:sta saatava yksittäisen tviitin metadata sisältää myös tviitin lähettäjän seuraajien lukumäärän. Näin voitiin havaita kuinka moni käyttäjä sai tviitin muilta käyttäjiltä tarkasteltavan ajanjakson aikana. Mitä suurempi joukko oli kyseessä, sitä suurempi oli myös käyttäjän vaikuttavuus. Tietyn käyttäjän lähettämän yksittäisen tviitin vaikuttavuuden määrä tietyssä tviittivirrassa voitiin laskea seuraavalla tavalla:

$$\text{Infr}(X, \text{Stream}) = p_1 * F_x + p_2 * RT_{m1} + \dots + p_{m_x} * RT_{mM_x}$$

Edellä olevassa laskentatavassa  $M_x$  tarkoittaa alkuperäisen käyttäjän  $X$  lähettämää tviittiä ja  $F_x$  tarkoittaa käyttäjän  $X$  seuraajien lukumäärää.  $RT_{mi} = F_{mk1} + F_{mk2} + F_{mki}$  tarkoittaa niiden seuraajien ( $X_{mk1} - X_{mki}$ ) lukumäärä yhteensä, jotka retviittasivat tviitin  $m_i$ .  $0 < i < M_x + 1$  tarkoittaa niiden potentiaalisten tviitin  $m_i$  saajien määrä otoksen koon ollessa  $k_i$ . Passiivisten seuraajien osuus voitiin havaita erillisellä kertoimella, jossa jokainen  $p_i$ ,  $0 \leq p_i \leq 1$ . (Veijalainen ym, 2015a.)

Alkuperäisen tviitin potentiaalisten saajien joukon koko saadaan asettamalla  $p_i = 1$   $i$ :n kaikille arvoille. Tällä tavoin voidaan vastata kysymykseen, kuinka moni käyttäjä sai tietyn käyttäjän lähettämän alkuperäisen tviitin. Samalla tämä kuitenkin yliarvioi käyttäjän vaikuttavuutta, sillä eri käyttäjillä voi olla samoja seuraajia ja näin ollen seuraajajoukot ovat limittäisiä. On myös huomioitava, että tviitin saajien seuraajajoukko voi olla todellisuudessa pienempi, sillä käyttäjät voivat saada saman tviitin retviitattuna myös muilta seuraajiltaan, eivätkä kaikki käyttäjät välttämättä edes lue kaikkia saamiaan tviittejä. Näin ollen yleensä pätee  $p_i < 1$ . Pienin mahdollinen arvo saadaan, kun lasketaan niiden seuraajien osuus, jotka retviittasivat tviitin. Tietyn  $p_i$ :n yläraja voitiin havaita

suorittamalla  $p_i = \frac{|F_x \cup F_{Mk1} \cup F_{Mk2} \dots \cup F_{Mki}|}{RT_{mi}}$ . Tällöin saatiin esimerkiksi poistettua eri käyttäjien yhteisten seuraajien vaikutus.  $P_i$ :n kertoimen vaikutuksen laskeminen vaatisi seuraajien keräämisen niiltä käyttäjiltä, jotka tviittasivat tai retviittasivat tviittivirrassa. (Veijalainen ym, 2015a.)

Mikäli kerroin  $p_i$  astetaan sille seuraajien osuudelle, joka retviittasi tviitin sen jokaisessa etenemisvaiheessa, voidaan saada vaikuttavuudelle sen matalin arvo. Tämä tuottaa tviitin etenemistä kuvaavan retviittipuun perusarvon ensimmäisestä retviittauksesta viimeiseen. Esimerkiksi, mikäli käyttäjä  $a$  lähettää tviitin  $Tw_1$ , joka retviitataan kymmenen käyttäjän  $a$  seuraajan ( $b_1 \dots b_{10}$ ) toimesta, sekä käyttäjän  $c_1$  toimesta, joka seuraa käyttäjää  $b_1$  eikä ketään muuta käyttäjää, voidaan tällöin retviittipuun todeta olevan arvoltaan 11. Olettaen, että tämän esimerkin käyttäjien seuraajilla ( $F_x$ ) on 100 seuraajaa, saadaan näin alkuperäisen tviitin lähettäjälle ja seuraajille kertoimet seuraavasti:  $p_a = 0,1$ ,  $p_{b1} = 0,01$  ja  $p_x = 0$ . Näin saadaan laskettua retviittipuun syvyys  $0,1 * 100 + 0,01 * 100 = 11$ . Nämä luvut voidaan laskea yhteen, jolloin saamme yhden käyttäjän vaikuttavuutta mittaavan kertoimen:  $0,1 + 0,01 + 0 + \dots + 0 = 0,11$ . Mitä pidempi tämä yhteenlaskettu summa on, sitä halukkaampia ovat käyttäjän seuraajat, seuraajien seuraajat jne. retviittaamaan tviitin. Jatkossa tähän edellä mainittuun vaikuttavuuden mittaamisen laskentakaavaan viitattaessa käytetään merkintää  $Infl\_3$ .

Vaikuttavuuden osalta laskettiin myös korkein mahdollinen arvo sen mukaan kuinka paljon tviitillä oli potentiaalisia lukijoita, jossa  $p_x = 1$  ja  $F_x = 100$ . Yksittäisen tviitin osalta voidaan käyttää seuraavaa esimerkkiä:  $1 * F_a + 1 * F_{b1} + 1 * F_{b2} + \dots + 1 * F_{b10} + 1 * F_{c1} = 100 + 10 * 100 + 100 = 1200$ . Jatkossa tähän vaikuttavuusarvon mittaamiseen viitataan merkinnällä  $Infl\_1$ . Potentiaalisten tviitin saajien lukumäärä voidaan näin ollen laskea käyttäen kerrointa  $p_i = 1$ . Tämä tosin yliarvioi vaikuttavuutta, sillä kaikki seuraajat, seuraajien seuraajat jne. eivät välttämättä lukeneet tviittiä tai retviittiä. Parametrejä  $p_i$  säätämällä tosin voidaan yrittää havaita todellinen vaikuttavuus niiden osalta, jotka oikeasti lukivat tviitin.

Jakamalla  $Infl\_1$  alkuperäisen tviitin lähettäjän seuraajilla, saatiin  $Infl\_2$ -arvo. Tällä tavoin pyrittiin löytämään retviittaavien seuraajien lukumäärän merkitys vaikuttavuuteen ilman alkuperäisen lähettäjän seuraajien määrää. Mikäli tällä tavoin saatiin korkeita arvoja, voitiin päätellä, että pääasiassa vain retweettaajat olivat tällöin halukkaita jakamaan alkuperäisen tviitin eteenpäin. Mikäli  $Infl\_2$  sai arvokseen yli 1, tarkoitti se sitä, että muut käyttäjät jakoivat tviittiä laajemmalle yleisölle nähtäväksi, kuin mitä alkuperäisen lähettäjän seuraajien lukumäärä on. Tällöin käyttäjät myös halusivat retviitata tviitin, mikä puolestaan tarkoittaa sitä, että todennäköisesti nämä käyttäjät myös itse lukivat tviitin.

### 5.3 Tviittianalyysi

Retviittipuun muodostettiin seuraajien relaatioista sekä tviittien ja retviittien aikaleimoista. Mikäli käyttäjä oli saanut retviitin useasta eri lähteestä, joiden etäisyys

poikkeaa alkuperäisestä tviitin lähettäjistä, hyödynnettiin tällöin sellaista käyttäjää, joka oli seuraajasuhteiden perusteella alkuperäistä lähettäjää lähimpänä. Tällä tavoin oli mahdollista havainnoida retviittien levinnäisyyttä verkostossa.

