Czech University of Life Sciences Prague

Faculty of Economics and Management

Department of Information Technologies

Bachelor Thesis

Impact and Analysis of Computational Propaganda

Changqi Sun

CZECH UNIVERSITY OF LIFE SCIENCES PRAGUE

Faculty of Economics and Management

BACHELOR THESIS ASSIGNMENT

Changqi Sun

Informatics

Thesis title

Impact and Analysis of Computational Propaganda

Objectives of thesis

The main objective of the thesis is to find and analyze the behavioral characteristics of social robots during social events.
Partial objectives:
-Analyzing the behavioral characteristics of social bots in different geographic regions.
-Investigating the impact of social bots on actual users.
-Propose methods to identify social bots.

Methodology

The methodology of the thesis is based on the processing and analysis of data, modeled by using Naive Bayes classifier and questionnaires for social populations, and the final conclusions are drawn based on the results.

The proposed extent of the thesis

40-50

Keywords

social media, fake news, social bots, algorithm, global issues

---

Recommended information sources

Emily Taylor, Stacie Hoffmann, 2020. Responses to Social Media Propaganda Manipulation. Retrieved from https://oxil.uk/publications/responses-to-social-media-propaganda-manipulation/

Giovanni Da San Martino, Stefano Cresci, Alberto Barron-Cedeno, Seunghak Yu, Roberto Di Pietro, Preslav Nakov, 2020. A Survey on Computational Propaganda Detection. Retrieved from https://arxiv.org/abs/2007.08024

Holly Schnader, 2018. The evolution of computational propaganda: Trends, threats, and implications now and in the future. Retrieved from https://commons.lib.jmu.edu/honors201019/713/

Samuel C. Woolley, Philip N. Howard, 2018. Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media. ISBN 9780190931407

Woolley, S and Guilbeault, D. 2017. Computational propaganda in the United States of America: Manufacturing consensus online. Retrieved from https://ora.ox.ac.uk/objects/uuid:620ce18f-69ed-4294-aa85-184af2b5052e

---

Expected date of thesis defence

2023/24 SS – PEF

The Bachelor Thesis Supervisor

Ing. Miroslav Brabec

Supervising department

Department of Information Technologies

Electronic approval: 24. 9. 2023

doc. Ing. Jiří Vaněk, Ph.D.

Head of department

Electronic approval: 3. 11. 2023

doc. Ing. Tomáš Šubrt, Ph.D.

Dean

Prague on 15. 03. 2024

Declaration

I declare that I have worked on my bachelor thesis titled "Impact and Analysis of Computational Propaganda" by myself and I have used only the sources mentioned at the end of the thesis. As the author of the bachelor thesis, I declare that the thesis does not break any copyrights.

In Prague on 15.3.2024 _____

Acknowledgement

I would like to thank to leader of my work Ing.Brabec Miroslav for suggestions, comments, valuable advices and patience, he has given me in processing my thesis. I also would like to thank my parents who are supporting me whole my study life.

Impact and Analysis of Computational Propaganda

Abstract

This thesis deals with the discussion of computational propaganda, from the rise of the concept until its current dissemination through social media and algorithms, as well as its discussion and application in the fields of computer science and social science. The thesis also lists historical events in the past where computational propaganda has occurred and explains them, as well as describes the occurrence of computational propaganda in different countries, as well as its application in war, and the theoretical part includes the ethical issues involved in computational propaganda at the present time.

In the practical part, a method for recognizing false information from social bots is proposed by training Naive Bayes classifier. As well as analyzing the behavioral characteristics of social bots and their impact on users with the help of surveys in the context of differences in geographic regions and cultural environments.

Keywords: social media, fake news, social bot, algorithm, global issues, false information, machine learning, propaganda

Dopad a analýza počítačové propagandy

Abstrakt

Tato práce se zabývá diskusí o počítačové propagandě od vzniku tohoto konceptu až po jeho současné šíření prostřednictvím sociálních médií a algoritmů, jakož i jeho diskusí a aplikací v oblasti informatiky a sociálních věd. Práce také uvádí historické události v minulosti, kdy se počítačová propaganda vyskytla, a vysvětluje je, dále popisuje výskyt počítačové propagandy v různých zemích a její uplatnění ve válce a teoretická část zahrnuje etické otázky spojené s počítačovou propagandou v současnosti.

V praktické části je navržena metoda rozpoznávání falešných informací od sociálních botů pomocí trénování Naive Bayesova klasifikátoru. Stejně jako analýza charakteristik chování sociálních botů a jejich vlivu na uživatele pomocí průzkumů v kontextu rozdílů v geografických oblastech a kulturních prostředích.

Klíčová slova: sociální média, falešné zprávy, sociální bot, algoritmus, globální problémy, falešné informace, strojové učení, propaganda

Table of content

List of tables

List of pictures

List of charts

1 Introduction

Human society is entering a fully digitalized era. The development of new technologies is having a profound impact on all aspects, and the disruptive effects of emerging technologies in the political sphere are particularly far-reaching. The combination of data and algorithms, which enable the calculation and prediction of human behaviour, has a powerful ability to influence and shape public opinion. Intelligent algorithms have made it possible to control society on a massive scale, creating many phenomenal events. The manipulation of algorithmic technology in the 2016 U.S. presidential election is an important sample of practice. And in the run-up to the 2020 U.S. election and the 2024 campaign, the left and right wings continue to deepen the role of algorithmic technology in politics. Traces of algorithmic manipulation of public opinion have also been found in other political events, such as the UK's exit from the European Union.

The first part of the thesis will provide an introduction to the concept of computational propaganda and describe its application in various fields, geographic regions and historical events.

The second part of the thesis focuses on the identification and human impact of disinformation, which can be influenced by different contexts and produce different results.

Computational propaganda has an important place in policy debates, political elections, national security and political crises, and is likely to become even more important as objective facts are less influential in shaping public opinion than appeals to emotions and personal beliefs.

2 Objectives and Methodology

2.1 Objectives

The main objective of the thesis is to find and analyze the behavioral characteristics of social robots during social events.

Partial objectives:

-Analyzing the behavioral characteristics of social bots in different geographic regions.

-Investigating the impact of social bots on actual users.

-Propose methods to identify social bots.

2.2 Methodology

The methodology of the thesis is based on the processing and analysis of data, modelled by using Naive Bayes classifier and questionnaires for social populations, and the final conclusions are drawn based on the results.

3 Literature Review

3.1 Introduction to computational propaganda

With the development of the times, online social media has developed rapidly. Social media such as X (Twitter), Facebook, and Instagram have become an important social way for everyone to share their life, news and knowledge. The new features of these emerging social media have brought people surprises and changes, making communication with each other simpler and more convenient. However, benefiting from its convenient means of dissemination, especially the speed of dissemination of popular topics like a chain reaction, it also brings convenience to the proliferation of false information. False information spreads widely on the Internet and generates a high level of discussion, which can cause some people to choose to believe in such hot news without confirming the authenticity of the information, which has a negative impact on society. People usually rely on their own common sense or rely on some media with authoritative information to judge the authenticity of information, but this may be affected by timeliness and the different professional directions involved, resulting in the failure to clarify false information in a timely manner. The generation of false information is mainly related to computational propaganda.

Computational propaganda consists of two parts: computation and propaganda, which is an emerging concept composed of two traditional words. The term computation in the first half was first used in early definitions in 1375-1425, and the word computation today refers more to The work of designing computers, and the second half of the term propaganda originated in New Latin, 1710-20, short for congregātiō dē propāgandā fidē congregation for propagating the faith. The term has several definitions:

1. Information, ideas, or rumors that are intentionally widely disseminated to help or harm an individual, group, movement, institution, nation, etc.

