Lappeenranta-Lahti University of Technology LUT

School of Business and Management

Business Analytics

Identifying controversy on social media using Finnish twitter messages

Author: Olli Laasonen

1st Supervisor: Pasi Luukka

2nd Supervisor: Jan Stoklasa

ABSTRACT

Author: Olli Laasonen

Title: Identifying controversy on social media using Finnish twitter messages

Faculty: LUT, School of Business and Management

Master's program: Business Analytics

Examiners: Pasi Luukka, Jan Stoklasa

Master's thesis: Lappeenranta-Lahti University of Technology

78 pages, 7 tables, 10 figures, and 10 appendices

Keywords: polarization, twitter, social media, retweet, quantifying, controversy, network models,

In recent years, there have been growing concern about how social media affects people's behavior on the internet. To better understand polarization, filtter bubbles, and social media, one of the first steps is to identify controversial topics from non-controversial ones. This paper examines the controversy and polarization on social media focusing on Twitter. This work focuses mainly on networks and how quantifying methods can be applied and used in these networks. The approach includes three main steps: data collection and graph building, partitioning graphs, and measuring the controversy using different measures. To get a broader understanding of data and topics, Twitter messages are also analyzed. The results suggest that even especially political topics or even ties to politics are resulting in more polarized results. Another observation that can be made from the results is that the partitioning part of the process does have an important role when identifying the controversy.

TIIVISTELMÄ

Tekijä: Olli Laasonen

Otsikko: Vastakkainasettelun tunnistaminen sosiaalisessa mediassa suomenkielisten Twitter keskusteluiden avulla

Tiedekunta: LUT, School of Business and Management

Maisteriohjelma: Business Analytics

Tarkastajat: Pasi Luukka, Jan Stoklasa

Pro gradu -tutkielma: Lappeenranta-Lahti University of Technology
78 pages, 7 tables, 10 figures, and 10 appendices

Hakusanat: Polarisaatio, Twitter, Sosiaalinen media, Retweet, Mitata, Kontroversia, Verkkomallit,

Sosiaalinen media ja se kuinka ihmiset käyttätyvät sosiaalisessa mediassa on herättänyt viime vuosina kasvavissa määrin huolta. Jotta sosiaalisen median polarisaatiota, filtteri kuplia ja käytöstä sosiaalisessa mediassa on mahdollista ymmärtää yksi ensimmäisistä askelista on pyrkiä tunnistamaan vastakkainasetteluita siellä. Tässä työssä tutkitaan vastakkainasettelua ja polarisaatiota keskittymällä käyttämään Twitter keskusteludataa. Tutkimus prosessi rakentuu kolmen eri päävaiheeseen, missä ensimmäiseksi kerätään dataa ja muodostetaan verkkomallit. Tämän jälkeen seuraa verkkojen jakaminen kahteen eri osaan ja viimeiseksi sovelletaan erilaisia vastakkainasettelua mittaavia mittareita. Mallien ja tulosten tulkitsemisen avuksia Twitter viestejä on pyritty analysoimaan eri näkulmista ja lähestymistavoin. Tuloksien perusteella vaikuttaisi siltä, että erityisesti polittiseet aiheet tai pienetkin viitteet politiikkaan tuottavat polarisoituneita tuloksia. Toisaalta toinen huomio, mikä tuloksista voidaan tehdä on se, kuinka suuri merkitys prosessin toisella, verkkojen jakamisvaiheella on tuloksiin.

TABLE OF CONTENTS

LIST OF APPENDICIES

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

The social media platforms like Twitter, Facebook, and Instagram have snowballed in the past decade (Kemp 2022). These platforms control significant parts of our everyday life, giving us many uses (Perrin 2015, 52-68). We can, for example, use social media to connect with old friends or get new friends, share our thoughts, and expand our views about the world. Social networks can also provide us access to recent news stories, review different kinds of things, and give recommendations. These networks with internet access have been seen to be one the key factors that have made the world more connected, and it has been agreed that social media have increased the diversity of information. We can mostly agree that social media's impact has been very positive and has given many new opportunities (Akram and Kumar 2017, 351-354).

However, in recent years there have been a lot of concerns about social media, and many negative side effects have come to life. One of them is the fact that even though the diversity of information has grown, social media have also been connected to the increased polarization in our society (Garimella, et al 2018a, 913-922). This increased polarization is not only limited to political issues but a wide variety of topics like science and healthcare (Bail, et al 2018, 9216-9221). Just a quick glance at the trending topics in social media exposes the fact that discussions are many cases, polarized, and people cannot reach a consensus.

The increased societal polarization has been seen to be originating from biased assimilation, which means that, as humans, we have a tendency to interpret information in a way that supports desired conclusions (Lord, Ross, and Lepper 1979, 2098). Biased assimilation combined with the huge amount of information that social media platforms are providing us leads very quickly to the situation in which ideas are reinforced by repetition inside a system (Flaxman, et al 2016a, 298-320). This phenomenon can be called an echo chamber. Another thing that has been suggested to tie social media to increased polarization is Eli Pariser's idea about filter bubbles (Pariser 2011). These filter bubbles originate from algorithmic bias, which means that algorithms are nowadays recommending and suggesting to us most of the content we see based on our own interests and what we might click on. This way, we are consuming only content that feeds our own believes and eventually might increase the polarization.

The importance of understanding, identifying, and researching controversy in social media is the fact that it is affecting things that we should not have heavy disagreements like facts about health and climate; people are not anymore capable of understanding the other side's arguments, and overall people are becoming angrier (Holone 2016, 298-301; Bail, et al 2018, 9216-9221). At the same time, societies are becoming more and more politicized, where everything depends on which side you support, and the truth has been completely forgotten (Prior 2013, 101-127). Therefore, getting a better understanding of polarization and controversies in social media makes it possible to protect and prepare us for these negative side effects. Also, companies that design these products can use an understanding of the negative effects to help them design a healthier digital environment.

From a business analytics perspective, this topic is important since developing different products and, for example, using recommendation systems or black box models can easily lead to models which feed the polarization without even our knowledge. So, there should be more emphasis on identifying polarization. Followed up with products that consider the human side of things. By studying the problem of controversy in social media, there is an opportunity to understand better social media platforms generally and how social media and opinions can be modeled using social network models. Getting a better understanding of these platforms should make it easier to consider polarization in the future and design healthier models.

1.1. Objectives

This research’s primary goal is to understand better social media platforms' dynamics using social network analysis. The intention is to focus on Finnish social media users and identify and analyze the controversy about different topics. I believe that situation in Finland is a bit different than in other countries that have been researched because we have a multi-party system. Meaning that Finns are not automatically divided into two different groups based on political views like in the US.

The first objective of this thesis is to find out how interactions between different users can be modeled using social network models and how to identify and measure controversies. Based on these findings, the second objective is to create conversation graphs representing discussion topics and then analyze these networks to identify possible controversies. One of the main focus points will be the measures developed to identify controversy based on network structure and how to deploy those with real-world data. Main research questions that will help reach the goals can be formulated in the following manner:

1. *What are the ways controversy can be identified from social media?*
2. *How is controversy identification applied to real-world data?*
3. *What type of topics appears to be controversial in Finland?*

The way this research is trying to get answers to the research questions is three-parted. The first part intends to conduct a literature review that explains existing literature about social network analysis and how discussions can be represented using a conversation graph. This part also includes different ways to quantify social media controversy. The second part includes building the models and implementing those using real data and the best practices determined in the first part. The last part consists of analyzing the networks so that what type of interactions is happening amongst Finnish users in social media can be defined.

The research questions are limited in that research focuses only on users using the Finnish language. This way, it is possible to research interactions between Finnish users. Another limitation that will be made is the fact that analyzed networks will be created from Twitter data, which means that other social media platforms are not researched in this thesis. Also, the topics that will be used to collect data will be limited in that there will be more controversial topics and not so controversial. The goal is also to pick issues that are not, at least at starters, political, such as science or sports-related topics.

1.2. Research Process

The data in this thesis can be collected using Twitter's API, which gives access to tweets. Using Twitter data to build networks have benefits over other social media platforms. The first thing is that Twitter's API provides relatively unrestricted access to data, mainly because all the tweets are public. This means studying user interactions is more accessible than, for example, on Facebook. Another thing is that the documentation for Twitter's API is extensive and makes it easy to find answers when facing problems. So, we have many options to get data, and I believe that the final decision will be made based on the duration it takes to get data. So, even though there are alternative ways to get older tweets, the intention is to collect real-time data as it is generated on Twitter.

Before analyzing network models, the plan is to analyze tweets on specific topics on a general level. This way, there will be some idea of what kind of words, language, and feelings a particular topic hold after this plan is to create networks and perform partitioning for these networks. There are many kinds of partitioning methods to detect communities from networks, like label propagation and multi-level and info-map algorithms. Partition networks will generate two different sides, which can be analyzed using controversial measures. Finding the best solution most likely means experiencing other options and picking the best one. To measure controversy several controversial measures have been developed to measure polarization in social networks. Garimella K et al. (2018) studied the controversial measures, and they concluded that scores like Random walk and Edge Betweenness seem to give reliable results about polarization(Garimella, et al 2018b, 23).

1.3. Structure of the thesis

The structure of this thesis is such that it includes six different chapters, which have been divided into their own smaller chapters. In the first chapter, the purpose was to introduce the research topic, what answers are tried to get, and how. The second chapter is a more detailed representation of the subject and the theoretical background on which the research is based. After the academic background is explained, the third chapter introduces the basic terminology attached to network theory which will be used later. The fourth chapter, Methodology, includes the descriptions of data and how the raw data has been collected and handled. Also, there are explanations of polarization measures that will be used to quantify controversy on networks. The results of the research and analysis of results are included in the fifth chapter, where the networks are firstly examined based on the content of tweets and the structure of networks. After that, polarization measures are analyzed. The last and sixth sections will wrap everything up and try to find possible new points of view for the future.

2. BACKGROUND

This thesis aims to identify the controversy and compare different topics from Finnish social media using network modeling. Reviewing the existing literature about social theories that cause polarization, the thesis will be placed in context. The review will also be going through different methods used to study polarization, giving justification for the later used methods. The focus will especially be on community detection and polarization quantification methods. Polarization studies combine many different fields like computer science, sociology, social psychology, statistics, and applied anthropology (Boccalettia, et al 2006, 251). Therefore, the idea of this chapter is not to go through everything and be a thorough review but instead give a sample of studies

Most of the literature review information was collected from the LUT Primo web library, to which all the students and staff from the Lappeenranta-Lahti University of Technology have access. The benefit of using Primo is that it searches articles from many databases like Scopus and Web of Sciences and ultimately combines them. In addition to Primo, Google Scholar was also used if Primo's links were old and did not open. Keywords that have been used are such as controversy, polarization, social network, and social media. The search was also limited, so the language needed to be English. The search from LUT Primo with boundaries resulted in 317 articles. From these articles, the final filtering to get suitable references were made based on an article's header abstract, introduction, and conclusions.

2.1. Polarization

Polarization has been studied more and more in past years, and since it combines many different fields, it has been defined in many different ways. From Oxford Dictionary, Polarization is defined as "the act of separating or making people separate into two groups with completely opposite opinions”. Polarization has been described as a process in which people increasingly describe politics and society in terms of us versus them (McCoy, et al 2018, 18). The term polarization is usually used to refer to political polarization, but the definition is more general than that and includes other forms of polarization like religious, cultural, and economic (Garimella Kiran 2018, 11-12). Using this more general definition, polarization is more of a situation in which the opinion distribution has two distinct tops around the neutral opinion (Baumann, et al 2020, 1).

Polarization in itself might not be a problem, especially in political issues, since people have different opinions and the right to disagree, topics can be expected to be divisive (Katsambekis and Stavrakakis 2013, 117-126; Enyedi 2016, 210). The concern which has arisen regarding polarization is the fact that people are aligned within clusters with mutually exclusive identities and at the same time distancing themselves from other clusters (Lozada 2014, 1-16). This development's consequence is that compromises and finding consensus between opposite sides are even more challenging, and ultimately, people might lose trust in public institutions (McCoy, et al 2018, 18). Especially, political polarization has been connected with problems like race issues and Anti-Muslim prejudice (Aimei Yang and Charles Self 2015, 46-69; Hout and Maggio 2021, 40-55). Another reason why polarization is studied is its connection to the internet because it has been seen as one reason why polarization in modern society is increasing more rapidly (Yardi and Boyd 2010, 326; Bessi, et al 2016, 7-8; Gilbert and Karahalios 2009, 218-219).

2.2 Causes of Polarization

Many different things can cause polarization. This chapter will focus on how polarization is increasing and how these factors can be connected to the internet and social media. Or at least it is theorized. One way which has been used to approach causes of polarization is different biases which can be divided into three different levels: individual-level biases, group-level biases, and system-level biases (Garimella Kiran 2018, 1-69). The individual-level does mean biases that are our own, which can be affected by group and system-level biases like like-minded people and a system's algorithm.

2.2.1 Individual

Festinger et al. (1962) developed the theory about cognitive dissonance in the book "A theory of cognitive dissonance." In it, he suggests that information that is confirming people's beliefs or decisions creates positive feelings. Furthermore, even though cognitive dissonance theory is quite old, it is still considered to provide explanatory power(Harmon-Jones and Harmon-Jones 2007, 7-16). The effect of cognitive dissonance is that people pick up mostly sources that agree with them, decreasing the diversity of the information sources (Festinger 1962, 1). The selective exposure theory has been seen to build up to cognitive dissonance theory since selective exposure can be used to reduce discomfort (Jeong, et al 2019, 236-237). Strategies to cope with dissonance have been studied, and strategies have been following Festinger's et al. (1962) original theory. One strategy identified in empirical strategies is that former opinion is tried to confirm by re-reading already known information (Taddicken and Wolff 2020, 213 - 214).

On an individual level, we can think that cognitive dissonance is the root cause of polarization, leading to different factors that can cause polarization. As earlier stated, people are trying to expose themselves to an agreement and reinforce their views, which creates a selective exposure process leading to homogeneous groups and homophily (Colleoni, et al 2014, 318 - 319). Homophily has been defined as a tendency to affiliate with individuals similar in particular attributes and can be seen as one of the polarization causes (Lazarsfeld, P. F., & Merton, R. 1954). These attributes can be, for example, age, gender, beliefs, education, or social status (Garimella Kiran 2018, 16; Colleoni, et al 2014, 319). This tendency of individuals on social networks can create a situation in which users mainly connect with people who have similar views creating so-called echo chambers (Garrett 2009, 279-280).

Confirmation bias is a term used to explain how people focus on searching and interpreting the information the way it confirms and supports earlier beliefs and hypotheses (Taddicken and Wolff 2020, 206-207). Confirmation bias means that a person is looking for information that supports the hypothesis, and if the information is against the beliefs person tries to disregard information that dismisses the hypothesis (Nickerson 1998, 210-211). It has been shown that confirmation bias affects how rumors and misinformation spread online, keeping the information in different groups separated and polarized (Michela Del Vicario, et al 2016, 557-558).

The theory that has been connected very closely to cognitive dissonance is selective exposure theory. This theory is not new, but some influential reviews and articles have not supported it, which is why there has not been much focus on studying selective exposure (Freedman 1965, 287-289; Sears and Freedman 1967, 194-213). The idea behind selective exposure is that people tend to favor information that does not contradict their existing opinions/interests (Klapper 1960, 19 - 20). The reason why selective exposure has been seen more and more the reason behind polarization is that today's society has much more sources of information (Natalie Jomini Stroud 2008, 346-347). New media have challenged traditional media and ultimately given people more choices; for example, the news is available in many different formats. The problem that selective exposure creates is that people communicate only with like-minded sources and forget other ideas, leading to one's original beliefs strengthening (Stroud 2010, 556-557; Garimella Kiran 2018, 16). Huckfeldt, Mendez, and Osborn (2004) suggest that, ultimately, a discussion between like-minded people has led to more polarized attitudes.

Selective exposure occurs in different forms and can be divided into different types based on beliefs that motivate exposure. In recent literature, four different types of selective exposure have been emphasized (Stroud 2017, 4-5). The first type of selective exposure compares entertainment and news against each other, and it seems that entertainment is preferred to news (Prior 2007, 94-141). Then there is such selective exposure in which a person selects the information with more personal connections (Young Mie Kim 2009, 276-277). The third type is where the emphasis is more on the medium than the content (Hwang, et al 2006, 476). The last type of selective exposure is selecting information-based like-minded beliefs, and in recent years, the focus has been on this type (Natalie Jomini Stroud 2008, 341-366; Stroud 2010, 556-576). This fourth type of selective exposure combined with confirmation bias leads to biased consumption and reinforces polarized attitudes (Natalie Jomini Stroud 2008, 349).

There have been proposed two additional concepts to selective exposure called selective exposure theory from selective exposure theory called selective perception and retention (Frey 1986, 41-80). These concepts mean that confronting unpleasant material, the action is just ignored, and on the other hand, retention means that some information is favored based on earlier beliefs (Garimella Kiran 2018, 16). Selective perception and retention are closely connected to the term biased assimilation. Meaning that a person gets information from every angle but interprets it the way it supports pre-existing beliefs (Lord, et al 1979, 2106). For example, suppose believers and skeptics read the same two opposing fictitious studies in which violent video games and increased aggression are studied. In that case, the result is that they become more confident in their views, and mixed evidence only widens the gap between the two sides. (Greitemeyer 2014, 5-6)

When discussing polarization, the term echo chamber often bounces up, which refers to a situation in which people are consuming only stuff that expresses their perspective. So, in some ways, they only "hear their voice." From a social media perspective, echo chambers mean that users read or watch material that users themselves agree with. Earlier addressed homophily and selective exposure is something echo chambers are closely related to and based on those theories' natural consequence (Garrett 2009, 279-278). Echo chambers have been studied quite widely in recent years and shown to exist in online media like blogs, forums, and social media sites (Adamic and Glance 2005, 43; Williams, et al 2015, 136; Cinelli, et al 2021, 5). Partly the interest in echo chambers has risen because currently, people have vast amounts of information available, which has made it more of a partisan choice (Garrett 2009, 279). When biases like confirmation bias and biased assimilation direct sources that users consume, echo chambers are more likely to be born (Flaxman, et al 2016b, 318; Garrett 2009, 275). The argument that social media is promoting or creating echo chambers has been challenged, and some studies do not fully support this point of view (Colleoni, et al 2014, 328; Barberá, et al 2015, 1539). Echo chambers are not problematic from a polarization point of view, but the spreading of dis/misinformation is another problem linked to the echo chambers (Michela Del Vicario, et al 2016, 557).

One suggested reason affecting selective exposure is information overload, which means people's difficulty understanding and effectively making decisions in a situation where the amount of information available is tremendous (Stroud 2011, 19-20). Social media and the internet have increased the overload and are suggested to escalate other biases like selective exposure (Garimella Kiran 2018, 17).

2.2.2 Group

The previous chapter described individual-level biases that can cause polarization. On the individual level, the focus is on the users' choices and how biases control those choices, and the individual level is supported by group-level biases (Garimella Kiran 2018, 15). Group-level biases are a representation of biases that originate from users who are similar to each other and their group membership.

Social identity theory which has been seen to be one of the reasons behind polarization, is a sociological theory interested in group dynamics. Social identity theory suggests that a group of any kind will awaken positive feelings for the in-group and, on the other hand, negative emotions towards the out-groups (Tajfel 1970, 96-103; Tajfel, et al 1971, 149-178). Examples of identities can be race, religion, or class, and typically person belongs to multiple groups, like a white Christian male. Experimental studies have suggested that something threatening the individual's identification increases the in-group's bias, and identifying more strongly will mean greater favorability in their evaluations (Branscombe, et al 1993, 386-387; Branscombe and Wann 1994, 654). Roccas and Brewer (2002) represented that the complexity of social identity (overlapping between different groups that a person is a member of) may help people tolerate out-groups and confront the threats against their identity. So social identity complexity is in many ways the same as homophily but now only on a group level.

The term in-group favoritism or bias has been used to refer to the situation that people can develop if, for example, a person identifies very strongly towards his in-group. The in-group bias means the tendency to favor one's group over other groups and represent itself even though the subject on hand is entirely meaningless (Tajfel, et al 1971, 149-

178). If the social identity complexity is very low, there is a higher chance of being exposed to in-group favoritism (Roccas and Brewer 2002, 103; Branscombe, et al 1993, 386-387). From social media and polarization perspective, the in-group bias can be detected so that own political ideology is evaluated positively and other ideologies are rejected more often (Roccas and Brewer 2002, 103). A good example of social media is where a political party carries out some form of actions that are defended heavily in social media. However, if some other party makes the same decision, it is not supported anymore.

Sunstein (2002) defined a term called group polarization following way: "members of a deliberating group predictably move toward a more extreme point in the direction indicated by the members' predeliberation tendencies." Group polarization has been meant to refer to a predictable shift within a group in how they move more toward the extreme and decrease the differences between group members (Sunstein 2002, 178). For example, studies have shown how pro-feminist women have become more strongly profeminist; after discussion (Myers 1975, 712-713). Alternatively, whites offered more negative responses to questions about white racism and the conditions of African-Americans in American cities after discussion (Myers and Bishop 1971, 389). Group polarization has been suggested to increase, and this way also extreme decisions when people have a similar goal or some unifying external factor like politics or race (Sunstein 2002, 181). This type of development is noticeable in political debates since voicing hostility for opposing party supporters is acceptable and quite extreme, especially in social media.

2.2.3. System

Systematic biases are such biases that are not controlled by the user or group, and most systematic biases have been created by institutions (Garimella Kiran 2018, 17). This means that a process has a natural tendency to favor some results over others results. System-level bias from the context of polarization is referred to in two terms;

media and algorithmic bias, which can serve as catalysts for individual- and group-level biases (Garimella Kiran 2018, 17).

The first term mentioned earlier, media bias, has many different types of definitions. Though the underlying idea of media bias is that media, especially in mass media and persons working for them (journalists and producers), are explicitly slanting to some ideological direction (Groseclose and Milyo 2005, 2-3; Gentzkow and Shapiro 2006, 282; Iyengar and Reeves 1997, 40-42). Media bias, in many cases, represents itself in the way that the same underlying fact is reported, but the choices of words, sources, and overall tone are entirely different (Gentzkow and Shapiro 2006, 282). In the context of polarization, media bias can be defined as explicitly favoring one side over another (Garimella Kiran 2018, 17). It is also important to note that journalists with ideological views do not automatically mean that reporting is slanted in some direction. The opposite of bias news does not necessarily mean objectivity, but it is more about neutrality and balance, which means that all sides are represented equally (Eberl, et al 2017, 1127).

Media bias has been studied widely in the United States, and for example, Groseclose and Milyo (2005) measured which way American news outlets slant by comparing news outlets' cites to Congresses' citations. Their results showed a tendency to slant the news to the left, which they speculated to be a consequence of journalists systematically slanting stories to the left (Groseclose and Milyo 2005, 42). Another famous media bias example studied Fox news and its effect on voting behavior (DellaVigna and Kaplan 2007, 1-52). They were able to show how conservative Fox News was able to affect the senate vote by being partisan and biased (DellaVigna and Kaplan 2007, 32). Media bias has been studied outside of the U.S. For example, in Germany, unemployment reporting seemed to be more biased toward negative reports than positive ones (Garz 2014, 499-515).

Another term, algorithmic bias, describes situations where algorithms behind applications like recommendation systems or search engines make bias decisions and ultimately create unfair outcomes (Cowgill and Tucker 2019, 2). The algorithmic bias has raised concerns since algorithmic systems have been used to control almost everything around us, such as the consummation of news, transportation, relationships, etc. Often, users and sometimes even creators are not aware of the biases, and the algorithm changes the options (Garimella Kiran 2018, 18). It is not just that algorithms have been used to guide, help, and make peoples' small everyday decisions. However, algorithms also guide decisions concerning criminal sentencing, lending decisions, and hirings, leading to quite life-changing consequences (Barocas and Selbst 2016, 679-680, 690). For example, there are algorithm bias cases where skin color has caused more severe sentences or ethnicity affecting credit scores and interest rates (Angwin, et al 2016, 139-159; Bartlett, et al 2019, 29).

From a polarization point of view, a biased algorithm creates a problem: it can quickly erase all the opinions that do not resonate with the user's preferences. Pariser (2011) introduced the filter bubble concept to explain how people can end up in a situation where algorithms only feed such information that agrees with the user's beliefs and ideologies and eventually isolates them into cultural and ideological bubbles. Filter bubbles are linked to the term pre-selected personalization driven by websites, advertisers, or social media platforms, and many times users do not have much control over this personalization(Zuiderveen Borgesius, et al 2016, 3). Filter bubbles have been presented to affect how people receive health information or impact the election results (Holone 2016, 298-301; Jackson Jasper 2017). In these situations, filter bubbles have been suggested to pull people apart and reinforce the differences between groups, making logical sense in theory. Still, the approach has been criticized for its lack of evidence (Boutin 2011; Hosanagar, et al 2013, 822; Vaccari, et al 2016, 9).

2.1.4. Social media and polarization

Previous chapters introduced the main factors that can lead to polarization causing, and it is quite clear that all these theories have similarities and connections to each other and interact complexly. In the end, users are getting trapped in the "cycle of polarization," where new choices and decisions only make the problem even worse. It has not been established unanimously in existing literature whether social media can increase polarization, so this chapter will introduce different views about the fact.

The more and more controversial takes on social media on topics like immigration, inequality, and race have increased interest in the role of the internet/social network and how it can affect the formation of beliefs. The most recent problems with misinformation and conspiracy theories have also increased the interest in social media platforms and how the information can be manipulated there (Michela Del Vicario, et al 2016, 554-559). The argument many studies have proposed that the internet and social media speed up the process of Polarization (Stroud 2010, 556-576; Pariser 2011, 1-304; Flaxman, et al 2016b, 298-320). Studies are arguing for the internet's negative impact on polarization based on three key reasons.

Firstly, the information available is so huge right now, leading to vicious media competition and very personalized media sources, rolling the responsibility of choosing news sources for people who select those sources that agree with them (Papacharissi 2002, 17). The other used reasoning is that people nowadays can easily use filters and read stuff they decide to avoid the discomfort that opposite views create (Van Alstyne and Brynjolfsson 2005, 865). It is essential to mention that filtering is not just used in social media, but filtering happens without clearly noticing the matter of fact. Still, altogether, filtering had suggested feeding confirmation bias. The last reason used to support the internet's possible adverse effects on polarization is how the internet has increased social feedback leading to reinforced group thinking, homogeneity, and group polarization (Sunstein 2002, 175-195).

As earlier stated, the literature is not unanimously in favor of the fact that the internet/social media are causing mass polarization in society. Arguments made against the internet's role are mainly premised on the idea of a broader range of choices on the internet (Algan, et al 2019, 19). It has also been stated that from 1996 to 2016, polarization increased most in groups that do not use the internet compared to the groups that use it (Boxell, et al 2017, 10616). Literature supporting the view that the internet is not increasing polarization but instead reduces it; see that internet and social media expose people to more different opinions leading to less biased and narrow views (Barbera 2015, 28; Dubois and Blank 2018, 740-741; Algan, et al 2019, 29).

Social media users have been believed to encounter more alternative political views through weak relationships that social media provides by giving access to an extensive network of people without only direct friends. This would mean that people are not stuck on their ideology, people are more centered, and holding extreme opinions is rarer (Algan, et al 2019, 25). Social media use seems to also promote cross-cutting discussion between individuals who lack political interest and politically active people (Heatherly, et al 2017, 1285). So cross-idealogical debate would not be limited only to people heavily invested in some idea and would promote even more the exchange of different thoughts.

Boulianne (2015) conducted a meta-analysis about social media use and participation, which found evidence of a positive relationship between social media use and participation in political life. Even though more than 80% of the coefficients were positive, the analysis questioned whether the effect was causal and transformative since only half of the coefficients were statistically significant (Boulianne 2015, 534).

2.3. Detecting polarization

The simple definition of polarization was described in section 1, how there are two sides with different opinions on a specific topic (Baumann, et al 2020, 1). Since polarization has been studied quite broadly while combining multiple fields, the result has been many different definitions. One way to define polarization is to recognize nine senses that can be measured separately using methods based on distributions of opinions (Bramson, et al 2016, 27). Since this work will apply network models and focus on structural aspects of polarization, polarization will be understood through a simple definition of two sides that have different opinions.

Controversies and polarization have been tried to identify and quantify, focusing on two types of elements. The first one is the content of texts, and another one is the structure of graphs. Using content to detect polarization can mean many different methods, which have been premised on surveys and measures of distributional properties like bimodality (DiMaggio, et al 1996, 690-755). The content-driven approach to studying polarization on social media and the internet has meant that the methods have focused on natural language processing (NLP) and sentiment analysis. Mejova et al. (2014) study focused on news articles and analyzed how different news outlets in the U.S. use polarized terms. Controversies on Twitter, on the other hand, were identified using different types of features that are connected to the tweet, such as words in a list of controversial topics from Wikipedia and sentiment of tweets (Pennacchiotti and Popescu 2010, 31). More recently, Jang et al. (2016) used probabilistic modeling to identify controversy from Wikipedia, the web, and News corpora.

The development of the internet and social media has changed the attention toward system-level approaches that take advantage of structural aspects of polarization in network representations (Baldassarri and Bearman 2017, 3-4). Typically, these structural polarization measures try to identify the interaction patterns between users in some groups compared to other groups in that same network (Conover, et al 2011, 89-96; Garimella, et al 2018b, 1-27). In this study, the main focus will be on these structural-based methods over content-based ones since, especially on Twitter, the text content is quite limited, and the structural-based approach provides a language-independent way to explore polarization. Next, in this chapter, the structural polarization measures have been examined more closely.

One of the first studies that utilized structural polarization methods to study polarization on the internet was conducted by Adamic and Glance (2005). Their study explored the links between U.S. political blogs, giving empirical evidence that political blogs were more likely linked to other blogs with similar political ideologies than those of others (Adamic and Glance 2005, 43). The interactions were measured simply by calculating the degree of interactions between democrats and republicans and analyzing the density of patterns these networks created (Adamic and Glance 2005, 40-41). A couple of years later, Hargittai, Gallo, and Kane (2007) conducted a similar study using the linkage between political blogs to study interactions that supported Adamic and Glance's (2005) results. To compare groups, they also used the E-I index value to tell the level of insularity in a group (Krackhardt and Stern 1988, 127). The E-I index has been prone to provide unreliable results to unequal group sizes, which have been tried to account for using solutions like the adaptive E-I index (Salloum, et al 2022, 11-12).

Conover et al. (2011) continued developing structural polarization research when they studied Twitter interactions before the 2010 mid-term elections. The focus of their study was to investigate political communication on Twitter, building retweets and mention networks revealing that retweet network is highly polarized (Conover, et al 2011, 95). Conover et al. (2011) extended the analysis of networks compared to the earlier work introducing the concept of modularity and graph partitioning to verify the controversy in the structure of graphs (Newman, M. E. J. and Girvan 2004, 69-84). It is important to note that modularity was not used to quantify polarization but only to identify it. Conover et al. (2011) also found that usage of neutral hashtags leads more probably to interactions with opposing communities (Conover, et al 2011, 95).

