Propagation structure feature of entertainment news in the Weibo online social network

ZILONG ZHAO(a)

School of Reliability and Systems Engineering, Beihang University - Beijing 100191, China and China Academy of Launch Vehicle Technology - Beijing 100076, China

received 13 January 2021; accepted in final form 29 April 2021
published online 15 September 2021

Abstract – Entertainment news caters to people's desires for leisure and gossip. Recently, with the Internet popularization, the online social network has become a convenient platform for netizens to receive and spread entertainment news. Previous study efforts of entertainment news focused on analyzing the psychological background and the economic significance of entertainment news. However, little attention was paid to realistic propagation structure features of entertainment news, which reflect the essential propagation mechanisms. This research analyzes different types of news in the online social network, and entertainment news and other news are compared with respect to different structural features. It is found that entertainment news propagates with more indirect re-postings and a polycentric structure. These unique characteristics are stable for different stages of propagation. This structural feature offers us a better understanding of entertainment propagation and may inspire the further analysis of entertainment news based on the propagation trace.

open access Copyright © 2021 The author(s)

Published by the EPLA under the terms of the Creative Commons Attribution 4.0 International License (CC BY). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Introduction. – The origin of entertainment news can be traced back to the 1830s when Benjamin Day established the first successful penny — The New York Sun in the United States in 1833, marking the beginning for the era of the penny press [1]. The major coverage of entertainment news on penny press added to its readership and success [2]. Alfred Harmsworth founded the Daily Mirror in 1903 and used “tabloid” to refer to it [3]. Since the paper was half the size of broadsheets with soft and entertainment news, the word “tabloid” gradually began to refer to the small-sized newspaper, reporting celebrities’ private life including emotional situation, sexual scandals, changes of the appearance, violent crimes, weird stuff and all kinds of advertisements [4]. Developing from the newspaper, entertainment news gradually penetrated into the field of other media like radio, television, and the online media nowadays [5]. The emergence and growth of the Internet also accelerated the spread and popularity of entertainment news [6]. Nowadays, the trend towards entertainment has overwhelmed the whole society in various perspectives, which offers general public relaxation and recreation. Acting as a complex system [7], the online social network also accelerates the spread of entertainment information and expands the entertainment market as well [8]. The giant social network companies in the USA, such as Facebook and Twitter, promote the explosive and overwhelming spread of entertainment news to reach a new level and propagate to the whole world [9]. For example, the Twitter social network has 500 million tweets per day, and 74% of Twitter users claim they use the network to get their news [10]. The most followed person on Twitter has already more than one hundred million followers. As for the Northeast Asian countries like China, Japan and South Korea, as the pillar industry, the entertainment industry helps entertainment news to penetrate into people's daily life [11]. By the end of 2018, Sina Weibo, the most famous social network site of China, announced it possessed 200 million daily active accounts in its official reports. And the entertainment news is the top reading topic for users [12]. The entertainment news propagation redefines the entertainment industry that engages in content innovation and media entrepreneurship to incubate media brands and aggregate fan communities [13].

(a)E-mail: zhaozilong@buaa.edu.cn (corresponding author)
In the new era of the Internet, studies about the entertainment news in the social networks are mainly from the psychology and economy area rather than the propagation pattern. For instance, with respect to the psychology of entertainment preferences, individuals have differences in preferences for such entertainment [14]. And the drive for entertainment on Facebook [15] and Wechat [16] is found to be related to extrovert personality and narcissism. And the entertainment satisfies people’s need for group behavior and conformity psychology [17]. Therefore, many users consider social network sites as a platform for fun and become sticking on the social network site. For example, a study based on the QQ zone of China found that entertainment motivation could be a major predictor of social network sites’ usage [18]. And in Belgium, Facebook use is positively related to entertainment-oriented purpose especially for adolescents [19]. Some social network sites even use entertainment news to attract and bind users: an up-to-date entertainment website named Celebrity Watch selects important readers and finds potential customers based on their browsing history [20]. The entertainment news not only changes the social network site but also influences the whole media industries [13]. Traditional entertainment spreads a one-directional message, whereas online social networks are the bi-directional platform where people congregate to communicate gossip and entertainment news [21].

