WORKING PAPERS

Political Data Science

Juan Carlos Medina Serrano, Simon Hegelich, Morteza Shahrezaye, Orestis Papakyriakopoulos

Social Media Report: The 2017 German Federal Elections

TUM.University Press

Juan Carlos Medina Serrano, Simon Hegelich,
Morteza Shahrezaye, Orestis Papakyriakopoulos

Social Media Report: The 2017 German Federal Elections

The German National Library has registered this publication in the German National Bibliography. Detailed bibliographic data are available on the internet at https://portal.dnb.de.

Imprint

Editor: Simon Hegelich
Layout design and typesetting: Juan Carlos Medina Serrano
Layoutguidelines for cover design: Designbuero Josef Grillmeier, Munich
Cover design: Caroline Ennemoser
Cover illustration: Juan Carlos Medina Serrano

TUM.University Press
Technical University of Munich
Arcisstrasse 21
80333 Munich

DOI: 10.14459/2018md1452635
ISBN print: 978-3-95884-015-7

www.tum.de

TABLE OF CONTENTS

1. Introduction

The emergence of social media has changed the way in which political discourse takes place. The dynamics between social media and politics have a direct impact on society. Therefore, it is crucial that researchers use new digital tools to investigate the interactions between users and social media platforms. Additionally, the dangers of online manipulation need to be considered, especially since social media channels allow information to spread quickly without the overview of traditional gatekeepers such as journalists and editors. Increasingly more media channels report on the threats of digital propaganda and disinformation campaigns. The Science report on fake news underlines the fact that insufficient research has been conducted in this area [13]. Another concern is that most of the previous work on social media interaction has centered on the US. Accordingly, the effect of social media on politics and social dynamics in other countries has to be further explored.

The goal of this study was to track the social media landscape during the period prior to the 2017 German federal election. According to the Reuter report on news from the Hans-Below Institute, 31% of the German population consume media through social media and for the demographic group of 18-34 years old it is their main source of information. By comparison, in the US approximately two-thirds of the population use social media as a source of news. This was reported by the PEW Research Center survey in 2017. Even though social media in Germany does not play the same dominant role as in other developed countries, the trend continues to show an increase in the importance of social media on public opinion. Moreover, the fast changes in social media mark a parallel with those in the political panorama. The 2017 election definitely brought some unprecedented changes in Germany's political sphere.

In 2017, a far-right party entered the German parliament for the first time in over half a century. With over 12 percent of the vote, the Alternative für Deutschland (AfD) emerged as the third strongest party in the German federal elections. This represented a schism in Germany's political sphere, especially considering the AfD's foundation only less than five years ago, as well as the limited electoral success of Germany's radical right in previous elections. Even after the elections, it took almost six months to form a government. This delay was mainly caused by the failure of a Jamaica coalition (an alliance between CSU, CDU, FDP and die Grünen), which was an unexpected surprise. The main differences between the parties included the topics of immigration, climate and the European integration process.

The intricacies of politics in social media can only be understood by monitoring the online user interactions. We decided to collect data from Facebook and Twitter during the months leading to the German federal election. We tracked Twitter for a period of seven months and collected Facebook data for one year. The Twitter data comprised tweets that belong to a sample of the German political online conversation. From Facebook, we retrieved the posts of the main political parties and their regional pages. For the analyses, we took into consideration the seven main political parties in Germany: CDU, CSU, AfD, Bündnis 90/Die Grünen, FDP, die Linke and SPD. For many of the plots throughout this study, we omit the names of the parties and simply show the colors that represent them.

Apart from a thorough analysis of the collected data, we decided to explore how users share German online media in social networks, especially regarding political news. The sharing behavior can offer insights into how users consume traditional media sources and new online media. Additionally, we established a partisanship score for each media to identify the preferences of supporters of different political parties. We present a map of online media in Germany, which can be directly compared to the media landscape in the US.

This is our final work in relation to the 2017 German federal elections. It expands our previously-published study entitled "Social Media im Wahlkampf" [9]. Some of the new content in this report includes fake news and Russian trolls on Twitter, text analysis and the evaluation of regional pages on Facebook, as well as the complete online media analysis.

2. Methods

In this study, we analyzed social media during the months leading up to the 2017 German federal election. We focused on the social media platforms Twitter and Facebook. Our methodology was selected to obtain relevant political data that could show the online interaction of users with politics in Germany. Table 2.1 shows an overview of the retrieved data and the methods we used for the analysis.

We started collecting Twitter data in March 2017 and continued the process until the election day. For this, we used the public Twitter Streaming API, which allows retrieving data by providing some key parameters, like hashtags, users or geolocations. However, it is not possible to obtain every tweet, since the API only returns at most a 1% sample of all the general Twitter data. In order to obtain a representative sample of German politics, we decided to gather tweets from 175 politically-relevant hashtags and 13,633 German Twitter accounts. The accounts include political parties, politicians, media portals, journalists, bloggers, redactors, correspondents, consultants and other important political actors. For these accounts, we also followed the retweets and mentions from users. From March 6 to the September 25, 2017, we collected 353,010,294 tweets.

With the help of the Twitter data, we looked for possible online manipulation. We followed trends during important events in Germany; for example, the TV duel and the day of the elections. We investigated the spread of some popular fake news stories and tried to find their impact, as well as identifying the most active accounts. In this study, we also aimed to identify social bots and their activity over time. Social bots are automated accounts that pretend to be

	Data	Methods
Twitter	**353,010,294** tweets with political relevant hashtags or from German Twitter accounts	- User network - Bot detection - URL analysis
Facebook	**37,152** posts from the German political parties' Facebook pages	- User network - Topic modeling
Online Media	**1,821,478** URLs extracted from tweets mentioning a German political party	- Graph analysis - Partisanship score

Table 2.1: *Data and methods used in the study.*

real users to manipulate public opinion and spread false information. For their detection, we used Botometer — a well-known bot detection tool from the Indiana University — and our own developed methods for identification.

In order to collect data from Facebook, we used the public Facebook API called Graph API. We obtained the data from the official Facebook pages of the seven political parties and their regional pages. Each party has an official page for each one of the sixteen German states. For each page, we collected all of the comments and reactions from the posts and their comments. The period of data extraction was one year, from September 2016 to September 2017. Overall, we collected 37,152 posts from 102 Facebook pages.

We used descriptive statistics techniques to investigate the interactions between the users and the political Facebook pages. Additionally, we performed text analysis on the user comments taken from the posts. We applied "topic modeling", a machine learning algorithm, to the posts and comments to obtain the main topics that were discussed during the election period. This algorithm allows analyzing large amounts of text data and discovering relevant topics.

Apart from the Facebook and Twitter data analysis, we were also interested in looking at the online media coverage and the relation between media sources on the web. For this purpose, we extracted all of the URLs that were shared in tweets mentioning any of the German political parties from our database. We analyzed 5,468,409 tweets and extracted 1,821,478 URLs. After filtering out the URLs pointing at tweets, Facebook posts or other social media platforms, we assigned the URLs to their corresponding site. The sites include a broad range of media sites, private corporations, blogs, campaign pages and government sites. Taking into consideration which media sources were shared by the same users, we were able to measure the proximity of the media sources. Moreover, we were able to assign a partisanship score for each media site, depending on the frequency of shares by the retweeters of the political parties. This analysis builds upon the methodology of the "Partisanship, Propaganda and Disinformation: Online Media and the 2016 U.S. Presidential Election" Harvard report. This report shows how the media in the US is heavily polarized between the right and left. We were able to make a comparison between the US and Germany's online media landscape.

Further details on the methodology are discussed in the following chapters and the bot detection methods are explained in the Appendix.

3. The German Twittersphere

During recent years, Twitter has become a channel for political communication. It is a place where many politicians engage with users and share their policies. They are keen to ascertain how to use Twitter to mobilize users for support. Research on the social dynamics in Twitter is vast and emphasis on the political context has been sought after. Twitter alone will not win elections, but it can nevertheless be an effective campaign tool.

Twitter was one of the main platforms of political discourse in the 2016 US campaigns, whereas in Germany the panorama is different. From the beginning of the service in 2006, Twitter did not connect with the German population as much as in the US or Britain. For example, in contrast to many world leaders, Chancellor Merkel has no Twitter account. It has been argued that the former 140-character limit was an impediment for Germans to express themselves. Nevertheless, Twitter should not be overlooked in analyzing the effect of social media on German politics. Important political events are discussed on a daily basis and journalists use this platform frequently as a source of information. Their perspective on Twitter defines what they perceive and affect their decisions on the information to be shared. Additionally, every German political party has its own Twitter account to communicate their messages to their followers in a direct and instant manner.

In order to gain a better understanding of Germany's political discussion on Twitter, we collected data from political German actors. At the same time, we followed a list of important hashtags and terms that relate to German politics. Our dataset contains 353,010,294 tweets. Each downloaded tweet contains much more than just text, as it can have more than 1,000

different features. These features include all of the information from the user, the tweet and the status of the tweet.

It is important to remark that our Twitter data is biased for two reasons. First, we only follow politically-relevant topics. We pursued a thorough investigation to have a complete list of users, terms and hashtags to follow, although small omissions may have occurred. The second reason is that Twitter only makes available a percentage of the complete data. The randomness of this sample has been called into question before [15]. Nevertheless, we have accumulated a significant number of tweets and we hope to have a representative sample that can show correct insights and give us a political panorama that cannot be appreciated only by looking at daily trends.

This chapter first focuses on showing the usage of Twitter by the political parties and the users who engage with them. Second, we discuss online manipulation techniques that were observed during the important events in the election period.

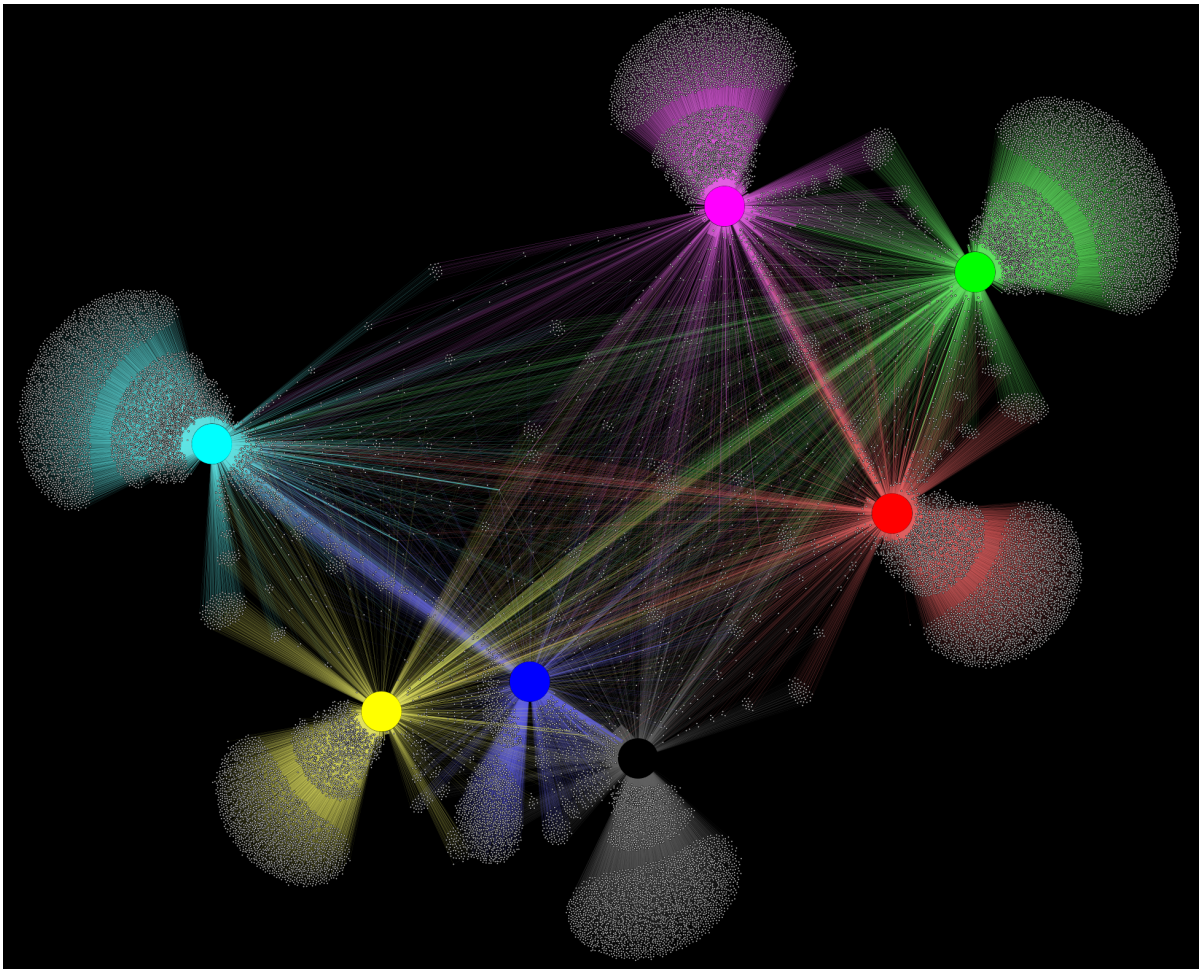

Figure 3.1: *Users who retweeted a political party. Each party is represented with its respective color.*

Partisan Users

What is the political sphere like on Twitter? Figure 3.1 aims to answer this question. It represents a network of users that have retweeted at least one of the accounts of the main German political parties. The color nodes represent the political parties and the smaller gray nodes the users. The position of the user depends on the number of retweets. The users closer to a party node are those who have retweeted the most. The fungi-shaped clusters around each of the color nodes represent the users who have only retweeted one political party. These can be seen as partisan users. Meanwhile, the users between the nodes have retweeted more than one party and could be cases of cross-partisanship.

We have used retweets as a measure of commitment to a party since previous studies [10] have shown that the act of retweeting mostly conveys support to a political party. By retweeting, the content is shared in the user's timeline and is seen by its followers. By contrast, mentioning a political actor (using @) can relate to expressing either support or dissatisfaction.

The network was modeled with Gephi's Force Atlas 2 algorithm, which makes nodes with many similar users come together and repels those that have less common users. Interestingly, the network algorithm puts the SPD, Die Grünen and Die Linke on one side and the CDU, CSU and FDP on the other. The latter ones are even closer to each other, showing that they have more common users. It was expected to have CSU and CDU close to each other, given their political union. Additionally, they were in a coalition, together with the FDP, until the 2013 election. From the two subgroups, the closest nodes are SPD and CDU, which are the parties of the grand coalition after 2013. We can see that on Twitter the user activity models Germany's political landscape to a certain extent.

Political Parties on Twitter

It is insufficient to look at the number of followers of the political parties on Twitter (Figure 3.2) to gain an understanding of the party's popularity. For example, the AfD is the political party with fewer followers on this platform, but as we will see in the following results it was the most dominant on Twitter during the election period.

Our main findings with respect to the political parties on Twitter are shown in Figure 3.3. The pie charts show four measures of party reachability: mentions of the party's name in tweets (volume of tweets), users' messages to the party's Twitter account (using @), retweets and the number of times the hashtag of each party was shared. The data corresponds to the seven-month period prior to the election. We observe that for three of the four measures the AfD has a larger percentage than the rest of the parties.

The AfD-related volume of tweets was greater than the tweets mentioning the other parties combined. On the other hand, die Grünen had the lowest amount of traffic, which does not reflect the fact that it has the account with the most followers. The chart with the number of hashtags in the collected tweets shows a similar pattern. For the collection process, we in-

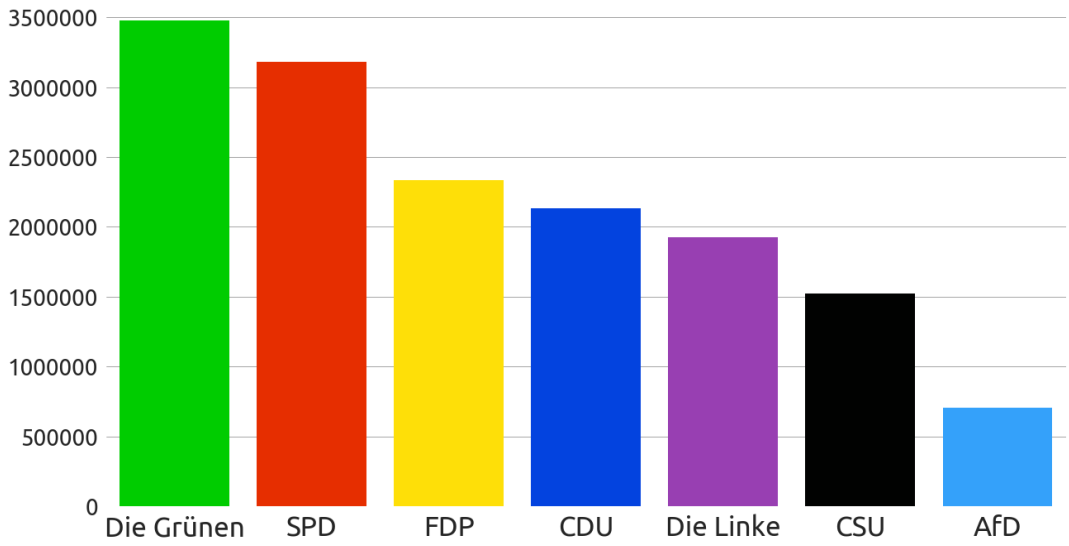

Figure 3.2: *Amount of followers for each German political party as of August 2017.*

cluded all possible hashtags for a party; for example, for Die Linke we followed both "dielinke" and "linke". For these first two measures, a tweet that included more than one hashtag or mentioned more than one party counted more than once. For this reason, the percentage is based on the number of occurrences rather than the number of tweets.

