Roles of Information Propagation of Chinese Microblogging Users in Epidemics: A Crisis Management Perspective

Li Sun

Department of Computer and Information Science, Faculty of Science and Technology, University of Macau, Taipa, Macao

Patrick Cheong-Iao Pang

Victoria University Business School, Victoria University, Melbourne, Australia

(Corresponding Author; Email: mail@patrickpang.net)

Yain-Whar Si

Department of Computer and Information Science, Faculty of Science and Technology, University of Macau, Taipa, Macao

Abstract

Purpose: The purpose of this research is to investigate the usage characteristics and the information propagation patterns of Chinese microblogs in different stages of an epidemic, given that the microblogging in China is different from other parts of the world. In addition, we aim to conceptualize the roles of different users and provide insights for using microblogging platforms to disseminate information in this context.

Methodology: We conducted an analysis on Sina Weibo microblogs about the African Swine Fever epidemic from August to October 2018. We firstly applied a label propagation algorithm to classify users into government, media, verified users and non-verified users. We analyzed several user metrics, traced the information propagation patterns of their microblogs and calculated the average speed of information propagation using computational approaches.

Findings: Our findings show that different types of users played different roles, such as supplying information, amplifying information, relaying information and engaging with other users. The microblogs posted by media dominated the propagation in most cases, but general users can propagate information faster. The direction of information propagation is one-way for the majority of microblogs, and few users repost earlier information. Additionally, microblogs attract more attention at the beginning and the middle phases of an epidemic. In the context of managing epidemics, we recommend governments and other verified users can work together to use microblogging platforms efficiently.

Originality/Value: This research is one of the few studies to investigate information propagation patterns of different user categories on a Chinese microblogging platform during an epidemic. Our work can be used by government agencies and public health authorities for disseminating information efficiently during epidemics or emergencies, especially in the early stages.

Keywords: Microblog, Epidemics, Sina Weibo, Information Propagation

1. Introduction

Epidemics, as a type of public health crises, are often unpredictable risks to the society, as they are difficult to detect and manage while they can cause severe life losses and significant economic damage. In such cases, governments and medical services take actions to manage the crises by tracking down the sources of the diseases, stopping the contamination, issuing warnings and guidelines, and providing the public information about symptoms, treatments and prevention (Reynords and Weeger, 2005). As a part of crisis management, communication plays an important role in raising the awareness and reducing the uncertainty of the public (Panagiotopoulos *et al.*, 2016). Research shows that microblogging is increasingly used in the management of various crises (e.g. Leong *et al.*, 2015; Liu and Xu, 2018).

Most of the research of the use of microblogging during crises focuses on the sharing of information (Graham *et al.*, 2015) and the community use of microblogs (Leong *et al.*, 2015). Little extant literature investigates the use of microblogs by governments and authorities in the context of crisis management (Guo *et al.*, 2020; Medaglia and Zheng, 2017). Microblogs posted by governments can provide opportunities for exchanging information about the crises (Liu and Xu, 2018), broadcasting updates (Abedin and Babar, 2018), as well as fighting against rumors (Liu *et al.*, 2016). In China, there are hundreds of thousands of microblog accounts at the local, regional and central level of governments (Zheng and Zheng, 2014), and these accounts sparked strong reactions from the Chinese society in past public health events (Fung *et al.*, 2013). As such, we believe that microblogs from governments can play a role in managing an epidemic. Therefore, understanding and modeling the dynamics between governments and other users are useful for the efficient use of microblogging platforms during these scenarios.

We are particularly interested in the interactions of governments' microblogs with media, verified users (V-Users) and other general users in this study. Media and news organizations are proven opinion leaders on Chinese microblogging platforms (Nip and Fu, 2016). While individual users can initiate a discussion on a microblogging platform, media users can amplify their voices and lead the discussion. In addition, the Chinese microblog users demonstrate a trend to follow those at a higher or similar social level (Chen *et al.*, 2012) and V-Users dominate the broadcast of information of medical and health topics (Han and Wang, 2015). Therefore, V-Users are also our focus in this research. Drawing on this background, we expect that these different categories of microblog users disseminate information in different ways, and the findings can help to take advantage of the strengths of different user categories for epidemic management.

Researchers have used different metrics to study microblogging during crises, for example, numbers of followers (Wang *et al.*, 2015), levels of participation (Guo *et al.*, 2020), and the degree of citizen engagement (Chen *et al.*, 2020), but they seldom use the characteristics of information propagation to investigate. On microblogging platforms, users spread information by reposting microblogs (or retweeting, in Twitter parlance). When a microblog is reposted, its information propagation pattern can be analyzed to understand which users are exposed during the lifetime of the post. Through the analysis of information propagation patterns, we can understand how information is propagated on microblogging platforms. Specifically, few scholars examine the information propagation patterns on Chinese microblogging platforms and the roles of their users during epidemics. This presents a gap for our research to conceptualize these patterns and better understand the interactions among microblogging users in the Chinese context.

In this paper, we analyzed Chinese microblogs related to the outbreak of African Swine Fever (ASF), which is a wide-spread and fatal disease of pigs. Although ASF does not directly affect human health, this first ASF epidemic in China caused a great amount of misinformation and panic, therefore this epidemic was still considered a public health event. On August 3rd, 2018, the Ministry of Agriculture and Rural Affairs of China confirmed the first occurrence of ASF in Liaoning Province and the outbreak caused a huge amount of discussion on the Internet. The active discussion of this incident provides a unique opportunity for researchers to study how microblogs are used to manage this crisis. In our study, we focused on Sina Weibo which is the most popular microblogging platform with 200 million daily users (Sina Weibo Data Center, 2018). We refer to Sina Weibo as Weibo, and Weibo posts as microblogs hereafter in this paper.

We analyzed 17,679 microblogs in our data collection period, which began from the start of the ASF epidemic and ended two weeks after the end of the outbreak in Liaoning Province, China. We used a semi-supervised algorithm to differentiate the microblogs posted in different stages of the epidemic. To study the effects of different user groups, we also classified the authors of these microblogs into four user categories: government, media, V-Users and general users (also known as non-V-Users). Their information propagation patterns and the characteristics of their microblogs were computed in order to derive insights into their use of microblogging in this event.

We discover that each user type demonstrates different usage characteristics and information propagation patterns during various stages of the epidemic. While government accounts are often considered as the source of information, media accounts and other V-Users can disseminate information to a wider range of receivers.

However, the average speed of information propagation of general users is faster than media in most cases, thus their role is mainly relaying information. Media users are also observed to interact and engage with their audience. Additionally, microblogs attract more attention at the beginning and action stages of an epidemic, whereas the microblogs related to the end of an epidemic capture less attention. In practice, governments and V-Users can work together to optimize the information dissemination during public health crises, especially in the early stages.

The remainder of this article is structured as follows. In the upcoming sections, we present the literature review followed by our research design. We also explain how our data were collected and how we analyzed the dataset. Then, we report on our results and discuss their implications. In the final section, we summarize the findings and the limitations, as well as propose the directions for future work.

2. Literature Review

In this section, we first introduce the literature of crises, crisis management and epidemic control. Then, we present the concepts of information propagation patterns, followed by the overview of the algorithms for analyzing a large quantity of microblogs and their applications in the current study.

2.1. Microblogging during Crisis and Epidemics

Crises are events that are unpredictable and creates uncertainty to high priority goals of an organization such as a group, a company and even the society (Seeger *et al.*, 1998). The unpredictable and disruptive nature of crises prompts the needs of crisis management, which aims to identify unsolved problems in crises and reduce harm (Bundy *et al.*, 2016). Rapid and mass communication is the key to crisis management and managing an outbreak (Fung *et al.*, 2013; Reynords and Weeger, 2005). Crisis communication is event-oriented and aims to respond to the immediate information needs of the public (Reynords and Weeger, 2005). Because of the speed of conveying information and the wide coverage of audience, social media and microblogging platforms are recognized as tools that can support and enhance crisis management (Tim *et al.*, 2017). To deal with different types of crises, Coombs (2014) defines three major stages in the crisis management process, in which each stage has different sub-stages and actions to be taken (Table 1). A previous study suggests that content posted on microblogging platforms can align with these different stages of a crisis, thus microblogs are useful resources for the stakeholders who need to take action against the crisis (Imran *et al.*, 2015). This three-stage model forms the basis for our study to analyze the microblogs throughout the different time points of the ASF outbreak.

Nevertheless, public health crises (such as epidemics) are often unforeseeable or have a short trigger (Spence *et al.*, 2016), and the pre-crisis is less relevant in this case. Thus, we put our main focus on the crisis and post-crisis stages.

