MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„If It Acts Like a Bot and Talks Like a Bot:

Computational Propaganda in the 2017

Turkish Constitutional Referendum"

verfasst von / submitted by

Beril Bulat

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2018 / Vienna, 2018

Studienkennzahl lt. Studienblatt/ degree programme code as it appears on the student record sheet:	A 066 550
Studienrichtung lt. Studienblatt degree programme as it appears on the student record sheet:	Communication Science
Betreut von / Supervisor:	Univ.-Prof. Dr. Hajo Boomgaarden, PhD

Acknowledgements

To my supervisor Prof. Dr. Hajo Boomgarden, thank you for placing your trust within me, for all your guidance, and for turning this study into a reality.

To my mother, thank you for always believing in me, for helping me dare to dream big, and for the courageous example you set.

To Kaan, thank you for your never-ending patience, your unconditional support, and for always knowing how to put a smile on my face.

I would not have made it this far without you.

Thank you.

Abstract

Political bots are fully or semi-automated algorithms that are politically oriented. They aim to manipulate public opinion and online political discussion by controlling social media accounts while mimicking human users in various ways. They are especially prevalent on Twitter due to structural platform features which enable them to deceive human users. Previous studies showed that bots differed from humans on basic user metrics. It has been argued, therefore, that these features could be used to identify bots. However, all of the studies were conducted in Western contexts. This study investigated if the same applies in a less democratic setting with limited media freedom and high levels of social media usage, hence with increased likelihood for automated social media manipulation. For this purpose, a data set with tweets posted during a week prior to the 2017 Turkish constitutional referendum was used to compare a total of 500 bot accounts with 500 human users, while taking their political attitudes towards the referendum into consideration. The expectation was that political bots would be much more visible and active in such contexts due to regime's dependency on online manipulation. A cross-disciplinary multi-method approach was adopted for data analysis. Results showed that the online conversation was largely dominated by bots that approved the referendum, and with humans that opposed the proposed changes. This was taken as the likely sign of orchestrated algorithmic propaganda in favor of the yes campaign. Results also showed that approving bots did not differ from humans as expected, which contradicted previous findings.

Key Words: Political bots, Computational Propaganda, Social Media, Online Manipulation, Twitter, Referendum

Abstract

„Political Bots" sind voll-oder teilweise automatisierte Algorithmen mit politischer Orientierung, die Social-Media-Accounts kontrollieren und das Verhalten menschlicher Nutzer imitieren, mit dem Ziel, öffentliche Meinung und politische Diskussion online zu manipulieren. Diese automatisierten Akteure sind besonders verbreitet auf Twitter, da bestimmte strukturelle Eigenschaften des Netzwerkes die Täuschung menschlicher Nutzer besonders begünstigen. Frühere Forschung fand, dass sich Bots und menschliche Nutzer in ihren Nutzungskennzahlen (User Metrics) distinkt unterscheiden und dass jene daher zur Identifizierung von Bots genutzt werden können. Jedoch beschränken sich die bisherigen Forschungserkenntnisse ausschließlich auf wetliche Länder. Die vorliegende Studie untersucht, ob ähnliche Unterscheidungen auch in weniger demokratischem Kontext getroffen werden können, wenn Medien- und Pressefreiheit eingeschränkt sind und die Nutzung sozialer Medien eine große Rolle spielt. Die Erwartung war, dass „Political Bots" deutlich sichtbarer und aktiver sein würden in autokratischem Kontext, aufgrund der Abhängigkeit des Regimes von derartiger Manipulation. Der der vorliegenden Untersuchung zugrundeliegende Twitter-Datensatz umfasst sämtliche Tweets, die eine Woche vor dem 2017 in der Türkei abgehaltenen konstitutionellen Referendum gepostet wurden. Anhand einer Stichprobe wurden 500 Bot-Accounts und 500 Accounts menschlicher Nutzer verglichen, unter Berücksichtigung der politischen Einstellungen gegenüber dem Referendum. Die Ergebnisse der Analysen zeigen, dass die Twitter-Konversation dominiert war von Bots, die das Referendum befürworteten, und menschlichen Nutzern, die es ablehnten -- ein Anzeichen für algorithmische Propaganda im Sinne der „yes campaign".

Key Words: Political bots, Computational Propaganda, Social Media, Online Manipulation, Twitter, Referendum

Table of Contents

Introduction

It was not so long ago that social media emerged as the catalyst of free speech, the likely savior of democracy. Proliferated through social media, movements of upheaval were emerging in the most unexpected parts of the world, which immensely boosted cyber-utopianism, the belief that emancipation would come from the transformation of communication channels. There was a consensus that revolution was not to be televised but tweeted. And for a little while, it was. The optimism, however, faded as the commercial businesses and then gradually political actors noticed the capacity of social media in facilitating mass outreach. They began to capitalize on the advancements in this new technology, largely at the expense of its democratic potential.

Social media is not only an advancement in the field of communication technologies; it is a fundamental transformation of the "preexisting media landscape and social structures guiding communication" (Gainous & Wagner, 2014, p.3). This transformation naturally brought new social rules, norms, and actors into consideration. While interactively connecting users, social media has changed our proximity to the media landscape and its surrounding structures. In addition to being continuously exposed to a ceaseless flow of news and information, users have shifted from being passive consumers of media to active content creators. A large part of digital communication no longer takes place between people, but between devices, in languages most individuals neither speak nor comprehend as users. The actual software and hardware that enable this communication remain invisible to the average consumer. Recently, we have also discovered that agents taking part in digital communication are not only humans.

The spotlight fell on social media bots after the 2016 presidential elections in the US. It was already well-known that large amounts of misinformation and fake news were circulating on social media platforms. When the automated agency behind this propaganda surfaced, an ever-escalating cross-disciplinary arms race has raged to detect, identify and categorize these automated actors, to be able to assess their impact on societies and democratic systems. This study is an attempt to contribute to this new line of research, which thus far has mainly focused on elections in Western settings with fully functioning democracies. The geographic bias in previous studies is a considerable shortcoming, as none of the Western societies are dependent on social media for political and informational purposes, hence, manipulation has a smaller impact. Focusing on a different political and societal context where there are considerably limited media liberty and freedom of expression, which translates into a significant dependency on social media, this study investigates political bot behavior and compares them with human user metrics on Turkish Twitter.

The structural features of Twitter, which simultaneously enable political discussion and manipulation, will be discussed in the next section through broadly adopted perspectives on public opinion, news consumption, and social movements. The next section will investigate the current state of bot studies and identify the research gaps. The third section will explain the reasons why Turkey provides the ideal context for a case study on bots. The hypothesis of this work will then be constructed based on the findings of previous research regarding the differences between bot and human metrics. The multi-method approach adopted in this work for analysis will be disclosed in the research methods section. The sixth section will include the data description and findings from the analysis. Finally, in the last section, the results and their implications will be discussed.

Literature Review

Social media platforms arguably have a more significant impact on our lives than any other technological advancement of the last century managed to make. However, the transformation they have brought is far from complete, and so is the debate on how they will affect our futures. This study focuses on Twitter with its asymmetric network form and digital agora-like structure that enables even the automated accounts to take part in the online conversation.

Twitter the Enabler

Twitter is a microblogging service that allows users to post 280-character entries. Once posted, these entries show up on users' profiles and followers' timelines as tweets. Due to its open-to-public logic, connecting with others on Twitter does not require users to confirm connection requests reciprocally. Publicly available profiles are not only the default profile setting, but also the norm. Via timelines, all users can access a real-time stream of tweets posted by the users they follow, and anyone can view these tweets as Twitter does not require registration except for tweeting. This provides increased exposure to a diverse array of opinions and views, which should ideally increase the likelihood of political engagement and discourse between individuals of opposing political views.

Indeed, several studies have found evidence of cross-ideological interactions on Twitter in different experimental settings (Conover et al., 2011; Himelboim, McCreery, & Smith, 2013; Liu & Weber, 2014; Yardi & Boyd, 2010). However, research has also revealed this exposure is rarely translate into rational exchange between parties (Conover et al., 2011; Liu & Weber, 2014; Yardi & Boyd, 2010). Users seldom interact with different minded people, and even when they do, they mainly attempt to change the opinions of others (Xiong & Lio, 2014). Moreover, this rare interaction hardly reflects onto political participation (Valenzuela, Kim, & Gil de Zuniga, 2012). As most users choose to follow and interact with like-minded people, constant partisan exposure in so-called echo chambers leads to increased polarization among users and reaffirmation of user bias (Conover et al., 2011; Sunstein, 2011; Yardi & Boyd, 2010). Owing to influential users and groups with high numbers of followers, opinions on new developments are formed very rapidly online in these echo chambers (Kascheskyet et al., 2011; Xiong & Lio, 2014). These findings suggest a gap between the asserted ideal and actual user practice on Twitter, one that could easily be exploited by bot masters to sway public opinion using the network's structural features.