The bottom left chart in Figure 3.3 shows the messages from the users to the Twitter account of each party. This is the only measure where the AfD is not ranked in first place. In fact, the percentages are very similar to the electoral results. However, this does not imply that it is possible to predict elections with Twitter data. For the 2009 German federal election, a study [21] ascertained that the election results could have been predicted by the percentage of the volume of tweets with an error of less than 2 percent. If we used their same predictive measure for 2017, the AfD would have been predicted to have over 50% of the vote. This represents another sign that the online panorama has changed since the AfD entered the German political system. We believe that simply by looking at Twitter statistics it is insufficient to predict election outcomes, although they can be used as a factor in a more complex model that includes socioeconomic and regional factors.

The final chart shows the retweets made to the parties' tweets. As previously mentioned, a retweet normally shows a sign of affiliation to a party. Retweets help messages to spread on Twitter and this is why it is a relevant measure of reachability. The right-wing party managed to have more retweets than the other parties, thus spreading their message in this way. AfD's dominance on Twitter was certainly helped by constantly being in the media limelight and sending provocative messages that violated mainstream social taboos in Germany [17]. Previous research has shown that anger spreads faster on social media [5], and the main resonated AfD topics related to discontent with the administration.

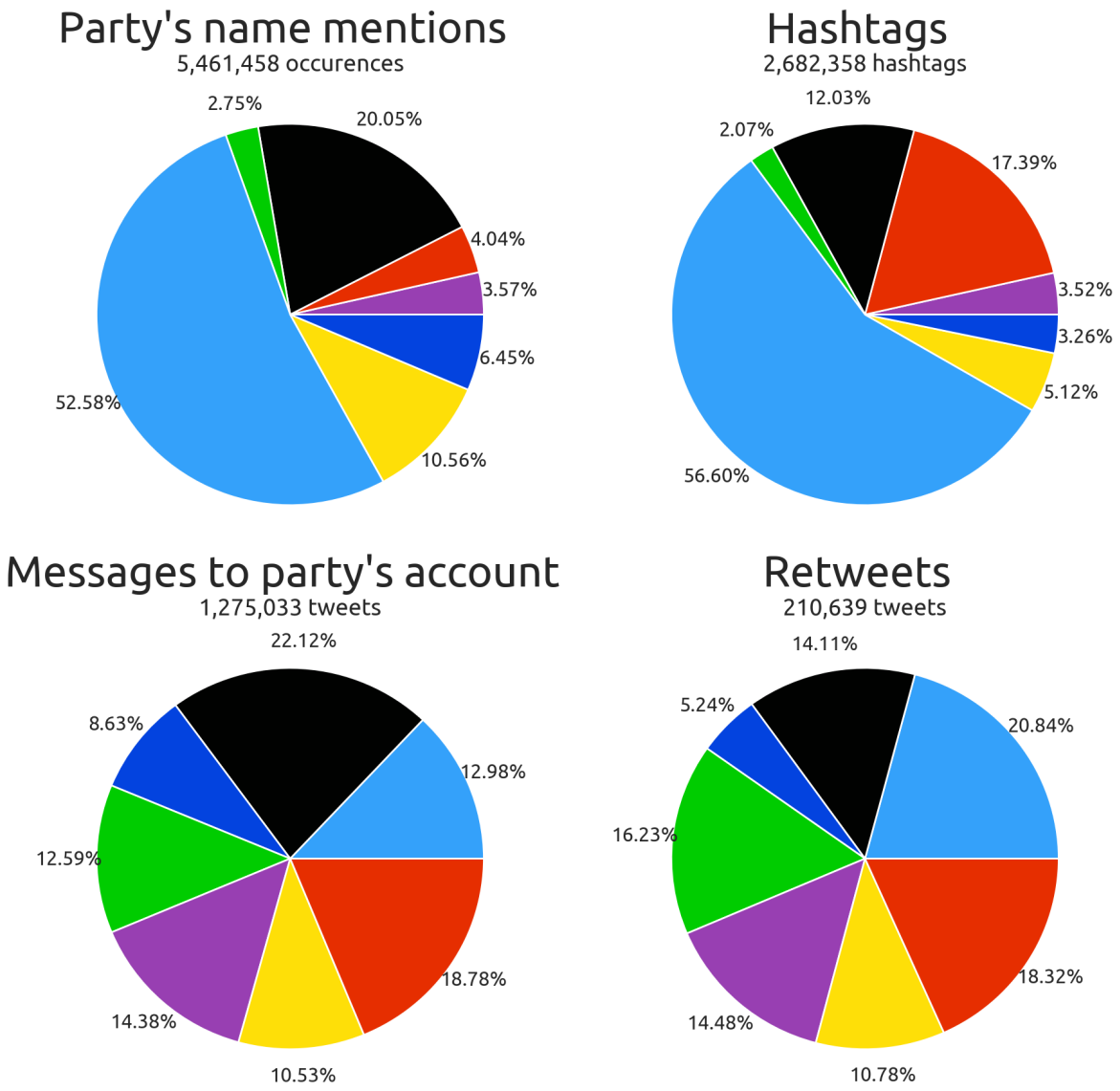

Figure 3.3: *Main statistics of reachability for the political parties. Each party is represented with its respective color. Top left: Amount of mentions. The percentage is based on the number of ocurrences and not the number of tweets. Top right: Hashtags in the collected tweets. For tweets with several hashtags, every hashtag counts. Bottom left: Number of tweets that mentioned an account, using @. Bottom right: Number of retweets from the political party accounts.*

In order to further understand how popular AfD was as a topic, we compared the trends over time concerning the usage of the word "AfD" with the words "Merkel" and "Schulz", the two main candidates in the election. We did this for three categories: German tweets, non-German tweets, and all tweets. We decided to divide tweets by language to have a comparison between the political discourse in Germany and the rest of the world. In order to have the completely accurate difference, we would need the geolocation of every tweet, however, less than one percent of all tweets include a geolocation [12] and in our collected tweets this value is around 0.05%. For this reason, focusing on the language is a better measure of discerning the difference in discourse within and outside the country. The time series for the three keywords are presented in Figure 3.4.

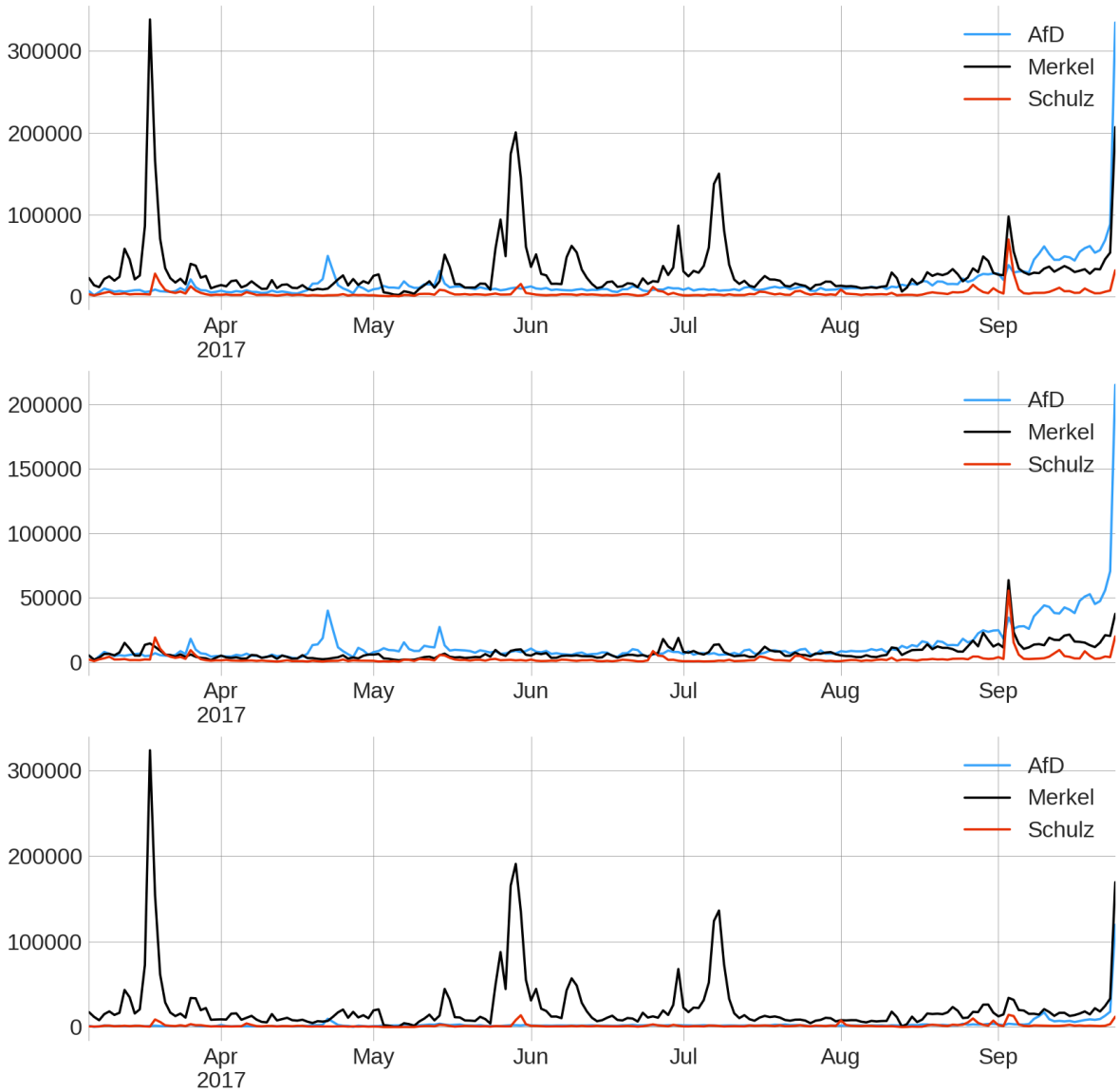

Figure 3.4: *Number of tweets mentioning AfD, Merkel and Schulz throughout the collection period. Top: Trends including all of the tweets. Middle: Only German tweets. Bottom: Trends excluding German tweets.*

The general trends show that Merkel was a more dominant topic on Twitter than the other two. The five peaks in Merkel's trend correspond to major events, namely: Merkel meeting US president Trump at the White House, then again at the NATO summit, the G20 summit, the TV duel and the election day. The volume of tweets corresponding to Schulz is never larger than Merkel throughout the seven months. Interestingly, starting in September, the number of tweets mentioning the AfD increases and takes over Merkel. This increase of Twitter volume during the period prior to the election could have been part of a strategy from the right-wing party to spread their political agenda. It can also be possible that Twitter users were simply more eager to talk about political parties. In order to have a better understanding of this phenomenon we investigated the possibility of online manipulation. Our results are discussed in the next section.

The panorama is different by looking at the trends using only German tweets. Here, the AfD's volume of tweets is usually above Merkel's trend. Moreover, the five peaks for the events concerning Merkel are not found. On the other hand, the trends with non-German tweets show that the conversation outside Germany focused mainly on Merkel. Therefore, we conclude that the AfD had a significant impact on the German Twittersphere, but not in the rest of Twitter.

Online Manipulation

One of the most powerful advantages of monitoring social media is to be aware of possible online manipulation. Although it is difficult to control this from the beginning, the efforts against it should focus on avoiding its propagation. By monitoring different online channels, it is also possible to inform the general public of current means of manipulation.

In Germany, online manipulation has emerged as an important political issue, fueled by fears of election meddling. In June 2017, German lawmakers passed a law against hate speech and fake news, the Netzwerkdurchsetzungsgesetz. This law enforces social network platforms to take down defamatory and hate speech content in a 24-hour period to avoid a fine of up to 50m Euro. Given that German policy-makers are eager to find measures of controlling mass manipulation online, it is important to understand what happened during the election period. For this section, we focus on three main aspects of online manipulation: fake news, social bots and intervention from foreign actors.

Fake News

One of the most controversial topics regarding online manipulation is the spread of fake news and its effect on political perception. Fake news is defined as "fabricated information that mimics news media content in form but not in organizational process or intent" [13]. Its goal can be to spread misleading information or simply false information to deceive people. There exists little research to date that sheds true light on the effect of fake news on political events. Additionally, there is no tool that can automatically classify between real and fake

news. Human curation is needed to understand which news is false or true, and even then it can be difficult to make a correct decision without the help of professional fact-checkers. For this reason, a complete review of all news shared in the millions of collected tweets is a difficult task. We decided to adopt two approaches to overcome this difficulty and have a general overview of the news that was shared during the election period.

The first approach involved looking at a relatively large sample of our data to find whether the most commonly-shared news included fake news. For this purpose, we selected the tweets that had mentioned at least one German political party and extracted the URLs that were shared in the tweets (the URL extraction procedure and the analysis of online media are the main topics of Chapter 5). The sample comprises more than 5 million tweets . We manually looked at the top 100 news shared in these tweets. Additionally, we checked the number of shares on Facebook for every media link found in the sample of tweets. We also manually analyzed the 100 most shared news on Facebook.

From the sample of tweets, none of the 100 most commonly-shared news was fake, but interestingly 51 of the top 100 were AfD-related. Most of them are polarizing stories that are meant to trigger political discussions. This reinforces the theory that populist parties thrive to be in the media limelight to gain as much publicity as possible [14]. We used the same URLs and checked their "popularity" on Facebook. The top results were different from those on Twitter. From the 100 most commonly-shared news, nine of them appear not to be completely accurate. As seen in Table 3.1, they are all related to the topic of refugees in Germany. Part of the message that is conveyed can be considered as misleading. They can be categorized as sensationalist stories with the purpose of triggering negative emotions to refugees. Most of them use certain facts incorrectly and show only one side of the story. Nevertheless, this is insufficient to categorize them as fake news. The study of fake news is not a trivial topic and requires extensive field knowledge.

Headline	Media	Shares
BKA vertuscht Straftaten von 600.000 Flüchtlingen!	Journalistenwatch	107,740
Syrer mit vier Frauen und 23 Kindern erhält monatlich ca. 30.030 €	Denken macht frei	97,228
Claudia Roth fordert mehr Flüchtlinge für Europa	Berlin Journal	92,021
Polizisten brechen Schweigen: Asylanten-Verbrechen werden auf Weisung von oben vertuscht	Unzensuriert.at	90,373
Merkel will in Afrika für Einwanderung nach Deutschland werben	Deutsche Wirtschafts Nachrichten	86,028
Merkel hofft auf 12 Millionen Einwanderer	Wochenblick	68,215
Ja, Asylbewerber bekommen wirklich kostenlosen Zahnersatz	Freie Zeiten	66,482
Flüchtlinge mit zwei Ehefrauen in Deutschland: Beide können Sozialhilfe bekommen	Epochtimes	61,942
Auf jeden neugeborenen Deutschen kommen fünf neue Migranten	Freie Zeiten	57,659

Table 3.1: *Misleading stories in the top 100 shared news on Facebook taken from the tweets' URLs.*

A similar study regarding the 2016 US election [6] concluded that “at least in the 2016 election [the fake news ‘framing’] seems to have played a relatively small role in the overall scheme of things.” According to the study, disinformation and propaganda from dedicated partisan sites played a stronger role in the election. By looking at the most-commonly shared news on both Twitter and Facebook, we argue that a similar conclusion applies to the German election. Our data suggest that the larger players in the online communication were propaganda topics on the AfD and critical opinions on the refugee crisis.

A second approach involved focusing on specific cases of fake news that were reported in the German press. We monitored how the stories evolved over time, which tweets had the most relevance and which users were most active in the discussion. We picked two cases: the Käßmann incident and the Von Sahringen story.

The Käßmann incident involves a speech made by Margot Kässmann, a Lutheran theologian, during the evangelical church day in May 2017. During the speech, she compared the program of the AfD — which pursues increasing the quota of German children without an immigrant background — with the Nazi ideologies. Her quote was shortened to make it incorrectly appear as if she had said that all Germans with two German parents and four German grandparents were Nazis. This generated fury on social media among AfD supporters.

Figure 3.5 shows how the story evolved during the end of May. The black area under the main trend corresponds to the tweets that not only referenced the story but also included the word Nazi. Most of them express rejection to Käßmann’s shortened quote. These selected tweets are more prominent at the beginning when the story breaks and then again during the last days of the story. In our dataset the five tweets with the most retweets were:

1 RT @SteinbachErika: Wenn ich nicht bereits aus der Partei namens EKD ausgetreten wäre, nach den Aussagen von Frau Käßmann wäre das jetzt fällig
2 RT @FraukePetry: #Käßmann blamiert sich mit ihren Aussagen auf ganzer Linie - Geschichtsrevisionismus in Reinform. #AfD
3 RT @uebermedien: Eine infame und verleumderische Kampagne gegen Margot Käßmann. @AfD_Bund @kirchentag_de
4 RT @SteinbachErika: Hervorragend! Imad Karim an Margot Käßmann: "Diesen ’Nazi’, der in Ihrem Kopf geistert, habe ich NIE getroffen"
5 RT @Joerg_Meuthen: Guten Morgen! Kann man allen Ernstes nüchternen Kopfes einen solchen Unsinn verzapfen? #Käßmann #AfD

Only one of the five tweets is debunking the story, while the rest are reacting against Käßmann’s statement. In our data, 4,590 users were tweeting or retweeting the story. The users that commented the most were:

1 e_pitzky
2 Rumsucher
3 PeterPa34083139
4 krippmarie
5 mrstone0856

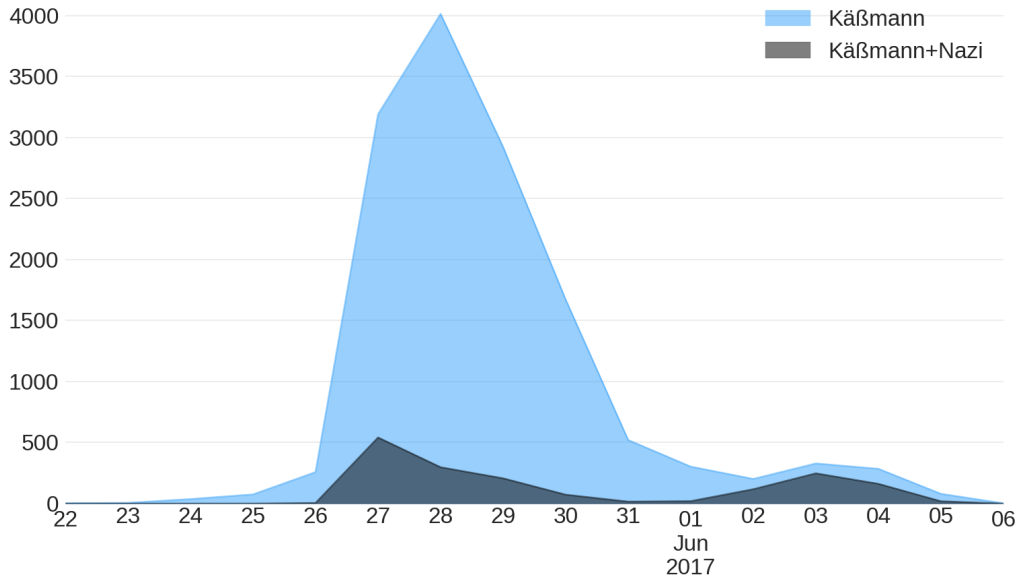

Figure 3.5: *Development of Käßmann's fake news story: Tweets that only included the word Käßmann and tweets that also included the word Nazi.*

All of them sent over 50 tweets related to the case. By looking at their profile, they seem to be AfD supporters. We cannot completely verify whether these accounts are automated accounts or not. However, it is not common that a normal user would send over 50 tweets about one topic. These accounts definitely had the purpose of making the information go viral.