Table 1. Three stages of crisis management (Coombs, 2014; Pan and Meng, 2016)

Stage	Sub-stages
Pre-crisis	1. Signal detection – identifying warning signs 2. Prevention – issues management, risk aversion, and reputation management 3. Preparation – crisis management plans, crisis teams, and crisis portfolios
Crisis	1. A triggering event marking the beginning of the crisis 2. Crisis recognition – the crisis event is labelled and accepted 3. Crisis containment – responding strategies, contingency plans and follow-up concerns
Post-crisis	1. Better prepared for the next crisis 2. Building a positive impression of crisis management efforts to the public 3. Confirming that the crisis is over

Researchers have been studying microblogs posted during crises and epidemics. Li and others analyzed the content of microblogs related to earthquakes in China, and have found that the topics of information related to the earthquake have a significant influence on the propagation patterns in terms of scale and topological characteristics (Li *et al.*, 2018). Others have analyzed the information flow of healthcare-related news stories in order to study the relationship between V-Users and non-V-Users on Weibo, and found that V-Users have a dominant position of broadcasting health information (Han and Wang, 2015). Vorovchenko and colleagues show that rumors and misinformation can spread rapidly across social media, and public health authorities need to monitor social media to understand people's perceptions and beliefs about health issues and to intervene in a timely manner (Vorovchenko *et al.*, 2017). Additionally, in the health context, studies suggest that Twitter users post about testable claims, personal experiences and privacy concerns about digital health technologies (Pang *et al.*, 2020; Pang and Chang, 2019). Most of the existing studies investigate the use of microblogging for the communities and individuals impacted by the situation (Yates and Paquette, 2011) but not for governments or crisis-related authorities. As they often play an important role in crisis management, viewing the issue from their angles is essential for exploring the efficient ways of using microblogging in catastrophic scenarios.

2.2. Information Propagation of Microblogs

Microblogging platforms are used for both social networking and news distribution purposes, and

therefore they are now an important tool for people to access information (Kwak *et al.*, 2010). News media can use microblogs to spread the information quickly and users can get the latest information in real-time through microblogs. In many microblogging systems, information is transmitted by reposting microblogs. When a user posts a new microblog, this microblog will be immediately pushed to the interface of the followers of this particular user. Then, these followers can repost the same post and the information is pushed to another level of followers (Boyd *et al.*, 2010; Kwak *et al.*, 2010). Such reposting actions propagate information further in the social network, and information propagation patterns can be derived based on these hops at which a single microblog has traveled. Modeling the information propagation in microblogging platforms may lead to more effective use of such platforms and provide insights into the underlying sociology (Yu *et al.*, 2015). However, there are subtle differences in the implementations of reposting on Weibo and Twitter. An example is that Weibo users can comment under a microblog without broadcasting the comment to their followers, whereas Twitter users cannot comment without retweeting or replying (Guan *et al.*, 2014). In this case, the information propagation model of Twitter may not be applicable to Weibo.

Research has also shown that there are cultural differences between Twitter and Weibo. Previous literature suggests that Weibo has a higher level of activities from organizational users, whereas a majority of actions are performed by individual accounts on Twitter (Lin *et al.*, 2016). Additionally, Chinese microblogging users are found to use fewer hashtags and hyperlinks than Twitter users (Gao *et al.*, 2012). They adopt "queuing reposts" (i.e. appending different users' words in a repost) to amplify inclusive and special messages which is less seen on Twitter (Guan *et al.*, 2014). On the other hand, figures reveal that microblogging is increasingly popular among Chinese Internet users. As of February 2019, the number of microblog users in China has reached 350 million, accounting for 42.3% of the number of Internet users in the country (China Internet Network Information Center, 2019). The prevalence of microblogging services makes them an effective tool for crisis management in China, particularly for communicating and conveying the right information to the right audience (Guo *et al.*, 2020).

Information propagation on microblogging platforms is a popular research topic. Scientists have proposed a number of models to describe information propagation processes such as the susceptible-infected-susceptible model (Shi *et al.*, 2008), the susceptible-infected-removed model (Moreno *et al.*, 2002), and the susceptible-infected model (Zhao *et al.*, 2012). These models can be used to predict the scale and speed of propagation. On the other hand, for investigating the information propagation, complex microblogging

networks can be transformed into graphs, and deeper inspections are performed through subgraphs and network graphs. Among these methodologies, patterns in microblogs are often investigated with three-node subgraphs, which have strong interpretability even with complex network structures (Bild *et al.*, 2015; Itzkovitz and Alon, 2005). Tang et al. (2014) have proposed a dynamic microblogging network model and found that active microblog users can promote the distribution of information. Another research reports that the number of followers is consistent with a power-law distribution with an exponent close to two (Li *et al.*, 2015). This literature can generate information propagation patterns which allow us to further examine the user characteristics, the information propagation patterns and the speed of propagation.

2.3. Classification and Categorization of Microblogs and Microblog Users

Various machine learning and text mining algorithms have been used to classify the content of microblogs and the types of microblog users. Latent Dirichlet Allocation (LDA) is an unsupervised algorithm that identifies hidden topical information in large document collections or corpora based on unmarked samples (Blei *et al.*, 2003). LDA has been used for many research purposes, including categorizing themes of social media posts (Karami *et al.*, 2018) and extracting concepts from documents (Grimmer, 2010). Prior work has suggested that LDA is an effective model for discovering topics of microblogs, e.g. Twitter posts (Hong and Davison, 2010). Besides of LDA, Naïve Bayes classifier (McCallum and Nigam, 1998) and support vector machines (Bishop, 2006) are other traditional supervised learning classification methods, which can learn existing labels (also known as "gold standard") from the training dataset, and predict the labels for testing samples. Label Propagation Algorithm (Blei *et al.*, 2003) is a semi-supervised learning classification method with the similar usage that uses a mix of labeled data and unlabeled data to train the classifier, which is then used for prediction on the testing dataset. Drawing on the applications of these algorithms, we can iteratively derive the categories of microblogs and users, which are represented as different sets of labels, for our subsequent analyses.

3. Methods

We detail our approaches of data collection, classifying microblogs and users, and the metrics used for analyzing information propagation patterns and the characteristics of microblogging in this section.

3.1. Data Collection

We used the search function provided by Weibo to obtain microblogs about the ASF epidemic using a web

crawler. Since not all users used the official name of ASF when they posted their messages, we had to use a more inclusive keyword "swine fever" to collect the data. Irrelevant posts were also found in our data. Therefore, we used classifier algorithms (as detailed in the next sub-section) to remove non-relevant posts. We began collecting data from the Chinese government's announcement on the first occurrence of ASF in Liaoning Province (August 3rd, 2018). Since users continued to post relevant microblogs even after the government announced that the first ASF epidemic had ended (September 29th, 2018), we continued to collect data until 2 weeks after the government announced that the end of the first ASF epidemic (October 15th, 2018). By using this approach, the microblogs posted within a full lifecycle of the first ASF epidemic can be captured.

3.2. Content Classification and Pre-processing

We adopted the stages of crisis management by Coombs (2014) to classify the microblogs into three categories, namely Start, Action and End, as listed in Table 2. We mainly focused on the crisis and post-crisis stages and did not include pre-crisis for two reasons: (1) the pre-crisis stage involves detecting signs of the outbreak and prevention/preparation, which was difficult in this outbreak as it spread within a short period of time; (2) there were few microblogs describing the content of pre-crisis while the mainstream of themes was related to the epidemic and its aftermath.

Table 2. Three categories of microblogs during the ASF outbreak

Name	Definition	Crisis Stage in Coombs (2014)
Start	Announcement of the start of a local ASF epidemic	Crisis Stage → Crisis Recognition Sub-stage
Action	Actions from governments against the ASF outbreak (e.g. inspections, killing infected animals, policy changes, etc.)	Crisis Stage → Crisis Containment Sub-stage
End	Announcement of the end of a local ASF epidemic	Post-crisis Stage → Confirming that the crisis is over

Figure 1 shows a structure of a microblog on Weibo. For the purpose of our research, we extracted the content (including the original content and the user-added comments after reposting) of each microblog related to ASF and use the textual information to classify the microblogs based on the stages of the outbreak. We used a semi-supervised label propagation algorithm proposed by Zhu and Ghahramani (2002) to classify and cleanse our data. As part of this process, an iterative approach was applied in which 4,000 microblogs were inputted to the algorithm in each iteration, and the algorithm subsequently assigned a label for each microblog. Then, a

researcher evaluated the correctness of the labels by reading the microblogs and correct the type manually if the computer-generated one did not reflect their meaning. The verified microblogs were used to train the algorithm again, and then applied to the next batch of the unprocessed microblogs. This process repeated until all microblogs in our dataset were processed and examined by one researcher to ensure the relevance and the quality of the classification. Duplicated or irrelevant microblogs were also removed from our dataset.