Alla esitettyjen algoritmien 1 ja 2 avulla rakennettiin retviittipuut. Algoritmia 1 hyödynnettiin havainnoimaan sitä, seuraako retviittaaja käyttäjää, joka lähetti alkuperäisen tviitin. Algoritmin tarkoituksena oli palauttaa arvot (retweeter\_id, retweeted\_id, 1), joista jälkimmäisin arvo kuvaa käyttäjien välistä etäisyyttä.

Algoritmi 1:

```

for tweet_id in tweets
{
  if not is_retweet(tweet_id) {
  continue
  }
  retweeter_id = obtain_retweeter_user_id(tweet_id)
  retweeted_id = obtain_retweeted_user_id(tweet_id)
  if follows(retweeter_id, retweeted_id){
  return (retweeter_id, retweeted_id, 1)
  }
}

```

Algoritmin 1 jälkeen suoritettiin Algoritmi 2, joka tarkastelee vastaako retviittipuun lukua 2, 3 vai 4.

Algorithm 2:

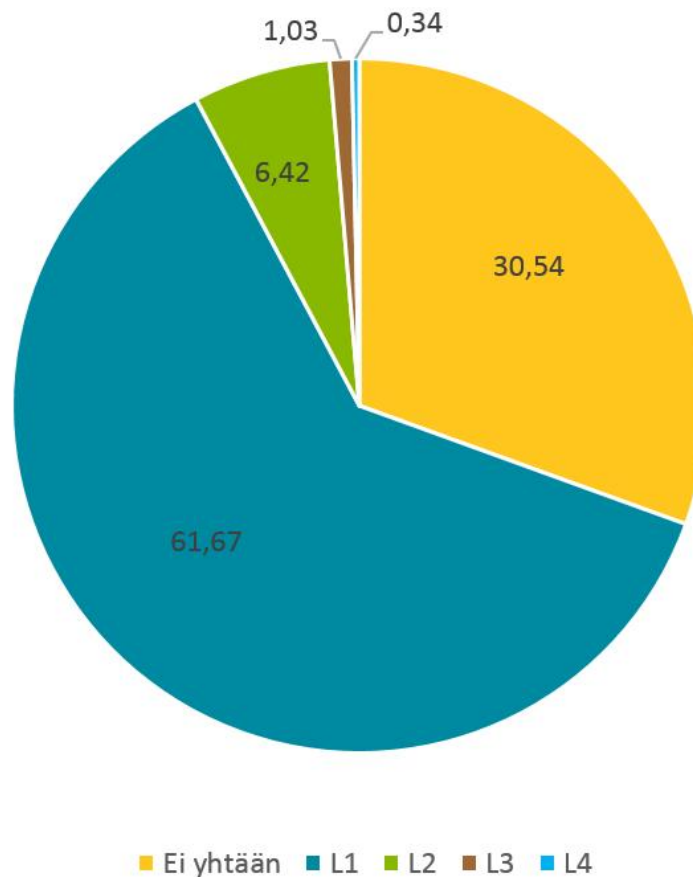
```

for distance in [2,3,4]{
  for tweet_id in retweets {
    user_ids = obtain_retweeted_user_ids(tweet_id)
    for user_id_ext in user_ids {
      for user_id_int in user_ids {
        ext_time = get_time(user_id_ext, tweet_id)
        int_time = get_time(user_id_int, tweet_id)
        if ext_time < int_time AND distance(user_id_ext, tweet_id) = distance - 1 AND distance(user_id_int, tweet_id) is NULL{
          if follows(retweeter_id, retweeted_id){
            return (retweeter_id, retweeted_id, distance)
          }
        }
      }
    }
  }
}

```

Tuloksena näistä saatiin, että 754549 retviitista 465364 retviitattiin alkuperäisen lähettäjän seuraajien toimesta. Tämän jälkeen 48405 retviittasi tviitin edelleen, jolloin puun syvyys vastasi arvoa 2. 7786 retviitattua tviittiä sai arvoksi 3 ja 2586 tviittiä sai arvoksi 4. Yhteensä 230408 tviitin osalta retviittipuun syvyyttä ei pystytty havainnoimaan kerättyjen seuraajarelaatioiden perusteella.

Tviittipuiden osalta noin 30% osalta syvyyttä ei pystytty määrittelemään, sillä merkittävältä osalta seuraajia ei saatu kerättyä seuraajatietoja. Valtaosa retviittauksista (n. 62%) olivat retviittipuun tasolla 1, eli tviitti oli retviitattu tällöin jonkun alkuperäisen lähettäjän seuraajan toimesta. Noin 6% tviiteistä olivat retviitattu alkuperäisen retviittaamisen jälkeen, noin 1% oli kolmannen tason retviittauksia, ja 0,34% neljännen tason retviittauksia. Mikäli oletamme, että retviittaamisen trendi pysyi sama myös tarkastelun ulkopuolelle jääneiden osalta, olisi 89% retviittauksista tasolla 1. Nämä retviittaukset ovat siis tällöin retviitattu esimerkiksi käyttäjän seuraajien toimesta. Samalla periaattella tasolla 2 olisi 9% retviittauksista, eli ne retviittaukset, jotka tulivat edelleen retviitatuiksi alkuperäisten retviittaajien seuraajien toimesta. Noin 1,5% retviittauksista olisi tasolla 3 jne. Nämä retviittaustasojen jakaumat havainnollistetaan kuviossa 5. Tehdyt havainnot tukevat Bakshyn ym. (2011) havaintoja siitä, että suurin osa tviiteistä ei leviä Twitterissä lainkaan tai sitten vain yhden hyppäyksen verran.



KUVIO 6 Havaittujen retviittipuiden syvyyksien osuus aineistossa

Sen retviittipuun osuuden kohdalla, jossa ei voitu tarkastella keneltä seurattavalta retviittaaja sai tviitin/retviitin, asetettiin retviitit niin lähelle retviittipuun alkua (alkuperäinen lähettäjä) kuin mahdollista. Mikäli retviittipuun koko oli

suurempi kuin tviitin alkuperäisen lähettäjän seuraajien lukumäärä, retviitit asetettiin tällöin yhdelle tai useammalle alkuperäisen lähettäjän seuraajalle. Myöhemmässä vaiheessa retviitanneiden käyttäjien voitiin olettaa seuraavan alkuperäistä retviittaajaa, mikäli tällä oli paljon omia seuraajia ja retviittaamiskäyttäytyminen jatkui voimakkaana käyttäjän retviittaamisen jälkeen. Näin ollen retviitatut tviitit liitettiin sen käyttäjän retviittiin, jolla oli eniten seuraajia ja mikäli aikaleima puolsi tätä. Nämä paljon seuraajia omaavat retviittaajat ovat useimmiten retviittipuun tasolla 1. Näin ollen retviittipuun rakentaa ilman, että tiedettiin tarkasti seuraajien suhteita toisiinsa tasoilla 1 ja 2. Tämän ei pitäisi vääristää liiaksi tarkastelun kohteena olleesta retviittipuusta tehtyjä havaintoja.

## 5.4 Vaikuttavuusanalyysi

Vaikuttavuutta tarkasteltiin aikaisemmin esiteltyjen mittaustapojen sekä seuraajien lukumäärän perusteella. Infl\_1:n avulla tarkasteltiin retviittipuun seuraajien lukumäärää, jossa  $\pi=1$  oli sijoitettu kaikkialle aineistoon. Tällöin käyttäjillä, joilla oli eniten seuraajia, oli myös eniten vaikuttavuutta. Aineiston mediaaniarvo Infl\_1:llä tarkasteltaessa oli 490, joka merkitsee mediaania niiden käyttäjien kohdalla, jotka tviitti mahdollisesti tavoitti. Keskiarvoksi kaikkien alkuperäisten noin 750000 tviitin kohdalla voitiin todeta olevan 16680. Eniten vaikuttavuutta tällä tavalla mitattuna on listattu taulukkoon 2.

