

Article

# Understanding Potential Cyber-Armies in Elections: A Study of Taiwan

Ming-Hung Wang <sup>1,\*</sup>, Nhut-Lam Nguyen <sup>1</sup>, Shih-chan Dai <sup>2</sup>, Po-Wen Chi <sup>3</sup> and Chyi-Ren Dow <sup>1</sup>

<sup>1</sup> Department of Information Engineering and Computer Science, Feng Chia University, Taichung 40724, Taiwan; lam@mail.fcu.edu.tw (N.-L.N.); crdow@mail.fcu.edu.tw (C.-R.D.)

<sup>2</sup> Department of Political Science, University of Massachusetts Amherst, Amherst, MA 01003, USA; shihchandai@umass.edu

<sup>3</sup> Department of Computer Science and Information Engineering, National Taiwan Normal University, Taipei City 11677, Taiwan; neokent@gapps.ntnu.edu.tw

\* Correspondence: mhwang@mail.fcu.edu.tw

Received: 17 February 2020; Accepted: 10 March 2020; Published: 13 March 2020



**Abstract:** Currently, online social networks are essential platforms for political organizations to monitor public opinion, disseminate information, argue with the opposition, and even achieve spin control. However, once such purposeful/aggressive articles flood social sites, it would be more difficult for users to distinguish which messages to read or to trust. In this paper, we aim to address this issue by identifying potential “cyber-armies/professional users” during election campaigns on social platforms. We focus on human-operated accounts who try to influence public discussions, for instance, by publishing hundreds/thousands of comments to show their support or rejection of particular candidates. To achieve our objectives, we collected activity data over six months from a prominent Taiwan-based social forum before the 2018 national election and applied a series of statistical analyses to screen out potential targets. From the results, we successfully identified several accounts according to distinctive characteristics that corresponded to professional users. According to the findings, users and platforms could realize potential information manipulation and increase the transparency of the online society.

**Keywords:** social networks; political propaganda; cyber-army; information manipulation

## 1. Introduction

Platforms of social media have been considered as beneficial for the deepening of democratization, since they allow users to engage in political discussions and deliberation more easily [1,2]. Sites for online political discussion lower the barriers of political participation, since it takes less time and effort to start new discussion topics, without the hassle of real-world organization and coordination. Social media also allows people to get together and share their thoughts on politics with others [3,4]. The anonymity of online platforms makes it less likely for users to face problems of social desirability or cross-pressure so that they can express their thoughts more freely, feel free to disagree with others, and even post hate speech [5–7]. Additionally, in the case of the Arab Spring uprisings in 2012, social media and online platforms enabled ordinary people to be engaged in political mobilization to fight against the oppressive state authorities [8,9]. Without the Internet and social media as a means to keep people connected and informed, the popular resistance against authoritarian regimes in the Arab Spring could hardly be imaginable [10,11].

All of the perspectives mentioned above seem to support the idea that social media and the Internet open up new possibilities for democracy and civil political participation to deepen. However,

we need to critically rethink the validity of this argument, which advocates technological determinism and optimism. Regardless of the advantages that the Internet and social media may bring to us for a more extensive degree of political participation amongst the public, some of the potential dangers that online platforms may cause to harm democracy need to be highlighted. Firstly, while the Internet makes it easier for people, on average, to gain more opportunities to take part in political discussions with others, it also, at the same time, silences certain groups of people who may not have access to or the ability to use this technology [12]. Secondly, it has been found that in cases of Scandinavian countries like Denmark and Sweden, online forums are used to spread anti-democratic propaganda and to help the rise of reactionary movements by fueling populist sentiments and spite [13].

One additional potential concern of using online platforms as the major means of political communication is that online political discussions may be manipulated by cyber-armies to shape how the public evaluates certain politicians. Political organizations and parties may recruit cyber-armies to monitor discussions, influence public opinion, and promote their candidates. If people do not understand the potential threat of information manipulation in online political discussions and trust the information and knowledge they gain in such discussions, these instances of cyber-armies will discredit the very idea of democracy, since people make their political judgments and voting decisions based on misinformation or manipulated arguments [14,15].

To avoid political propaganda and opinion shaping using recruited cyber-armies on online platforms, identifying such users and expelling them is an intuitive approach. However, currently, only governmental investigation can confirm the existence and identities of cyber-armies. To provide a systematic and efficient method for detecting these users, in this study, we attempt to detect potential cyber-armies by identifying aggressive accounts during election campaigns. In the first step, we hypothesized that recruited users are very active and stay firm in their political stances. They may support specific candidates and attack others. In order to examine these characteristics, we address the first Hypothesis H1.

**Hypothesis 1.** *There is a group of users giving a lot of negative/positive ratings to articles talking about particular candidates.*

In a second step, as cyber-armies are recruited users by certain groups, we hypothesized that they should spend much more time than ordinary users to monitor messages, promote their candidates, and respond to attacks in a speedy manner. To identify such users, we describe the second Hypothesis H2.

**Hypothesis 2.** *There is a group of users who can rapidly respond to any articles talking about particular candidates.*

In a third step, from a behavioral pattern viewpoint, we investigate the daily activities of users. As recruited users may follow regular work times to reply to and rate articles, we would like to address the third Hypothesis H3 and identify users corresponding to this characteristic.

**Hypothesis 3.** *There is a group of users who are active on weekdays and inactive on weekends.*

To address the above hypotheses, we first propose a series of systematic methods for identification. Next, we collect our dataset from the most prominent political discussion forum in Taiwan, where a national election was held in November 2018. The dataset includes more than 25,000 articles published between May and November 2018. Statistical methods are employed to investigate commenting behaviors of the articles. The findings help us to better understand the phenomenon of online political information manipulation and to start thinking about possible ways to counter the negative effects of cyber-armies and restore the function of online forums as a public sphere for democratic deliberation and discussion.

The remainder of this paper is organized as follows. Section 2 discusses the related works. The overview of the proposed approach and data collection are presented in Section 3. In Section 4, we

exhibit in-depth analysis results according to our method. The validation of the proposed method and results is presented in Section 5. The main issues to address in the future and the conclusion of our research are presented in Section 6.

## 2. Related Works

### 2.1. Political Propaganda on Social Networks

Social networks have become a vital tool for information diffusion and exchanges of political opinions [16–19]. Taufiq et al. [20] conducted a survey of students at the University of Narowal, Pakistan, and figured out that the majority of the students have been using social media for political discussions. In [21], by investigating Twitter messages related to the German federal election, the authors concluded that Twitter has actually been used for political discussions and that the messages truly reflect the election results. Emotional analysis of online messages is also receiving much attention from researchers. In [22], the authors examined political tweets on Twitter to identify the factors that affect users' retweet behaviors. They found that the retweet behavior of a user is strongly affected by the tweet's emotional and political affiliation. In [23], the authors provided a hybrid method for predicting the election results based on the number of articles, the online ratings toward candidates, and the sentiment scores of articles. In [24], the authors focused on reddit, a popular American-based social forum, to investigate different characteristics of controversy in political discussions. Regarding political leaning, a political leaning inference scheme was proposed in [25] based on tweets and retweets. In this paper, the authors assumed that tweet and retweet behaviors on Twitter are consistent. To identify influential spreaders on Twitter during the Malaysian General Election in 2013 and to clarify whether they have effects on election results, Sun et al. [26] proposed an influential spreader-detection scheme based on  $k$ -shell decomposition. The authors concluded that both political and non-political Twitter accounts have potential influences on the election results, especially the non-political accounts. However, the party which these non-political accounts support or attack is not addressed in this paper. In addition, these accounts may manipulate the public opinion, which made a significant impact on the final election results.