The second story that we followed is about a vote-rigging claim days before the German election. It started with a tweet from the account *@von_Sahringen* that stated: "I was called to be an election helper. On Sunday the votes for the AfD will be made invalid". Hours later, the account for Germany's official election bureau responded saying that this act was prohibited. By then, the hashtag #wahlbetrug ("election fraud") had started spreading on Twitter. The hourly trends of tweets mentioning Von Sahringen and Wahlbetrug can be seen in Figure 3.6. The largest peak for Von Sahringen occurred after the post from the election bureau. After this peak, the discussion on *Wahlbetrug* began to take off. The trend continued until the election, even though several media sites[1] had already reported that the account was a fake account with a modified picture of a Pakistani actress.

By looking at the most active users in the conversation, we find that bots were part of spreading the election fraud narrative. From the top five accounts mentioning @Von_Sahringen, two were closed by Twitter: A_Flicklgruber and ExilFury. Two other accounts changed their identity: EmperorFawful, a male user, became Julia Bathory, a female user; and Ouando_MdB, a supposedly AfD representative in the German parliament became KasimirQY, a shower curtain salesman in Estonia. This is typical behavior of bots that change their identity (description, image, location) to serve new purposes. On the other hand, the top five users that tweeted about the election fraud were all closed by Twitter. We assume that since the case was extensively reported on news media, Twitter acted against the accounts that were spreading the story. The next subsection continues the discussion on bots and their impact

[1] https://medium.com/dfrlab/electionwatch-final-hours-fake-news-hype-in-germany-cc9b8157cfb8

on election-related events.

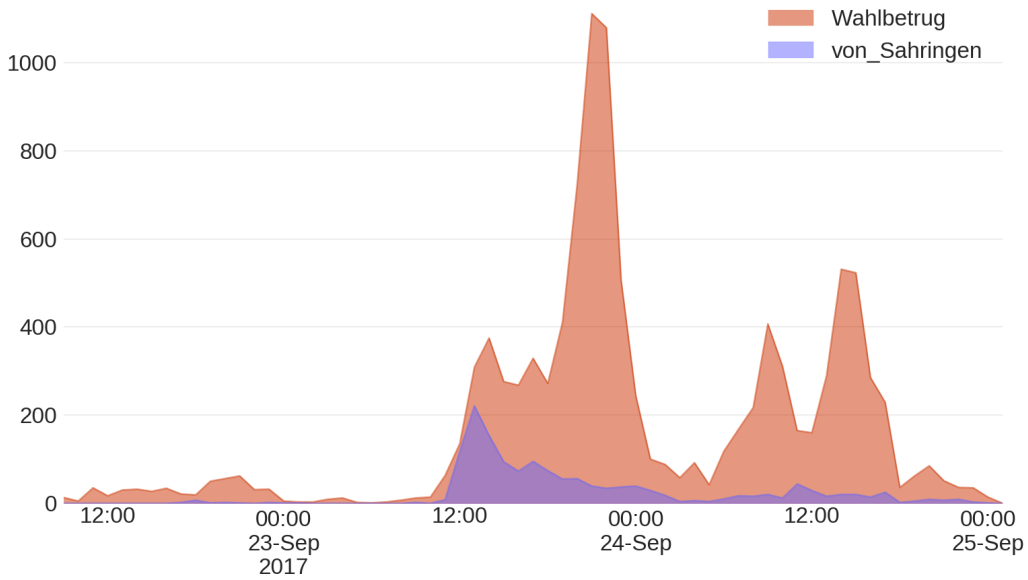

Figure 3.6: *Development of the Wahlbetrug story: Tweets that mentioned the fake Twitter account and tweets that included the word Wahlbetrug.*

Social Bots

Social bots are automated accounts that try to emulate human behavior to influence normal users. They produce content and interact with humans on social media [7]. It has not been proven, whether social bots have been successful in the past in influencing people, but they are known to have manipulated trends in social networks. This could alter the perception of what is happening online and then influence the platform´s algorithms into showing the falsified content to more users.

During the US election, it was estimated that around 19% of the tweets connected to election topics were created by bots [2]. Different media sources were speculating how much influence bots would have in the German federal election[2]. Unfortunately, a percentage number is not completely reliable, since it is unclear which measure from automatization makes an account a bot. Indeed, the term social bot is loosely defined. Additionally, the existing detecting methods only identify accounts with certain predefined characteristics and bots are constantly evolving to avoid getting "caught" (account closed by the platform). In order to have a better understanding of social bots in Germany, our task was to use a similar study as conducted in [2], while also using other detection methods for comparison.

We used three approaches to identify bots in our dataset: two heuristic-based methods and one machine learning method. Heuristic methods are based on rules deciding whether a tweet comes from an automated account or not. The rules depend on specific properties

[2]http://www.sciencemag.org/news/2017/09/social-media-bots-tried-influence-us-election-germany-may-be-next

from a tweet. For example, a tweet that originates from a verified account is directly treated as non-automated. We selected four different properties that have been used in previous works to identify tweets with a bot-like behavior: tweets coming from a suspicious source, tweets that are text duplicates and not retweets, tweets from users with excessive amount of tweets per day and tweets from users with a ratio of friends and followers close to 1. The two heuristic approaches differ only in the number of rules that have to be true to categorize a tweet. A simple approach includes all tweets that fall into any of the four categories. A second, stricter approach involves categorizing tweets as bots when they comply with at least two of the heuristics.

On the other hand, machine learning methods automatically find patterns with the help of hand-coded data, called training data. These methods are useful to make predictions on data that are similar to the training data. For the machine learning approach, we used the open source tool Botometer, which was developed at the Indiana University [22] and has been used extensively in the literature, including the aforementioned US study. This tool categorizes Twitter accounts depending on 1,150 features that are used in a machine learning model. Botometer gives a score between 0 and 1 per user and all of the users with a score larger than 0.5 are considered to be bots. Botometer specializes in English tweets and thus it is not as accurate for the German language. For our analysis, accounts that were explicitly closed by Twitter cannot be analyzed by Botomer and they are directly considered as bots. Concrete explanations regarding the three selected methods of bot detection are included in the Appendix.

For the analysis on bots, we first focused on the tweets that had mentioned the AfD during the complete period of observation. As already mentioned before, during the month of September the discussion on AfD increased in comparison to the main candidates (Figure 3.4). For this reason, we decided to use this sample of data and try to understand whether this effect was caused by bots. We selected only the German tweets since the AfD effect is only present in the German tweets. The dataset comprises 2,747,193 tweets from 195,779 users.

The results of the three methods are shown in Figure 3.7[3]. Each method gives a different percentage average of bot accounts. This originates from the fact that social bots are not a well-defined category and each method detects different kinds of suspicious accounts. The simple heuristic approach gives an average of 14.98 percent, the strict heuristic 1.5 percent and with the Botometer tool — including closed accounts — 9.9 percent. The results of the simple heuristic and Botometer are similar to previous research on bots, where around 10-15 percent of the conversation on Twitter is attributed to automated accounts.

More insights can be obtained from observing the percentage of bots over time. This percentage is based on the total number of tweets, which means that the 21% of the simple heuristic on the first of April corresponds to 1,037 tweets and the 13% on the first of September to 3,235 tweets. For the three methods, the percentage remains constant throughout the first months and then declines around June. During the month of September, it further declines in both heuristic methods. For the Botometer approach, the percentage remains under 10 percent for most of the time during the last two months. The five peaks above 10 percent are

[3] We had previously shown in [9] the Botometer results without including the closed accounts.

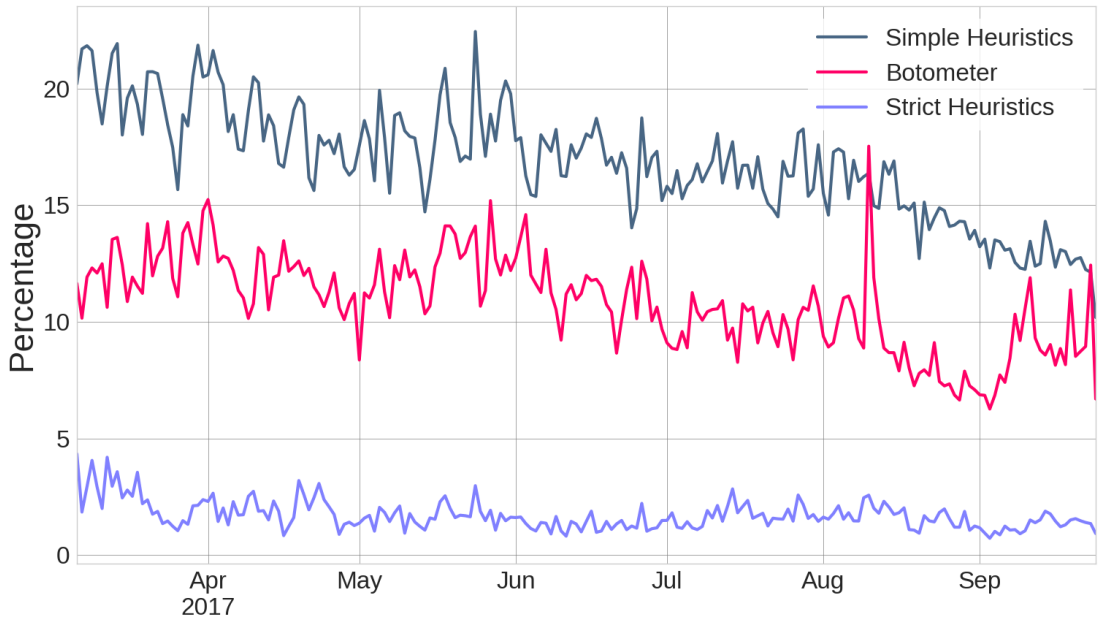

Figure 3.7: *Percentage of tweets from bots in the collection period according to three methodologies. The evaluated sample comprises all the German tweets that included the word AfD.*

attributed to now-closed accounts. Even though there were overall more suspicious accounts tweeting about the AfD during the month of September, their activity was insufficient to strongly influence the overall conversation. By contrast, during the month of September, the normal user accounts generated more content that included the word AfD than bots. Accordingly, we conclude that automated accounts were not the main cause of the increase in conversation on the AfD.

We decided to analyze other datasets and compare with the AfD results to have a better measure of bot effectiveness. There were two specific cases in the election period where the media explicitly reported on bot attacks. The first one[4] was during the TV debate on September 3 between Merkel and Martin Schulz, the candidates from CDU and the SPD, respectively. During the debate, the hashtag #verräterduell ("traitors' duel") became prominent in the Twitter conversation, although it did not make it to Twitter's trends. Nevertheless, it was considered an attack from far-right groups, as shown by a journalistic investigation[5]. The second case was part of the aforementioned fake news story from the election fraud[6]. After the Von Sahringen story broke, there were many accounts using the hashtag #wahlbetrug ("election fraud"). According to the investigation, a Russian network of bots was responsible for spreading tweets with AfD propaganda that included the election fraud hashtag.

We collected 3,615 tweets that included the hashtag #verräterduell and 8,617 tweets with the hashtag #wahlbetrug from our data. The number of tweets is too small to use the heuristic

[4]https://medium.com/dfrlab/botspot-memes-target-der-spiegel-merkel-678a2fc52b05

[5]https://www.buzzfeed.com/karstenschmehl/willkommen-in-der-welt-von-discord-teil1?utm_term=.laYM2EkLKz&bftwdenews#.fiXJ84ayXY

[6]https://www.focus.de/politik/deutschland/bundestagswahl_2017/bundestagswahl-analyse-hunderte-fake-twitter-profile-verbreiten-beitraege-von-afd-unterstuetzern_id_7631486.html

methods, so we analyze the data only with the Botometer approach, which focuses on users and not on tweets. The results for the different events are presented in Figure 3.8. The case with the most bots is the *Wahlbetrug* with 58 percent of tweets coming from automated accounts. Among this 58 percent, 51 percent of the tweets come from closed accounts by Twitter. This case was very present in German media and happened in the days leading up to the election, which probably motivated Twitter to be effective in closing suspicious accounts. By comparison, 28 percent of tweets belong to bot accounts in the *Verräterduell* case and 11 percent come from already-closed accounts. For the AfD, among the 2,747,193 tweets, 4.5 percent of the 9.9 percent belong to closed accounts. This is much lower than the events known for having bot attacks.

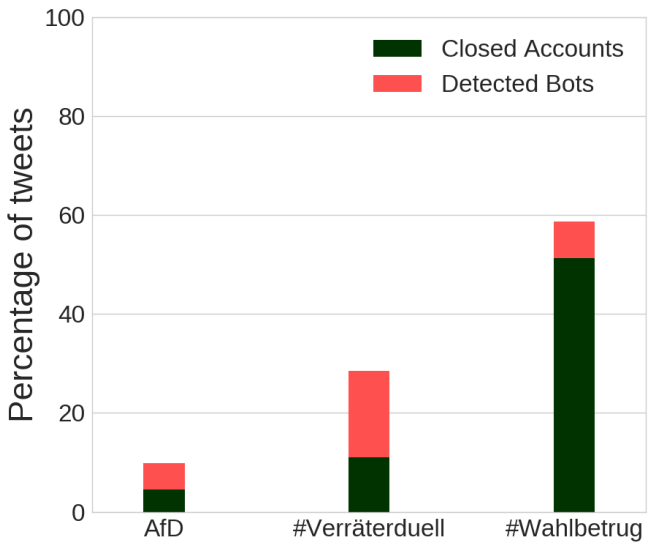

Figure 3.8: *Percentage of tweets from bots for the three evaluated cases. Each case is divided in detected bot accounts by Botometer and closed accounts by Twitter.*

Our analyses confirm that social bots were indeed depleted during the election period as an online manipulation tool. These attacks were able to gain attention from the media, although a real effect on voter intention cannot be quantified from the results. Apart from the coordinated bot attacks, we do not see evidence of social bots having as much effect on the overall political conversation as expected by media experts. However, the phenomenon that we observe is that the right-wing opposition party AfD is dominant on Twitter, and most of the bots we found were working in their favor. Our work is consistent with [16], where 1 million tweets of German political content were collected in a ten-day period.

Foreign Intervention

The last form of online manipulation that we researched was the intervention of foreign actors in the German elections. This is a broad topic that has to be approached carefully since the plausibility of tracing back the origins of online manipulation is extremely limited. For this subsection, we focus on actors that have been speculated to have attempted to influence the elections through manipulation.

The first actor under investigation is Russia, primarily, since it was in the media limelight during the US elections. The US Congress continuously investigated the allegation of Russian efforts to meddle in their elections and evaluated whether the usage of Russian propaganda had an effect on the outcome. In Germany, it was also speculated whether Russia could have had plans to target Germany[7].

The first way to evaluate intervention is to focus on previously-disclosed efforts. In October 2017, Twitter released a list of 2,752 Twitter accounts that the company identified as being connected to the Russian Internet Research Agency[8]. The identified Russian trolls were analyzed in [23] and [1]. The latter study collected tweets related to the US election for two months and subsequently found 221 Russian trolls in the data. However, until now it has not been shown whether these accounts were also active in topics outside the US elections. We proceeded to explore whether the already-deactivated accounts appeared in our Twitter dataset. Surprisingly, we found 23,595 tweets from 458 of the accounts tied to the IRA. As explained in the collection methodology, the tweets need to have a connection to a topic in German politics. From the 458 accounts, 98 tweeted in German and generated 13,932 German tweets from the total of 23,595 tweets. This means that the identified Russian trolls were also trying to spread information to the German users.

The activity over time of the accounts separated in German and non-German tweets (Figure 3.9) shows that before May most of the accounts were tweeting in languages other than German, which changes at the beginning of May. In May and June most of the tweets are German and in July and August there is a similar quantity of German and non-German tweets. In September, the month of the election, almost all the tweets are German, which corresponds with the idea of the troll accounts trying to influence Twitter during the election period.

