# Data Mining and Mathematical Modeling in Big Data: Partial-Rogue Lump Ripple Model on Sina Weibo Public Opinion

Weiwei Wu, Xiqin Ao and Wanjun Yao

*School of Big Data and Artificial Intelligence, AnHui XinHua University, 555 Wangjiang St., Hefei, Anhui, China*

Abstract:     There has been a tremendous rise in the growth of Chinese Microblog --- Sina Weibo in the past few years. Sina Weibo contains many interesting popular issues. The mining of these vast amounts of data and the corresponding mathematical models is an interesting question. In particular, the trend of some popular topics can be described by the soliton wave model based on some hypothesis and parameter settings. The evolution trend of the sub-events belonging to a super popular topic is stated by employing the partial-rogue lump ripple solution theory of the nonlocal Kadomtsev-Petviashvili equation.

## 1 INTRODUCTION

Sina Weibo is a major Chinese microblogging platform. It was founded in 2009 and has since become one of the most popular social media platforms in China, with millions of active users. Weibo allows users to share posts, images, videos, and more in real-time, similar to Twitter. However, the platform is more open and allows for longer posts, making it more suitable for discussions and in-depth content. Weibo hashtags are also very popular, often trending topics that users can follow or participate in discussions around. One unique feature of Weibo is its integration with traditional media. Many news outlets and celebrities have official accounts on Weibo, using it as a platform to share news, updates, and interact with their followers. This has made Weibo a powerful tool for both information dissemination and social commentary. In addition to its popularity in China, Weibo has also gained a significant user base among overseas Chinese and Chinese speakers around the world. The platform offers multiple language options, allowing users to switch between Simplified and Traditional Chinese, as well as English and other languages. Overall, Weibo is a dynamic and vibrant social media platform that offers a window into Chinese culture, society, and current affairs. It continues to evolve and adapt to the changing needs of its users, making it a must-follow platform for anyone interested in understanding the Chinese perspective.

The public opinion events are often reflected on Weibo. Weibo is the most important observation window for online public opinion. The study of public opinion on Weibo has attracted a lot of attentions[1-6]. It is worth mentioning that the evolution and the control for the popular issues on Sina Weibo public opinion is the main problem in the research field. There exists such a popular topic in online public opinion, which encompasses numerous sub-events, each with a high level of popularity. We can refer to this as a popular topic matrix. We have discovered this trend in the popular topic of the film "Full River Red". Then we use the partial-rogue lump ripple solution theory of the nonlocal Kadomtsev-Petviashvili equation to describe the system and state the appropriate complete strategic classifications.

This paper is organized as follows. The next section deals with the partial-rogue lump ripple waves in the development process of the popular spots on Sina Weibo public opinion. The nonlocal Kadomtsev-Petviashvili equation is introduced in section 3, which is 2+1 dimension model. The partial-rogue lump ripple solutions of the nonlocal Kadomtsev-Petviashvili equation are also presented in this section. Section 4 is devoted to conclusions and discussions.

## 2 DATA MINING

The film Full River Red is a suspense comedy directed by Zhang Yimou and starred by Shen Teng, Yiyang Qianxi, Zhang Yi, Lei Jiayin and Yue Yunpeng. Full River Red was released in Chinese Mainland on January 22, 2023. It was released in North America on March 17th. Full River Red tells the story of Qin Hui leading his troops to talk with the Jin State four years after Yue Fei died in Shaoxing of the Southern Song Dynasty. On the eve of the meeting, the envoy of the Jin Dynasty died at the prime minister's residence, and the secret letter he carried disappeared without a trace. The story of a soldier and the deputy commander of the personal battalion who coincidentally become entangled in this huge conspiracy. Prime Minister Qin Hui orders the two to search for the culprit and the truth. Man Jianghong's box office revenue reached 4.544 billion yuan. Full River Red won the 36th China Film Golden Rooster Award Award, the 18th China Changchun Film Festival, the 30th University Student Film Festival and other honorary awards. This is a film that has achieved great success, but also has a lot of controversy. After its official release, Full River Red quickly became a popular topic in online public opinion, with a lot of controversial discussions related to it. Since its release on January 22, 2023, with the box office of the film Full River Red rising all the way, there have also been public opinion disputes about it, including the landmark events, namely, box office fraud, plagiarism, prosecuting the social media influencers, publicizing errors, cultural and creative peripheral products disputes and so on. We will introduce the landmark events of the Full River Red, which is a popular topic of online public opinion.