As mentioned earlier, the initial work of polarization studies focusing on online communities concentrated on identifying polarization from network structures. The subsequent natural development was to develop measures that identify and quantify the polarization as well as language- and domain-independent metrics. One of the first ones to do this was Guerra et al. (2013), who created the metric called Boundary Polarization (BP). This measurement aimed to analyze the boundary between two possible polarized communities since these are the individuals interacting with the (potential) opposing group (Guerra, et al 2013, 219). It was also demonstrated that boundary polarization overcame the drawbacks of modularity but still left much hope for improvement and was limited to a small set of users (Guerra, et al 2013, 221-222). The new metric was explored using real-world data from Twitter and Facebook, including fields like politics and sports. Guerra et al. (2013) analysis of real-world data revealed that polarized networks tend to have a low number of individuals with a high degree of the boundary between two communities.

A couple of years later, Morales et al. (2015) released a new way to measure polarization on a social network inspired by the electronic dipole moment, called Dipole polarization (DP). Dipole polarization works very similarly as a community detection method called label propagation which has been used very widely (Morales, et al 2015, 2). Dipole polarization requires so-called elite users whose opinions will be determined manually. After that, the ideas are determined to listener users based on the network structure and iterations (Morales, et al 2015, 2-3). In the end, The polarization metric will be summarized into one value which represents mainly the distance between positive and negative clusters. To validate the new polarization metric, twitter data about Hugo Chavez's death announcement was used (Morales, et al 2015, 5-6). The analysis results showed that the polarization seemed to decrease right after the information, but after a couple of days, it had increased higher than before (Morales, et al 2015, 7). The shortcoming of using dipole polarization is connected to the fact that the distribution of opinions happens through selected groups, causing some users not to play their role in the distribution process (Emamgholizadeh, et al 2020, 5).

One of the latest polarization metrics was developed by Garimella et al. (2018) when they introduced the metric called random walk controversy (RWC). It is similar to earlier introduced Boundary and Dipole polarization metrics in how it quantifies polarization based on the graph structure of social interactions. However, the algorithm works entirely differently from the earlier ones since it relies on random walks and walks probability of ending up in communities' influential individuals (Garimella, et al 2018b, 11). The polarization score is a summary of multiple random walks based on the starting and ending point distributions. Garimella et al. (2018) have modified how the metric should be calculated since the straightforward way of using the Monte Carlo simulation is quite computationally heavy (Garimella, et al 2018b, 1-27). The same article as was random walk controversy introduced by Garimella et al. (2018) also introduced the betweenness centrality polarization score. The simplified idea of this metric is to compare the difference in the edge betweenness centrality scores of external and internal links (Garimella, et al 2018b, 13). The experiments with a random walk and other methods revealed that random walk seemed to work better than other structure polarization metrics like dipole or boundary polarization with Twitter data (Garimella, et al 2018b, 23). On the other hand, the limitations still remain that random walk relies on graph partitioning, which can be overfitting and polarization's bound (Garimella, et al 2018b, 23-24). The experiments with different methods also suggested that if used Twitter data, the graphs should be built based on retweets or following, but the content and mention graphs do not seem to give reliable results (Garimella, et al 2018b, 16-17).

The Random walk controversy measure has been suggested to give lower values for larger graphs with the same average degree and to remove this dependency suggestion to use an adaptive random walk controversy metric (Salloum, et al 2022, 9). The early results and experiments suggest that the adaptive random walk controversy could solve community size problems and give more accurate results (Salloum, et al 2022, 14). All in all, all the structural polarization measures have some downsides and are not entirely perfect. However, compared to the other options like content-based methods, the most significant benefit of structural measures is that quantifying can be made automatically from communication systems, is language-independent, and perform better in many cases (Garimella, et al 2018b, 21-22).

3. NETWORKS

This chapter intends to give a brief background to the network theory and give a basic understanding of general terms and structural factors used to discuss networks and this thesis. The introduced terms will be such, which are going to reappear in the methodology parts of the text. Still, it is better to introduce them now because it is easier to understand methods when terms are familiar. Also, some terms have been used a little bit differently depending on the field of study, and this way, concepts will be the same for everyone.

3.1. Background of Network theory

Systems that can be represented using networks and are taking the form of networks are countless, but maybe one of the most famous systems usually formed using a network is the World Wide Web. Other networks can be transportation networks, food and egological webs, and industrial networks (Dorogovtsev and Mendes 2001, 1). Even though networks can have many different use cases and interpretations, there is always the same fundamental concept. A network has vertices or nodes representing some items that have connections between them called edges (Newman, M. E. J. 2003, 167).

In understanding network theory, it is essential to clarify where the study of networks is coming from. Historically, the study of networks has been a branch of discrete mathematics that has been called graph theory (Boccalettia, et al 2006, 177). The first proof of graph theory that has been often cited is Euler's 1735 solution to the Königsberg bridge problem (Newman, M. E. J. 2003, 2). In his publication, Euler tried to find a round trip using all the bridges of the city of Köninsberg exactly once. Even though social networks have not been studied for a long time, networks' structure and functions have been studied for quite a long time.

Watts and Strogatz's (1998) seminal paper about small-world networks and Barabasi and Albert's (1999) paper about scaling in random networks created new interest and research in the study of networks. The study of networks focused on complex networks that were irregular, complex, and dynamic. Complex networks meant that models proposed in mathematical graph theory were not good enough anymore. Moreover, it had to be developed to mirror Watts and Strogatz's (1998) idea of network properties (Newman, M. E. J. 2003, 4). Also, the interest in real networks raised, and new network models started to appear (Boccalettia, et al 2006, 177-178).

It might seem that graph and network theory are synonyms of each other, but this is not quite the case. The difference-maker between these two is that network theory is a set of techniques for analyzing graphs (Albert and Barabási 2001, 1). On the other hand, graph theory is more the framework for the mathematical treatment of networks (West 2001).

3.2. Terminology of networks

Since networks and graph theory have been used to study a wide range of different fields, the terms might be confusing, and sometimes some words have a little bit different definitions. This chapter will explain the basic terms used in this thesis to better understand the methods and results that this thesis will later discuss. The focus of this chapter will be mainly on the structural properties of a network since these are the factors that create the possibility of studying polarization on social media.

3.2.1 Vertices, edges & direction

Two of the most critical pieces that networks need to have are vertices and edges. Vertices are often called nodes if the study focuses on computer science, and in sociology, vertices are actors (M. E. J. Newman 2003, 173). In figures, the vertices are usually represented using dots connected using edges that are lines between nodes. The line between vertices can be called links in computer science and ties in sociology (Newman, M. E. J. 2003, 173). It is important to note that it does not matter how the dots and lines are drawn in visual representations. The only thing that matters is which node has the edge over which node (Boccalettia, et al 2006, 179).

One of the most common properties of a network is the direction of edges. Edges can be directed or undirected, which means that a network is directed or undirected (Dorogovtsev and Mendes 2001, 2-3). The difference between these two is that edges are always both ways between nodes in an undirected network and do not represent a direction (West 2001, 53). In some rare cases, directed edges can be called arcs (Newman, M. E. J. 2003, 173). From visual representations, the directed edges are defined using arrows. For example, communication in social media is directed since messages travel only in one direction and do not connect both ways. In the same way, on Twitter, following somebody is a directed relationship. Even though the visual representation of networks helps to illustrate phenomena and understand their properties, real networks are usually represented using adjacency matrices.

3.2.2. Degree

Node degree is the number of edges connected to a node (Newman, M. E. J. 2003, 173). There can be more than one edge between two nodes, so the degree does not necessarily equal the number of nodes connected to the node. A directed network means that outgoing (out-degree) and ingoing (in-degree) edges can be separated from, and this way, more information can be analyzed (Boccalettia, et al 2006, 181). In some sources' node degrees have been called connectivity, but it already has meaning in graph theory, so it is less complicated to use the term degree. Calculating node degree is one of the first statistics calculated and analyzed in studies focusing on the internet and social media (Hayes 2000, 12-13). Exploring degrees on social media usually means the possibility of unveiling those actors who influence other actors the most.

3.2.3. Shortest path lengths & diameter

The shortest path is another very commonly used term when networks are discussed. Sometimes called geodesic distance, the shortest path means the distance between two vertices through the optimal route (Dorogovtsev and Mendes 2001, 3). In many cases, especially if the network is extensive, there is not just one shortest path but multiple (Newman, M. E. J. 2003, 173). Distribution of the shortest path lengths with the average shortest path is a summary statistic often calculated to analyze networks (Dorogovtsev and Mendes 2001, 3). From the distribution of the shortest path, it is possible to get the maximum value called the diameter, representing the longest geodesic path between any two vertices (Boccalettia, et al 2006, 182). It has been discovered that the average shortest path seems to be small even for large networks, and this is called the small-world effect (Newman, M. 2000, 819).

Understanding optimal paths can be extra beneficial for transportation and communication. In these fields, marginals are small, and minor changes can lead to significant financial savings, making studying the shortest paths quite interesting for the private sector. The shortest path can be understood in social media networks how similar the actors are, and that's why it has been used in some quantification methods like Garimella et al.'s (2018) betweenness measure.

3.3. Network types

As stated earlier in chapter 3.1, graph theory is the mathematical framework for networks. Networks are often called graphs (especially in mathematical literature) (Newman, M. E. J. 2003, 2). The simplest example is a set of nodes connected by edges (Brian Hayes 2000, 9). However, many times the nodes and edges can have different types of properties. For example, in a social network of people, nodes can represent nationality, gender, or income (Newman, M. E. J. 2003, 3-4). At the same time, edges can represent friendship or geographical proximity (M. E. J. Newman 2003, 3-4). Studies about polarization on social media, nodes are usually representing users. The edges are some interactions, like a retweet, following, replies, or like.

In this thesis, the networks will be unweighted, which means that the connection between nodes either exists or not, but the distinction between unweighted and weighted networks is good to understand (Boccalettia, et al 2006, 198). A weighted network would be a situation in which the intensity of the relationship can differ, and each edge carries a value representing the strength of the connection (Banavar Jaynath, et al

1999, 130-132; Berlow 1999, 330-334). The real-world examples of weighted networks could be weighted transportation networks' nature or the strength of social links.

When handling real-world networks, the situation can be such that a node is connected to another node by more than one link called a multigraph (Hayes 2000, 10). Multigraphs can also originate if a node has a link to itself called a loop (Boccalettia, et al 2006, 180). Call graphs are one example of a network having multigraph properties since telephone numbers are connected by calls made between numbers, and calls can be completed multiple times between the same callers. A bit more complicated graph type is hypergraph which changes how edges are understood. In hypergraphs, the edges are called hyperedges that can join more than two nodes together (Newman, M. E. J. 2003, 172)

4. METHODOLOGY

The following chapter goes through the whole research process, from the data collection to the quantifying phase. There are also introductions to the controversy measures and what type of thinking measures are based upon.

4.1. Process of quantifying polarization

This part's primary focus is how raw data is collected and manipulated. This part aims to clarify what kind of steps must be taken to deploy controversial measures. The point of the chapter is to give an overall look at the process and not to dive deeply into the theoretical side of things

4.1.1 Data collection and restrictions of process

Most of the analysis that has been done focusing on social media controversy has followed a similar structure (Garimella, et al 2018b, 1-27; Chen, et al 2021, 1-27; Conover, et al 2011, 89-96). Even though the end goal might have been slightly different, the investigation started from almost always the collection of the data part. If this part had not been done, it meant that some other party had already collected the data. In the context of Twitter, the data is a compound of tweets and all the information that is connected to these tweets. This means that a tweet is not just an anonymous message. There is data about the user who has created the tweet, discussion around a topic form of replies and retweets, how many and who have liked the tweet, etc.

The most common way to collect and analyze Twitter data is to use companies own Rest API, which provides a programmatic way to access Twitter. Even though the data collection is quite simple and easy through API, some limitations restrict how data can be collected from Twitter when a basic-level account is used. Also, you must apply access to the API from Twitter, and additional applications must be sent if requirements are met, and broader access is needed. It is possible, for example, to create or like tweets, but in this work, API is used to do searches and store information about tweets that pop off from searches.

The significant restrictions affecting this work are that you cannot access the archive fully, meaning that the search gives you only one week's worth of tweets (Twitter 2022). The condition is not that big of a deal if inquiries are made repeatedly and duplicates are deleted before making any analysis. Another limitation set is the number of tweets you can retrieve, which is 2M per month (Twitter 2022). Since the research focuses on Finnish Twittersphere conversations that were not growing that big, this limitation has been crossed.

4.1.2 Tweets to Graphs

Collecting tweets using Twitter's API has basically the same idea as using the regular advanced search in the Twitter application. Meaning that you create a query where you can set criteria for what type of tweets you want to search, and by repeating the search requests, you can go through tweets that met the particular standards and have been posted in the past seven days. This search query could be simple, like a single keyword like "#Russia." The data, in this case, would consist of all the tweets with "#Russia" in them. But the problem occurs when a hashtag is written form, for example, “#russia” which means that the tweet is not included. That's why it is necessary to use queries that consist of multiple words. This thesis will use a straightforward topic model called NMF to solve this problem since it seems to work quite well with the tweet type of texts, is easy to implement, and is not computationally heavy (Egger and Yu 2022, 4; Shi, et al 2018, 1111).

The way NMF is used step is that there is some seed word that is used to do the initial search of tweets. From the first patch of tweets, NMF is used to create a topic meaning a group of different keywords. So, for example, in Russia, example “seed word” #Russia could be linked to words like #russia, Russia, russia, and Ukraine based on NMF topic modeling. The search is done again when the query contains words connected to each other and are relevant. Using more than one keyword to execute a search, networks have a broader viewpoint and do not pick only one side's talking points since one side might use different hashtags (Garimella, et al 2018b, 16) (Garimella, et al 2018, 16). Tweets form a network where vertices are users who have been participating in the conversations by retweeting some tweets, and edges are representations of the agreement or shared point of view even though some profiles have added a disclaimer to the status that retweets are not endorsements, the earlier studies have suggested that retweets are still mainly used as an endorsement (Metaxas, et al 2015, 661).

4.1.3. Graph to partitioning

In partitioning, the idea is to prepare the network so that polarization measures can be completed. So partitioning is, in some ways, interphase. A graph is divided into two parts in the way the partitioning algorithm gives every vertex its label. In this case, we are studying polarization from the point of view that there are only two sides, so there are only two different labels. The result of partitioning represents which side a user is "landing" on based on the structure of the network and activity on Twitter.

Basically, when assuming that there are two sides and how tightly are these sides connected (Garimella, et al 2018b, 1-27)? Then the question is whether these two sides disagree or agree, meaning that the network is polarized or not. A network structure where the two sides are very loosely connected would mean that users disagree, which is the core focus of controversy measures.

4.1.4. Partitioning to measuring controversy

So the last part is the actual measuring of the controversy, which would provide a piece of information on whether or not the topic has been polarized. The polarization measures take a graph as input and put it together with results from the partition. The controversy measure algorithms are built differently, but every measure's goal is to give one number that represents the controversy. Various measurements have been developed to quantify disagreement on a network graph, and that's why multiple measures are tested. The measures are described in a more detailed way in chapter 5.

4.2. Data and Graph building

This chapter provides a more detailed look at the data and how the graphs are composed, also giving some critical statistics about network sizes and how the posts were made.

4.2.1. Data collection

Even though there are multiple social media sites online to research controversy on the internet, there are fewer options. The most natural ones to the polarization studies are Twitter and Facebook since both sites are more focused on the textual format of communication than, for example, Instagram or Snapchat. In the latter ones, the posts are more pictures and videos, which makes it challenging to build graphs. Picking Twitter over Facebook has a couple of reasons: tweets are all public, and getting access to them is much easier and less complicated than posts on Facebook. Another reason to choose Twitter for this type of study is that Twitter is used to debate current news much more often than Facebook.

On Twitter, earlier studies suggested that the networks built from retweets were the most reliable, and also most of the earlier studies focused on comparing retweet networks (Conover, et al 2011, 95; Garimella, et al 2018b, 23; Salloum, et al 2022, 1-33). Garimella, et al. (2018) used the following statuses to build a graph, but since it has a heavy computational process, which makes it more convenient to use retweet graphs. So all the graphs are based on retweet activity, and any other user interactions are not used to build graphs.

The raw data was collected mainly between July 2022 and October 2022 from the topics that were trending during that time. However, some tweets are gathered at the beginning of September. Since the study deals with real-time events, all the topics are not collected during the whole period. This means that if the topic is such that people are talking for a short time, the gathering period is also shorter.

The main focus is on the point when users were interacting the most. On the other hand, the more general topics and discussions happening during a more extended period, the more data is gathered longer. Since the data had to be collected within seven days from its original creation and if the data were collected for two straight weeks, there might be minor interruptions in the timeline.

The queries used to make searches are simple, without any complicated definitions on top of search words. One addition to the rules is that the language has to be fi, meaning that only tweets that contain the Finnish language are accepted. This way, the focus is on only Finnish Twitter usage. The only other rule is the so-called tweet mode set to the extended. Changing the mode gives access to whole retweets and other additional information.

4.2.2 Construction of topics

The graph building starts by generating the topics. The earlier studies have defined topics as only one hashtag or multiple hashtags (Salloum, et al 2022, 7; Garimella, et al 2018b, 6). The problem with using only one hashtag to generate a topic is that the graph which will be built can be one-sided and leave some parts out. Also, Garimella et al. (2018) raised this issue when building the graphs. In this work, topic modeling is not the most critical part, but it is mainly a tool to give an idea of what type of language is used and what keywords should be used to search tweets.

This topic can be created using many tactics, one being straightforward manual defining, meaning that the keywords are defined without different algorithms or statistics. In this work, the topic decided to use the Non-Negative matrix Factorisation algorithm, which was introduced first by Paatero and Tapper (1994). Since the topic modeling and text data analysis can be dispersed, the final decision to pick some was made manually, which means that all the words which are part of some topic modeling results are not used to build graphs.

The NMF-based algorithms' basic idea is to decompose the term-document matrix (Shi, et al 2018, 1111). Term document matrix, in this case, means matrix representation of corpus. The goal of decomposing is that the result is two low-rank factor matrices (Lee and Seung 1999, 788-791). The output of NMF is usually named W and H, whereas, in topic modeling, W contains the topics found and H the weights of the topics.

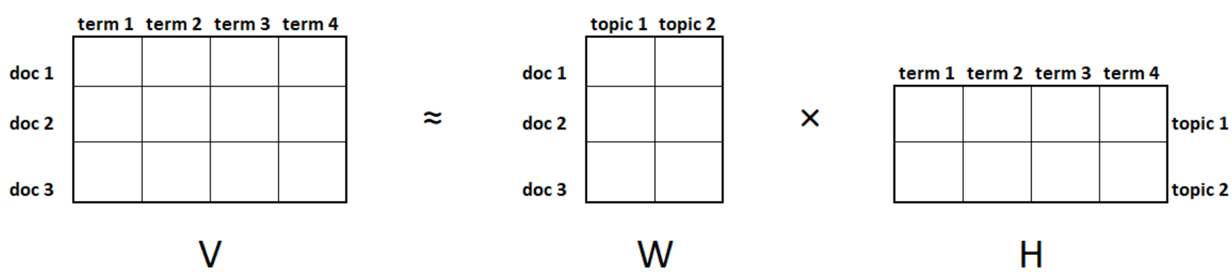


Figure 1 Illustration of Non-Negative Matrix Factorixation model for topic modeling

If the original matrix is called V, the basic idea is to try to find two non-negative matrices, W and H.

$$V = WH \tag{1}$$

Thus the matrix V is n × m matrix where the n is dimensional data vectors and m is the number of examples in the dataset (M'sik and Casablanca 2020, 5757). In the process, the matrix V is approximately factorized by two matrices: W, n by t matrix, and H, t by m matrix (Lee and Seung 1999, 790). The whole point is to get variable t less than n or m, which will result in the compression of the original data matrix. From a topic modeling perspective, NMF is portrayed in Figure 1. The document-term matrix is V, the W (n by t) matrix is the document-topic matrix, and the H matrix is the topic-term matrix (t by n) where the variable t is the topics (M’sik and Casablanca 2020, 5757).

4.2.3 Retweet graphs

After all, the keywords forming the topic are chosen, and search queries are completed, every case is evaluated. The first thing is to assess whether or not there are enough tweets to construct a graph, and if the volume of tweets is low, then the topic is dropped. Some tweets might be the results of multiple search queries; all the duplicate tweets were dropped off using the tweet's unique tweet id object. The final set of keywords that were used to build the graphs with how many related tweets there were at the end of the day and the periods when the tweets were gathered are described in Table 1.

As earlier described, from every topic, there is retweet graph G in which a node represents the users tweeting about the subject and edges the user's retweet from another user. To reduce noise, Garimella et al. (2018) noticed that it is best to form the edges between two users that were created if the edge existed at least two times, so if the user had retweeted another person's tweet two or more times. This way, there is a smaller chance that a retweet does not represent the endorsement. The threshold of two retweets seems to offer a good balance between not too much filtering and noise reduction (Garimella, et al 2018, 8).

4.2.4. Description of topics

This chapter describes different topics and what keyword has been used as a seed word to start the creation of topics. After this chapter, topics are especially in visuals referred to using only the keyword which has been seed since it is good compression of a topic.

The Finnish word Venäjä means Russia, which has been used to create one of the topics. As a whole, the topic contains very generic words which are connected to Russia's attack on Ukraine. The entire query used to create a topic about Russia's war in Ukraine contains words like Ukraina (Ukraine) and venäläinen (Russian). The idea of this topic is to represent discussions during the time Russia decided to announce mobilization.

Quite closely connected to Russia's war in Ukraine is a topic of the Nordtstream explosion, which is built using the keyword #Nordstream. The time window for gathering this topic is a little after Russia's mobilization news. The most significant difference between the general Russia topic and the Nordstream explosion is that another Nordstream attack is a closer news story for the Finnish people. One more topic related to the war in Ukraine but a more national issue is a conversation on tourist visas (#Viisumit) and whether visas should be granted to Russian citizens.

Table 1. Descriptions of the topics

Keyword used to start search	No of Tweets	Description of Topic	Time period
Fortum	25 495	*Speculation of Fortum's saving packet and its actualisation*	10.7.-28.7.2022
Fortum	19 251	*Speculations of Fortum's future with Uniper and possible decisions*	31.7.-3.9.2022
Fortum	11 730	*Germany reaches deal to nationalise troubled gas giant Uniper*	17.9.-24.9.2022
Huumeet	8 745	*Citizens' initiative on supervised drug use rooms and reaching 50000 signatures*	15.7. - 31.7.2022
Iran	31 991	*Protests and civil unrest against the government of Iran*	17.9. - 1.10.2022
Marin	128 505	*Finland PM Sanna Marin partying and backlash*	17.8.-4.9.2022
#sectorallyfinland	2 641	*WRC Secto Rally Finland 2022*	3.8.-8.8.2022
#RallyNew Zealand	2 819	*WRC Rally New Zealand 2022*	26.9.-2.10.2022
Venäjä	26 964	*Russia's decision to start mobilization.*	19.9.-24.9.2022
Yle	51 597	*Discussions about Finland's national brodcasting company and current affairs programs*	27.8.-11.9.2022
#Sähkönhinta	8 572	*Discussions on high energy prices and future energy crisis*	20.8.-3.9.2022
Viisumit	11 390	*Restrictions on the entry of Russian citizens*	18.9.-1.10.2022
#NordStream	26 820	*Nordstream gasline explosions*	26.9-1.10.2022
Huuhkajat	3 497	*Finnish football national team playing in Uefa Nations league*	20.9.-26.9.2022
#Jodi	4 993	*The Ministry of Social Affairs and Health decision to change recommendation of iodine intake*	9.10.-13.10.2022
#meidänkaikkienasia	6 730	*The office of Local Government and County Employers (KT) and Finnish nurses agreed upon salaries and other benefits*	27.9.-4.10.2022

The topic of the protests in Iran is different from any other because the location of events is far from Finland and does not significantly impact Finnish people's everyday lives. Iranian women's protests arose news in September 2022, and Finnish social media discussions were collected using the keyword Iran. Most of the tweets were collected close to the time Mahsa Amini was killed, which started the protests.

The topic of Finnish news and traditional media is investigated using the keyword YLE, a Finnish nationally owned media company. This topic has been collected for precisely two weeks, but search words in query seem to be somewhat general since the search resulted in over 50000 tweets. The subject is quite interesting since YLE has often been accused of being politically biased towards the left or right side, depending on the topics discussed on the news.

One of the topics related very heavily to politics and at the same time linked to one person is leaked material where Sanna Marin is partying and all the other things associated with this partying. The single thing that can be brought up from that topic is that the discussion was very active. Discussion around drugs (Huumeet) focused the conversation on drugs and drug consumption rooms during the time citizens' initiative was open. Even though the topic is political, it has mixed support without classic left and right division.

Another political topic connected to the economy is the topic of Fortum. This Fortum discussion is divided into three parts because Fortum has been discussed so much and long. In periodical order, the first Fortum topic is about the emergency packet and Fortum/Uniper relation speculations. The second Fortum topic's collection period is more related to the speculation of economic problems and all the problems that Fortum faced with Uniper. The last one is about Fortum's decision to sell Uniper, which was picked because the discussion was relatively inactive during that time. The search query for all the Fortum-related topics is the same.

Three national news that trended during the time topics were created were news of the ending of job actions between Finnish local authorities and nurses (#meidänkaikkienasia), iodine (#Jodi), and energy prices (#Sähkönhinta) in Finland. Classically the issue would be very left/right divided, but now it was not that clear because the left-leaning government had made some decisions that were making nurses' job actions harder. When the news broke, there were no real answers to what had been agreed upon, and the results created confusion on social media. Most of the discussion around energy prices focused on the new, high prices which resulted from Russia's attack on Ukraine. Discussion around the topic of iodine originated from the fact that iodine tablets were sold out because the ministry of social and health in Finland changed the recommendation for iodine intake.

#Huuhkajat, #RallyNewZealand, and #sectorallyfinland are all Finnish sports topics that were trending when tweets were gathered. All the cases are linked to nationality since Huuhkajat is a Finnish national football team nickname. Kalle Rovanperä drove his best season and won a world title in WRC class.

4.2.5. Tweet datasets

See all the daily counts of tweets about each topic from appendix 1 and all the average number of users tweeting per day from appendix 2. The total number of tweets between different topics fluctuates quite much, and one of the biggest reasons is that various

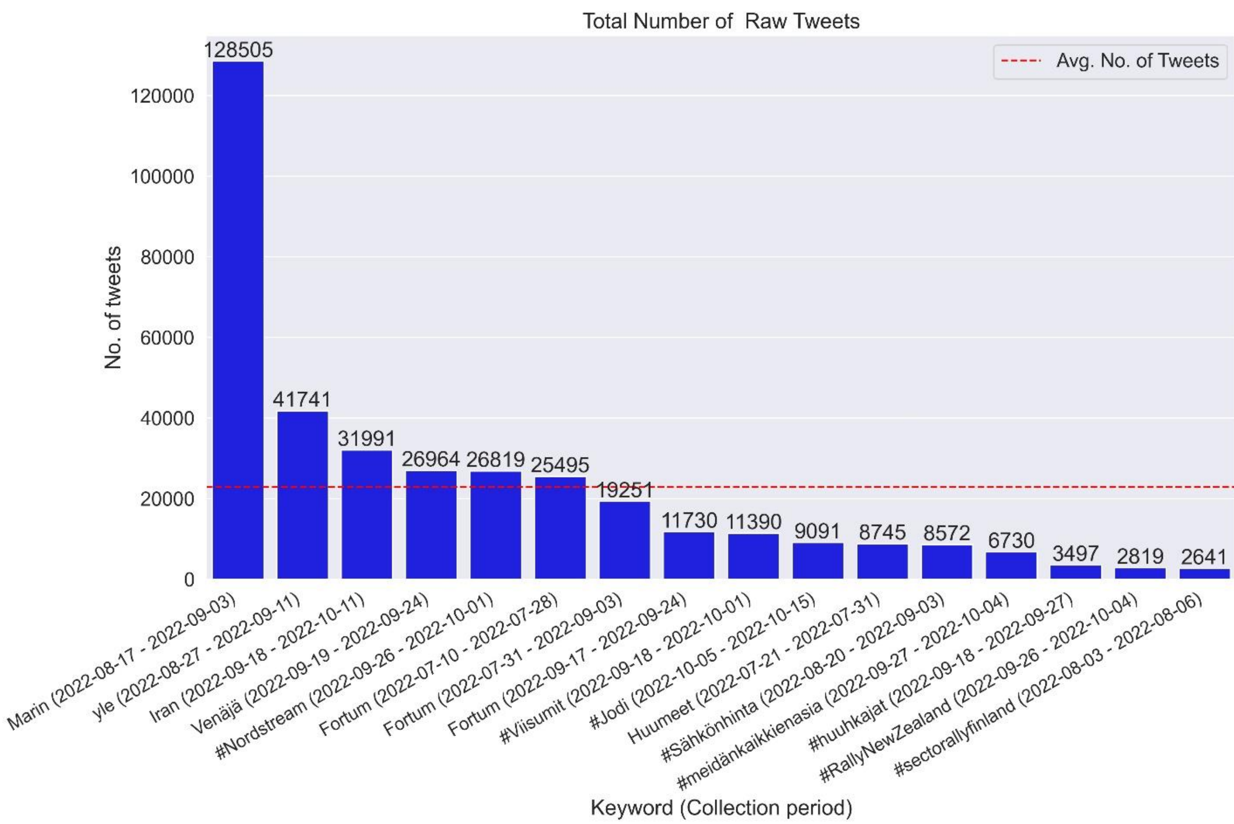


Figure 2 Total Number of Raw Tweets

topics have different collection periods (Figure 2). This is because some cases have more interest in social media, resulting in data being collected longer time than others.

From the total number of tweets perspective, it is clear that the topic of Sanna Marin's partying inspired much more than any other topic (Appendix 1). Even though the tweets were collected for quite a long time compared to other topics, it is still quite remarkable how many tweets were posted during two and half week timespan. The second highest number of tweets gathered is topic handling about YLE. The high number of tweets might be because the search query had some very vague search words outside the honest discussion around YLE and its current affairs programs.

The topics with more than average tweets seem to be linked to matters that are internationally big news if we count Yle out. Even though the tweets are collected from Finnish users, the news or event itself has been noticed outside Finland. For example, Russia, Nordstream, and Fortum are topics that affect Finnish people but are also internationally big news. Below average, most of the news is somewhat local news stories that are no that broadly trending in other countries.

Using real-world data and only having a 7-day window to collect data means that data is heavily affected by real-world events, and topics heavily emphasize these events. As earlier mentioned, tweets were composed between July 2022 and October 2022, meaning that issues connected to Russia's invasion of Ukraine were often trending. Topics straightly related to Russia's war were not only trending often, but also the discussions around it were very active. The fact can be noticed Table 2. The searches started using Venäjä (Russia), or #Nordstream keywords had almost 4500 tweets per day.

A little bit indirectly connected to the war in our data are discussions around the topic of visas and whether or not Finland should permit Russian people to have tourist visas. The tweets were collected for a more extended period, but at the highest topic organized, only 1405 tweets per day, significantly less than other topics related to the war.