### 2.2. Cyber-Armies and Online Information Manipulation

Nowadays, the Internet serves as an essential platform for every aspect of our daily lives; the potential threat of cyber-armies cannot be over-emphasized [27]. Given the lack of a comprehensive scholarly discussion on what constitutes cyber-armies in previous literature, the meaning of this term changes in response to the context in which it is used [28]. Common characteristics of cyber-armies include a coordinated action taken by a group of individuals to shape the way that people think about politics and politicians [29]. Cyber-armies can mainly be categorized into two broad types, which differ in their targets and goals. The first type of cyber-army is highly associated with the concept of terrorism and other malicious and criminal activities [30]. In these instances, the goal of the hacking attacks is to create panic, fear, or disruption in people's daily lives, and these render the targeted political system unstable. These instances of terrorist cyber-attacks have been discussed in literature on international relations and security studies [28,31]. The second type of cyber-army aims to manipulate domestic public opinion and to sway electoral outcomes to favor certain partisan groups and political parties [32,33]. This term has become especially salient in Taiwanese politics after the 2014 local election, where instances of cyber-armies were reported to shape public perceptions on the two mayoral candidates of the capital Taipei City [34]. However, currently, there is no efficient and accurate method except for judicial investigation by the government, which can assure who the cyber-armies are. In [35], the authors identified social bots and their behaviors during Japan's 2014 general election by using a corpus-linguistic technique. They showed that a cyber-army of bots, who favored Shinzo Abe, played an important role in his success in this election. This paper only considered repeated tweets as a criterion for a social bot detection algorithm; however, other features,

such as rating behaviors, are also essential factors in identifying public opinion. In addition, in [36], the authors investigated reddit using network analysis and comment intervals to distinguish bot accounts. However, activities of bots or programmable accounts may have regularity compared to human-operated accounts. Leveraging human features, such as political preference and working-time patterns, for human-operated cyber-army detection remains scant. Thus, in this paper, we attempt to identify such users by recognizing active accounts on the most popular social forum in Taiwan during the 2018 Taiwanese capital election. We pursue our analysis from various perspectives and show a number of distinctive users with obvious political leanings, thousands of comments, minute-long response times, and regular behaviors on weekdays and weekends.

The importance of the Internet as a source of political information has been increasing rapidly in the last two decades [37–39]. However, greater access to the Internet as a means to obtain political information also implies that there are higher risks of misinformation, fake news, and political manipulation [40]. Many studies have been conducted in economics and commerce research to explore how exposure to online comments and public opinion would affect the way that people perceive certain products and companies. It has been found that higher proportions of online negative consumer reviews will make it likely for consumers to adopt the opinions of the reviewers [41–43]. Although information manipulation of online discussions has been found effective in the area of business, there has been little systematic analysis of whether online information manipulation impacts vote decisions in real-world elections [21,44]. Furthermore, before discussing the impacts of cyber-armies, a fundamental question is whether we can detect their existence in the first place. Therefore, this paper aims to find the behavioral patterns of online professional commenters who attempt to shape public opinion and the way that people perceive specific candidates.

### 3. Methodology

To address our hypotheses, we start by distinguishing users based on their online behaviors, including comments, ratings, and other activities. Compared with regular users, “recruited users” have distinctive characteristics, as follows. (1) These users are active in commenting on political articles. (2) They are eager and tenacious in showing their support/rejection of specific candidates. (3) These users may respond to articles quickly and spend too much time on the forum. To verify the existence of such users, we collected the articles from the most significant political discussion forum in Taiwan before the national election, which was held in November 2018.

#### 3.1. Dataset

The dataset was crawled from the most influential forum in Taiwan, namely the PTT Bulletin Board System (PTT). PTT consists of over 20,000 boards discussing about various topics. “Gossiping,” one of the most popular boards, is concentrated on news discussions, especially for political issues. We collected all articles and comments in the Gossiping Board from 24 May to 24 November, 2018 (the election day), a six-month-long observation timeframe. Each article posted on PTT consists of the following information:

1. **Author information:** IP address, author ID, and nickname.
2. **Article metadata:** Publication time.
3. **Article content:** Textual part of the article.
4. **User comment and rating:** User’s comment and positive/neutral/negative ratings.

The crawled data was grouped into three subdatasets according to the three major candidates, including Wen-Je Ko (the current mayor of Taipei city), Shou-Chung Ting (nominated by the major opposition party, Kuomintang), and Wen-Chih Yao (nominated by the ruling party, the Democratic Progressive Party). To concentrate on articles focusing on particular candidates, articles containing multiple candidates’ names are not included in our dataset. Table 1 summarizes our dataset.

**Table 1.** A summary of our dataset.

Candidate	# Articles	# Commenters	# Authors	# Comments
Ko	19,316	68,334	4549	1,632,681
Yao	3928	29,348	1709	230,438
Ting	1903	21,197	973	105,539

### 3.2. Users with Obvious Political Preferences

To address Hypothesis H1, we demonstrate a series of formalized methodologies as follows. Hypothesis H1 can be interpreted as identifying active users with a strong political preference/disfavor toward certain candidates. To address this hypothesis, we define the user polarity to measure the attitude of a user toward a candidate, as in Definition 1.

**Definition 1.** For each commenter  $u$ , we compute the total number of positive and negative ratings toward candidate  $c$ , denoted as  $PR_{u,c}$  and  $NR_{u,c}$ , respectively. The rating polarity of commenter  $u$  toward candidate  $c$  is denoted as  $polarity_{u,c}$ .

$$polarity_{u,c} = PR_{u,c} - NR_{u,c} \quad (1)$$

Next, we transform the user's attitude toward candidates into a polarity point in an  $n$ -dimensional space. The definition of the polarity point is described in Definition 2.

**Definition 2.** Assume that there are  $C$  candidates; the polarity point of  $u$  is defined as  $p_u$  in a  $C$ -dimensional space, where  $p_u = (polarity_{u,1}, polarity_{u,2}, polarity_{u,3}, \dots, polarity_{u,C})$ .

From the definition, every user will be given a point in a  $C$ -dimensional space. Next, we define a pivot point, representing the median polarity of users toward candidates, as a reference in this space. If some users are relatively far from the pivot point, they should reveal obvious attitudes toward candidates. We describe the definition of the pivot point in Definition 3.

**Definition 3.** The median value of the polarity scores of  $N$  users toward candidate  $c$  is denoted as  $polarity.med_c$ . For a total of  $C$  candidates, the pivot polarity point of  $C$  candidates is defined as  $p.med$  in a  $C$ -dimensional space, where  $p.med = (polarity.med_1, polarity.med_2, polarity.med_3, \dots, polarity_C)$ .

To identify the "outlier" users, we measure the political preferences of users by calculating the Euclidean distances between the user point  $p_u$  and the pivot point  $p.med$ . The political preference of user  $u$  is calculated according to Equation (2):

$$pref_u = \sqrt{\sum_{k=1}^C (polarity_{u,k} - polarity.med_k)^2} \quad (2)$$

According to the above measurement, we can assess the distance between each user and the pivot points; thus, we can systematically investigate users matching the criteria described in Hypothesis H1.

### 3.3. Users with Rapid Response Times

To address Hypothesis H2, the best way is retrieving the online time or login/logout time of each user. However, most of the social service sites do not reveal these data due to user privacy. Therefore, to address this hypothesis, we turn to measure the response time of each user. If a user wishes to respond to an article within a few minutes, this user should spend more time online to keep focusing on certain posts.

According to the above assumptions, we analyze the comment activity and measure the response time of every user in our dataset. The definition of the response time is described in Definition 4.

**Definition 4.** For user  $u$  who has commented on article  $a$  published at  $publish.time_a$  a total of  $t$  times, the comment timestamps are denoted as  $comment.time_{u,a,1}, comment.time_{u,a,2}, \dots, comment.time_{u,a,t}$ . We take the first timestamp to calculate the response time, since it is the first activity of user  $u$  on article  $a$ . Thus, the response time of  $u$  to  $a$  is defined as  $resp.time_{u,a}$ .

$$resp.time_{u,a} = comment.time_{u,a,1} - publish.time_a \quad (3)$$

The set of articles related to candidate  $c$  is denoted as  $\mathbb{A}_c$ .

To measure the activity of users toward articles about particular candidates, we compute the response time of every user toward each candidate. According to Definition 4, the response time of user  $u$  toward candidate  $c$  is denoted as  $resp.cand_{u,c}$ , and it can be expressed as Equations (4) and (5).

$$R_{u,a} = \{resp.time_{u,a} \mid a \in \mathbb{A}_c\} \quad (4)$$

$$resp.cand_{u,c} = median(R_{u,a}) \quad (5)$$

$$resp.min_u = \min\{resp.cand_{u,1}, \dots, resp.cand_{u,C}\} \quad (6)$$

For  $C$  candidates, we identify the users who rapidly respond to articles according to Equation (6). From the method, these users are who we wish to recognize in Hypothesis H2.