2. A specific doctrine or principle promoted by an organization or movement

3. An organization that disseminates propaganda or sports. [1].

But computational propaganda now most often used in a political context, especially to refer to operations supported by a government or political group. When the same method is used for a business or product, it is often referred to as marketing or advertising. A social behavior that uses various symbols to spread certain concepts to influence people's thoughts and actions. Different forms of propaganda can also be distinguished by color: white propaganda is propaganda that openly identifies the source of information, black propaganda is propaganda that deliberately hides the true source of information, and grey propaganda is propaganda that does not directly identify the source of information.When the two words are combined, computational propaganda is formed.

In the era of paper media, information can only be spread unilaterally from the media to the masses. Due to the physical distance of communication, limited discussion of events can only be achieved. It was limited to a small scope, but then the advent of the Internet age brought huge changes to this phenomenon. Thanks to the ease of social media, the dissemination of information is not limited to the unilateral dissemination of information from the media to the masses; every user can be a source of information, and thanks to the global reach of the Internet, the discussion of events can break the limits of physical distance and is no longer limited to a certain regional area.

Originally propaganda was described as the neutral dissemination of information, while the changes brought about by the Internet era have led to its being increasingly considered pejorative and used to manipulate public opinion about it. Computational propaganda is an up-to-date, global propaganda method, which has been widely used in many countries and regions. As human society is gradually entering the era of comprehensive digitalization, thanks to the development of emerging technologies, through Artificial intelligence, algorithms, and big data are used to analyze and identify target groups, creating new conditions for effective dissemination of propaganda, and propaganda methods have gradually shifted from manual to automation, intelligence and anonymity.

3.2 Areas covered by computational propaganda

At present, research on computational propaganda includes two fields, computer science and social science. Computer science mainly focuses on the technical issues involved in computational propaganda, and its main purpose is to detect, identify, and track the existence of social robots. It is a collection of artificial intelligence, big data, algorithms and other technologies. Emerging technologies will greatly enhance the ability of computing propaganda. Technologies such as artificial intelligence chatbots, artificial intelligence synthesized speech and images, automated operation tools, machine deep learning, emotional computing tools supported by big data, and psychometric analysis will make efficient, autonomous, and accurate computational propaganda a reality.

Social science mainly regards computational propaganda as a To analyze and explore the various effects of computational propaganda on the economic, cultural, and political spheres, in order to use new technologies to manipulate public opinion. Technically, we can define computational propaganda as a collection of social media, autonomous agents, algorithms, and databases that are characterized by:

1. Automation: Allows the scale of propaganda strikes to expand

2. Scalability: Allows huge and fast coverage within the content distribution

3. Anonymity: Allows perpetrators to remain unknown.[2]

3.2.1 Computer science field

Supervised machine learning is a common method for identifying social robots. This method selects characteristic parameters from samples, establishes a discriminant function, and classifies unrecognized samples. It can effectively use data information to form a classification model that conforms to the characteristics. K nearest neighbour algorithm, linear regression, naive Bayes algorithm, support vector machine, and decision tree have all been used to identify social robots.

Botometer (formerly BotOrNot) is the first publicly available interface for detecting social bots from X (twitter). The system utilizes X's search interface to collect the 200 most recent posts and 100 most recent mentions of the account to be detected, and determines the likelihood that the account belongs to a social bot in terms of six categories of features: network, user, friends, events, content, and sentiment.

3.2.2 Political field

Computational propaganda is widely used in the fields of politics, economy and culture, among which the influence of computational propaganda on the political field is the main research direction at present. In politics, computational propaganda is often used to sway national elections, influence public opinion, incite public protests, and denigrate and attack other countries. One of the most important components of program advocacy is a bot, an automated program that can be used to perform simple and repetitive tasks. The work of bots is very useful, and most of the internal links that allow us to browse Wikipedia are created and maintained by bots. [3]

It is estimated that bots make up about 9% of the user base on social media platforms.[4] Social robots deployed on social media can automatically generate content and interact with humans on social media in an attempt to change human behavior. Political bots are social bots that are used for political manipulation. There are three common types of political bots:

1. Propaganda bots: used to spread true, half-truth or completely false information in large quantities.

2. Follow bots: Provide false popularity and trending for content that made by propaganda bots.

3. Barrier Bots: Disrupting ongoing conversations by diverting topics, etc.

Bots start their connections through interactions with users, these bots connect with each other in a separate network and get information from each other's interactions, which creates a layered effect in the bot network, through which bots interact with a network of nodes interactive. Politicians are increasingly using political bots to pretend to be more popular on social networks or attack rivals to create electoral advantage. [错误!未找到引用源。] Such computational propaganda is often in the hands of powerful and well-resourced political actors. Computational propaganda common manipulation tactics in politics consist of the following components:

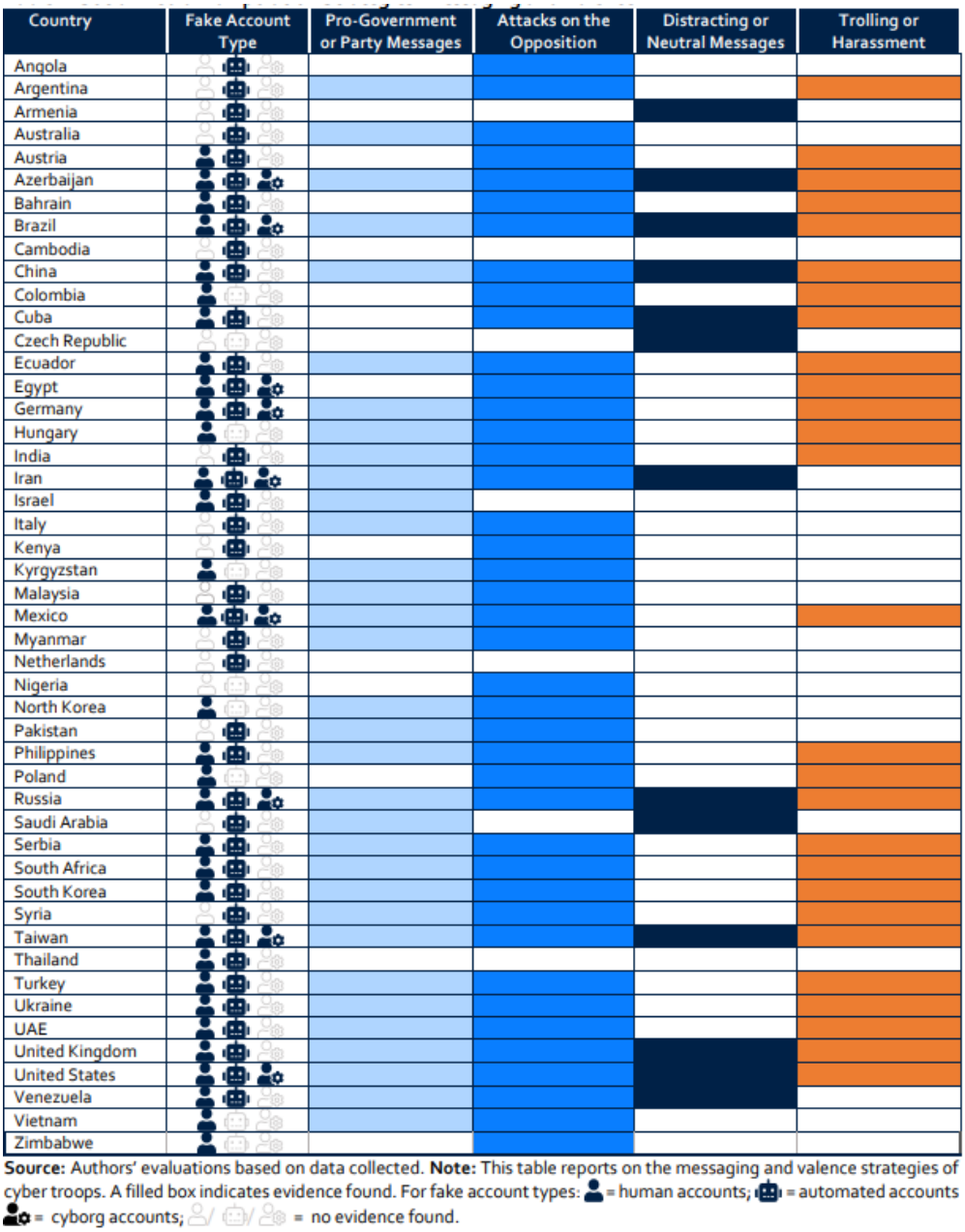
1. Deployment of political bots to attack foreign countries or political opponents on social media, to propagate inaccurate information about political opponents or to spread negative information about political opponents, and to drown out political opponents' issues in order to deter political opponents.