The contents of the tweets offer more insights into how these accounts operated. Interestingly, 11,730 of the German tweets included an URL, which means that the majority of the tweets (~84%) had the intended purpose of amplifying information from other sources. The five media sources that were shared the most by the accounts were the Dresdner Neueste Nachrichten, Bild, Die Welt, Tagesspiegel and the Berliner Zeitung. We find that most of the tweets include links to what can be referred to as mainstream media. From more than 1,000 shared media links, only ninteen refer to sputnik.com, nine to Russia Today's German portal and two to sputnik.de. The supposedly Russian trolls were only amplifying messages from normal media sites, which differs from their activity in the US discourse where sites like Breitbart.com and thegatewaypundit.com had considerable diffusion [1].

The influence of the troll accounts can be traced back by evaluating retweets. From the total tweets, language independent, 11,571 were original tweets and they originated 92,043 retweets in our database from 46,198 users. Only 37 bot accounts retweeted contents between each other. As previously mentioned, the collected data is just a sample from the complete Twitter traffic and it is biased towards political topics. Nevertheless, a replication factor of 8 in our sample suggests that the troll accounts were not isolated in Twitter and their activity was

[7] http://www.spiegel.de/netzwelt/web/bundestagswahl-2017-debatte-um-moegliche-manipulationen-durch-russland-a-1165520.html

[8] https://www.recode.net/2017/11/2/16598312/russia-twitter-trump-twitter-deactivated-handle-list

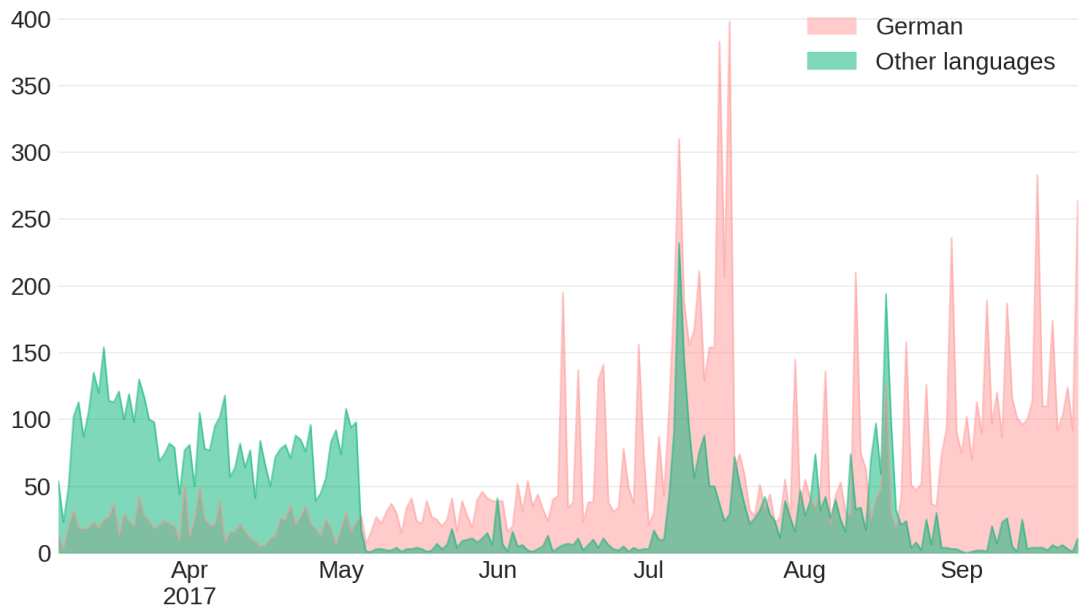

Figure 3.9: Number of tweets from accounts belonging to the list of detected Russian trolls. The tweets are separated between German and non-German tweets.

noticed by other Twitter accounts. The top retweeted accounts were media accounts, i.e die Welt, Tagesschau, Stuttgarter Nachrichten, Tagesspiegel and Spiegel.

We further analyzed the 13,932 German tweets. In the texts, the top five hashtags were Merkel, AfD, CDU, SPD and G20. The first four are included in the list of hashtags that we are following to collect tweets and the fifth hashtag corresponds to the high activity during the days of the 2017 G20 conference (Figure 3.9). Regarding the topics discussed in the tweets, we cleaned the data by eliminating stop words, smileys and URLs and created a word cloud (Figure 3.10) of the most discussed topics. The word Merkel is the most common one, followed by AfD. Also interesting are the words, flüchtlinge (refugees), polizei (police) and the hashtag merkelserfolge (Merkel's successes), which was mainly used in a negative, sarcastic tone on Twitter.

Figure 3.10: Word cloud of the tweets from the detected Russian trolls.

A similar analysis on Russian propaganda is conducted with live data by the

Alliance for Securing Democracy. This agency published an online tool[9] that follows 500 Twitter accounts related to Russian media outlets, propaganda accounts, and probable bot accounts. The list of accounts was created by looking at the followers of the German account for Sputnik News *@de_sputnik*. The complete methodology for the account selection and bot detection is found on the website. The tool, called *Artikel 38*, comprises a dashboard that follows the information being spread by the selected Twitter accounts. The dashboard includes the top hashtags, topics and URLs, similar to the presented results on the Russian trolls. Since the complete list of followed accounts is undisclosed, we are unable to make a direct comparison with our data. However, an analysis of the users who retweeted the main Russian outlets in Germany is possible.

We compared the number of retweets between the two main Russian outlets in Germany, Sputnik and RT, and two main German media outlets, die Welt and Spiegel (Figure 3.11). For the Russian media, we included the main account and the German account. Among the total retweets, one-quarter of them belong to the Russian media. This means that on Twitter the Russian outlets' spread is significant and comparable to the mainstream media accounts. The possibility of influence relies on how these media outlays portray the information. At the London School of Economics, a renowned working group found out that in the German-language offerings of Russian media, the AfD is "presented in a positive light, while other German officials and institutions are shown negatively".

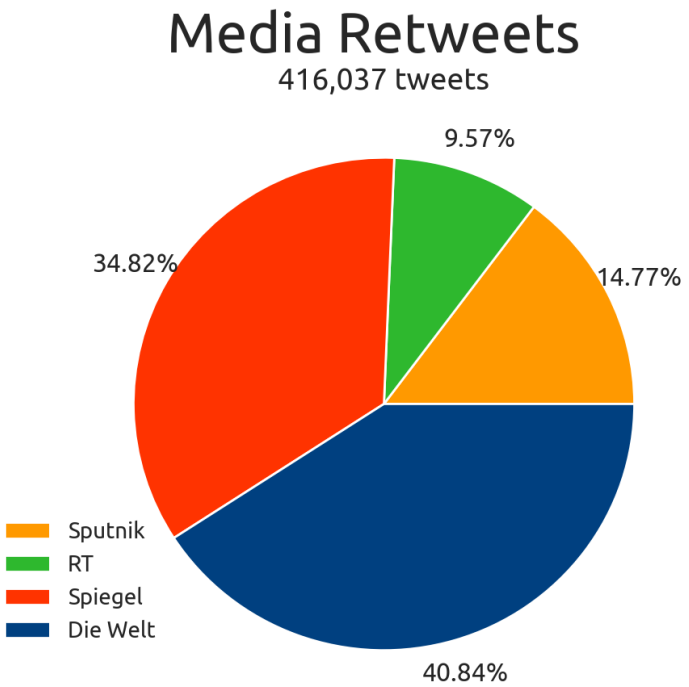

Figure 3.11: *Retweets from four different media channels, including two Russian media outlets.*

Apart from Russia, we considered a group that partly shaped the 2016 US election's discourse, namely the extreme right. As shown in [6], it influenced the polarization of the US media, where the online media power from the right went from the mainstream conservative channels like Fox News to an even more conservative side that mainly included Breitbart. This

[9] http://dashboard-german.securingdemocracy.org/

internet site often supported the alt-right (short for alternative right), which is a group that encompasses extreme ideas from the right-wing spectrum. The alt-right is known for anti-Semitism, anti-Islam and neo-Nazi ideologies. Given that this group had an influence on the US elections, we explored whether this was also the case for Germany. During the election period, we observed an increase of alt-right connections in the German Twitter conversation, especially through a newly-founded social media platform, Gab.

Gab (https://gab.ai/) is a social media platform created in August 2016. It has a similar interface to Twitter, whereby users can post 300-character messages — called "gabs"— which appear in a vertically-scrolling timeline format. At the same time, the platform provides the option to upvote or downvote each post, as in Reddit. Gab promotes itself as a supporter of free speech. Nevertheless, it has been used by the alt-right and white supremacists to spread their ideas since hate speech and harassment are allowed. Each Gab account can be optionally linked to a Twitter account for cross-platform posting and these are the accounts that we can find in our data.

We looked in our dataset for tweets from users who had Gab in their description or their profile URL and for tweets that were posted directly from Gab (cross-platform posting). We filtered only German tweets, which we presume to have originated from German users. Overall, we analyzed 568,118 tweets. At the beginning of 2017, there were few German tweets related to Gab, but starting in June there is a significant increase in activity (Figure 3.12). This increase continues until the election day. This shows that the connection between the alt-right community and German users has increased and that with the help of new online spaces there is a higher chance of communication between the two groups.

We further constructed a word cloud with the tweets with a connection to Gab (Figure 3.13), using the same procedure as before. It portrays some similarities with the word cloud from the Russian trolls. The most commented topics are AfD and Merkel and several terms from right-wing discourse such as immigrants, refugees, terror and Islam. One interesting term is Dora Bromberger, the name of a famous 20th century Jewish German painter. As reported

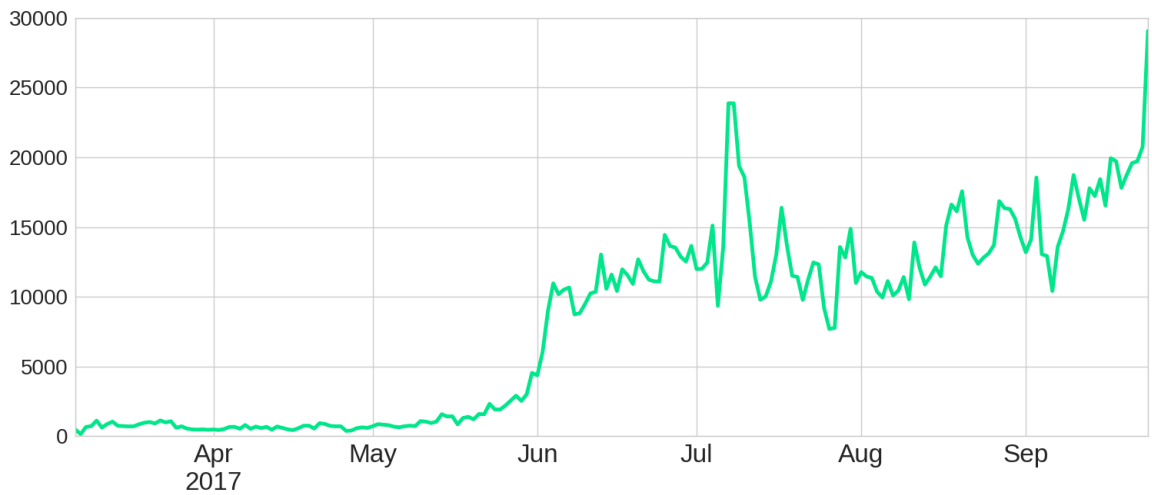

Figure 3.12: *German tweets from Twitter accounts related to Gab.*

in the media[10], a supposedly fake account with this name was tweeting fake statements about the increase of crimes by refugees during the Munich beer festival. This was another minor case of fake news and false accounts. Indeed, the three discussed topics on online manipulation are close to each other and overlap in the methods that are implemented to influence online users.

Figure 3.13: *Worldcloud from the German tweets by the reported Russian trolls.*

[10]http://www.sueddeutsche.de/muenchen/soziale-medien-die-wiesn-bilanz-wird-auf-twitter-ins-fremdenfeindliche-verdreht-1.3696532

4. Political Campaigns on Facebook

Facebook is nowadays the largest social network. In the first quarter of 2018, Facebook reported having 2.2 billion monthly active users. In Germany, this platform is also the most popular one with around 31 million active users. For this reason, the German political parties use Facebook as one of the main channels of online communication. Facebook pages allow politicians and their parties to reach out to possible voters and communicate their agenda.

The analysis of political groups on Facebook is fundamental to understand how digital technologies can contribute to political participation. Already in 2012, a Facebook study in Nature [4] showed that the platform's tools were successful in mobilizing people to vote. Indeed, in recent years politicians have used Facebook to target groups of users who are more likely to support them. There have been recent debates on how the users' data is being used by third-party companies for political manipulation. The 2018 Cambridge Analytica scandal brought Mark Zuckerberg — Facebook's founder and CEO — to testify in front of the US Senate. The data mining firm had used private data illegally to profile voters and target them.

For this study, we only retrieved publicly-available data. Overall, we collected 37,152 posts from the German political parties' Facebook pages. The period of the posts comprises one year. We obtained the data from the main party pages and the regional pages. The data includes all comments, comments on comments, reactions and shares. For this purpose, we used the Facebook Graph API. We obtained the data before the changes introduced in the API in February 2018. With the new version of the API, it is no longer supported to obtain information about which users reacted to a post.

In contrast to the last chapter, we are unable to conduct an analysis of online manipulation in Facebook. The data for such analysis is private data. What users share and what they post to their friends remains private. Moreover, we have no access to the online ads that were shown in the users' feeds. Despite observing strange behavior on the political party pages (e.g. users who commented on or liked more than 1,000 posts), we cannot obtain the data to assess whether it is part of a manipulation campaign. For this reason, this chapter focuses instead on the political campaigns carried out on Facebook by the German parties. Additionally, we analyzed the political interaction between users and the posts from the party pages.

The Facebook Political Constellation

The Facebook political sphere is shown in Figure 4.1. The methods that we implemented are similar as in the Twitter case of Figure 3.1. The main difference is that instead of retweets, the measure for analysis is the number of likes from users in the posts of each political party. We selected this measure because the amount of support to the party can be quantified. In this case, the number of users in the data was more than 100 times larger than in the Twitter data. For this reason, we limited Figure 4.1 to only show the users who have liked more than nine posts from a page. It is more probable that users who have liked many posts are true supporters of the political party and these are the users of our interest.

The larger, colorful nodes on Figure 4.1 represent the political parties and the smaller nodes the Facebook users. The closeness of a user to a party node is proportional to the number of

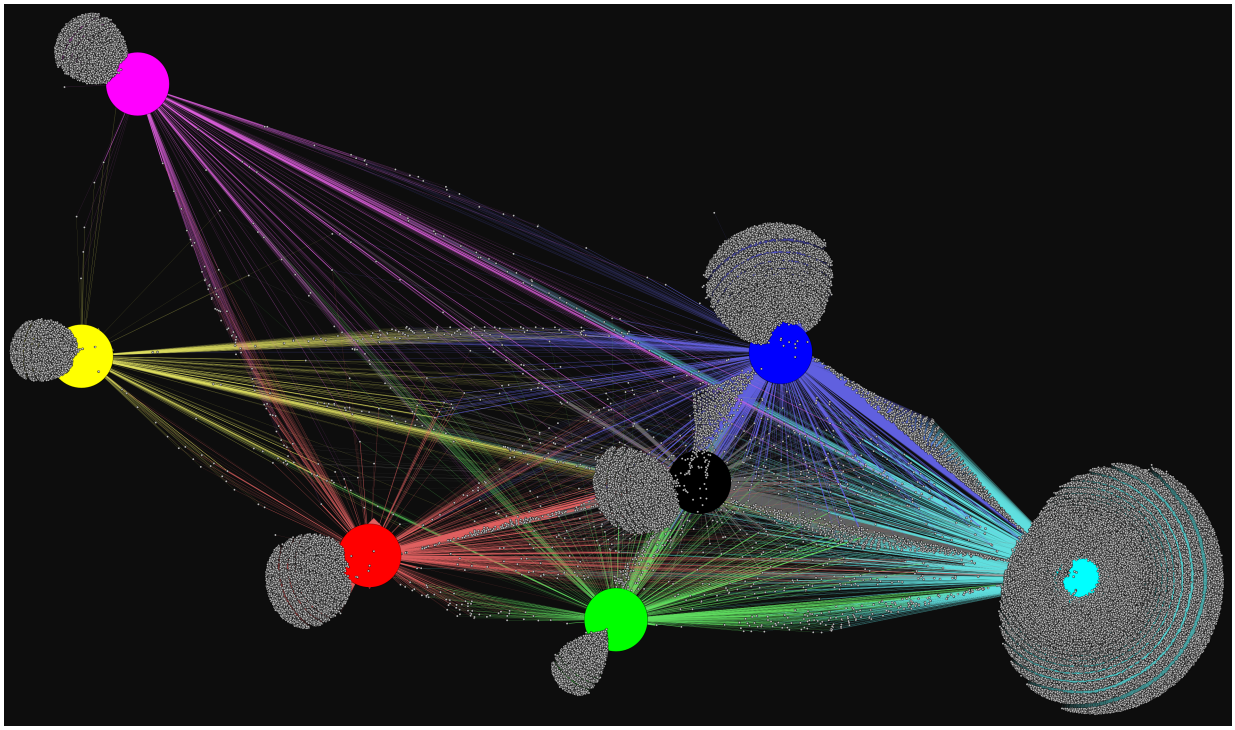

Figure 4.1: *Active users who liked the page of a political party. Only users that liked more than nine posts are included. Each party is represented with its respective color.*

likes on the posts of the party. The users clustered around a party node can be regarded as partisan users, whereas those between nodes as users who support more than one political party. The bubble of partisan fans around the AfD is larger than for the rest of the parties. Additionally, the AfD shares many active users with the CDU and CSU. As expected, these two sister parties also share many common users. Active users on more than one political party could be targeted by political parties to convince them to vote for them. This is a technique of political microtargeting, a topic that is explored in [18].