TAULUKKO 2 Käyttäjien vaikuttavuus seuraajien perusteella

Käyttäjä	käyttäjä_id	infl_1
justinbieber	27260086	138500607
AP	51241574	91785137
Harry_Styles	181561712	63905937
ReutersUS	15 108 530	60064143
Louis_Tomlinson	84279963	44351359
GGPolitics	24138795	43784429
coopnytimes	65431946	43709053
AC360	227 837 742	42989109
BBCNewsUS	69 329 527	41658888
MichaelSkolnik	24165761	41237339
ddlovato	21111883	40600958
selenagomez	23375688	33988871
piersmorgan	216299334	26999580
CNNSitRoom	34310801	26474445
caketin85	18933433	26133302
GlobalGrind	1496971	25923685
ABC	28785486	25637212
HuffPostCrime	317393966	25567659
HuffPostParents	16581734	25482959

Tulos muuttui, mikäli vaikuttavuutta skaalattiin alkuperäisen lähettäjän seuraajien lukumäärän mukaisesti hyödyntäen Infl\_2:sta. Tällöin eniten vaikuttavuutta sai käyttäjätunnus ClarkeEtc041, jolla oli tarkasteluhetkellä vain yksi seuraaja, vaikuttavuusarvon ollessa tällöin kuitenkin 389689. Tämä johtuu siitä, että kyseisen käyttäjän lähettämä tviitti retviitattiin yhteensä neljä kertaa sellaisten käyttäjien toimesta, joilla itsellään oli miljoonia seuraajia. Näin tarkasteltaessa sellaiset käyttäjät nousivat vaikuttavuuden osalta muiden yläpuolelle, joiden seuraajilla tai myöhemmin seuraajaketjussa olleilla käyttäjillä oli itsellään paljon seuraajia. Sellaiset käyttäjät, joilla oli paljon seuraajia, kuten esimerkiksi julkisuuden henkilöt ja erilaiset mediayhtiöt, vajosivat puolestaan tässä vaikuttavuustilastoissa alaspäin. Eniten vaikuttavuutta omaavat käyttäjät tällä tavalla mitattuna on esitetty taulukossa 3. Mediaaniarvoksi tässä kohdin saatiin 0,85, mikä tarkoittaa sitä, että käyttäjä tavoittaa retviittien kautta periaatteessa vähemmän muuta yleisöä kuin mitä tällä itsellään on seuraajia. Huomioitavaa on, että tutkimusaineistosta havaitut 4,35 miljoonaa retviittausta pohjautuvat ainoastaan 0,75 miljoonaan alkuperäiseen tviittiin. Tämä tarkoittaa sitä, että keskimäärin yksi tviittaus retviitattiin 6 kertaa. Varianssi on tältä osin laaja, sillä tällaisen perusanalyysin pohjalta voitiin nähdä, että vain 13%:lla käyttäjistä vaikuttavuus oli kasvanut, kun mittarina käytettiin retviitattuja tviittejä ja mikäli vaikuttavuus pohjautui heidän seuraajiensa lukumäärään.



TAULUKKO 3 Käyttäjien vaikuttavuus retviittausten perusteella

Käyttäjä	käyttäjä_id	infl_2	Seuraajien lkm
ClarkeEtc041	1005135091	389689	1
aakashdeeparora	404020377	341710	4
rafaqatgohar	106923024	169850	4
caketin85	18933433	89192.2	293
VineTheMagazine	220850478	53927.3	10
angiewatson36	1327309309	29985	1
ccamia	45445240	28781.4	274
iEndorseYouHoz	79702155	18903.8	875
TuuliaHokkanen	1261733839	17156	28
ziapappi	19136376	15320	600
ReginaPraga	1354896439	11781.2	12
KathllenHonda	939217482	11101.6	11
DuronJonathan	142237961	9903.94	145
NuuYourkCo	1319678292	9519	2
AaronKatersky	125358141	9292.65	1859
jdawsey1	38936142	8740.82	1924
madkar23	630209606	7427.55	31
ATXYouthAcademy	1262204772	6891.78	9
edithhonan	257168815	6864.77	945
DrGhettella	489570995	6546.04	149

Jokainen retviitattu tviitti mitattiin hyödyntämällä Infl\_3:a, jolloin laskettiin arvo jokaiselle aliketjulle, joka sijoittui sen alkupisteestä katsottuna kohtaan 1, 2, 3 tai 4. Tämän jälkeen arvot laskettiin juuresta alkaen yhteen. Tarkasteluaineiston osalta keskiarvoksi saatiin 0,00767 ja vain 91 puun juurta sai arvokseen 1 tai enemmän. Taulukossa 4 voidaan nähdä vaikuttavimmat käyttäjät tällä tavalla mitattuna. Esimerkiksi Justin Bieber (käyttäjä\_id: 27260086) sekä muut useita seuraajia omaavat käyttäjät sijoittuvat korkealle vaikuttavuuden osalta tällä tavalla tarkasteltuna, ja vain yksi käyttäjä, jolla oli vain kolme seuraajaa, on mukana taulukossa. Tällä tavoin pystymme toteamaan mittauksen korreloivan positiivisesti retviittipuun koon kanssa, joka perustuu alipuiden laskettuihin arvoihin.

TAULUKKO 4 Käyttäjien vaikuttavuus retviittipuiden arvon mukaan

Tviittivirta	käyttäjä_id	Seuraajien lkm	infl_3
6352	582907105	1119231	6.14957
26124	829673054	46490	18.3443
49830	84279963	10087488	8.90494
24771	21111883	13300606	7.64086
161	1160668878	3	18.3443
132	256158353	107008	3.54796
11165	18753325	816322	3.94178
8147	40927173	1151655	3.89011
89337	27260086	58529773	31.8984
17307	1156695804	3263	4.45337

## 5.5 Johtopäätökset

Tässä kohdassa on esitelty tutkimuskysymykset, joihin on annettu vastaukset tieteellisen kirjallisuuden ja empiirisen osuuden perusteella, sekä tehty tarvittavat johtopäätökset. Tutkimuksessa on pyritty selvittämään sitä, miten vaikuttavuus näyttäytyy ja muodostuu sosiaalisissa verkostoissa informaation diffuusion vaikutuksesta. Tämän tutkimuksen keskeisin tutkimuskysymys oli seuraava: *Mikä on informaation diffuusion merkitys käyttäjän vaikuttavuuden muodostumiselle Twitterissä?* Päätutkimuskysymyksen rinnalle asetettiin seuraavat avustavat tutkimuskysymykset:

- *Miten sosiaalinen vaikuttavuus syntyy sosiaalisissa verkoissa?*
- *Voidaanko vaikuttavuus Twitterissä yksilöidä käyttäjäkohtaisesti siten, että joku käyttäjä on vaikuttaja ja joku ei?*

Tutkimuksessa pyrittiin ensinnäkin selvittämään mitä vaikuttavuus on ja miten se näyttäytyy sosiaalisissa verkostoissa, joita myös erilaiset sosiaalisen verkoston sivustot eli yhteisöpalvelut edustavat. Kirjallisuuden pohjalta vaikuttavuuden voidaan sanoa noudattavan potenssilakia: vaikuttajia on pieni joukko massasta ja samalla on suuri osuus niitä henkilöitä, joilla on vain vähäinen vaikuttavuus, kun tarkastellaan koko henkilöiden muodostamaa sosiaalista verkostoa. (Junquero-Trabado ym., 2011.) Vaikka kirjallisuudessa vaikuttajista puhuttaessa ei yleensä tarkoiteta esimerkiksi julkisuuden henkilöitä vaan ”tavallisia” yksilöitä, voidaan tämän tapaustutkimuksen tulosten pohjalta todeta, että Twitterissä vaikuttajista puhuttaessa huomio kuitenkin kiinnittyy nimenomaan niihin yksilöihin, jotka ovat keskiössä omassa sosiaalisessa verkostossaan omaten paljon seuraajia. Twitterissä tällaisia käyttäjiä, joilla on paljon seuraajia, ovat yleensä nimenomaan julkisuuden henkilöt.