Apart from the psychology and economy research mentioned above, the study of structural features is also indispensable and provides the opportunity to understand the infectious mechanism of entertainment. For example, many works analyze information propagation by the epidemic model [22–25]. Some researches focus on the propagation structure features of a specific kind of news. For example, some research studies the fake news and find its structural features are significantly different from those of the other news [26–28]. The propagation of fake news on Twitter is found to be farther, faster, deeper, and broader than the real news. And fake news inspired fear, disgust, and surprise [27]. Interestingly, some structural features may emerge at an early stage [29]. Apart from the fake news we mentioned, the entertainment news is another kind of infectious news, which has a great influence on people’s modern lives. The topology of entertainment news propagation reflects the essential properties and mechanism of how this influence spreads. However, the structural features of entertainment news have rarely been analyzed.

On the one hand, the studies of entertainment news are mainly from psychology and economy perspectives rather than propagation structure. On the other hand, the research about structural features is rarely focused on entertainment news. As a result, in this work, we analyze the entertainment news of the social networks from the structural perspective. We select realistic entertainment news and compare it with other news in the perspective of topology. The entertainment news networks are found to have particular structural features: the ratio between the indirect and direct re-postings of entertainment news is relatively larger than other types of news and the networks of entertainment news often have a smaller Gini coefficient [30]. This indicates that entertainment news is driven by multiple users, which describes the propagation pattern. We also test these features in the early stage of propagation. These unique structural features of entertainment news will help us to understand its infection mechanism better.

**Methods.**

*Dataset description.* The dataset of anonymous users and their retweeting traces are collected through the Sina Weibo social network platform, which is the largest social network of China. In this work, we choose relatively larger networks (greater than or equal to 200 re-postings) to analyze for they have a certain influence. Hence, we study 1691 propagation news with 1.5 million re-postings from 01.01.2016 to 08.30.2018. The users are selected randomly without considering the types of users. More details are shown in table 1.

**Definition of entertainment news.** Acting as the top popular reading topic for users [12], entertainment news mainly reports celebrities’ gossip, scandals, and aims to arouse the ordinary audience’s interest and caters to their taste. Specifically, in this work, the dataset containing 1691 news networks is classified according to their content: politics news, economy news, social news, entertainment news as well as education and knowledge news. Due to the importance of news topics for this work, we carefully classify the dataset manually, with a three-person voting model to avoid subjective bias. If three judges fail to give the same conclusion, then this model observes the principle of minority obeying majority. If the content of one network contains information with several topics, the manual classification could select the main type conveniently from the reader’s perspective. Because of the importance of news topic and the convenience of selecting the most important topic, we choose a manual three-person voting model rather than natural language processing in this work. Due to a huge dataset, other automatic approaches such as the natural language processing are required. After the classification, there are 479 entertainment news

| Number of users | Entertainment news | The other news |
|-----------------|-------------------|----------------|
| Number of re-postings | 581311 | 990174 |
| Number of propagation networks | 479 | 1212 |
networks. Because the other four kinds of news have similar structural features, we study them together in order to compare them to entertainment news more conveniently. Hence there are 1212 other news networks. The information about the number of anonymous users and re-postings are shown in the table 1.

**Network establishment.** We plot the schematic diagram to explain how to establish the re-posting trace network as shown in fig. 1(a). In the propagation network, the node stands for user and the edge stands for re-posting relationship. This propagation network is a directed network with the direction of information spreading (from source to receiver). The direction and other concepts such as the layer, direct re-postings and indirect re-postings are shown in this figure. This example network contains eight users as well as seven re-postings among them.

**Ratio between the indirect and direct re-postings.** The layer describes group re-postings that share the same distance from the creator. As shown in fig. 1(a), the seven re-postings could be divided into layer one, layer two and layer three. The re-postings from the first layer are actually the direct re-postings (out-degree) from the creator, whereas the re-postings from the rest of the layers such as layer two and layer three represent the indirect re-postings. Therefore, the ratio between the indirect and direct re-postings is defined as

$$r = \frac{\text{size of indirect re-posting}}{\text{size of direct re-posting}} = \frac{n_i}{n_d},$$  \hspace{1cm} (1)

where \(n_d\) is the size of direct re-posting and \(n_i\) is the size of indirect re-posting.