Figure 1. An example of a microblog on Weibo

3.3. User Classification

As shown in Figure 2, the profile pages of Weibo users contain some basic information, but it cannot be directly used for analysis without further processing. For V-Users, their identities are confirmed by the microblogging platform and a short text-based description is assigned to each V-User to illustrate the nature of the account. Taking Figure 2 as an example, the translation of the verified description is “Official Weibo of CCTV News Center”. However, this field was found inconsistent among V-Users, e.g. some users had only the registered legal names of the organization/business which were hard to inform the groups in which these users belonged to. Also, this field could contain arbitrary text which increases the difficulty of classification. Apart from these limitations, there was no dedicated fields in user profile pages to show the categories of the users. Therefore, we could not use trivial approaches (e.g. keyword matching or regular expressions) to differentiate and classify these Weibo accounts. As such, we adopted a similar iterative approach described in

Section 3.2 to classify users using a machine learning algorithm. In this process, we made use of the verification description and the user-provided description as inputs for the semi-supervised algorithm to classify different types of V-Users. The combined text had a longer length which provided a better context for user classification.

Figure 2. User information on Weibo

As listed in Figure 3, Weibo users could be divided into V-Users and non-V-Users. For the purposes of our research, we further classified V-Users into three categories: *media, government* and *other V-Users* (who did not fall in either *media* or *government*), in order to compare their characteristics with the ones of non-V-Users. We particularly chose to separate *media* and *government* into standalone categories because the literature review has demonstrated their roles and functions in different crises. In terms of the nature of the content under this classification, *media* users are the ones who post news stories and communicate updates on Weibo about an epidemic, whereas *government* (both local and central governments) accounts mainly post official announcements, guidelines and updates.

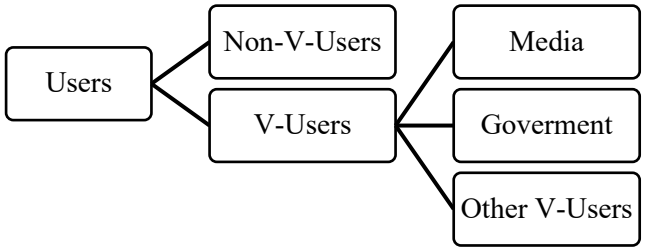


Figure 3. Relationship between non-V-Users and V-Users

3.4. User Characteristics

Microblog users demonstrate different characteristics in terms of their usage. Shen et al. (2014) have provided several measurements to quantify this usage. In this paper, we adopt three of them, namely the ratio of followee to follower (RFF), the ratio of original posts (ROP), and the proportion of nighttime posts (PNP). These calculations are based on the microblogs and the information on user profile pages collected from Weibo.

3.4.1. Ratio of Followee to Follower (RFF)

On microblogging platforms, the number of followees (*Followee_Count*) means the number of users who are interested in a particular microblog user; and the number of followers (*Follower_Count*) means the number of other users who follow the feed of a user. Equation 1 gives the formula to calculate the ratio between these two numbers of a user *u*. RFF ratio was adopted in our approach because it may have an implication on the propagation of microblogs.

$$RFF(u) = \lg \frac{Followee\_Count + 1}{Follower\_Count + 1} \tag{1}$$

3.4.2. Ratio of Original Posts (ROP)

Equation 2 represents the ratio between the number of original posts (*Original_Posts_Count*) and the number of total posts (*Total_Posts_Count*) of a particular user *u*. A lower ratio means that a user reposts more other users' posts instead of authoring their own. By using ROP, we can infer whether a user tends to publish their own content or reposts others with this metric.

$$ROP(u) = \frac{Original\_Posts\_Count + 1}{Total\_Posts\_Count + 1} \tag{2}$$

3.4.3. Proportion of Nighttime Posts (PNP)

Equation 3 shows the formula for PNP, which shows the ratio of nighttime posts (from 1:00 am to 7:00 am, as defined by Shen and others). This includes the reposted content of others' microblogs during the nighttime.

This metric can reflect whether a particular user is active at night.

$$PNP(u) = \frac{Night\_Posts\_Count + 1}{Total\_Posts\_Count + 1} \tag{3}$$

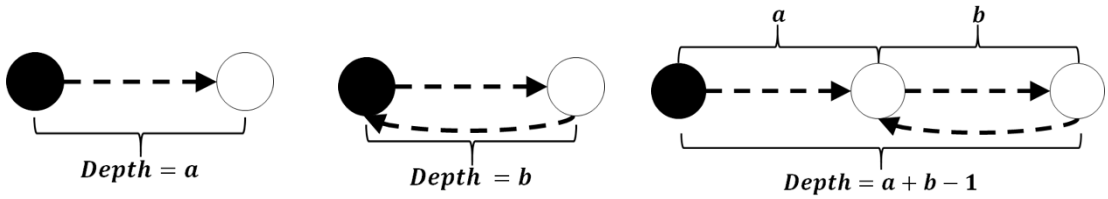
3.5. Information Propagation Patterns and Speed

In order to study the information propagation pattern, we use directed graphs to represent the reposting behavior of users. Figure 4 shows an example of such a notation. When a user (User A) posts a microblog and the microblog is reposted by another user (User B), the information flows from the author to the user reposting the microblog, which is illustrated by a block arrow in Figure 4(a). In contrast, a standalone node represents that a user's posts have never been reposted, as shown in Figure 4(b).

(a) User A has been reposted by User B (b) User A has never been reposted

Figure 4. Representing retweets using graph nodes

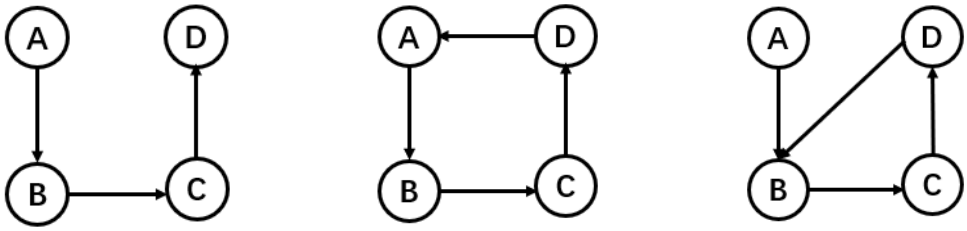
When a particular user posts a microblog, the followers of this user can see the microblog. Subsequently, if this microblog is further reposted, then the followers of the user who reposts can also see the same microblog. Microblogs can be propagated to the downstream followers across the platform in this way. When a microblog is reposted multiple times, an information propagation pattern can be derived in order to track who have read and reposted the microblog. Previous research has proposed a three-node connection subgraph to denote the detailed information propagation among three microblog users (Li *et al.*, 2018). We further extend this model and define three types of patterns of microblog retweets as shown in Figure 5. Among these patterns, Pattern 1 shows the path for one-way reposting, that is, a microblog is reposted by different users, and the repost stops after the number of hops; Pattern 2 illustrates that the microblog is reposted by two different users, and then the original author reposted the same microblog from the user at the end of the path; Pattern 3 is the combination of Patterns 1 and 2, in which a microblog is reposted by a series of users, and then the same microblog is reposted by a user in the middle of the chain. Depth indicates the number of users (i.e. hops) that a microblog has reached throughout its journey.



(a) Pattern 1 (b) Pattern 2 (c) Pattern 3

Figure 5. Patterns of information propagation

Figure 6 provides some examples for different information propagation patterns. Figure 6(a) is the simplest form of information propagation, which illustrates that a microblog posted by User A is reposted by User B, followed by C and D. This microblog passes through the original author and other three users, and its information propagation has a depth of 4. In this case, there is no interaction between the original author and other retweeting users and the reposting direction is one-way. The example in Figure 6(b) is similar to Figure 6(a), except that this microblog is reposted by the original author (User A). The depth is also 4 because four users in total are involved in the information propagation chain. We can observe that a closed loop is formed which shows that the original author of the microblog interacts with the last user in this information propagation. Figure 6(c) demonstrates a special case of patterns in which a microblog is reposted by a user in the middle of the chain. A triangular path is formed in this case.



(a) Pattern 1 (b) Pattern 2 (c) Pattern 3

Figure 6. Examples of information propagation patterns

By using the depth and the elapsed time of a microblog propagation, we can calculate the speed of the propagation given any information propagation pattern. Let *D* be the depth of a pattern and *T* be the elapsed time in hours, Equation 4 defines the speed *S* of a microblog *i*:

$$S_i = \frac{D_i}{T_i} \tag{4}$$

4. Results

We firstly describe an overview of our collected dataset, then report on the results of our data analyses in this section, including the analysis of user characteristics, the information prorogation patterns, and the speed

analyses of different types of microblogs.