Tarkasteltaessa vaikuttavuuden syntymistä, voidaan käyttäjien keskinäisellä viestintäprosessilla nähdä olevan siinä erityinen rooli. Yksilöt luovat sekä jakavat informaatiota vuorovaikuttaen keskenään oman verkostonsa sisällä, ja heikkojen yhteyksien kautta viestit välittyvät myös tehokkaasti verkostojen ulkopuolelle. (Rogers, 2003, s. 5-6). Tällaista informaation leviämistä kutsutaan monissa yhteyksissä informaation diffuusioksi. Verkostoissa solmuina kuvattavien käyttäjien voidaan nähdä olevan joko aktiivisena tai ei-aktiivisena riippuen siitä onko informaatio tavoittanut tämän. Mikäli solmu on aktiivinen, on se silloin myös kykenevä levittämään saamaansa informaatiota eteenpäin. (Yang & Leskovec, 2010.) Twitterissä käyttäjän lukiessa seurattavansa lähettämän tviitin, voidaan käyttäjää kuvaavan solmun tällöin todeta aktivoituneen. On myös mahdollista, että käyttäjä retviittaa saamansa tviitin eteenpäin omille seuraajilleen, jolloin tarjoutuu mahdollisuus ketjureaktiolle, jossa retviitti leviää käyttäjien keskuudessa ja näin aktivoi laajasti muita käyttäjiä verkostossa. Vaikuttajilla on tässä kohdin erityinen rooli, sillä heidän kauttaan on koettu mahdolliseksi saavuttaa tällaisia ketjureaktiota informaation diffuusion keinoin. (Rogers, 2003, s.26-28.)

Tapaustutkimuksen avulla oli tarkoitus myös selvittää onko tällaisen vaikuttavuuden syntyminen vain sattumaa käyttäjästä riippumatta vai voidaanko havaita, että tietyillä Twitterin käyttäjillä on enemmän vaikuttavuutta kuin toisilla. Tähän liittyen kirjallisuuden kautta esitettiin kaksi vastakkaista teoriaa, joista Gladwell (2002) esitti, että tietyillä henkilöillä on sellaisia erityisiä ominaisuuksia, joiden vuoksi heillä on merkityksellisesti vaikuttavuutta verrattuna muuhun populaatioon. Hänen mukaansa nämä vaikuttajat pystyisivät ominaisuuksiensa myötä aikaansaamaan eräänlaisia ”sosiaalisia epidemioita”, jotka voidaan rinnastaa laajamittaiseen informaation diffuusioon.

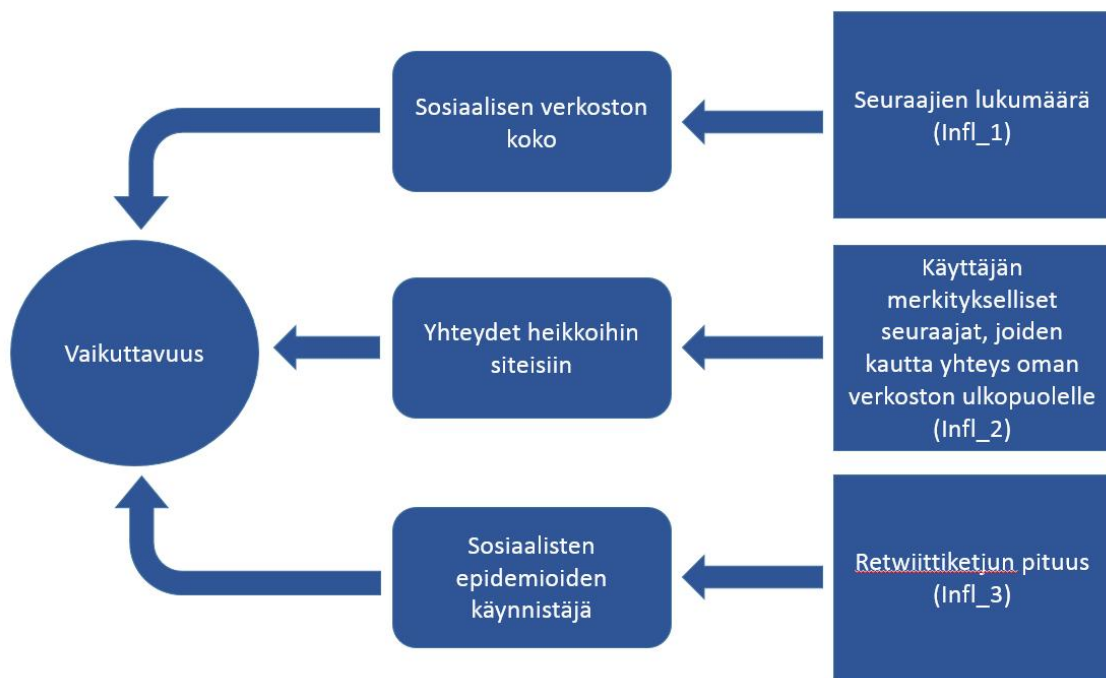
Watts (2007) on puolestaan esittänyt vastakkaisen näkemyksen vaikuttavuudesta. Hänen mukaansa vallitsevilla olosuhteilla olisi enemmän merkitystä kuin tietyillä yksilöillä, ja tästä johtuen esimerkiksi tämän tutkimuksen keskiössä olevaa informaation diffuusiota ei voitaisi ennustaa tarkastelemalla pelkästään sen alullepanijaa. Tällöin kyse olisi pikemminkin sattumasta, joka tapahtumana olisi sattumanvaraisesti riippuvainen ajasta ja paikasta ilman, että tietyillä yksilöillä olisi tässä kohdin ennustettavaa merkitystä. Twitteriä tarkasteltaessa on tosin havaittu informaation siirtyvän tiettyjen välikäsien kautta suurille massoille (Wu ym., 2011), jolloin myös vaikuttajat olisivat ennalta-arvattavampia kuin mitä Watts esittää. Myös tämän tutkimuksen tulokset tukevat tätä selitysmallia: kun vaikuttavuutta skaalattiin alkuperäisen lähettäjän seuraajien lukumäärän mukaisesti, välittyi informaatio tällöin laajemmalle seuraajajoukolle nimenomaan sellaisten käyttäjien toimesta, joilla itsellään oli paljon seuraajia.

Tämän tapaustutkimuksen tulokset tukevat myös kirjallisuuskatsauksesta esille nousseita havaintoja siltä osin, että käyttäjän seuraajien määrällä voidaan nähdä olevan vaikuttavuuden muodostumisessa merkityksellinen rooli. Seuraajat luovat käyttäjän lähettämille tviiteille yleisön, jolloin myös potentiaalisten retviittausten mahdollisuus kasvaa. Mikäli tviitti tulee retviitatuksi, on myös

mahdollista, että retviittipuu pitenee ja leviää muiden käyttäjien seuraajien vaikutuksesta, jolloin myös käyttäjän vaikuttavuus kasvaa.