2. Post pro-government or political party messages, exaggerate follower numbers, use bot accounts to tweet in large numbers, and help generate false trends or political consensus.

3. Mix political bots and manipulated accounts to make interactions feel more real and avoid detection.

3.3 Representation of computational propaganda in different geographic areas

Computational propaganda is a global trend, but it has different manifestations for different regions, not just social robots, planned propaganda also includes fake accounts manipulated by humans, which method has higher utilization Depends on political, economic and cultural differences between regions. Looking at the figure1 below showing fifty selected countries and counted propaganda presence in that area, we can see that only a few places have all the fake account types present. The following will analyze the computational advocacy characteristics of different regions.



Source: Authors' evaluations based on data collected. **Note:** This table reports on the messaging and valence strategies of cyber troops. A filled box indicates evidence found. For fake account types: = human accounts; = automated accounts = cyborg accounts; = no evidence found.

Figure 1:Social Media Manipulation Strategies: Messaging and Valence (Bradshaw and Howard, 2018)[19]

3.3.1 United States

The most widely discussed case is the role played by bots used to spread disinformation and attack rivals during the 2016 election, with research showing that the use of political bots reached an all-time high at a pivotal moment in the election. 错误!未找到引用源。 Bot traffic accounts for more than 60% of all online traffic. [7] On Election Day, the share of social bots supporting Trump is about five times that of Hillary Clinton, with the top 100 bots posting about 500 tweets per day, or about 18% of all presidential election-related tweet traffic. [8]

By extracting a set of Trump-related tags or Clinton-related tags and building a botnet, as shown in the figure 2 below, Trump's botnet consists of 944 bots, while Clinton's botnet has only 264 bots, Trump's Pu's botnet is almost four times larger than Clinton's.[9] Automated and fake accounts on Twitter and Facebook promoted allegations of corruption against Hillary Clinton and drove a flood of junk news coverage of events such as her and pedophilia. [10]

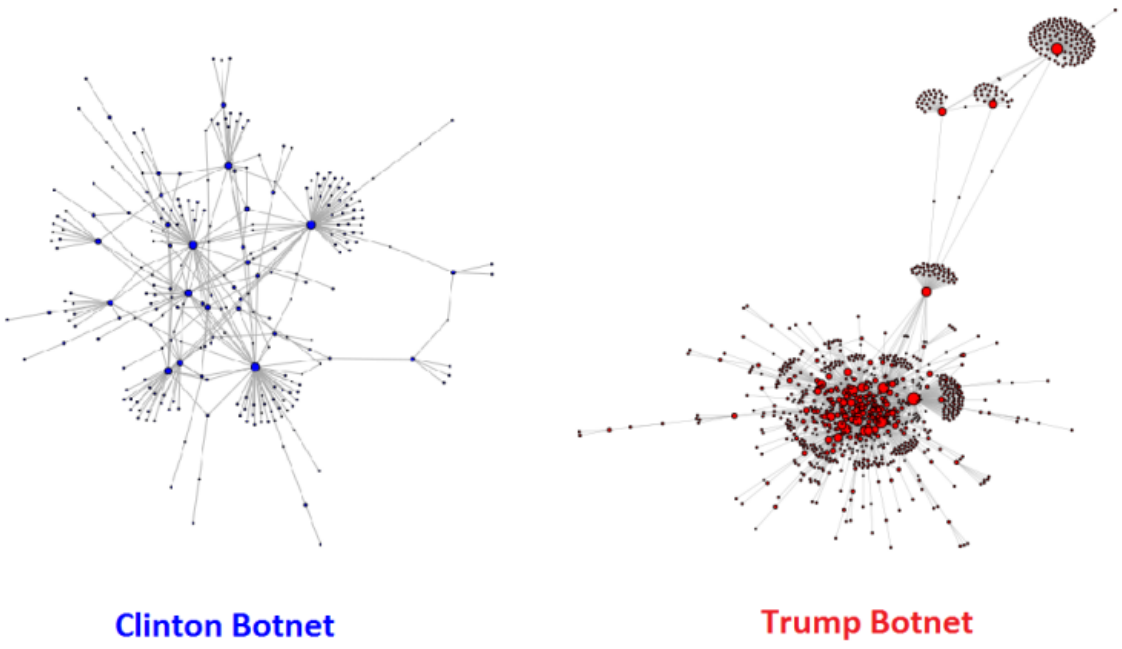


Figure2: Comparing the Largest Botnet within the Retweet Networks for Trump vs Clinton Related Hashtags.(Woolley and Guilbeault, 2017)[9]

3.3.2 United Kingdom

The massive use of computational propaganda in the UK came during the Brexit referendum, during which social robots were heavily used to support the idea that the UK should leave the EU. In June 2016, 77,000 bots signed an online petition for a second Brexit referendum. [11] There are 13,493 twitter bots forming the online campaign in support of leaving the EU. [12] During the week of 5-12 June 2016, accounts that tweeted more than 100 times using event-related hashtags generated 32% of all Twitter traffic about Brexit. This number of posts was generated by less than 2,000 of the more than 300,000 users, meaning that less than 1% of accounts generated nearly one-third of all content. Ranking the traffic generated on the topic of "Brexit" almost certainly shows that 7 of the top 10 accounts are bots.[13]

3.3.3 Ukraine

In 2015, Ukrainian media ain.ua published an interview on the "Akmetov's Robot" incident, in which a former staff member claimed to have managed an online business under one of Ukraine's largest oligarchs, Rinat Akmetov. In the interview, he recounts the instructions he and his colleagues receive every day: where to post what kind of comments.[14] We usually interpret computational propaganda as the autonomous action of social bots, but the “bots” here are actually fake accounts centrally managed by paid service personnel.

For interviewees within the industry, one stakeholder described that he usually gets political clients from intermediaries, is tasked with distributing certain messages, neutralizing negative news or attacking clients' competitors, and company bosses often target students when hiring teams , which typically post up to 200 comments a day, and received about $100 a week during the 2012 parliamentary election campaign. [15] In the wake of the official 2016 investigation into the tragic crash of Malaysia Airlines flight MH17, one Twitter user discovered that whenever a person tweeted the hashtag "#MH17" in Russian, a bot would join the conversation, And reply to a link to a fake article questioning the findings. [15]

3.3.4 China

The computational propaganda project of the Internet Research Institute of Oxford University found that no computational propaganda phenomenon was found in China. There is little evidence of automation in comments on posts, and most of the comments left by the most-posted users in this dataset are generic attacks on other users, and are scattered across comments on multiple posts.