Table 2. Description statistics of collected Tweets

Seed word (Collection Period)	No. of days collected data	Avg. No of tweets per day	Std No of tweets per day	Min No. of tweets per day	Max No. of tweets per day
Venäjä (2022-09-19-2022-09-24)	6	4494,0	2534,2	1897	8879
#Nordstream (2022-09-26-2022-10-01)	6	4470,0	2559,6	579	7642
Huumeet (2022-07-21-2022-07-31)	10	874,5	566,0	310	2188
Marin (2022-08-17-2022-09-03)	18	7139,2	4753,1	1519	16350
#RallyNewZealand (2022-09-26-2022-10-04)	9	313,2	481,5	8	1532
Iran (2022-09-18-2022-10-11)	24	1333,0	640,3	433	2563
#Jodi (2022-10-05-2022-10-15)	11	826,5	1107,6	193	3114
Fortum (2022-07-31-2022-09-03)	29	663,8	767,1	31	2877
#meidänkaikkienasia (2022-09-27-2022-10-04)	8	841,3	686,6	319	2474
Fortum (2022-09-17-2022-09-24)	8	1466,3	1095,5	476	3166
#Viisumit (2022-09-18-2022-10-01)	14	813,6	383,8	76	1405
#Sähkönhinta (2022-08-20-2022-09-03)	15	571,5	370,1	65	1121
#sectorallyfinland (2022-08-03-2022-08-06)	4	660,3	500,6	182	1364
yle (2022-08-27-2022-09-11)	16	3224,8	1301,3	787	6275
#huuhkajat (2022-09-18-2022-09-27)	10	349,7	376,8	42	1194
Fortum (2022-07-10-2022-07-28)	19	1341,8	1080,6	151	4443

One of the most interesting topics raised from the daily amount of tweets is the topic related to Sanna Marin, and I started collecting the keyword, Marin. Daily averages of how many tweets were posted are much more than any other topic. Sanna Marin's partying inspired many tweets, and an interesting fact about this is that even though the daily average is very high, the tweets were collected in 18 days (Table 2). This shows that the discussion did not end very quickly but continued much more extended time than other topics. The extended period is likely because the new information was coming out just as the old news started to fade away. This also can be noticed from Appendix 1 because on August 19, the daily amount of Tweets peaked, but the 24^{th} number of tweets was almost identical.

Time periodically, when the tweets have been collected, it is possible to notice that topics about the Nordstream explosion and Iranian women's protest are overlapping quite heavily. With the amount of discussion around Nordstream compared to the protest of Iranian women, it seems that the topic closer to Finnish people collected more attention and interest on Twitter. Interestingly it also appears that the discussion somehow projected the traditional media in Finland since the article covering explosions was more covered than protests in Iran.

The common factor between different sports events is the fact that there is a nationality factor that evolved in those. All three sports events that were trending and chosen to be researched also seem to have similar profiles when looking at the bar plots of Tweets per day. These events peaked one day, and then the discussion was much less inactive on other days. The days when the number of tweets was the highest follow the fact that Finland has been performing well since the Finnish national football team won on September 26. On the other hand, Kalle Rovanperä won the world championship on October 2.

The statistics about how many tweets have been tweeted show how much content has been produced on Twitter. Aside from that, it is good to highlight the fact of how different individual users have been creating the content on Twitter, meaning how many tweets one user has been tweeting on average. This type of descriptive stats helps to understand if a small group has been creating all the tweets and how topics have attracted engagement from many different people. The descriptive statistics about daily Twitter users are represented in Table 3.

Table 3. Statistics of Individual Users Tweeting per Day

Seed word (Collection Period)	Avg. Users Tweeting per day	Min Users Tweeting per day	Max Users Tweeting per day
Venäjä (2022-09-19-2022-09-24)	2 457	1 083	4 356
#Nordstream (2022-09-26-2022-10-01)	2 233	434	3 191
Huumeet (2022-07-21-2022-07-31)	521	223	1 207
Marin (2022-08-17-2022-09-03)	3 030	890	5 359
#RallyNewZealand (2022-09-26-2022-10-04)	162	3	848
Iran (2022-09-18-2022-10-11)	838	339	1 350
#Jodi (2022-10-05-2022-10-15)	454	142	1 656
Fortum (2022-07-31-2022-09-03)	416	26	1 531
#meidänkaikkienasia (2022-09-27-2022-10-04)	454	209	1 318
Fortum (2022-09-17-2022-09-24)	897	350	1 750
#Viisumit (2022-09-18-2022-10-01)	505	68	813
#Sähkönhinta (2022-08-20-2022-09-03)	421	56	839
#sectorallyfinland (2022-08-03-2022-08-06)	391	142	767
yle (2022-08-27-2022-09-11)	2 330	544	5 412
#huuhkajat (2022-09-18-2022-09-27)	226	26	624
Fortum (2022-07-10-2022-07-28)	686	99	1 860

With the average number of users tweeting per day, it is not surprising that the topics that have attracted more tweets on average have more individual users tweeting on average. This means that topics about war, YLE, and Sanna Marin are "outperforming" all the other topics. Something that can be elevated from these stories is that Sanna Marin's news attracted more of the same people to tweet again than war stories.

The topics about Fortum seem to perform a bit differently than any other topic since the average number of tweets is significantly higher than any different topic. However, the average number of users tweeting daily is still relatively low. This would mean that topics have been produced in tighter groups, and a broad audience has not been so interested in these topics.

4.3. Graph partitioning

After a graph is ready, a topic has been chosen, and a network is formed; the next step is to divide the graph, as explained in chapter 4.1. In this work, the focus is on polarization, meaning that only two communities are formed. The actual partition can be made using many different algorithms.

The logic of all the partitioning algorithms is the define every node to one community in the way that it is closer to that community compare to other communities. At the end of the day, earlier studies have suggested that the partition algorithm does not make that big of a difference and gives similar results (Salloum, et al 2022, 10; Garimella, et al 2018b, 10).

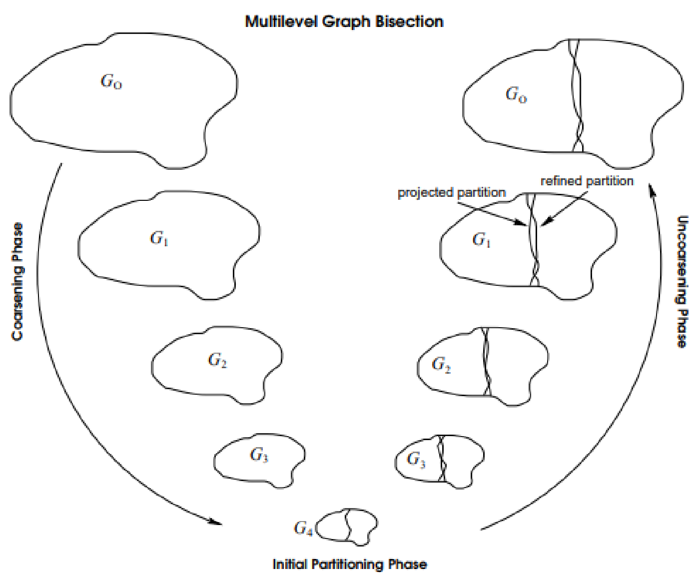


Figure 3. The first phase is the coarsening phase, where the size of the graph is decreased then the partition is done to a smaller graph and finally, partition is refined to a larger graph (Karypis and Kumar 1995, 2).

One of the most used algorithms seems to be an algorithm called METIS which was developed by Karypis and Kumar (1995). METIS optimizes the partitioning by trying to find two communities and simultaneously minimizes the connections between them (Salloum, et al 2022, 4). On its basic level, the process works by first creating a smaller graph representing the original graph and computing the bisection from that (Karypis and Kumar 1995, 2). The method of decreasing the size of a graph is done step by step, and increasing after partitioning is also handled in steps. After this, the partitioning is projected back towards the original graph piece by piece completing the partition simultaneously (Karypis and Kumar 1995, 2). This whole process is represented in Figure 3.

4.4. Measures

This chapter of the thesis introduces all the measures used and briefly describes the mathematical theory behind the polarization measures. Overall eight different scores are used to evaluate the polarization of networks. The common factor is that every score expects two different disjoint sets determined in the partition stage. These different sets represent opposing views on various topics.

4.4.1. Modularity

Modularity is one of the most popular measures when the goal is to explain division in the network. In the context of social media, Conover et al. (2013) used modularity to detect controversy, but the modularity score was used only to verify controversy. Later on, modularity has also been used to quantify controversy on social media (Waugh, et al 2009, 11-21; Salloum, et al 2022, 1-33).

The modularity score measures how different the communities are compared to the other communities in the network. Modularity is quite popular when evaluating neighborhoods because it is fast to calculate and, therefore, computationally easy to calculate (Newman, Mark EJ 2006, 8581). When modularity is used to quantify polarization, a calculation is relatively straightforward since it is only the formula of modularity that is used to evaluate communities.

$$P_Q = \frac{1}{2Card(E)} \sum_{\mathrm{ij}} \left(W_{\mathrm{ij}} - \frac{k_i k_j}{2Card(E)} \right) \delta(c_i, c_j), \tag{1}$$

where

E *is set of edges in network and* $Card(E)$ is the number of edges,

W_{ij} is the element of *the* adjacency matrix and

k_i is the degree of node i

When the nodes i and j belong to the same community, the value $\delta(c_i,c_j)$ equals one; otherwise, it is zero.

Modularity results are such that values are between -0,5 and 1, where the higher values imply that communities are separated and, therefore, more polarized (Waugh, et al 2009, 11). In many cases, the modularity score is combined with an optimization problem. Since the community partition has been made by using METIS, the optimization problem is not relevant anymore.

4.4.2. E-I Index & Adaptive E-I index

The E-I index is probably the most straightforward measure since it only calculates the ratio between the number of internal friendship links within the community and the number of external friendship links within the community. It has been introduced Krackhardt and Stern (1988) to argue that some social network structures are more effective than most organizations when confronting a crisis. That's why the E-I index might have been called the Krackhardt E/I ratio.

The E-I index and Modularity are somewhat different measures from others because, with these two, it would also be possible to quantify controversy in more than two communities if wanted. The E-I index formula can be written in a couple of different ways:

$$E - I\ index = \frac{EL - IL}{EL + IL},$$

(2)

where

$$EL\ is\ the\ number\ of\ external\ friendship\ links$$

$$IL\ is\ the\ number\ of\ internal\ friendship\ links$$

or

$$E - I\ \text{index} = \frac{Card(C)}{Card(C')}$$

(3)

where

$$E \; is \; the \; set \; of \; edges \; in \; Network$$

$$C \text{ is the cut } of \; set \; \{(s,t) \in E \mid s \in A, t \in B\}$$

$$C' \text{ is the complement of that set } (C' = E/C)$$

The E-I index fluctuates between -1 and 1 depending on the network (Krackhardt and Stern 1988, 127). As we can see from mathematical representation, the closer the index comes to -1 more separate the communities are since the internal links overpower the external links (Krackhardt and Stern 1988, 127). More internal than external links would mean communities are their units or, in the social media context, bubbles with a more polarized atmosphere. For example, the E-I index has been recently used to measure party polarization in the Netherlands (Esteve Del Valle, et al 2022, 736-755).

The E-I index is easy to understand, but it has some downsides, one being that it is not very good in situations where community sizes are heavily skewed towards others. Salloum et al. (2022) have extended the method to how the E-I index considers unequal communities and tackles this issue using the density of links. The extension is called the Adaptive E-I index, which accounts for the different group sizes by using the thickness of the ties within each block (Salloum, et al 2022, 6).

$$P_{\mathrm{AEI}} = \frac{\sigma_{\mathrm{AA}} + \sigma_{\mathrm{BB}} - (\sigma_{\mathrm{AB}} + \sigma_{\mathrm{BA}})}{\sigma_{\mathrm{AA}} + \sigma_{\mathrm{BB}} + (\sigma_{\mathrm{AB}} + \sigma_{\mathrm{BA}})}, \tag{4}$$

where

$$\sigma_{\mathrm{AA}} \text{ is the ratio of actual and potential links within } the \text{ community } A(\sigma_{\mathrm{BB}} \text{ similarly})$$

$$\sigma_{\mathrm{AB}} \text{ are links between communities } A \text{ and } B \text{ divided by all the}$$

$$\text{potential links (similarly } \sigma_{\mathrm{BA}})$$

Compared to the original E-I index, the Adaptive E-I index also ranges between -1 and 1, having the same type of interpretation. This means smaller values mean a more polarized network. If the situation is such that both communities are equal from a size perspective, then the score would be identical to the original E-I index.

4.4.3. Dipole Polarization

The first one to use the dipole polarization measure as a context for quantifying polarization on social media was Morales et al. (2015), which originated from physics. The dipole polarization score is based on label propagation, where a minority group of influential individuals propagates opinions through the social network for all the individuals (Morales, et al 2015, 2).

The assumption that the measure is built upon is the idea that there are two same-size and completely polarized communities. There are two sub-groups in each of these groups, S being elites, and L listeners, where the elites are so-called influencers and contain the top-k% highest degree nodes (Morales, et al 2015, 2). The first subgroup S, the elites, gets an extreme opinion value of -1 or 1, represented by X_s, and the second subgroup is labeled as neutrals meaning that the opinion value X_l is 0.

The so-called opinion generation part is such that elites influence through the network G listeners' opinions by changing their opinion values. The actual opinion values are calculated as the mean opinion value of incoming opinion values. The iterations are repeated so many times that the opinion value starts convergence.

The calculation of the controversy measure is done by first figuring out that the number of vertices $Card(V)$ with positive (n^+) and negative scores (n^-) and the absolute difference of their normalized size $\Delta A = \left|\frac{n^+ - n^-}{Card(V)}\right|$(Garimella, et al 2018, 14). From this can be calculated gc^+ (gc^-), which are average opinion/gravity center values among positive (n^+) and negative scores (n^-). These are used to figure out the distance between the two gravity centers, which compose set $d = \frac{|gc^{+} - gc^{-}|}{|X_{\max} - X_{min}|} = \frac{|gc^{+} - gc^{-}|}{2}$ being half of the gc^{+} and gc^{-} absolute difference (Morales, et al 2015, 3). From the combination of **Δ**A, the

absolute difference of n^{+} and n^{-} normalized size and d, the Dipole polarization is defined as MBLB = $(1 - \Delta A)d$.

The idea behind MBLB is that if two communities are parted, then label propagation assigns communities different extreme values X. Perfectly polarized situation means that MBLB equals 1. On the other hand, 0 would indicate a neutral position. Since variable **Δ**A, MBLB can be perfectly polarized only if community sizes are equal and the unbalance of communities decreases the polarization score (Morales, et al 2015, 4). Morales et al. (2015) used MBLB with real-world data when researching Venezuelan elections, and MBLB could detect polarization from network structure.

4.4.4. Boundary Connectivity

Boundary connectivity has been developed by Guerra et al. (2013). The idea of this measure is to examine vertices, especially the ones located within the boundary of two communities (Guerra, et al 2013, 215-224). The thought is that the polarization is more negligible if a social network graph is formed in such a way that most of the high-degree nodes are located at the edge of communities. The high-degree nodes at the edge of communities would indicate that the different communities are more connected and not separate entities.

The node u belongs to the boundary of X if and only if it is connected to at least one node of the other partition Y. After meeting this condition, the node u also has to be connected to at least one node in community X, which is not linked to any nodes in community Y. From this it is possible to define sets B_X, B_Y, which are containing all the boundary nodes in the community. This means that the union of these sets $B = B_X \cup B_Y$ is all the boundary nodes. The internal nodes $I_X \, and \, I_Y$ are a contradiction between all the community's (X or Y) nodes and boundary nodes. The union of internal sets $(I = I_X \cup I_Y)$ represents all the inner nodes.

The rationalization of Boundary polarization is that if two partitions represent two sides of a controversy, boundary nodes are more strongly connected to the internal nodes than opposite boundary nodes. Mathematically this can be formulated following way:

$$P_{\mathrm{BP}} = \frac{1}{Card(B)} \sum_{s \in B} \frac{d_I(s)}{d_B(s) + d_I(s)} - 0.5 \tag{5}$$

where

d_B is the number of edges between the node s and nodes in B

d_I is the number of edges between node s and nodes in I

The highest polarization score can rise by 0.5, and at its lowest, -0.5. As we can see from the mathematical formula, when controversy occurs, the value number of edges between a node and boundary nodes is low, resulting in higher polarization scores. More specifically, if the measure is more than zero, then on average, boundary nodes tend to connect internal nodes more often than opposite sides boundary nodes.

Below zero, the situation is the opposite, so boundary nodes are more connected to the other side than their own community's internal nodes. So higher values of boundary polarization measure indicate a more polarized social network. Also, Guerra et al. (2013) showed that polarized networks tend to have a low amount of high-degree nodes in the boundary between two sides.

4.4.5. Betweenness Centrality Controversy

Betweenness Centrality Controversy has been introduced by Garimella et al. (2018), which tries to measure the distribution of edge betweenness centralities. When a network is divided into two different communities, for example, X and Y, there are forming edges that are cut off. In the context of this work, a cut would mean the partitioning the

METIS algorithm produces. Betweenness centrality (BC) of a particular edge, for example, e, would be defined following way:

$$\mathrm{bc}(e) = \sum_{s,t \in V} \frac{\sigma_{s,t}(e)}{\sigma_{s,t}},$$

(6)

where

$$V \ is \ the \ set \ of \ nodes \ in \ the \ network$$

$$\sigma_{s,t} \ is \ the \ total \ no. \ of \ shortest \ paths \ between \ nodes \ s \ and \ t \ in \ the \ network$$

$$\sigma_{s,t}(e) \ is \ the \ shortest \ paths \ which \ include \ edge \ e$$

In a well-separated network, the structure is such that the cut should consist of edges that are "bridges" for the so-called structural holes. The rationalization of this measure is the idea that the shortest paths connecting two sides should go through edges that are only linked to the other side by these bridge edges. Betweenness centrality's definition makes it possible to notice that this structure would cause high BC values for the cut edges on the boundary. When comparing all the graph's edges and the cut set of edges in a polarized situation, cut edges should have higher BC values.

Following this same reasoning, in a situation where a network is not well separated, the cut set of BC values is compared to all the BC values. The only difference in a polarized case would be that BC values should be similar. Both distributions, edge centralities for edges in the cut and the distribution of edge centralities for the rest of the edges, are used to calculate the Kullback-Leibler (KL) divergence d_{KL}. After that, KL-divergence d_{KL} has calculated the PDFs for KL are estimated by kernel density estimation. Measure the BCC, which aims to quantify polarization by comparing the centralities of boundary and non-boundary links, can be represented following way:

$$\mathrm{BCC} = 1 - e^{-d_{\mathrm{KL}}}$$

where

$$d_{KL} = \sum_{x \in X} p(x) ln \frac{p(x)}{q(x)},$$

$p(x) and\ q(x)$ *being two probability distributions of a discrete random variable x* (7)

When the polarization increases, the distributions are more different, which eventually increases the factor d_{KL}. From formula 7, it is possible to notice that a polarized situation would mean values close to one. On the other hand, values close to zero would mean that distributions are more similar and that network would be less polarized.

4.4.6 Random Walk Controversy & Adaptive Random walk controversy

Garimella et al. (2018) have also introduced another metric to quantify network controversies. This measure is called Random Walk Controversy (RWC), and it uses random walks to establish polarization in a network structure. The idea of RWC starts from an assumption that large degree nodes are evidence of so-called influencers or authoritative users (Garimella, et al 2018, 1-27). On controversial topics, different sides are believed to have authoritative users of their own, and RWC measures the probability of being exposed opposite side's influencer (Garimella, et al 2018, 11).

Graph G, built in this case from retweets, is divided into two communities, X and Y, by METIS at the second part of the process. Calculating the RWC score starts by defining the influencers, some k-highest degree nodes from both sides. So a little bit same type of step when calculating the Dipole moment. The degree in both measures is considered a sign of popularity since these users have more endorsements than others. To choose k, there are not any recommendations made by Garimella et al. (2018), but in some later works have been used a single value k=10, so it will also be used in this work (Salloum, et al 2022, 23).

After authoritative users have been determined, a node where a random walk starts must be picked. This happens firstly by giving each side 50-50 chances and then randomly picking one user from that community. From there, the random walk continues as long as it encounters one of the influencer nodes and stops. It does not make any difference if an influencer is on the same side or the opposite side.

Garimella et al. (2018) define the RWC measure in the following way: *"Consider two random walks, one ending in partition X and one ending in partition Y, RWC is the difference of the probabilities of two events: (i) both random walks started from the partition they ended in, and (ii) both random walks started in a partition other than the one they ended in."* The formula for this measure can be created following way.

$$\mathrm{RWC} = P_{\mathrm{XX}}P_{\mathrm{YY}} - P_{\mathrm{XY}}P_{\mathrm{YX}} \tag{8}$$

where

P_{XX} is the probability for a random walk ending side X when it has started from side X

P_{XY} is the probability for a random walk ending side X when it has started from side Y

The scores of this measure fluctuate between -1 and 1, where the higher values mean a more polarized network. In a fully polarized situation, the right side of subtraction in formula seven would be one leaving the left side as zero. So all the random walks would end up inside their community without ever meeting another side's influencers.

As earlier stated, the number of influencers (k) affects the result and can play quite an important role, but there are no guidelines on choosing the k value. Salloum et al. (2022) have developed an adaptive version of the RWC measure to tackle this issue. They suggest that the k value should be changed based on the number of influencers in a community. This would happen the way that K would represent the fraction of influencers from each side, and then selecting k for community X would happen $\frac{K}{n_x}$, and the same way would k be set for community Y $\frac{K}{n_y}$. Since the Adaptive RWC (RWC, too) is also sensitive to the k-value, the comparisons of the results are challenging. Still, on the other hand, qualitatively, the final scores would be more stable and not that sensitive to small changes in the actual parameter value K.

4.4.7 Spearman rank correlation

When the term correlation is used usually it is referred to linear relationship between two continuous variables which means more accurately Pearson correlation (Schober, et al 2018, 1763). The relationship can be such that if one variable increases the other increase too or such that increase of one variable means that other variable decreases. The way Spearman rank correlation can be described is such that it is a Pearson correlation coefficient calculated ranks of the values two variables instead their actual values (Schober, et al 2018, 1766). So the difference is the fact that values are not anymore continuous but ordinal. This means the use cases of Pearson and Spearman correlation are little bit different.

The results of Spearman rank correlation fluctuate between -1 and 1 where more closer the result is ±1 stronger the relationship is (Mukaka 2012, 69). Spearman rank correlation between two variables can be calculated using following formula

$$r = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2 - 1)} \tag{9}$$

where

$$d_i \text{ is the difference in ranks for } x_i \text{ and } y_i$$

5. RESULTS

This chapter includes an analysis of the data and results which have been generated. The results chapter has three parts, each with different points of view to examine social networks. The first part consists more of content perspective, the second part is visual examination, and the third chapter has the results of polarization measures.

5.1. Descriptive statistics of Graphs

The first step in understanding possible polarization in different topics is to examine descriptive statistics of various topics. The amount of tweets makes it hard to inspect all the individual cases, which is why the examination focuses on the most influential parts. This way, it is possible to understand the subject through the most influential parts of networks. In this chapter, the hope is to understand what type of and how words and text are used, the most popular tweets based on retweets, and who are the most influential users. The focus of this part is to focus on the topics themselves so that the graphs are not only nodes and edges.

5.1.1. Hashtags

Hashtags are used to categorize a tweet posted by a specific user, which means that hashtags are an excellent way to find different categories. Also, adding the hashtag makes it easier for other like-minded users to find the tweet and be liked or retweeted it. It is expected that keywords used to form topics are more common than other words since almost every tweet has to contain one of the search words used in queries. The most used hashtags usually find a way to the trending page, which adds to the popularity of these words. The lists of the most used hashtags can be found in Appendix 3.

In some cases, the hashtag used can clearly describe a user's sentiment. On the other hand, in some situations, the hashtag is very neutral and does not highlight any view. In this work, good examples of these neutral hashtags are something like #Venäjä (in English #Russia), which primarily refers to the country called Russia. An example of a more sentiment hashtag is #meidänkaikkienasia, which refers to the shortage of healthcare workers in Finland as a matter for all the citizens. It is much more supportive than, for example, #hoitajapula (lack of healthcare workers).

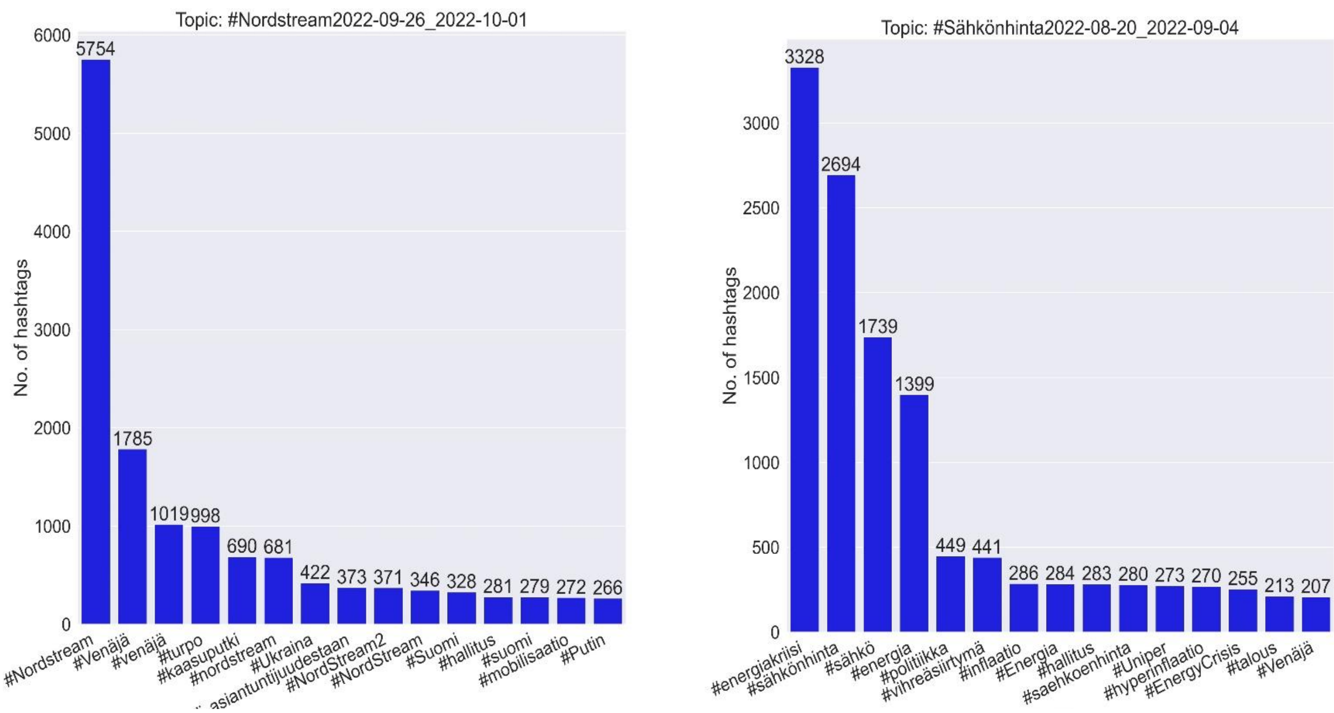


Figure 4. Top 15 hashtags on topics Nordstream and high electricity prices

Regarding the hashtags of the topics, it is possible to notice that usually, the most used hashtags are short and compact. More often, the more complex hashtags seem not to be supported by big audiences. The available amount of hashtags seems to be somewhere between 2 and 3, and it appears that trending topics have 2 or 3 more used hashtags, and other hashtags are much less prevalent.

Indeed some cases have much more popular hashtags, which might result from the topic having many different storylines and creating more different types of hashtags. A good example is Sanna Marin partying since the new revelations meant that Twitter users developed new hashtags to describe the situation. So first, the hashtag called #Jauhojengi, which was a reference that some users heard from the leaked video, peaked at first, but usage did fall. After the first leaked video, another leak created #Kesärata, a word from the second video. The usage of #Marin was used much more evenly, resulting in #Marin being a more used hashtag at the end.

It is expected that the words used in search queries appear more often on tweets that were more or less general words about the trending topics. But still, outside of these search words, general and descriptive words without too many statements or opinion about the situation seems to be more popular. What is meant by this is that, for example, in the topic of the NordStream explosion, the most used hashtags are general hashtags about Russia, Ukraine, the Baltic sea, and security politics (Figure 4). Hashtags about shutting the border or slurs about Russia are much uncommon. So it seems that the collected datasets contain more composed tweets than hateful ones, at least from the hashtag point of view.

Even though the hashtags seem relatively neutral in many cases, hashtags associated with Finnish politics are prevalent and often used in topics that might imply politics on some level, such as electric prices (Figure 4). So topics about sports do not contain political hashtags. Political hashtags appearing almost always could be the implication that many issues politicize quickly, and this way creates the borderlines between different viewpoints.

The absolute number of used hashtags is quite heavily connected to the number of collected tweets, which is not surprising given that more tweets and retweets are linked to the number of tweets collected. More interestingly, the unique hashtags are not related to the number of hashtags, and it looks like more broad search queries and topics where the topics were not that specific have resulted in more unique hashtags. On the other hand, it seems that issues more or less originated from social media (or hashtags have been the first keyword), like topics about citizens' initiative of drugs, talk about iodine or nurses' collective agreement. Hashtag usage is more constricted in these cases. If the collected topic has had some hashtags which have been identified, the use of the same popular hashtags seems to have been much more regular.

The correlation matrices of popular hashtags present the way hashtags are combined when used in tweets. All the correlation matrices can be found from appendix 4.To limit the number of hashtags, the popular hashtags, in this case, have over 100 appearances. The negative correlation between hashtags seems to be more heavily focused on those hashtags where the same thing has been written using different appearances (Appendix 4). So for example, #Tehy and #tehy or #huuhkajat and #Huuhkajat.