Toisaalta tutkimusaineistosta havaittiin myös, että pelkästään seuraajien lukumäärä ei ole merkittävin tekijä vaikuttavuuden syntymisessä vaan se, ketkä seuraavat käyttäjää ja ketkä retviittaavat tämän lähettämiä tviittejä eteenpäin. Vaikka alkuperäisen tviitin lähettäjällä ei itsellään ole kuin vähän seuraajia, mutta tviitti tulee retviitatuksi sellaisten käyttäjien toimesta, joilla itsellään on paljon seuraajia, on tviitillä tällöin mahdollisuus saada sellainen yleisö Twitterissä, jota sillä ei alkuperäisen tviitin lähettäjän omien seuraajien perusteella vielä ollut. Tällaisella käyttäjällä voidaan todeta olevan tässä kohdin huomattavasti enemmän vaikuttavuutta verrattuna sellaisiin käyttäjiin, jolla muutoin on enemmän seuraajia. Tämä myös edellyttää, että ne vähäiset käyttäjän seuraajat huomaavat ja lukevat tviitin, mikä ei välttämättä ole suuren tviittivirran seasta aina oletusarvoista. Tviitin kiinnostavuuteen sen lukijan näkökulmasta eikä seuraajien lukumäärään ole käyttäjällä juurikaan vaikutusmahdollisuutta. Näin ollen voidaan tässä kohdin puhua myös Wattsin mainitsemasta sattumavaikuttajasta.

Tämän tutkimuksen pohjalta voidaan päätutkimuskysymykseen vastattaessa todeta informaation diffuusiolla ja sen osatekijöillä olevan merkityksellinen rooli vaikuttavuuden muodostumisessa. Gladwellin (2002) mukaan vaikuttajat ovat motivoituneita informaation välittäjiä, joka Twitterissä ilmenee siten, että käyttäjä yleensäkin päättää retviitata saamansa tviitin muille käyttäjille. He kokevat saamansa informaation merkitykselliseksi, joka tällöin motivoi heitä jakamaan tietosisällön eteenpäin. Heillä on myös erityisen paljon merkityksellisiä yhteyksiä verkostossaan. Tämä näyttäytyy yleensä seuraajien lukumäärän paljoutena, mutta myös siten, että seuraajien joukosta löytyy niitä merkityksellisiä käyttäjiä, jotka pystyvät jakamaan tviitit eteenpäin aina vain laajemmalle yleisölle. Toisaalta vaikuttavuuden on todettu syntyvän vastavuoroisen prosessin kautta, joka puolestaan vaatii käyttäjän seuraajien aktiivista osallistumista informaation diffuusioprosessiin retviittaamalla käyttäjän lähettämiä tviittejä. Nämä vaikuttavuuden muodostumisen kannalta keskeiset tekijät operationalisoitiin tämän tutkimuksen teoriaosuuden luvussa 3.2.1 ja ne on tutkimustuloksista tehtyjen havaintojen pohjalta havainnollistettu kuviossa 7. Kuvioista ilmenee, että käyttäjän vaikuttavuuden muodostumiseen vaikuttaa henkilön seuraajien määrä sekä yhteydet merkityksellisiin heikkoihin siteisiin. Vaikuttavuus kuitenkin syntyy vastavuoroisen prosessin tuloksena, jolloin seuraajien retviittauksista muodostuvien retviittipuiden syvyydellä voidaan nähdä olevan vaikuttavuuden muodostumisen kannalta olennainen merkitys. Kuvioon on myös merkitty jokaisen vaikuttavuuden muodostumiseen vaikuttavan tekijän kohdalle ne mittaamistavat, joiden avulla vaikuttavuutta tässä tutkimuksessa mitattiin. Näitä mittaamistapoja olivat Infl\_1, infl\_2 ja Infl\_3.



KUVIO 7 Vaikuttavuuden muodostumiseen osatekijät

On kuitenkin huomioitava, että tässä tutkimuksessa havaitut tulokset saattaisivat poiketa merkittävästi, mikäli vastaavanlainen aineisto kerättäisiin eri ajan-kohtana sekä eri asiayhteydessä. Varsinkin ne käyttäjät, jotka tämän tutkimuksen aineiston havaintojen pohjalta ovat vaikuttajia, eivät välttämättä ole vaikuttajia muissa yhteyksissä, sillä vain vähän seuraajia omaavien käyttäjien yksittäiset tviitit eivät välttämättä tule samalla tavalla retviitatuksi jokaisessa havaintoaineistossa. Havaitut tulokset kuitenkin tukevat osaltaan sitä tosiasiaa, että ne Twitterin käyttäjät, jotka ovat oman sosiaalisen verkostonsa keskiössä aktiivisten retviittaajien ympäröiminä, ovat myös niitä todennäköisiä vaikuttajia.

## 6 YHTEENVETO

Tässä tutkimuksessa pyrittiin selvittämään aikaisempiin tutkimuksiin pohjautuvan tieteellisen kirjallisuuskatsauksen sekä empiirisen tapaustutkimuksen avulla, mikä merkitys informaation diffuusiolla on Twitterissä käyttäjien vaikuttavuuden muodostumisessa. Lisäksi tutkimuksen tarkoituksena oli selvittää, voidaanko vaikuttavuutta käyttäjien välillä yleensäkin luotettavasti mitata ja onko vaikuttavuus jollain tavoin ennustettavaa, kun tarkastellaan käyttäjien välistä vuorovaikutusta sosiaalisessa verkostoissa. Tässä tutkimuksessa vaikuttavuutta ja informaation diffuusiota sosiaalisissa verkostossa tarkasteltiin tieteellisen kirjallisuuskatsauksen kautta ja tällä tavoin pyrittiin saamaan vaikuttavuuden määritelmä mitattavaan muotoon empiiristä tutkimusosiota varten. Tämän tutkimuksen empiirisessä osuudessa tutkittiin huhtikuussa 2013 Bostonin maratonilla tapahtuneisiin pommi-iskuihin liittyviä tviittipäivityksiä ja retviittauksia. Aineisto käsitti yli kahdeksan miljoonaa tviittiä ja retviittiä, jotka oli lähetetty yli neljän miljoonan Twitter-käyttäjän toimesta tapahtumaan liittyen. Tutkimusaineiston kautta pyrittiin havaitsemaan sitä, kuinka informaatio liikkuu sosiaalisessa verkostossa käyttäjältä toiselle, sekä tarkastelemaan, kuinka vaikuttavuus muodostuu informaation diffuusion kautta käyttäjäkohtaisesti.