4.1. Overview

Originally, we downloaded 40,952 microblogs in our data collection period, and 8,565 among them were discarded because they were duplicates of other microblogs. After the classification and data pre-processing steps, 17,679 unique posts were found to be ASF-related. The composition of their topics is shown in Table 3. These microblogs were posted by 9,224 users and we further categorized these users into four different types, which are listed in Table 4.

Table 3. Results of the classification of microblogs

Type	Number of Microblogs	Percentage
Start	8,789	49.7%
Action	8,635	48.8%
End	255	1.4%
Total	17,679	100.0%

Table 4. Results of user classification

	Number of Users	Percentage
Non-V-Users	4,950	53.7%
Media	1,805	19.6%
Government	1,097	11.9%
Other V-Users	1,372	14.9%
Total	9,224	100.0%

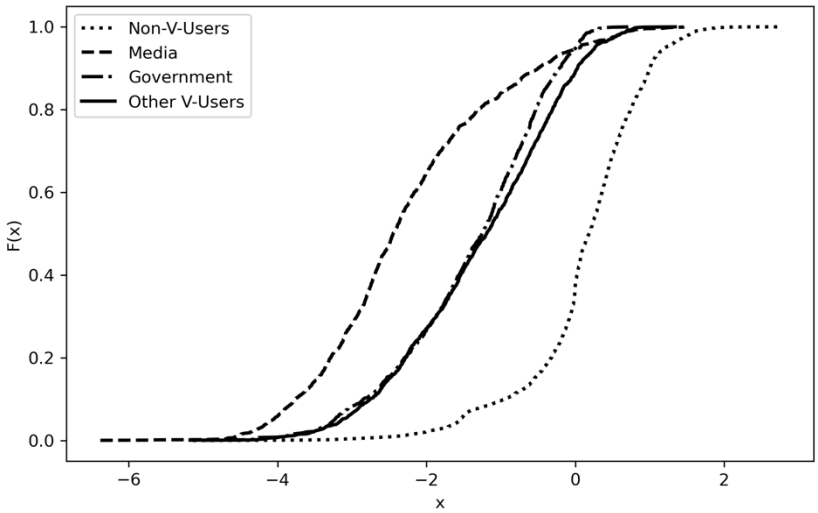
4.2. User Characteristics

In this sub-section, we present the results of various user characteristics, such as RFF, ROP and PNP. Table 5 shows the average values of these characteristics of different user categories. We can observe that the average RFF value of V-Users was negative while the average RFF value of non-V-Users was positive. These indicated that the number of followers was generally higher than the number of followees for V-Users, and vice versa for non-V-Users. The ROP numbers suggested that media and government accounts posted more original posts among V-Users, and non-V-Users had almost the same level of original posts to other V-Users. Finally, by referring to the higher PNP value of non-V-Users, we can identify that they had a higher ratio of nighttime posts than all V-Users.

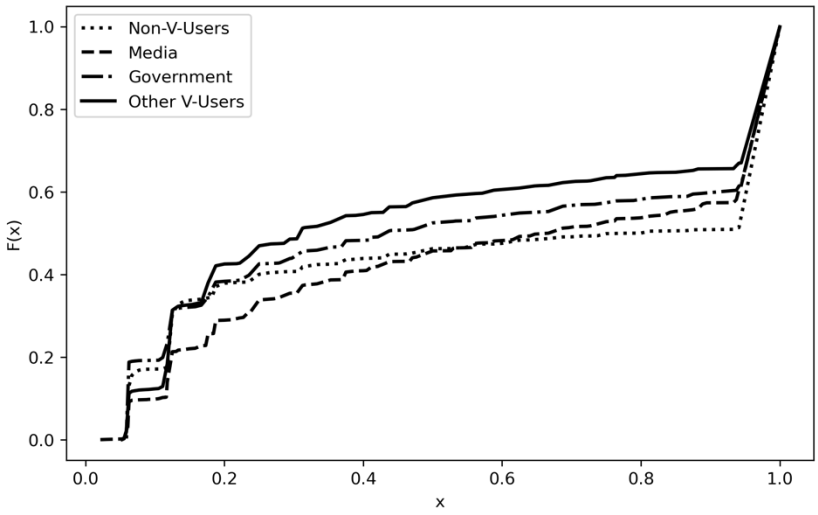
Table 5. Average values of user characteristics

User Features	Non-V-Users	Media	Government	Other V-Users
Ratio of Followee to Follower (RFF)	0.112	-2.275	-1.410	-1.291
Ratio of Original Posts (ROP)	0.491	0.596	0.540	0.497
Proportion of Nighttime Posts (PNP)	0.187	0.082	0.070	0.105

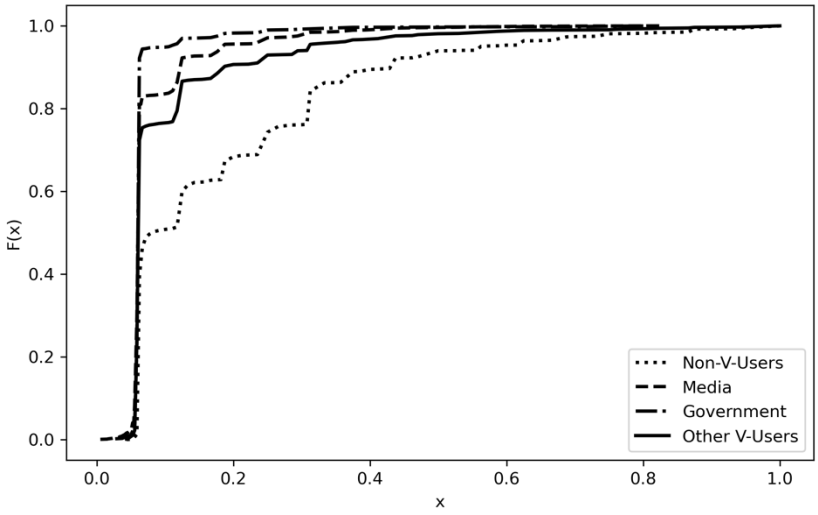
Figure 7 presents the plots of the cumulative distribution function (CDF) of each corresponding user characteristic. The CDF diagram of RFF (Figure 7(a)) shows that more than 90% of V-Users had RFF values lower than zero, which means that a very large portion of V-Users had more followers than followees; in contrast, only 40% of non-V-Users were the same. Among different types of V-Users, media users appeared to have the lowest RFF ratios than other categories of V-Users, i.e. they had larger differences between the numbers of followers and followees. Besides, as seen in Figure 7(b), about 60% media V-Users had lower ROP values than other user types, meaning that this proportion of media users had more original posts; whereas other users had higher ROP values which indicated that their reposts were more than self-authored microblogs. Finally, for PNP values shown in Figure 7(c), more than 80% of V-Users were associated with very low PNP values, reflecting that they were less likely to post at night; non-V-Users demonstrated a different pattern that they were more active at nighttime comparing with other user types.



(a) RFF



(b) ROP



(c) PNP

Figure 7. The CDF diagrams of different features of user categories

We further analyzed the time of posting microblogs in terms of different user types. Figure 8 presents the numbers of posts grouped by the hour of the day. We find that the numbers of microblogs posted at 1:00 am to 7:00 am were relatively low. On the other hand, for V-Users, the numbers of posts surged in the morning and the afternoon, while it went quieter at noon and in the evening. In contrast, the pattern of non-V-Users was more balanced with slight peaks during the morning and in the afternoon.