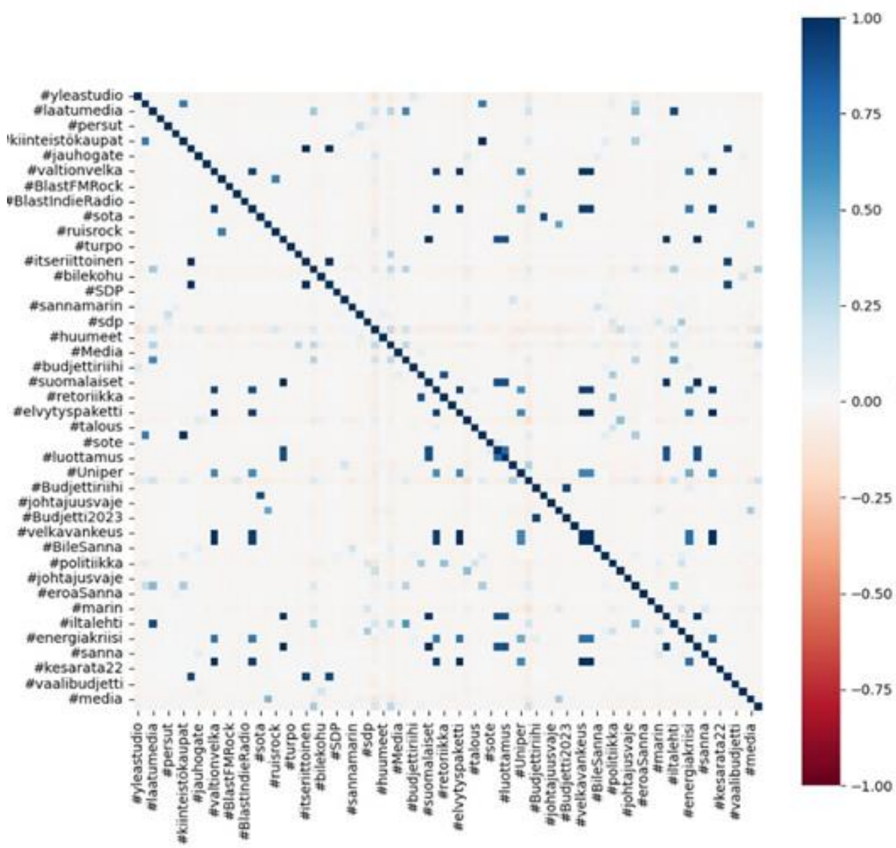


Figure 5. Correlation matrix (Marin2022-08-17_2022-09-05)

Another fact about the correlation matrices as a whole is that only very few hashtags would not have any positive correlation to any other popular hashtag. So the more used hashtags are very often used with other popular hashtags. One exception is the Iranian riot's topics since popular hashtags do not correlate. In the subject of Iranian riots, no positive correlation would suggest that people are more often only using one single hashtag or some less popular hashtag. The case of Sanna Marin is also quite different since it has very little negative correlation, and the correlation is positive, or it does not exist (Figure *5*).

5.1.2. Tweets and Retweets

If the hashtags are part of the tweets that categorize the tweet and make it easier for other users to find and classify it, it cannot give the whole story of the tweet's content. In many cases, the hashtags can be relatively neutral and not provoke an opinion, but the content can be aggressive.

Overall, the number of retweets on each topic is entirely aligned with how many tweets have been gathered, which is not surprising since the number of tweets is all the original tweets and retweets. The average number of retweets per tweet is 1.21, so there is over one retweet for every tweet created. Especially topics about Russian tourist visas and Finnish drug policies, the amount retweets is significantly higher than the number of original tweets. So most people are not creating any new content on Twitter but just re-peating smaller groups of people's points.

Figure 6. Worclouds of Fortum and Russian tourist visas

Sports-themed tweets have many English-written tweets that have been popular from a retweet perspective. On other topics, the international tweets that Finnish users have retweeted are much fewer. The most popular tweets based on retweets contain many that focus on criticizing some political person, party, or group. Tweets about Fortum are a perfect example of a situation where the critic concentrates heavily on either the right or left political side. The subject of these popular tweets about Fortum is almost unilaterally trying to find somebody to blame. The blame for these Fortum cases is directed toward politicians. Even though there are tweets about the parties, Prime Minister Sanna Marin and Minister of European Affairs and Ownership Steering Tytti Tuppurainen are often recurring.

The topic about Fortum and the bad decisions around it is not the only topic where the most popular tweets-based retweet count seems to contain a negative sentiment. It appears that if the sports topics are not considered, the subjects of tweets are very accusatory and try to discredit some sides. The tweets about new Iodine recommendations sound plain on paper. Still, this topic has also been tweeted about how the ministry of social affairs and health is deceiving the Finnish people.

Another fact that is also worth mentioning is that the most retweeted tweet about the rally in Finland is about the fact that Finnish climate activists were attacked. So interestingly tweet that is dividing people very effectively and is not about the rally itself is getting more attention than any other tweet. The tweets about the Iranian protest have quite interesting content since the most popular tweets focus on women in general, and multiple tweets do not favor current developments. The tweets about actual protests or situations in Iran seem to have a minor role on Finnish Twitter.

The issue about the collective agreement of Finnish nurses seems to have the most positive tweets about the whole situation, and the focus of tweets is primarily celebratory. Tweets highlight that the money is finally going to the right place, and there seems to be quite a unified opinion about it. The discussion around the new drug policy in Finland has a bit different theme compared to other topics since the most popular tweets are more about the topic and less about attacks towards something or somebody. Also, there are clear messages and encouragement to act and sign the citizens' initiative.

The analysis of the most retweeted tweets exposes the fact that popular tweets are usually offensive towards some party (not necessarily political) and try to bring out some mistake that has been made. For example, Nordstream's most retweeted tweet has significantly more retweets than any other tweet, and it only shares an old video about politicians who have previously supported the Nordstream gas line. The same type of theme has tweets about Fortum between September 17th and 25th, where the Fortum CEO's statement about Russia and how they are not going use energy as a weapon.

The supportive tweets have not brought as much attention to the tweet as criticizing collected topics, making it less appealing to tweet this type of tweet in some ways. Evaluating the likes of most retweeted tweets are almost always higher than retweets, but in some topics, the count of likes and retweets are not aligned. In these situations where the more retweeted tweet has fewer likes than another tweet, the more liked tweet has some more extreme opinion or confrontation.

For example, in the iodine discussion, where the most liked tweet is about how older people have most likely hoarded all the pills even though they didn't benefit from the additional amounts, based on retweets, the tweet is only the fifth most popular. Another example where this type of behavior comes true topic about Fortum. The tweet about Sanna Marin only partying during the help package negotiations and not taking care of Finland's interests collected more likes than a tweet about Fortum's careless decision to buy Uniper. Even though based on retweets, the order was different. So if the tweet's popularity is measured based on retweets or likes, it is not just retweets encouraging to create more controversial tweets, but likes seem to create the same type of conditions.

Changing a bit of perspective and focusing more on the tweet's creators gives a better understanding of what type of users usually get more attention and if it matters to have, for example, a verified account or a specific number of followers. The user information was checked during October 2022, so the changes to Twitter's verification policies do not affect user data. So the verifications in this work mean that verified account is active, notable, and authentic accounts of public interest that Twitter has verified based on specific requirements. The analysis focuses on the 15 most retweeted users in each topic, where the first thing noticed is that five users have been banned. These users have been prohibited between the data collection period and the end of October 2022.

From the most retweeted users, only about every fifth user has a verified Twitter account. So it seems that it does not make that big of a difference if you have a verified account or not, and indeed, the market is much smaller, which certainly affects Twitter's way of granting verified marks. From the follower counts, it is easy to discover because over a hundred thousand followers have only 22 users in the data. At the same time, for example, Barack Obama has about 133 million or Elon Musk 114 million and counting. This scale difference means that everything is much more local, and worldwide impacts are rare.

The most retweeted accounts are very heavily emphasized users with over a thousand followers but not much over 20 000 followers. On average, there are about 16 000 followers. But even though there are under thousand follower accounts, there does not seem to be any clear pattern in favor of more followed accounts. Accounts with just over a thousand followers can have the same amount of retweets as accounts with tens of thousands.

Table 4. The most retweeted users on Marin Case

Topic (Collection Period)	User	Count of retweets	Followers	Followed	Verified	Banned Account
Marin2022-08-17_2022-09-05	pbyrokraatti	2 205	40 753	996	FALSE	-
Marin2022-08-17_2022-09-05	AuteroMiia	882	11 028	1 789	FALSE	-
Marin2022-08-17_2022-09-05	Dimmu141	628	47 201	8 822	FALSE	-
Marin2022-08-17_2022-09-05	iinapalo	533	8 725	244	FALSE	-
Marin2022-08-17_2022-09-05	VilleTavio	532	23 878	260	TRUE	-
Marin2022-08-17_2022-09-05	Jrvinen_J	512	15 187	16 188	FALSE	-
Marin2022-08-17_2022-09-05	NuuniUnna	504	2 218	2 836	FALSE	-
Marin2022-08-17_2022-09-05	petteri_vaara	499	1 843	1 301	FALSE	-
Marin2022-08-17_2022-09-05	MatsUotila	461	23 184	997	FALSE	-
Marin2022-08-17_2022-09-05	seiska	461	49 906	408	FALSE	-
Marin2022-08-17_2022-09-05	HannuJarvinenPS	443	2 711	2 791	FALSE	-
Marin2022-08-17_2022-09-05	HuhtasaariSaara	438	12 243	757	FALSE	-
Marin2022-08-17_2022-09-05	Valavuori	435	152 060	516	FALSE	-
Marin2022-08-17_2022-09-05	aj_hirsi	434	7 633	1 181	FALSE	-
Marin2022-08-17_2022-09-05	JariTervo1	412	106 328	4 391	FALSE	-

One common factor there is with especially political topics. The accounts that actively participate in political discussions and exercise political commentary are very much part of the majority of most retweeted accounts. So these users are not all politicians but journalists or sideline commentators. On some direct political topics, users who collected the most retweets quite unilaterally also represent the same values and have the same type of ideas, and the theme of tweets is quite populist.

An excellent example of this is Sanna Marin's leaks of partying, where most of the accounts are conservative and more or less right-leaning (Table 4). This same dynamic repeats with Fortum and the prices of energy topics in which similar users represent the most retweeted accounts, as in Sanna Marin's case. It looks that in data and political discussions, the more or less oppositive view of account has been much more popular than supportive.

The drug topic is also quite interesting because there are much smaller accounts on top of the most retweeted chart. Multiple issues have fewer tweets and retweets, like prices of energy, nurses' new collective agreement, or sports, than the topic of drugs, and all these have, on average higher follower counts. It is fitting because the subject of drugs is mainly about the citizen's initiative; in a way, it would be expected that the accounts are smaller and less well-known. Russia's most retweeted accounts seem to have the opposite type of situation where the amount of retweets and tweets is not the highest. Still, on average, if the international accounts are not considered, the most retweeted accounts are over 50 000 followers when any other topic does not get over 20 000.

5.2. Structures of Graphs

Before entering to analyze polarization measures, the networks themselves are worth investigating. By investigating the networks, it is possible to get an idea of the type of structure they have and what type of interactions there are. Analyzed networks being retweets networks means that nodes of these networks are accounts/users, and edges between these nodes are retweet interactions between users.

In this part of the work, the visualizations are made using Gephi and some of its built-in features. The graphs are partitioned using METIS, and the results are imported to the Gephi as labels of nodes. The benefit of doing this step is that the nodes can be colored to represent their actual side cluster. Even though the Gephi is a convenient tool for getting an idea of what types of networks are dealt with, the differences between different networks can be tricky to tell, especially if there are many nodes and edges. So to help understand the characteristics of networks, some metrics will be calculated to help describe networks.

5.2.1. Visualisation of Networks

The networks from a visual point of view and statistics later on reveal very quickly that the networks are much smaller than the data described earlier. All the visual representations of networks can be found in Appendix 5. The size reduction is happening because of the earlier assumption that the connection (edge) between two users (nodes) is only made if there are two or more retweets between users. Most significantly, this impacts those users who are on more informative channels and do not have that much interaction and engagement with other users. An excellent example of this impact is an account called @Puolustusvoimat (Finnish Defence Forces), which almost completely vanishes from the Russia topic even though it was the most retweeted account. Also, sports topics have many of these types of reports which accounts have retweeted only once in the dataset.

Another reason that diminishes the network sizes is that real-world data creates small components around the more extensive and connected component, which can be called a giant component. Even though the smaller components could be a very isolated group of users, the sizes of these small components are so small compared to the main network that it would be tough to make partitions and calculate polarization measures. That’s why these smaller components, which are not connected to the main network, are just increasing the noise of the graph, and that’s why it is most beneficial to focus on only the most significant part of the connected network.

One thing which is good to note from looking at the networks is that the sizes differ quite a much. This is to be excepted considering the amount of data but is good to raise since it makes the characters of networks also quite different. The three sports topics are the three smallest graphs based on nodes. Since the number of nodes and edges is relatively small, it is easy to see spot connections between nodes and structure. Mainly the issues about New Zealand rally and Huhkajat have similar systems in which one to three nodes have a significantly higher degree, almost like hubs, and around these are all the other nodes.

Especially the topic of the Finnish national football team has a node, which has the edge over 81% of nodes. Something interesting about the New Zealand rally network is that there are three nodes with a higher degree, but these top three nodes are not linked to each other. Based on structure, these smaller graphs don’t seem to have much of that theoretical look where two clusters are weakly attached, making the partition to two groups look quite forced.

Moving on to the more extensive graphs, it becomes much more challenging to notice similarities and differences between graphs. Still, the visualizations give an excellent opportunity to see individual networks and how the communities will form. Having the hubs does not seem to vanish even though the networks are growing.

Figure 7. Retweet Network about Topic Marin

Based on the network structure, firstly, groups are portrayed more or less in the traditional structure of polarization, where two groups are connected loosely (Figure 7). This design's topics are tweets about Fortum in July and September, videos about PM Sanna Marin, women’s protests in Iran, and Russia’s attack on Ukraine. However, the last one has a relatively small minority compared to another group of nodes. Especially the topics which have raised many interactions are very clearly creating two opposing sides, which are in itself very tightly connected. From the labels, it is hard to assess partitions because artificial change can significantly impact the visual representation, which means how the nodes are scattered in the graphs or how the edges are drawn. Nevertheless, it would seem that greater networks, such as the topics about Sanna Marin and Russia, are more aligned with institutions.

Figure 8. Retweet Network about topic Russia

The structure which is close to this classic structure is such networks where two different groups can be identified. Still, one of the groups is almost entirely connected to one node with connections to the other (Figure 8). A structure of this type would mean that most users have tweeted only one person's tweets without interacting with each other, but another class still would have more broad interactions. The topics with this structure are the Nordstream attack, Fortum in August, and Russia.

As earlier described, the smaller graphs where the amount of nodes is small are well connected, but from more extensive networks, some graphs are also quite well connected. So addition to the sports topics is energy prices and YLE, which are compactly packed and more or less represent the traditional way of a non-polarized network. Meaning that the network does not contain loosely connected two distinctive clusters. Especially topic of YLE is interesting because the nodes have been gathered around one node (Appendix 5). One reason could be that the subject was collected using relatively sparse keywords and a prolonged period creating many smaller sub-graphs. The unconnected sub-graphs were removed when picked the giant component of the network.

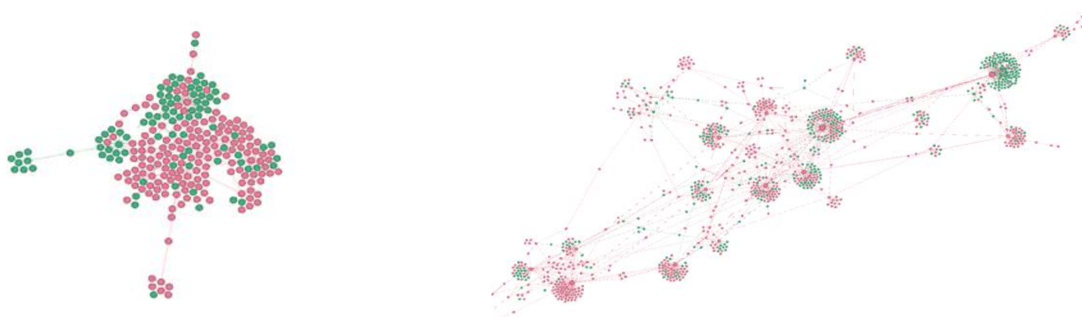


Figure 9. Retweet Networks about Finnish electric prices and Russian tourist visas

There are also topics, such as the discussion about Ionide or Russian tourist visas, which visually portray some ways each other but do not have similarities with earlier graphs (Figure 9). In these networks, there is almost a middle part or center where many nodes are connected tightly together and, from this central part, form subgroups that do not connect to the other side's subgroup. In these situations, it is almost like there would be three different discussions or groups forming where the middle part connects two sides. In these situations, it seems that it does not align with the traditional view of polarization, where there would be two almost disconnected clusters. Since the measure are mainly designed to measure polarization between two groups, the design of graphs might raise some problems with measures.

5.2.2. The key statistics of graphs

The graphs are a good way to inspect and get an understanding of networks. Still, the comparisons are hard to make since visualization can look very different depending on how nodes are settled down or which edges are focused on. Also, the size can make it hard to find similarities between the graphics. That's why some statistical numbers are calculated to support visual inspections and can be found in Table 5.

The size of the graphs can be defined based on two different and simple measures: the number of nodes and edges. All the graphs in this work are undirected, meaning that the edge between two nodes does not have a direction property, and two nodes can

Table 5. Network summary statistics

Topic (Collection Period)	Number of Nodes	Number of Edges	Average Degree	Density	Diameter	Triadic Closure
Venäjä, 2022-09-19 - 2022-09-24	824	1 236	1,500	0,00365	12	0,00246
#Nordstream, 2022-09-26 - 2022-10-01	2 092	2 863	1,369	0,00131	11	0,00076
Huumeet, 2022-07-21 - 2022-07-31	171	270	1,579	0,01858	8	0,10971
Marin, 2022-08-17 - 2022-09-03	2 852	8 432	2,957	0,00207	13	0,02439
#RallyNewZealand, 2022-09-26 - 2022-10-04	174	192	1,103	0,01276	7	0,00247
Iran, 2022-09-18 - 2022-10-11	2 006	3 050	1,520	0,00152	20	0,00706
#Jodi, 2022-10-05 - 2022-10-15	1 022	1 209	1,183	0,00232	11	0,00086
Fortum, 2022-07-31 - 2022-09-03	1 001	1 573	1,571	0,00314	10	0,00245
#meidänkaikkienasia, 2022-09-27 - 2022-10-04	206	376	1,825	0,01781	6	0,07822
Fortum, 2022-09-17 - 2022-09-24	324	407	1,256	0,00778	13	0,01307
#Viisumit, 2022-09-18 - 2022-10-01	1 162	1 616	1,391	0,00240	10	0,00259
#Sähkönhinta, 2022-08-20 - 2022-09-03	249	302	1,213	0,00978	10	0
#sectorallyfinland, 2022-08-03 - 2022-08-06	66	90	1,364	0,04196	7	0,03279
yle, 2022-08-27 - 2022-09-11	636	775	1,219	0,00384	11	0,00520
#huuhkajat, 2022-09-18 - 2022-09-27	95	96	1,011	0,02150	7	0,00101
Fortum, 2022-07-10 - 2022-07-28	1 466	2 655	1,811	0,00247	11	0,00938

have only one link. The amount of gathered data does not give the best picture of the networks' size since data cleaning diminishes networks significantly. The most significant deflection is the topic of YLE, which initially contained many tweets which were not about the Finnish public broadcasting studio but more or less linked to K-pop. In total, 7/16 topics have over a thousand nodes, so over a thousand different users. And all of these topics plus one has over a thousand edges meaning over a thousand links (two or more retweets) between users. There are also two networks under 100 nodes and edges at the end. These topics should be interpreted carefully later on since the result of polarization measures can be vulnerable to tiny changes in small graphs.

The shortest paths of a network can be summarized using a measure called the diameter, which is simply the longest of all the shortest paths. This measure aims to give a sense of the network’s overall size. From diameter, there is a possibility to illustrate the size of a network and shape from the other point of view. A network must be connected to calculate the diameter.

The topic with the highest diameter is the Iran protest, which is not the biggest surprise since the graph structure has some very disconnected users. From the issues collected, the higher diameter figure seems to get also Marin (13), Fortum in September (13), and Russia’s attacks (12). But all in all, nine topics have a diameter between 10 and 13. Something interesting that can also be noticed is that Fortum collected in September has significantly fewer nodes and edges than other top four topics. The minor diameter figure is held by the topic of Finnish nurses’ agreeing on a new collective agreement (#meidänkaikkienasia). So in this topic, six interactions separate even the farthest-located users making the most compact network of users.

The number of nodes and edges can be calculated degree of nodes and the average degree for the network, which is the average ratio of edges and nodes. The only topic which has over two average degrees is Marin’s partying. On that network, there would be, on average, almost three (2.96) edges from every node (Table 5). Compared to the average degrees, the nurse’s new collective agreement, Fortum, and drug policy have higher average degrees over 1.5 edges (Table 5). It would suggest that in these networks, the interactions between nodes (users) are more likely, so users are more actively promoting others' tweets. From the bottom of the table are sports topics where the average degree is close to one, and links between two users are much less likely.

Something to measure how knit the networks are, there is an option to calculate a metric called density which measures the ratio of actual edges that which network contains and all the possible ones. It can be quickly noticed that the options are high, especially when the number of nodes goes up. The number of nodes that make the density depends on the size of a graph since the possibility of having an edge is more limited for smaller graphs. Still, looking at the whole, the densities are less than 0.04, indicating that the collected retweet networks are not even close to perfectly connected networks where the density would be 1. The low density shows well that users do not widely share the support, and most likely, one of the reasons is that there is so much content.

The last structural calculation and figure calculated is triadic closure, which supposes that if there are connections between two people, A and B, to one person, C, there are most likely connections between these two people. Meaning that the triadic closure is all about the complete triangles which are forming the networks. In measuring triadic closure, a simple measure has been developed called transitivity. Transitivity represents a ratio between all triangles on a network and all possible triangles getting values between 0 and 1. As well as densities, the transitivity numbers are low in many cases, even though some topics are getting higher numbers than the density. Low density could explain this since there are fewer possibilities to form triangles.

The highest transitivity is getting topics about drug policy and also nurses' collective agreement. From the visual representations of the graphs, these topics are more tightly packed networks, and also statistics seem to support that fact. Transitivity also states that the topic of electricity prices has zero transitivity. The result could only be possible if the actual network does not contain triangles. Overall the low numbers of transitivity also indicate that even though users are retweeting some person's tweet much less often, they retweet other people who have retweeted the same tweet.

5.3 Polarization measures

The two previous chapters create the image of what type of networks are dealt with and help the understanding of polarization measures. In the beginning, it is good to note that the networks are different sized, and the features of networks are different. This makes it hard to compare graphs, which means there is no way of telling that some topic is more polarized than others. But still, it is possible to explore the different polarization measures and how the results of these measures portray the networks' polarization.

Table 6. Polarization score summary

Topic (Collection Period)	Betweenness	E-I Index	Adaptive E-I Index	Boundary Connectivity	Dipole Moment	Modularity	RWC	ARWC
#huuhkajat, 2022-09-18-2022-09-27	0,796	0,729	0,685	0,133	0,168	0,234	0,740	0,580
#Jodi, 2022-10-05-2022-10-15	0,812	0,909	0,911	0,097	0,628	0,425	0,861	0,667
#meidänkaikkienasia, 2022-09-27-2022-10-04	0,905	0,697	0,658	0,089	0,389	0,250	0,806	0,275
#Nordstream, 2022-09-26-2022-10-01	0,760	0,864	0,858	0,166	0,829	0,391	0,837	0,837
#RallyNewZealand, 2022-09-26-2022-10-04	0,619	0,906	0,906	0,197	0,819	0,437	0,951	0,729
#sectorallyfinland, 2022-08-03-2022-08-06	0,566	0,778	0,769	0,167	0,363	0,326	0,905	0,552
#Sähkönhinta, 2022-08-20-2022-09-03	0,506	0,821	0,816	0,306	0,461	0,359	0,873	0,353
#Viisumit, 2022-09-18-2022-10-01	0,781	0,874	0,866	0,187	0,833	0,381	0,877	0,846
Fortum, 2022-07-10-2022-07-28	0,729	0,892	0,889	0,225	0,816	0,419	0,785	0,716
Fortum, 2022-07-31-2022-09-03	0,643	0,898	0,885	0,235	0,750	0,342	0,899	0,857
Fortum, 2022-09-17-2022-09-24	0,941	0,966	0,966	0,285	0,774	0,457	0,945	0,540
Huumeet, 2022-07-21-2022-07-31	0,855	0,830	0,811	0,211	0,312	0,318	0,875	0,306
Iran, 2022-09-18-2022-10-11	0,914	0,932	0,932	0,204	0,911	0,457	0,889	0,882
Marin, 2022-08-17-2022-09-03	0,918	0,979	0,973	0,281	0,637	0,312	0,879	0,886
Venäjä, 2022-09-19-2022-09-24	0,844	0,864	0,850	0,159	0,545	0,338	0,764	0,680
yle, 2022-08-27-2022-09-11	0,865	0,928	0,928	0,142	0,622	0,422	0,887	0,727

5.3.1. Modularity

Starting from one of the most popular scores used to describe social networks, we can evaluate the discrepancy in a social network. Looking at the networks, it is possible to notice that modularity scores do not climb over 0.5 but are still between 0.3 and 0.45. Other studies have stated that modularity scores between 0.3 and 0.7 in practice suggest that the network has a strong community structure (Newman, M. E. J. and Girvan

2004, 8). A strong community structure would mean that the users are more affected by selective exposure and living in their bubbles. The topics clearly below this range and would not contain separate communities based on modularity score are networks about a Finnish football team and nurses' new collective agreement. From the visual representation, it also seems to be quite hard to notice any separations between communities (Table 6).

Modularity scores being high for networks of Iran protests and Fortum (9/2022) seems, also to be aligned with the visual representations. Single detail but some ways quite interesting that pop up from the modularity scores are the cases about Sanna Marin. The score (0.3124) is very close to the lower threshold based on Newman & Girvan's (2004) definition of solid communities in practice. Even though the content of the tweets and visuals, the network seems to have relatively clear signs of different blocks. The other thing that is good to point out is that no topic would get over 0.5. If the maximum score for modularity is 1 and 0.7 in practice, then subjects would not be very heavily divided based on modularity.

5.3.2. E-I index & Adaptive E-I index

The results of E-I indices are quite problematic and give quite contradictory signs compared to the modularity and other scores. The E-I index is calculated, making it very heavily dependent on the METIS and how the nodes are divided and labeled to the groups. From the visual representations, it is clear that the partition is not perfect, and maybe one of the reasons is the fact that METIS is forced to divide the network into two different clusters and not 3 or 4. Nevertheless, since the E-I index calculates the sums of internal and external links and bases the final score on that is not working very well.

Basically, the scores are telling that the external links, so relations between two different groups where the users have been divided, are heavily dominating the link types. This means that both E-I indexes are almost all between 0.7 and 1, the maximum score for a heterogeneous network. Looking at the adaptive E-I index scores and comparing those to the original measure, the differences are slight, so considering the network size doesn’t change the results.

5.3.3. Dipole polarization

The dipole polarization results are much closer to the modularity scores, but the fluctuating or differences in values are bigger. For the dipole polarization, one means that the network is perfectly polarized and zero a non-polarized network. The measure variation on collected topics is 0.1684 and 0.9106, so based on the dipole moment, there are some non-polarized networks and some polarized ones. The topic of the Finnish national team gets the lowest value of the topics and adds proof that the topic would not be polarized. The topic of drugs seems to be also getting quite low value as well as Secto rally in Jyväskylä and nurse’s collective agreement. The visual representations of all four networks are quite tight, and differentiating two noticeable groups are quite hard. Also, the modularity scores are quite low.

The topics that are set up closer to the one and being polarized topics based on the measure seem to be Iran protests, Russian tourist visas in Finland, Nordstream attacks, and, a little bit surprisingly, Rally New Zealand. Looking at the visualizations of the networks in the first three topics, there seem to be those two communities that are weakly connected to each other. From the subject's point of view, the Russian tourist visas and Iran protests could be expected to be divisive topics. From the subject or structural point of view, Rally New Zealand getting a high dipole polarization score and being labeled as a polarized topic is quite unexpected. However, the modularity score of the topic is also giving indications of a strong community structure. Fortum stories are also getting quite a high dipole polarization score and could be described as polarized than non-polarized based on the score.

5.3.4. Boundary Connectivity

The boundary polarization scores all over zero indicate that in every network, the nodes in the boundary tend to establish more in-group connections (Guerra, et al 2013, 221). Especially those cases where the score is significantly higher than zero. Looking at the boundary polarization scores, some changes are following the modularity and dipole polarization scores. Meaning the scores seem to give somewhat similar results whether or not the network structure has signs of polarization.

The topics with boundary polarization clearly over zero are electricity prices in Finland, discussing Fortum in September, and Sanna Marin’s leaks. The score in these cases indicates that, on average, nodes, the boundary tends to connect internal nodes rather than nodes from the other group, indicating antagonism. From these topics, which have over 0.25 scores, Fortum and Marin are visually giving the indications of polarized networks, but from the electricity price graph, it is quite hard to find a polarized structure. The size of the graph can be one of the things which makes it hard to find a divided structure visually.

Overall the scores of boundary polarization are quite aligned with earlier results, but something a little bit contradictory results is the topic of drug policy. Dipole polarization does classify the structure of the network as non-polarized, but boundary polarization defines the network as more polarized. Visually the network looks quite well connected, making it hard to find any split to communities. The other topics which were earlier defined as non-polarized are topics where nodes on the boundary have the same likelihood of establishing their own group or opposite group.

5.3.5. Betweenness centrality controversy

Betweenness centrality controversy (BCC) which based on the fact of how the cut is affecting betweenness and whether or not betweenness increases or not. The BCC scores for the topics are between 0.5 and 1, which would indicate that most of the values are identified as more polarized than non-polarized. Nonetheless, there are multiple values that are very close to 0.5, and for these topics, it is not so clear whether or not the topics are polarized since the scale of measure zero to one, closer one meaning polarized.

The topics falling in the middle of interval zero to one are the Secto Rally in Finland and electric prices. Looking at the earlier measures’ results, Secto Rally is not that big of a surprise since the results are quite well aligned with earlier measures' results and not finding polarization. On the other hand, the topic of electric prices has gotten a little bit more mixed results, and especially with the boundary controversy, the BCC score's interpretation is the opposite.

Moving on to the topics that could be identified as polarized based on the BCC score. The scores are overall quite high, and there are a total of 11 topics that have over 0.75 scores. From these topics, there are some which are quite well aligned with earlier measures, and good examples of these topics are protests in Iran, Fortum discussions in September, and Russian tourist visas. There are also those topics which, based on BCC score, can be defined as polarized and are more or less contradictory results of other measures indications. For example, topics about the Finnish national team and nurses' collective agreement had been classified as non-polarized or not clearly polarized topics based on boundary polarization, dipole polarization, and modularity. Noticing from these both topics' visuals, it is quite hard to notice the typical polarized structure, and both networks seem to be well-connected.

5.3.6. Random walk controversy & Adaptive Random walk controversy

Looking at the random walk controversy (RWC) scores overall, all the values are over zero, indicating that a random walk more likely stays on the same side and does not cross the boundary. From the results, it is also quite clear that RWC values are more than zero, meaning that the probability of crossing is much more unlikely than staying on its side. The theory is that if the chance to end up other side is small, then the graph structure is such that users are not introduced to different and new ideas.

The topics that had been scoring highly earlier also seem to get high scores of RWC. For example, the Iran protests, the New Zealand rally, and Fortum in September had high Modularity scores and Dipole polarization. So this way, it seems that the RWC scores follow other measures. On the other hand, RWC is giving contradictory results because some networks have been identified as non-polarized, but the RWC scores state differently. Examples of these topics are #huuhkajat, #meidänkaikkienasia, and huumeet. So it would appear that RWC scores classify all the network structures as polarized, although the RWC scores are lower if the other metrics had classified networks as non-polarized. And it does clearly raise the question of whether or not the measure can identify the non-polarized networks from the polarized network.