**Concentration ratio.** We measure the concentration ratio of the propagation network by the Herfindahl-Hirschman Index [31] from the network perspective. One given network has one concentration ratio, and it is defined as the sum of the squares of the out-degrees of the nodes within the given propagation network. The concentration ratio is defined as

$$\text{Concentration ratio} = \sum_{j=1}^{N} \left( \frac{k_i^{\text{out}}}{L} \right)^2,$$

$$L = \sum_{j=1}^{N} k_j^{\text{out}},$$  \hspace{1cm} (2)

where \(N\) is the number of nodes in the network and \(k_j^{\text{out}}\) is the out-degree of the node number \(j\).

**Gini coefficient [30].** The Gini coefficient is initially used as a single parameter aimed at measuring the degree of inequality of income distribution. Based on this, we analyze the distribution of the out-degree of nodes within the given network to plot the Lorenz curve. The Lorenz curve is originally a graph showing the proportion of overall income or wealth assumed by the bottom \(x\)% of the people. Here it is shown for the bottom \(x\)% of nodes what percentage (\(y\)% of the total out-degree these

**Fig. 1:** The typical propagation networks of entertainment news. (a) A schematic diagram of direct and indirect re-postings of the propagation networks. The creator is marked especially by node a. Node b reposts node a’s message, hence producing a direct re-posting, whose direction is the information spreading direction (from source to receiver). Also, node e reposts node b’s message, creating an indirect re-posting accordingly. The re-postings could be divided into different layers by the distance from the creator. (b) This network contains 454 nodes and 514 edges, and its content is about publicity shots of a famous Chinese actor. The creator “i” is marked. We use Pajek software with the 2D Fruchterman-Reingold layout to plot these figures. (c) Another typical example for the propagation network of entertainment news with 722 nodes and 732 edges. The creator “j” is marked. This news is a video about a Korea TV series.
nodes have. Theoretically, the Gini coefficient can range from zero (complete homogeneity) to one (complete heterogeneity). Furthermore, we use the complementary Gini coefficient to replace the Gini coefficient in order to better demonstrate its distribution by logarithmic coordinates. The complementary Gini coefficient is defined as

\[
\text{complementary Gini coefficient} = 1 - G,
\]

where \( G \) is the Gini coefficient

\[
G = \frac{A}{A + B} = 1 - 2B,
\]

where \( A \) is the area that lies between the line of equality line and the Lorenz curve and \( B \) is the area under the line of the Lorenz curve,

\[
B = \int_0^1 ydx,
\]

with \( x \) the cumulative portion of nodes ranked by out-degree and \( y \) the cumulative portion of out-degrees.

Results. – To establish the news propagation network, we apply an approach based on re-posting relationships among users regarding the same news [29]. Based on the identified propagation networks, we demonstrate the propagation networks of entertainment news in fig. 1. For example, in fig. 1(b), the network is evolving with the participation of not only the creator “i” but also other influential re-posters. Apart from the creator, the latter re-posters could also play an important role in the further spreading of entertainment news. To check this, the proportion of indirect re-postings in entertainment networks is studied.

The investigation of the layer in typical propagation networks helps to understand the distribution of entertainment news. For the whole lifespan (fig. 2(a)) and the early stage (fig. 2(b)), we study the layer distribution for entertainment news and other news, respectively. Although the first layer is dominating, entertainment news networks tend to have a relatively smaller first (direct) layer and larger indirect layers. Moreover, considering the influential re-posters (their out-degree is above 5% of the whole size of the given propagation network) rather than all the re-postings, we first count the number of influential nodes and all the nodes for the different layers. And then we calculate the proportion of influential nodes. As shown in fig. 2(c), the proportions of influential re-posters are generally higher in entertainment news than the other news. As a result, the layer distributions show different evolution paths between entertainment and other news. Whereas in fig. 2(d), at the early stage of propagation, the proportions of influential spreaders of entertainment and other news are low and have a minor difference considering the re-postings within one day. This indicates that for the entertainment news, the influential spreaders may appear at the later stage of spreading. For earlier stages such as within one hour and within five hours, the proportions of influential spreaders are even lower, so we do not show them here.