In contrast, a large amount of anti-China automated propaganda was found in the hashtags related to Chinese politics on Twitter. The researchers collected 1,177,758 tweets, of which nearly 30% were from the top 100 accounts with the most tweets, No pro-China accounts were found among the 100 accounts, and half were automated accounts posting anti-China content. [16] The top 100 account-related tags with the highest posting volume are shown in the table 1 below. The likely reason why China has not launched computational propaganda is because Chinese propaganda has always been dominated by state media, and the technical capabilities needed to automate propaganda are beyond the capabilities of traditional media workers. Anti-Chinese content on Twitter has had little effect, targeting only a small group of people such as Chinese and international students settled overseas.

	Number of accounts	Number of posts	Percentage of posts in dataset
Anti-Chinese-state bots			
1989 group	22	117,578	10%
Pan-Asia group	22	44,678	4%
Independent anti-Chinese-state bots	5	7,969	0.68%
Both anti-Chinese-state and commercial content	1	1,090	0.09%
Other political bots			
Professional news bots	10	39,239	3%
“Fake news” bots	4	10,213	0.87%
Other non-political bots			
Commercial bots	8	34,860	3%
Job bots	6	8,592	0.73%
Other bots (non-political)	4	6,620	0.56
Account suspended			
Account suspended	18	64,170	5%
TOTAL	100	335,009	28.44%

Table 1: Top 100 highest-posting accounts within China related hashtags (Bolsover, 2017)[16]

3.4 Terrorism-related computing propaganda

A large number of robots were also found in terrorist propaganda activities. From 2012 to 2016, ISIS uploaded videos to YouTube and then cross-posted them to Facebook and X for dissemination. They created thousands of X accounts, including human fans and bots, and used this social media to recruit followers one-on-one, creating labels about terrorist attacks and gloating about the misfortune. The anti-terrorism organization ‘Ghost Security Group’ has stated that ISIS uses almost all online social media to communicate and share its propaganda, and calculates the propaganda to achieve the maximum effective spread. When researchers investigated activity on x in December 2014, Isis was operating between 46,000 and 70,000 accounts.

3.5 Computational propaganda in war

For a long time, the role of the media in international conflicts has been to serve as a weapon and a tool of war, to conduct propaganda and psychological warfare, and during the First and Second World Wars, both opposing sides used various media as a means of boosting their own popularity and morale, and of stigmatizing the enemy. For example, during the World War, propaganda was mainly broadcast and paper-based; internally, newspapers, posters and radio speeches were used for propaganda, while externally, airplanes and balloons from various countries were used to drop pamphlets on the front lines of the war, which generated a great deal of fake news during the war.

3.5.1 World war I

In 1928, Arthur Ponsonby published the book ‘Falsehoods in Wartime’. The book used a large number of documents to examine the ins and outs of fake news during World War I, including many famous international fake news at the time, such as the Allied Powers' report ‘Germany used the bodies of soldiers to make grease, pig feed and fertilizer’ , it was later discovered that the propagandists deliberately translated the German word 'Kadaver' (corpse, used only to refer to animal carcasses) as 'corpse' when reporting the news.

Germany, as a core member of the Axis Powers, is indoctrinating its domestic people with messages such as "Reuters is the war lie maker" and "Nirthecliffe is the minister of lies". From all aspects, the war during World War I was Current news is no different from propaganda, and the truth has become a victim of war. Objective and true information is difficult to find, but it is limited by the technology at the time. Most of the international fake news in this period was spread through paper. Both scope and depth are limited.

3.5.2 World war II

On August 31, 1939, Adolf Hittler created a pretext for war and sent the Nazi SS to launch a False flag attack on the Sender Gleiwitz Radio Station in Gleiwitz. He called the incident a Polish attack on Germany to justify Nazi Germany's invasion of Poland. It became the trigger for World War II. Later, Hitler and Goebbels integrated foreign propaganda into the wartime management system.

Goebbels served as the Minister of Mass Education and Propaganda in Nazi Germany. He attached great importance to the role of propaganda. For example, many of his famous sayings: "A lie repeated a thousand times becomes the truth" "News is a weapon of war" "News" The purpose is to aid the war effort rather than provide information', which are fully reflected in his propaganda ideas and strategies.

Like Nazi Germany, other countries participating in World War II also used various ways to spread wartime fake news. For example, the secret agency PWE (Political warfare executive) established by Britain in World War II was set up for the purpose of undermining the morale of the countries allied with Nazi Germany, and its main activities were to carry out covert propaganda in the occupied countries, distributing leaflets and underground Its main activities included secret propaganda in the occupied countries, distribution of leaflets and underground publications, rumor campaigns, and the production of forgeries in order to lower the morale of the enemy. In short, international fake news became an important part of State propaganda during the two world wars, and the production and dissemination of fake news as a means of attacking the enemy and promoting oneself became the norm in wartime journalism.

3.5.3 Russian invasion of Ukraine

On February 24, 2022, Russia began to invade Ukraine. This invasion was the largest war to hit a European country since World War II. It has already caused hundreds of thousands of casualties on both sides. Computer technology is developing rapidly in the 21st century. Computational propaganda has been widely used in this war. The degree of propaganda is several orders of magnitude higher than the fake news propaganda in World War I and World War II. The Internet has become has become the main battlefield in the war of public opinion.

In the early days of the war, the Russian media made a big deal out of the fact that Ukrainian President Zelensky had long since fled Kiev, and that Zelensky, in an attempt to prove that he was defending Kiev along with the Ukrainian people, had posted several selfie videos with the city of Kiev in the background, only to be challenged by the Russian media that the background behind him had been synthesized by a computer.

The BBC also exposed a Russian propaganda campaign involving the dissemination of false information about Ukraine on TIKTOK, such as false accusations that senior Ukrainian officials and their families purchased luxury cars and villas abroad after Russia invaded Ukraine. TIKTOK said it had deleted it More than 12,000 fake accounts originating from Russia. As a major military conflict in the Internet era, computational propaganda has become a key force that can influence the battlefield of public opinion. This can help obtain international aid and support by occupying moral advantages, and can also achieve the effect of isolating and disintegrating opponents.

3.6 Ethics and Challenges of Computational Propaganda

Human society has entered a period of transition to digitalization. The development and application of information technology not only brings about industrial changes at the tool level, but also promotes human society to accelerate its entry into the digital society. The current rise and development of computational propaganda technologies such as social media robots is considered an inevitable result, and changes are accompanied by emerging governance challenges.

Computational propaganda has caused countries, users, companies and other parties to encounter an unprecedented crisis, conflicts in the public sphere continue, and public rationality is gradually lost. The early detection method is to classify a single node in the network as malicious or legitimate, but the disadvantage of this method is that it is difficult to detect complex robots. The current research trend does not focus on a single account, but on account groups as a Overall and analysis of key characteristics followed by modeling.[17]

This requires the participation of social platforms. After a slow start, several platforms have made a lot of self-discipline reactions, especially Google, YouTube, Twitter and Facebook. [18] A series of problems brought by technological development is that the governance of computing propaganda is imminent. With the development of globalization, power has begun to be redistributed, and the form of governance has begun to be a global model of joint governance by governments, citizens, institutions, and enterprises.

4 Practical Part

4.1 Naive Bayes classifier

The primary research uses our collected dataset to train the Naive Bayes classifier algorithm model, classifies the data as training and test data, and finally evaluates the model.