Twitter increasingly provides the setting to analyze public opinion on political issues, functioning as an opinion polling resource. (Anstead & O'Loughling, 2014; Jungherr & Schoen, 2016) Politicians are, therefore, among the highest beneficiaries of the platform. Capitalizing on big data, they manage to orchestrate campaigns that can potentially change the electoral outcomes (Jungherr & Schoen, 2016). Through user reactions, they get constant feedback regarding their popularity among voters (Ekman & Widholm, 2015). Obama's famous 2008 election campaign was the first in which social media had a decisive role in the elections. Around 55% of the entire adult population had participated online in the political process by either accessing news, communicating with others, or sharing information (Smith, 2009). Since then, social media platforms have become a battleground for politicians, where they promote their persona, spread their cause, raise funds, and retaliate negative campaigning efforts.

But how closely can social media data actually reflect the reality? Evidence shows that there is a strong correlation between Twitter's real-time data and opinion polls (Galley et al., 2016). During federal elections in Germany, party-specific Twitter traffic was found close to election polls in reflecting results (Tumasjan, Sprenger, Sandner, & Welpe, 2010). However, it has also been shown by Gayo-Avello et al (2011). that Twitter data perform poorly at times, due to the "demographic bias in user base" which is posited as “an important electoral factor” (p.3). Twitter data was also used in psychological profiling to help campaigns connect with their audience in particular ways that are likely to lead them to take action. An article published last year in *The Guardian* exposed president Trump’s campaign payments to one such private company, which “models target audience groups and predict the behavior of like-minded people” (Doward & Gibbs, 2017). Although it is hard to estimate the real impact of this strategy in terms of offline political action, social media users are likely to believe the false stories that are supportive of their choice of candidate “especially if they have ideologically segregated media networks” (p. 215) as Alcott & Gentzkow (2017) suggests. Evidence clearly shows that the data collected on our digital actions already renders our political systems vulnerable to manipulation. Inevitably, bot intrusion of online political communication brings forward the accuracy of aggregated data into question.

With the rise of social media also came a significant change in the way we access, select, and consume news, as the increased adoption of mobile devices has led conventional means of news consumption into obsolescence. Consuming news became a "shared social experience" for social media users, who themselves became active content producers (Pew Research Center, 2015). News organizations nowadays embed links in contents they post on social media in an attempt to attract a part of the online traffic to their websites (Flaxman, Goel, & Rao, 2016). It is not unusual anymore to read an article where the author references tweets as primary source (Brands & Graham, 2017). The underlying argument in favor of the modern means of news consumption is that it is beneficial, for both the readers and organizations, as the internet reduces the cost of news production and increases access (Flaxman, Goel, & Rao, 2016). Nevertheless, wider access to a media environment where professional roles become fluid and organizational boundaries are blurred can be harmful at times, as we have seen in the case of recent US Presidential Elections.

Many young people get their news from social media platforms. Reports on digital news consumption shows that 40% of young people between the ages of 18-24 and over 25% of people between the ages of 25-34 use social media websites as their main source of news (Reuters Institute, 2017). However, greater access to news content on social media makes individuals susceptible to misinformation and fake news, whether or not social media is their primary method of getting news. It is well established that frequent exposure to a statement, regardless of its validity, only increases its perceived accuracy and plausibility (Hasher, Goldstein, & Toppino, 1977). Recent research shows that the same effect applies to social media environments. In an experiment conducted with 639 Twitter users where they were shown contents labeled either as news or rumor, Hyegyu & Jung (2017) showed that independent of the label "believability and intention to share were stronger for a tweet with a high number of retweets" (p. 4). Considering a well-infiltrated bot network's immense capacity to retweet and thus increase visibility, bots have the potential to effortlessly alter the way we perceive any news content and information.

Social media platforms employ algorithms to provide their users with personally tailored results (Flaxman, Goel, & Rao, 2016). Seemingly, this helps users navigate their feeds, by showing them the content that best suits their individual preferences. However, customization may also undermine civic discourse by providing users with information that reaffirms their biases, thus promoting the online echo-chambers (Sunstein, 2001). It is known that there is a strong positive association between social network use for news consumption and online political participation (Gil de Zuniga, Jung, & Valenzuela, 2012). This aspect becomes more problematic when it is considered that even the unintentional news exposure positively affects both online and offline political participation (Kim, Chen, & Gil de Zuniga, 2013). The broader picture shows that sufficient algorithmic orchestration could actually yield individuals to take action, regardless of the content validity, or the individual's initial intention.

Networked protests, or connective action networks, are among the most praised transformations brought by the rise of social media (Bennett & Segerberg, 2012). They differ from earlier social movements in terms of organization, logistics, coordination and information diffusion (Anduiza, Cristancho, & Sabucedo, 2014; Teocharis, Lowe, Deth, & Garcia, 2015; Tufekci, 2017). Twitter allows social movements to spread their message and connect with possible participants early on, regardless of their geographic location, which leads to increased visibility and mobilization. However, due to the decentralized and horizontal structure of networked movements, they lack the infrastructure and decision-making capacity that is crucial for generating social or political change (Tufekci, 2017; Teocharis et al., 2015).

Some scholars criticize networked movements for decreasing offline activism by reducing centrally organized high-risk protest movements to the level of simple, one-click activism on social media (Gladwell, 2010; Morozov, 2009). However, research suggests otherwise. Social media activism, or slacktivism, indeed helps social movements by widening their reach and raising awareness (Berbera, et al., 2015). In fact, as illustrated by the Arab Spring, in authoritarian countries where even the smallest expression of insurgency counts, slacktivism is crucial for mobilization and capacity building (Tufekci & Wilson, 2012). Nevertheless, the dependence of networked movements on social media bares the risk of exposure to movement engineering by external forces. In political contexts where online platforms seem to provide the safest means of unsurveilled communication, social movements actually face the highest risk of confronting orchestrated manipulation.

Despite its structural biases and defaults, Twitter has an unfulfilled capacity to increase political dialogue, to provide a polarized but free space for political discussion, and to give rise to mass social mobilization. However, it is also structurally vulnerable to outside influences such as computational propaganda, which could be detrimental for fragile democracies. For scholars, therefore, it is of utmost importance to understand the impact of Twitter, especially in contexts where it has a significant role in online political communication. A shortcoming of prior research in this regard is the way online political communication has been perceived as a human-to-human interaction, although we now know that is not the case. Overlooked by previous studies, automated social media accounts are increasingly participating in the online political discussion in various ways.

Terminology

Although we have only recently become aware of their existence, bots have been around for quite a while. In fact, almost half of all online traffic last year was generated by bots (Zeifman, 2017). Bots are algorithms designed to fulfill specific tasks online, and the nature of these tasks defines how they engage in activities on social media. There are monitoring bots that supervise the health and security of websites (Zeifman, 2017) or journalistic bots that curate content and post information online (Lokot & Diakopoulos, 2016). Then there are malignant bots, like the ones used in distributed denial of service attacks (DDOS) or those employed in identity theft (Zeifman, 2017). The type of bots that are at the focus of this work is automatized social media accounts capable of mimicking human behavior for a particular political agenda.

A variety of terms are in use referring to automated social media accounts, each highlighting different features. At times, the same term is used to describe various phenomena, depending on the authors' disciplinary backgrounds and perspectives. Some scholars emphasize the interactive aspect of bots and employ the term "social bot" for all accounts run by an algorithm (Ferrara et al. 2016), but this feature applies to a variety of bots that are active on social media. Hegelich and Janetzko (2016) underline the element of imitation and define social bots as "automatic programs that are mimicking humans" (p.579), although this definition also overlooks the difference in bot tasks. Boshmaf et al. (2013) highlight the opponent aspect and describe a socialbot as "an automation software that controls an adversary-owned or hijacked account on a particular OSN and has the ability to perform basic activities such as posting a message and sending a connection request" (p. 556). Even though the reference to opponents connotes a political context, this definition is also insufficient as it overlooks the human imitation element.