In the same way, as RWC and adaptive RWC (ARWC) scores are over zero, the difference now is that there is much more variation in results. The results are smaller than RWC scores, and the drops are significant, especially in contradictory cases. The situation changes the way that even though the results are positive, the results are much closer to zero than one. Closer to zero would indicate that the likelihood of crossing boundaries or staying in the same community is fifty-fifty. Mirroring ARWC results to the other polarization measures results seems to be much more aligned.

Surely the ARWC scores are not falling under zero, so the probability of leaving your own community is smaller than staying. The ARWC scores of networks that are less clearly labeled polarized are electricity prices in Finland, nurses' collective agreement, and new drug policy. The same topics have been scoring more unclear or non-polarized results. Also, the clearer polarized topics like Iran protests, Sanna Marin, Russian tourist visas, and Nordstream explosion, based on earlier measures, are still scoring high ARWC scores.

5.4. Summary of polarization measures

Finally, in Table 7, there are all the measures summarized by labeling results in either the Polarized, Debatable, or Unpolarized categories. Debatable in this case, meaning the results which cannot be very clearly labeled to either Polarized or Unpolarized class. From the summary table, it is quite evident that measures E-I Indeces and RWC measures are labeling topics much more easily as polarized. It would seem that E-I indices and RWC measures rely on partitions to be clear. Now as can be noticed from visual representations in Appendix 5, the partitions are not clearly cut into two clusters, making it hard to E-I indices and RWC measures to separate polarized and unpolarized topics. On the other hand, those measures which are less relying partition are giving more changes in results, and unpolarized topics also start to stand out.

From the topics, there are a couple that seems to have polarized results across the board, and these are Nordstream attacks, Iranian protests, and Russian tourist visas in

Finland and Fortum in July and September. From these topics, the Nordstream attack and protest in Iran are a little bit surprising, but based on the content of tweets, it would seem that the Iran protest was linked to domestic policy, and the Nordstream attacks generated conversations about some old decisions made by Finnish politicians. There is not any clearly unpolarized topics, but some topics, like the Finnish nurse collective agreement and Secto Rally Finland, seems to have more variation on the results, and seem to lean on more unpolarized.

To help better understand polarization measures the spearman rank correlation will be calculated in the way that all the topics are ranked from the highest polarization score to the lowest. These results are summarized in Appendix 7. After this, there is a possibility to calculate the spearman correlation, which can give us information about the way different measures quantify polarization. Also, this helps us understand if there are some similar measures that might “react” in similar ways to the characteristics of a network.

Table 7. Summary of polarization measure results

Topic (Collection Period)	Betweenness	E-I Index	Adaptive E-I Index	Boundary Connectivity	Dipole Moment	Modularity	RWC	ARWC
#huuhkajat, 2022-09-18-2022-09-27	Polarized	Polarized	Polarized	Debatable	Unpolarized	Unpolarized	Polarized	Polarized
#Jodi, 2022-10-05-2022-10-15	Polarized	Polarized	Polarized	Debatable	Polarized	Polarized	Polarized	Polarized
#meidänkaikkienasia, 2022-09-27-2022-10-04	Polarized	Polarized	Polarized	Debatable	Unpolarized	Unpolarized	Polarized	Polarized
#Nordstream, 2022-09-26-2022-10-01	Polarized	Polarized	Polarized	Polarized	Polarized	Polarized	Polarized	Polarized
#RallyNewZealand, 2022-09-26-2022-10-04	Debatable	Polarized	Polarized	Polarized	Polarized	Polarized	Polarized	Polarized
#sectorallyfinland, 2022-08-03-2022-08-06	Debatable	Polarized	Polarized	Polarized	Unpolarized	Debatable	Polarized	Polarized
#Sähkönhinta, 2022-08-20-2022-09-03	Debatable	Polarized	Polarized	Polarized	Debatable	Polarized	Polarized	Polarized
#Viisumit, 2022-09-18-2022-10-01	Polarized	Polarized	Polarized	Polarized	Polarized	Polarized	Polarized	Polarized
Fortum, 2022-07-10-2022-07-28	Polarized	Polarized	Polarized	Polarized	Polarized	Polarized	Polarized	Polarized
Fortum, 2022-07-31-2022-09-03	Debatable	Polarized	Polarized	Polarized	Polarized	Debatable	Polarized	Polarized
Fortum, 2022-09-17-2022-09-24	Polarized	Polarized	Polarized	Polarized	Polarized	Polarized	Polarized	Polarized
Huumeet, 2022-07-21-2022-07-31	Polarized	Polarized	Polarized	Polarized	Unpolarized	Debatable	Polarized	Polarized
Iran, 2022-09-18-2022-10-11	Polarized	Polarized	Polarized	Polarized	Polarized	Polarized	Polarized	Polarized
Marin, 2022-08-17-2022-09-03	Polarized	Polarized	Polarized	Polarized	Polarized	Debatable	Polarized	Polarized
Venäjä, 2022-09-19-2022-09-24	Polarized	Polarized	Polarized	Debatable	Debatable	Debatable	Polarized	Polarized
yle, 2022-08-27-2022-09-11	Polarized	Polarized	Polarized	Debatable	Polarized	Polarized	Polarized	Polarized

Looking at the results in Figure 10. Not surprisingly, Adaptive and E-I indices have a high correlation since both topics did have similar results in almost every topic. On the other hand, the correlation is negative toward all the other measures, which also, in many ways, is expected since the results are quite clearly opposite compared to other measures. The positive correlation results are forming between dipole moment and modularity as well as adaptive random walk controversy and dipole moment. In both these cases also the p-values seem to show that correlations are statistically significant. All the p-values can be found in Appendix 6. All in all, measures are not having much positive correlations and mostly the results are neutral.

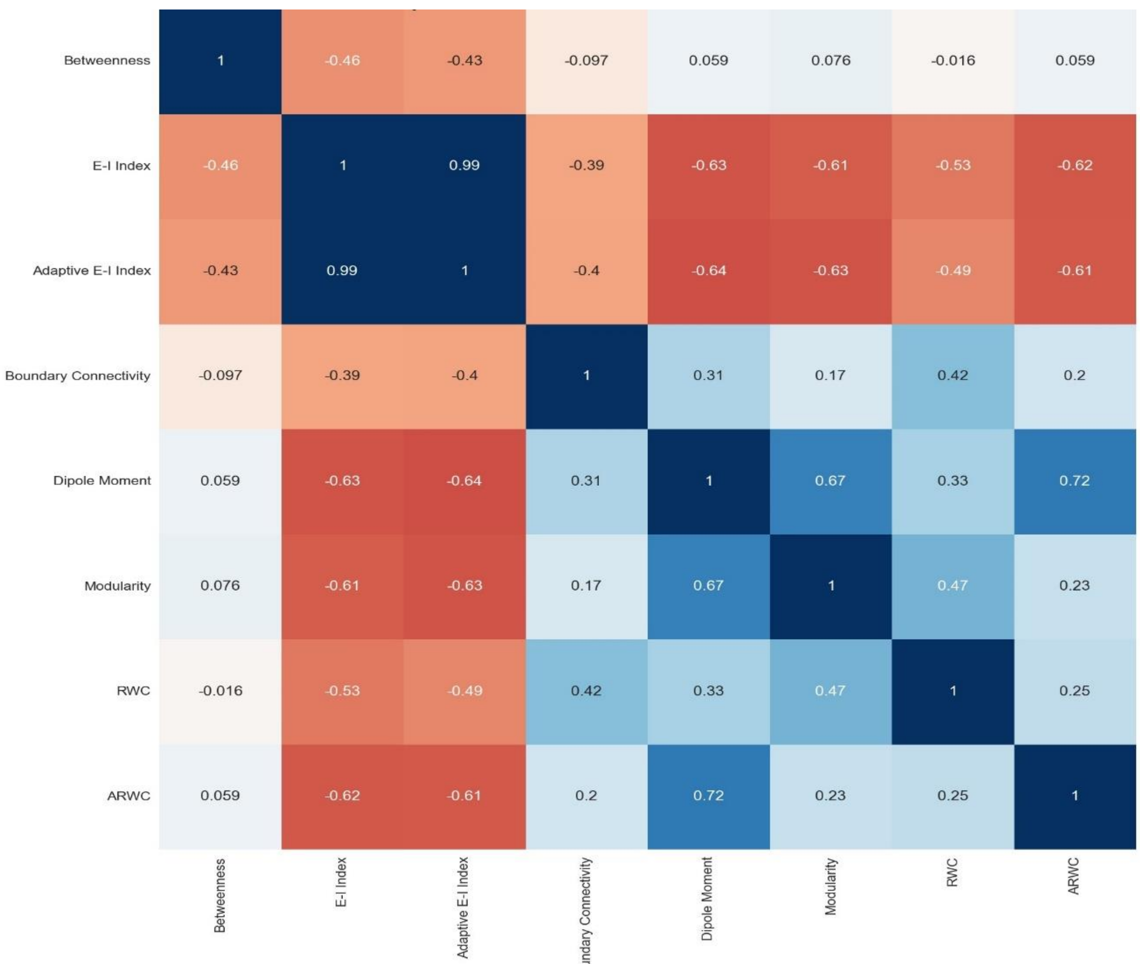


Figure 10. Spearman rank correlation coefficients

6. CONCLUSION

This thesis's final and sixth chapters will conclude the study by summarizing the findings connected to the aims and questions introduced in the first chapter. The chapter also consists of a review of limitations in the research and proposes the possible new directions where polarization in social media studies could be taken.

6.1 Summary

In recent years, multiple ways to study controversy on social media have been developed, and approaches to finding out disagreements have been quite different. The main two directions to identifying polarization could be noticed where one is more focused on the content and other structures of networks which, for example, social media creates. Focusing more on the content usually involves applying natural language processing or sentiment analysis (Mejova, et al 2014; DiMaggio, et al 1996, 690-755). The second approach, where polarization is studied through networks and how polarization could be explained based on network structure, has become popular recently (Garimella, et al 2018b, 1-27; Guerra, et al 2013, 215-224; Morales, et al 2015, 9-10). The development is happening mainly because the internet and the whole world have been developing in the direction where large networks can be analyzed (Baldassarri and Bearman 2017, 3-4).

The structural approach follows quite closely the same kind of path where from source data, a network is created where from nodes are it is tried to find different groups and ultimately measure the closeness to other groups (Conover, et al 2011, 89-96; Garimella, et al 2018b, 1-27; Morales, et al 2015, 9-10). The measures which can be used to identify and summarize polarization in a network can be something like modularity (Conover, et al 2011, 89-96), dipole polarization (Morales, et al 2015, 9-10), or random walk controversy (Garimella, et al 2018b, 1-27).

Dealing with real-world data from social media like Twitter or some other platform, adding some parts of content analysis does seem to give a better basis for examining polarization measures. Since the structural approach relies heavily on the structure of a network and real-world data being diverse, understanding the system helps, which could be made by using visual illustrations and network statistics. The theory behind each measure is a little different, and that's why using multiple criteria and examining combinations of these seems to give a better understanding of polarization with real-world data. Real-world data is messy, and there are no perfect representations of polarized or non-polarized situations, and that's why combining as much information does help to understand the dynamic of network and polarization. The polarization measures seem mostly aligned except for a couple of exceptions E-I index and RWC, which have difficulty separating more connected and divided networks. However, the RWC was suggested to be one of the more reliable measures (Garimella, et al 2018b, 1-27).

Based on the polarization measures, the themes that are farther from the politics, such as sports, are getting non-polarized results, and the networks are also visually packed. Sports-themed topics are national-level events, and the tweets are limited to the Finnish language, which may affect the fact that the results are different from studies in which interactions have been between rival teams (Garimella, et al 2018b, 1-27; Guerra, et al 2013, 215-224).

Mainly more political the topics get, the more divided the networks are, and the scores are also turning polarized from non-polarized, which seems to be aligned with the results of Conver et al. (2011) and Guerra et al. (2013). The topic that does have scores that are consistently non-polarized is the Finnish nurses' collective agreement, even though it can be considered political. Based on the content of the tweets, it seems that the reception to the decision was quite positive, and the whole network is quite connected. One common factor with topics scoring high polarization scores is where somebody/something can be blamed for failure. Based on the content of tweets, there are the blaming and defending sides of users in these kinds of situations. Maybe one of the most noticeable ones is the Sanna Marin case.

6.2. Limitations and future research

When researching social media, there are plenty of different sites and platforms, making it quite hard to cover all of them since every platform works differently. This research, it had been decided to use openly available Twitter data. Still, it is most certainly a limitation of this study because, in every social media, there are slightly different dynamics in the interactions. That is why it would be essential to get more studies focusing on various social media platforms, even though it might be hard since social media companies are setting barriers to data access. Another significant fact that should be pointed out is how data was chosen and the quality of data for this research. The data is manually picked from Twitter; since it had to be collected during specific time periods, it is easy to concentrate on particular topics. The manual process also makes topics easily skewed toward the researcher's interests. Also, the number of topics in this research is quite limited, and more research with different issues from different starting points is needed.

Maybe one of the most significant limitations of this study is how the networks have to be parted, which highlights the role of partition algorithms. Future research should focus on the pipeline partition parts and how to graph partitioning can be optimized by testing different types of mechanics. All the networks in this research are divided into two groups since the polarization measures can mostly work with two groups leaving out all the possible multisided controversies. From the visualizations, it is clear that it would be beneficial to study multisided debates in some cases, as Garimella et al. (2018) suggest in their work. More development of polarization scores would need to be directed to these multisided quantifying measures.

In the grand scheme of all, social media dominates our everyday lives, making it more and more critical and urgent to understand how it affects our lives so that users can be served healthier products. It means that more studies with real-world data and practical applications are needed.

REFERENCES

Adamic, Lada and Natalie Glance. (2005) The political blogosphere and the 2004 U.S. election. *Proceedings of the 3rd international workshop on link discovery* 36-43.

Aimei Yang and Charles Self. (2015) Anti-Muslim prejudice in the virtual space: A case study of blog network structure and message features of the 'Ground Zero mosque controversy'. *Media, war & conflict* 8, 1, 46-69. Available: DOI: 10.1177/1750635214541030.

Akram, Waseem and Rekesh Kumar. (2017) A study on positive and negative effects of social media on society. *International Journal of Computer Sciences and Engineering* 5, 10, 351-354.

Albert, Réka and Albert-László Barabási. (2001) *Statistical Mechanics of Complex Networks.*

Algan, Yann, et al. (2019) Friendship Networks and Political Opinions: A Natural Experiment among Future French Politicians Centre for Economic Policy Research. [https://spire.sciencespo.fr/hdl:/2441/5b1u7t12sb8ikq4csvts6tsa3h].

Angwin, Julia, Jeff Larson, Surya Mattu, and Lauren Kirchner. (2016) Machine bias. *ProPublica, May* 23, 2016, 139-159.

Bail, Christopher A., Lisa P. Argyle, Taylor W. Brown, John P. Bumpus, Haohan Chen, MB F. Hunzaker, Jaemin Lee, Marcus Mann, Friedolin Merhout, and Alexander Volfovsky. (2018) Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences* 115, 37, 9216-9221.

Baldassarri, Delia and Peter S. Bearman. (2017) Dynamics of Political Polarization Columbia University. [https://search.datacite.org/works/10.7916/d8fx7hbq].

Banavar Jaynath, R., Maritan Amos, and Rinaldo Andrea. (1999) Size and form in efficient transportation networks. *Nature* 399, 130-132.

Barbera, Pablo. (2015) *How social media reduces mass political polarization. Evidence from Germany, Spain, and the U.S.* Available: http://pablobarbera.com/research.html.

Barberá, Pablo, J. T. Jost, J. Nagler, R. Bonneau, and J. A. Tucker. (2015) *Tweeting From Left to Right Is Online Political Communication More Than an Echo Chamber.*

Barocas, Solon and Andrew D. Selbst. (2016) Big Data s Disparate Impact. 104, 3, 671-732. Available: DOI: 10.15779/z38bg31.

Bartlett, Robert P., et al. , (2019) *Consumer-Lending Discrimination in the FinTech Era.* Cambridge, MA: National Bureau of Economic Research.

Baumann, Fabian, Philipp Lorenz-Spreen, Igor M. Sokolov, and Michele Starnini. (2020) *Modeling Echo Chambers and Polarization Dynamics in Social Networks.* : American Physical Society (APS).

Berlow, E. L. (1999) Strong effects of weak interactions in ecological communities Arctic Alpine Resilience (ArcAlpNet) View project Structure and dynamics of Mutualistic Networks View project Eric L Berlow Vibrant Data. *Nature* 398, 330-334.

Bessi, Alessandro, Fabiana Zollo, Michela Del Vicario, Michelangelo Puliga, Antonio Scala, Guido Caldarelli, Brian Uzzi, and Walter Quattrociocchi. (2016) *Users Polarization on Facebook and Youtube.* : Public Library of Science (PLoS).

Boccalettia, S., V. Latora, Y. Moreno, M. Chavez, and D. -. Hwang. (2006) Complex networks Structure and dynamics. *Physics Reports* 424, 4-5, 175-308.

Boulianne, Shelley. (2015) Social media use and participation: a meta-analysis of current research. *Information, communication & society* 18, 5, 524-538. Available: DOI: 10.1080/1369118X.2015.1008542.

Boutin, Paul. (2011) *Your Results May Vary.* Available: https://online.wsj.com/article/SB10001424052748703421204576327414266287254.html?reflink=desktopwebshare_permalink.

Boxell, Levi, Matthew Gentzkow, and Jesse M. Shapiro. (2017) Greater Internet use is not associated with faster growth in political polarization among US demographic groups. *Proceedings of the National Academy of Sciences - PNAS* 114, 40, 10612-10617. Available: DOI: 10.1073/pnas.1706588114.

Bramson, Aaron, Patrick Grim, Daniel J. Singer, Steven Fisher, William Berger, Graham Sack, and Carissa Flocken. (2016) Disambiguation of social polarization concepts and measures. *The Journal of mathematical sociology* 40, 2, 80-111. Available: DOI: 10.1080/0022250X.2016.1147443.

Branscombe, Nyla R. and Daniel L. Wann. (1994) Collective self-esteem consequences of outgroup derogation when a valued social identity is on trial. *European journal of social psychology* 24, 6, 641-657. Available: DOI: 10.1002/ejsp.2420240603.

Branscombe, Nyla R., Daniel L. Wann, Jeffrey G. Noel, and Jason Coleman. (1993) In-Group or Out-Group Extemity: Importance of the Threatened Social Identity. *Personality & social psychology bulletin* 19, 4, 381-388. Available: DOI: 10.1177/0146167293194003.

Brian Hayes. (2000) Computing Science: Graph Theory in Practice: Part I. *American scientist* 88, 1, 9-13.

Chen, Ted H. Y., Ali Salloum, Antti Gronow, Tuomas Ylä-Anttila, and Mikko Kivelä. (2021) Polarization of climate politics results from partisan sorting: Evidence from Finnish Twittersphere. *Global Environ.Change* 71, 1-27.

Cinelli, Matteo, Gianmarco De, Francisci Morales, Alessandro Galeazzi, Walter Quattrociocchi, and Michele Starnini. (2021) *The echo chamber effect on social media.*

Colleoni, Elanor, Alessandro Rozza, and Adam Arvidsson. (2014) Echo Chamber or Public Sphere? Predicting Political Orientation and Measuring Political Homophily in Twitter Using Big Data. *Journal of communication* 64, 2, 317-332. Available: DOI: 10.1111/jcom.12084.

Conover, Michael, Jacob Ratkiewicz, Matthew Francisco, Bruno Goncalves, Alessandro Flammini, and Filippo Menczer. (2011) Political Polarization on Twitter. 5, 1, 89-96. Available: DOI: 10.5281/zenodo.4589065.

Cowgill, Bo and Catherine E. Tucker. (2019) Economics, Fairness and Algorithmic Bias. *SSRN Electronic Journal* Available: DOI: 10.2139/ssrn.3361280.

DellaVigna, Stefano and Ethan Kaplan. (2007) The Fox News Effect: Media Bias and Voting. *The Quarterly Journal of Economics* 122, 3, 1187-1234. Available: DOI: 10.1162/qjec.122.3.1187.

DiMaggio, P., J. Evans, and B. Bryson. (1996) Have American's Social Attitudes Become More Polarized? *The American journal of sociology* 102, 3, 690-755. Available: DOI: 10.1086/230995.

Dorogovtsev, S. N. and J. F. F. Mendes. (2001) *Evolution of networks.*

Dubois, Elizabeth and Grant Blank. (2018) The echo chamber is overstated: the moderating effect of political interest and diverse media. *Information, communication & society* 21, 5, 729-745. Available: DOI: 10.1080/1369118X.2018.1428656.

Eberl, Jakob-Moritz, Hajo G. Boomgaarden, and Markus Wagner. (2017) One Bias Fits All? Three Types of Media Bias and Their Effects on Party Preferences. *Communication research* 44, 8, 1125-1148. Available: DOI: 10.1177/0093650215614364.

Emamgholizadeh, Hanif, Milad Nourizade, Mir S. Tajbakhsh, Mahdieh Hashminezhad, and Farzaneh N. Esfahani. (2020) A framework for quantifying controversy of social network debates using attributed networks: biased random walk (BRW). *Social Network Analysis and Mining* 10, 1, 1-20.

Enyedi, Zsolt. (2016) Populist Polarization and Party System Institutionalization. *Problems of post-communism* 63, 4, 210-220. Available: DOI: 10.1080/10758216.2015.1113883.

Esteve Del Valle, Marc, Marcel Broersma, and Arnout Ponsioen. (2022) Political interaction beyond party lines: Communication ties and party polarization in parliamentary twitter networks. *Soc.Sci.Comput.Rev.* 40, 3, 736-755.

Festinger, Leon. , (1962) *A theory of cognitive dissonance.*: Stanford university press.

Flaxman, Seth, Sharad Goel, and Justin M. Rao. (2016a) Filter bubbles, echo chambers, and online news consumption. *Public Opin.Q.* 80, S1, 298-320.

Flaxman, Seth, Sharad Goel, and Justin M. Rao. (2016b) Filter bubbles, echo chambers, and online news consumption. *Public opinion quarterly* 80, S1, 298-320. Available: DOI: 10.1093/poq/nfw006.

Freedman, Jonathan L. (1965) Preference for dissonant information. *Journal of personality and social psychology* 2, 2, 287-289. Available: DOI: 10.1037/h0022415.

Frey, Dieter. "Recent Research on Selective Exposure to Information." In *Advances in Experimental Social Psychology*, edited by Anonymous . New York, N.Y: Elsevier Science & Technology, 1986.

Garimella Kiran. "Polarization on Social Media.", Aalto University, 2018.

Garimella, Kiran, Gianmarco D. F. Morales, Aristides Gionis, and Michael Mathioudakis. (2018a) Political Discourse on Social Media: Echo Chambers, Gatekeepers, and the Price of Bipartisanship.

Garimella, Kiran, Gianmarco D. F. Morales, Aristides Gionis, and Michael Mathioudakis. (2018b) Quantifying controversy on social media. *ACM Transactions on Social Computing* 1, 1, 1-27.

Garrett, R. K. (2009) Echo chambers online?: Politically motivated selective exposure among Internet news users1. *Journal of computer-mediated communication* 14, 2, 265-285. Available: DOI: 10.1111/j.1083-6101.2009.01440.x.

Garz, Marcel. (2014) Good news and bad news: evidence of media bias in unemployment reports. *Public Choice* 161, 3/4, 499-515. Available: DOI: 10.1007/s11127-014-0182-2.

Gentzkow, Matthew and Jesse M Shapiro. (2006) Media Bias and Reputation. *The Journal of political economy* 114, 2, 280-316. Available: DOI: 10.1086/499414.

Gilbert, Eric and Karrie Karahalios. (2009) Predicting tie strength with social media. *Proceedings of the SIGCHI Conference on human factors in computing systems* 211-220.

Greitemeyer, Tobias. (2014) I Am Right, You Are Wrong: How Biased Assimilation Increases the Perceived Gap between Believers and Skeptics of Violent Video Game Effects. *PloS one* 9, 4, e93440. Available: DOI: 10.1371/journal.pone.0093440.

Groseclose, Tim and Jeffrey Milyo. (2005) A Measure of Media Bias. *The Quarterly Journal of Economics* 120, 4, 1191-1237. Available: DOI: 10.1162/003355305775097542.

Guerra, Pedro, Wagner Meira Jr, Claire Cardie, and Robert Kleinberg. (2013) A measure of polarization on social media networks based on community boundaries. 7, 1, 215-224.

Harmon-Jones, Eddie and Cindy Harmon-Jones. (2007) Cognitive Dissonance Theory After 50 Years of Development. *Zeitschrfit für Sozialpsychologie* 38, 1, 7-16. Available: DOI: 10.1024/0044-3514.38.1.7.

Hayes, Brian. (2000) Computing science: Graph theory in practice: Part I. *Am.Sci.* 88, 1, 9-13.

Heatherly, Kyle A., Yanqin Lu, and Jae K. Lee. (2017) Filtering out the other side? Cross-cutting and like-minded discussions on social networking sites. *New media & society* 19, 8, 1271-1289. Available: DOI: 10.1177/1461444816634677.

Holone, Harald. (2016) The filter bubble and its effect on online personal health information. *Croatian medical journal* 57, 3, 298-301. Available: DOI: 10.3325/cmj.2016.57.298.

Hosanagar, Kartik, Daniel Fleder, Dokyun Lee, and Andreas Buja. (2013) Will the Global Village Fracture into Tribes? Recommender Systems and their Effects on Consumer Fragmentation "Will the global village fracture into tribes?" -P. Resnick. *Management Science* 60, 2, 805-823.

Hout, Michael and Christopher Maggio. (2021) Immigration, Race & Political Polarization. *Daedalus (Cambridge, Mass.)* 150, 2, 40-55. Available: DOI: 10.1162/daed_a_01845.

Hwang, Hyunseo, Michael Schmierbach, Hye-Jin Paek, Homero Gil de Zuniga, and Dhavan Shah. (2006) Media Dissociation, Internet Use, and Antiwar Political Participation: A Case Study of Political Dissent and Action Against the War in Iraq. *Mass communication & society* 9, 4, 461-483. Available: DOI: 10.1207/s15327825mcs0904_5.

Iyengar, Shanto and R. Reeves. , (1997) *Do the media govern? Politicians, voters and reporters in America.* Thousand Oaks, CA: Sage.

Jackson Jasper. (2017) *Eli Pariser: activist whose filter bubble warnings presaged Trump and Brexit.* Available:
https://www.theguardian.com/media/2017/jan/08/eli-pariser-activist-whose-filter-bubble-warnings-presaged-trump-and-brexit.

Jeong, Myeongki, Hangjung Zo, Chul H. Lee, and Yasin Ceran. (2019) Feeling displeasure from online social media postings: A study using cognitive dissonance theory. *Computers in human behavior* 97, 231-240. Available: DOI: 10.1016/j.chb.2019.02.021.

Katsambekis, G. and Y. Stavrakakis. (2013) "Populism, anti-populism and European
democracy: A view from the south.", [accessed
https://www.opendemocracy.net/en/can-europe-make-it/populism-anti-populism-and-european-democr/.

Kemp, S. (2022) "DIGITAL 2022: APRIL GLOBAL STATSHOT REPORT .", [accessed 19.10. 2022]. https://datareportal.com/reports/digital-2022-april-global-statshot.

Klapper, Joseph T. , (1960) *The effects of mass communication.* New York, NY, US: Free Press.

Krackhardt, D. and R. N. Stern. (1988) Informal Networks and Organizational Crises: An Experimental Simulation. *Social psychology quarterly* 51, 2, 123-140. Available: DOI: 10.2307/2786835.

Lee, Daniel D. and H. S. Seung. (1999) Learning the parts of objects by non-negative matrix factorization. *Nature* 401, 6755, 788-791.

Lord, Charles G., Lee Ross, and Mark R. Lepper. (1979) Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of personality and social psychology* 37, 11, 2098-2109. Available: DOI: 10.1037/0022-3514.37.11.2098.

Lozada, Mireya. *Us or Them? Social Representations and Imaginaries of the Other in Venezuela.*

M’sik, Ben and Ben M. Casablanca. (2020) Topic modeling coherence: A comparative study between lda and nmf models using covid’19 corpus. *International Journal* 9, 4,.

McCoy, Jennifer, Tahmina Rahman, and Murat Somer. (2018) Polarization and the Global Crisis of Democracy: Common Patterns, Dynamics, and Pernicious Consequences for Democratic Polities. *The American behavioral scientist (Beverly Hills)* 62, 1, 16-42. Available: DOI: 10.1177/0002764218759576.

Mejova, Yelena, Amy X. Zhang, Nicholas Diakopoulos, and Carlos Castillo. (2014) Controversy and Sentiment in Online News.

Metaxas, Panagiotis, Eni Mustafaraj, Kily Wong, Laura Zeng, Megan O'Keefe, and Samantha Finn. (2015) What do retweets indicate? Results from user survey and meta-review of research. 9, 1, 658-661.

Michela Del Vicario, Alessandro Bessi, Fabiana Zollo, Fabio Petroni, Antonio Scala, Guido Caldarelli, H. Eugene Stanley, and Walter Quattrociocchi. (2016) The spreading of misinformation online. *Proceedings of the National Academy of Sciences - PNAS* 113, 3, 554-559. Available: DOI: 10.1073/pnas.1517441113.

Morales, Alfredo J., Javier Borondo, Juan C. Losada, and Rosa M. Benito. (2015) Measuring political polarization: Twitter shows the two sides of Venezuela. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 25, 3, 9-10.

Mukaka, Mavuto M. (2012) A guide to appropriate use of correlation coefficient in medical research. *Malawi medical journal* 24, 3, 69-71.

Myers, David G. (1975) Discussion-Induced Attitude Polarization. *Human relations (New York)* 28, 8, 699-714. Available: DOI: 10.1177/001872677502800802.

Myers, David G. and George D. Bishop. (1971) Enhancement of dominant attitudes in group discussion. *Journal of personality and social psychology* 20, 3, 386-391. Available: DOI: 10.1037/h0031920.

Natalie Jomini Stroud. (2008) Media Use and Political Predispositions: Revisiting the Concept of Selective Exposure. *Polit Behav* 30, 3, 341-366. Available: DOI: 10.1007/s11109-007-9050-9.

Newman, M. (2000) Models of the Small World. *Journal of Statistical Physics* 101, 3, 819-841. Available: DOI: 10.1023/A:1026485807148.

Newman, M. E. J. and M. Girvan. (2004) Finding and evaluating community structure in networks. *Physical review. E, Statistical, nonlinear, and soft matter physics* 69, 2 Pt 2, 69-84. Available: DOI: 10.1103/PhysRevE.69.026113.

Newman, M. E. J. (2003) The Structure and Function of Complex Networks. *SIAM review* 45, 2, 167-256. Available: DOI: 10.1137/S003614450342480.