Twitter on yksi suosituimmista verkkoyhteisöpalveluista, jossa käyttäjät lähettävät seuraajilleen lyhyitä tekstipohjaisia tilapäivityksiä eli tviittejä pyrkien palvelun itsensä mukaan kertomaan mitä käyttäjä sillä hetkellä tekee tai mitä sillä hetkellä tapahtuu. Informaation diffuusio ilmenee Twitterissä muiden käyttäjien retviitatessa käyttäjän lähettämiä tviittejä omille seuraajilleen. Tviittien leviäminen verkostossa noudattaa kuitenkin potenssilakia, eli vain harvat tviitit tulevat retviitatuksi ja leviävät laajasti käyttäjältä toiselle useamman hyppäyksen verran. Tviitit kuitenkin leviävät nopeasti käyttäjältä toiselle sosiaalisessa verkostossa, mikäli tviitti tulee retviitatuksi ensimmäisen kerran. Retviitattu tviitti voi saavuttaa hetkessä lukuisan käyttäjäjoukon riippumatta siitä, kuinka paljon alkuperäisen tviitin lähettäjällä on seuraajia. Tällöin tviitin potentiaalisten lukijoiden määrän voidaan nähdä kasvavan nopeasti jokaisen retviittauksen vaikutuksesta, eikä käyttäjän omien seuraajien määrästä voida päätellä mikä lähetetyn tviitin todellinen yleisö on. Tässä myös piilee retviittauksen teho informaation diffuusion näkökulmasta: periaatteessa jokaisella käyttäjällä on mahdollisuus olla vaikuttaja ja potentiaalinen ”sosiaalisen epidemian” aiheuttaja. Twitterissä ”sosiaalinen epidemia” kuitenkin vaatii vertaiskäyttäjien aktiivista retviittaamista, jolloin vaikuttavuus voidaan nähdä vastavuoroisen prosessin kautta tapahtuva ilmiönä.

Twitterissä käyttäjien muodostama verkosto on asymmetrinen, jolloin käyttäjän tviittien seuraaminen ei vaadi vastavuoroisuutta tai velvoitetta seurata takaisin. Tästä johtuen retviittaaminen voidaan nähdä tehokkaaksi silloin, kun käyttäjä on sosiaalisen verkostonsa keskiössä omaten paljon seuraajia, jolloin hän täyttää verkossa olevia rakenteellisia aukkoja eri verkoston tiiviiden käyttäjäryhmien eli klusterien välillä. Tällöin käyttäjä toimii samalla myös eräänlaisena heik-

kona siteenä eri klusterien välillä, jolloin heillä on mahdollisuus siirtää informaatiota niiden verkostojen kesken, joihin hän itse vaikuttaa. Käyttäjän omien seuraajien määrä ei kuitenkaan vielä tee käyttäjästä vaikuttajaa vaan vaikuttavuuden muodostuminen vaatii myös vertaiskäyttäjien aktiivisuutta.

Vaikuttavuutta sosiaalisissa verkostoissa on tutkittu erilaisten mallintamistapojen kautta, joissa informaation diffuusio on prosessina riippuvainen vertaiskäyttäjien päätöksistä. Tällöin vaikuttavuus määräytyy sen perusteella, miten verkostossa solmuina kuvatut käyttäjät aktivoituvat ja näin omaksuvat saamaansa informaatiota. Twitterissä tämä ilmenee siten, että vertaiskäyttäjät päättävät retviitata saamansa tviitin eteenpäin omille seuraajilleen. Twitterissä vaikuttavuuden mittaamista on pyritty tarkastelemaan erilaisten mittareina toimivien indikaattorien avulla. Esimerkiksi käyttäjien lähettämällä tviiteillä, muiden käyttäjien retviiteillä sekä seuraajien lukumäärällä on nähty olevan merkitystä vaikuttavuuden muodostumisen osalta.

Myös tämän tapaustutkimuksen kautta nousi esille, että sellaiset käyttäjät, joilla on paljon seuraajia, ovat usein myös todennäköisiä vaikuttajia. Pelkästään seuraajien määrä ei kuitenkaan määrittele vaikuttajaa vaan vaikuttavuus on pitkälti riippuvainen aktiivisista seuraajista, jotka retviittaavat käyttäjän lähettämiä tviittejä eteenpäin ja joilla itsellään on paljon seuraajia. Tällöin yleensä julkisuuden henkilöt nousevat vaikuttavuustilastoissa korkealle seuraajien lukumäärän myötä.

Tutkimustuloksista pystyttiin havaitsemaan, että myös sattumalla on oma osuutensa, kun yksittäinen tviitti saattaa levitä sosiaalisessa verkostossa vain muutaman merkityksellisen käyttäjän retviitausten kautta. Tällaisissa tapauksissa on käytetty termiä sattumavaikuttaja, kun diffuusioprosessia ei voida enustaa pelkästään sen alullepanijasta vaan se on pikemminkin sattumanvaraisesti tapahtuva ilmiö. Tällaisten sattumavaikuttajien kohdalla vaikuttavuus saattaa vaihdella havainnonitijaksosta riippuen. Mikäli vastaavanlainen aineisto kerättäisiin eri ajankohtana sekä eri asiayhteydessä, saattaisi myös havaitut tulokset poiketa merkittävästä tässä tutkimuksessa havaituista tuloksista, eivätkä tällöin nämä samat käyttäjät välttämättä olisi vaikuttajia muissa yhteyksissä. Tämä vuoksi voidaan olettaa, että tämän tutkimuksen kautta saadut havainnot eivät välttämättä ole yleistettävissä. Tutkimuksesta tehdyt havainnot tukevat kuitenkin osaltaan sitä tosiasiaa, että ne Twitterin käyttäjät, jotka ovat oman sosiaalisen verkostonsa keskiössä aktiivisten retviittaajien ympäröiminä, ovat myös todennäköisiä vaikuttajia.

Jatkotutkimukset vaikuttavuuden muodostumisesta informaation diffuusion seurauksena voisivat keskittyä tviittisisältöjen merkitykseen, jota tässä tutkimuksessa ei otettu vaikuttavuuden muodostumiseen vaikuttavana indikaattorina huomioon. Tämän tutkimuksen osalta empiirinen aineisto käsitti viiden päivän aikana lähetetyt tviitit liittyen yksittäiseen reaali maailman tapahtumaan. Mielenkiintoista olisi tarkastella havaintoja, jossa analysoitavana on ajallisesti laajamittaisempi aineisto. Tällöin myös olisi mahdollisuus pienentää sattumavaikuttajien osuutta tutkimustuloksissa.

## LÄHTEET

- Abdullah, S. & Wu, S. (2011). An epidemic model for news spreading on Twitter. Teoksessa 23rd IEEE International Conference on Tools with Artificial Intelligence (s. 163-169). Boca Raton, FL, November 7-9, 2011. IEEE Computer Society.
- Acemoglu, D., Ozdaglar, A. & Yildiz, E. (2011). Diffusion of innovations in social networks. Teoksessa 2011 50th IEEE Conference on Decision and Control and European Control Conference (s. 2329-2334). Orlando, Florida, December 12-15, 2011.
- ATK-sanakirja (2008). Termit, määritelmät ja vastineet eri kielillä. (14. painos) Helsinki: Talentum
- Bacile, T.J. (2013). The Klout challenge: Preparing your students for social media marketing. *Marketing Education Review*, 23(1), 87-92.
- Backstrom, L., Boldi, P., Rosa, M., Ugander, J. & Vigna, S. (2012). Four degrees of separation. Teoksessa Proceedings of the 3rd Annual ACM Web Science Conference (WebSci'12) (s.33-42). Evanston, IL, USA, June 22-24, 2012.
- Bakshy, E., Hofman, J.M., Mason, W.A. & Watts, D.J. (2011). Everyone's an influencer: quantifying influence on Twitter. Teoksessa Proceedings of the fourth ACM international conference on Web search and data mining (WSDM'11) (s.65-74). Hong Kong, China, February 9-12, 2011.
- Banerjee, N., Chkraborty, D., Dasgupta, K., Mittal, S., Joshi, A., Nagar, S., Rai, A. & Madan, S. (2009). User interests in social media sites: an exploration with micro-blogs. Teoksessa Proceedings of the 18th ACM conference on Information and knowledge management (s. 1823-1826). Hong Kong, China, November 2-6, 2009.
- Berry, J., and Keller, E. (2003). *The Influentials: One American in Ten Tells the Other Nine How to Vote, Where to Eat, and What to Buy*. Free Press.
- Bhattacharya, D. & Ram, S. (2012). Sharing news articles using 140 characters: A diffusion analysis on Twitter. Teoksessa 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (966-971). Istanbul, Turkey, August 26-29, 2012.
- Boase, J., & Wellman, B. (2001). A Plague of Viruses: Biological, Computer and Marketing. *Current Sociology*, 49(6), 39-55.
- Bolton, R.U., Pararusaman, A., Hoefnagels, A., Migchels, N., Kabadayi, S., Gruber, T., Komarova Loureiro, Y. & Solnet, D. (2013). Understanding generation Y and their use of social media a review and research agenda. *Journal of Service Management*, 24(3), 245-267.
- Boyd, D., Golder, S. & Lotan, G. (2010). Tweet, tweet, retweet: conversational aspects of retweeting on Twitter. Teoksessa Proceedings of the 2010 43rd Hawaii International Conference on System Sciences (s.1-10). Washington DC: IEEE Computer Society.