### 3.4. Users Acting as Office Workers

According to Hypothesis H3, we try to distinguish if there are some users that behave like office workers and are active mostly during regular hours (e.g., Monday to Friday). To address Hypothesis H3, we first compute the daily activities of each user in the last six months of observation (i.e., 27 weeks). Next, we take the average of the behavior of each user regarding the number of comments, the comment polarity toward candidates, and the response time according to 27 groups of weekdays and weekends.

**Definition 5.** For user  $u$  during week  $w$ , we calculate the total number of comments on weekdays and weekends, denoted as  $com.weekday_{u,w}$  and  $com.weekend_{u,w}$ . The observation period is  $N$  weeks, and the total number of active weeks of user  $u$  is denoted as  $N_u$ .

From the above definition, the average activity difference of each user between weekdays and weekends is computed as follows:

$$diff_u = \frac{\sum_k^{N_u} \left( \frac{com.weekday_{u,k}}{5} - \frac{com.weekend_{u,k}}{2} \right)}{N_u} \quad (7)$$

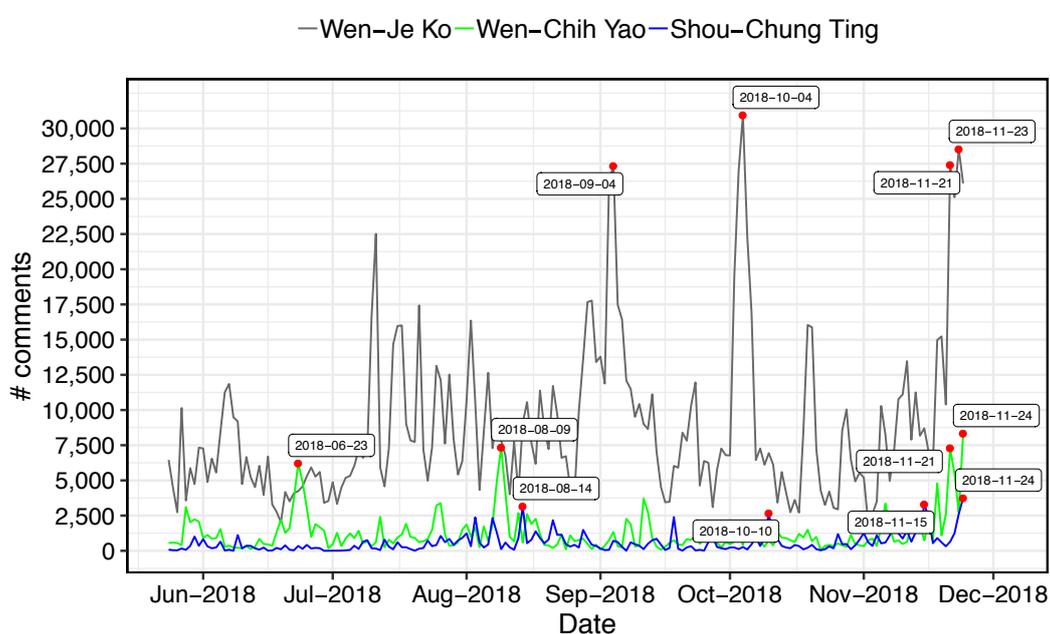
Users selected according to Equation (7) may contain those who are only active for a very short period of time (e.g., two weeks). To identify those who have a long-term activity pattern, we only study those users who have activities in at least  $P\%$  (in this study, we use 30% as the example threshold, though the value can be adjusted according to the application situation) of our  $N$ -week-long observation.

## 4. Analysis

In this section, we analyze our data collection in terms of the numbers of comments, the polarities of ratings given by users, and the response times of users. These metrics are employed to find users with apparent political preferences and those who can respond to articles in a brief period. We also investigate the daily activities to find commenters with an apparent behavioral difference between weekdays and weekends.

#### 4.1. Daily Quantity

Figure 1 demonstrates the daily number of comments for each candidate during the six-month-long observation. From the figure, we find the incumbent, Mayor Wen-Je Ko, is in heated discussions, with many more comments than the other two candidates. We also denoted the top four peaks of the frequencies for each candidate in the same figure. Except for the election day/eve, we find that the other three peaks in Ko's line (2018-09-04, 2018-10-04, and 2018-11-21) are located in the last three months of the election cycle, and have obvious growths compared with those of the first three months of the observation. In contrast, for the other two candidates, Ting and Yao, some of their peaks are situated in the first three months; however, they were not able to bring the high volumes at the initial stage of the campaign to the final months. From the number of comments observed on a daily basis, we can see significant differences between each candidate. To further investigate if a large number of comments are posted by particular groups of users, in the following paragraphs, we discuss the individual behaviors of top commenters.



**Figure 1.** Daily number of comments about each candidate during our observation.

#### 4.2. Candidate Popularity

Table 1 shows that the most popular candidate is the incumbent, Taipei city mayor Wen-Je Ko. Figure 2 presents the complementary cumulative distribution function (CCDF) of comments of users for each candidate. In the figure, we take the logarithm of both the x- and y-axes for better readability. We observe that the probability difference of the three candidates becomes more substantial when the number of comments is over 100. The probability of having users with high numbers of comments about Ko is bigger than for the other two candidates, which shows that Ko gets much more attention from active users than the other two candidates. From another point of view, the distribution of comments about each candidate follows the power-law distribution, indicating that a majority of commenters have very few comments and a significant number of comments are posted by a tiny portion of commenters. In addition, we observe that more points of Ko are farther from the distribution pattern (as shown in the circle in Figure 2) than those of the other two candidates. The results indicate that active users are much more likely to comment on Ko's articles rather than on other candidates'.

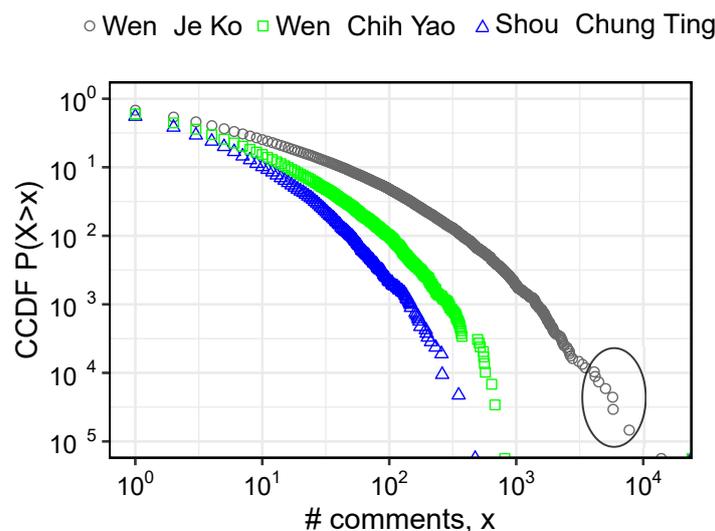


Figure 2. The distribution of comments given by commenters.

We present the network graphs of commenters and authors for the three candidates in Figure 3. For each candidate, we only consider the top ten commenters with the highest numbers of comments. The top ten commenters are denoted as the labeled nodes, and the unlabeled nodes represent the authors. We do not plot the authors whose articles do not receive comments from these commenters. There will be a link between commenter  $u$  and author  $v$  if commenter  $u$  has commented on an article posted by author  $v$ . The size of the labeled node is determined by its degree; the bigger the node, the higher its degree. The weight of an edge denotes the number of comments that commenter  $u$  has given to author  $v$ . As shown in the figure, the number of nodes and the density of Ko's network are much higher than those of the other two candidates, which again shows that Ko is the most discussed candidate. In terms of the active users in the networks, user 001 is the most active user among the three candidates. This may imply that this user is the most influential user on the three networks. However, considering each network individually, we find that each network has its own major users. For example, user 022 in Yao's network, which does not belong to the top ten commenters of Ko and Ting, gives many comments to Yao's articles. Similarly, user 130 comments a lot on Ting's articles. From another perspective, there are strong links between user 130 and other users in the top ten commenters of Ting's articles, which shows that this user gives or receives many comments from the other top ten users in Ting's group.