Bayes classifier is a general term for a class of classification algorithms, which are all based on Bayes' theorem and are collectively known as Bayes classifier (Bayes classifier is supervised learning, the so-called supervised learning that is, from the known feature information in the sample data to speculate on the possible outputs in order to complete the classification, and vice versa clustering problems are known as unsupervised learning). One of the most widely used classification algorithms, the Naive Bayes approach is a simplification of the Bayes algorithm, where attributes are assumed to be conditionally independent of each other given a target value. This means that no attribute variable has a greater or lesser weight for the decision outcome. Although this simplification reduces the classification effect of the Bayes classification algorithm to a certain extent, it greatly simplifies the complexity of the Bayes method in real-world applications. The formula is:

$$P(A|B)=P(B|A)P(A)/P(B)$$

P(A|B): the probability of event A conditional on B. In Bayes' theorem, the conditional probability is also referred to as the posterior probability, i.e., our reassessment of the probability of event A after event B has occurred.

P(B|A): the probability of event B conditional on A.

P(A) and P(B) are known as prior probabilities (also known as marginal probabilities), i.e., an inference we make about the probability of event A before event B occurs (without taking into account any event B aspects)

However, when we use only Bayes classifier for the process, if the sample size is not sufficiently large it will result in an error with a computed probability of 0. This error is caused by insufficient training, which will cause the quality of the classifier to be greatly reduced. In order to solve this problem, we introduce Laplace Smoothing, which has the simple effect of adding 1 to the count of all divisions under each category, so that if the training sample set is sufficiently large, it does not affect the results and solves the case of the error with a frequency of zero. After the introduction of Laplace Smoothing the Naive Bayes formula is:

$$P(w'|positive) = \frac{\text{number of reviews with w' and y = positive } + \ \alpha}{\text{N } + \ \alpha * K}$$

Figure 3: Laplace Smoothing

Firstly we process the collected dataset, the dataset itself has been manually completed with label categorization, the data labels are classified into 6 categories based on the degree of falsity, which are patently-false, false, barely-true, half-true, mostly-true, and true. Next, we remove emoticons in the content, images, the hashtags #, @, and web links, and other irregular text content, which can cause problems in text content recognition. After data processing, we get 18000 valid data and the data cases are as follows:

ID	author	claim	label
37	John Chiang	"A majority of Americans now live in states where they have decided to legalize cannabis,"" including recreational and medical marijuana."	true
3461	Allen West	"After Patrick Murphy’s charges were dropped, Murphy’s father gave the prosecutor a huge campaign donation."	barely-true
8601	Paul Krugman	"A recent report from the president's deficit commission was ""completely empty"" on controlling health care costs."	false

Table 2: data cases

We use python for data processing and it is well known that python has very wide application in machine learning. Firstly take out the content column and then set the criteria for determination, we define false content as 0 and true content as 1.

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.naive_bayes import MultinomialNB
from textblob import TextBlob
import nltk
nltk.download('punkt')


data = pd.read_csv('E:/fake_test/fact_checking_train.csv', sep='\t')


content = data['claim']


data.loc[data.loc[:,'label'] == 'true', 'real_or_bot'] = 1
data.loc[data.loc[:,'label'] == 'mostly-true', 'real_or_bot'] = 1
data.loc[data.loc[:,'label'] == 'half-true', 'real_or_bot'] = 1
data.loc[data.loc[:,'label'] == 'false', 'real_or_bot'] = 0
data.loc[data.loc[:,'label'] == 'pants-fire', 'real_or_bot'] = 0
data.loc[data.loc[:,'label'] == 'barely-true', 'real_or_bot'] = 0

```

Figure 4: Set the criteria for determination

Then we select the stop words, there are many words in the text content that have no meaning for the data analysis, and stop them will not affect the results, we stop these meaningless words to help reduce the amount of calculations.

```
stopwords=[]
with open('E:/stopwords.txt','r') as f:
    lines=f.readlines()
    #print(lines)
    for tmp in lines:
        line=tmp.strip()
        #print(line)
        stopwords.append(line)
stopwords = list(set(stopwords))
#print(stopwords)
```

Figure 5: Set the stopwords

Next we need to further process the text content into a standardized format. We use the Textblob library for text processing, which is used to separate the words in each sentence.

```
35  content_list = []
36
37  for tmpp in content:
38      #print(tmpp)
39      blob = TextBlob(tmpp)
40      words = blob.words
41      token_text = ''.join(words)
42
43      content_list.append(token_text)
44  #print(content_list[:5])
```

Figure 6: Set standardized format.

Then we transform the stop words into word frequency matrix and calculate the number of occurrences of each word by the function. Then we prepare the training set and test set, we divide the first 10,000 data in the dataset into the training set, and the remaining 8,000 data in the back into the test set.

```
47  con = CountVectorizer(stop_words = stopwords)
48
49  x = con.fit_transform(content_list)
50
51  #print(X.toarray())
52
53  x_train = x.toarray()[:10000,:]
54  y_train = data['real_or_bot'][:10000]
55
56
57  x_test = x.toarray()[10000:,:]
58  y_test = data['real_or_bot'][10000:]
```

Figure 7: Train and test data set.

Finally, construct the Naive Bayes classifier algorithm to train the data and predict the data, at this time Laplace Smoothing is introduced and then the model is evaluated and the obtained prediction correctness is 81.3%

```
60    mb = MultinomialNB(alpha=1)
61    mb.fit(x_train,y_train)
62
63    y_pre = mb.predict(x_test)
64
65    print(mb.score(x_test,y_test))
66
```

Figure 8: Train algorithm.