- Bettencourt, L.M.A., Cintron-Arias, A., Kaiser, D.I. & Castillo-Chavez, C. (2006). The power of a good idea: quantitative modeling of the spread of ideas from epidemiological models, *Physica A: Statistical Mechanics and its Applications*, 364, 513-536.
- Brown, E.B. & Feng, J. (2011) Measuring User Influence on Twitter Using Modified K-Shell Decomposition. Teoksessa Fifth International AAAI Conference on Weblogs and Social Media. Barcelona, Catalonia, Spain, July 21, 2011.
- Budak, C., Agrawal, D. & El Abbadi, A. (2012). Diffusion of information in social networks: is it all local? Teoksessa IEEE International Conference on data mining (s. 121-130). Brussels, Belgium, December 10-13, 2012. IEEE Computer Society.
- Burt, R.S. (2004). Structural Holes and Good Ideas. *The American Journal of Sociology*, 110(2), 349-399. 32
- Capurro, R. & Hjørland, B. (2003), The concept of information. Teoksessa B. Cronin (toim.), *Annual Review of Information Science and Technology*, 37(1) (s.343-411). Information Today Inc.
- Cha, M., Haddadi, H., Benevenuto, F. & Gummadi, K.P. (2010). Measuring User Influence in Twitter. The Million Follower Fallacy. Teoksessa Proceedings of the 4th International AAAI Conference on Weblogs and Social Media. Washington DC, USA, May 23-26, 2010.
- Chang, H.-C. (2010). A new perspective on Twitter hashtag use: diffusion of innovation theory. *Proceedings of the American Society for Information Science and Technology*, 47(1), 1-4.
- De Choudhury, M. (2011). Tie formation on Twitter: Homophily and structure of egocentric networks. Teoksessa 2011 IEEE International Conference on Privacy, Security, Risk, and Trust, and IEEE International Conference on Social Computing, (s. 465-479). Boston, USA, October 9-11, 2011. IEEE Computer Society.
- De Choudhury, M., Sundaram, H., John, A., Seligmann, D.D. & Kelliher, A. (2010). "Birds of a Feather": Does User Homophily Impact Information Diffusion in Social Media. *ArXiv*, 1006(4).
- Dretske, F. I. (1986). Minds, machines and meaning. Teoksessa C. Mitcham, & A. Huning (toim.), *Philosophy and technology II: Information technology and computers in theory and practice* (97-109). Dordrecht, The Netherlands: Reidel.
- Feitelson, D.G., Frachtenberg, E. & Beck, K.L. (2013). Development and deployment at Facebook. *IEEE Internet Computing*, 17(4), 8-7.
- Gladwell, M. (2002). *The tipping point: How little things can make big difference*. New York: Back Bay Books.
- Goldenberg, J., Han, S., Lehman, D.R. & Hong, J.W. (2009). The role of hubs in the adoption process. *Journal of marketing*, 73, 1-13.
- Goldenberg, J., Libai, B. & Muller, E. (2001). Talk of the network: A complex systems look at the underlying process of word-of-mouth. *Marketing Letters*, 12(3), 211-223.

- Golder, S. & Yardi, S. (2010). Structural predictors of tie formation in Twitter: transitivity and mutuality. Teoksessa Proceedings of the 2010 IEEE Second International Conference on Social Computing (s. 88-95). Washington DC: IEEE Computer Society.
- Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6), 1360-1380.
- Granovetter, M. (1978). Threshold Models of Collective Behavior. *American Journal of Sociology*, 83(6), 1420-1443.
- Guille, A. & Hacid, H. (2012). A predictive model for the temporal dynamics of information diffusion in online social networks. Teoksessa 21st international conference companion on World Wide Web (s. 1145-1152). Lyon, France, April 16-20, 2012.
- Gruhl, D., Guha, R., Liben-Nowell D. & Tomkins, A. (2004). Information diffusion through blogspace. Teoksessa World Wide Web (s.491-501). New York, USA, May 17-22, 2004.
- Hirsjärvi, S., Remes, P. & Sajavaara, P. (2007). Tutki ja kirjoita. (13. uud. painos). Keuruu: Otavan Kirjapaino Oy.
- Iribarren, J.L. & Moro, E. (2011). Affinity paths and information diffusion in social networks. *Social Networks*, 33(2), 134-142.
- Jackson, M.O. & Leeat, Y. (2005). Diffusion on social networks. *Économie Publique*, 16(1), 3-16. 33
- Java, A., Song, S., Finin, T. & Belle, T. (2007). Why we twitter: understanding microblogging usage and communities. Teoksessa Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis (s. 56-65). San Jose, CA, August 12, 2007.
- Junquero-Trabado, V., Trench-Ribes, N., Aguila-Lorente, M.A. & Dominguez-Sal, D. (2011). Comparison of influence metrics in information diffusion networks. Teoksessa Computational Aspects of Social Networks (CASoN) (s. 31-36). Salamanca, Spain, October 19-21, 2011.
- Järvinen, P. & Järvinen, A. (2004). Tutkimustyön metodeista. Tampere: Opinpajan Kirja.
- Kaplan, A.M. & Haenlein, M. (2011). The early bird catches the news: Nine things you should know about micro-blogging. *Business Horizons*, 55(1), 105-113.
- Katz, E., & Lazarsfeld, P. (1955). Personal influence: The part played by people in the flow of mass communication. Glencoe, IL: Free Press.
- Kawamoto, T. (2013). A stochastic model of the tweet diffusion on the Twitter network. *Physica A: Statistical Mechanics and its Applications*, 392(16), 3470-3475.
- Keller, E. & Berry, J. (2003). The influentials: One American in ten tells the other nine how to vote, where to eat, and what to buy. New York: Free Press.
- Kempe, D., Kleinberg, J. & Tardos, E. (2003). Maximizing the spread of influence through a social network. Teoksessa 9th ACM SIGKDD international conference on Knowledge discovery and data mining (s. 137-146). Washington DC, USA, August 24-27, 2003.