#### 4.3. Commenter Polarity

Using the polarity of users towards candidates, we attempt to address the Hypothesis H1. We first select the top 100 commenters in our dataset according to their comment quantity to investigate.

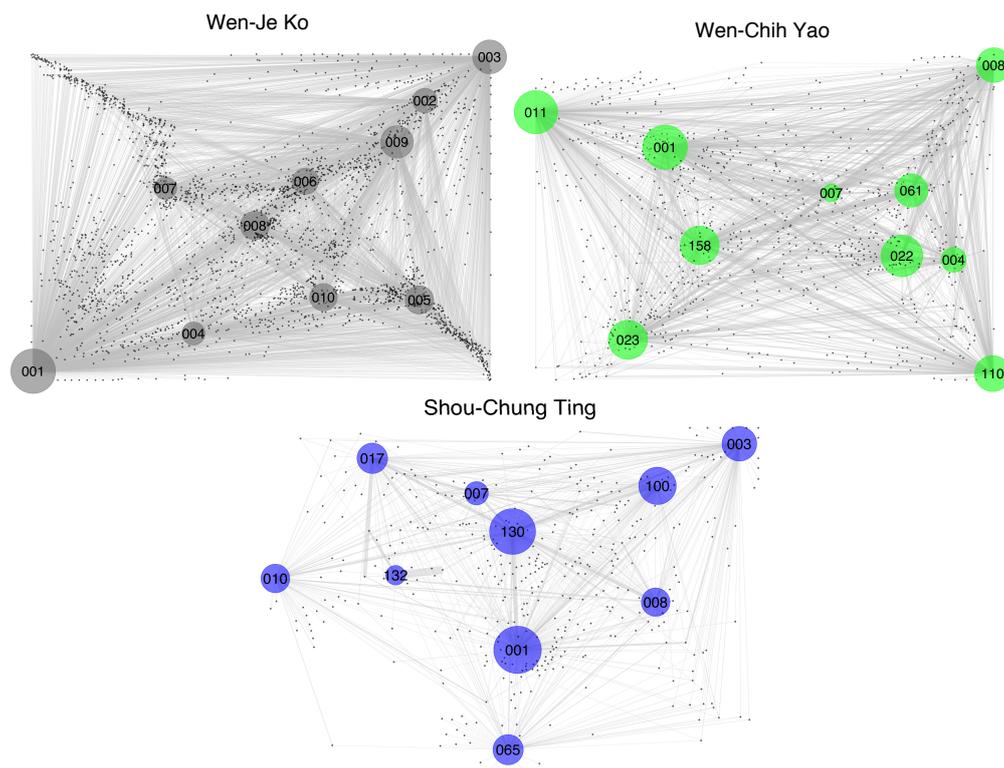
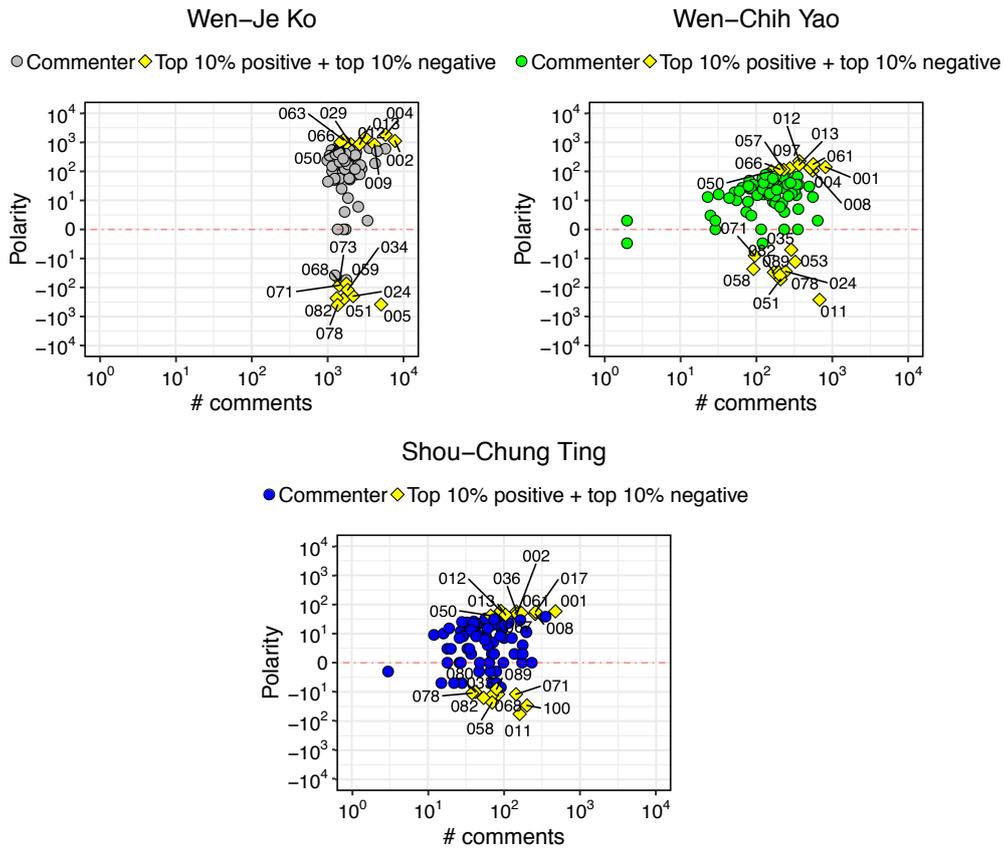


Figure 3. Network graph of the top ten commenters for each candidate.

#### 4.3.1. Comment Polarity and Quantity of Active Commenters

Figure 4 shows the top commenters for the three candidates. The top 10% of the positive users and the top 10% of the negative users are marked in yellow diamonds. This figure shows that these commenters behave distinctly among the three candidates. From the horizontal locations of points, we can observe that the top commenters are significantly more active in Ko's discussions than in the other two candidates'. We also find that active users are more polarized in Ko's articles than in Yao's and Ting's. In Figure 4, in Ko's distribution, points are more distant from the horizontal line representing polarity = 0. These results may be attributed to the fact that Ko is the incumbent and the most-discussed candidate; thus, the favors and disavors toward him are more obvious than for the other two challengers.

Furthermore, we calculate the Pearson's correlation between the polarity and the number of comments of the users for each candidate. The Pearson's correlations for Ko, Yao, and Ting are 0.458, 0.139, and 0.343, respectively. From the results, we observe that a higher number of comments does not reflect higher online ratings, particularly for Yao. According to the above statistics, we find that the correlation coefficient in Ko's dataset is considerably higher than in Yao's and Ting's. This observation demonstrates that the active commenters are likely to give more negative ratings to Yao's and Ting's articles than to Ko's articles, while Ko receives more positive ratings than the other two candidates. In addition, these coefficients correspond to a series of poll results during the campaign and to the final election results ([https://en.wikipedia.org/wiki/2018\\_Taiwanese\\_municipal\\_elections](https://en.wikipedia.org/wiki/2018_Taiwanese_municipal_elections)), where Ko had much more support than Ting and Yao during the campaign, and won the election.