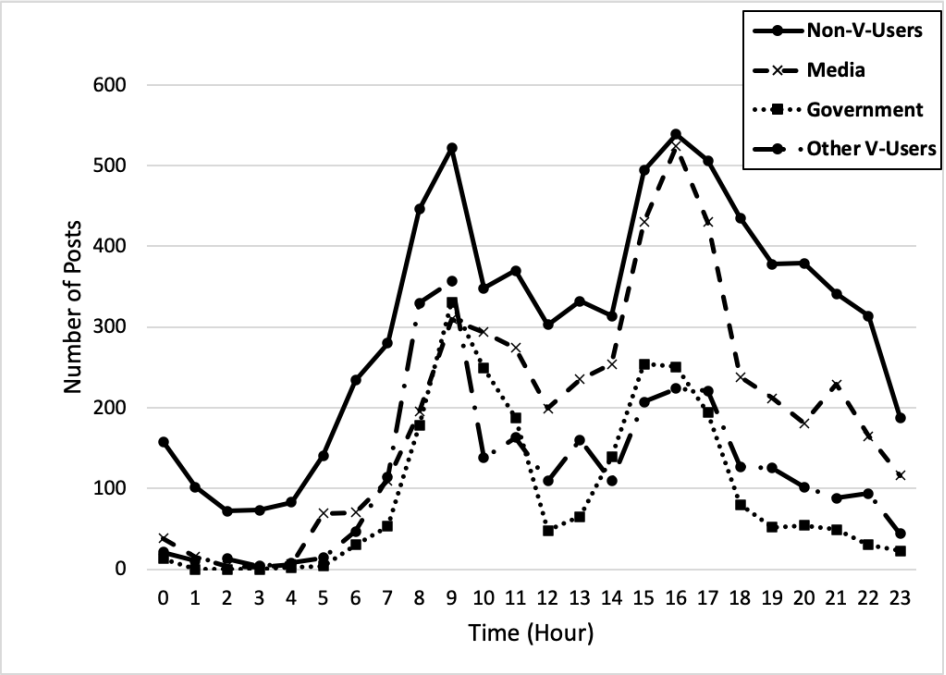


Figure 8. Microblog post time (in hour) of all types of users

4.3. Information Propagation Patterns

In this sub-section, we present the analyses on the three information propagation patterns with different stages and from different user categories. Table 6 details the numbers of microblogs in each type of patterns, grouped by different content types (i.e. the start of the epidemic, actions and the end of the epidemic) as well as the types of users.

Table 6. The numbers of microblogs of different information propagation patterns

User Type	Pattern 1 (One-way)			Pattern 2 (Self-repost)			Pattern 3 (Patterns 1 and 2 Co-exist)		
	Start	Action	End	Start	Action	End	Start	Action	End
Non-V-Users	78	47	0	26	10	0	1	0	0
Media	951	686	22	28	21	0	59	28	0
Government	167	365	1	2	8	0	2	0	0
Other V-Users	272	161	6	18	13	1	5	3	1

It can be observed that Pattern 1 (one-way reposting) is the type with the highest number of patterns regardless of the type of content of the microblog. Also shown in Table 6, media users dominated the posts of ASF microblogs in almost every stage and pattern type, while non-V-Users composed only a small number of

microblogs in this context. In addition, the numbers of microblogs in the *End* category (signaling the end of the epidemic) was very small compared with other content types.

In Figure 9, we plot the depth of these information propagation patterns for different types of users in various crisis stages. Pattern 1 had the highest number of Weibo in quantity among all types of patterns. We can also observe that the first peak was formed at the depth of 2 or 3, which means that information was propagated to two or three users and then the propagation stopped, i.e. no one continued to repost. On the other hand, for the microblogs posted by *media*, no matter in which stages, their maximum depth was larger than the other three types of users. In most of the cases, Pattern 1 demonstrated higher depth when comparing with other patterns. Finally, in terms of the crisis stages, microblogs in the *Action* stage could reach a higher depth, which means that each microblog in this stage was reposted more times than other stages; whereas the posts in the *End* stage did not attract many reposts in all cases.

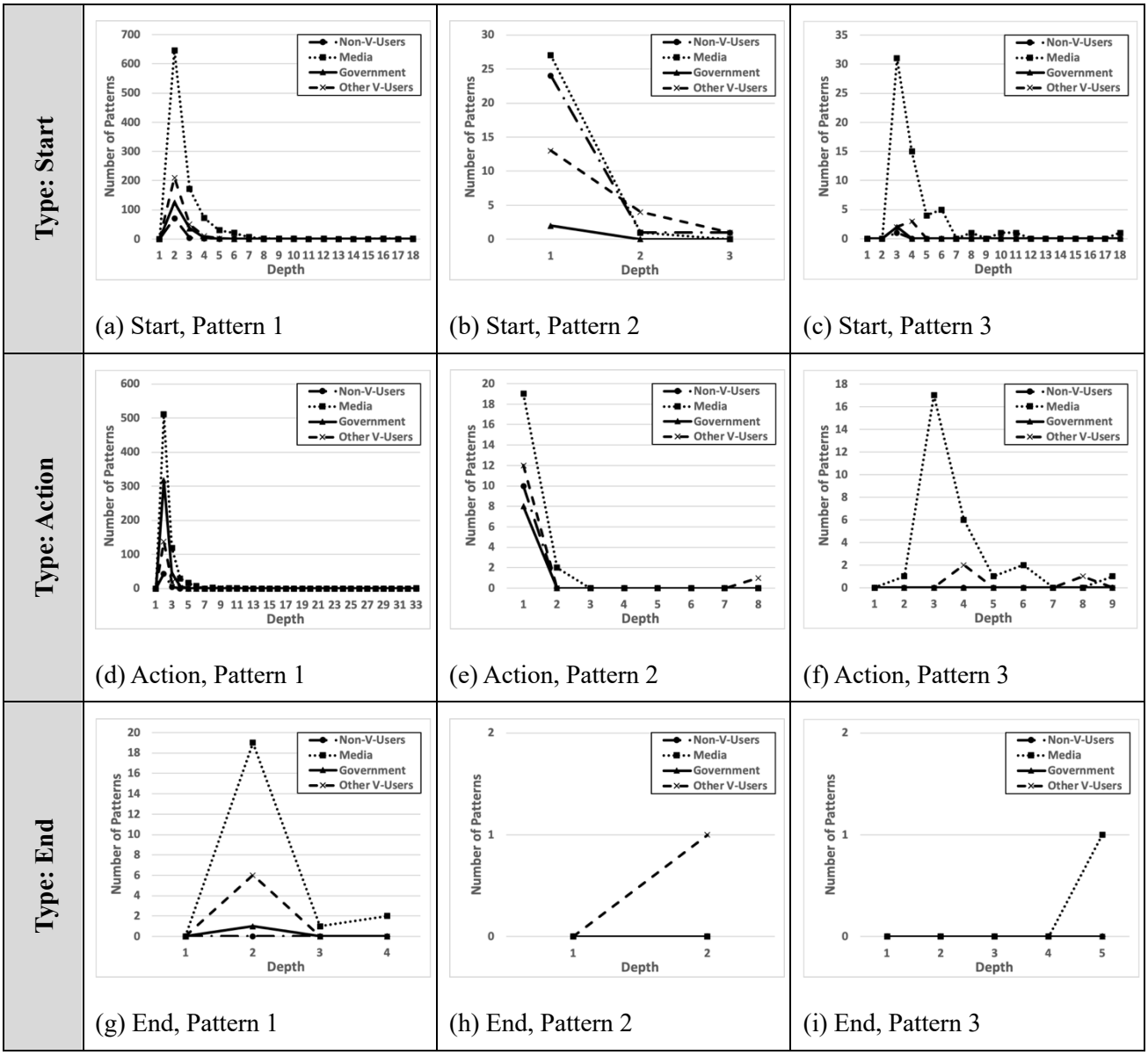


Figure 9. Results of the depth analysis of patterns

4.4. Average Propagation Speed

In this sub-section, we present our results on speed analyses of different information propagation patterns. The speed value depicts the depth (i.e. the number of users) a microblog travelled in an hour. A higher value means that a microblog can be seen by more users within a certain period of time. For each kind of information propagation patterns, we calculated the average speed of the microblogs with these patterns, average, standard deviation, and the 10% (slow tier) and the 90% (fast tier) values. In some cases, there was no sufficient data to calculate the speed and N/A is listed.

Table 7 shows the speed of Pattern 1 (i.e. one-way reposting). The speed at the 10th percentile was similarly low across different user types, representing that the speed had no large difference among users for slow

microblogs. For the microblogs in the *Start* stage, the mean and the 90th percentile values of non-V-Users were the highest, indicating that the microblogs written by non-V-Users were generally faster. Media coverage within the *Action* and the *End* stages exhibited higher speed than other types of users, showing that media had speed advantage for such content.

Table 7. Information propagation speed of Pattern 1 (higher is better)

Content Type	User Type	10th Percentile	90th Percentile	Mean	Standard Deviation	Number
Start	Non-V-Users	0.09	2.86	2.83	14.22	78
	Media	0.11	2.50	1.90	8.70	951
	Government	0.10	1.69	1.54	9.35	167
	Other V-Users	0.10	1.88	0.97	2.36	272
Action	Non-V-Users	0.10	1.11	0.46	0.47	47
	Media	0.10	2.35	1.91	9.53	686
	Government	0.10	1.97	1.14	3.22	365
	Other V-Users	0.10	1.06	1.04	3.47	161
End	Non-V-Users	N/A	N/A	N/A	N/A	0
	Media	0.10	2.22	0.85	1.46	22
	Government	0.10	0.10	0.10	N/A	1
	Other V-Users	0.10	0.35	0.18	0.11	6

As shown in Table 8, for Pattern 2 (self-reposting), the average speed of microblogs in the *Start* stage from non-V-Users, as well as the microblogs in the *Action* and *End* stages from other V-Users was generally faster. This means that these user categories quickly reposted their microblogs after another user reposted their content.