Some scholars differentiate bots based solely on the level of human assistance they require and adopt the term "bots" for fully automated accounts, and "cyborgs" for semi-automated accounts (Chu, Gianvecchio, Wang, & Jajodia, 2012). There are also those who use the term "sock puppets", emphasizing the false identities used by automated accounts, but this term does not specifically suggest a political agenda (Bastos & Mercea, 2017). In contrast, Howard and Wooley (2016) use the term "political bots", as a sub-category of Ferrara's (2016) social bots and define them as "algorithms that operate over social media, written to learn and mimic real people to manipulate public opinion across a diverse range of social media and device networks" (p. 4885). Although this term and the definition is the most encompassing regarding the goals of this work, it overlooks the human agency required in bot operations in some contexts.

Clearly, the consensus in this field is lacking, and leads to a terminological ambiguity stemming from theoretical foundations. All of the terms mentioned above could indeed be appropriate since bots differ in nature depending on their tasks and goals. For this research, however, I chose to adopt the term political bots, as I focus solemnly on the type of bots designed to achieve a specific political goal. Hence, the following definition is proposed, one that incorporates different factors identified in the definitions presented above: Political bots are fully or semi-automated algorithms that are politically-oriented, that aim to manipulate public opinion, and pollute the online discussion by controlling social media accounts and mimicking human users in sophisticated ways.

Research on Bots

A reciprocal relationship exists between political bots and social media platforms. Different to humans, political bots on Twitter can only perform actions through the Twitter Application Programming Interface (API) (Howard & Kollanyi, 2016). This interface requires a "code-to-code connection to enable real-time posting and parsing of information" (Howard & Kollanyi, 2016, p.1). Genuine users access Twitter through the front door. But the actions taken through the API, whether by a developer or an algorithm, mostly remain invisible (Howard & Kollanyi, 2016; Murthy, et al., 2016). This structural advantage increases the persuasive capacity of political bots and enables them to exploit platforms (Everett, Nurse, & Arnau, 2016), deceiving humans fifty percent of the time (Edwards, Edwards, Spence, & Shelton, 2014). Furthermore, bots are also perceived to be as credible as human users, even when users know they are interacting with bots (Edwards, Edwards, Spence, & Shelton, 2014). As bots become more adept at deceiving human users, it becomes increasingly important that we are able to identify them, understand their behaviors, and measure the extent of their influence on social media platforms, especially considering the platform vulnerabilities identified in the previous section.

Anyone with sufficient social, technical, and financial capital could employ and sustain an army of politicalbots (Murthy, et al., 2016). This suggests that the means to control the content and direction of the online discussion can be owned. In the hands of political actors who have access to sufficient capital, politicalbots transform into agents of suppression, propaganda, and misinformation (Wooley & Howard, 2016). This phenomenon is broadly termed as "computational propaganda" by Wooley & Howard (2016), in reference to the "assemblage of social media platforms, autonomous agents and big data" (p.4886) that enables large-scale public opinion manipulation. How this manipulation works vary depending on the country and sociopolitical context. Although, there are also some commonalities that were seen across digital borders. Strategies such as attacking the opposition, spreading pro-government or pro-candidate messages, increasing a candidate's or a leader's social capital among followers, and demobilizing oppositional movements online can often be observed in countries where political bots are active (Bastos & Mercea, 2017; Bessi & Ferrara, 2016; Ferrara, Varol, Davis, Menczer, & Flammini, 2016; Stieglitz, et al., 2017; Wooley & Howard, 2016). However, the level of their impact on the online political discussion remain unresolved as a question, while the research continues to investigate the extent of automated manipulation.

Political bots gained substantial public attention due to their roles in the 2016 US Presidential Elections and Brexit. In both cases, significant levels of automation and hyper-partisan content was detected, which severely influenced the flow of information and discussion (Bessi & Ferrara, 2016; Woolley & Guilbeault, 2017; Gallacher et al., 2017; Bastos & Mercea, 2017). Research in this field gained momentum upon these findings. The Oxford Internet Institute's Computational Propaganda Project examined political bot presence in a number of countries, results of which are only available in working papers currently. In 2017 German federal elections, only twenty-two accounts showed adequately suspicious activity levels, mostly owing to precautionary regulations and civil society initiatives that prevented automated manipulation (Neudert, 2017). During the French presidential elections, temporal patterns were investigated during a week with 842 thousand tweets, although only six percent of the data set was found to be automated (Howard et al.,2016). In Poland, with a dataset of fifty thousand tweets and ten thousand unique accounts, five hundred potential bot accounts were detected utilizing heuristics. However, these automated agents accounted for 20% of the entire conversation on Twitter. (Gorwa, 2017). Another study that investigated bot activity in Canada found a relatively higher number of bots tasked with amplifying particular political positions, although researchers concluded that their influence on online political discussion remained very much limited (McKelvey & Dubois, 2017). Other research initiatives that focused on European countries, such as Germany (Brachten et al., 2017), and Austria (Kusen & Strembeck, 2017), reported similarly subtle amounts of bot activity.

A limited number of data science studies, however, indicated that the case might be different in other parts of the world. In the Ukrainian Twittersphere, a botnet of 1740 unique political bot accounts was detected, (Hegelich & Janetzko, 2016). Utilizing text-mining and unsupervised learning, several behavioral patterns such as window dressing and amplification were found, and it was shown that political bots with advanced level algorithms could act autonomously at times (Hegelich & Janetzko, 2016). Another study examined Syrian civil war-related tweets by following botnet activities for thirty-five weeks and detected efforts to pollute online discussion through “smoke-screening” and “misinformation” (Abokhodair, Yoo, & McDonald, 2015). In Venezuela, political actors were observed over the course of a year (Forelle et al., 2015). Results showed a small amount of bot activity among politicians yet revealed that bots were mainly employed by the opposition (Forelle et al., 2015). During the 2016 elections in Japan, in a sample of 550 thousand tweets, more than a third of the dataset was found to be automated, spreading propaganda for the Japanese right-wing (Schafer, Evert, & Heinrich, 2017).

In another study of the U.S. Presidential elections, Bessi and Ferrera have gathered more than twenty million election-related tweets to investigate the presence of bots (2016). While inferring political partisanship through hashtag adoption, the authors analyzed the data through the machine learning framework BotorNot Phyton API, which they had previously designed for bot detection. Results showed that almost twenty percent of all election-related conversation on Twitter was generated by roughly seven thousand bot accounts. A similar dataset containing tweets collected a week before the US elections was used in a study which investigated behavioral differences between political bot and human users through metrics (Stieglitz, et al., 2017). According to our knowledge, this study is the first one to compare bot behavior to human behavior within the same data set, instead of features taken from the literature as was the case in earlier studies.

For bot detection, authors primarily employed heuristical approaches such as excluding verified accounts, filtering users based on number of followers and source platforms used for tweeting. Despite the potential shortcoming in their bot detection methods, results complied with previous research, as humans and bots differed in all metrics but one. Comparing bot metrics, which are taken as strong indicators of potential bot influence, to that of humans, Stieglitz et al. found that bots had a lower number of followers, @ symbols, and higher number of URL links in their tweets (2017), which matches with the results of Chu et al. (2012). Moreover, based on the study of Cossu et al. (2016) who proposed that the number of retweets and followers of an account are directly related to its influence they confirmed the relationship between the number of followers and retweeting behavior, showing that the more followers a bot account has, the more influence it has over a broader network through retweets (Stieglitz, et al., 2017).

Clearly, when it comes to studying computational propaganda, the political context and social media dependency matters. In the West, values such as freedom of speech, freedom of the press, the right to information, and the right to protest are all under constitutional protection. Western countries are not dependent on social media for any political reason. However, in a less democratic setting where the state control is all-pervasive, and where media freedom is limited, individuals are dependent on social media to access undistorted news content or engage in non-surveilled political discussion. Hence, controlling the social media environment becomes vital for regimes, and bots provide the most efficient and effective means for such a task. To further contribute to this field of research, therefore, this study investigates the differences in bot and human user behaviors on Twitter in an authoritarian context. Ranking 100th among 167 countries on the Economist's Intelligence Unit Democracy Index (2017), Turkey provides an excellent case for this study.