The propagation is highly influenced by the proportion of the first layer since it is the majority. As a result, we
Propagating structure feature of entertainment news in the Weibo online social network

Fig. 4: The concentration ratio at different stages of propagation. (a) For the whole lifespan, the entertainment news has relatively larger concentration ratios. The p-value of the Mann-Whitney U test between entertainment and other news is below 0.01. Similar to the r, we also emphasize the early stage and plot the concentration ratio distribution including one hour from the first re-posting (b), five hours (c) and one day (d). The p-values are also below 0.01.

Apart from the creator, we further study the out-degree of all the nodes by the concentration ratio. It is originally used in economics: it measures the market share of the industry and illustrates the oligopolistic degree. Here we use the concentration ratio to measure the inequality degree of nodes property (out-degree) from the network perspective. Specifically, the concentration ratio calculates the sum of the squares of the out-degrees (see the section “Methods”). The concentrated network has a relatively larger concentration ratio. As shown in fig. 4(a), the peak of entertainment news is obviously lower than other news and the corresponding concentration ratio of the peak is relatively large. This indicates that entertainment news networks turn out to have smaller concentration ratios. Other kinds of news have few nodes which contribute a lot in the structural diffusion process, since they have broadcast-like diffusion processes [33,34]. Compared with other news, the entertainment news propagates in a viral way. Similar to the methods of the r, we check the stability of this phenomenon from the temporal perspective. And this finding remains at the early stage (fig. 4(b), (c) and (d)). The concentration ratio reflects the inequality from the network perspective, and when we look deeper into the node perspective, the Gini coefficient demonstrates a more obvious difference between types of news as follows.

The ratio between the indirect and direct re-postings and the concentration ratio we mentioned above is a kind of global parameter. For example, the r divides edges of a network into two groups: direct re-postings and indirect re-postings, which does not consider the microscopic difference of propagation connection. And the concentration ratio studies the concentrated degree from the network perspective. More in detail, from the node perspective, we analyze the Gini coefficient of a given network to measure...
the out-degree heterogeneity of each node. For example, the random regular network is one of the most homogeneous networks since every node has the same out-degree. The Gini coefficient of the random regular network is zero. However, the star network is the most heterogeneous network and the Gini coefficient of the star network is one. The realistic propagation networks are between these two limiting cases. Therefore, we use the Gini coefficient to measure the heterogeneity degree (see the section “Methods”). In fig. 5(a), the complementary Gini coefficient of entertainment news is considerably larger than that of other news. Entertainment news networks have a viral-like structural diffusion process since their propagation is more homogeneous and involves some influential broadcasters. On the contrary, other news demonstrates a broadcast-like layout due to one dominant creator acting as the unique influential disseminator. Again, we test whether this difference is stable at the early stages by analyzing the distribution of the complementary Gini coefficient within one hour (fig. 5(b)), within five hours (fig. 5(c)) and within one day (fig. 5(d)). And it also shows a difference, with a p-value below 0.01.

Discussion. – Recently, entertainment news has gained great popularity among different types of news media [12], which influences the entire society in all perspectives. Here, we study the propagation structure feature based on the complex network theory. Considering different driving factors including social contagion behaviors or strategic considerations [35], the topology of the network helps to explore the hidden structure of entertainment news. Compared with the popularity of entertainment, studies of its propagation mechanism did not draw enough attention. The underlying mechanism of entertainment news still has many unsolved problems currently, indicating the significance and necessity of study on propagation features. In this work, we focus on the re-posting ratio and the heterogeneity level of networks. The entertainment news has a peculiar distribution of these two features and these characteristics could appear at the early stage of spreading. During the entertainment news propagation, the masses are willing to post entertainment news to satisfy the psychological need of curiosity, conformity or narcissism [16], and the entertainment practitioners add fuel to the gossip fire deliberately. These enthusiastic audiences and entertainment reporters could constitute influential re-posters. As a result, the entertainment news spreads often by the contribution of both the creator and influential latter re-posters, which reflects its multiple-spreader structure compared to the single dominant spreader in other news. And the r of entertainment news is larger, demonstrating that entertainment news has a stronger infectivity as a result of cascading dynamics [36-38]. The fan economy is a kind of economic mode that obtains benefits from the relationship between fans and the people who are followed like stars, idols, and celebrities. The mechanism that entertainment news generates special structural features will be useful in the study of the fan economy, which could drive further research in the future. Besides, negative and unhealthy entertainment news can be intentionally and effectively supervised and controlled according to these structural features.