Event 1. Box office fraud (BOF). On the evening of January 24, 2024, a netizen reported that his movie ticket for The Wandering Earth II was returned by the cinema on the grounds of "equipment failure", but the netizen found that Full River Red was shown in the same screening hall at the same time of the day. After the incident broke out, some netizens questioned that Full River Red might have bad competition behaviors such as "stealing box office".

Event 2. Plagiarism (PLA). During the film's popular release, some netizens questioned that the film Full River Red copied the TV series Longmen Escort Agency released in 2013. In addition to this plagiarism controversy, there is also the issue of plagiarism in the ending theme song. Han Hong, a famous singer, is suspected of plagiarizing the song "Look" at the end of the movie Full River Red.

Event 3. Prosecuting the social media influencers (PSI). In the evening of January 28, 2024, the official blog of Full River Red posted five blog posts in succession, intended to respond to and clarify the negative accusations of "box office fraud", "ghost field", "plagiarism", "black screen" and other negative accusations that continued to surround it. Subsequently, the film company issued another article announcing the lawsuit against the relevant rumormongers. From the attached screenspopular, it can be seen that the defendants are four internet influencers.

Event 4. Publicizing errors (PCE). Because in the prosecution, the official Weibo of the film Full River Red used the "false" argument from Qin Hui to refute the negative accusations, which attracted netizens' ridicule, and other mistakes in its publicity were also fermented. Some netizens have found that the official blog post previously released, which stated "Do not forget the shame of Jingkang" and "loyalty and righteousness are yet to be exhausted," mistakenly disconnected the fixed term "Jingkang". We found that the Full River Red poster produced by Wanda Movie also had typos, in which simple and complex characters were mixed, and the "hair with rage" was written as rich, not "hair". In addition, when the official Weibo celebrates the box office breaking through 2.5 billion yuan, the use of copywriting such as "Step through the Mountains and Rivers, Celebrate the New Year" and the use of Song Huizong's thin gold font to write the names of the main actors are quite inappropriate.

Event 5. Cultural and creative peripheral products disputes (CPD). On February 13, 2023, the film Full River Red cultural and creative surrounding crowdfunding project was launched. Soon, netizens around the shelves complained about "hell", like "around Qin Hui". For example, two "Prime Minister's Tongling" incense expanding stones based on the "Qin Hui Token" were named "Yaoqin" and "Qingmei", which were the names of the two female singers who died tragically in Qin Hui's hands. In addition, a creative hourglass that turns cherry red easily reminds people of the "Storm Pavilion" where Yue Fei was wrongly killed. In addition, unlocking Qin Hui's Q bookmarks in Fulili has become a point of roast. On February 14, with the relevant queries on the popular search, the film Full River Red officially released a statement of off shelf: "The crowdfunding activities of derivatives around the film Full River Red have a deviation from some netizens' understanding of the original

intention of product design and creativity, which has had a bad impact on online public opinion. We apologize for this and decide to off shelf the crowdfunding activities.

We collected the newly added messages per 2 days on Sina Weibo from January 22, 2023 to February 20, 2023. The followed table shows the data of the three cases.

Table 1 (a): The data of the new messages per 2 days (Unit: Thousand).

|         | 24 | 26 | 28  | 02  | 04   | 06   | 08   |
|---------|----|----|-----|-----|------|------|------|
| Event 1 | 9  | 35 | 130 | 326 | 371  | 763  | 962  |
| Event 2 | 5  | 27 | 145 | 289 | 332  | 651  | 828  |
| Event 3 | 0  | 0  | 22  | 87  | 259  | 354  | 787  |
| Event 4 | 0  | 0  | 8   | 34  | 121  | 208  | 369  |
| Event 5 | 0  | 0  | 0   | 9   | 25   | 89   | 126  |
| Whole   | 14 | 62 | 305 | 745 | 1108 | 2065 | 3072 |