The main difference between the Facebook and Twitter spheres is that the FDP and die Grünen have their places interchanged. In Figure 4.1, the FDP and Die Linke are close to each other but they share almost no users in common. This means that for these two parties the closeness is just a visualization artifact. Even though Facebook likes and Twitter retweets cannot be directly compared, they are reliable measures of party affiliation. The visual representations of both social media platforms show a similar pattern that parallels Germany's political landscape. The next subsections further explore the interactions of the users with the Facebook political pages.

Online Political Campaigns

The seven political parties in Germany used Facebook constantly for the 2017 elections. The online campaign on Facebook centralized the social media presence of the parties. All seven main political parties had similar activity (Figure 4.2). Both Die Grünen and Die Linke had the lowest number of posts, whereas the CSU posted the most during the one-year period. Nevertheless, the number of posts between the parties was similar. The post activity of the political parties over time is shown in Figure 4.3. The activity during the month of the elections significantly increased, which corresponds to similar campaign strategies between parties.

The contents of the posts can be generally understood by looking at the most commonly-discussed topics. For this reason, we implemented a topic modeling algorithm[1] on the posts. The algorithm creates 65 topics[2] from all of the posts and then assigns each post to one of the topics. The top three topics per party are shown in Table 4.1. Each topic contains five keywords and the percentage of documents that belong to it. The keywords shown belong among the top twenty words per topic and are those that in our consideration explained the topic better.

For each party, there is one topic that relates to the party's candidate or candidates, while the other two topics are mostly connected to the main policies of their corresponding campaign. For example, the AfD has a topic of migrants and borders, the CSU has the topic of family and working place and for the FDP the topic related to digitalization and education is on the list. We conclude that the German political parties were using Facebook to push their agenda.

[1] Topic modeling is a natural language processing algorithm that uses a statistical model to find topics in text data. The algorithm we used was LDA. Further information can be found in [3]

[2] The optimal number of topics according to the statistical analysis

Apart from setting the agenda on the main party pages, the German political parties were also active at the state level. Germany is a federal republic comprising sixteen states, for each of which there exists a Facebook page from all of the parties. Figure 4.4 shows the total number of posts sent in the regional pages per party. The total activity of the regional political pages is different from that of the main pages. The AfD was significantly more active in the regional pages, with over 2,000 posts more than the FDP and over 4,000 more than die Grünen.

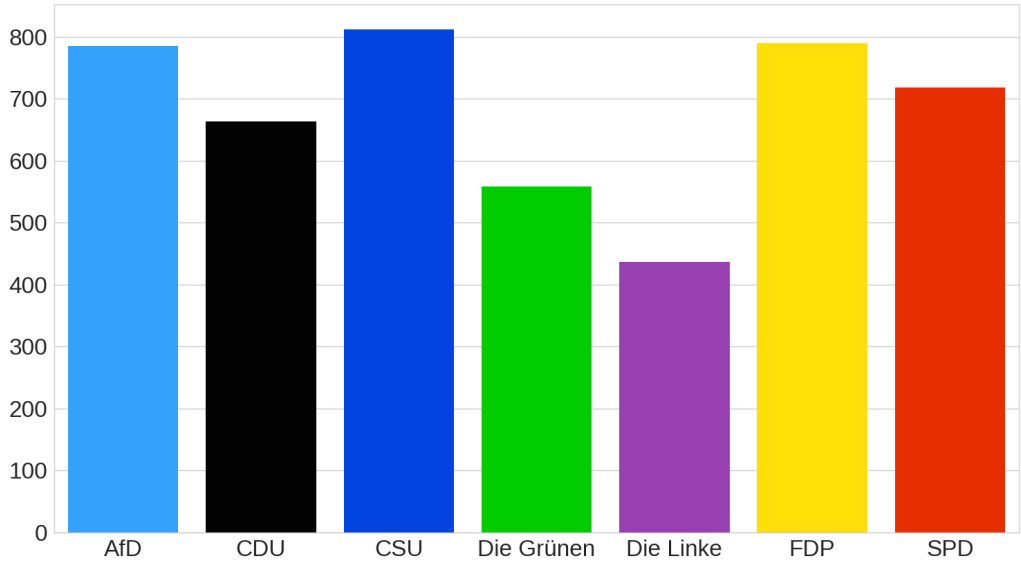

Figure 4.2: *Number of posts sent by each political party during the year prior to the election.*

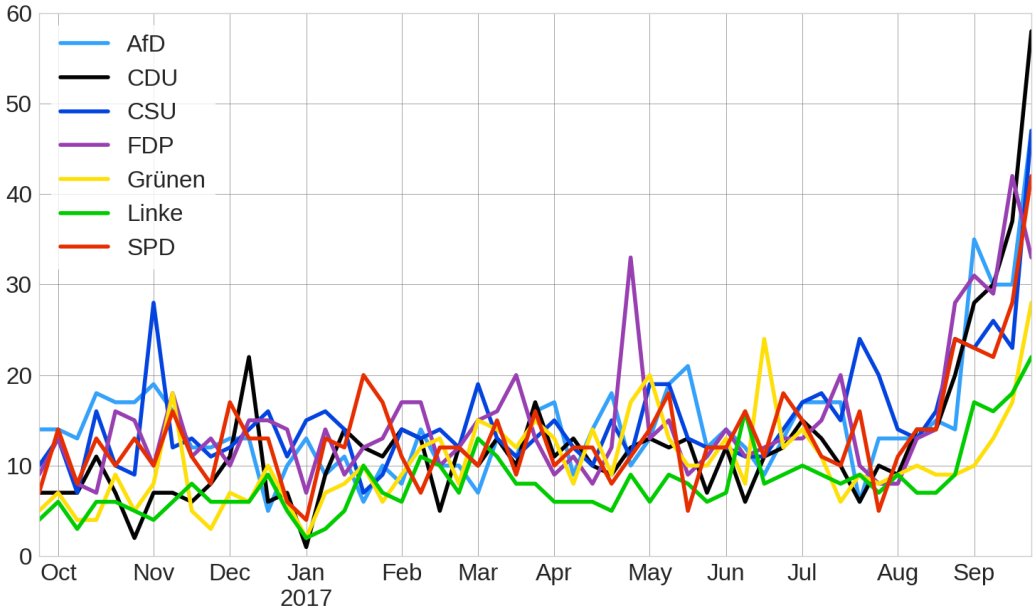

Figure 4.3: *Number of posts per week during the year prior to the election.*

AfD		
#1 (12%)	**#2** (12%)	**#3** (10%)
Weidel	Deutschland	AfD
AfD	Asylbewerber	SPD
Petri	Migranten	CDU
Gauland	Grenzen	Deutschland
Storch	Abschiebung	Altparteien
CDU		
#1 (21%)	**#2** (17%)	**#3** (14%)
CDU	Tauber	Merkel
Union	Generalsekretär	Angela
Regierungsprogram	Digitalisierung	Kanzlerin
Sicherheit	Themen	unterstützten
Wahlkampf	Pressekonferenz	Vertrauen
CSU		
#1 (16%)	**#2** (15%)	**#3** (12%)
Klartext	Löwenstark	Einschalten
Seehofer	Bayernplan	fragcsu
Scheuer	Familien	ARD
Obergrenze	Arbeitsmarkt	ZDF
Grundsatzprogram	Heimat	Live
Die Linke		
#1 (35%)	**#2** (7%)	**#3** (6%)
Linke	sozial	Pflege
Berlin	zusammen	fordern
Wagenkecht	Wahlkampf	gerecht
Gerechtigkeit	unterstützen	Gesellschaft
gemeinsam	Kraft	Reichen
FDP		
#1 (29%)	**#2** (20%)	**#3** (8%)
Lindner	Digitalisierung	Freiheit
Christian	Bildung	Rechtstaat
FDP	Chancen	finanziert
Steuern	Bürger	Deutschland
Entlastung	selbstbestimmt	Wohlstand
Die Grünen		
#1 (40%)	**#2** (19%)	**#3** (5%)
Grüne	Live	Rente
Klimaschutz	Fragen	Arbeit
Umwelt	Özdemir	Euro
Trump	Görig-Eckardt	Altersarmut
Bundesregierung	Wahlkampf	steigen
SPD		
#1 (24%)	**#2** (13%)	**#3** (9%)
Schulz	Europa	AfD
Martin	Demokratie	gegenhalten
SPD	gemeinsam	StimmefürVernunft
Gerechtigkeit	kämpfen	Deutschland
Bildung	Frieden	Zukunftskongress

Table 4.1: *Top three topics of the party posts for each party. Each topic includes five keywords from the top twenty words and the percentage of documents that belong to that topic.*

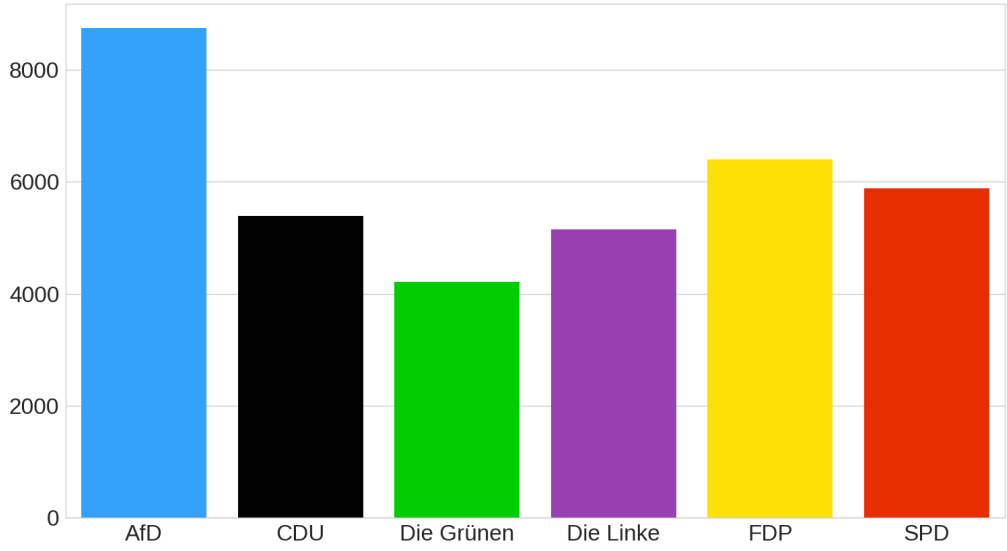

Figure 4.4: *Total number of posts in the 16 regional Facebook pages per party during the year prior to the election.*

User Interactions

The users on Facebook interact with the pages of the political parties in three consecutive ways: first, by liking the page, which makes the user a fan of the page; second, by having the political posts appear in their timelines; and third, by reacting and commenting on the posts. Figure 4.5 displays the number of fans per party on October 2017, the month after the elections. The AfD has the page with the largest number of fans, almost double the number of fans of the CDU and the SPD page. This contrasts with the number of followers on Twitter, where AfD has the lowest number (Figure 3.2).

However, in order to interact with a Facebook page, it is not necessary to be a fan of the page. The interactions of users with the political parties can be quantified by looking at the number

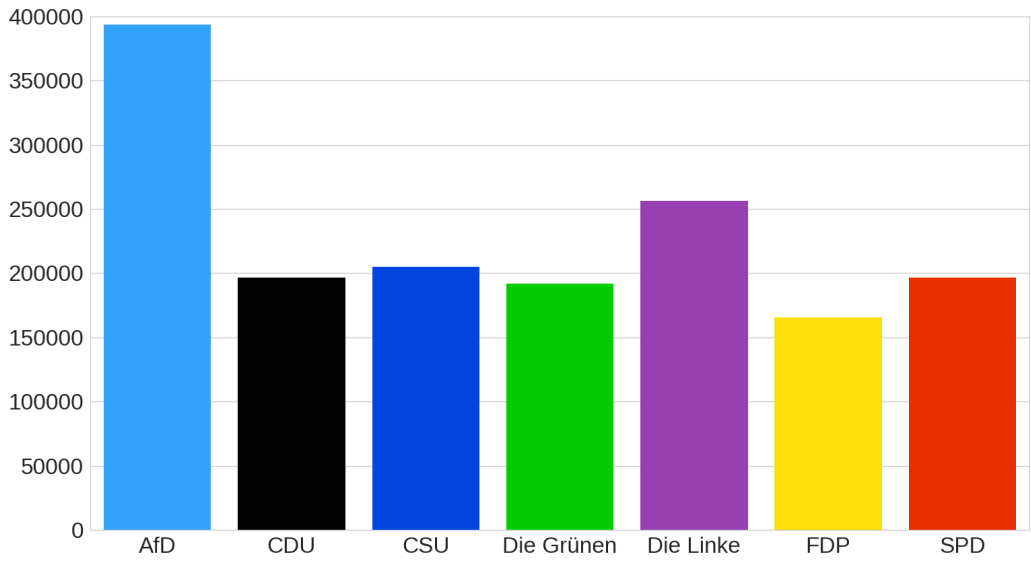

Figure 4.5: *Number of Facebook fans for each German political party as of October 2017.*

of users who have liked or commented on the posts (upper left plot of Figure 4.6). The AfD is the party with the most active users. It has four times the activity of the CDU, whose page has the least activity. Remarkably, the CSU, which is only represented in the state of Bavaria, occupies second place.

We further looked at three other measures of party reachability: number of likes, number of comments and number of shares per page. The results are shown in the other three plots of Figure 4.6. For these measures, the dominance of the AfD is clear. As with the number of followers, the CSU also occupies second place in terms of likes and comments. The content of the comments can be positive or negative, and therefore the number of comments is not a direct measure of support. Nonetheless, the topics discussed by the AfD show more engagement from the users.

For the number of shares, the distance between the AfD and the rest of the parties is even larger. The right-wing party obtained more than five times the number of shares. By sharing a post, a user makes the content available to his friends. Therefore, more shares translate into a larger spread of a political party's message. We consider this to be the most powerful measure for party reachability.

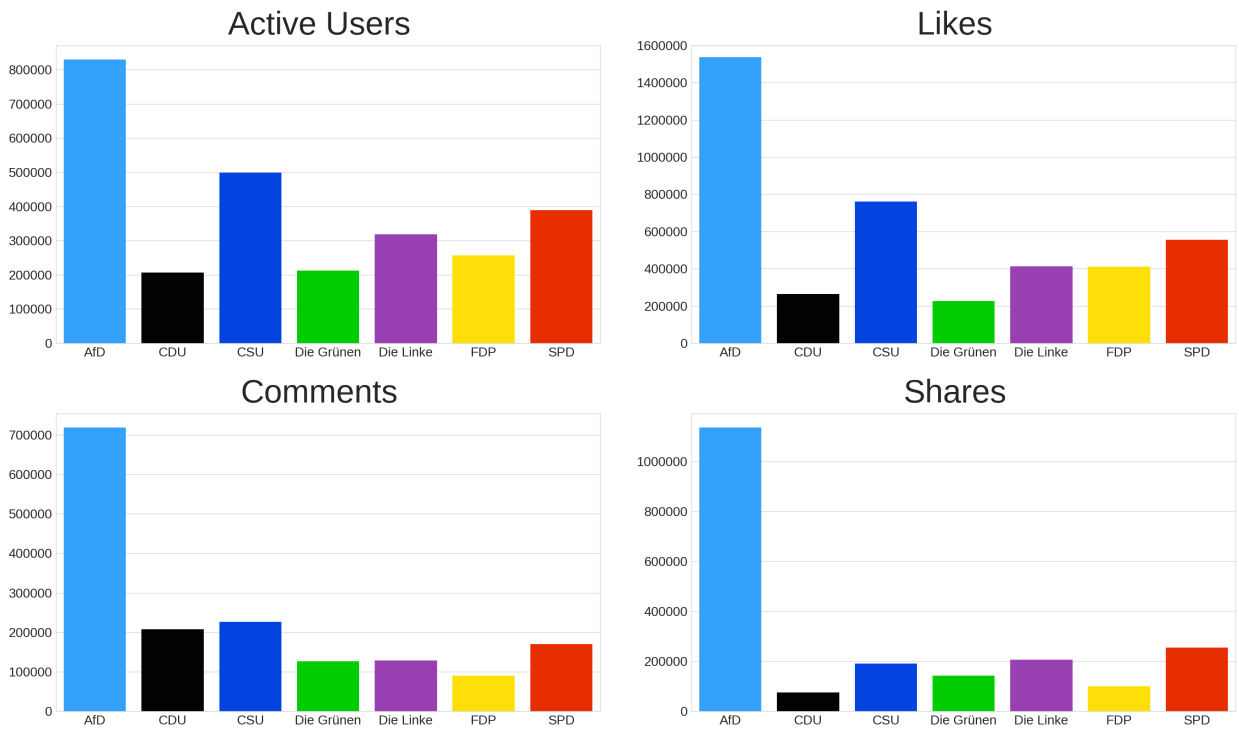

Figure 4.6: *Party reachability measures from September 2016 to September 2017. Upper left: Number of unique users that have either commented or liked a post in the Facebook pages. Upper right: Number of likes per party. Bottom left: Number of comments, including comments on comments. Bottom right: Number of times that the party posts were shared.*

Figure 4.7 illustrates the behavior of users and likes for the AfD and the CDU. The number of users who made a given number of likes is represented on a logarithmic scale. While most of the users liked a few posts, there were accounts that made more than 400 likes during the

one-year period. These users on the right side of the plot are hyperactive users. Their activity could be part of a strategy to make the Facebook page more popular by trying to influence the algorithm that decides what to show to other users. Our analyses show that hyperactive users are present on all political party pages.