```
PS C:\Users\97800\Desktop> & C:/Users/97800/AppData/Local/Microsoft/WindowsApps/python3.11.exe c:/User
s/97800/Desktop/test.py
[nltk_data] Downloading package punkt to
[nltk_data]     C:\Users\97800\AppData\Roaming\nltk_data...
[nltk_data]   Package punkt is already up-to-date!
0.812576453
PS C:\Users\97800\Desktop>
```

Figure 9: Model Prediction Accuracy.

4.2 Survey

The secondary research was completed by distributing questionnaires through the Internet. This survey was to understand people's cognition, opinions and coping methods of social robots in different countries and regions. To build the survey, we used the service 'survey circle' 'survey swap' 'wenjuanxing' which allows the creation of internet surveys. The authors used English and Chinese as the survey languages to distribute to respondents in different regions. In order to show the differences between different language audiences differences, the author will translate the Chinese results into English and compare the results with English questionnaire respondents.

Due to the closed Internet environment in the Chinese region, Internet users in the Chinese region use different online social media, so the authors decided to link online social media that offer similar functions:

1. Weibo-X, Weibo is a social network media website that shares and disseminates short information like X(Twitter).

2. Tieba-Reddit, tieba and Reddit provide similar community services. Users can post posts and participate in discussions based on different topics.

3. Douyin-Tiktok, both are short video apps developed by Bytedance, just distinguish between Chinese and international versions.

4. Xiaohongshu-Instagram, the services provided by these two social applications are geared towards photo sharing.

5. Bilibili-youtube, Bilibili is a website that allows users to upload homemade videos. Its service content is similar to Youtube.

6. Zhihu-Quora, Zhihu is an online question and answer website that provides services similar to Quora.

The questionnaire survey adopts a non-probability sampling method, mainly because the respondents are not selected based on the principle of random, but based on whether they are interested in the survey topic. At the end of the questionnaire survey, a total of 186 valid responses were collected from many regions around the world, including 114 responses to the Chinese questionnaire and 72 responses to the English questionnaire. The questionnaire consisted of 14 questions. The following figure shows the regional distribution of the respondents.

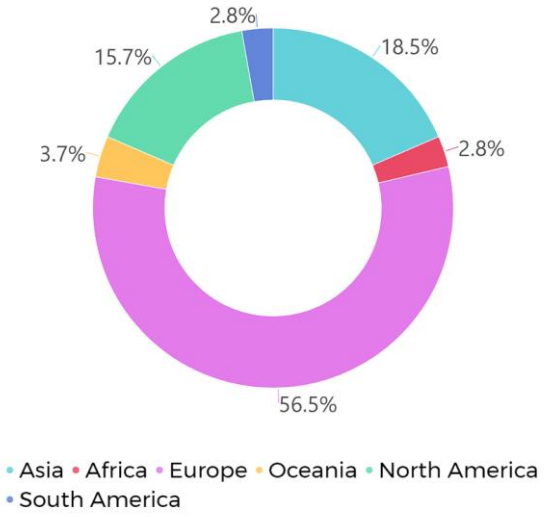


Chart 1: Geographical distribution of respondents.

Q1: Do you know what is social bot?

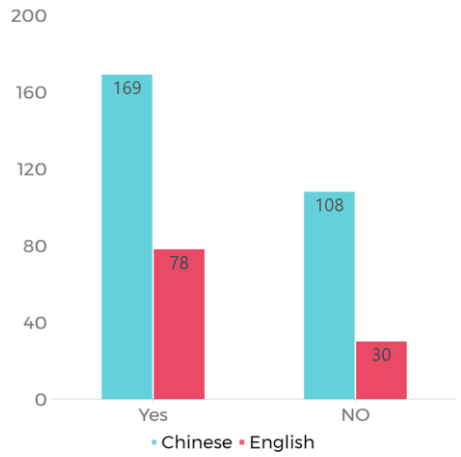


Chart 2: Respondents' knowledge of social bot.

Respondents first answered the question about their basic knowledge of social robots, i.e., whether they were clear about the definition of social robots. 72% of respondents to the English-language questionnaire said they were sure, with a slightly lower proportion of Chinese respondents choosing ‘Yes’ to this question, at 61%. Respondents who chose the NO option would withdraw from the questionnaire.

Q2: Are you able to recognize whether a social account is a bot or a real person?

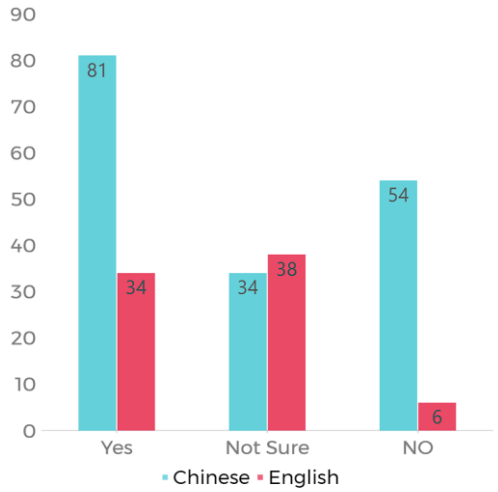


Chart 3: Whether respondents can recognize social bot.

This question was designed to investigate respondents' ability to recognize bot accounts, whereby respondents judged whether or not the controller of the account was a real person based on their own experience. There was a large difference in the responses to this question between the two languages, with 43% of English respondents choosing ‘Yes’, 8% of English respondents choosing ‘No’ and 49% choosing ‘Not sure’, and 48% of Chinese respondents choosing ‘Yes’, 32% of Chinese respondents choosing ‘No’, and 20% of Chinese respondents choosing ‘Not sure’. A higher percentage of Chinese respondents could not distinguish the authenticity of the accounts. Respondents who chose the NO option would withdraw from the questionnaire.

Q3: Have you ever interacted with a bot account? For example: reply, like or shae content?

This question is related to the previous one in that respondents are more likely to recognize the real identity of the person they are interacting with if they can identify the account being manipulated by the bot. The majority of English respondents had interacted with a bot account 76% of the time, with 10% and 14% choosing 'No' and 'Maybe' respectively, compared to 50%, 39% and 11% of Chinese respondents. This may be due to the fact that more Chinese respondents were unable to recognize a bot account in the previous question.

Q4: On which social media or apps do you usually encounter bot accounts?

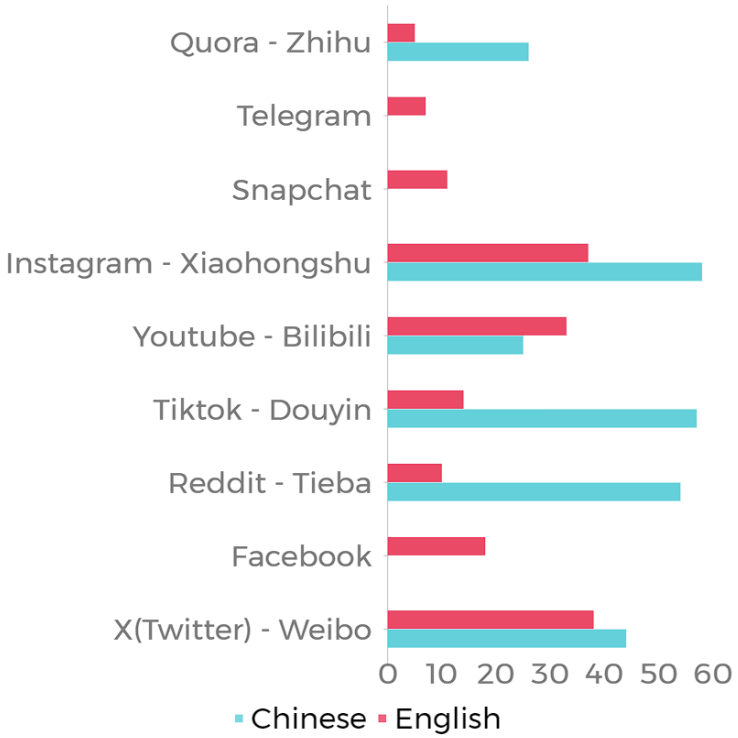


Chart 4: Social media where respondents believe social bot appear the most.

This question asked respondents to answer in which online social media they see bot accounts more frequently, and the three most frequent choices in the English respondents were X, Youtube, and Instagram. Tieba, Xiaohongshu, and Douyin were the most frequent choices in Chinese social media.

Q5: Do you think social bots have positive or negative impact on online discussions or social interactions?

Half of the English respondents have a positive attitude towards this issue, and 1/4 of the respondents have opposing views. Chinese respondents were slightly more negative with 43%, positive with 29%, and neutral with 28%.

Q6: What do you think is the main role of bot accounts on social media?

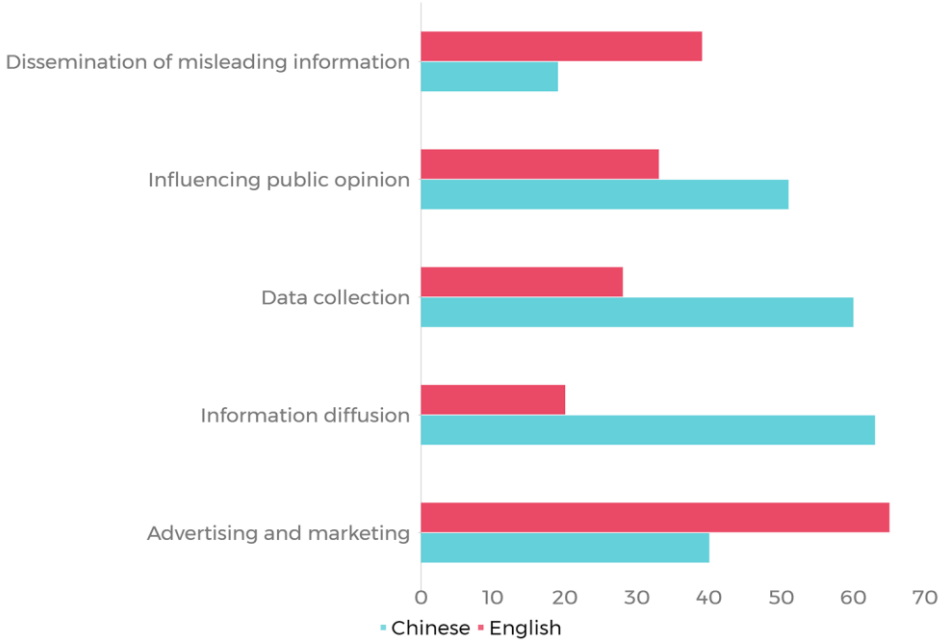


Chart 5: The main behaviors of social bots as perceived by the respondents.

The three roles with the highest number of selected by English respondents in this question were Advertising and marketing, Influencing public opinion, and Dissemination of misleading information. The Chinese consider Information diffusion, Data collection, Influencing public opinion are the main role.

Q7: Are you concerned about bot accounts spreading fake information or misleading propaganda on social media?

This question was asked as an extension of question 6, and respondents in both languages expressed the same opinion, with most having concerns about bot accounts spreading false information.