Table 1 (b): The data of the new messages per 2 days (Unit: Thousand).

|         | 10   | 12   | 14   | 16   | 18  | 20  |
|---------|------|------|------|------|-----|-----|
| Event 1 | 868  | 687  | 471  | 352  | 324 | 219 |
| Event 2 | 734  | 814  | 627  | 583  | 347 | 201 |
| Event 3 | 587  | 231  | 152  | 149  | 67  | 25  |
| Event 4 | 501  | 324  | 217  | 173  | 88  | 63  |
| Event 5 | 342  | 405  | 268  | 231  | 97  | 65  |
| Whole   | 3032 | 2461 | 1735 | 1488 | 923 | 573 |

Denote time by t. The following figure is used to describe the variation of the six indexes as Whole (general description of the film Full River Red) and Event 1-5.
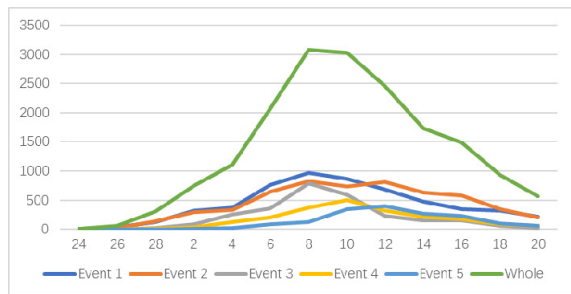


Figure 1: The trend of six indexes.

The Figure 1 illustrates that Full River Red corresponds to multiple sub-topics within this popular topic. However, this figure only reflects their static data and does not reveal the underlying evolutionary mechanisms. To further understand this process, we will introduce nonlocal Kadomtsev–Petviashvili equation to simulate it. Partial-rogue

lump ripple theory, one topic of the soliton theory, is one of the most important areas fields of mathematical physics. Today, soliton theory has multiple applications in plasma physics, matrix model of string theory, non-linear optics and light wave communication technology.

# 3 PARTIAL-ROGUE LUMP RIPPLE MODEL

## 3.1 Assumptions

(1) The selected popular topics diffuse in a plane, i.e. the space of the evolution of the popular spot on Weibo contains two variables x and y;
(2) The new daily messages about popular topics on Weibo will be zero when $t \to +\infty$. This assumption is reasonable and very important, and the reason will be further explained in the end of this section.

## 3.2 Nonlocal Kadomtsev-Petviashvili Equation Model

The nonlocal Kadomtsev-Petviashvili equation is introduced in the form as

$$\begin{cases} A_{xt} + \dfrac{\alpha}{4}(A_x + B_x)(3A_x - B_x) + \dfrac{\alpha}{4}(A - B)B_{xx} \\ + \dfrac{\alpha}{4}(3A + B)A_{xx} + \beta A_{xxxx} + \gamma A_{yy} = 0, \\ B = \hat{P}_S^x \hat{P}_S^y \hat{T}_d A = A(-x, -y, -t). \end{cases} \quad (1)$$

Note that $A = A(x, y, t)$, $B = B(x, y, t)$, $\hat{P}_S^x \hat{P}_S^y \hat{T}_d$ denotes parity with a shift of the space variable x and y reversal with a delay, and $\alpha, \beta, \gamma$ are coefficient. The function $A = A(x, y, t)$ is the amplitude of the wave of the popular topics as a function of three real variables, space x, y and time t.

## 3.3 Solution of Nonlocal Kadomtsev-Petviashvili Equation

Following Hirota's bilinear transform method for the integrable equations, we can employ the variable transformation as

$$\begin{aligned} A &= \frac{12\beta}{\alpha}(\ln f)_{xx} + b_1(\ln f)_{xxt} + b_2(\ln f)_{xxy}, \\ B &= \frac{12\beta}{\alpha}(\ln f)_{xx} - b_1(\ln f)_{xxt} - b_2(\ln f)_{xxy} \end{aligned} \quad (2)$$

to transform the nonlocal Kadomtsev-Petviashvili equation into Hirota's bilinear form as

$$f_{xt}f - f_x f_t + \beta\left(f_{xxxx}f - 4f_{xxx}f_x + 3f_{xx}^2\right) + \gamma\left(f_{yy}f - f_y^2\right) = 0,$$