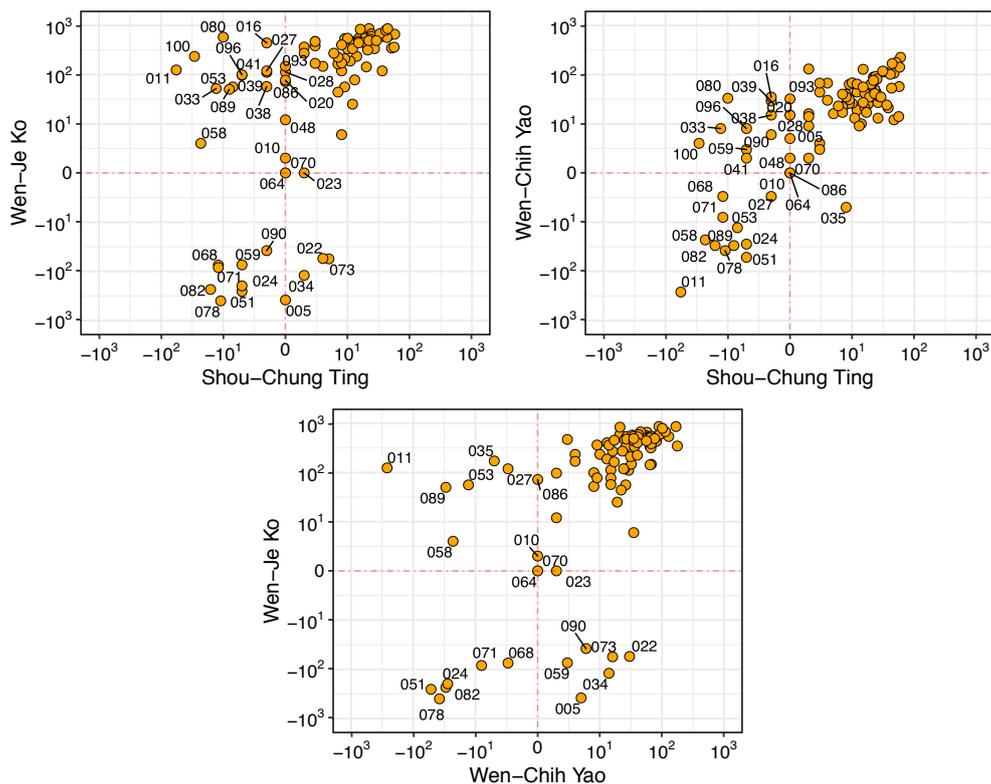


**Figure 4.** Analysis of the number of comments and the polarity of the top 100 commenters for the three candidates.

From the viewpoint of the individual users shown in Figure 4, we observe some of them whose behaviors of online commenting are distinctive from those of the majority of users, and they are treated as targets to address H1; for example, users **004** (most positive) and **078** (most negative) in Ko’s articles, **012** (most positive) and **011** (most negative) in Yao’s, and **012** (most positive) and **011** (most negative) in Ting’s. Note that user **011** gives the most negative ratings to both Yao’s and Ting’s articles, but not to Ko’s, demonstrating an apparent political preference. These users correspond to our first hypothesis that there are some users that give extremely polarized ratings to certain candidates. However, from only the polarity perspective, it is not convincing enough to make the judgment that candidates recruit these users. In the next section, we present the analysis of the comment polarities of the top commenters between the three candidates.

#### 4.3.2. Analysis of Top Commenters’ Polarities towards Candidates

A major objective of users recruited by a political campaign is to promote their candidates and to attack other candidates. To figure out the user preferences towards candidates, we demonstrate the polarities of the top commenters for the three candidates, as shown in Figure 5. We first discuss the third quadrant, indicating negative ratings for both candidates. We observe that there are more points in the third quadrant of the middle sub-figure (Yao-Ting). The results also reveal that more users tend to give Ko positive ratings, but not for the other two candidates.



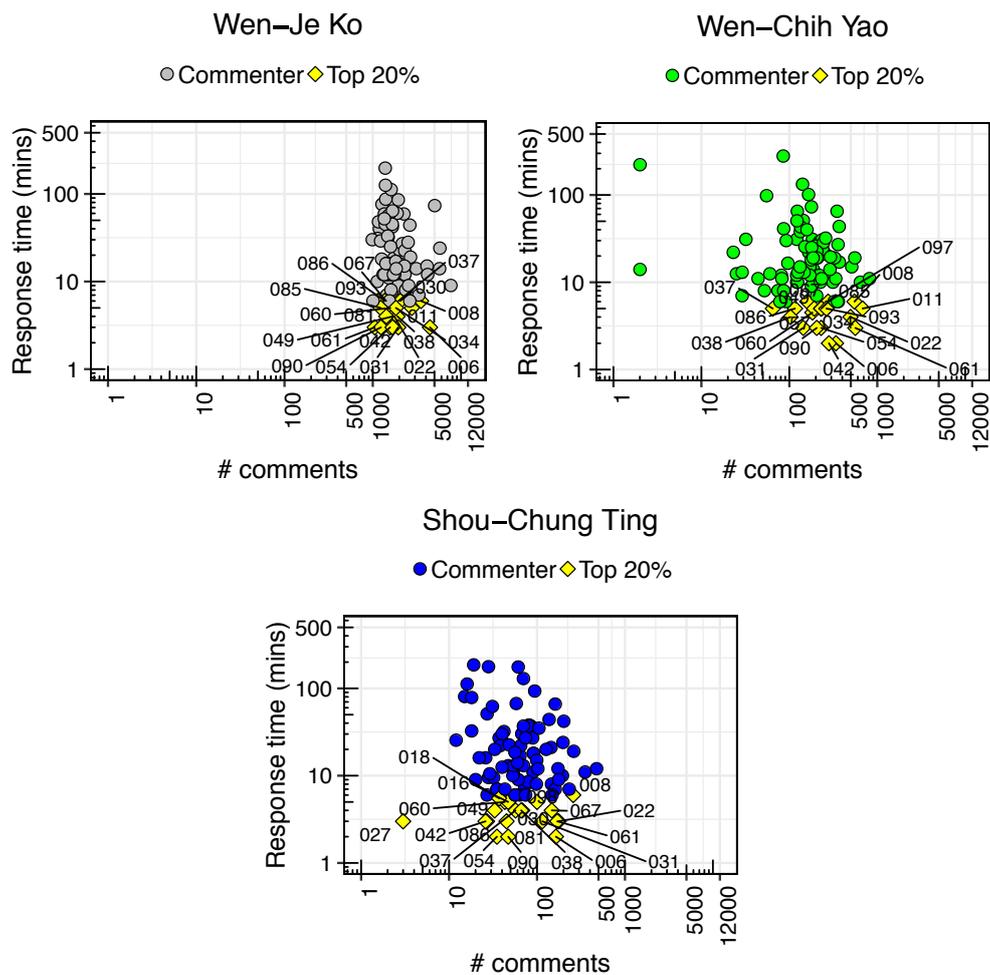
**Figure 5.** Analysis of the polarities of comments of popular commenters for the three candidates.

For individuals, we find that some users have obvious political preferences. For example, **053** (Ko: 55, Yao: -12, Ting: -6), **089** (Ko: 49, Yao: -29, Ting: -7), and **027** (Ko: 118, Yao: -2, Ting: -1) rate positively on Ko's articles and negatively on Ting's and Yao's articles. In addition, user **011** (Ko: 123, Yao: -263, Ting: -56) gives positive ratings to Ko's articles, but gives significantly negative ratings to the other two candidates' articles. According to the results, the users that we identified have strong political preferences, and they can be categorized into four types. (1) Users **004**, **001**, **012**, **063**, **002**, and **066** hold positive opinions about three candidates, but they give extremely high ratings to Ko's articles compared with the other two candidates', as shown in Table 2. (2) User **011** has strong negative ratings for Yao's and Ting's articles and positive ratings for Ko's. (3) User **005** has strong negative ratings for Ko's articles but stays neutral for the other two candidates' articles. (4) Users **078** and **082** are negative for the three candidates' articles but are more obvious in Ko's.

The users mentioned above reveal political support/rejection of candidates. Thus, based on the overview and the case studies in the previous results, we identify some users with frequent activities and strong political preferences; they are the users for Hypothesis H1.

#### 4.4. Commenters' Response Times

To address the Hypothesis H2, we focus on measuring the response times of each user toward candidates' articles. Figure 6 shows the median value of the response times of the top 100 commenters. It is noted that, in the figure, we mark the top 20% of the commenters with the smallest response times in yellow diamonds. It can be seen from the figure that users **006** and **054**, filtered from the previous analysis, reply rapidly to the three candidates' articles. Users **090** and **042** also respond extremely quickly to the three candidates' articles. We also find that a significant number of users give a lot of comments to each candidate and respond to the articles within a short period (the bottom-right points).



**Figure 6.** Analysis of the number of comments and the response times (median) of the top 100 commenters for the three candidates

In Table 2, the top 20 users with the shortest response times for the three candidates are shown. From the table, we find that these users responded very quickly to candidates' articles. Most of these users replied within 2–7 min after articles were published. For example, users 006 and 054 respond very rapidly (the median of the response time  $\leq 3$  min) to candidates' articles, which means that at least 50% of their comments are published in 3 min. Compared with ordinary users who casually glance at the board and comment, the results imply that these users spend a lot of time monitoring and discussing related articles. From the statistics, we consider that these users correspond to the target of H2.