Table 8. Information propagation speed of Pattern 2 (higher is better)

Content Type	User Type	10th Percentile	90th Percentile	Mean	Standard Deviation	Number
Start	Non-V-Users	0.05	4.00	2.01	7.11	26
	Media	0.05	3.33	1.03	2.88	28
	Government	0.14	0.36	0.25	0.16	2
	Other V-Users	0.05	10.00	1.62	3.56	18
Action	Non-V-Users	0.04	0.34	0.16	0.13	10
	Media	0.10	0.59	0.34	0.20	21

	Government	0.04	0.61	0.31	0.16	8
	Other V-Users	0.04	1.11	1.00	2.72	13
End	Non-V-Users	N/A	N/A	N/A	N/A	0
	Media	N/A	N/A	N/A	N/A	0
	Government	N/A	N/A	N/A	N/A	0
	Other V-Users	0.11	0.11	0.11	N/A	1

For Pattern 3 (Table 9), our results show that the microblogs by media accounts, and they demonstrated the highest speed among different content types. However, there were not many microblogs mentioning the end of the epidemic in this pattern. Therefore, the speed of the *End* stage was not representative.

Table 9. Information propagation speed of Pattern 3 (higher is better)

Content Type	User Type	10^{th} Percentile	90^{th} Percentile	Mean	Standard Deviation	Number
Start	Non-V-Users	0.71	0.71	0.71	N/A	1
	Media	0.18	5.14	1.91	3.06	59
	Government	0.18	0.79	0.49	0.43	2
	Other V-Users	0.20	0.45	0.32	0.11	5
Action	Non-V-Users	N/A	N/A	N/A	N/A	0
	Media	0.20	1.88	3.17	11.27	28
	Government	N/A	N/A	N/A	N/A	0
	Other V-Users	0.20	4.71	2.01	2.38	3
End	Non-V-Users	N/A	N/A	N/A	N/A	0
	Media	1.50	1.50	1.50	N/A	1
	Government	N/A	N/A	N/A	N/A	0
	Other V-Users	N/A	N/A	N/A	N/A	0

To sum up, for Pattern 1 and Pattern 2, the microblogs describing the start of the epidemic authored by non-V-Users were disseminated faster. On the other hand, for the microblogs about taking actions and announcing the end of the epidemic, the microblogs posted by media were spread faster. For Pattern 3, all types of content were propagated faster with media accounts, however, this observation may not be representative due to the small amount of data in this pattern. In most of the patterns, governments' microblogs exhibited the lowest speed among V-Users.

5. Discussion

We discuss both the theoretical and practical implications of our results in this section, then conclude by presenting the limitations of our study.

5.1. Theoretical Implications

Our empirical study analyzed a Weibo dataset during an ASF epidemic in China. We have demonstrated that different categories of V-Users (i.e. media, government and other V-Users) as well as non-V-Users demonstrate distinct characteristics in their information propagation during this epidemic, and these characteristics can be conceptualized and provide insights for these different parties to better use microblogging for crisis management. Based on our results, we synthesize the characteristics of these users and propose their roles of information propagation in Table 10, followed by the elaboration of our conceptualization.

Table 10. Characteristics and roles of different types of users

User Type	Main Characteristics	Role
Non-V-Users	• Lower ratio of original posts • Active in the daytime and at night • High propagation speed in the Start stage • One-way information propagation	Relaying Information
Media	• Higher ratio of original posts • Active during business hours • High propagation speed in all stages with high depth • Multiple patterns of information propagation	Amplifying Information and Engaging with Other Users
Government	• Higher ratio of original posts • Active during business hours • Low speed and low depth • One-way information propagation	Supplying Information
Other V-Users	• Lower ratio of original posts • Active during business hours • High propagation speed in Start and Action stages • One-way information propagation but sometimes repost their own microblogs • Little presence in the End stage	Amplifying Information

Non-V-Users: This group of users post fewer original microblogs compared with other types of users. This implies that they do not compose much of their content but repost a significant number of microblogs from

other sources. They have the highest information propagation speed among all users in the beginning stage but not in other stages, reflecting that they are more interested in the information at the beginning of the epidemic and forward the information about the occurrence of the crisis to other people as quick as possible. Based on these observations, we propose that their role is to *relay* information during the information propagation process. Additionally, their information propagation is one-way (shown by the very few instances of Pattern 2 and 3) and they do not interact with older microblogs. This also indicates that they tend to relay and inform others, rather than engaging in discussions. Another specialty of this user group is that their active time is more scattered and can be seen throughout the day including evenings and nighttime. This contradicts the general findings of Liu et al. (2016) that Weibo users are more active in the daytime. Due to the rapid-changing nature of an epidemic, general users may wish to broadcast the updates of the epidemic regardless of the time of the day. Further research is needed to identify the causes of different usage patterns in a public health crisis.

Media: Media serves the purpose of reporting the facts and the development of a crisis and has consistent presence in all stages of an epidemic, which is different from other users who only shows interests in certain stages. Media users demonstrate the lowest RFF ratio among all types of users, reflecting their potential to broadcast information to a substantial number of followers during an epidemic. They also demonstrate their roles of content creators with higher ROP values, which represents that they post their own content more than reposting the information from others. In addition, media accounts have posted a large portion of microblogs of Pattern 2 or 3. This suggests that the information posted by media is more likely to stimulate discussions and interactions among downstream users by reposting, quoting and commenting. Previous literature finds that media and news outlets are often opinion leaders on Weibo (Nip and Fu, 2016), and our findings extend its applicability to the context of public health crisis management. A report of the World Health Organization (WHO) highlights the importance of effective communication and discussion through mass media in public health emergencies (World Health Organization, 2005). We share the similar point of view by reviewing our results and suggest that media can actually play the role of *reporting and engaging* with the public in the virtual world during a crisis.

Other V-Users: This group of users composes fewer original posts and reposts more information from other people. We can also observe that their information propagation patterns fall in Patterns 1 and 2. In addition to one-way propagation, this represents that they repost their own earlier posts and use them to provoke newer discussions. However, their activities can only be observed in the *Start* and *Action* stages during an epidemic

and they do not appear to have much interests in the *End* stage. With their relative fast information propagation speed, we propose that other V-Users can help to *amplify* information for reaching to a wider audience, especially in the early stages of a public health crisis. However, unlike media accounts, these V-Users do not engage with their audience and do not provide coverage consistently throughout the development of a crisis. In addition, the negative RFF value for this user group confirms the preliminary findings proposed by Chen et al. (2012) which suggests that the Chinese microblogging platforms are hierarchical in where users have preference towards V-Users.

Government: We observe that government accounts mainly *supply* information to other parties without many other activities and interactions on microblogging platforms, demonstrated by the minimal number of instances of Pattern 2 and 3. Although they show a negative ratio of RFF (i.e. having more followers than followees), their information propagation speed and depth are relatively low which affects their abilities to disseminating information as a result. Our analysis suggests that, in the Chinese microblogging context, the role of government accounts is to post announcements and first-hand updates about a crisis, because they control the resources (such as first responders and medical facilities) for providing accurate information at the scene. Instead of propagating the information widely by themselves, we argue that information announced by governments is picked up by media and other V-Users, and they subsequently broadcast the information, lead the discussions and engage with their followers.

In summary, our theoretical contributions include the four roles of information propagation in an epidemic: suppling, relaying, reporting and engaging, and amplifying respectively. Different user groups play a different role to allow information accessible by a wider audience.

5.2. Practical Implications

When conducting a campaign on microblogging platforms to fight against an epidemic, the above findings can be useful for understanding the capabilities of different users and leverage their strengths in various phases of an epidemic. Below we highlight some implications that can be leveraged to improve the information propagation of microblogs during public health crises.

Firstly, we can align the difference between media and non-V-Users (who are mostly general users) to disseminate information efficiently. Although media accounts can reach potentially more users by having a higher number of followers and greater depth of information propagation, their speed is not as fast as general users as shown in our results. Also, the activities of general users can be seen throughout the day while V-Users

are mainly active in the business hours only. In this case, we can leverage the round-the-clock activities of general users to propagate the content further. More specifically, if media can extend their service to the evening, other audience can pick up the information and it can be forwarded to other people fairly quick. This will greatly relieve the uncertainty of people during epidemics as timely information is important on such occasions. However, the strengths of media users can sometimes become their shortcomings. If inaccurate or false information is disseminated through media on microblogging platforms, such information can be spread quickly and may create undesirable outcomes. Therefore, government authorities and media need to ensure the accuracy of their content so that misinformation cannot easily be propagated to create unnecessary panic.