Newman, Mark E. (2006) Modularity and community structure in networks. *Proceedings of the national academy of sciences* 103, 23, 8577-8582.

Nickerson, Raymond S. (1998) Confirmation Bias. *Review of general psychology* 2, 2, 175-220. Available: DOI: 10.1037/1089-2680.2.2.175.

Papacharissi, Zizi. (2002) The virtual sphere. *New media & society* 4, 1, 9-27. Available: DOI: 10.1177/14614440222226244.

Pariser, Eli. , (2011) *The filter bubble.* New York: Penguin Press.

Pennacchiotti, Marco and Ana-Maria Popescu. (2010) *Detecting controversies in Twitter: a first study.* : Association for Computational Linguistics.

Perrin, Andrew. (2015) Social media usage. *Pew research center* 125, 52-68.

Prior, Markus. (2013) Media and political polarization. *Annual Review of Political Science* 16, 101-127.

Prior, Markus. , (2007) *Post-Broadcast Democracy.* New York: Cambridge University Press.

Roccas, Sonia and Marilynn B. Brewer. (2002) Social Identity Complexity. *Personality and Social Psychology Review* 6, 2, 88-106.

Salloum, Ali, Ted H. Y. Chen, and Mikko Kivelä. (2022) Separating Polarization from Noise: Comparison and Normalization of Structural Polarization Measures. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW1, 1-33.

Schober, Patrick, Christa Boer, and Lothar A. Schwarte. (2018) Correlation coefficients: appropriate use and interpretation. *Anesthesia & analgesia* 126, 5, 1763-1768.

Sears, David O. and Jonathan L. Freedman. (1967) Selective exposure to information: A critical review. *The public opinion quarterly* 31, 2, 194-213.

Shi, Tian, Kyeongpil Kang, Jaegul Choo, and Chandan K. Reddy. (2018) Short-text topic modeling via non-negative matrix factorization enriched with local word-context correlations. 1105-1114.

Stroud, Natalie J. , (2011) *Niche News : The Politics of News Choice.*: Oxford University Press.

Stroud, Natalie J. (2010) Polarization and Partisan Selective Exposure. *Journal of communication* 60, 3, 556-576. Available: DOI: 10.1111/j.1460-2466.2010.01497.x.

Stroud, Natalie J. "Selective Exposure Theories." In *The Oxford Handbook of Political Communication*, 1 ed., edited by Anonymous : Oxford University Press, 2017.

Sunstein, Cass R. (2002) The Law of Group Polarization. *The journal of political philosophy* 10, 2, 175-195. Available: DOI: 10.1111/1467-9760.00148.

Taddicken, M. and L. Wolff. (2020) 'Fake News' in Science Communication_ Emotions and Strategies of Coping with Dissonance Online. *Media and Communication* 9, 1, 206-217.

Tajfel, Henri. "Experiments in Intergroup Discrimination." *Scientific American* 223 (Nov 1, 1970): 96-103, MEDLINE.

Tajfel, Henri, M. G. Billig, R. P. Bundy, and Claude Flament. (1971) Social categorization and intergroup behaviour. *European journal of social psychology* 1, 2, 149-178. Available: DOI: 10.1002/ejsp.2420010202.

Twitter. (2022) "Twitter.", [accessed https://developer.twitter.com/en/docs/twitter-api/getting-started/about-twitter-api#v2-access-level.

Vaccari, Cristian, Augusto Valeriani, Pablo Barberá, John T. Jost, Jonathan Nagler, and Joshua A. Tucker. (2016) Of Echo Chambers and Contrarian Clubs: Exposure to Political Disagreement Among German and Italian Users of Twitter. *Social media + society* 2, 3, 205630511666422. Available: DOI: 10.1177/2056305116664221.

Van Alstyne, Marshall and Erik Brynjolfsson. (2005) Global Village or Cyber-Balkans? Modeling and Measuring the Integration of Electronic Communities. *Management science* 51, 6, 851-868. Available: DOI: 10.1287/mnsc.1050.0363.

Waugh, Andrew S., Liuyi Pei, James H. Fowler, Peter J. Mucha, and Mason A. Porter. (2009) Party polarization in congress: A network science approach. *arXiv preprint arXiv:0907.3509* .

West, D. B. , (2001) *Introduction to graph theory.* 2 ed.: Pearson education Inc.

Williams, Hywel T. P., James R. McMurray, Tim Kurz, and F. Hugo Lambert. (2015) Network analysis reveals open forums and echo chambers in social media discussions of climate change. *Global environmental change* 32, 126-138. Available: DOI: 10.1016/j.gloenvcha.2015.03.006.

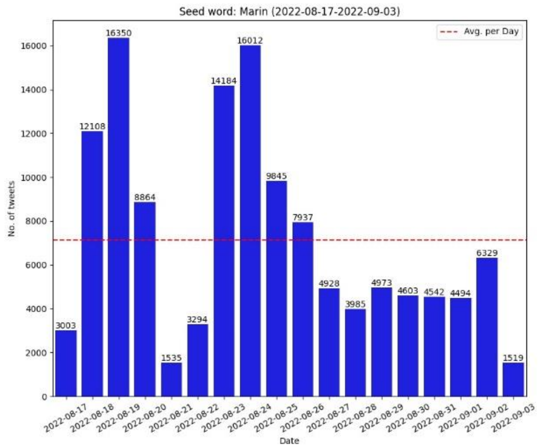
Yardi, Sarita and Danah Boyd. (2010) Dynamic Debates: An Analysis of Group Polarization Over Time on Twitter. *Bulletin of science, technology & society* 30, 5, 316-327. Available: DOI: 10.1177/0270467610380011.

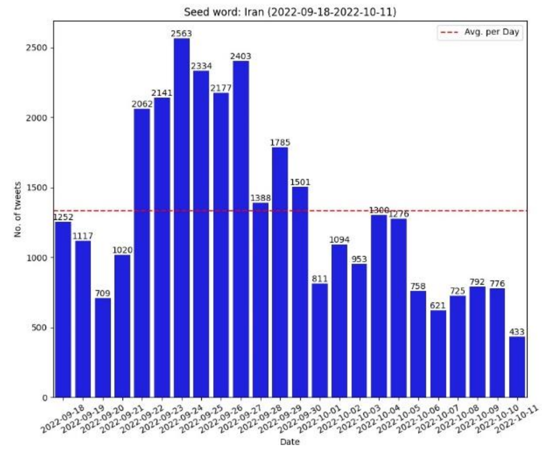
Young Mie Kim. (2009) Issue Publics in the New Information Environment. *Communication research* 36, 2, 254-284. Available: DOI: 10.1177/0093650208330253.

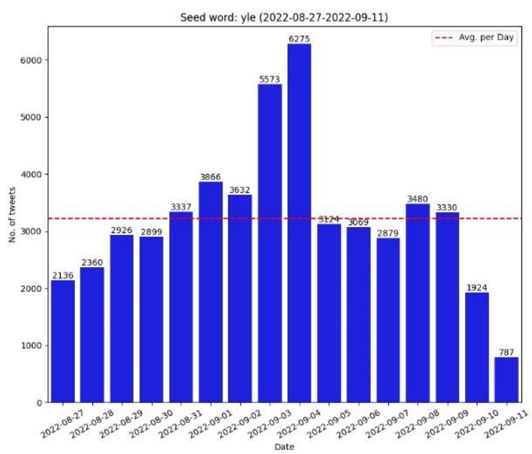
Zuiderveen Borgesius, F. J., D. Trilling, J. Möller, B. Bodó, C. H. de Vreese, and N. Helberger. (2016) Should We Worry about Filter Bubbles? *Internet policy review* 5, 1,Available: DOI: 10.14763/2016.1.401.

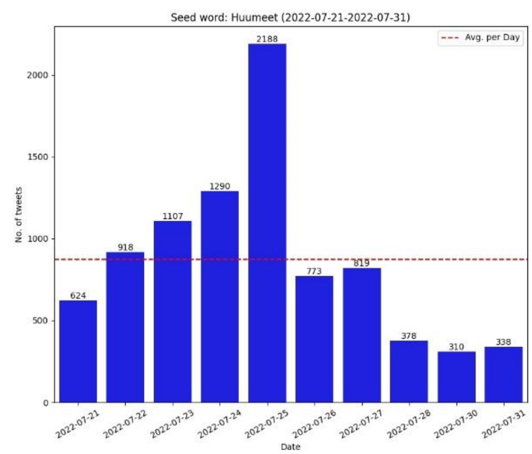
APPENDICIES

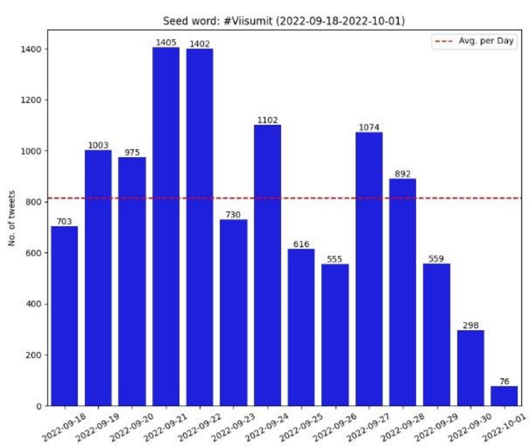
Appendix 1. Number of Tweets per Day

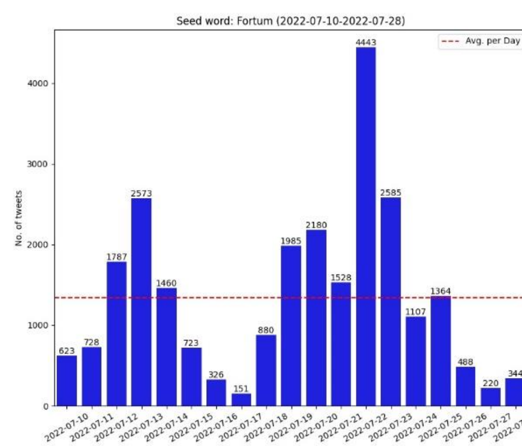


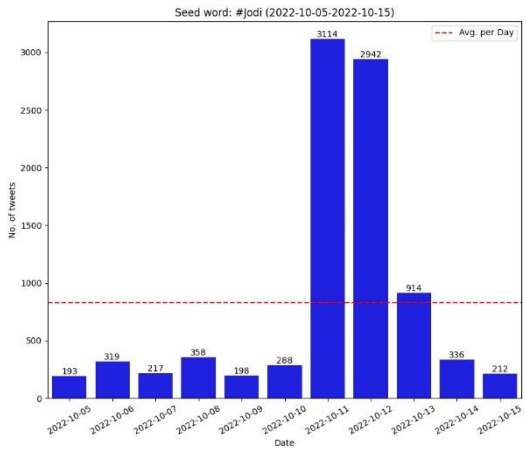


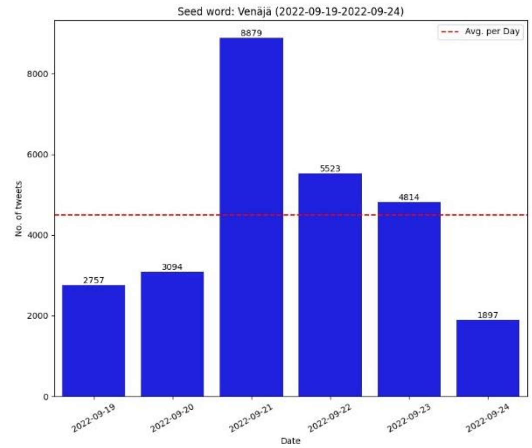


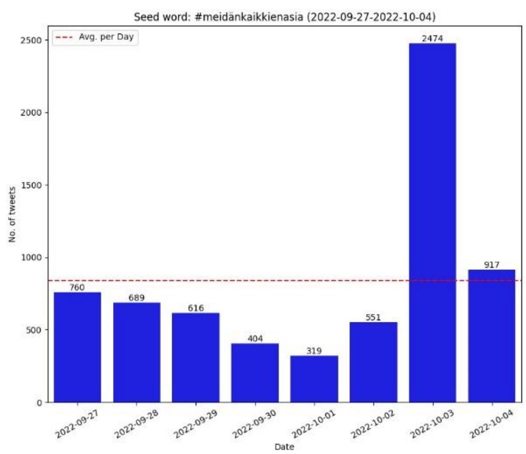


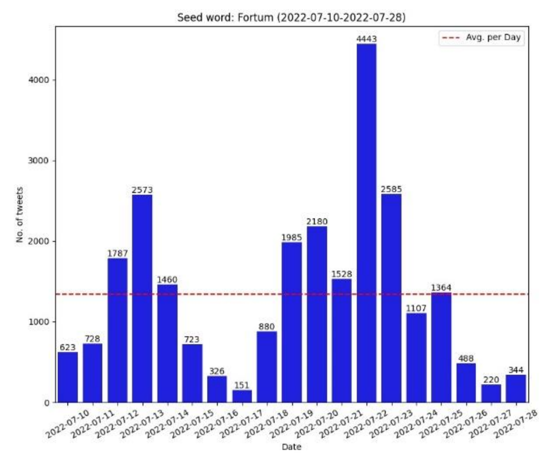


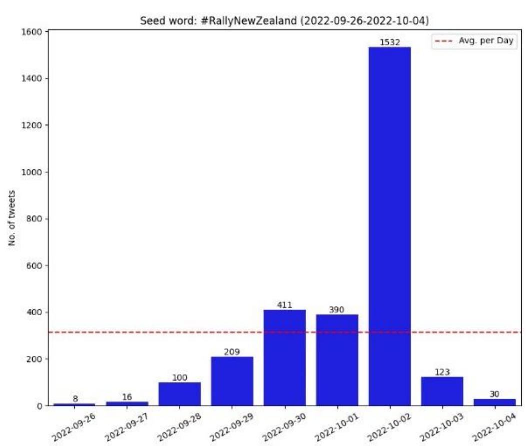


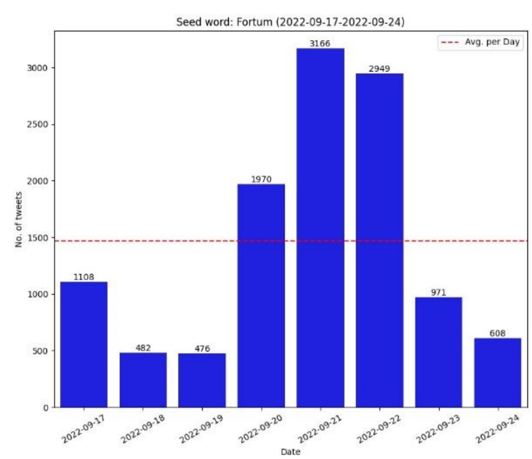

Seed word: #Jodi (2022-10-05-2022-10-15)
Avg. per Day
No. of tweets
Date
193
319
217
358
198
288
3114
2942
914
336
212

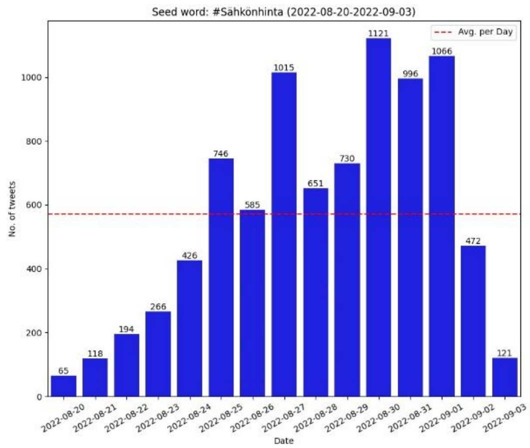

Seed word: Venäjä (2022-09-19-2022-09-24)
Avg. per Day
No. of tweets
Date
2757
3094
8879
5523
4814
1897

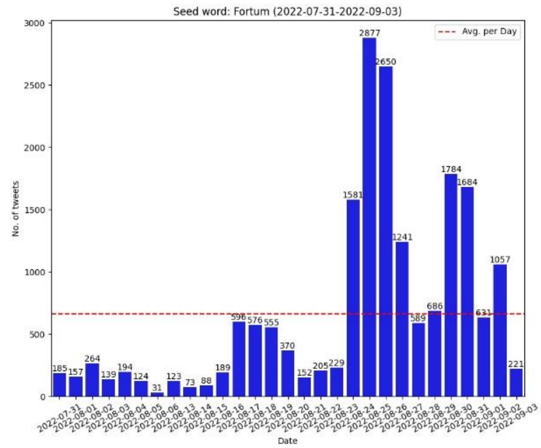

Seed word: #meidänkaikkienasia (2022-09-27-2022-10-04)
Avg. per Day
No. of tweets
Date
760
689
616
404
319
551
2474
917

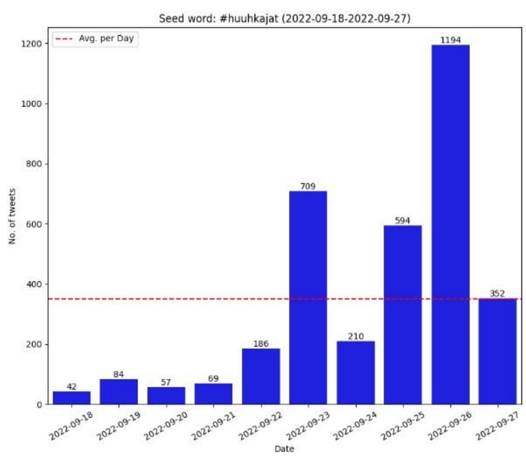

Seed word: Fortum (2022-07-10-2022-07-28)
Avg. per Day
No. of tweets
Date
623
728
1787
2573
1460
723
326
151
880
1985
2180
1528
4443
2585
1107
1364
488
220
344

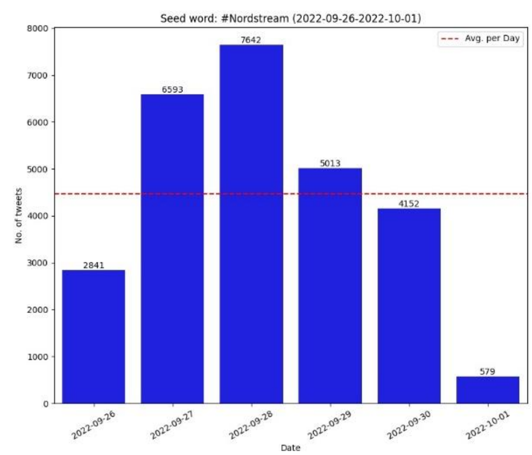

Seed word: #RallyNewZealand (2022-09-26-2022-10-04)
Avg. per Day
No. of tweets
Date
8
16
100
209
411
390
1532
123
30

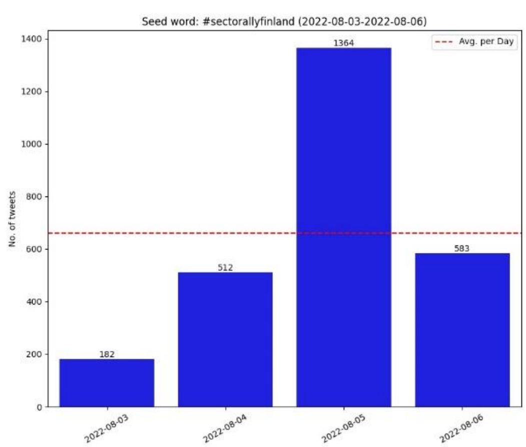

Seed word: Fortum (2022-09-17-2022-09-24)
Avg. per Day
No. of tweets
Date
1108
482
476
1970
3166
2949
971
608


Seed word: #Sähkönhinta (2022-08-20-2022-09-03)
Avg. per Day
No. of tweets
Date
65
118
194
266
426
746
585
1015
651
730
1121
996
1066
472
121

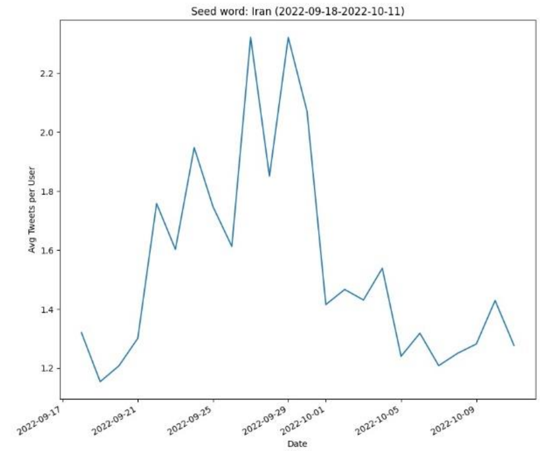

Seed word: Fortum (2022-07-31-2022-09-03)
Avg. per Day
No. of tweets
Date
2877
2650
1784
1684
1581
1241
1057

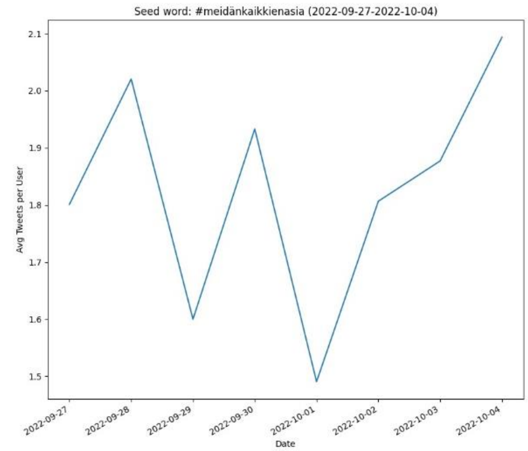

Seed word: #huuhkajat (2022-09-18-2022-09-27)
Avg. per Day
No. of tweets
Date
42
84
57
69
186
709
210
594
1194
352

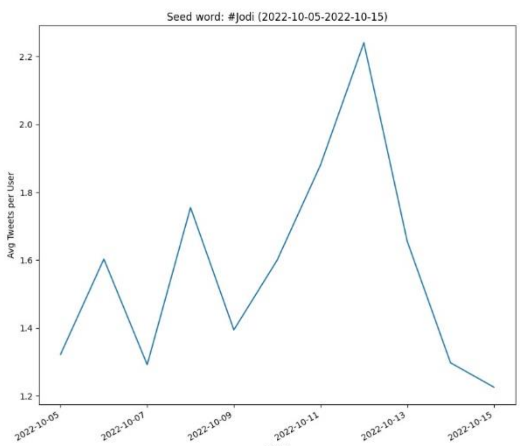

Seed word: #Nordstream (2022-09-26-2022-10-01)
Avg. per Day
No. of tweets
Date
2841
6593
7642
5013
4152
579

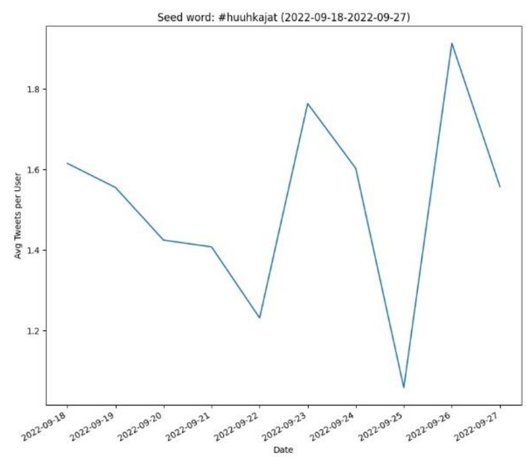

Seed word: #sectorallyfinland (2022-08-03-2022-08-06)
Avg. per Day
No. of tweets
Date
182
512
1364
583

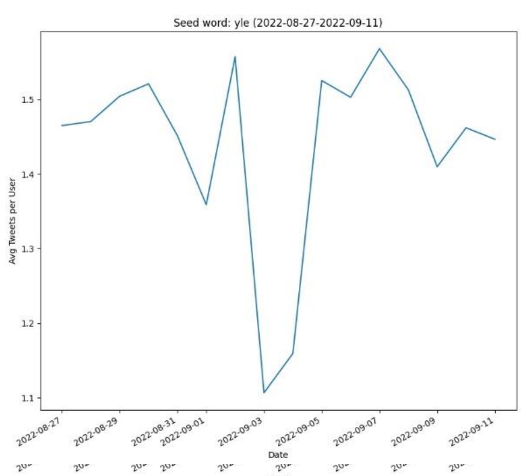
Appendix 2. Average Number of Users Tweeting per day

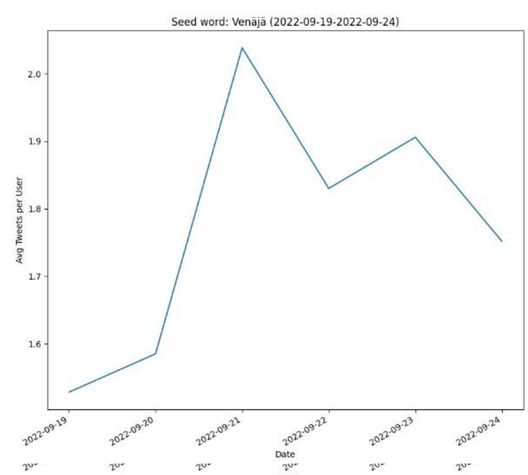


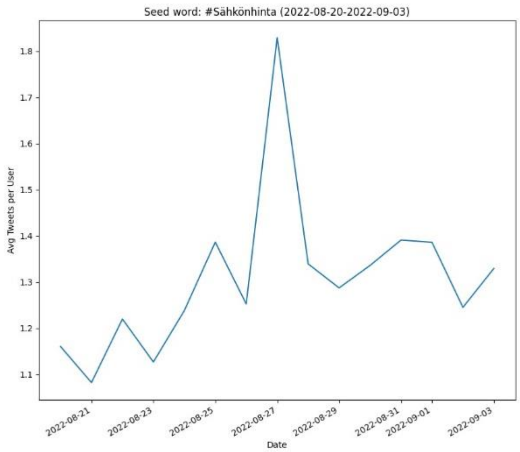


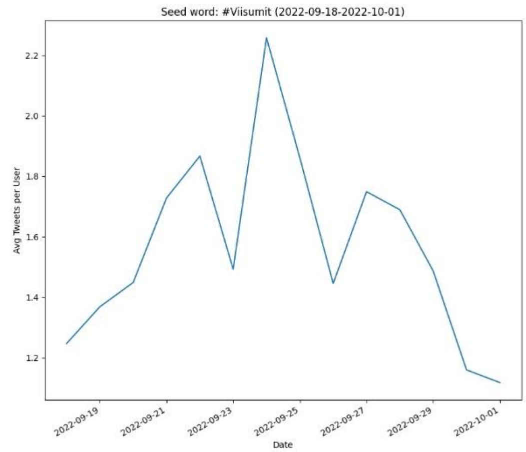


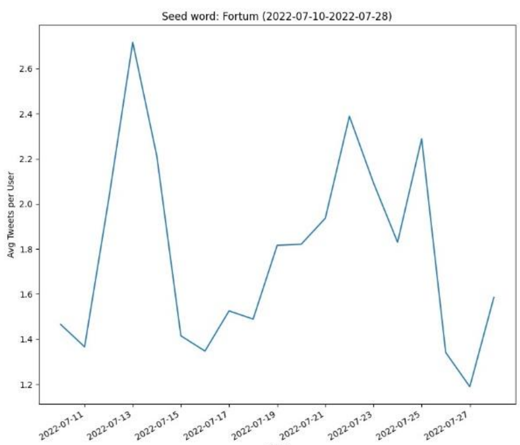


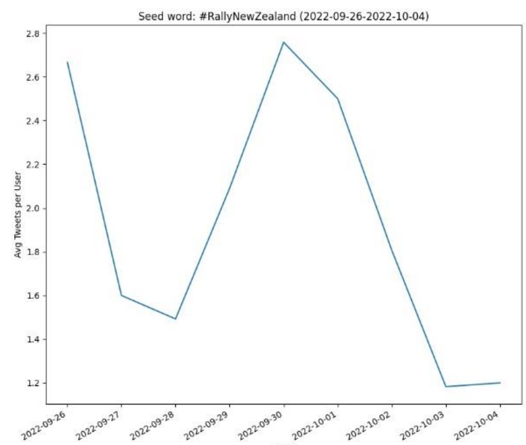


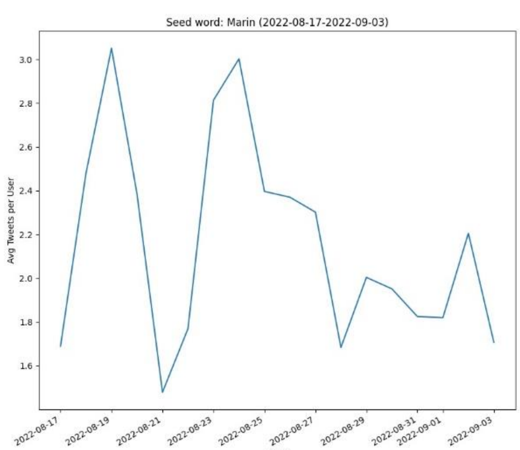


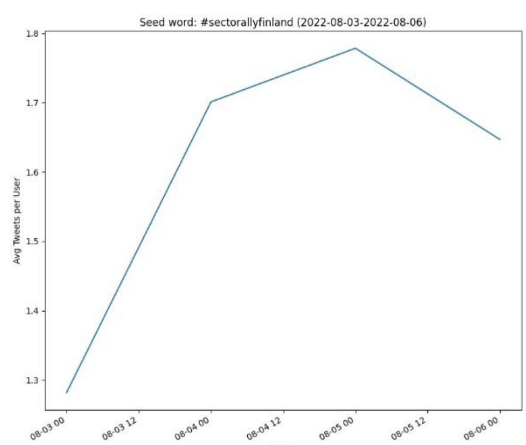

Seed word: #Sähkönhinta (2022-08-20-2022-09-03)
Avg Tweets per User
1.1
1.2
1.3
1.4
1.5
1.6
1.7
1.8
2022-08-21
2022-08-23
2022-08-25
2022-08-27
2022-08-29
2022-08-31
2022-09-01
2022-09-03
Date

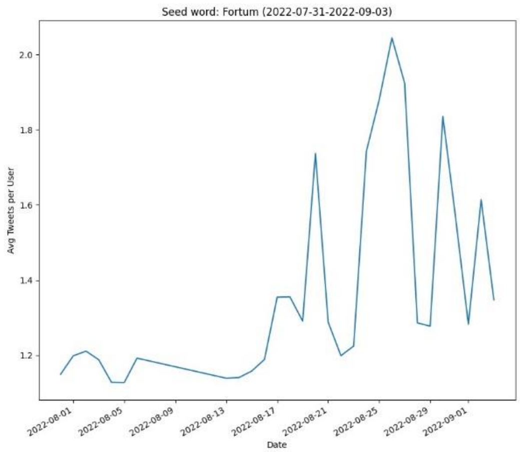

Seed word: #Viisumit (2022-09-18-2022-10-01)
Avg Tweets per User
1.2
1.4
1.6
1.8
2.0
2.2
2022-09-19
2022-09-21
2022-09-23
2022-09-25
2022-09-27
2022-09-29
2022-10-01
Date