Turkish Political Background

Once a secular stronghold in the region, then a model of well-tuned balance between democracy and Islam, Turkey has steadily drifted towards authoritarianism in the recent years. The beginning of Turkey's transformation can be traced back to the Gezi Park protests of 2013, the Turkish equivalent of the Arab Spring. While traditional media sources completely turned a blind eye to the largest demonstrations in the country's history, social media facilitated diffusion of information and mass mobilization, and in so doing became the government's biggest enemy (Tucker, et al., 2016). On the second day of the month-long protests, then Prime Minister Erdogan declared Twitter to be “the worst menace to society” (Letsch, 2013). This statement was followed by the violent suppression of the protests and Internet censorship. Since then, the government has resorted to selective social media bans at times of potential political turmoil, as in the case of the ban on Twitter during the corruption scandal in 2014 (Bender, 2014), or when elected Kurdish leaders were arrested in 2016 (Roberts, 2016).

Censorship was not introduced to Turkey by the current government. A 1983 press law restricts the media from “reporting information deemed to fall within the sphere of national security” (Metz, 1996). However, the scope of censorship has radically broadened to include any unfavorable criticism on any platform. According to the Committee to Protect Journalism’s recent numbers, 253 journalists are currently imprisoned in Turkey, and the country ranks 155th on the Press Freedom Index among 180 states (Reporters Without Borders, 2017). As a result, a large percentage of the Turkish population is reported to distrust traditional media sources and prefers social media platforms for news consumption and sharing information (Reuters Institute, 2017). Surveys show that sixty percent of the population of Turkey are active social media users, and Twitter is the fourth most popular platform after Facebook, YouTube, and Instagram (Digital in 2017 Global Overview Report, 2017). Turkish social media users circumvent temporary bans by changing their DNS settings or by employing virtual private networks (VPN). Although the government also banned the most popular VPN services used in the country in 2016, the Global Web Index report shows the country has the third highest share of Internet users employing VPNs to access restricted sites and social networks, with forty-five percent of active users (Kingsley, 2017).

Since states cannot have full control over social media, the next best thing to do for them would be to manipulate social media, unless they ban access to the Internet in its entirety. And as explained in the previous section, political bots are one of the most advanced and effective methods for social media manipulation. I believe that in an authoritarian context, where bots are deployed not only to influence elections but to manipulate users over an extended period, bots and humans will exhibit different behaviors to those identified in the findings of Stieglitz et al. (2017). Moreover, I argue that the political identification of users, whether a genuine user or a political bot, which was overlooked in the previous study, could make a difference in behaviors that are under investigation.

To be able to compare the results with those of existing literature, I examine user behaviors on Turkish Twitter during the constitutional referendum of 2017. This referendum sought approval of 18 amendments to the Turkish constitution, as proposed by the ruling party. The main goal was to replace the existing system with an executive presidency, which concentrates extraordinary powers in the hands of the president. Held under a decree of emergency, the referendum was a showdown between the ruling party, which holds an overwhelming parliamentary majority, and the two opposition parties.

Hypotheses

Considering the political climate and the level of polarization in the society, it is expected to see a different number of political bots and human users depending on their attitudes towards the referendum, which, in this case, could either be “approving” or “opposing” the constitutional referendum:

H1: Approving bots will be higher in number than opposing bots.

H2: Opposing human users will be higher in number than approving human users.

Since political bots are expected to influence users for a political purpose, and since their behaviors are investigated during a constitutional referendum, it is reasonable to expect bot and human users from different political sides to behave differently. The literature suggests that standard features such as number of followers, friends, retweets, @ symbols (suggesting mentions), and URL counts differentiate between political bots and humans, and that these features could be used to identify bots, and to predict their online influence (Stieglitz, et al., 2017; Chu, Gianvecchio, Wang, & Jajodia, 2012).

Stieglitz et al. (2017) argue that the number of followers could be an important indicator of a bot, as bots were found to have a lower number of followers than human users. However, in a well-organized and long-lasting bot scheme that was not injected into the network but organically grown from within, findings are expected to differ:

H3a: Approving bots will have a higher number of followers than the other three groups.

Based on this assumption which suggests a vast network, it is consequently proposed that:

H3b: Approving bots will have a higher number of retweets than the other three groups.

Previous findings suggest that bots generate less original content than humans and thus use @ symbols less frequently (Chu, Gianvecchio, Wang, & Jajodia, 2012; Stieglitz, et al., 2017). In this case, however, political bots are expected to control public opinion through extensive intimidation, which requires some level of human agency in original content creation and mentioning other users frequently. Therefore, it is hypothesized:

H3c: Approving bots will have a higher rate of @ symbols in their tweets than opposing bots.

Moreover, since social media is the only medium of free speech available for the opposition, it is expected that opposition supporters will interact more frequently than the pro-government human users. Hence, it is proposed:

H3d: Opposing human users will have a higher rate of @ symbols in their tweets than approving human users.

Stieglitz et al. suggest that political bots tend to use more URL links in their posts to spread information and generate influence. In an authoritarian setting where the entire media is already under strict government control, political bots would not necessarily be vital for spreading misinformation. In contrast, social media would indeed be a crucial instrument for opposition human users to access and spread alternative news sources and information.

H3e: Opposing human users will have a higher number of URL links in their tweets than the other three groups.

Two important predictors of potential influence on Twitter are the number of followers and the number of retweets. Stieglitz et al. have identified a significant relationship between these features. To test if the same applies in a different political context, it is lastly proposed:

H4a: The number of followers of a bot account will be positively related to the number of retweets of the account for both groups.

H4b: The number of followers of a human user account will be positively related to the number of retweets of the account for both groups.

Data Collection & Methods

Funded by the University of Vienna's Computational Communications Lab, the data was collected through Twitter's Historical Power Track API. The dataset contained roughly 230 thousand unique accounts and around 1.1 million tweets that were posted during the last week of the referendum. The meta data included information such as user id, user location, preferred language of the profile, number of followers and friends of the profile, total status count of the profile, tweet content, replies to the content, retweet count of the content, and URL located in the content. The dataset was called through eight different most popular campaign hashtags used during the referendum process. Two of them were neutral and generally employed by both sides (#Referandum2017, #Oyver), while the remaining six were either pro-approval (#Evet, #EvetZaferMilletindir, #EvetGelecektir) or pro-opposition (#Hayir, #HayirdaHayirVar, #BugunHayirCikacak), three per each side. The Historical PowerTrack API yielded all of the tweets, retweets and quote tweets in which the hashtags were used.

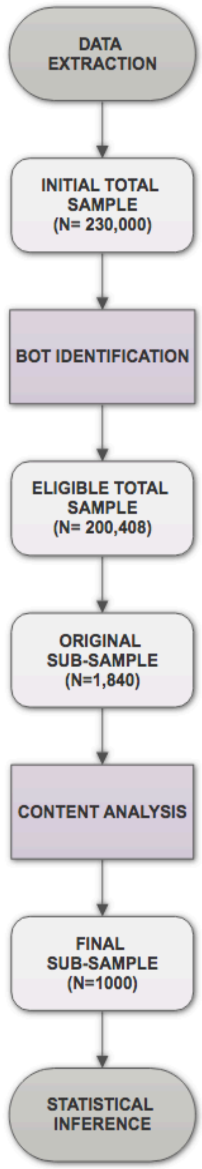
The study adopted a cross-disciplinary multi-method approach that incorporated a machine learning framework for bot detection, content analysis for user classification and quantitative analysis for statistical inference. The first step, therefore, involved a state-of-the-art detection mechanism, the BotOrNot machine learning framework, which was also the preferred method in the previous studies of 2016 US Presidential Elections (Bessi & Ferrara, 2016; Woolley & Guilbeault, 2017). According to its creators at the Indiana University, BotOrNot models were trained with 5.6 million tweets and over 30 thousand profiles for classification (Davis, Varol, Ferrera, Flammini, & Menczer, 2016). In its assessment, BotOrNot classification algorithm employs more than 1,150 features that it has previously learned, which are grouped into six main classes: network diffusion, user profile information, friends, temporal patterns, content, and sentiment. The algorithm then returns two scores. The language-based score considers the classes of content and sentiment for tweets that are in English. For this work, due to linguistic limitations, only the language-independent score was taken into account, which is estimated by ignoring the content and sentiment classes.

After parsing the user information from the data set, the entire list of unique users was sent to BotOrNot Python API. Once it is fed account information, BotOrNot retrieves the account's activities and features on Twitter in real time. Then it reports a score as it completes its assessment, which represents the probability of the account being automated, where higher scores indicate higher probability of automation. However, a severe limitation was faced in the identification process due to BotOrNot evaluation taking place in real time. Results showed that 25% of the unique accounts, around 63 thousand in total, was not suitable for assessment. A portion of these accounts had no tweets on their timelines anymore (%0.3 of the entire data set), some were deleted by the account owners entirely (% 9.1), some were set as private (%12) and some were suspended by Twitter due to suspicious activity (%2.4). Since it was not possible to assess the automated activity behind these accounts with the study method, they were removed from the final sample. This impediment also rendered around 300 thousand tweets unusable, which were removed from the sample as well (Table 1).