Secondly, while both governments and media have their presences on microblogging platforms, the latter has a stronger impact and a wider audience in terms of delivering information. Governments and public health authorities are normally considered as genuine information sources during epidemics. Since their abilities of propagating information are not as strong as media accounts, they can consider working with media and leverage their capacity to broadcast updates about epidemics, not only because they have more followers but also their microblogs are likely to be reposted more. This is also beneficial to the media because partnering with credible information sources in crises is known as a best practice (Seeger, 2006). On the other hand, because the functional role of media is to facilitate reassurance and reduce crisis-related uncertainty during crises, they should cooperate with governments to establish a rapid communication system for rapid response and recovery from crises (Pan and Meng, 2016).

We have observed earlier that most of the information propagation patterns are Pattern 1. This means that most of the information travels in one direction, that is consistent with the research of other crises (White and Fu, 2012). The lack of Pattern 2 and 3 also implies that Weibo users do not repost or engage with the microblogs posted earlier, which poses a challenge for engaging with the public during an epidemic. Therefore, when there are updates regarding an ongoing epidemic, the author should post new microblogs instead of adding extra information to existing threads, which are not revisited by the majority of users. Additionally, the average depth for different user categories and patterns ranges from 2 to 3, which shows that most of the microblogs will not be reposted after reaching two or three users along the information propagation path. This value is lower than the average depth found in other work (Yu *et al.*, 2015). As new information appears quickly and new microblogs are posted frequently in a crisis situation, and people focus on fresh information and stop reposting these microblogs after a short timeframe. Therefore, upstream users in the information propagation chain (such as

governments and media outlets) should keep their information frequently pushed to the public. In the context of risk communication, this approach can also help these microblog users to gain credibility during a crisis (Spence *et al.*, 2016).

Finally, the stage of an epidemic has an empirical effect on information propagation. Reflected by the higher maximum depth of the microblogs within the *Start* and *Action* stages, our study has revealed that microblog users have higher engagement with the microblogs about the start of epidemics and less with the microblogs about other stages. While media users have greater potential to push information further to other audience, people have more interactions with the microblogs about the start of epidemics, demonstrated by more reposts from non-V-Users. It appears that general users are more sensitive to this type of announcements. In line with other research, the use of microblogging platforms in the early stages of extreme events can reduce people's anxiety and minimize the opportunities of rumors spreading (Oh *et al.*, 2010; Vorovchenko *et al.*, 2017). Therefore, it is preferable to use microblogging platforms to broadcast the discovery of epidemics and make sure accurate information can be propagated, particularly at the beginning of an epidemic.

5.3. Limitations

The study has its limitations as with other empirical studies. First of all, the dataset obtained contained only the microblogs and the user information related to the first ASF epidemic in China. As the epidemic was ongoing, our dataset might not reflect the latest development and user responses. Secondly, we only included tweets related to announcements of the ASF outbreak and the actions taken against the outbreak in our analysis. Other microblogs such as personal reactions and commercial advertisements were excluded but they might exhibit different information propagation patterns. Future research may also include an analysis of the data posted by non-V-Users to gain more insights about the content they posted and their sentiment during public health events.

6. Conclusion

The main aim of this paper is to understand the information propagation patterns on China's largest microblogging platform during a public health crisis. As such, we used a series of metrics to investigate microblogging data by analyzing their characteristics and information propagation patterns at different stages of the crisis. In this research, we found that four types of users (i.e. government, media, other V-Users and non-V-Users) demonstrate different information propagation patterns and roles such as supplying, amplifying

and relaying information. We observe that most of the microblogs are mostly posted during the beginning and the middle stages of an epidemic. Similar to other research, media users are found to be highly effective in information propagation due to their higher numbers of followers and the greater depth of reposting. However, our findings add that general users show a higher speed in broadcasting information and are active round-the-clock. In addition, information travels mainly in one-way which makes updating inaccurate messages and engaging with audience challenging. Based on the analysis results provided in this article, governments and authorities can leverage the power of media and individual users on microblogging platforms to disseminate information effectively for public health crisis management, especially during the early stages.

7. Acknowledgements

This research was funded by the Research Committee of University of Macau, under grant MYRG2018-00246-FST. Author Pang completed part of this work during his appointment at the School of Computing and Information Systems of the University of Melbourne, and he felt grateful for their limited support for this research. The authors would also like to thank Prof. Kin Sun Chan from the University of Macau for his advice and contributions in this paper.

8. References

Abedin, B. and Babar, A. (2018), "Institutional vs. Non-institutional use of Social Media during Emergency Response: A Case of Twitter in 2014 Australian Bush Fire", *Information Systems Frontiers*, Vol. 20 No. 4, pp. 729–740.

Bild, D.R., Liu, Y., Dick, R.P., Mao, Z.M. and Wallach, D.S. (2015), "Aggregate Characterization of User Behavior in Twitter and Analysis of the Retweet Graph", *ACM Transactions on Internet Technology*, Vol. 15 No. 1, p. 4:1–4:24.

Bishop, C.M. (2006), *Pattern Recognition and Machine Learning*, Springer-Verlag, New York, NY, USA.

Blei, D.M., Ng, A.Y. and Jordan, M.I. (2003), "Latent Dirichlet Allocation", *Journal of Machine Learning Research*, Vol. 3 No. 1, pp. 993–1022.

Boyd, D., Golder, S. and Lotan, G. (2010), "Tweet, tweet, retweet: Conversational aspects of retweeting on twitter", *Proceedings of the 43rd Hawaii International Conference on System Sciences*, pp. 1–10.

Bundy, J., Pfarrer, M.D., Short, C.E. and Coombs, W.T. (2016), "Crises and Crisis Management: Integration, Interpretation, and Research Development", *Journal of Management*, Vol. 43 No. 6, pp. 1661–1692.

Chen, Q., Min, C., Zhang, W., Wang, G., Ma, X. and Evans, R. (2020), "Unpacking the black box: How to promote citizen engagement through government social media during the COVID-19 crisis", *Computers in Human Behavior*, Vol. 110, p. 106380.

Chen, Z., Liu, P., Wang, X. and Gu, Y. (2012), "Follow Whom? Chinese Users Have Different Choice", *ArXiv Preprint*, Vol. 12120167.

China Internet Network Information Center. (2019), *The 43rd Statistical Report on Internet Development in*

China.

Coombs, W.T. (2014), *Ongoing Crisis Communication: Planning, Managing, and Responding*, Sage Publications, Thousand Oaks, CA.

Fung, I.C.-H., Fu, K.-W., Ying, Y., Schaible, B., Hao, Y., Chan, C.-H. and Tse, Z.T.-H. (2013), "Chinese social media reaction to the MERS-CoV and avian influenza A (H7N9) outbreaks", *Infectious Diseases of Poverty*, Vol. 2 No. 1, p. 31.

Gao, Q., Abel, F., Houben, G.-J. and Yu, Y. (2012), "A Comparative Study of Users' Microblogging Behavior on Sina Weibo and Twitter", in Masthoff, J., Mobasher, B., Desmarais, M.C. and Nkambou, R. (Eds.), *User Modeling, Adaptation, and Personalization*, Springer Berlin Heidelberg, Berlin, pp. 88–101.

Graham, M.W., Avery, E.J. and Park, S. (2015), "The role of social media in local government crisis communications", *Public Relations Review*, Vol. 41 No. 3, pp. 386–394.

Grimmer, J. (2010), "A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases", *Political Analysis*, Vol. 18 No. 1, pp. 1–35.

Guan, W., Gao, H., Yang, M., Li, Y., Ma, H., Qian, W., Cao, Z., *et al.* (2014), "Analyzing user behavior of the micro-blogging website Sina Weibo during hot social events", *Physica A: Statistical Mechanics and Its Applications*, Vol. 395, pp. 340–351.

Guo, J., Liu, N., Wu, Y. and Zhang, C. (2020), "Why do citizens participate on government social media accounts during crises? A civic voluntarism perspective", *Information & Management*, p. 103286.

Han, G. and Wang, W. (2015), "Mapping user relationships for health information diffusion on microblogging in China: A social network analysis of Sina Weibo", *Asian Journal of Communication*, Vol. 25 No. 1, pp. 65–83.