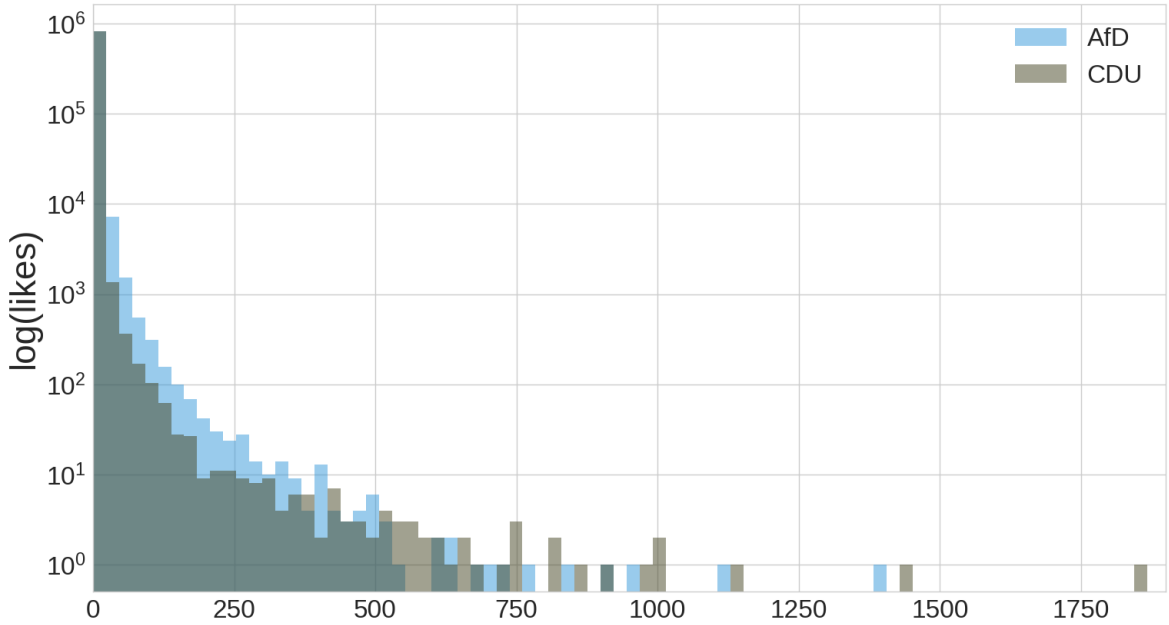

Figure 4.7: *Distribution of the number of likes per user for the AfD and CDU. The logarithmic y-scale counts the number of users who had a given number of likes during one year.*

Apart from likes, there are other five smiley reactions that a user can use: *love*, *haha*, *wow*, *sad* and *angry*. These reactions can shed a light on how the users perceive the political parties. The results are presented in Figure 4.8. When interpreting reactions, it is important to take into consideration two aspects: first, an angry reaction does not portray anger towards the party, but rather a reaction to a post that made the user angry, e.g. the post could contain news from other political parties that would make the supporters enraged; and second, a *haha* reaction could mean that the user finds the post funny, but it could also demonstrate sarcasm.

Die Grünen, Die Linke, SPD and FDP have *love* as the most common reaction, whereas for the AfD and the CSU there are fewer love reactions than the rest. For the CDU and CSU, the most commonly-used reaction is *haha*. Given that the content of both of these pages is serious, the funny reaction could be mostly related to sarcasm. For the AfD, the angry reaction overtakes the rest. This corresponds to the characteristics of a populist party that tries to promote the resentment of people against the "establishment". For all of the parties, the *wow* and *sad* reactions are used the least, making them less meaningful than the other ones for the analysis. Understanding how people react to the parties' online activity is relevant for designing successful online political campaigns. This analysis shows that the German users react in a different manner to the political parties and the communication strategies should thus be adapted correspondingly.

Figure 4.8: *Number of reactions to the Facebook posts for the German political parties. Likes are excluded. The SPD (not shown in the figure) has a similar distribution to the FDP.*

As the final part of our investigation of the main party pages, we evaluated the comments made by the users. We applied the same methodology as with the post, whereby the narrative between party posts and user comments can be compared. Table 4.2 includes the top three topics per party of the users' comments. The first topic of the AfD is different from the others since it includes only three words. Most of these comments included either simply the word "AfD" or a combination of "wählen"(to vote) and "AfD".

In contrast to Table 4.1, there are topics that belong to the top three topics of more than one party. For instance, the second topic of the AfD, CDU and CSU is the same and it includes comments related to migrants coming to Europe. This topic on migration also appears as die Grünen's third top topic, albeit with only 5% of the user comments belonging to it. The debate on the migrant crisis was one of the main discussions during the election period. Interestingly, in the posts, the topic of migrants is only relevant for the AfD.

AfD		
#1 (31%)	**#2** (27%)	**#3** (4%)
AfD alternative wählen	Deutschland Europa Flüchtling Islam Demokratie	Deutschland Volk Kinder Kultur Grundgesetz
CDU		
#1 (21%)	**#2** (17%)	**#3** (14%)
CDU CSU Union Wahlkampf Obergrenze	Deutschland Europa Flüchtling Islam Demokratie	Merkel Erfolg BRD freuen herzlich
CSU		
#1 (16%)	**#2** (15%)	**#3** (12%)
CDU CSU Union Wahlkampf Obergrenze	Deutschland Europa Flüchtling Islam Demokratie	reden versprechen handeln Klartext umsetzen
Die Linke		
#1 (35%)	**#2** (7%)	**#3** (6%)
Linke DDR Wagenknecht Antifa Freien	Partei wählen schaffen Mitglied Unterstützung	Geld Steuern Wirtschaft Arbeitsnehmer Industrie
FDP		
#1 (29%)	**#2** (20%)	**#3** (8%)
Deutschland Volk Kinder Kultur Grundgesetz	Geld Steuern Wirtschaft Arbeitsnehmer Industrie	Glückwunsch Ergebnisse Einsatz verdient Sieg
Die Grünen		
#1 (40%)	**#2** (19%)	**#3** (5%)
Grünen Özdemir Umwelt Autoindustrie Energie	Partei wählen schaffen Mitglied Unterstützung	Deutschland Europa Flüchtling Islam Demokratie
SPD		
#1 (24%)	**#2** (13%)	**#3** (9%)
SPD Linken Koalition Programm Katastrophe	Martin Kanzler Demokratisch GroKo Mehrheit	Schulz Wissen Bundeskanzler Brüssel Lösung

Table 4.2: *Top three topics from the user comments. Each topic includes five keywords from the top ten keywords and the percentage of documents that belong to that topic.*

Regarding the user interactions in regional party pages, we limit ourselves to reporting the analysis on the AfD pages. The four heat maps of Germany in Figure 4.9 present the results. The upper right map illustrates the election results for the AfD in each state. The range of colors span from white to red. The red states represent the places where the percentage of voters choosing AfD was larger. The states that previously belonged to the German Democratic Republic are those where the AfD received most of the votes. The other three maps in Figure 4.9 show the number of followers, the number of comments on the posts and the total number of shares. We normalized these metrics to make a comparison to the election results. For this purpose, we divided each measure by the number of residents in each state.

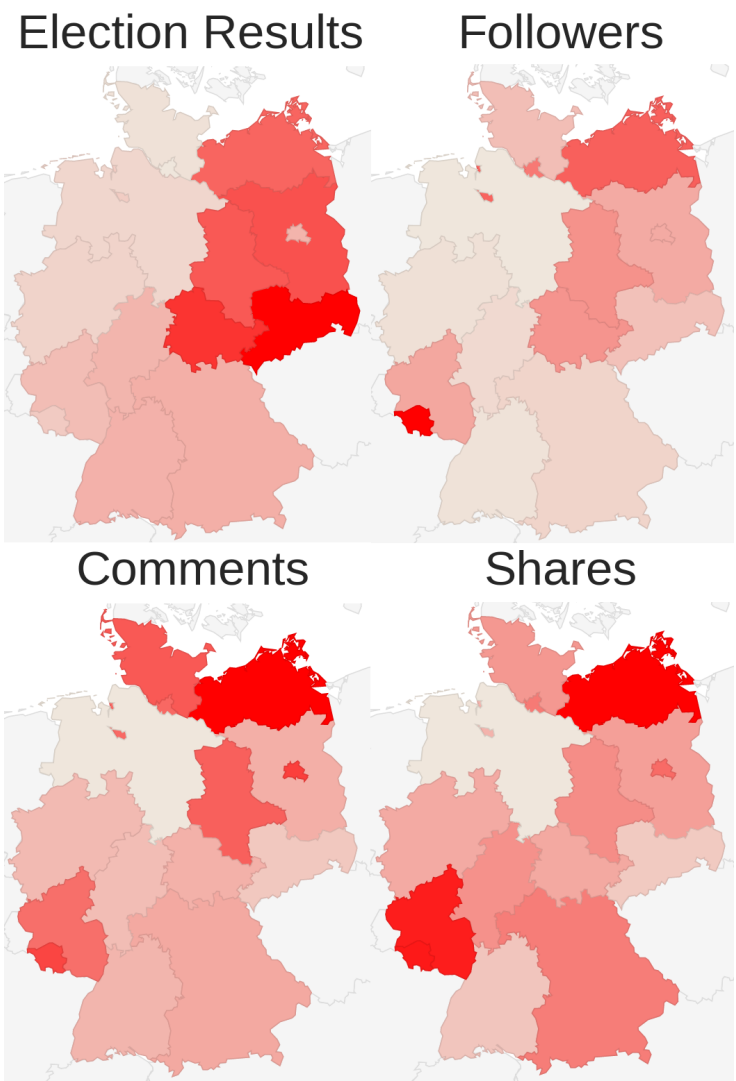

Figure 4.9: *Election results for the AfD in each German state and the Facebook measurements for the AfD's regional pages. The color scale spans from white to red, with red indicating the highest values.*

The states that had a better voting result have also a higher normalized percentage of Facebook followers. The two exceptions are Saarland and Rhineland-Palatinate where the proportional number of followers is high in comparison to the election results. For these two states, the same pattern is seen in the comments and shares. The heat maps of comments and shares

also show high proportional engagement in Mecklenburg-West Pomerania and Schleswig-Holstein. The results show that higher engagement with the regional pages does not directly translate to better election results.

We can evaluate the importance of regional pages by comparing the number of followers between the main page and the regional pages combined. AfD´s main Facebook page had 393,798 followers just after the election, whereas the regional pages combined followers amounted to 382,816. This is a small difference of only 10,982 followers. Indeed, users interested in political parties obtain information from both main and regional party pages.

5. The Online Media Landscape

After having provided an overview of the two main social media platforms in the German election, the final part of this study focuses on the news media that the users shared in both platforms. This information is valuable since it can offer an overview of how online media is consumed by users with different political views. The applied methodology is similar to that presented in a 2017 Harvard report [6], which shows the US online media panorama in times of the 2016 presidential election.

The Harvard report took as a starting point a sample of tweets mentioning the political candidates. For this study, we decided to start with the tweets that had mentioned a German political party. In German politics, there is a stronger focus on the parties and campaign proposals, rather than only on the candidates. In the data, we collected a total of 5,468,409 tweets that mentioned a political party. These tweets included 1,821,478 URLs. Many of them were shortened URLs, and therefore we extracted the extended URLs. Subsequently, we implemented a cleaning process in which we left out the URLs related to other tweets, Facebook posts, Instagram photos, images on the web, Amazon products, YouTube videos and Google and Bing searches. The final set comprised 900,027 URLs.

The last preprocessing step involved finding similar domains and group them together. For example, the domain *spon.de* and the mobile version domain *m.spiegel.de* both correspond to the media source *Der Spiegel*. In order to observe which media sites were more dominant, we established a ranking system. The ranking was not only based on the number of domain appearances since it would then just show the media shared by hyperactive users. To consider

all users, the ranking that we used gave every user a score of one, which was then divided between the media that they shared. For example, a user that shared 5 articles from *Der Spiegel* and two from *Die Zeit* would give 3/5 weight to the first media site and 2/5 to the second one. This methodology gives every user the same importance and avoids the effects of hyperactive users.

Table 5.1 shows the top ten shared Twitter domains in the sample of tweets mentioning the German political parties. *Die Welt* is the domain that was shared the most. The sixth and seventh place belong to two nationalist media and correspond to the high levels of right-wing conversation in Twitter (see Chapter 3).

With the information on the URLs, we constructed an online media network graph using Gephi (Figure 5.1). The nodes correspond to the online domains that appeared on the tweets

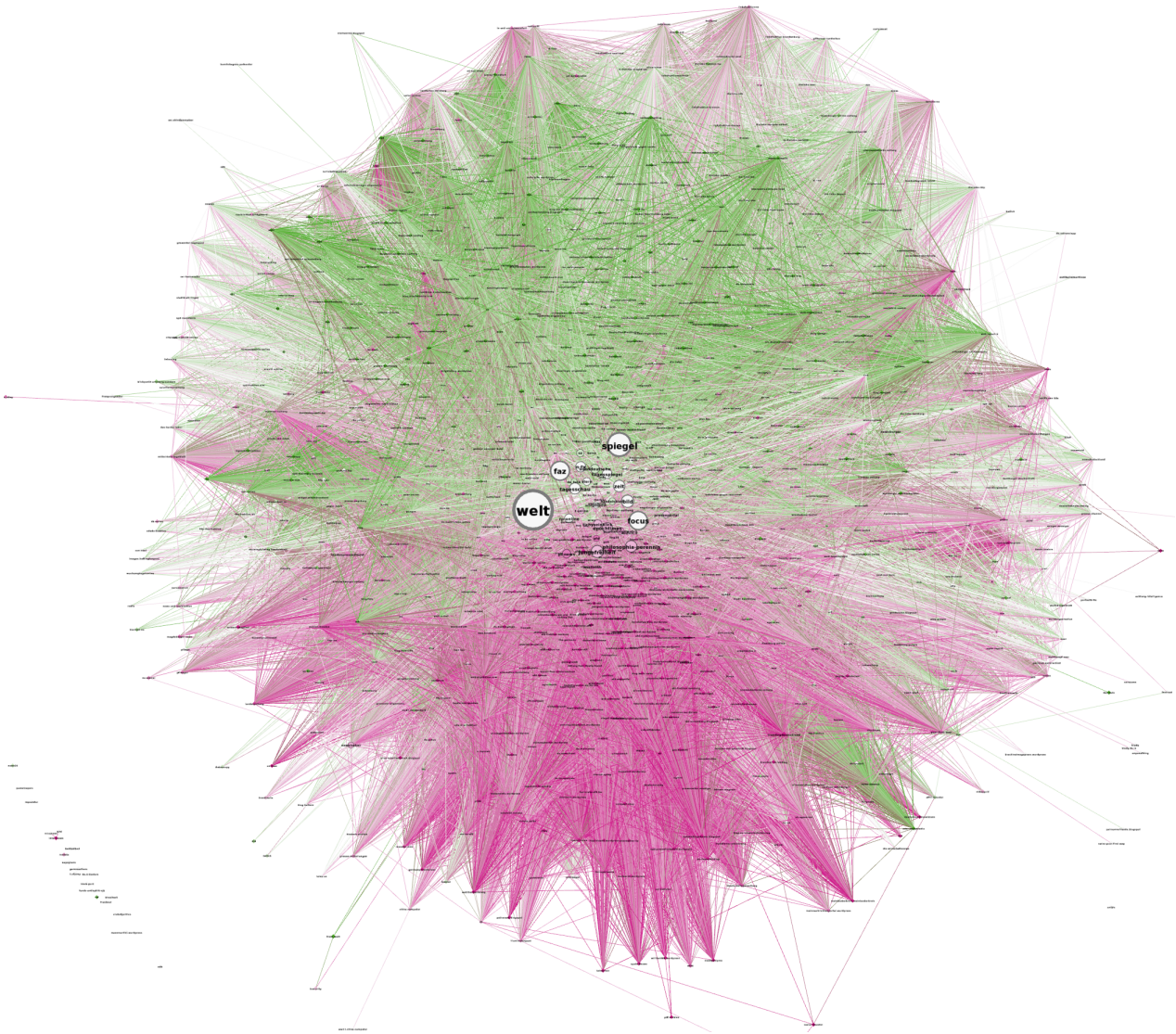

Figure 5.1: *Media network based on the Twitter shares from tweets mentioning political parties. Green corresponds to media mostly shared by supporters of Die Grünen, whereas purple corresponds to media mostly shared by supporters of the AfD*

	Twitter	Facebook
1	welt.de	welt.de
2	spiegel.de	focus.de
3	faz.net	spiegel.de
4	focus.de	bild.de
5	bild.de	zeit.de
6	jungefreiheit.de	der-postillon.com
7	philosophia-perennis.com	fr.de
8	zeit.de	epochtimes.de
9	tagesschau.de	faz.net
10	epochtimes.de	n-tv.de

Table 5.1: *Top ten most commonly-shared media sites per party on Twitter and Facebook.*

mentioning the political parties. The size of the node is based on the number of times that it was shared on Twitter. Only the most commonly-shared domains are visible, as they dominated the Twitter conversation and are the ones included in Table 5.1. The distances between nodes rely on the sharing behavior of the users. Two nodes are connected in case there were users that shared URLs from both media domains. The weight of the edge was calculated as the number of users that shared the domains of both nodes. The Gephi algorithm uses the weight of the edges to calculate closeness. Media domains that were commonly shared by the same users have edges with a larger weight and thus they are close to each other. Accordingly, the position of each media domain and its surroundings give a clear online panorama of how users interact with German online media.

The colors on the network reflect the attention level that the supporters of the AfD and Die Grünen have on media sources. The purple color for AfD and the green for Die Grünen. The white colored nodes represent media that are shared equally by supporters of both political parties. The colors are based on a partisanship score, which relies on the sharing behavior of retweeters from the seven political parties. The score was calculated by first taking all the

	# of retweeters	# of URLs
CDU	29,259	861,274
CSU	10,870	844,251
FDP	22,175	709,429
Die Grünen	33,124	934,567
Die Linke	30,006	935,476
AfD	41,020	1,301,884
SPD	37,513	791,761

Table 5.2: *Number of retweeters per party and the number of URLs extracted from their tweets. All of the information was accessed from the tweets we collected.*

Twitter accounts that had retweeted a German political party. Subsequently, all of the tweets belonging to these accounts were taken from the complete collected tweets and the URLs that had been shared were extracted. Table 5.2 shows how many users retweeted each political party and how many URLs were extracted from the users. The users correspond to the nodes on Figure 3.1 from Chapter 3.