Table1

Summary of the data set based on BotOrNot scores

	N_{USERS}	N_{TWEETS}
Total Sample	262,834	1,096,605
Eligible Sample	200,408	793,764
≥ %75	923	3411
≤ %25	137,335	412,068



Once the accounts were removed, remaining eligible total sample from the data set included roughly 793 thousand tweets posted by 200 thousand unique accounts. Two subsamples were created out of this population (Figure 1). First, in order to separate bots from genuine human users, a threshold of 25% was set based on user scores received from BotorNot. This means that the accounts with a score of 25% or less were accepted to be most likely-humans, whereas those who had 75% or higher were accepted as most likely bots. The population of analysis for bots included 923 likely-bot accounts and 3411 tweets. Whereas for humans, the sample size with a score of 25 or lower turned up to be much larger (Table 1). Therefore, before moving on to content analysis for categorization, 917 users were randomly selected among from most-likely human sample, to ensure comparability for our first two hypothesis. Then, based on content analysis results, a second sub-sample was created to test the remaining hypothesis. Details of this process are as disclosed below.

Figure 1: Data preparation and analysis steps.

Content analysis was conducted by two independent native speaker coders who classified each original tweet content into pre-defined categories of approval, opposition and neutral. During analysis, the one-to-one protocol was implemented, thus each coder was required to classify each tweet strictly into only one of the three categories. Coders double-coded the entire data set of tweets and then based on these initial results, each user was further classified into one of three categories. All of the tweets that showed support and approval for the referendum, any of the proposed amendments, government and its members, the president, and also the tweets that clearly disapproved the opposition campaign and any of the opposition leaders were categorized as "approving". The posts that clearly showed disapproval of the proposed amendments, government and its members, or the "yes campaign", and those that supported the opposition campaign or members were categorized as "opposing". Any tweet with non-sided informational content, or with ambiguous content that lacked any clear identification or attitude towards the referendum was categorized as "neutral". Discrepancies between coders were discussed and a consensus was reached for each tweet and user before they were added to the analysis.

Table 2

Summary of populations based on content analysis

	Bots	Humans
Approving	485	266
Opposing	301	537
Neutral	137	134
Total	923	917

The initial subsample included all of the accounts that were subject to content analysis. This subsample was used to test the hypotheses 1 and 2. Then, 250 accounts were randomly selected from each category and a final subsample of 1000 unique accounts was compiled to statistically analyze and test our hypothesis 3 and 4 regarding user metrics. In this work, R version 3.5.0 was used for statistical inference.

Analysis & Results

The 250 identified opposing bots had the highest number of followers amongst the four groups with 2764.37 followers on average per account (SD = 27,097.72), followed by approving bots with 2357.62 followers (SD = 9892.67). In contrast, approving human users had 256.77 followers on average (SD = 460.84) and opposing human users had 448.91 followers (SD= 1687.64). Approving bots also had the highest mean score when it came to the usage of @ symbol with an average of 3.96 per account (SD = 8.629), whereas opposing bots used mentions 1.75 times (SD = 5.91). Among approving users, the average was 2.64 mentions per account (SD = 7.90). And for opposing users 2.74 (SD = 5.49). The number of URL shares were surprisingly low in the dataset, with an average of .54 per account (SD = 2.90) for approving bots, and .36 (SD = 1.49) for opposing bots, compared to .23 (SD = .623) for approving users, and .78 (SD = 4.46) for opposing users. When it came to the number of retweets in the data set, it was again approving bots that had the highest mean scores with 3.42 (SD = 6.19) per account, whereas opposing bots had 1.48 (SD = 4.642) retweets on average. Approving users had 1.91 (SD = 6.08) retweets compared to 2.51 (SD = 5.104) average retweets of opposing human users per account (Appendix).

In testing our first two hypotheses, we used the original subsample that had 923 bot accounts and 917 human accounts. This sample included 485 approving bot accounts in contrast to 301 opposing, and 537 opposing human users, compared to approving 266 (Table 2). A Pearson's Chi-square analysis was performed to test if there is a significant difference between the expected and observed frequencies in samples. Results showed that there was indeed a significant difference between groups, and that the user type was not independent of the political attitude in the referendum X^2(2, N = 1840) = 132.42 , p <.001. Upon investigating the standardized residuals, it is concluded that there were significantly more approving bots (S_R = 10.271) than expected, compared to opposing bots (S_R = -11.175), and significantly more opposing human users (S_R = 10.271), compared to approving users (S_R =11.175) (Figure 2). The hypothesis 1 and 2 were therefore supported.

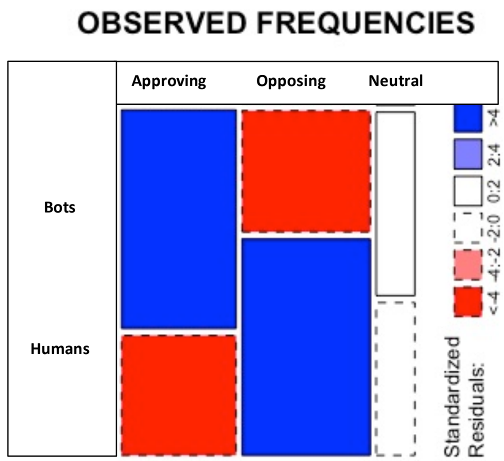


Figure 2: Visualization of the observed frequencies per category, color-coded based on standardized residual values. Large blue areas indicate extreme frequency in the data ($S_R > 4$)

For the following hypothesis, mean values of each category were compared for the variables "followers", "@ count", "retweet count" and "URL count" with the final subset that included 250 accounts per category. As the data set lacked a normal distribution and no equality of variance was observed, a non-parametric permutation test was performed for mean comparison. Permutation tests reshuffle categories and recompute mean differences each time to see where the observed value falls among all the possible values. This makes an interference possible regarding the likelihood of getting the observed value due chance. The data set in this work was permuted 5000 times using a two-sided confidence interval at %2.5 and %97.5 for each hypothesis.

For hypothesis 3a, it was argued that approving bots would have more followers than the other groups. Permutation results showed approving bots had more followers than both approving users (M_D = 2072.38) and opposing users (M_D = 1887.98), but fewer followers than opposing bots (M_D = -329.15). However, all of these values fell in between the confidence interval (-2375.36 - 2351.99) meaning none of them were extreme compared to possible mean values, therefore the results were non-significant. The hypothesis was rejected (Figure 3).

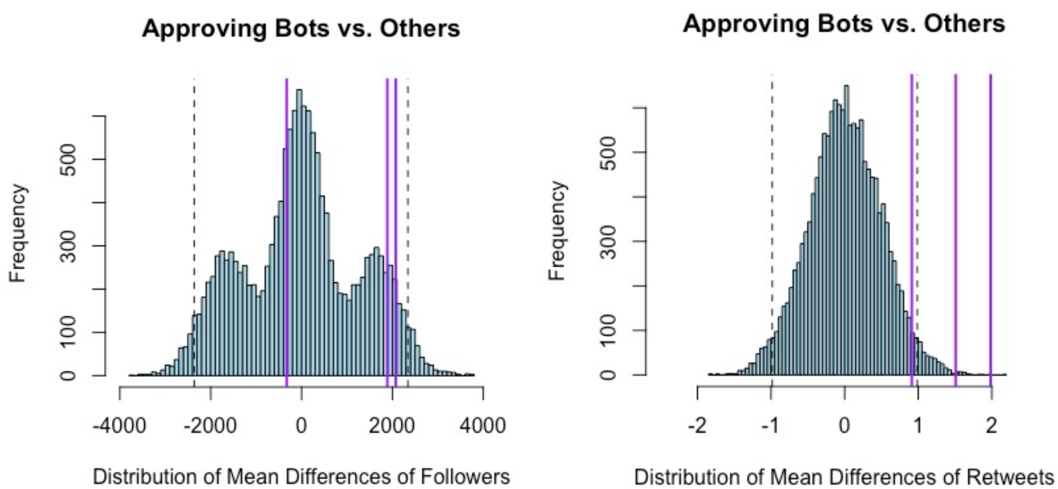


Figure 3: Permutation results for hypothesis 3a on the left and 3b on the right. Straight lines illustrate the observed values among the distribution of the possible mean differences for each case of comparison between categories. Dotted lines are the limits of confidence intervals. For 3b the mean difference between approving bots is significantly different than opposing bots and approving humans.