Q8: Are you concerned that bot accounts may influence public opinion on a topic or issue?

A higher proportion of English respondents chose the affirmative option for this question than Chinese respondents, probably because there are fewer historical cases of this kind in China.

Q9: Do you think bot accounts are disruptive to real users?

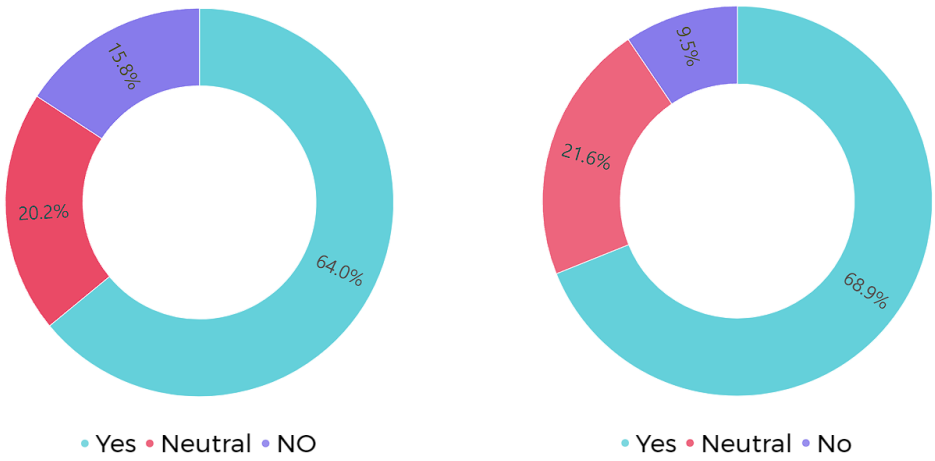


Chart 6: Respondents' perceptions of social bot.

The same percentage of respondents in both languages answered in the affirmative, suggesting that bot accounts are interfering with the behaviour of real users around the world.

Q10: Are you concerned about your online privacy or data being collected by bot accounts?

This question, which was used to survey respondents about the importance of privacy and data security, produced a large difference in responses between the two languages, with more than 80% of English-speaking respondents expressing concern, compared to 54% of Chinese-speaking respondents. One of the reasons for this difference may be due to the impact of the Facebook data breach a few years ago.

Q11: How much do you trust the content of the bot account?

All respondents rated the content credibility of bot accounts low, which suggests that the content generated by bot accounts is not valuable and that they are mostly doing the same thing over and over again.

Q12: Do you support limiting or banning certain types of bot accounts on social media?

In this question, respondents expressed their opinion on whether or not to restrict bot accounts, with roughly half of the respondents in both languages in favour, and a somewhat higher percentage of English-speaking respondents choosing neutrality for the remaining option.

Q13: Do you think bot accounts should be clearly identified in their profiles or postings?

Nearly all respondents agreed that there is a need to mark non-personalized accounts.

Q14: Who do you think should manage or regulate the activities of bot accounts?

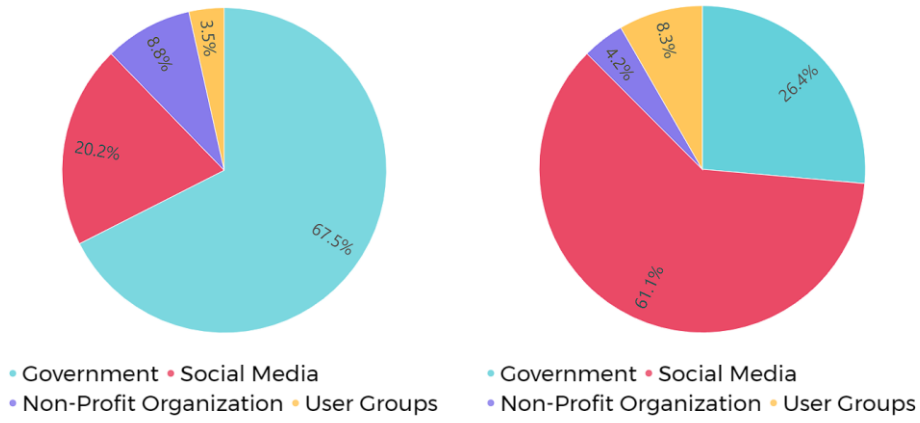


Chart 7: Respondents' perceptions of regulatory responsibility for social bot.

This question produced interesting differences in choices between the two language respondents, which is not surprising; China's system is one in which the government dominates economic activity, so respondents believe the government should take more responsibility. Most of the online social media used by English-speaking respondents are global service providers, and it is difficult for a single government to fully regulate them, so respondents believe that social media itself should regulate bot accounts.

5 Results and Discussion

Focusing on false message identification for bots, we can see that by training the data, the model built using Naive Bayes has a high recognition accuracy, but before that it needs to manually process a large amount of data, which is obviously not going to be a common way to manage social bots in the future. For more efficient and intelligent identification of social bots, we need to rely on more advanced algorithms through deep learning, while social bots will also iterate by learning from human behaviour, which is a confrontation that will last for the foreseeable future.

We hypothesized that the behavioural characteristics of social bots differ under different geographical conditions, and the findings confirm this hypothesis when dealing with a survey that compared to the English questionnaire respondents who found that social bots were more often used for advertising and marketing, dissemination of misleading information, Chinese questionnaire respondents found that social bots were more often used for information diffusion and data collection in Chinese Internet social media. These differences in results further influence the perception of social bots by respondents in different regions, in the case of past history. In historical cases, more social bot activity can greatly increase negative public attitudes.

Throughout all the survey items, we found that even though there are differences due to different environments, cultures, languages, and regions, people are generally affected by the negative impacts of social bots, which are exacerbated by more social events, and also enrich the public perception of social bots and computational propaganda, which is a new challenge that all humans are facing.

6 Conclusion

Computational propaganda is the product of traditional propaganda in the age of artificial intelligence and a new tool to change the geopolitics of information. In the first part of the thesis, we have already analysed the areas involved in computational propaganda and the different characteristics it has shown in different countries, but they are all mainly oriented to the political and commercial levels to attack political and commercial opponents. The beginnings of computational propaganda are also visible in historical events, as a way of disseminating information that has evolved with the development of technology until it has now become an important means of attack in the cultural sphere.

The use of computational propaganda in the war on propaganda is a further argument for this point of view. From the paper propaganda of World War I to the radio propaganda of World War II, and to the present day, with the nature of the Internet and the development of the field of computer science, which has reached alarming proportions, computational propaganda is pervasively intruding into the lives of the general public, and how to cope with it is a serious challenge that we face at the present time.

In the second part of the thesis, we deal with the other objectives of this work. For the detection of social bots, the use of machine learning is currently the dominant approach, and its detection is achieved with high accuracy by Naive Bayes classifier, but at the same time we note the tediousness of the work when it comes to the preprocessing of the information. Through the Internet research, we dealt with the rest of the objectives of this work, social bots are generally having a negative impact on all people and concerns about them are gradually growing day by day. And we concluded that social bots active on the Chinese Internet focus on data collection and diffusion of information in a more covert manner than the behaviour of social bots on the English-speaking Internet around the world, which is biased towards advertisement and marketing and spreading of false news.