Hong, L. and Davison, B.D. (2010), "Empirical study of topic modeling in Twitter", *Proceedings of the First Workshop on Social Media Analytics - SOMA '10*, pp. 80–88.

Imran, M., Castillo, C., Diaz, F. and Vieweg, S. (2015), "Processing Social Media Messages in Mass Emergency: A Survey", *ACM Computing Surveys*, Vol. 47 No. 4, available at:https://doi.org/10.1145/2771588.

Itzkovitz, S. and Alon, U. (2005), "Subgraphs and network motifs in geometric networks.", *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, Vol. 71 No. 2 Pt 2, p. 026117.

Karami, A., Dahl, A.A., Turner-McGrievy, G., Kharrazi, H. and Shaw, G. (2018), "Characterizing diabetes, diet, exercise, and obesity comments on Twitter", *International Journal of Information Management*, Vol. 38 No. 1, pp. 1–6.

Kwak, H., Lee, C., Park, H. and Moon, S. (2010), "What is Twitter, a social network or a news media?", *Proceedings of the 19th International Conference on World Wide Web*, pp. 591–600.

Leong, C.M.L., Pan, S.L., Ractham, P. and Kaewkitipong, L. (2015), "ICT-enabled community empowerment in crisis response: Social media in Thailand flooding 2011", *Journal of the Association for Information Systems*, Vol. 16 No. 3, p. 1.

Li, L., Zhang, Q., Tian, J. and Wang, H. (2018), "Characterizing information propagation patterns in emergencies: A case study with Yiliang Earthquake", *International Journal of Information Management*, Vol. 38 No. 1, pp. 34–41.

Li, Y., Qian, M., Jin, D., Hui, P. and Vasilakos, A.V. (2015), "Revealing the efficiency of information diffusion in online social networks of microblog", *Information Sciences*, Vol. 293, pp. 383–389.

Lin, X., Lachlan, K.A. and Spence, P.R. (2016), "Exploring extreme events on social media: A comparison of user reposting/retweeting behaviors on Twitter and Weibo", *Computers in Human Behavior*, Vol. 65, pp. 576–581.

Liu, F. and Xu, D. (2018), “Social Roles and Consequences in Using Social Media in Disasters: a Structurational Perspective”, *Information Systems Frontiers*, Vol. 20 No. 4, pp. 693–711.

Liu, Y., Wang, B., Wu, B., Shang, S., Zhang, Y. and Shi, C. (2016), “Characterizing super-spreading in microblog: An epidemic-based information propagation model”, *Physica A: Statistical Mechanics and Its Applications*, Vol. 463, pp. 202–218.

McCallum, A. and Nigam, K. (1998), “A comparison of event models for Naive Bayes text classification”, *Proceedings of AAAI-98 Workshop on Learning for Text Categorization*, Vol. 752, pp. 41–48.

Medaglia, R. and Zheng, L. (2017), “Mapping government social media research and moving it forward: A framework and a research agenda”, *Government Information Quarterly*, Vol. 34 No. 3, pp. 496–510.

Moreno, Y., Pastor-Satorras, R. and Vespignani, A. (2002), “Epidemic outbreaks in complex heterogeneous networks”, *The European Physical Journal B-Condensed Matter and Complex Systems*, Vol. 26 No. 4, pp. 521–529.

Nip, J.Y.M. and Fu, K. (2016), “Challenging Official Propaganda? Public Opinion Leaders on Sina Weibo”, *The China Quarterly*, Vol. 225, pp. 122–144.

Oh, O., Kwon, K.H. and Rao, H.R. (2010), “An Exploration of Social Media in Extreme Events: Rumor Theory and Twitter during the Haiti Earthquake 2010”, *Proceedings of International Conference on Information Systems (ICIS 2010)*, Vol. 231, pp. 7332–7336.

Pan, P.-L. and Meng, J. (2016), “Media Frames across Stages of Health Crisis: A Crisis Management Approach to News Coverage of Flu Pandemic”, *Journal of Contingencies and Crisis Management*, Vol. 24 No. 2, pp. 95–106.

Panagiotopoulos, P., Barnett, J., Bigdeli, A.Z. and Sams, S. (2016), “Social media in emergency management: Twitter as a tool for communicating risks to the public”, *Technological Forecasting and Social Change*, Vol. 111, pp. 86–96.

Pang, P.C.-I. and Chang, S. (2019), “The Twitter Adventure of #MyHealthRecord: An Analysis of Different User Groups During the Opt-Out Period”, *Studies in Health Technology and Informatics*, Vol. 266, pp. 142–148.

Pang, P.C.-I., McKay, D., Chang, S., Chen, Q., Zhang, X. and Cui, L. (2020), “Privacy Concerns of the Australian My Health Record: Implications for Other Large-scale Opt-out Personal Health Records”, *Information Processing and Management*, Vol. 57 No. 6, p. 102364.

Reynords, B. and Weeger, M.W. (2005), “Crisis and Emergency Risk Communication as an Integrative Model”, *Journal of Health Communication*, Vol. 10 No. 1, pp. 43–55.

Seeger, M.W. (2006), “Best practices in crisis communication: An expert panel process”, *Journal of Applied Communication Research*, Vol. 34 No. 3, pp. 232–244.

Seeger, M.W., Sellnow, T.L. and Ulmer, R.R. (1998), “Communication, Organization, and Crisis”, *Annals of the International Communication Association*, Vol. 21 No. 1, pp. 231–276.

Shen, Y., Yu, J., Dong, K. and Nan, K. (2014), “Automatic Fake Followers Detection in Chinese Micro-blogging System”, *Proceedings of Pacific-Asia Conference on Knowledge Discovery and Data Mining 2014*, pp. 596–607.

Shi, H., Duan, Z. and Chen, G. (2008), “An SIS model with infective medium on complex networks”, *Physica A: Statistical Mechanics and Its Applications*, Vol. 387 No. 8–9, pp. 2133–2144.

Sina Weibo Data Center. (2018), *2018 Weibo User Development Report*, Beijing, China.

Spence, P.R., Lachlan, K.A., Edwards, A. and Edwards, C. (2016), “Tweeting Fast Matters, But Only if I Think About It: Information Updates on Social Media”, *Communication Quarterly*, Vol. 64 No. 1, pp. 55–71.

Tang, M., Mao, X., Yang, S. and Zhou, H. (2014), "A dynamic microblog network and information dissemination in '@' mode", *Mathematical Problems in Engineering*, Vol. 2014, p. 492753.

Tim, Y., Pan, S.L., Ractham, P. and Kaewkitipong, L. (2017), "Digitally enabled disaster response: the emergence of social media as boundary objects in a flooding disaster", *Information Systems Journal*, Vol. 27 No. 2, pp. 197–232.

Vorovchenko, T., Ariana, P., van Loggerenberg, F. and Amirian, P. (2017), "#Ebola and Twitter. What Insights Can Global Health Draw from Social Media?", in Amirian, P., Lang, T. and van Loggerenberg, F. (Eds.), *Big Data in Healthcare*, Springer, Cham, pp. 85–98.

Wang, H., Lei, K. and Xu, K. (2015), "Profiling the followers of the most influential and verified users on Sina Weibo", *Proceedings of IEEE International Conference on Communications (ICC 2015)*, pp. 1158–1163.

White, J.D. and Fu, K.-W. (2012), "Who do you trust? Comparing people-centered communications in disaster situations in the United States and China", *Journal of Comparative Policy Analysis: Research and Practice*, Vol. 14 No. 2, pp. 126–142.

World Health Organization. (2005), *Effective Media Communication during Public Health Emergencies: A WHO Field Guide*, World Health Organization, Geneva.

Yates, D. and Paquette, S. (2011), "Emergency knowledge management and social media technologies: A case study of the 2010 Haitian earthquake", *International Journal of Information Management*, Vol. 31 No. 1, pp. 6–13.

Yu, Z., Wang, Z., He, H., Tian, J., Lu, X. and Guo, B. (2015), "Discovering Information Propagation Patterns in Microblogging Services", *ACM Transactions on Knowledge Discovery from Data*, Vol. 10 No. 1, p. 7.

Zhao, L., Guan, X. and Yuan, R. (2012), "Modeling collective blogging dynamics of popular incidental topics", *Knowledge and Information Systems*, Vol. 31 No. 2, pp. 371–387.

Zheng, L. and Zheng, T. (2014), "Innovation through social media in the public sector: Information and interactions", *Government Information Quarterly*, Vol. 31 Supplement 1, pp. S106–S117.

Zhu, X. and Ghahramani, Z. (2002), *Learning from Labeled and Unlabeled Data with Label Propagation (Technical Report CMU-CALD-02-107)*, Carnegie Mellon University, Pittsburgh, PA.