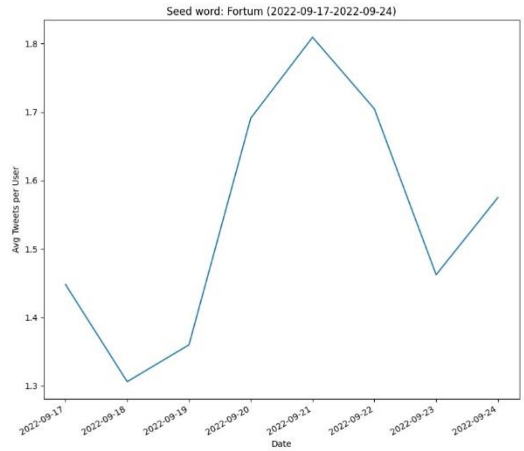

Seed word: Fortum (2022-07-10-2022-07-28)
Avg Tweets per User
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2022-07-11
2022-07-13
2022-07-15
2022-07-17
2022-07-19
2022-07-21
2022-07-23
2022-07-25
2022-07-27
Date

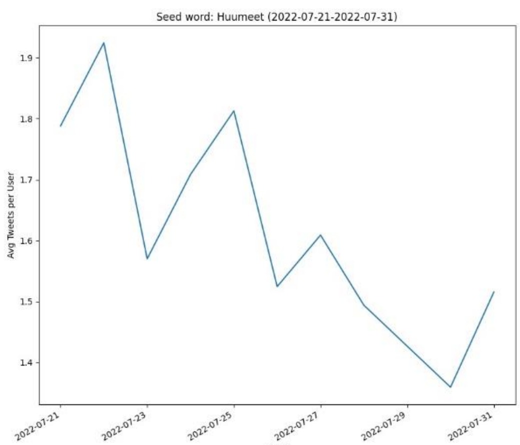

Seed word: #RallyNewZealand (2022-09-26-2022-10-04)
Avg Tweets per User
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
2022-09-26
2022-09-27
2022-09-28
2022-09-29
2022-09-30
2022-10-01
2022-10-02
2022-10-03
2022-10-04
Date

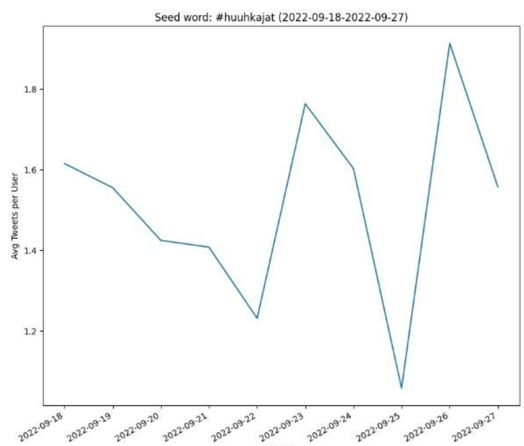

Seed word: Marin (2022-08-17-2022-09-03)
Avg Tweets per User
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
2022-08-17
2022-08-19
2022-08-21
2022-08-23
2022-08-25
2022-08-27
2022-08-29
2022-08-31
2022-09-01
2022-09-03
Date

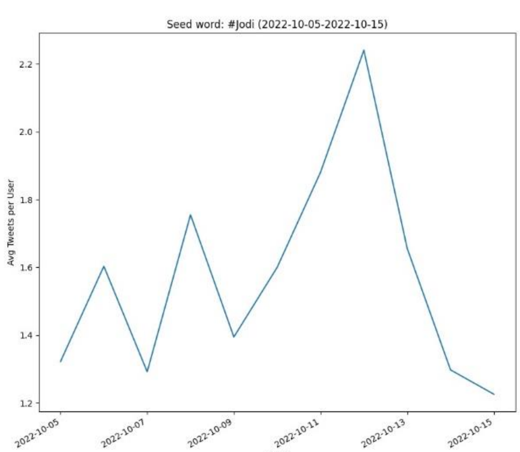

Seed word: #sectorallyfinland (2022-08-03-2022-08-06)
Avg Tweets per User
1.3
1.4
1.5
1.6
1.7
1.8
08-03 00
08-03 12
08-04 00
08-04 12
08-05 00
08-05 12
08-06 00
Date

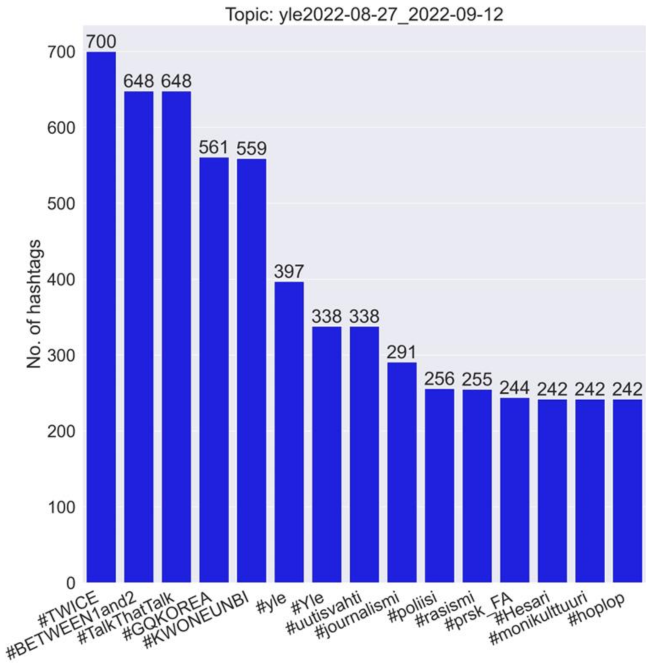

Seed word: Fortum (2022-07-31-2022-09-03)
Avg Tweets per User
Date

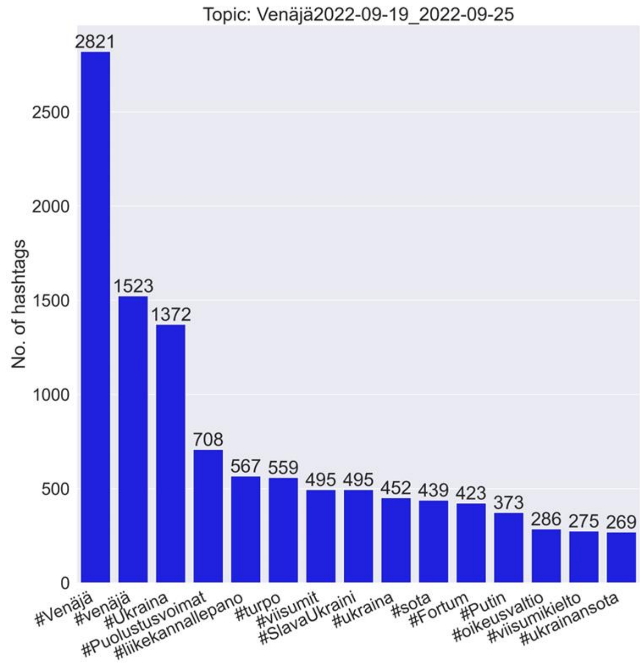

Seed word: Fortum (2022-09-17-2022-09-24)
Avg Tweets per User
Date


Seed word: Huumeet (2022-07-21-2022-07-31)
Avg Tweets per User
Date

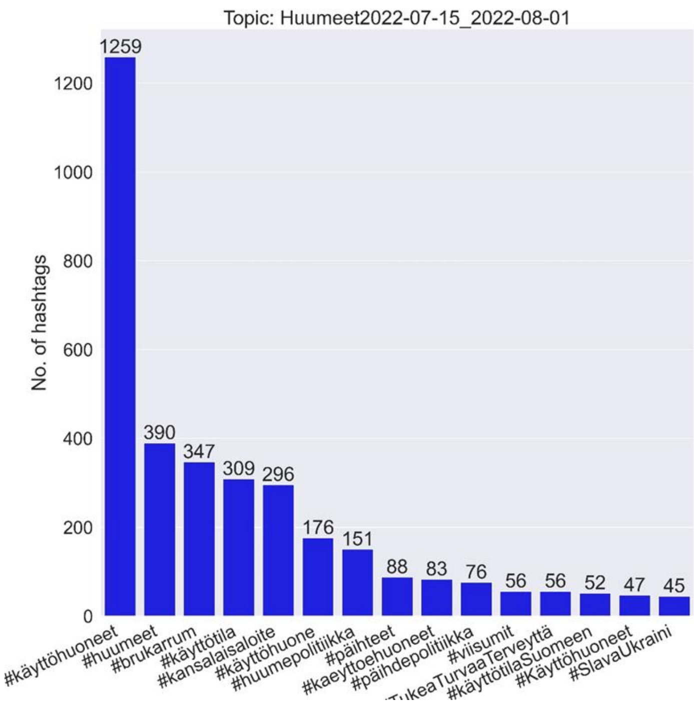

Seed word: #huuhkajat (2022-09-18-2022-09-27)
Avg Tweets per User
Date

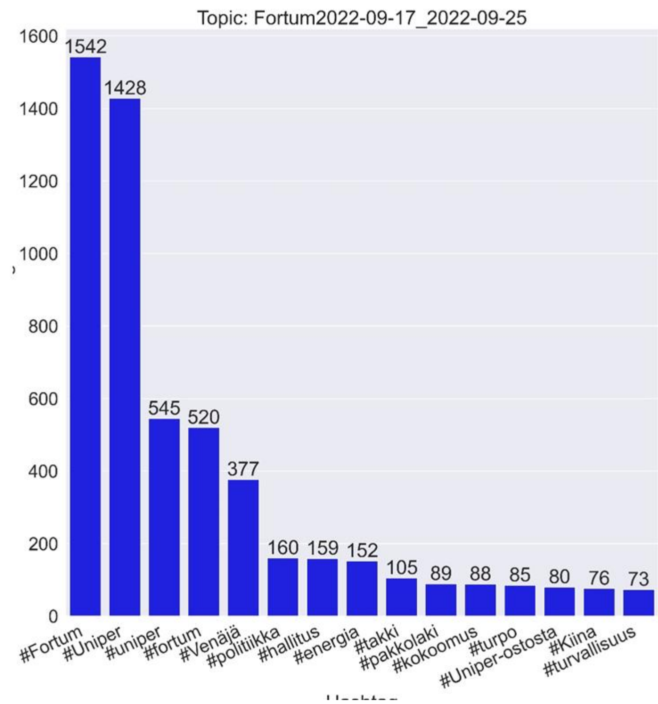

Seed word: #Jodi (2022-10-05-2022-10-15)
Avg Tweets per User
Date

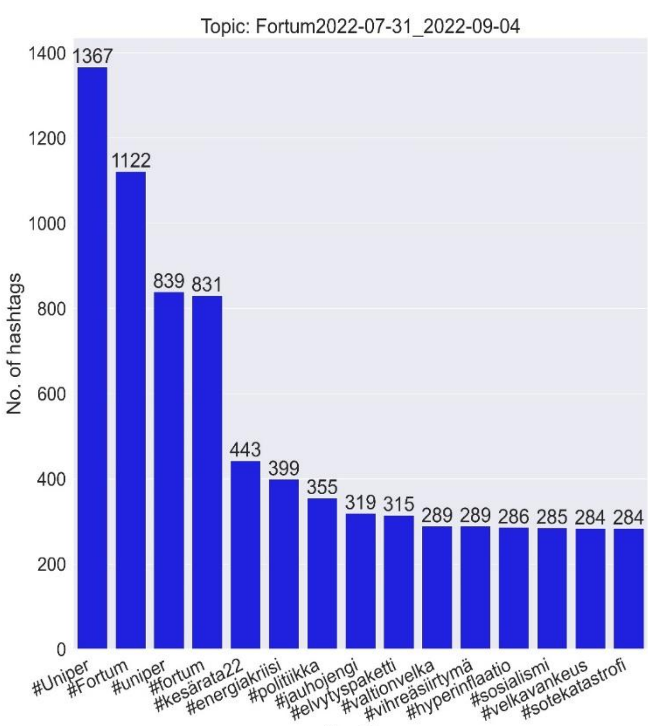


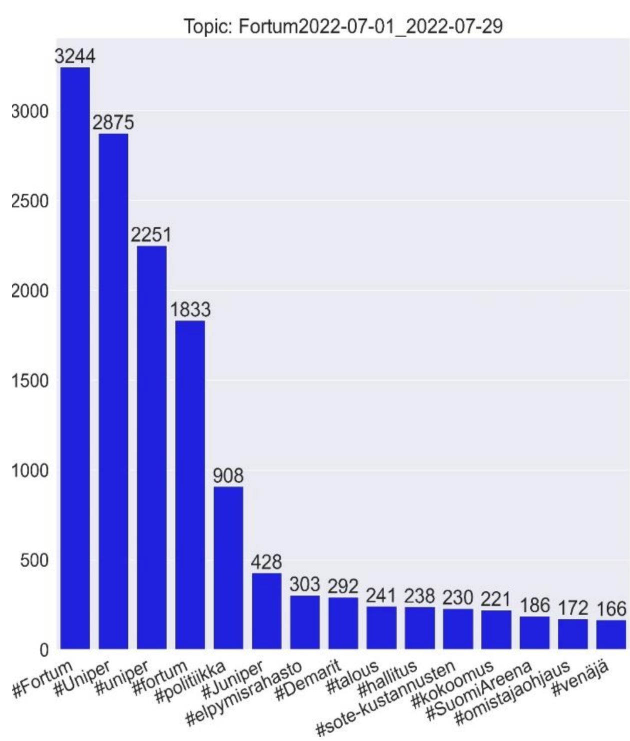


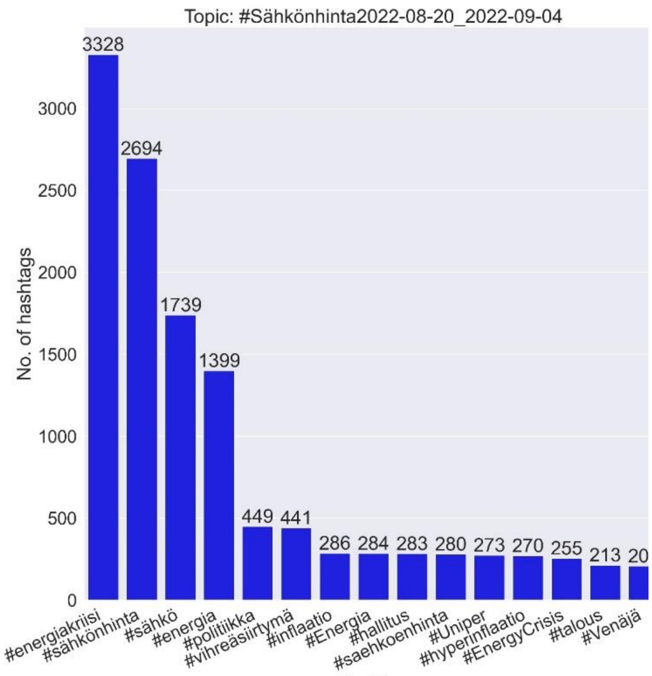
Appendix 3. The 15 most popular hashtags by topic

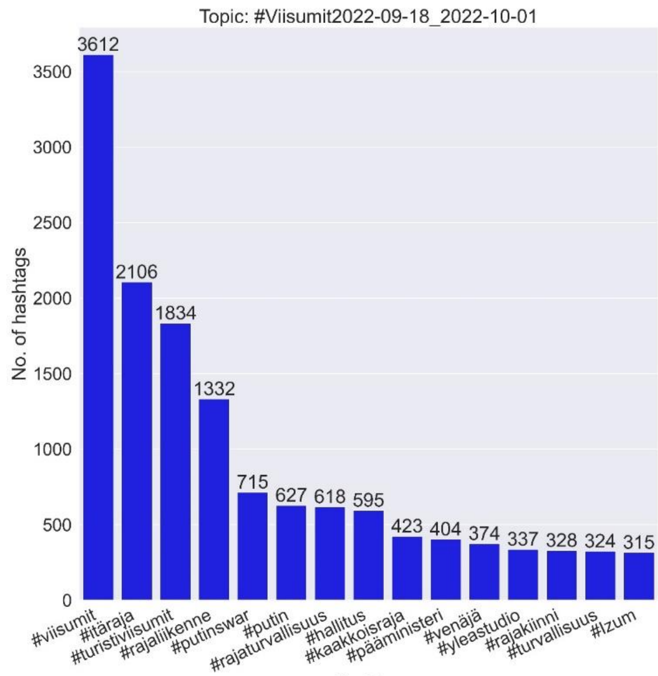

Topic: Huumeet2022-07-15_2022-08-01
No. of hashtags
0
200
400
600
800
1000
1200
1259
390
347
309
296
176
151
88
83
76
56
56
52
47
45
#käyttöhuoneet
#huumeet
#brukarrum
#käyttötila
#kansalaisaloite
#käyttöhuone
#huumepolitiikka
#päihteet
#kaeyttoehuoneet
#päihdepolitiikka
#viisumit
#käyttötilaSuomeen
#Käyttöhuoneet
#SlavaUkraini

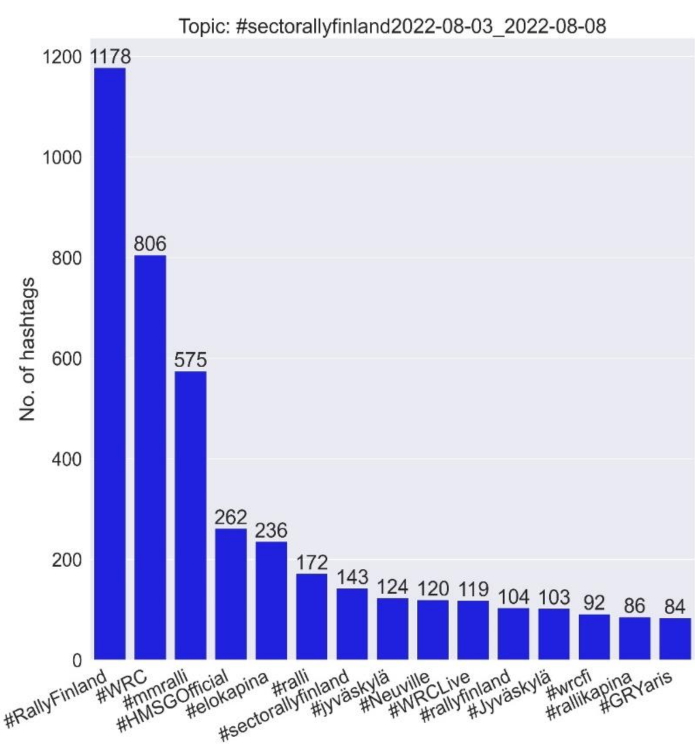

Topic: Fortum2022-09-17_2022-09-25
0
200
400
600
800
1000
1200
1400
1600
1542
1428
545
520
377
160
159
152
105
89
88
85
80
76
73
#Fortum
#Uniper
#uniper
#fortum
#Venäjä
#politiikka
#hallitus
#energia
#takki
#pakkolaki
#kokoomus
#turpo
#Uniper-ostosta
#Kiina
#turvallisuus

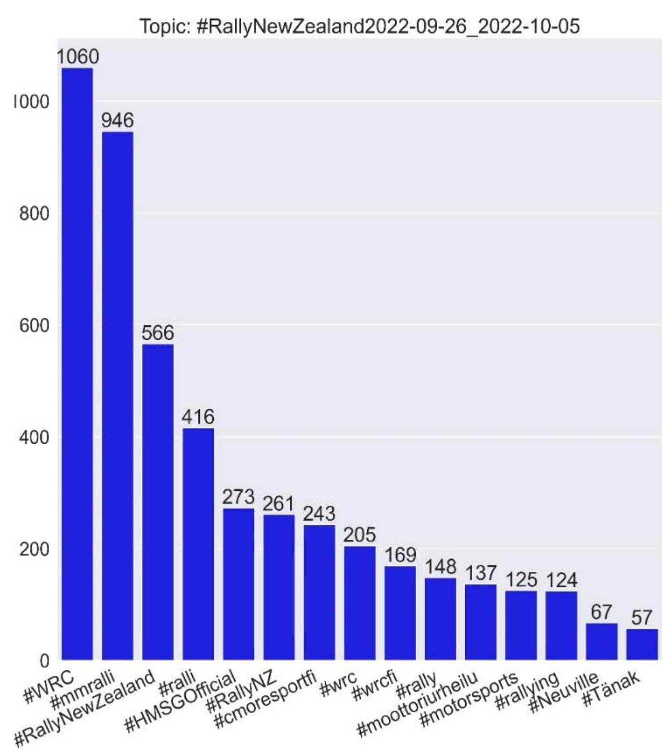

Topic: Fortum2022-07-31_2022-09-04
No. of hashtags
0
200
400
600
800
1000
1200
1400
1367
1122
839
831
443
399
355
319
315
289
289
286
285
284
284
#Uniper
#Fortum
#uniper
#fortum
#kesärata22
#energiakriisi
#politiikka
#jauhojengi
#elvytyspaketti
#valtionvelka
#vihreäsiirtymä
#hyperinflaatio
#sosialismi
#velkavankeus
#sotekatastrofi

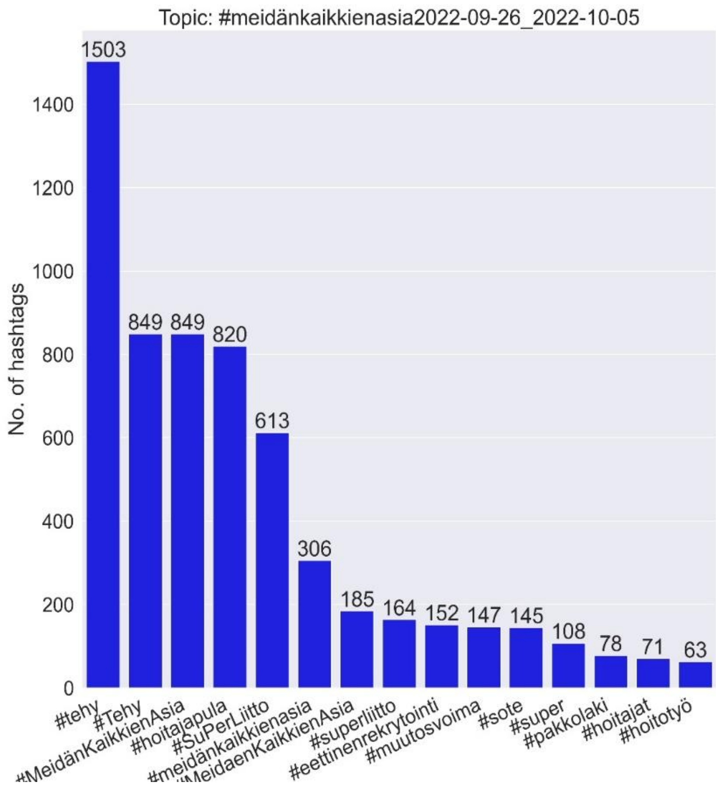

Topic: Fortum2022-07-01_2022-07-29
0
500
1000
1500
2000
2500
3000
3244
2875
2251
1833
908
428
303
292
241
238
230
221
186
172
166
#Fortum
#Uniper
#uniper
#fortum
#politiikka
#Juniper
#elpymisrahasto
#Demarit
#talous
#hallitus
#sote-kustannusten
#kokoomus
#SuomiAreena
#omistajaohjaus
#venäjä

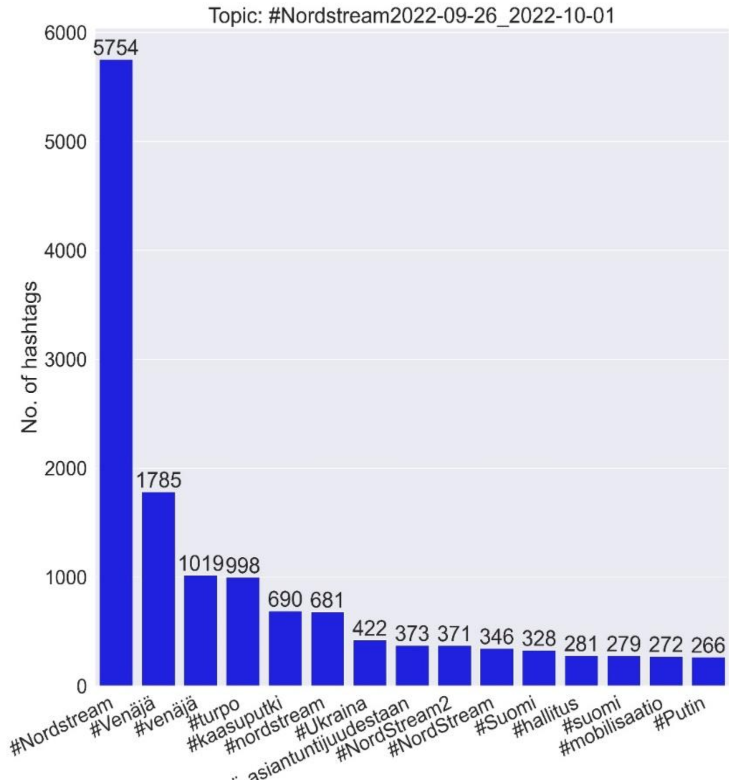

Topic: #Sähkönhinta2022-08-20_2022-09-04
No. of hashtags
3328
2694
1739
1399
449 441
286 284 283 280 273 270 255 213 20
#energiakriisi
#sähkönhinta
#sähkö
#energia
#politiikka
#vihreäsiirtymä
#inflaatio
#Energia
#hallitus
#saehkoenhinta
#Uniper
#hyperinflaatio
#EnergyCrisis
#talous
#Venäjä

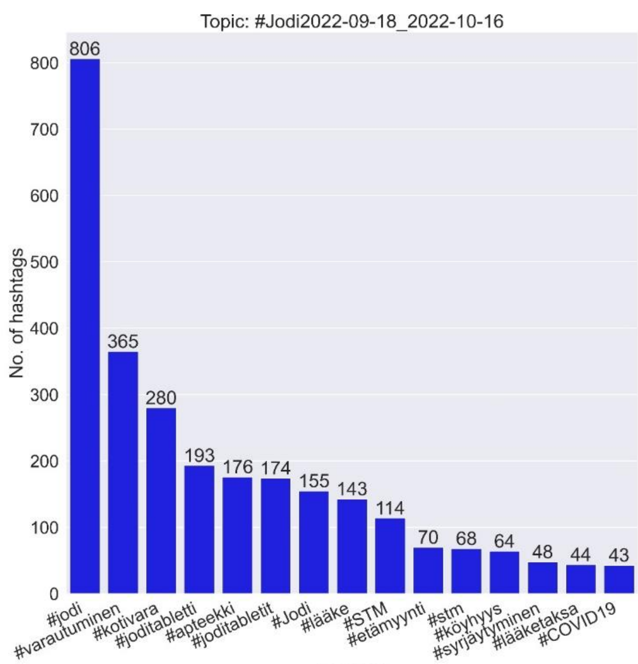

Topic: #Viisumit2022-09-18_2022-10-01
No. of hashtags
3612
2106
1834
1332
715 627 618 595
423 404 374 337 328 324 315
#viisumit
#itäraja
#turistiviisumit
#rajaliikenne
#putinswar
#putin
#rajaturvallisuus
#hallitus
#kaakkoisraja
#pääministeri
#venäjä
#yleastudio
#rajakiinni
#turvallisuus
#lzum

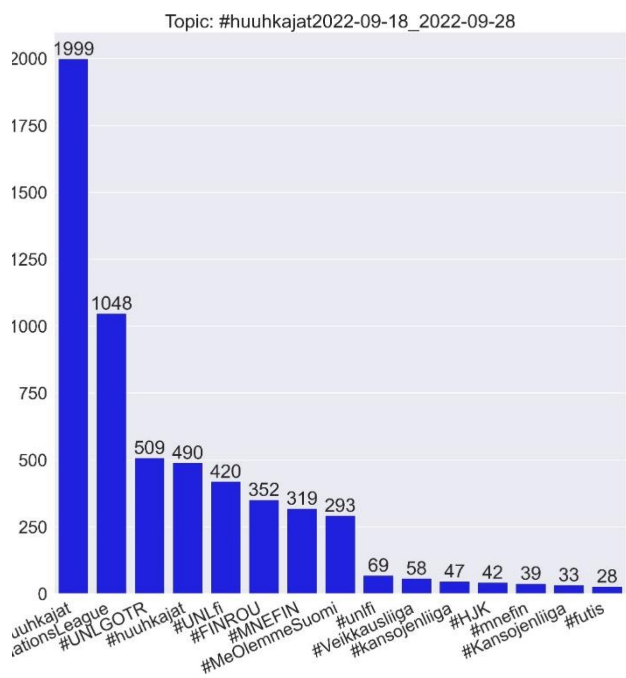

Topic: #sectorallyfinland2022-08-03_2022-08-08
No. of hashtags
1178
806
575
262 236
172 143 124 120 119 104 103 92 86 84
#RallyFinland
#WRC
#mmralli
#HMSGOfficial
#elokapina
#ralli
#sectorallyfinland
#jyväskylä
#Neuville
#WRCLive
#rallyfinland
#Jyväskylä
#wrcfi
#rallikapina
#GRYaris


Topic: #RallyNewZealand2022-09-26_2022-10-05
1060
946
566
416
273 261 243 205
169 148 137 125 124 67 57
#WRC
#mmralli
#RallyNewZealand
#ralli
#HMSGOfficial
#RallyNZ
#cmoresportfi
#wrc
#wrcfi
#rally
#moottoriurheilu
#motorsports
#rallying
#Neuville
#Tänak

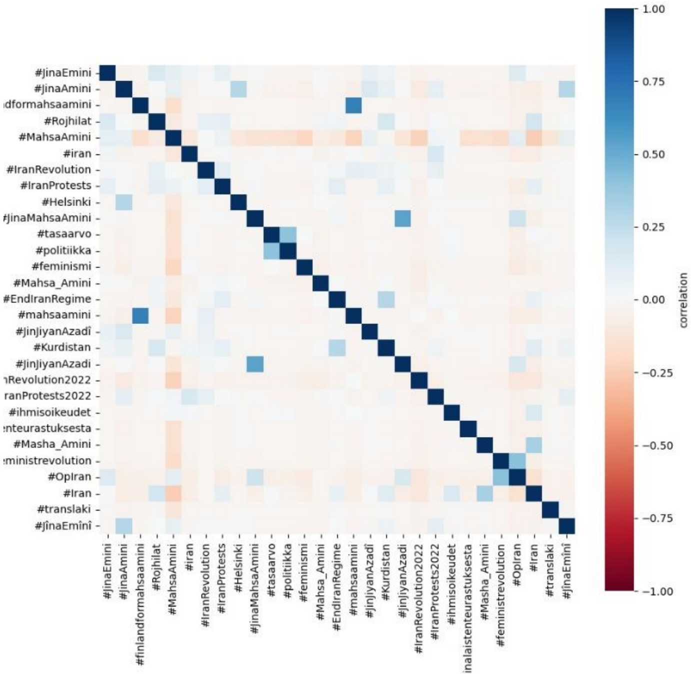

Topic: #meidänkaikkienasia2022-09-26_2022-10-05
No. of hashtags
0
200
400
600
800
1000
1200
1400
1503
849
849
820
613
306
185
164
152
147
145
108
78
71
63
#tehy
#Tehy
#MeidänKaikkienAsia
#hoitajapula
#SuPerLiitto
#meidänkaikkienasia
#superliitto
#eettinenrekrytointi
#muutosvoima
#sote
#super
#pakkolaki
#hoitajat
#hoitotyö