The rise of the phenomenon of computational propaganda affects people's participation in public life, and accurately identifying and capturing social bots, and grasping their behavioral characteristics and motivations are crucial for us to better govern them. The public's level of awareness and ability to recognize computational propaganda and fake news should be raised as much as possible to minimize the negative impact.

7 References

[1] Dictionary [Online]
https://www.dictionary.com/browse/propaganda

[2] Rahmawati, Nurul Aulia. "COMPUTATIONAL PROPAGANDA ON TikTok AS THE 21st CENTURY PROPAGANDA MODEL." International Review of Humanities Studies 7.1 (2022): 21.

[3] Woolley, Samuel C., and Philip N. Howard, eds. Computational propaganda: Political parties, politicians, and political manipulation on social media. Oxford University Press, 2018.

[4] Schnader, Holly. "The evolution of computational propaganda: Trends, threats, and implications now and in the future." (2018).

[5] Howard, Philip N., Samuel Woolley, and Ryan Calo. "Algorithms, bots, and political communication in the US 2016 election: The challenge of automated political communication for election law and administration." Journal of information technology & politics 15.2 (2018): 81-93.

[6] Bessi, Alessandro, and Emilio Ferrara. "Social bots distort the 2016 US Presidential election online discussion." First monday 21.11-7 (2016).

[7] Condliffe, J. "Over 60 Percent of Internet Traffic Is Now Driven By Bots." Gizmodo", https://www. gizmodo. com. au/2013/12/over-60-percent-of-internettraffic-is-now-driven-by-bots 9 (2013): 2021.

[8] Howard, Philip N., Bence Kollanyi, and Samuel Woolley. "Bots and Automation over Twitter during the US Election." Computational propaganda project: Working paper series 21.8 (2016).

[9] Woolley, Samuel, and Douglas Guilbeault. "Computational propaganda in the United States of America: Manufacturing consensus online." (2017).

[10] Howard, Philip N., Samuel Woolley, and Ryan Calo. "Algorithms, bots, and political communication in the US 2016 election: The challenge of automated political communication for election law and administration." Journal of information technology & politics 15.2 (2018): 81-93.

[11] BBC, "EU referendum petition hijackedby bots" http://www.bbc.com/news/technology-36640459.

[12] Bastos, Marco T., and Dan Mercea. "The Brexit botnet and user-generated hyperpartisan news." Social science computer review 37.1 (2019): 38-54.

[13] Howard, Philip N., and Bence Kollanyi. "Bots,# strongerin, and# brexit: Computational propaganda during the uk-eu referendum." arXiv preprint arXiv:1606.06356 (2016).

[14] AIN.UA Kak rabotayut internet-trolli i kak ikh raspoznat: intervyu s byvshim “akhmetovskim botom” (How Internet trolls work and how one can recognize them: an interview with a former “Akhmetov’s bot”)
Как работают интернет-тролли и как их распознать: интервью с бывшим «ахметовским ботом» – AIN.UA

[15] Zhdanova, Mariia, and Dariya Orlova. "Computational Propaganda in Ukraine: Caught between external threats and internal challenges." (2017).

[16] Bolsover, Gillian. "Computational propaganda in China: An alternative model of a widespread practice." (2017).

[17] Martino, Giovanni Da San, et al. "A survey on computational propaganda detection." arXiv preprint arXiv:2007.08024 (2020).

[18] Taylor, Emily, and Stacie Hoffmann. "Industry responses to computational propaganda and social media manipulation." (2019).

[19] Bradshaw, Samantha, and Philip N. Howard. "Challenging truth and trust: A global inventory of organized social media manipulation." The computational propaganda project 1 (2018): 1-26.

Appendix

The Impact of Social Bots(bot account) on Users

This questionnaire was used to investigate the impact of social bots on users and consisted of 15 questions, thank you for your answers!

* Indicates required question

1. Where are you from? *

   Mark only one oval.

   - ◯ Asia
   - ◯ Africa
   - ◯ Europe
   - ◯ Oceania
   - ◯ North America
   - ◯ South America

2. Do you know what is social bot? *

   Mark only one oval.

   - ◯ Yes
   - ◯ No

3. Are you able to recognize whether a social account is a bot or a real person? *

   Mark only one oval.

   - ◯ Yes
   - ◯ NO
   - ◯ Not sure

4. Have you ever interacted with a bot account? For example: reply, like or shae content *

Mark only one oval.

- ◯ Yes
- ◯ No
- ◯ Maybe

5. On which social media or apps do you usually encounter bot accounts? *

Check all that apply.

- ☐ X(Twitter)
- ☐ Facebook
- ☐ Reddit
- ☐ Tiktok
- ☐ Youtube
- ☐ Instagram
- ☐ Weibo
- ☐ Snapchat
- ☐ LinkedIn
- ☐ Telegram
- ☐ Other: _____

6. Do you think social bots have positive or negative impact on online discussions or social interactions? *

Mark only one oval.

- ◯ Positive
- ◯ Negative
- ◯ Neutral

7. What do you think is the main role of bot accounts on social media? *

Check all that apply.

- [ ] Advertising and marketing
- [ ] Information diffusion
- [ ] Data collection
- [ ] Influencing public opinion
- [ ] Dissemination of misleading information
- [ ] Other: _____

8. Are you concerned about bot accounts spreading fake information or misleading propaganda on social media? *

Mark only one oval.

- ◯ Yes
- ◯ No
- ◯ Neutral

9. Are you concerned that bot accounts may influence public opinion on a topic or issue? *

Mark only one oval.

- ◯ Yes
- ◯ No
- ◯ Neutral

10. Do you think bot accounts are disruptive to real users? *

Mark only one oval.

- ◯ Yes
- ◯ No
- ◯ Neutral

11. Are you concerned about your online privacy or data being collected by bot accounts? *

Mark only one oval.

- ◯ Yes
- ◯ No
- ◯ Neutral

12. How much do you trust the content of the bot account? *
(1 being very distrustful and 5 being very trusting)

Mark only one oval.

	1	2	3	4	5	
very	◯	◯	◯	◯	◯	very trusting

13. Do you support limiting or banning certain types of bot accounts on social media? *

Mark only one oval.

- ◯ Yes
- ◯ No
- ◯ Neutral

14. Do you think bot accounts should be clearly identified in their profiles or postings? *

Mark only one oval.

- ◯ Yes
- ◯ No
- ◯ Neutral

15. Who do you think should manage or regulate the activities of bot accounts? *

Mark only one oval.

- ◯ Government
- ◯ Social Media
- ◯ Non-Profit Organization
- ◯ User Groups
- ◯ Other: _____