Hypothesis 3b claimed that approving bots would have more retweets in the data set than the other groups. Permutation results showed approving bots had significantly more retweets than opposing bots (M_D = 1.98), and approving users (M_D = 1.51). However, for opposing users, no significant difference was found (M_D = 0.91). The observed value, in this case, fell in between the confidence interval (-.98 – .98). The hypothesis was only partially supported (Figure 3).

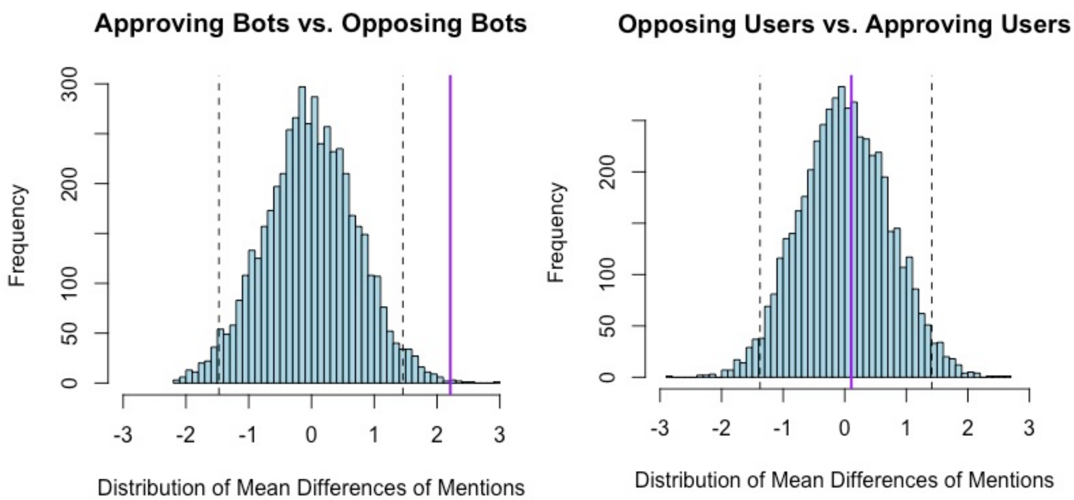


Figure 4: Results for the hypothesis 3c on the left, and hypothesis 3d on the right. The single line is the observed value among the distribution of possible mean differences of mentions.

For the hypothesis 3c, it was argued that approving bots would have more @ symbols in their tweets as an indication of mentions, than opposing bots. Permutation results supported this assumption ($M_D = 2.12$, $p < .01$). Hypothesis 3c was therefore supported (Figure 4). However, hypothesis 3d, which argued that opposing human users would have more @ symbols in their tweets than approving users, was rejected. No significant difference was observed between group means ($M_D = 0.18$), the observed value found to be in between the confidence intervals (-1.375410 - 1.413915) (Figure 4).

For the hypothesis 3e, it was argued that opposing human users would share more URL links than all the other three groups. Permutation results showed the opposing users shared more URLs than approving bots ($M_D = .24$), opposing bots ($M_D = .42$) and approving users ($M_D = .54$), but the difference was only significant for approving users and not for any of the bot groups (Figure 5). The hypothesis was thus only partially supported.

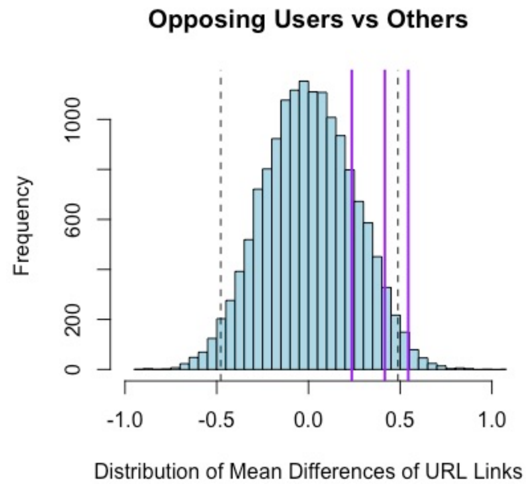


Figure 5: Permutation results for hypothesis 3e. Lines illustrate the observed values among the distribution of possible mean differences of URL shares.

For hypothesis 4a and 4b, due to the non-normality and unequal variance in the data set, nonparametric Spearman's rank-order correlation analysis was performed to test whether the number of followers of an account was positively related to the number of retweets. Among bots, a moderate but significant correlation was found between the number of followers and retweets ($r_s = .13$, $p < .001$) which meant the number of retweets increased with the number of followers and vice versa. Therefore, hypothesis 4a was supported. However, this was not the case for humans. Results showed there was no association between the number of followers and retweets of a human user ($r_s = .08$, $p = .07$). Hence, 4b was rejected.

Discussion

The goal of this work was to compare human and bot account metrics to test if the assumptions about bots would hold in a different socio-political context. Specifically, this study investigated the differences between human and bot account metrics during a constitutional referendum in a less democratic context where freedom of speech and media are dramatically abused, and where there are high levels of social media usage. The main argument was that high levels of social media dependency would increase the likelihood of manipulation by political actors, and this would translate into bots and humans behaving differently, based on their attitudes towards the referendum.

Results show that there were significantly more approving bots than expected compared to opposing bots, and significantly more opposing human users compared to approving users. This means that the user type is not independent from the attitudes towards the referendum and a significant number of bots were approving the referendum, whereas human users were largely opposing. It is clear that social media manipulation was performed to boost the government campaign during the referendum, and a significant part of human users used the platform to support the opposition as expected.

Following hypothesis investigated the main differences between bot and human user accounts while taking their attitudes in the referendum into consideration. Contrary to expectations and to previous research, bot account metrics did not significantly differ from human account metrics in most cases. Regarding the number of followers, approving bots had less followers than opposing bots. Moreover, both groups had a higher number of followers on average than human users, although the difference was not significant. The same applied to the number of retweets, as approving bots had more retweets in the data set than all three groups, although this difference was only significant for opposing bots and approving users. When it comes to the generally held assumption about bots using less @ symbols due to their lesser ability to create original content, it was argued that approving bots used significantly more @ symbols in their tweets than opposing bots, which was approved. Humans did not differ from each other regarding usage of @ symbols. For the variable URL link shares, the only significant difference was found between opposing users and approving users.

The literature suggests that bots and humans differ on basic behavioral user metrics, and that metrics thus could be used to identify possible bot accounts (Chu, Gianvecchio, Wang, & Jajodia, 2012; Stieglitz, et al., 2017). However, findings clearly show that bots in my data set did not differ from humans, neither as I expected., nor as it was proven in previous studies. In fact, approving bots were mostly human-like in terms of number of followers, number of retweets, number of @ symbols and URL link shares. In this sense, study results are interesting as they indicate what we know about bots may not apply to those found in less democratic contexts. Considering the empirical work of Edwards (2014), which showed that genuine users mistake bots for humans 50% of the time, the similarity of bots to humans on Turkish Twitter could even be seen as alarming. Another important outcome of the study from this perspective is that it shows user metrics by themselves are not liable for bot detection purposes, and they should be supplemented with rather more advanced methods, or they are likely to lead to inaccurate conclusions.

The last two hypotheses were concerning the relationship between the number of followers of the account types and the number of retweets. The assumption about bots was supported, as there was a modestly positive and significant relationship. This suggested that the more followers a bot had, the more retweets it got. However, the same did not apply to human user accounts, which showed no association. These findings matched with the previous literature (Stieglitz, et al., 2017), which argues that the number of followers and retweets of bots are suggestive of possible influence, meaning the higher the numbers, the higher the impact and reach (Cossu, Labatut, & Dugue, 2016; Stieglitz, et al., 2017). One likely explanation could then be that bots are largely tasked with retweeting and following other bots to amplify each other's influence over the network. In contrast, human users are more selective, and their decision is based on personal preference. If political bots are indeed mainly interacting with other bots as the literature suggests, results could then be indicative of a bot-network in Turkey, which is capable of behaving similar to human users.

Conclusion, Limitations & Future Research

This study compared a total of 500 political bots to 500 human users taken from the same data set which included tweets posted a week prior to the 2017 Turkish constitutional referendum. To my knowledge this was the first study that compared bot and human user metrics while considering political attitudes. The goal was to show that in a less democratic context where basic freedoms are curtailed and where heavy state surveillance prevails, bot agents would be a lot more active than previous work in this field suggested.