Topic: #Nordstream2022-09-26_2022-10-01
No. of hashtags
0
1000
2000
3000
4000
5000
6000
5754
1785
1019
998
690
681
422
373
371
346
328
281
279
272
266
#Nordstream
#Venäjä
#venäjä
#turpo
#kaasuputki
#nordstream
#Ukraina
#NordStream2
#NordStream
#Suomi
#hallitus
#suomi
#mobilisaatio
#Putin

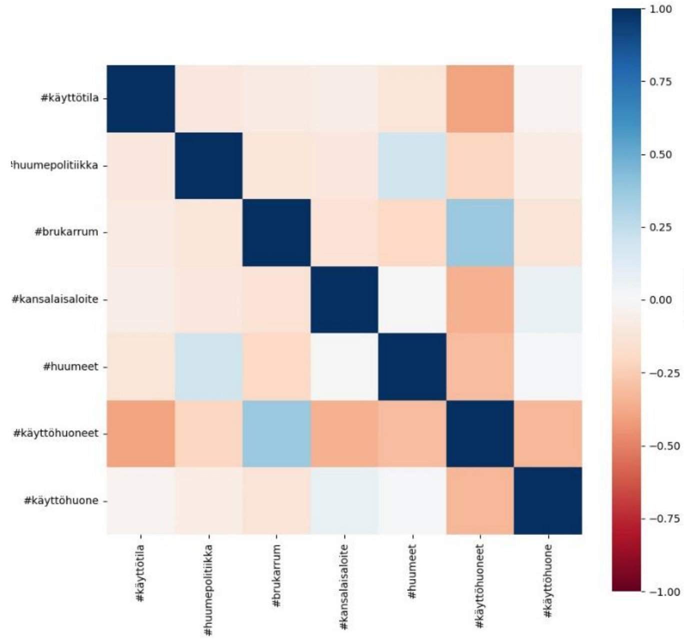

Topic: #Jodi2022-09-18_2022-10-16
No. of hashtags
0
100
200
300
400
500
600
700
800
806
365
280
193
176
174
155
143
114
70
68
64
48
44
43
#jodi
#varautuminen
#kotivara
#joditabletti
#apteekki
#joditabletit
#Jodi
#lääke
#STM
#etämyynti
#stm
#köyhyys
#syrjäytyminen
#lääketaksa
#COVID19


Topic: #huuhkajat2022-09-18_2022-09-28
0
250
500
750
1000
1250
1500
1750
2000
1999
1048
509
490
420
352
319
293
69
58
47
42
39
33
28
#UNLGOTR
#huuhkajat
#UNLfi
#FINROU
#MNEFIN
#MeOlemmeSuomi
#unlfi
#Veikkausliiga
#kansojenliiga
#HJK
#mnefin
#Kansojenliiga
#futis

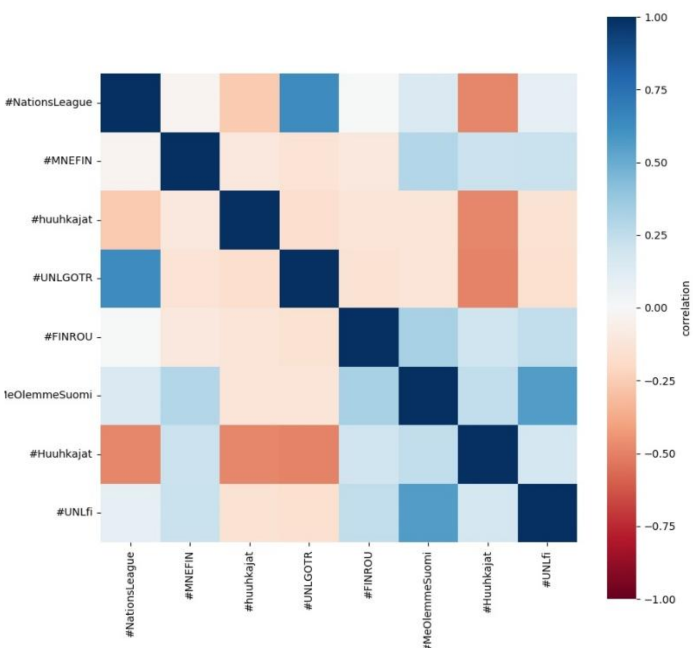
Appendix 4. Correlation matrices



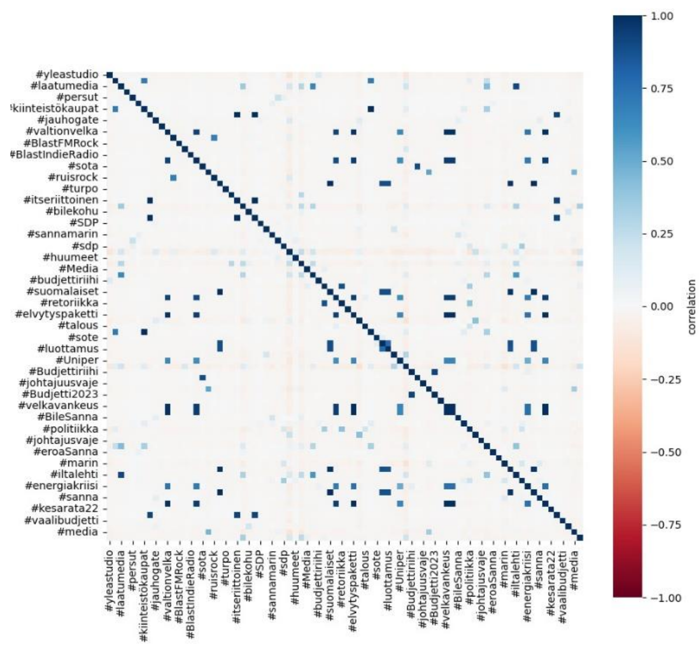
Iran2022-09-18_2022-10-13



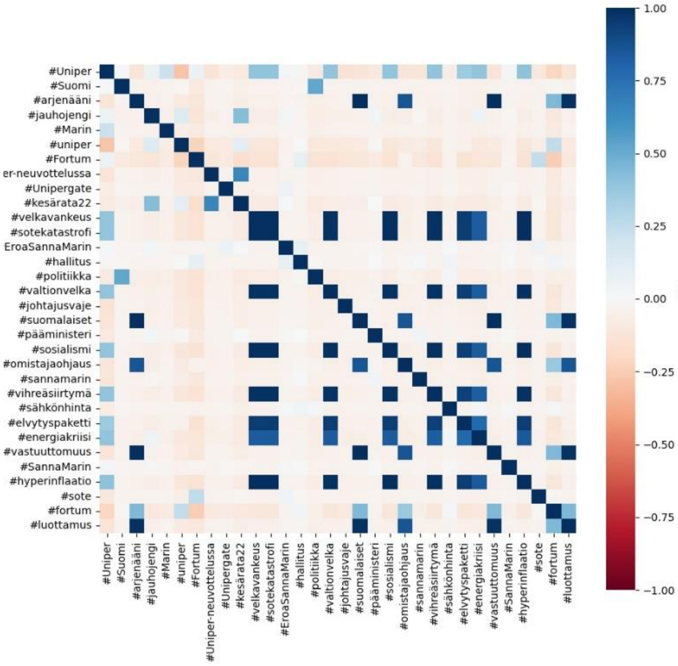
Huumeet2022-07-15_2022-08-01



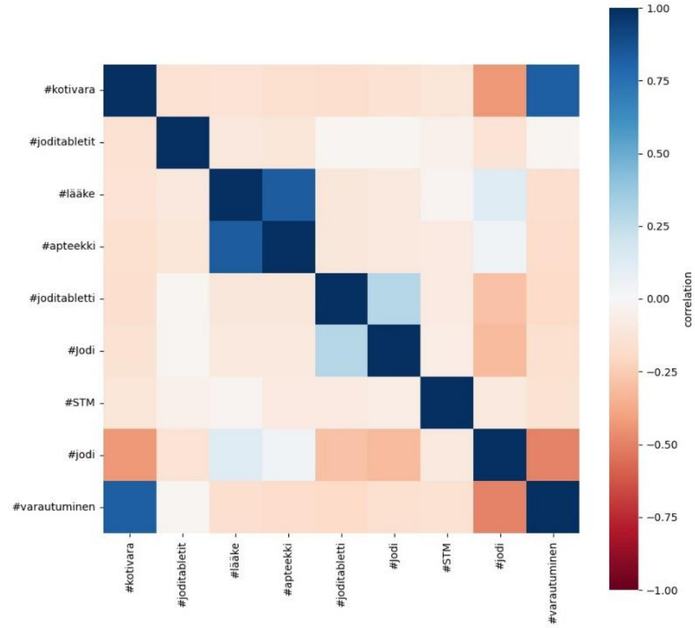
#huuhkajat2022-09-18_2022-09-28



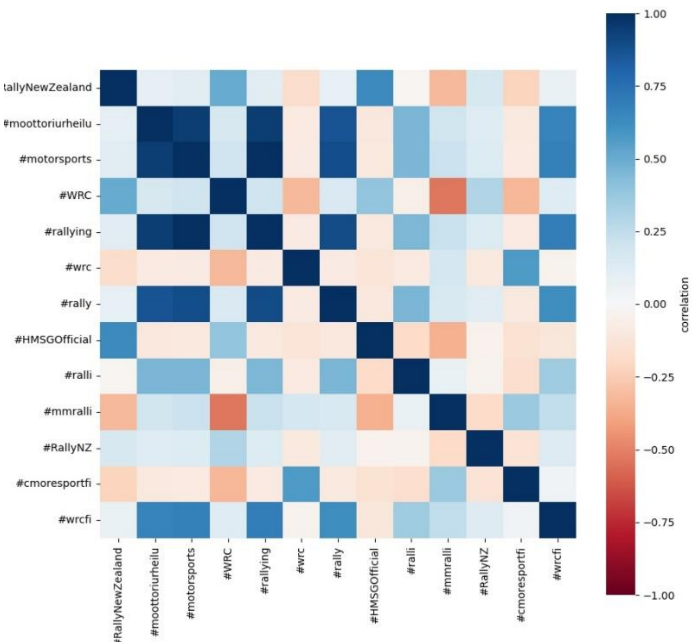
Marin2022-08-17_2022-09-05



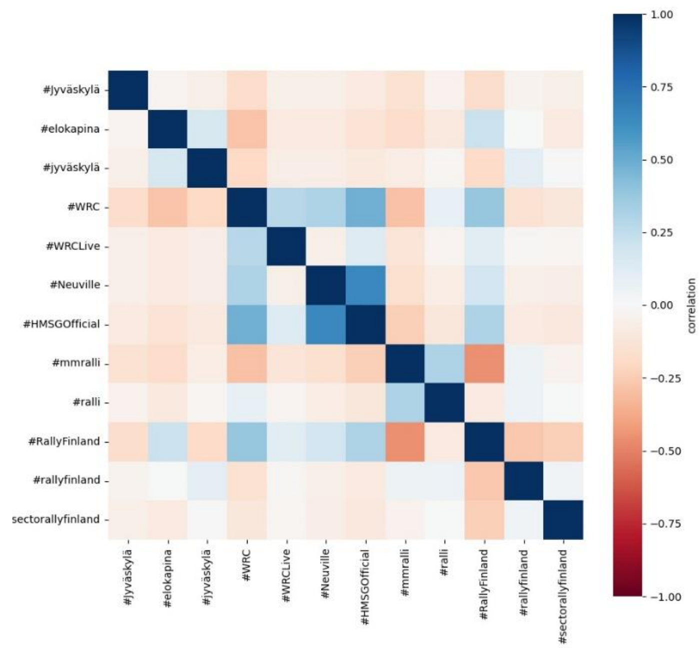
Fortum2022-07-31_2022-09-04



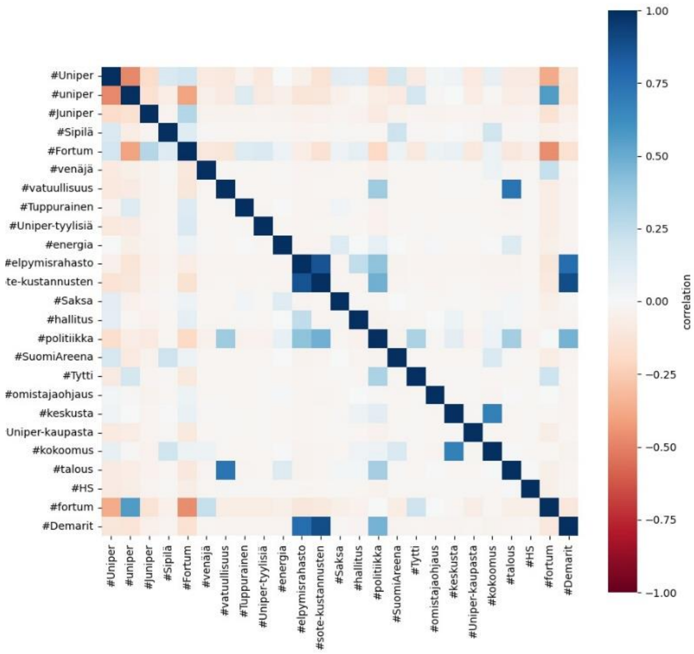
#Jodi2022-09-18_2022-10-16



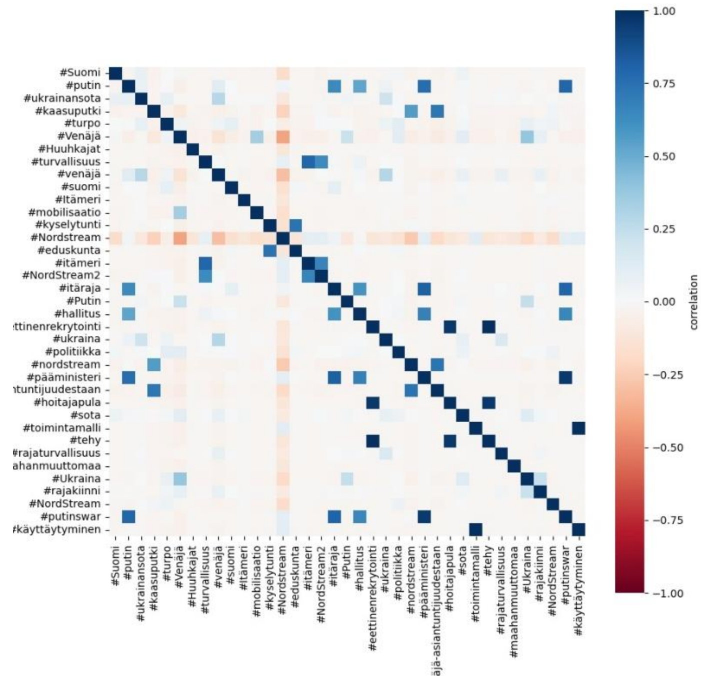
#RallyNewZealand2022-09-26_2022-10-05



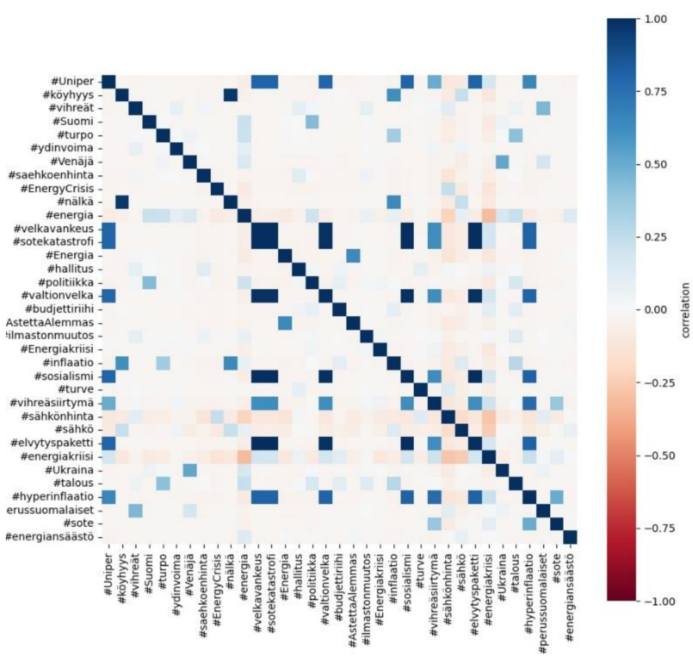
#sectorallyfinland2022-08-03_2022-08-08



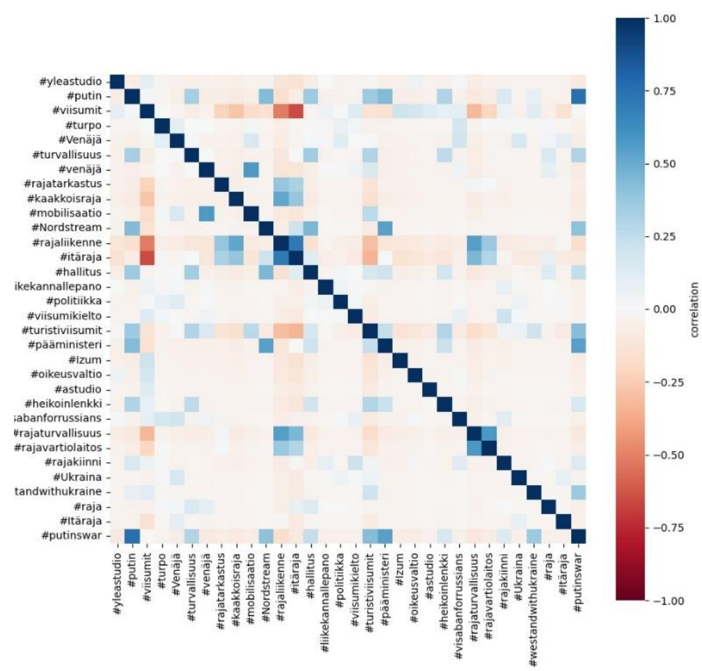
Fortum2022-07-01_2022-07-29



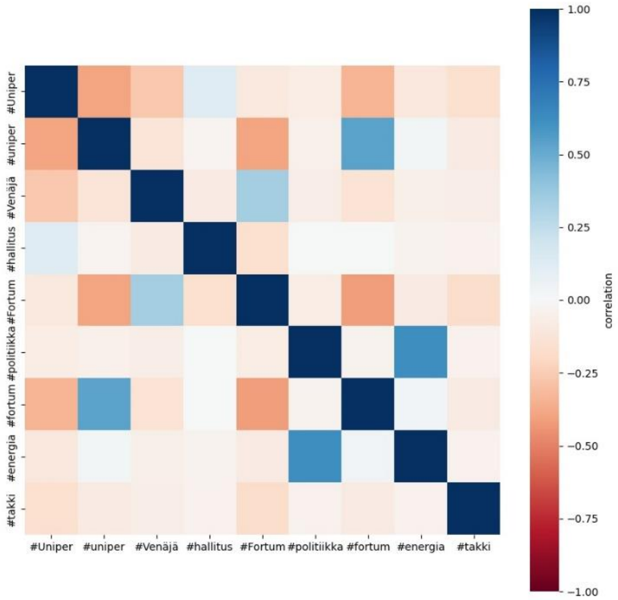
#Nordstream2022-09-26_2022-10-01



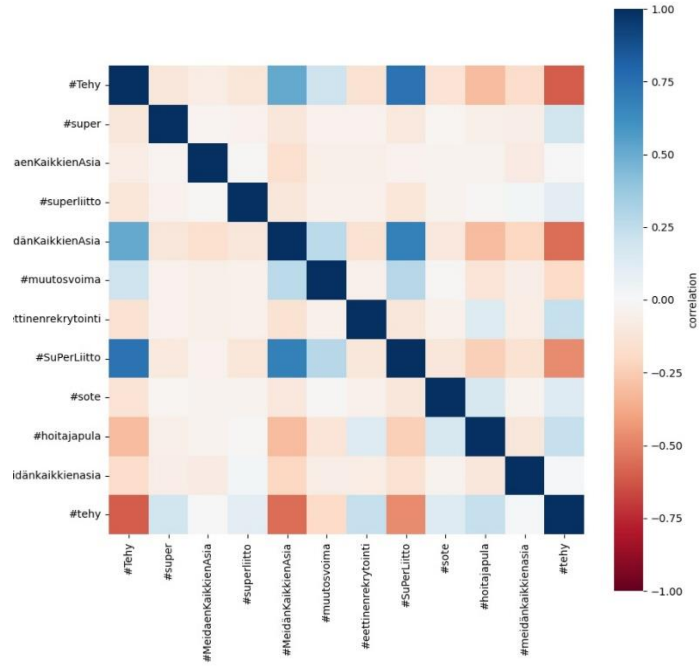
#Sähkönhinta2022-08-20_2022-09-04



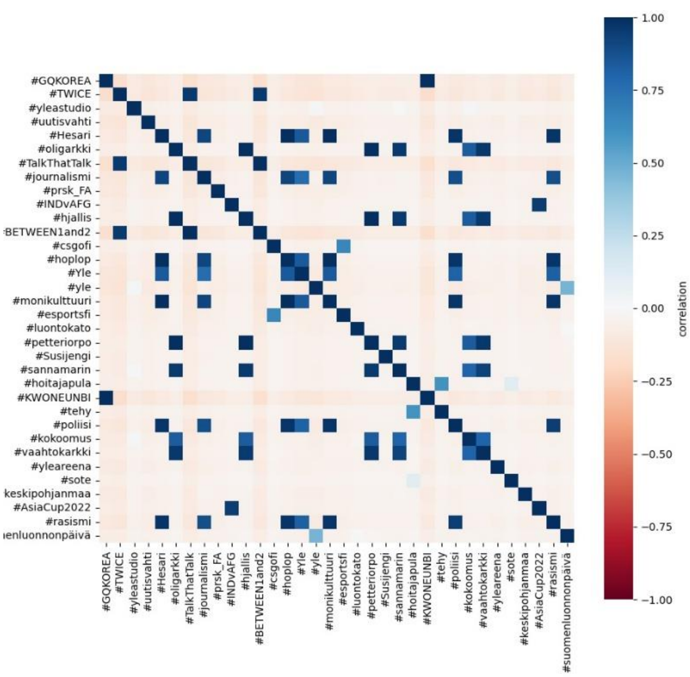
#Viisumit2022-09-18_2022-10-01



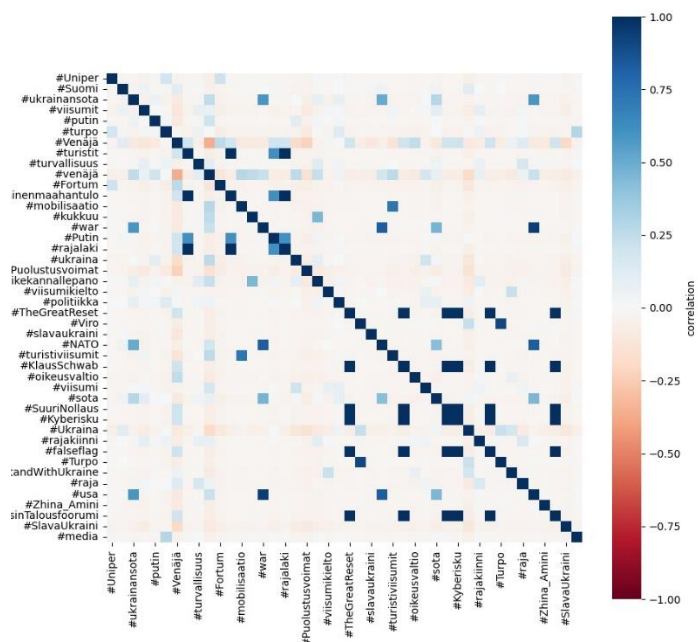
Fortum2022-09-17_2022-09-25



#meidänkaikkienasia2022-09-26_2022-10-05

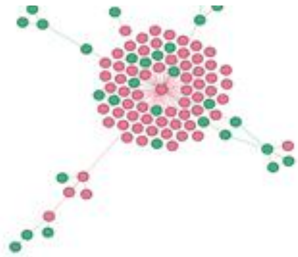


yle2022-08-27_2022-09-12



Venäjä2022-09-19_2022-09-25

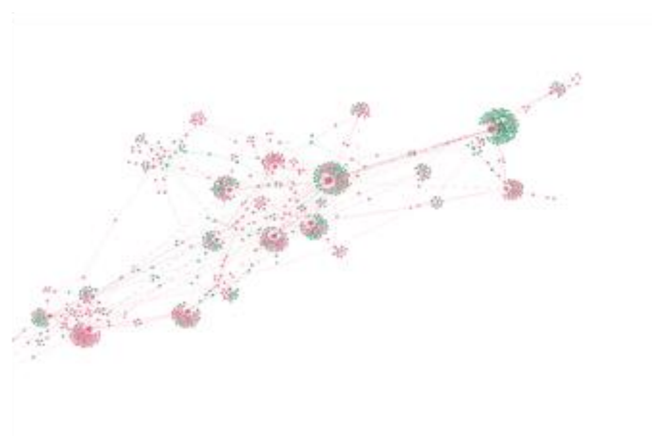
Appendix 5. Gephi representation of each network



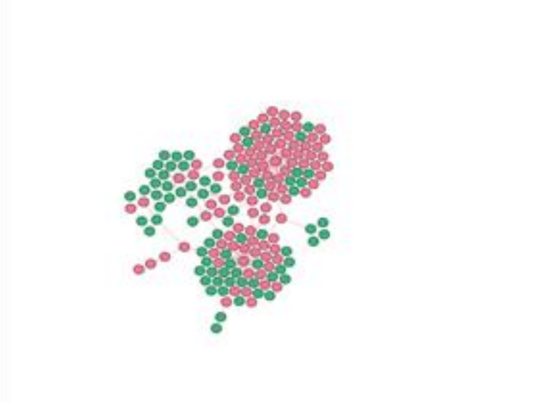
#huuhkajat2022-09-18_2022-09-28

Huumeet2022-07-15_2022-08-01

Iran2022-09-18 2022-10-13



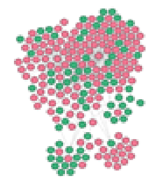
#Jodi2022-09-18_2022-10-16



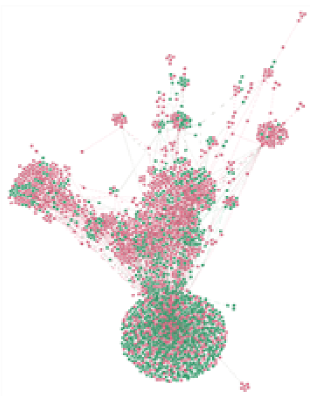
#RallyNewZealand2022-09-26_2022-10-05

#Sähkönhinta2022-08-20_2022-09-04

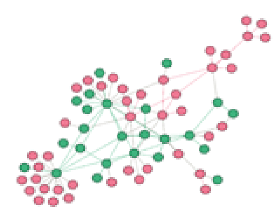
Marin2022-08-17_2022-09-05



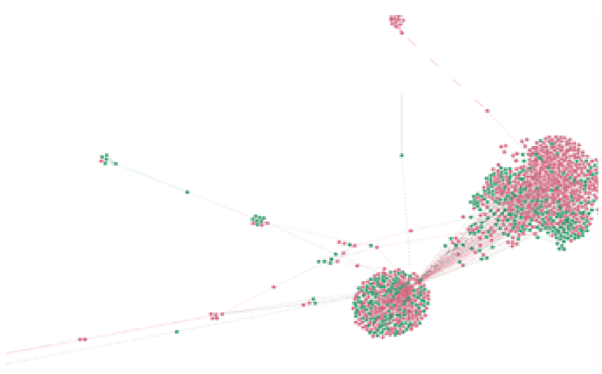
#meidänkaikkienasia2022-09-26_2022-10-05



#Nordstream2022-09-26_2022-10-01



#sectorallyfinland2022-08-03_2022-08-08

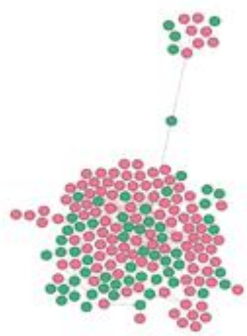


Fortum2022-07-31_2022-09-04

yle2022-08-27_2022-09-12

Venäjä2022-09-19_2022-09-25

#Viisumit2022-09-18_2022-10-01



Fortum2022-09-17_2022-09-25

Fortum2022-07-01_2022-07-29

Appendix 6. P-values of Spearman rank correlation

	Betweenness	E-I Index	Adaptive E-I Index	Boundary Connectivity	Dipole Moment	Modularity	RWC	ARWC
Betweenness	1,000	0,076	0,094	0,721	0,829	0,778	0,953	0,829
E-I Index	0,076	1,000	0,000	0,134	0,009	0,012	0,037	0,010
Adaptive E-I Index	0,094	0,000	1,000	0,122	0,007	0,009	0,053	0,012
Boundary Connectivity	0,721	0,134	0,122	1,000	0,244	0,528	0,103	0,458
Dipole Moment	0,829	0,009	0,007	0,244	1,000	0,004	0,219	0,002
Modularity	0,778	0,012	0,009	0,528	0,004	1,000	0,064	0,387
RWC	0,953	0,037	0,053	0,103	0,219	0,064	1,000	0,347
ARWC	0,829	0,010	0,012	0,458	0,002	0,387	0,347	1,000

Appendix 7. Rank table

Topic (Collection Period)	Betweenness	E-I Index	Index	Boundary Connectivity	Dipole Moment	Modularity	RWC	ARWC
#huuhkajat, 2022-09-18-2022-09-27	9	2	2	14	16	16	16	11
#Jodi, 2022-10-05-2022-10-15	8	12	12	15	9	4	11	10
#meidänkaikkienasia, 2022-09-27-2022-10-04	4	1	1	16	13	15	13	16
#Nordstream, 2022-09-26-2022-10-01	11	7	7	10	3	6	12	5
#RallyNewZealand, 2022-09-26-2022-10-04	14	11	11	8	4	3	1	6
#sectorallyfinland, 2022-08-03-2022-08-06	15	3	3	11	14	11	3	12
#Sähkönhinta, 2022-08-20-2022-09-03	16	4	5	1	12	8	10	14
#Viisumit, 2022-09-18-2022-10-01	10	8	8	9	2	7	9	4
Fortum, 2022-07-10-2022-07-28	12	9	10	5	5	9	14	8
Fortum, 2022-07-31-2022-09-03	13	10	9	4	7	12	4	3
Fortum, 2022-09-17-2022-09-24	1	15	15	2	6	1	2	13
Huumeet, 2022-07-21-2022-07-31	7	5	4	6	15	13	9	15
Iran, 2022-09-18-2022-10-11	3	14	14	7	1	2	5	2
Marin, 2022-08-17-2022-09-03	2	16	16	3	8	14	7	1
Venäjä, 2022-09-19-2022-09-24	6	6	6	12	11	10	15	9
yle, 2022-08-27-2022-09-11	5	13	13	13	10	5	6	7