#### 4.5. Users with the Three Characteristics

According to the above analyses of comments, political preferences, and response times, we select those active users with many comments, strong favorites, and fast replies. We rank the 100 active users by the three features and find fourteen users ranking in the top 50% in all of the characteristics, as shown in Table 3. From the results, we consider the users who satisfy the three features we proposed as professional users.

**Table 2.** Top 20 users for number of comments, polarity, and response time.

# Comments					Polarity					Response (min)				
ID	Ko	Yao	Ting	Total	ID	Ko	Yao	Ting	Pref	ID	Ko	Yao	Ting	Minimum
001	13,849	813	475	15,137	004	1817	129	1	1698.967	006	3.0	2.0	2.0	2.0
002	7652	44	148	7844	001	1579	140	57	1463.485	054	3.0	3.0	2.0	2.0
003	5699	312	352	6363	012	1440	223	59	1335.758	090	3.0	3.0	2.0	2.0
004	5744	511	37	6292	063	1200	80	24	1080.265	042	3.0	2.0	3.0	2.0
005	5019	357	18	5394	002	1111	11	46	989.893	061	3.0	3.0	3.0	3.0
006	4446	339	164	4949	066	1009	117	30	893.455	031	3.0	3.0	3.0	3.0
007	4176	554	197	4927	013	858	166	43	752.746	022	4.0	4.0	3.0	3.0
008	3521	550	258	4329	029	860	88	21	741.996	081	4.0	6.0	3.0	3.0
009	4106	60	60	4226	009	834	20	15	712.119	086	6.0	5.0	3.0	3.0
010	3322	352	232	3906	050	790	100	40	674.603	027	15.0	14.0	3.0	3.0
011	2772	680	161	3613	057	677	112	32	564.455	038	4.0	4.0	4.0	4.0
012	3103	369	91	3563	043	670	44	23	549.205	049	4.0	5.0	4.0	4.0
013	2624	368	105	3097	036	659	56	56	541.223	037	6.0	5.0	4.0	4.0
014	2670	293	127	3090	078	-399	-37	-10	523.581	067	6.0	6.0	4.0	4.0
015	2509	258	89	2856	008	622	105	46	510.125	030	6.0	6.0	4.0	4.0
016	2355	360	47	2762	005	-382	4	0	504.096	060	5.0	4.5	5.0	4.5
017	2298	177	261	2736	049	599	91	23	483.709	085	5.0	5.0	6.0	5.0
018	2620	78	37	2735	003	593	20	37	472.207	034	5.0	5.0	6.0	5.0
019	2540	74	20	2634	099	589	41	15	467.968	011	5.0	5.0	7.0	5.0
020	2372	147	27	2546	080	577	32	-9	455.582	093	6.0	5.0	6.0	5.0

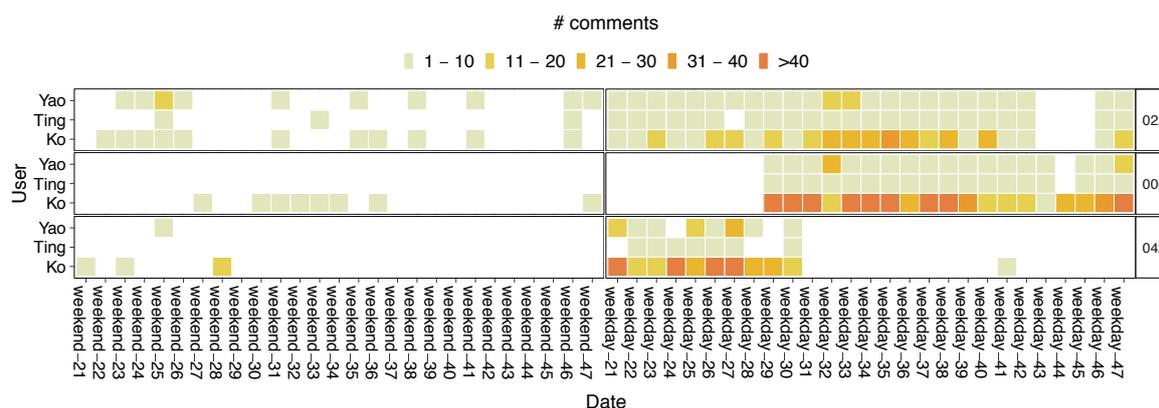
**Table 3.** The overlapping users in the top 50% for number of comments, polarity, and response time.

# Comments					Polarity				Response (min)			
ID	Ko	Yao	Ting	Total	Ko	Yao	Ting	Pref	Ko	Yao	Ting	Minimum
001	13,849	813	475	15,137	1579	140	57	1463.485	10.0	11.0	12.0	10.0
002	7652	44	148	7844	1111	11	46	989.893	9.0	11.0	6.0	6.0
003	5699	312	352	6363	593	20	37	472.207	14.0	10.0	11.0	10.0
006	4446	339	164	4949	490	34	28	369.379	3.0	2.0	2.0	2.0
008	3521	550	258	4329	622	105	46	510.125	6.0	6.0	6.0	6.0
011	2772	680	161	3613	123	-263	-56	282.448	5.0	5.0	7.0	5.0
016	2355	360	47	2762	439	34	-1	317.734	8.0	6.0	5.0	5.0
026	2223	81	90	2394	538	47	15	417.532	11.0	12.0	11.0	11.0
030	2052	91	68	2211	545	34	18	423.752	6.0	6.0	4.0	4.0
031	1944	144	115	2203	526	40	22	405.301	3.0	3.0	3.0	3.0
036	1833	121	146	2100	659	56	56	541.223	11.0	10.0	8.0	8.0
037	1954	65	57	2076	539	26	25	417.731	6.0	5.0	4.0	4.0
044	1736	171	53	1960	482	50	22	362.344	14.0	15.0	10.0	10.0
049	1689	192	33	1914	599	91	23	483.709	4.0	5.0	4.0	4.0

#### 4.6. Daily Behavior

In addition to the previous studies, we address the third hypothesis by analyzing the daily behavior. To understand if some users act differently between weekdays and weekends, we calculate the weekday and weekend difference according to Equation (7). From the results, we present the commenting behavior of the three users with a significant difference between weekdays and weekends, as shown in Figure 7.

From the figure, we notice the obvious difference in commenting activities from the colored area. The three users frequently comment on the three candidates during weekdays, but discuss at few times during weekends. For example, users 006 and 042 comment actively on weekdays during the six-month observation, and they are almost inactive on weekends. For user 022, we also find a significant difference between weekdays and weekends. These users match the characteristics described in H3.



**Figure 7.** The top users with significant differences in comment quantity between weekdays and weekends (left panel: Weekends; right panel: Weekdays).

#### 4.7. Sockpuppet Analysis

Identifying sockpuppets (multiple accounts belonging to a single user) in social media is also a major task for confronting potential information manipulation. In this study, we use behavioral analysis and find that two accounts, users **006** and **042**, shown in Figure 7, have very similar behavioral patterns. Firstly, these two users have comparable activities, but are just perfectly separated by weeks 30 to 31. User **042** commented frequently before week 31 and hibernated after that time; in contrast, user **006** started to behave actively on week 30. Both of them are very active in commenting on the three candidates, especially for Ko, only on weekdays. Secondly, the IDs of the two users are  $VV^{***}$  and  $W^{***}$  (the \* digits are the same). These two IDs are very similar and only have a difference between W and VV. Finally, from IP investigation, we find that the two accounts have five mutual IP addresses. Such evidence may imply that the two IDs could belong to the same user.

### 5. Validation with IP Address Tracking

Currently, retrieving the ground truth of cyber-armies is a difficult task, as only formal investigation by the government can confirm if an account is a professional user who is part of a cyber-army. To validate our proposed method and results, in this study, we compare our work with the results of IP address tracking, which is used in identifying sockpuppet accounts on Wikipedia [45,46].