	CDU	CSU	FDP	AfD
1	welt.de	welt.de	welt.de	welt.de
2	cdu.de	spiegel.de	fdp.de	jungefreiheit.de
3	spiegel.de	bild.de	spiegel.de	bild.de
4	faz.net	faz.net	faz.net	focus.de
5	tagesschau.de	focus.de	liberale.de	philosophia-perennis.com
6	bild.de	tagesschau.de	tagesschau.de	tichyseinblick.de
7	focus.de	cdu.de	zeit.de	spiegel.de
8	zeit.de	tichyseinblick.de	bild.de	epochtimes.de
9	sz.de	jungefreiheit.de	sz.de	faz.net
10	tagesspiegel.de	zeit.de	focus.de	deutsch.rt.com

Table 5.3: *Top ten most commonly-shared media sites per party on Twitter.*

	Die Grünen	Die Linke	SPD
1	spiegel.de	spiegel.de	spiegel.de
2	welt.de	welt.de	welt.de
3	tagesschau.de	tagesschau.de	tagesschau.de
4	zeit.de	zeit.de	zeit.de
5	sz.de	taz.de	spd.de
6	tagesspiegel.de	sz.de	sz.de
7	faz.net	die-linke.de	faz.net
8	taz.de	faz.net	spdlink.de
9	gruene.de	tagesspiegel.de	tagesspiegel.de
10	sueddeutsche.de	dasND.de	bild.de

Table 5.4: *Top ten most commonly-shared media sites per party on Twitter.*

A consecutive step to obtain the partisanship score consisted in assigning a ranking system with the information on the URLs per party. We used the same methodology as for ranking media, although this time for the URLs shared by each party's retweeters. The top ten media sources per party are presented in Table 5.3 and Table 5.4. A partisanship score between -1 and 1 was given to each media depending on the ranking position for each party. The score of zero corresponds to media that had a similar position on both parties ranking. In order to construct the network in Figure 5.1, media domains that had a high ranking on the AfD list

and a low ranking on Die Grünen got a score close to -1. If the contrary occurred, the media domain received a score close to 1. This score translates directly to the intensity of the color. For example, *die Welt* holds the first position for the ranking of both political parties. Hence, it appears as a white node, which corresponds to score zero. On the other hand, the *taz* (Die Tageszeitung) is only on the top ten of Die Grünen and has a much lower place in the AfD ranking and thus it appears in color green. Indeed, the taz is regarded as a left-leaning media and is expected to have low sharing from AfD supporters. Another interesting example is the *RT deutsch*, which appears only for the AfD on the top ten shared media and consequently has a strong purple color. These results can be seen more clearly in Figure 5.2, which takes a closer look at the center of the media network.

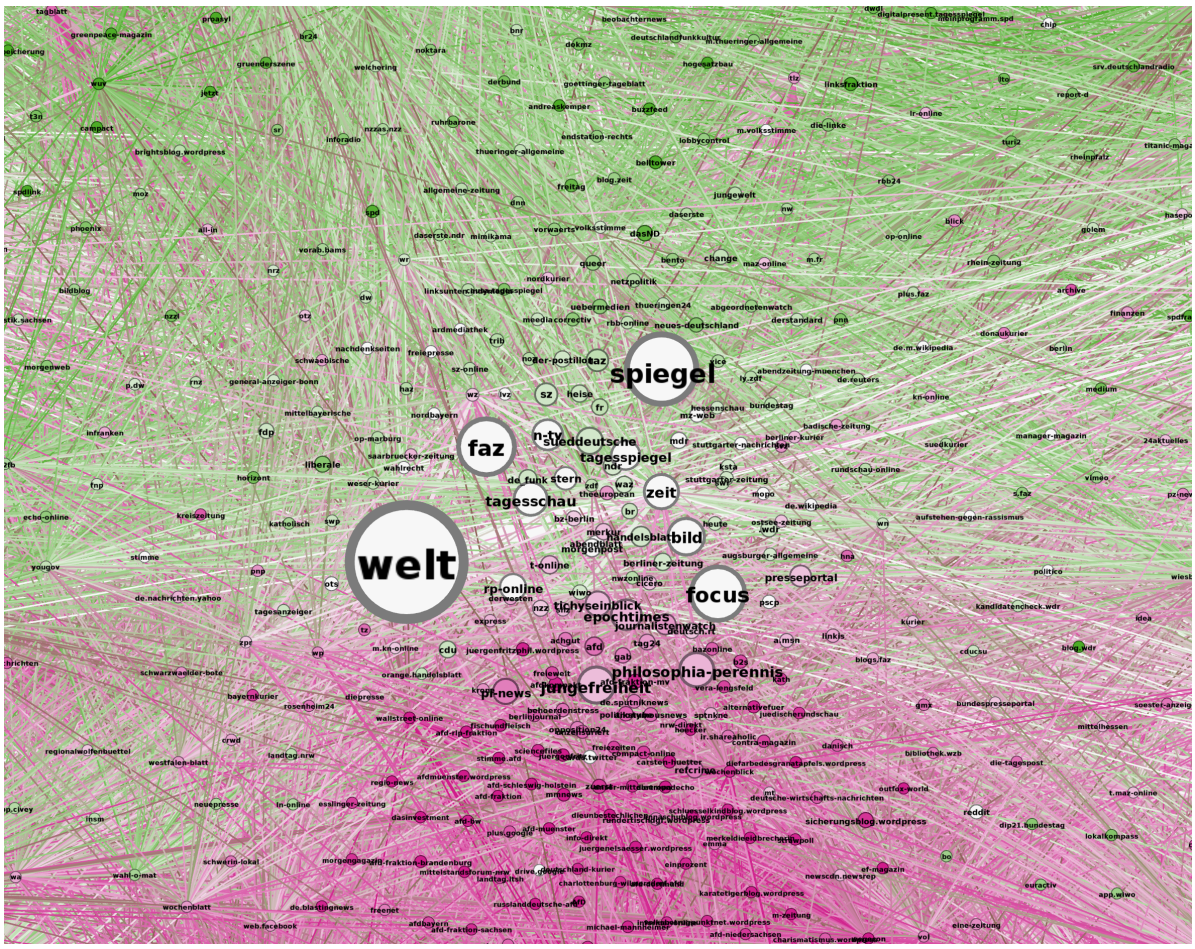

Figure 5.2: *A closer look to the AfD/Grünen media network.*

It is evident that the media ecosystem is divided in two, with the upper part having more attention from supporters of Die Grünen and the lower part from supporters of the AfD. This translates to a polarization in online media between the supporters of the two parties. However, the mainstream media nodes are in the center and are mostly colored white. This means that the users from both political sides still share common ground, where they show interest in general news from mostly non-partisan media sites. This speaks against separated echo chambers[1].

[1]Description of the situation where "[citizens] do not hear the arguments of the opposite side, but are rather surrounded by people and news sources who express only opinions they agree with." [8]

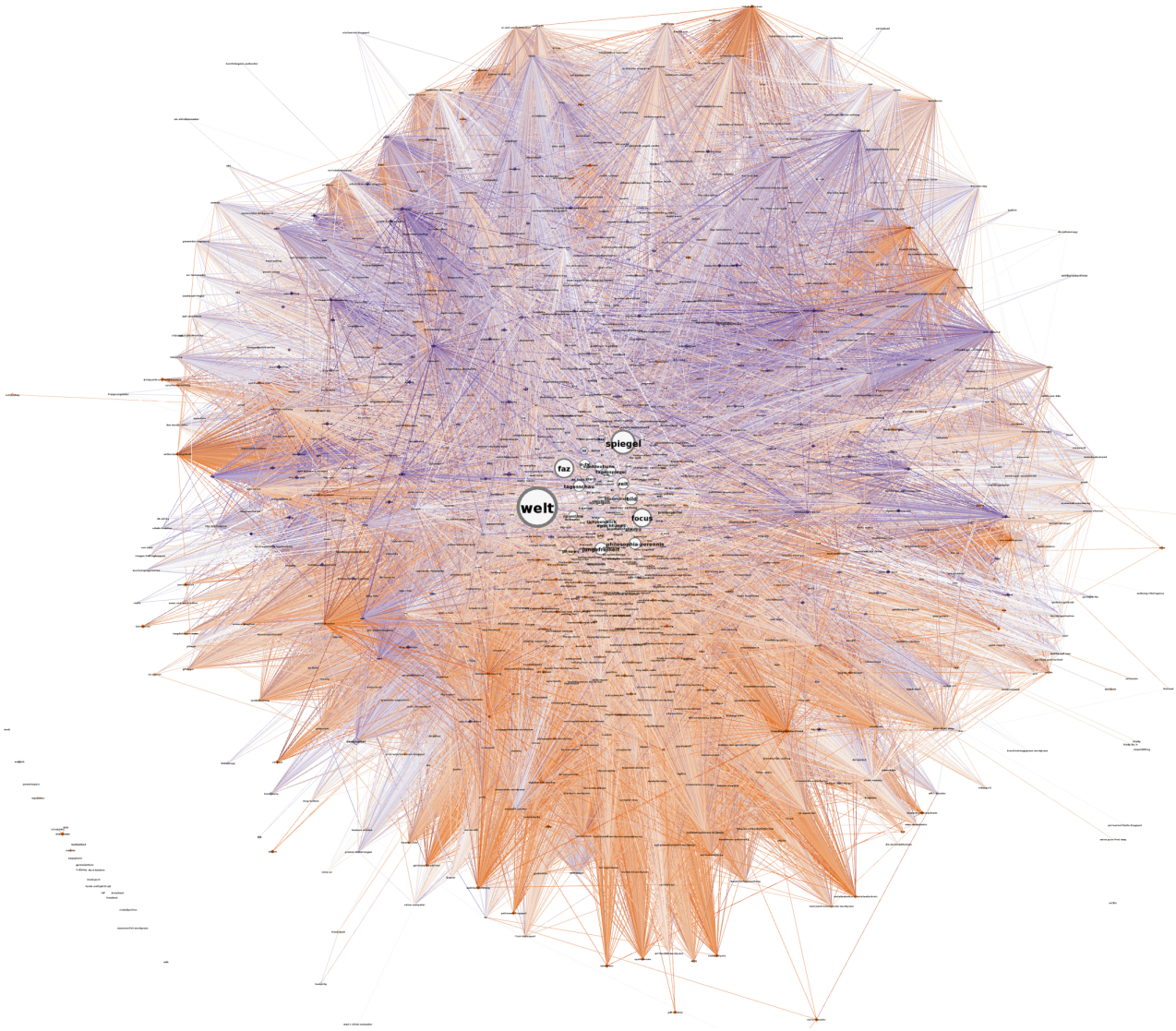

Figure 5.3: *Media network based on the Twitter shares. Purple is used for CSU and orange for the AfD.*

We used the same procedure to compare the media ecosystem between the AfD and CSU (Figure 5.3 and Figure 5.4). The orange represents the AfD and the purple the CSU. This media network is similar to the one with Die Grünen, given that it has the same division that establishes the AfD on the lower half. However, the colors are not as strong and the division is less defined. In this case, most of the main media in the center of the graph are white nodes, many more than in Figure 5.1. Moreover, many of the right-leaning sites appear also in white. This parallels the fact that in the German Twittersphere (Figure 3.1) the CSU and AfD nodes are closer to each other than to Die Grünen. The results are as expected, since the CSU and the AfD belong to the right political specter.

Furthermore, we performed a comparison of the spread of online media on Twitter and Facebook. As previously mentioned, the size of the nodes on Figure 5.1 corresponds to the number of times the media links were shared on Twitter. In order to make the comparison, we adjusted the size of the nodes to the number of times the same media links were shared

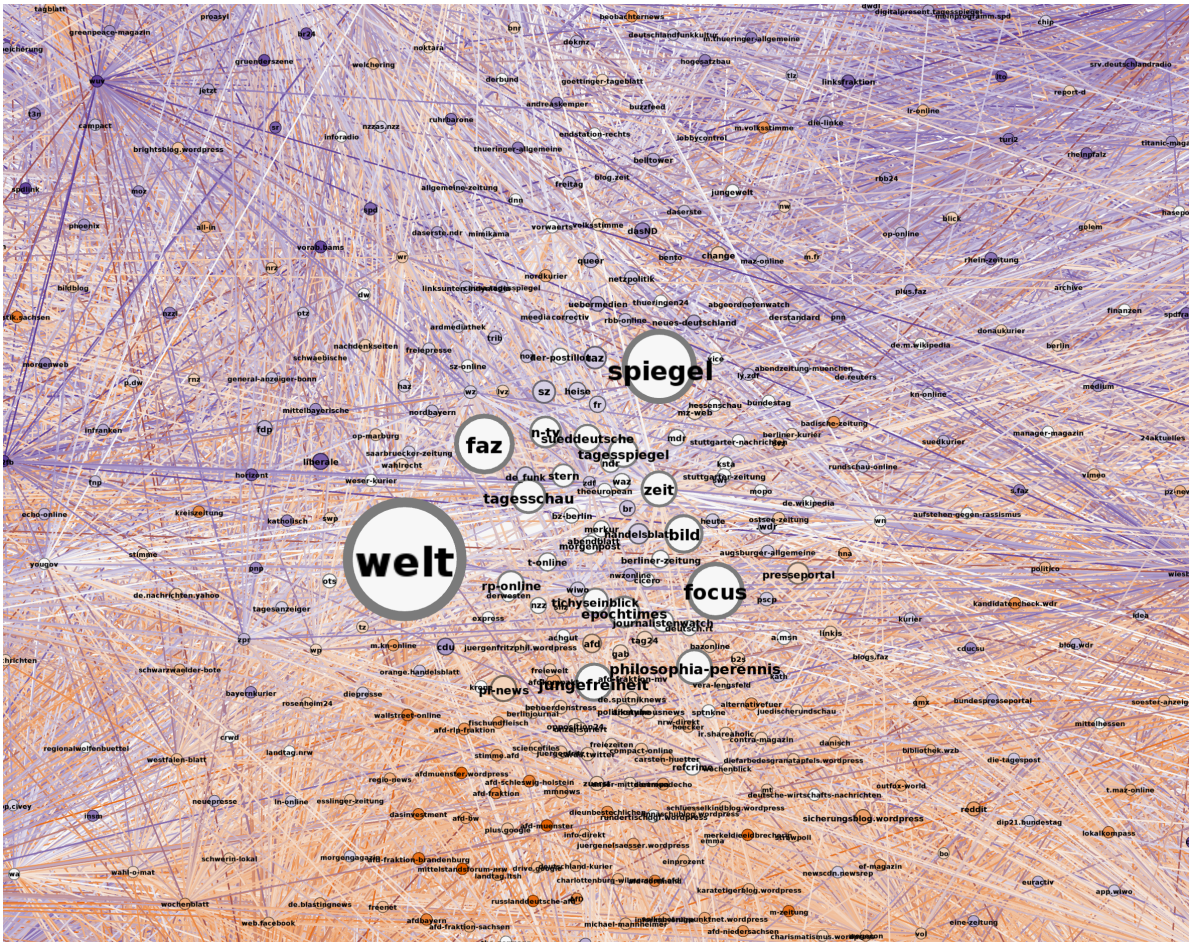

Figure 5.4: A closer look to the AfD/CSU media network.

on Facebook. The Facebook API allows to check the number of times a story was shared on the platform. We checked each of the unique URLs from the 900,027 extracted and cleaned URLs. The stories belonging to the same domain were aggregated and their respective shares were summed up to get a final Facebook share score per media domain. The results for the Facebook shares are shown on Figure 5.5. Only the size of the node is affected by the new results. The position and the colors are still based on the Twitter sharing behavior.

Both Figure 5.2 and Figure 5.5 have nodes with very similar sizes. The largest node is *Die Welt* on both platforms. Some other media have a stronger impact on Facebook than Twitter, like *Focus* and *Der Postillon*, a satirical site similar to *The Onion*. Table 5.1 shows the top 10 media sources shared on Facebook based on the shared URLs with political content. As with the networks, both top ten lists are similar. Accordingly, we conclude that the German online media landscape is similar on both Facebook and Twitter.

A trimmed network with only the top shared media on Facebook is shown in Figure 5.6. In this case, the network corresponds to that of the AfD and CSU. Most of the edges are white since they connect white-colored nodes. The position of the nodes is based only on how the users shared links on Twitter. A Facebook-based graph network taking into account the sharing behavior of users would make a better comparison. Nevertheless, this is not possible to construct, since the user profiles and the media they share are private data.