By categorizing accounts based on their political attitudes towards the referendum, this study showed there were significantly more bots that were supportive of the amendments, and significantly more humans that were in opposition. This was a clear sign of algorithmic manipulation in favor of the government's campaign. It was also shown that bots who were supportive of the constitutional change did not lack behind human users in terms of standard features. Overall, these findings are important in understanding political bots and their influence over social networks in less democratic contexts.

One limitation of this study was the number of ineligible accounts that were filtered out prior to analysis. Due to bot detection algorithm running real time, around one-fourth of the entire data set obtained from Twitter were rendered unusable by the time the data set was accessed. This was a serious impediment, as around six thousand accounts were already suspended by Twitter due to suspicious activity. Future research should collect and filter the data set in real time without any delays to be able to capture and identify all of the bot accounts that were active during the data collection process. Moreover, this work was focused on the week prior to the referendum day, which might have possibly diminished the level of bot activity from the start by limiting its scope with temporal restrictions. However, social media manipulation in less democratic contexts is not limited to elections or referendums. To get a better picture of computational propaganda on social media platforms, future research should consider investigating the entire campaigning processes.

References

Abokhodair, N., Yoo, D., & McDonald, D. (2015). Dissecting a Social Botnet: Growth, Content and Influence in Twitter. *Proceedings of the 2015 ACM International Conference on Computer-Supported Cooperative Work and Social Computing* (pp. 839-851). Association for Computing Machinery.

Alcott, H., & Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives, 31*(2), 211-236.

Anduiza, E., Cristancho, C., & Sabucedo, J. (2014). Mobilization through online social networks: the political protest of the indignados in Spain. *Information, Communication & Society, 17*(6), 750-764.

Anstead, N., & O'Loughling, B. (2014). Social Media Analysis and Public Opinion: the 2010 UK general Elections. *Journal of Computer Mediated Communication, 20*(2), 204-220.

Bastos, M., & Mercea, D. (2017). The Brexit Botnet and User-Generated Hyperpartisan News . *Social Science Computer Review* .

Bender, J. (2014, March 21). *10 Shameful Facts About Censorship in Turkey*. Retrieved October 28, 2017, from Business Insider: http://www.businessinsider.com/10-facts-about-censorship-in-turkey-2014-3?_ga=2.10136181.1743567774.1514953729-1721344500.1514510303&IR=T

Bennett, L., & Segerberg, A. (2012). The Logic of Connective Action. *Information, Communication & Society, 15*(5), 739-768.

Berbera, P., Wang, N., Bonneau, R., Jost, J., Nagler, J., & Tucker, J. (2015, November 30). The Critical Periphery in the Growth of Social Protests. *Plos One, 10*(11).

Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 U.S. Presidential election online discussion. *First Monday, 21*(11).

Boshmaf, Y., Muslukhov, I., Beznosov, K., & Ripeanu, M. (2011). The Socialbot Network: When bots socialize for fame and money. *Proceedings of the 27th Annual Computer Security Applications Conference* (pp. 93-102). New York: ACM.

Brachten, F., Stieglitz, S., Hofeditz, L., Kloppenborg, K., & Annette, R. (2017). Strategies and Influence of Social Bots in a 2017 German State Election - A case study on Twitter. *Australasian Conference on Information Systems.* Hobart.

Brands, B. J., & Graham, T. (2017). Social Media Sourcing Practives: How Dutch Newspapers Use Tweets on Political News Coverage. *Managing Democracy in the Digital Age*, 159-178.

Chu, Z., Gianvecchio, S., Wang, H., & Jajodia, S. (2012, August 23). Detecting automation of twitter accounts: Are you a human, bot, or cyborg? *IEEE Transactions on Dependable and Secure Computing, 9*(6), 811-824.

Conover, M. D., Ratkiewicz, J., Francisco, M., Goncales, B., Flammini, A., & Menczer, F. (2011). Political Polarization on Twitter. *Proceedings of The Fifth International Conference on Weblogs and Social Media.* Barcelona: The AAAI Press.

Cossu, J.-V., Labatut, V., & Dugue, N. (2016). A Review of Features for the Discrimination of Twitter USers: Application to the PRediction of Offline Influence. *Social Network Analysis and Mining, 6*(1), 25.

Davis, C., Varol, O., Ferrera, E., Flammini, A., & Menczer, F. (2016). BotOrNot: A system to Evaluate Social Bots. *Proceedings of the 25th International Conference Companion on World Wide Web.* WWW '16 Companion .

Doward, J., & Gibbs, A. (2017, March 4). *Politics.* Retrieved October 20, 2017, from The Guardian: https://www.theguardian.com/politics/2017/mar/04/nigel-oakes-cambridge-analytica-what-role-brexit-trump

Edwards, C., Edwards, A., Spence, P., & Shelton, A. (2014). Is that a bot running the social media feed? Testing the differences in perceptions of communication quality for a human agent and a bot agent on Twitter. *Computers in Human Behavior, 33*, 372-376.

Ekman, M., & Widholm, A. (2015). Politicians as Media Producers: Current Trajectories in the relation between journalists and politicians in the age of social media. *Journalism Practive, 9*(1), 78-91.

Everett, R., Nurse, J., & Arnau, E. (2016). The anatomy of online deception: what makes automated text convincing? *Proceedings of the 31st Annual ACM Symposium on Applied Computing* (pp. 1115-1120). NYC: ACM .

Ferrara, E. (2015). Manipulation and abuse on social media. *ACM SIGWEB Newsletter*.

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016, July). The Rise of Social Bots . *Communications of the ACM, 59*(7), 96-104.

Flaxman, S., Goel, S., & Rao, J. (2016, March 22). Filter Bubbles, Echo Chambers, and Online News Consumption. *Public Opinion Quarterly, 80*(1), 298-320.

Forelle, M., Howard, P., Hernandes, A., & Savage, S. (2015, July 26). Political Bots and the Manipulation of Public Opinion in Venezuela.

Gainous, J., & Wagner, K. (2014). *Tweeting to Power: The Social Media Revolution.* NY: Oxford University Press.

Gallacher, J., Kaminska, M., Kollanyi, B., & Howard, P. N. (2017). *Junk News and Bots during the 2017 UK General Election: What Are UK Voters Sharing Over Twitter?* University of Oxford, Computational Propaganda Project. University of Oxford.

Galley, J., Rossini, P., Hemsley, J., Kenski, K., Zhang, F., Bryant, L., & Semaan, B. (2016). SOCIAL MEDIA, U.S. PRESIDENTIAL CAMPAIGNS, AND PUBLIC OPINION POLLS: DISENTANGLING EFFECTS. *Selected Papers of AoIR 2016: The 17th Annual Conference of the Association of Internet Researchers.* Berlin: AoIR.

Gayo-Avello, D., Metaxas, P., & Mustafaraj, E. (2011). Limits of Electoral Predictions Using Twitter. *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media.* Association for the Advancement of Artificial Intelligence.

Gil de Zuniga, H., Jung, N., & Valenzuela, S. (2012). Social Media Usa for News and Individuals' Social Capital, Civic Engagement and Political Participation. *Journal of Computer Mediated Communication , 17*(3), 319-336.

Gladwell, M. (2010, October 10). Small Change: Why the revolution will not be tweeted. *The New Yorker*.

Gorwa, R. (2017). *Computational Propaganda in Poland: False Amplifiers and the Digital Public Sphere.* University of Oxford , Computational Propaganda Project . Working Paper.

Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the Conference of Referential Validity. *Journal of Verbal LEarnng and Verbal Behavior, 16*, 107-112.

Hegelich, S., & Janetzko, D. (2016). Are Socialbots on Twitter Political Actors?" . *Proceedings of the 10th International Conference on Web and Social Media* , 579-582.

Himelboim, I., McCreery, S., & Smith, M. (2013, January 25). Birds of a Feather Tweet Together: Integrating Network and Content Analyses to Examine Cross-Ideology Exposure on Twitter. *Journal of Computer-Mediated Communication, 18*(2), 154-174.

Howard, P. N., & Kollanyi, B. (2016, June 20). Bots, #Strongerin, and #Brexit: Computational Propaganda During the UK-EU Referendum.

Howard, P., Bradshaw, S., Kollanyi, B., Clementine, D., & Bolsover, G. (2017). *Junk News and Bots during the French Presidential Elections: What are French Voters Sharing Over Twitter?* Data Memo 2017.3, University of Oxford , Computational Propaganda Project .