#### 5.1. Users Sharing the Same IP Address

To identify users using the same IP address, we measure the similarity between the used IP address sets of two users using the Jaccard index. The Jaccard index for each pair of users is calculated as follows:

$$J(u_i, u_j) = \frac{|\Gamma(u_i) \cap \Gamma(u_j)|}{|\Gamma(u_i) \cup \Gamma(u_j)|}, 0 \leq J(u_i, u_j) \leq 1, \quad (8)$$

where  $\Gamma(u_i)$  and  $\Gamma(u_j)$  are the sets of used IP addresses of user  $u_i$  and user  $u_j$ , respectively.

We only consider the users sharing the same IP address with the Jaccard index  $\geq 0.5$ , as shown in Table 4. We observe that users **006** and **042**, the accounts that we suspected to be owned by the same person, use five and six different IP addresses, respectively, and, not surprisingly, five of the six IP addresses are exactly the same. In addition, the response times of these two users are also equal (2 min). In terms of polarization, it is obvious that the two users have similar political preferences. From these analyses, we can confirm that the users **006** and **042** belong to the same owner. Similarly, users **028** and **031** use almost exactly the same IP address (Jaccard index = 0.970). The polarities of these two users are also similar. Unlike users **006** and **042**, the users **028** and **031** have different response times, but the difference is not too big. Therefore, the results imply that these two accounts belong to the same user or, at least, have similar political stances. Another pair of users using the same IP address includes the

users **018** and **070**. The IP addresses that these two users used are 70% similar. The response times of the two users are nearly the same, 5.5 min for **018** and 8 min for **070**. However, the political reference scores of these two users are quite different; user **018** gives positive ratings to the three candidates, while user **070** stays neutral.

**Table 4.** A sample group of users sharing the same IP addresses.

ID	# Used IP	# Shared IP	Jaccard Ind. of IP	Resp. Time (min)	# Comments			Polarity		
					Ko	Yao	Ting	Ko	Yao	Ting
028	33	32	0.970	11.0	1764	336	173	112	14	0
031	32			3.0	1944	144	115	526	40	22
018	34	28	0.700	5.5	2620	78	37	77	8	12
070	34			8.0	1346	233	98	0	0	0
006	6	5	0.833	2.0	4446	339	164	490	34	28
042	5			2.0	1679	282	26	224	39	6
049	2	2	0.500	4.0	1689	192	33	599	91	23
079	4			25.5	1381	150	40	452	16	16

### 5.2. Comparison between the Proposal and the IP Tracking Method

In Table 3, we list 14 mutual users with a high rank in the three features: Comment quantity, preference score, and response time. These users can be considered as potential cyber-armies, identified according to our proposal. Next, we compare if these users can be validated as sockpuppets using IP tracking. From Tables 3 and 4, users **006**, **031**, and **049** are included in the results of both our proposal and IP tracking. This outcome demonstrates that our results can detect a number of active sockpuppets identified by the existing methods. In addition, our method can provide more potential candidates with professional user characteristics for further cyber-army investigations.

## 6. Conclusions

To identify potential targets of recruited users in online spaces, we conducted a behavioral analysis according to an online observation of a popular forum during the 2018 election in Taiwan. First, by analyzing the commenting and endorsement behaviors of users toward candidates, we identified several users who regularly rate positively/negatively for specific politicians. In addition, we examined the time intervals between publishing articles and posting comments to measure the response times of users. Several users were found with response times within only 2–7 min of articles being posted. Moreover, a number of users commented more than 1000 times, but they were only active during weekdays. With a verification using IP tracking, we identified groups of active accounts with very similar IP histories. Through this work, we demonstrate a series of approaches to identifying and validating potential recruited/professional users. To conclude this study, three main contributions are described as follows:

1. We collected a dataset consisting of over 25,000 articles and 73,000 users during a national election from PTT, a large-scale social platform in Taiwan.
2. We investigated the dataset according to multiple behavioral features to distinguish cyber-armies, and several potential targets are recognized from our results.
3. We validated the identified accounts using the IP tracking method. From the results, we found that groups of users shared a large number of IP addresses when posting and commenting.

With the increasing popularity and influence of online social platforms, messages on social media could be powerful and representative in both online and conventional mass media. Thus, issues about recruiting users or employing bots for online information manipulation are becoming crucial for modern society. In this paper, we strive to identify such professional users and to provide a series of numerical results to address the hypotheses proposed. According to the research outcomes, we

believe that this work could benefit the government as well as organizations to understand online user behaviors using an overview and case studies. Moreover, online platforms could leverage the methods for identifying potential opinion shaping, especially during elections and large-scale social events. For the next steps of this research, we aim to improve the proposed approaches to be adopted in different platforms and countries. In addition, verification approaches for the identified accounts based on content analysis or other clues are required. Through these approaches, we hope to increase the transparency and democracy of online information and discourse for next-generation communication paradigms.

**Author Contributions:** Conceptualization, M.-H.W.; methodology, M.-H.W.; software, M.H.W. and M.-H.W.; validation, M.-H.W., N.-L.N., S.-C.D., and P.-W.C.; formal analysis, M.-H.W. and N.-L.N.; investigation, M.-H.W., N.-L.N., and S.-C.D.; resources, M.-H.W., P.-W.C., and C.-R.D.; data curation, M.-H.W. and N.-L.N.; writing—original draft preparation, M.-H.W., N.-L.N., and S.-C.D.; writing—review and editing, M.-H.W., N.-L.N., S.-Ch.D., and P.-W.C.; visualization, M.-H.W. and N.-L.N.; supervision, M.-H.W., P.-W.C., and C.-R.D.; project administration, M.-H.W.; funding acquisition, M.-H.W., P.-W.C., and C.-R.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by the Ministry of Science and Technology, Taiwan, under the Grant MOST 107-2218-E-035-009-MY3, MOST 107-2218-E-003-002-MY3, and the Taiwan Information Security Center at the National Sun Yat-sen University (TWISC@NSYSU).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Stanley, W.; Weare, C. The Effects of Internet Use on Political Participation: Evidence from an Agency Online Discussion Forum. *Adm. Soc.* **2004**, *36*, 503–527. [[CrossRef](#)]
2. Stieglitz, S.; Dang-Xuan, L. Social media and political communication: A social media analytics framework. *Soc. Netw. Anal. Min.* **2013**, *3*, 1277–1291. [[CrossRef](#)]
3. Polat, R.K. The Internet and Political Participation: Exploring the Explanatory Links. *Eur. J. Commun.* **2005**, *20*, 435–459. [[CrossRef](#)]
4. Halpern, D.; Gibbs, J. Social media as a catalyst for online deliberation? Exploring the affordances of Facebook and YouTube for political expression. *Comput. Hum. Behav.* **2013**, *29*, 1159–1168. [[CrossRef](#)]
5. Santana, A. Virtuous or Vitriolic: The Effect of Anonymity on Civility in Online Newspaper Reader Comment Boards. *J. Pract.* **2014**, *8*, 18–33. [[CrossRef](#)]
6. Mondal, M.; Silva, L.A.; Benevenuto, F. A measurement study of hate speech in social media. In Proceedings of the 28th ACM Conference on Hypertext and Social Media, Prague, Czech Republic, 4–7 July 2017; pp. 85–94.
7. Datta, S.; Adar, E. Extracting inter-community conflicts in reddit. In Proceedings of the International AAAI Conference on Web and Social Media, Munich, Germany, 11–14 June 2019; Volume 13, pp. 146–157.
8. Khondker, H.H. Role of the New Media in the Arab Spring. *Globalization* **2011**, *8*, 675–679. [[CrossRef](#)]
9. Steinert-Threlkeld, Z.C. Spontaneous collective action: Peripheral mobilization during the Arab Spring. *Am. Political Sci. Rev.* **2017**, *111*, 379–403. [[CrossRef](#)]
10. Hermida, A.; Lewis, S.C.; Zamith, R. Sourcing the Arab Spring: A case study of Andy Carvin’s sources on Twitter during the Tunisian and Egyptian revolutions. *J. Comput.-Mediat. Commun.* **2014**, *19*, 479–499. [[CrossRef](#)]
11. Castells, M. State’s Response to an Internet-Facilitated Revolution: The Great Disconnection. In *Networks of Outrage and Hope: Social Movements in the Internet Age*; John Wiley & Sons: Hoboken, NJ, USA, 2015; pp. 62–67.
12. Norris, P. Understanding the Digital Divide. In *Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide*; Cambridge University Press: Cambridge, UK, 2001; pp. 26–38.
13. Askanius, T.; Mylonas, Y. Extreme-Right Responses to the European Economic Crisis in Denmark and Sweden: The discursive Construction of Scapegoats and Lodestars. *Javnost Public* **2015**, *22*, 55–72. [[CrossRef](#)]
14. Kaye, B.; Johnson, T. Online and in the Know: Uses and Gratifications of the Web for Political Information. *J. Broadcast. Electron. Media* **2002**, *46*, 54–71. [[CrossRef](#)]