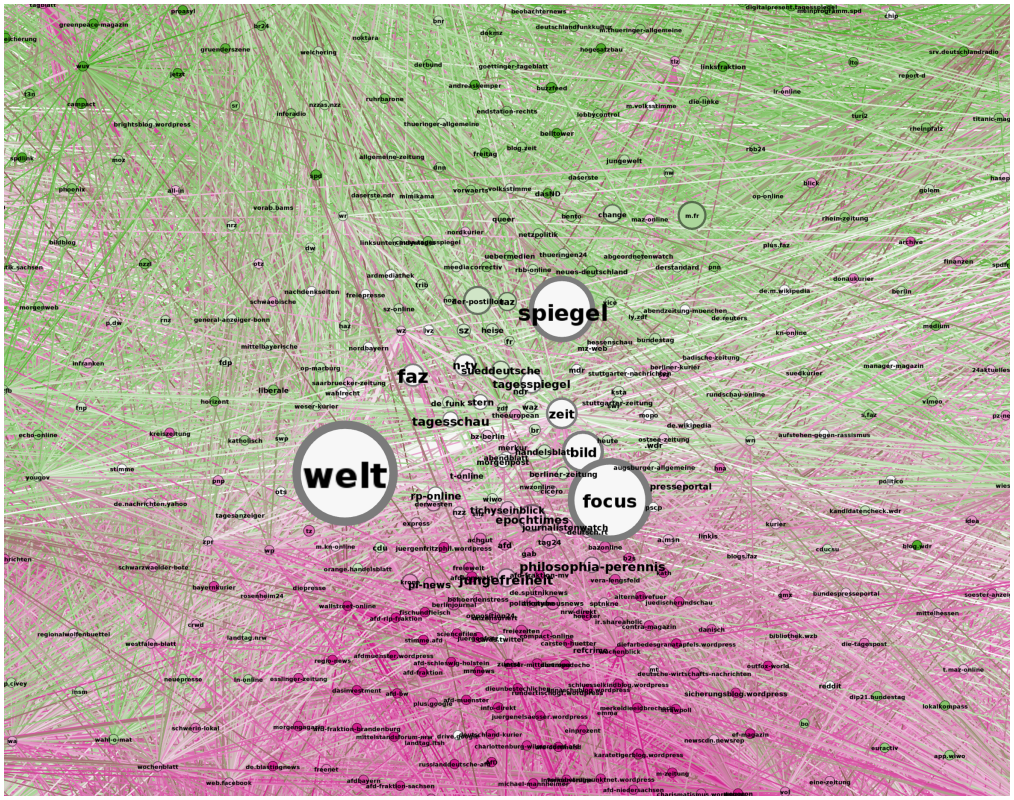

Figure 5.5: The AfD/Grünen media network with modified node sizes based on Facebook shares.

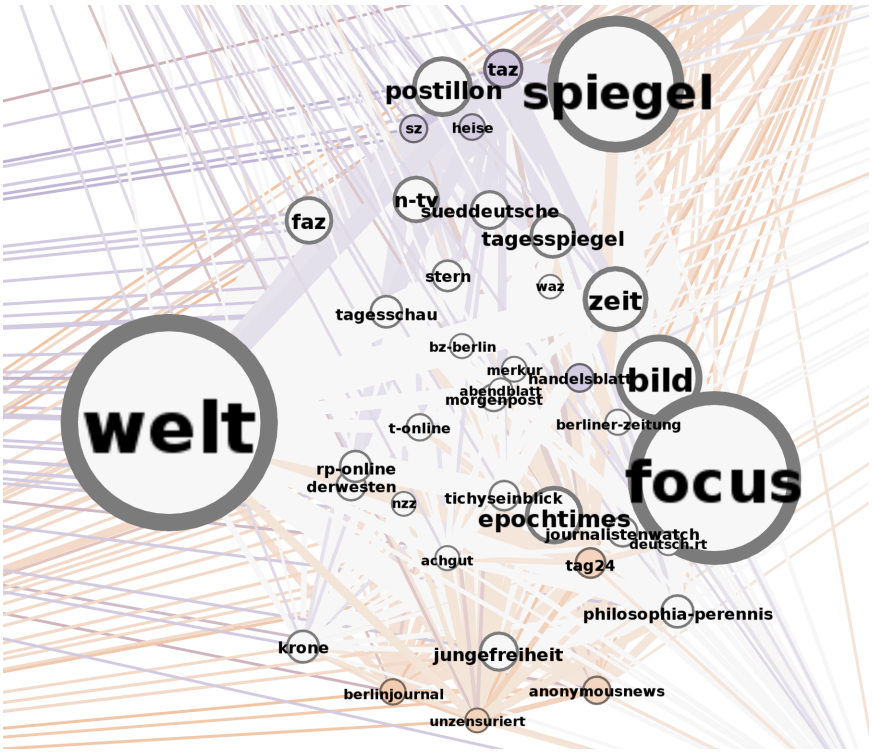

Figure 5.6: The trimmed AfD/CSU media network with only the most commonly-shared media on Facebook. The size of the nodes corresponds to the number of Facebook shares.

Finally, we can make a direct comparison between the German and the US media (Figure 5.7), given that the process to construct the presented graph networks was based on the methods in [6]. In the US there is a clear separation of the right-wing media sphere (colored red), where Breitbart played a central role, and the center-left and left media sphere. The US mainstream media is encompassed in the left side, with the exception of Fox News on the right. By contrast, the German media network contains the main media outlets in the center. Moreover, the right-wing sphere (Die Junge Freiheit, philosophia perennis, Tichys Einblick and others) is close to the mainstream media. These differences relate directly to the degree of polarization and segregation of the population regarding media consumption. Although the polarization is seen in the German media network, it is not as prominent as in the US media landscape.

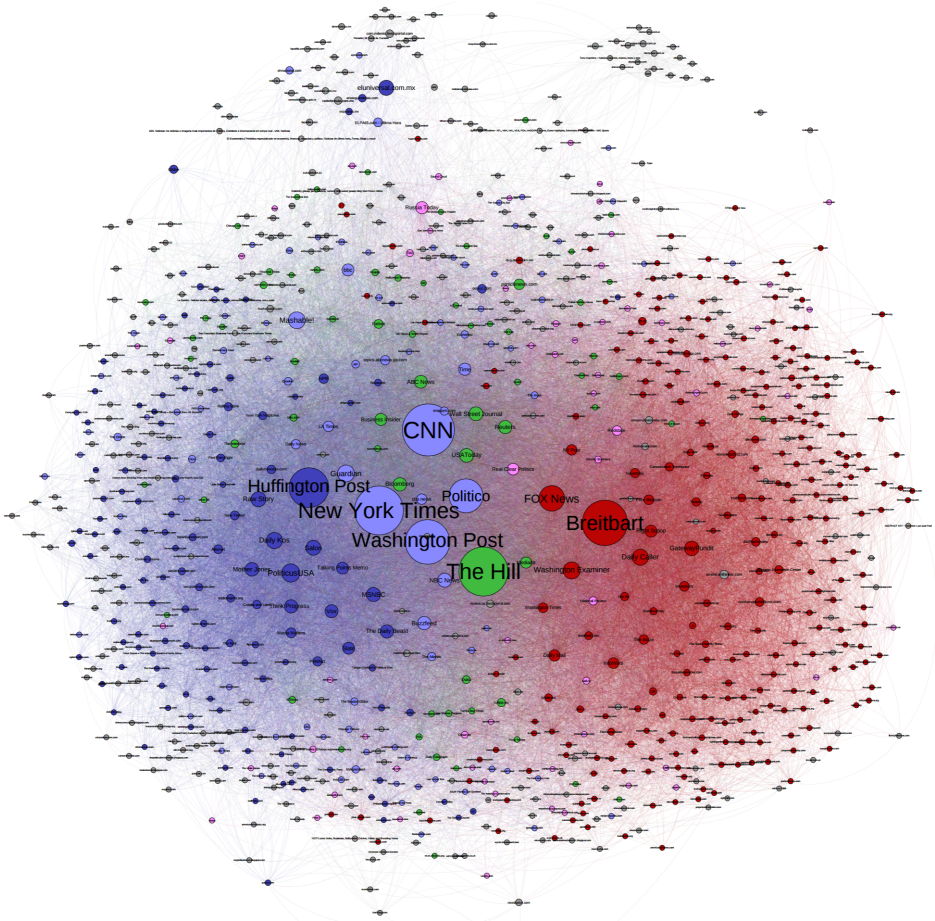

Figure 5.7: *The US media landscape on Twitter during the 2016 US elections. Source: The "Partisanship, Propaganda, and Disinformation: Online Media and the 2016 U.S. Presidential Election" Harvard report [6]*

6. Conclusion

This study has presented a detailed analysis of social media during the months leading up to the 2017 German federal elections. The uniqueness of this work relies on the amount and quality of data that was collected. To the authors' best knowledge, there is no other report on social media and politics in Germany with the amount of data that we tracked. Most analyses found in the literature are based on up to 5 million tweets or consider shorter periods of time.

The outcome of our research can be summed up by four major findings:

The first finding is that the AfD dominated in social media. On both Twitter and Facebook, the right-wing political party managed to spread their message to more users. There is a possibility that part of their success on the 2017 elections relates to these results. Already in 2016, Schelter et al. [20] formulated that "the rise of the AfD can be associated with an amount of social media coverage and user engagement that is unprecedented in the German political landscape".

The second finding is that online manipulation mechanisms existed that targeted the German election process on Twitter. Nevertheless, the observed amount was less than expected by experts. It is difficult to measure the effects that the detected social bots, fake news stories and foreign intervention techniques had on the German public. However, the results are consistent with Neudert et al. [16], which also found that the bots were working in favor of the AfD and with Saengerlaub et al. [19], who presented an analysis on fake news in Germany.

The third finding is that the German public is less prone to being affected by online misinformation than the US public. The closeness of right- and left-wing media in Germany to the mainstream media shows that citizens of different political parties are consuming information from validated sources. We further conclude that false news did not play a major role in the conversation regarding the election. The top shared news on Facebook and Twitter connected to political parties had only a few misleading stories and no completely fabricated news. The news items related to migration were those that had the most misleading facts.

The last finding is that the German political structure is correctly represented in the social media platforms. This is confirmed by looking at the behavior of the Twitter users retweeting the political parties and the Facebook users liking the posts of the pages. Moreover, the online media landscape mirrors the media consumption by users with different political views. The comparison with the US allowed us to understand the differences in polarization and media consumption between the citizens of these two nations.

These findings provide an overview of the dynamics in Germany between social media and politics. Further research could focus on content analysis, sentiment analysis, and network dynamics. We encourage researchers of different disciplines to undertake social media, politics, and digitalization as a combined field of study. Especially since the dangers of political manipulation are still yet to be understood and solutions need to be found. The future is digital and together with the capabilities of social media, the political discourse and political interactions will continue to evolve over time.

Appendix

Bot Detection Mechanisms

Botometer

Botometer is a Twitter bot detection framework from the University of Indiana. The framework is based on a machine learning algorithm trained from tens of thousands of labeled examples [22]. It connects directly with the Twitter API and for a given account, it returns a score from 0 to 1. This score determines how probable it is that the account is an automated account. Accounts with high scores exhibit a more bot-like behavior. For our analysis, we used 0.5 as a threshold to classify bots.

The machine learning algorithm analyses 1,150 features, which are divided into six categories. For our data, we only used the features from four categories. The omitted categories can only be used with English data since they are related to the content, language features, and sentiment features from the account's tweets.

Heuristics Approach

Heuristic approaches are based on rules obtained from theoretical analysis and expert knowledge. The main advantage of using heuristics is that the rules can be adapted to different settings, which is a limitation with trained data in machine learning. At the same time, the disadvantage is that the algorithm is not able to learn from new data. In our case, we used four different heuristics:

- **Analysis of the source:** Every tweet has a record in its metadata called source. It shows the platform from where a tweet was sent. The most common source is the Twitter platform, which appears when a user tweets directly from the Twitter website. Some platforms, like IFTTT, are known services that allow to automatically send tweets. The

analysis of the source can be used to find suspicious sources, which are commonly used by bot accounts.

- **Friend-follower-ratio:** Our research has found that often bots on Twitter have a similar number of friends and followers to avoid detection. Moreover, many bots follow each other and they automatically add any account that follows them to their list of friends. A friend-follower-ratio of approximately one can be used to detect bots. For accounts with few followers, it is more probable that by chance the friend-follower-ratio is close to one. Therefore, we only classified accounts with more than 100 followers and an equalized friend-follower ratio as suspicious.
- **Number of tweets per day:** One of the most common criteria for detecting bots is the average number of tweets an account has posted. This is calculated by dividing the total number of tweets by the number of days since the creation of the account. The Oxford Internet Institute has proposed 50 tweets per day as a threshold to identify bots [11]. For this heuristic, we used a more conservative approach and only took into consideration the 5% of the most active users.
- **Text duplicates:** This heuristic is based on the tweet's text. A tweet, whose text is duplicated from another tweet and is not a retweet could have been created by an automation process. Bots are known for copying tweets from other accounts to display a normal user behavior.

The combination of the heuristics defines if a tweet is categorized as originated from a bot account or not. In contrast to Botometer, this approach is based on data contained in individual tweets and not on the whole information of the user. We decided to categorize all the tweets that complied with any of the heuristics in a simple bot category. Tweets that belonged to more than one heuristic were categorized in a strict bot category. The tweets that were created by verified accounts were taken out of the bot analysis.

Bibliography

[1] A. Badawy, E. Ferrara, and K. Lerman, *Analyzing the digital traces of political manipulation: The 2016 Russian interference Twitter campaign*, arXiv preprint, arXiv:1802.04291 (2018).

[2] A. Bessi and E. Ferrara, *Social bots distort the 2016 U.S. presidential election online discussion*, First Monday, 21 (2016).

[3] D. M. Blei, A. Y. Ng, M. I. Jordan, and J. Lafferty, *Latent Dirichlet allocation*, Journal of Machine Learning Research, 3 (2003), pp. 993–1022.

[4] R. M. Bond, C. J. Fariss, J. J. Jones, A. D. I. Kramer, C. Marlow, J. E. Settle, and J. H. Fowler, *A 61-million-person experiment in social influence and political mobilization*, Nature, 489 (2012), pp. 295–298.

[5] R. Fan, J. Zhao, Y. Chen, and K. Xu, *Anger is more influential than joy: Sentiment correlation in Weibo*, arXiv preprint, arXiv:1309.2402 (2013).

[6] R. Faris, H. Roberts, B. Etling, N. Bourassa, E. Zuckerman, and Y. Benkler, *Partisanship, propaganda, and disinformation: Online media and the 2016 U.S. presidential election*, 2017.

[7] E. Ferrara, O. Varol, C. A. Davis, F. Menczer, and A. Flammini, *The rise of social bots*, Communications of the ACM, 59 (2016), pp. 96–104.

[8] K. Garimella, G. D. F. Morales, A. Gionis, and M. Mathioudakis, *Political discourse on social media: Echo chambers, gatekeepers, and the price of bipartisanship*, arXiv preprint, arXiv:1801.01665 (2018).

[9] S. Hegelich, *Die digitale Revolution der Meinungsbildung Social Media im Wahlkampf*, Hans Seidel Stiftung, 2018.

[10] S. HEGELICH AND M. SHAHREZAYE, *The communication behavior of German MPs on Twitter: Preaching to the converted and attacking opponents*, European Policy Analysis, 1 (2015), pp. 155–174.

[11] P. N. HOWARD AND B. KOLLANYI, *Bots, #strongerin, and #brexit: Computational propaganda during the UK-EU referendum*, arXiv preprint, arXiv:1606.06356 (2016).

[12] D. JURGENS, T. FINETHY, J. MCCORRISTON, Y. T. XU, AND D. RUTHS, *Geolocation prediction in Twitter using social networks: A critical analysis and review of current practice*, in ICWSM, AAAI Press, 2015, pp. 188–197.

[13] D. M. J. LAZER, M. A. BAUM, Y. BENKLER, A. J. BERINSKY, K. M. GREENHILL, F. MENCZER, M. J. METZGER, B. NYHAN, G. PENNYCOOK, D. ROTHSCHILD, M. SCHUDSON, S. A. SLOMAN, C. R. SUNSTEIN, E. A. THORSON, D. J. WATTS, AND J. L. ZITTRAIN, *The science of fake news*, Science, 359 (2018), pp. 1094–1096.

[14] G. MAZZOLENI, J. STEWART, AND B. HORSFIELD, *The media and neo-populism: A contemporary comparative analysis*, Greenwood Publishing Group, 2003.

[15] F. MORSTATTER, J. PFEFFER, H. LIU, AND K. M. CARLEY, *Is the sample good enough? Comparing data from Twitter's streaming API with Twitter's Firehose*, in ICWSM, 2013.

[16] L.-M. NEUDERT, B. KOLLANYI, AND P. N. HOWARD, *Junk news and bots during the German parliamentary election: What are German voters sharing over Twitter?*, 2017.

[17] S. OBERMÜLLER, *Manipulation durch Sprache in der politischen Rede des Nationalsozialismus. Parallelen zur politischen Rede von AfD und PEGIDA*, GRIN Verlag, 2016.

[18] O. PAPAKYRIAKOPOULOS, M. SHAHREZAYE, A. THIELTGES, J. C. MEDINA SERRANO, AND S. HEGELICH, *Social Media und Microtargeting in Deutschland*, Informatik-Spektrum, 40 (2017), pp. 327–335.

[19] A. SÄNGERLAUB, M. MEIER, AND W.-D. RÜHL, *Fakten statt Fakes. Verursacher, Verbreitungswege und Wirkungen von Fake News im Bundestagswahlkampf 2017*, 2017.

[20] S. SCHELTER, F. BIESSMANN, M. ZOBEL, AND N. TENEVA, *Structural patterns in the rise of Germany's new right on Facebook*, in ICDM Workshops, IEEE Computer Society, 2016, pp. 440–445.

[21] A. TUMASJAN, T. SPRENGER, P. SANDNER, AND I. WELPE, *Predicting elections with Twitter: What 140 characters reveal about political sentiment*, in Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, 2010, pp. 178–185.

[22] O. VAROL, E. FERRARA, C. A. DAVIS, F. MENCZER, AND A. FLAMMINI, *Online human-bot interactions: Detection, estimation, and characterization*, arXiv preprint, arXiv:1703.03107 (2017).

[23] S. ZANNETTOU, T. CAULFIELD, E. D. CRISTOFARO, M. SIRIVIANOS, G. STRINGHINI, AND J. BLACKBURN, *Disinformation warfare: Understanding state-sponsored trolls on Twitter and their influence on the web*, arXiv preprint, arXiv:1801.09288 (2018).