- Kimura, M., Saito K. & Nakano, R. (2007). Extracting influential nodes for Information diffusion on a social networks. Teoksessa 22nd national conference on Artificial intelligence (s. 1371-1376). Vancouver, Canada, July 22-26, 2007.
- Kwak, H., Lee, C., Park, H. & Moon, S. (2010). What is Twitter, a social network or news media? Teoksessa Proceedings of the 19th international conference on World wide web (s. 591-600). Raleigh, North Carolina, April 26-30, 2010.
- Kwon, J. & Han, I. (2013). Information diffusion with content crossover in online social media: an empirical analysis of the social media transmission process in Twitter. Teoksessa 46th Hawaii International Conference on System Sciences (s. 3292-3301). Wailea, Hawaii, January 7-10, 2013.
- Lee, C., Kwak, H., Park, H. & Moon, S. (2010). Finding influentials based on the temporal order of information adoption in Twitter. Teoksessa Proceedings of the 19th international conference on World Wide Web (s.1137-1138). Raleigh, USA, April 26-30, 2010.
- Letierce, J., Passant, A., Decker, S. & Breslin, J.G. (2010). Understanding how Twitter is used to spread scientific messages. Teoksessa Web Science Conference (WebSci10). Raleigh, USA, April 26-27, 2010.
- Leskovec, J., Adamic, L.A. & Huberman, B.A. (2007). The dynamics of viral marketing. *ACM Transactions on the Web* 1(1).
- Leskovec, J., Singh, A. & Kleinberg, J. (2006). Patterns of influence in a recommendation network. Teoksessa 10th Pacific-Asia conference on Advances in Knowledge Discovery and Data Mining (s. 380-389). Singapore, April 9-12, 2006.
- Magnani, M., Montesi, D. & Rossi, L. (2013). Factors enabling information propagation in a social network site. Teoksessa T. Özyer, J. Rokne, G. Wagner & A.H.P. Reuser (toim.), *The influence of technology on social network analysis and mining* (s. 411-426). Wien: Springer.
- McPhearson, M., Smith-Lovin, L. & Cook, J.M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27, 415-444.
- Metsämuuronen, J. (2002). *Mittarin rakentaminen ja testiteorian perusteet*. (2. painos) Helsinki: International Methelp Ky.
- Milgram, S. (1967). The small-world problem. *Psychology Today*, 1(1), 61-67.
- Mischaud, E. (2007). *Twitter: Expression of the whole self. An investigation into user appropriation of a web-based communications platform*. Pro gradu-tutkielma. London School of Economics and Political Science.
- Najar, A., Denoyer, L. & Gallinari, P. (2012). Predicting information diffusion on social networks with partial knowledge. Teoksessa Proceedings of the 21st international conference companion on World Wide Web (s. 1197-1204). Lyon, France, April 16-20, 2012.
- Newman, M.E.J. (2001) The spread of epidemic disease on networks. *Physical Review E*, 66(1).
- Niiniluoto, I. (2002). *Johdatus tieteenfilosofiaan*. (3. painos). Keuruu: Otava
- Osborne, M.J. (2003). *An introduction to Game Theory*. Oxford University Press.

- Perrin, A. (2015). Social Media Usage: 2005-2015. Haettu 29.5.2016 osoitteesta [http://www.pewinternet.org/files/2015/10/PI\\_2015-10-08\\_Social-Networking-Usage-2005-2015\\_FINAL.pdf](http://www.pewinternet.org/files/2015/10/PI_2015-10-08_Social-Networking-Usage-2005-2015_FINAL.pdf)
- Raux, S., Grünwald, N. & Prieur, C. (2011). Describing the Web in less than 140 characters. Teoksessa Fifth International Conference on Weblogs and Social Media. Barcelona, Spain, July 17-21, 2011. The AAAI Press.
- Rogers, E. (2003) Diffusion of innovations (5. painos). New York: Free Press.
- Romero, D.M., Galuba, W., Asur, S. & Huberman, B.A. (2011). Influence and passivity in social media. Teoksessa D. Gunopulos, T. Hofmann, D. Malerba & M. Vazirgiannis (toim.), Machine learning and knowledge discovery in databases (s. 18-33). Berlin: Springer.
- Romero, D.M. & Kleinberg, J.M. (2010). The directed closure process in Hybrid social information networks. Teoksessa Proceedings of the Fourth International Conference on Weblogs and Social Media (ICWSM 2010), Washington DC, USA, May 23-26, 2010.
- Romero, D.M., Meeder, B. & Kleinberg, J. (2011). Differences in the mechanics of information diffusion across topics: idioms, political hashtags and complex contagion on twitter. Teoksessa 20th international conference on World wide web (s. 695-604). Hyderabad, India, March 28-April 1, 2011.
- Schelling, T.C. (1978). Micromotives and Macrobehavior. New York: Norton.
- Scott, J. (1991). Social Network Analysis: A Handbook. Sage Publications Ltd.
- Semenov, A. & Veijalainen, J. (2013). A modelling framework for social monitoring. International Journal of Web Engineering and Technology 8 (3), 217-249.
- Statista. (2016). Hours of video uploaded to YouTube every minute as of July 2015. Haettu 29.5.2016 osoitteesta <http://www.statista.com/statistics/259477/hours-of-video-uploaded-to-youtube-every-minute/>
- Suh, B., Hong, L., Pirolli, P. & Chi, E.H. (2010) Want to be retweeted? Large scale analysis on factors impacting retweet in Twitter network. Teoksessa Social Computing (SocialCom) (s.177-184). Minneapolis, USA, August 20-22, 2010.
- Twitter Blog. (2009, 30. syyskuuta). Soon to Launch: Lists. Haettu 28.10.2013 osoitteesta <https://blog.twitter.com/2009/soon-launch-lists>
- Twitter (2016). Company. Haettu 29.5.2016 osoitteesta <https://about.twitter.com/company>
- Uusitalo, H. (2001) Tiede, tutkimus ja tutkielma: Johdatus tutkielman maailmaan. (7. painos). Juva: WSOY
- Veijalainen, J., Semenov, A. & Reinikainen, M. (2015a). User influence and follower metrics in a large Twitter dataset. Teoksessa Proceedings of the 11th International Conference on Web Information Systems and Technologies (s. 487-487). Lisbon, Portugal, May 20-22, 2015.
- Veijalainen, J., Semenov, A. & Reinikainen, M. (2015b). Influence propagation through retweets in Twitter. Manuscript.

- Watts, D. (2007). Challenging the Influentials Hypothesis. *Word of Mouth Marketing*, 3, 200-211.
- Watts, D.J. & Strogatz, S.H. (1998), Collective dynamics of "small-world" networks. *Nature*, 393, 440-442.
- Weimann, G. (1991). The influentials: Back to the concept of opinion leaders. *The Public Opinion Quarterly*, 55(2), 267-279.
- Weng, J., Lim, E.-P., Jiang, J. & He, Q. (2010). TwitterRank: Finding topic-sensitive influential Twitterers. *Teoksessa Third International Conference on Web Data Mining (WSDM) (s.261-270)*. New York, USA, February 2-6, 2010.
- Wu, F., Huberman, B.A., Adamic, L.A. & Tyler, J.R. (2004). Information flow in social groups. *Physica A: Statistical and Theoretical Physics*, 337(1-2), 327-335
- Wu, S., Hofman, J.M., Mason, W.A. & Watts, D.J. (2011). Who says what to whom on Twitter. *Teoksessa Proceedings of the 20th international conference on World wide web (WWW'11) (s.705-711)*. Hyderabad, India, March 28-April 1, 2011.
- Yang, J. & Leskovec, J. (2010). Modeling information diffusion in implicit networks. *Teoksessa 2010 IEEE 10th International Conference on Data Mining (ICDM) (s.599-608)*. Sydney, Australia, December 13-17, 2010.
- Yin, R.K. (1989) *Case study research: Design and methods*. Sage Publ., Beverly Hills Ca.
- Zhao, D. & Rosson, M.B. (2009). How and why people Twitter: The role that micro-blogging plays in informal communication at work. *Teoksessa Proceedings of the ACM 2009 international conference on Supporting group work (s. 243-252)*. Sanibel Island, Florida, 10-13, 2009.