15. Morgan, S. Fake news, disinformation, manipulation and online tactics to undermine democracy. *J. Cyber Policy* **2018**, *3*, 39–43. [[CrossRef](#)]
16. Guille, A.; Hacid, H.; Favre, C.; Zighed, D.A. Information diffusion in online social networks: A survey. *ACM Sigmod Rec.* **2013**, *42*, 17–28. [[CrossRef](#)]
17. Xu, B.; Liu, L. Information diffusion through online social networks. In Proceedings of the 2010 IEEE International Conference on Emergency Management and Management Sciences, Beijing, China, 8–10 August 2010; pp. 53–56.
18. Agrawal, D.; Budak, C.; El Abbadi, A. Information diffusion in social networks: Observing and affecting what society cares about. In Proceedings of the 20th ACM International Conference on Information and Knowledge Management, Glasgow, UK, 24–28 October 2011; pp. 2609–2610.
19. Li, Y.; Qian, M.; Jin, D.; Hui, P.; Vasilakos, A.V. Revealing the efficiency of information diffusion in online social networks of microblog. *Inf. Sci.* **2015**, *293*, 383–389. [[CrossRef](#)]
20. Ahmad, T.; Alvi, A.; Ittefaq, M. The Use of Social Media on Political Participation Among University Students: An Analysis of Survey Results From Rural Pakistan. *SAGE Open* **2019**, *9*, 2158244019864484. [[CrossRef](#)]
21. Tumasjan, A.; Sprenger, T.; Sandner, P.; Welpe, I. Predicting elections with Twitter: What 140 characters reveal about political sentiment. In Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, Washington, DC, USA, 23–26 May 2010; pp. 178–185.
22. Hoang, T.A.; Cohen, W.W.; Lim, E.P.; Pierce, D.; Redlawsk, D.P. Politics, sharing and emotion in microblogs. In Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Niagara, ON, Canada, 25–28 August 2013; pp. 282–289.
23. Wang, M.H.; Lei, C.L. Boosting election prediction accuracy by crowd wisdom on social forums. In Proceedings of the 2016 13th IEEE Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, USA, 9–12 January 2016; pp. 348–353.
24. Guimaraes, A.; Balalau, O.; Terolli, E.; Weikum, G. Analyzing the Traits and Anomalies of Political Discussions on Reddit. In Proceedings of the International AAAI Conference on Web and Social Media, Munich, Germany, 12–14 June 2019; Volume 13, pp. 205–213.
25. Wong, F.M.F.; Tan, C.W.; Sen, S.; Chiang, M. Quantifying political leaning from tweets, retweets, and retweeters. *IEEE Trans. Knowl. Data Eng.* **2016**, *28*, 2158–2172. [[CrossRef](#)]
26. Sun, H.I.; Ch'ng, E.; See, S. Influential spreaders in the political Twitter sphere of the 2013 Malaysian general election. *Ind. Manag. Data Syst.* **2019**, *119*, 54–68. [[CrossRef](#)]
27. Abomhara, M.; Køien, K. Cyber Security and the Internet of Things: Vulnerabilities, Threats, Intruders and Attacks. *J. Cyber Secur.* **2015**, *4*, 65–88. [[CrossRef](#)]
28. Andress, J.; Winterfeld, S. Definition for Cyber Warfare. In *Cyber Warfare: Techniques, Tactics and Tools for Security Practitioners*; Elsevier: Amsterdam, The Netherlands, 2013; pp. 2–4.
29. Gandhi, R.; Sharma, A.; Mahoney, W.; Sousan, W.; Zhu, Q.; Laplante, P. Dimensions of Cyber-Attacks: Cultural, Social, Economic, and Political. *IEEE Technol. Soc. Mag.* **2011**, *30*, 28–38. [[CrossRef](#)]
30. Geers, K. The Challenge of Cyber Attack Deterrence. *Comput. Law Secur. Rev.* **2010**, *26*, 298–303. [[CrossRef](#)]
31. Starr, S. Towards an Evolving Theory of Cyberpower. In *The Virtual Battlefield: Perspectives on Cyber Warfare*; IOS Press: Amsterdam, The Netherlands, 2009; pp. 18–52.
32. Marwick, A.; Lewis, R. *Media Manipulation and Disinformation Online*; Data & Society Research Institute: New York, NY, USA, 2017.
33. Bradshaw, S.; Howard, P. *Troops, Trolls and Troublemakers: A Global Inventory of Organized Social Media Manipulation*; University of Oxford: Oxford, UK, 2017.
34. Ko, M.C.; Chen, H.H. Analysis of Cyber Army's Behaviours on Web Forum for Elect Campaign. In *Asia Information Retrieval Symposium*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 394–399.
35. Schäfer, F.; Evert, S.; Heinrich, P. Japan's 2014 general election: Political bots, right-wing internet activism, and prime minister Shinzō Abe's hidden nationalist agenda. *Big Data* **2017**, *5*, 294–309. [[CrossRef](#)]
36. Hurtado, S.; Ray, P.; Marculescu, R. Bot Detection in Reddit Political Discussion. In Proceedings of the Fourth International Workshop on Social Sensing, Montreal, QC, Canada, 15 April 2019; pp. 30–35.
37. Johnson, T.; Kaye, B. Using is Believing: The Influence of Reliance on the Credibility of Online Political Information among Politically Interested Internet Users. *J. Mass Commun. Q.* **2000**, *77*, 865–879. [[CrossRef](#)]

38. Gil De Zuniga, H.; Puig-I-Abril, E.; Rojas, H. Weblogs, traditional sources online and political participation: An assessment of how the Internet is changing the political environment. *New Media Soc.* **2009**, *11*, 553–574. [[CrossRef](#)]
39. Pierce, D.R.; Redlawsk, D.P.; Cohen, W.W. Social influences on online political information search and evaluation. *Political Behav.* **2017**, *39*, 651–673. [[CrossRef](#)]
40. Berinsky, A. Rumors and Health Care Reform: Experiments in Political Misinformation. *Br. J. Political Sci.* **2017**, *47*, 241–262. [[CrossRef](#)]
41. Di Caprio, D.; Santos-Arteaga, F. Strategic Diffusion of Information and Preference Manipulation. In *Management Theories and Strategic Practices for Decision Making*; IGI Global: Hershey, PA, USA, 2013; pp. 40–58.
42. Li, X.; Hitt, L. Self-Selection and Information Role of Online Product Reviews. *Inf. Syst. Res.* **2008**, *19*, 456–474. [[CrossRef](#)]
43. Hu, N.; Bose, I.; Koh, N.S.; Liu, L. Manipulation of Online Reviews: An Analysis of Ratings, Readability, and Sentiments. *Decis. Support Syst.* **2012**, *52*, 674–684. [[CrossRef](#)]
44. Sarafidis, Y. What Have You Done for me Lately? Release of Information and Strategic Manipulation of Memories. *Econ. J.* **2007**, *117*, 307–326. [[CrossRef](#)]
45. Solorio, T.; Hasan, R.; Mizan, M. A case study of sockpuppet detection in wikipedia. In Proceedings of the Workshop on Language Analysis in Social Media, Atlanta, GA, USA, 13 June 2013; pp. 59–68.
46. Sockpuppet Investigations. 2019. Available online: [https://en.wikipedia.org/wiki/Wikipedia:Sockpuppet\\_investigations](https://en.wikipedia.org/wiki/Wikipedia:Sockpuppet_investigations) (accessed 20 May 2019).



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).