Hyegyu, L., & Jung, O. H. (2017). Normative Mechanism of Rumor Dissemination on Twitter. *Cyberpsychology, Behavior and Social NEtworking, 20*(3).

Ju, A., Jeong, S. H., & Chyi, I. (2014). Will Social Media Save Newspapers. *Journalism Practice, 8*(1).

Jungherr, A., & Schoen, H. (2016). The Mediation of Politics through Twitter: An Analysis of Messages posted during the Campaign for the German Federal Election 2013. *Journal of Computer Mediated Communication*(21), 50-68.

Kaschesky, M., Sobkowicz, P., & Bouchard, G. (2011, July). Opinion mining in social media: modeling, simulating and visualizing political opinion formation in the web. *Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times*, 317-326.

Kim, Y., Chen, H.-T., & Gil de Zuniga, H. (2013). Stumbling upon news on the Internet: Effects of incidental news exposure and relative entertainment use on political engagement. *Computers in Human Behavior, 29*(6), 2607-2614.

Kingsley, P. (2017, June 10). Turks Click Away, but Wikipedia Is Gone. *New York Times, 2018*(October), p. 28.

Kusen, E., & Strembeck, M. (2017). *An Analysis of the Twitter Discussion on the 2016 Austrian Presidential Elections.* Complexity Science Hub Vienna. Vienna: arXiv preprint arXiv:1707.09939.

Letsch, C. (2013, June 2). Turkish Protests: Prime Minister Blames Opposition. *The Guardian*.

Liu, Z., & Weber, I. (2014). Is Twitter a Public Sphere for Online Conflicts? A Cross-Ideological and Cross-Hierarchical Look. In L. M. Aiello, & D. McFarland (Ed.), *Social Informatics: 6th International Conference on Social Informatics.* Barcelona: Springer.

Lokot, T., & Diakopoulos, N. (2016). News Bots: Automating news and information dissemination on Twitter. *Digital Journalism, 4*(6), 682-699.

McKelvey, F., & Dubois, E. (2017). Computational Propaganda in Canada: The Use of Political Bots. *Working Paper* .

Metz, H. C. (1996). *Turkey: A Country Study .* Washington D.C. : Federal Research Division, Library of Congress.

Metzger, M., & Flanagin, A. (2013). Credibility and trust of information in online environments: The use of cognitive heuristics. *Journal of Pragmatics, 59*, 210-220.

Morozov, E. (2009). Iran: Downside to the "Twitter Revolution". *Dissent, 56*(4), 10-14.

Murthy, D., Powell, A., Tinati, R., Anstead, N., Carr, L., & Halford, S. W. (2016). Automation, Algorithms, and Politics| Bots and Political Influence: A Sociotechnical Investigation of Social Network Capital. *International Journal Of Communication, 10*.

Neudert, L.-M. (2017). *Computational Propaganda in Germany: A Cautionary Tale .* Working Paper No. 2017.7, University of Oxford, Computational Propaganda Project.

Pennycook, G., & Rand, D. (2017). The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Stories Increases Perceived Accuracy of Stories Without Warnings . *Working Paper* .

Pew Research Center. (2015). *Understanding the participatory news consumer: How internet and cell phone users have turned news into a social experience.*

Porter, J. (2009, March 29). *Bokardo .* Retrieved July 2, 2017, from Relationship Symmetry in Social Networks: Why Facebook will go Fully Asymmetric: http://bokardo.com/archives/relationship-symmetry-in-social-networks-why-facebook-will-go-fully-asymmetric/

Reporters Without Borders. (2017). *Press Freedom Index.*

Reuters Institute. (2017). *Digital News Report .* University of Oxford.

Roberts, H. (2016, April 11). *The Turkish Government reportedly blocked WhatsApp and other social media sites*. Retrieved October 28, 2017, from Business Insider : http://www.businessinsider.de/social-media-and-messaging-sites-blocked-in-turkey-2016-11?r=UK&IR=T

Schafer, F., Evert, S., & Heinrich, P. (2017). Japan's 2014 General Election: Political Bots, Right-Wing Internet Activism, and Prime Minister Shinzo Abe's Hidden Nationalist Agenda. *Big Data, 5*(4).

Smith, A. (2009). *The Internet's Role in Campaign 2008.* PEW Research Center , Internet & Technology.

Stieglitz, S., Brachten, F., Berthele, D., Schlaus, M., Venetopoulou, C., & Veutgen, D. (2017). Do Social Bots Still Act Different to Humans? Comparing Metrics of Social Bots with Those of Humans. In M. G., *Social Computing and Social Media. Human Behaviour. Lecture Notes in Computer Science* (Vol. 10282, pp. 379-395).

Sunstein, C. (2001). *Echo Chambers : Bush vs. Gore Impeachment and Beyond.* Princeton, NJ: Princeton Univeristy Press.

Sunstein, C. (2011). *Going to Extremes: How Like Minds Unite and Divide.* New York City, NY: Oxford University Press.

Teocharis, Y., Lowe, W., Deth, J., & Garcia, G. (2015, August 12). Using Twitter to mobilize protest action: online mobilization patterns and action repertoires. *Information, Communication & Society, 18*(2), 202-220.

The Economist Intelligence Unit. (2017). *Democracy Index 2017.* The Economist.

Tucker, J. A., Nagler, J., Metzger, M., Barbera, Pablo, Penfold-Brown, D., & Bonneau, R. (2016). Big Data, Social Media and Protests. In M. Alvares, *Computational Social Science Discovery and Prediction* (pp. 199-224). NY: Cambridge University Press.

Tufekci, Z. (2017). *Twitter & Tear Gas: The Power and Fragility of Networked Protest.* New Haven: Yale University Press.

Tufekci, Z., & Wilson, C. (2012). Social Media and the Decision to Participatein Political Protest: Observations From TahrirSquare. *Journal of Communication, 62*, 363-379.

Tumasjan, A., Sprenger, T., Sandner, P., & Welpe, I. (2010). Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media.* Munich : Association for the Advancement of Artificial Intelligence.

Valenzuela, S., Kim, Y., & Gil de Zuniga, H. (2012). Social Networks That Matter: Exploring the Role of Political Discussion for Online Political Participation. *International Journal of Public Opinion Research , 24*(2), 163-184.

We Are Social . (2017). *Digital in 2017 Global Overview Report.*

Wooley, S., & Howard, P. (2016). Political Communication, Computational Propaganda, and Autonomous Agents — Introduction. *International Journal of Communication, 10*, 4882-4890.

Woolley, S., & Guilbeault, D. (2017). *Computational Propaganda in the United States of America: Manufacturing Consensus Online.* Working Paper , University of Oxford, Computational Propaganda Research Project.

Xiong, F., & Lio, Y. (2014). Opinion Formation on Social Media : An empirical approach. *Chaos: An Interdisciplinary Journal of Nonlinear Science, 24*(1).

Yardi, S., & Boyd, D. (2010, October). Dynamic Debates: An Analysis of Group Polarization Over Time on Twitter. *Bulletin of Science Technology & Society, 30*(5), 316-327.

Zangerle, E., & Specht, G. (2014). "Sorry, I was hacked" A classification of Compromised Twitter Accounts. *Proceedings of the 29th Annual ACM Symposium on Applied Computing*, (pp. 587-593). Gyeongju.

Zeifman, I. (2017). *Bot Traffic Report 2016*. Retrieved July 1, 2017, from Incapsula: https://www.incapsula.com/blog/bot-traffic-report-2016.html

Appendix

Descriptive Statistics

GROUP		Followers	Mentions	URL Links	Retweets
Approving Bots	Mean	2357.62	3.96	.54	3.42
	N	250	250	250	250
	Std. Deviation	9892.669	8.629	2.901	6.189
	Variance	9.79E+7	74.461	8.418	38.301
Opposing Bots	Mean	2764.37	1.75	.36	1.48
	N	250	250	250	250
	Std. Deviation	27097.7	5.913	1.486	4.642
	Variance	7.34E+8	34.958	2.207	21.544
Approving Humans	Mean	256.77	2.64	.23	1.91
	N	250	250	250	250
	Std. Deviation	460.841	7.911	.623	6.077
	Variance	212375	62.586	.388	36.932
Opposing Humans	Mean	448.91	2.74	.78	2.51
	N	250	250	250	250
	Std. Deviation	1687.637	5.493	4.463	5.104
	Variance	2848120	30.175	19.917	26.050