The entertainment industry is influenced by its local culture. For example, since the last century, the UK and US had tabloid culture, which served as a kind of newspaper to spread entertainment news. And the entertainment industry has become a pillar industry in the Northeast Asia countries like China, Japan and South Korea, where entertainment news plays an indispensable role in people’s life. Additionally, different social network platforms also have different operating mechanisms including users’ category, text language and length, multimedia attachment and so on. This diversity of culture and platform causes the difficulty of studying entertainment news only by user property analysis or natural language processing. However, with respect to the propagation structure feature, for example, we find that entertainment news propagates with more indirect re-postings and a polycentric structure in this work. And these structure properties could study deep into their “DNA” and break the shackle of different cultures and platforms. Moreover, the structural propagation findings of this work require verification from other platforms of other cultures. In our previous work about fake news, similar features are found for both Weibo and Twitter platforms from different cultures [29]. Due to data reason, this work only studies the dataset from Weibo, and the analysis of other platforms may drive further research in the future.

REFERENCES

[1] Schudson M., Discovering the News: A Social History of American Newspapers (Basic Books) 1981.
[2] Keenan K. L., Public Relat. Rev., 22 (1996) 215.
[3] Bird S. E., For Enquiring Minds: A Cultural Study of Supermarket Tabloids (University of Tennessee Press) 1992.
[4] Bai M., All the Truth Is Out: The Week Politics Went Tabloid (Alfred A. Knopf) 2014.
[5] Postman N., Amusing Ourselves to Death: Public Discourse in the Age of Show Business (Penguin) 2006.
[6] Ensink T., Communications-Sankt Augustin Then Berlin, 23 (1998) 383.
[7] Strogatz S. H., Nature, 410 (2001) 268.
[8] McQuail D., Mass Communication Theory: An Introduction (Sage Publications, Inc.) 1987.
[9] Craig D. and Cunningham S., Social Media Entertainment: The New Intersection of Hollywood and Silicon Valley (New York University Press) 2019.
[10] Twitter by the Numbers: Stats, Demographics & Fun Facts, https://www.omnicoreagency.com/twitter-statistics/.
[11] Ahn J., Oh S. and Kim H., Korean pop takes off! Social media strategy of Korean entertainment industry, in 2013 10th International Conference on Service Systems and Service Management (IEEE) 2013, pp. 774–777.
Weibo user development report of 2018, http://data.weibo.com/report/reportDetail?id=433.

Cunningham S., Media Int. Aust., 169 (2018) 5.

Rentfrow P. J., Goldberg L. R. and Zilca R., J. Personal., 79 (2011) 223.

Ong E. Y., Ang R. P., Ho J. C. et al., Personal. Individ. Differ., 50 (2011) 180.

Wang D., Comput. Hum. Behav., 66 (2017) 138.

Wu C. M., Key Eng. Mater., 474-476 (2011) 1573.

Wang J.-L., Jackson L. A., Wang H.-Z. et al., Personal. Individ. Differ., 80 (2015) 119.

Theocharis Y. and Quintelier E., New Media Soc., 18 (2016) 817.

Ali O., Flagnias I., De Be T. et al., Celebrity Watch: Browsing news content by exploiting social intelligence, in Joint European Conference on Machine Learning and Knowledge Discovery in Databases (Springer, Berlin, Heidelberg) 2011, pp. 613–616.

Clemons E. K., Decis. Support Syst., 48 (2009) 46.

Pastor-Satorras R. and Vespignani A., Phys. Rev. Lett., 86 (2001) 3200.

Liu M., Li D., Qin P. et al., PLoS ONE, 10 (2015) e0120701.

Li D., Qin P., Wang H. et al., EPL, 105 (2014) 68004.

Liu C., Li D., Fu B. et al., EPL, 107 (2014) 68003.