where $b_1$ and $b_2$ are arbitrary constants and $f = f(x, y, t)$, satisfying the following condition as

$$f(x, y, t) = \hat{P}_S^x \hat{P}_c^y \hat{T}_d f(x, y, t) = f(-x, -y, -t).$$

Then, the following Wronskian determinant are first introduced

$$f_N = W\left(\phi_1, \phi_2, \cdots, \phi_N\right) = \begin{vmatrix} \phi_1 & \phi_1^{(1)} & \cdots & \phi_1^{(N-1)} \\ \phi_2 & \phi_2^{(1)} & \cdots & \phi_2^{(N-1)} \\ \cdots & \cdots & \cdots & \cdots \\ \phi_N & \phi_N^{(1)} & \cdots & \phi_N^{(N-1)} \end{vmatrix}, \phi_j^{(k)} = \frac{\partial^k \phi_j}{\partial x^k}.$$

More specifically, the nonlocal Kadomtsev-Petviashvili equation admits the following solutions

$$f = f_N = |a_{jk}|_{1 \le j,k \le N}, a_{jk}$$
$$= \delta_{jk}$$
$$+ \int_x \psi_j^+(x', y, t)\psi_k^-(x', y, t)dx'$$

with a range functions $\psi_j^+ = \psi_j^+(x, y, t)$ and $\psi_k^- = \psi_k^-(x, y, t)$ satisfying

$$i\psi_{k,y}^- - \sqrt{-\frac{3\beta}{\gamma}}\psi_{k,xx}^- = 0, \psi_{k,t}^- + 4\beta\psi_{k,xxx}^- = 0, 1 \le$$
$$k \le N, \gamma < 0, \quad i\psi_{j,y}^+ + \sqrt{-\frac{3\beta}{\gamma}}\psi_{j,xx}^+ = 0, \psi_{j,t}^+ +$$
$$4\beta\psi_{j,xxx}^+ = 0, 1 \le j \le N, i^2 = -1,$$

where $\delta_{jk}$ is Kronecker function.

Taking $N = 1$, we get the partial-rogue lump ripple solution as

$$f = \delta_{11} + \int_x \left[\frac{\partial^2}{\partial x_1^2}\psi_1^+ \cdot \frac{\partial^2}{\partial x_1^2}\psi_1^-\right] dx'$$
$$= \delta_{11}$$
$$+ \sigma_1^2 \frac{\sqrt[3]{12\beta}}{48\beta\gamma t^{\frac{7}{3}}}\{[Ai'(H_1)]^2$$
$$- H_1 Ai^2(H_1)\}$$
$$+ \frac{\rho_1^2}{18^{\frac{2}{3}}\beta^{\frac{1}{3}}t}\{[H_1 Ai(H_1)]^2 - H_1[Ai'(H_1)]^2$$
$$- 2Ai(H_1)Ai'(H_1)\},$$

where $H_1$ is given as

$$H_j = \frac{18^{\frac{1}{3}}\left[4\gamma\left(x - x_j\right)t + \left(y - y_j\right)^2\right]}{24\gamma\beta^{\frac{1}{3}}t^{\frac{4}{3}}}.$$

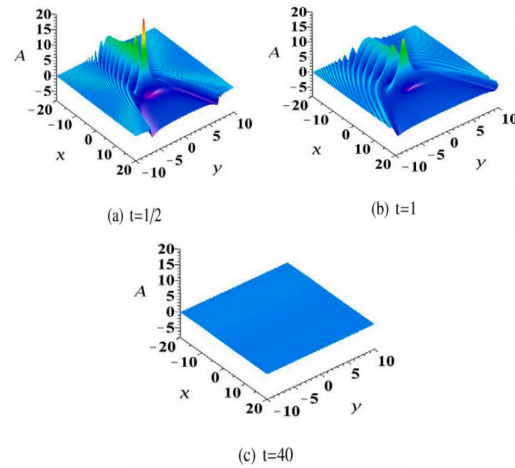The partial-rogue lump ripple solution is composed of a range of lumps as shown in Fig. 2.



(a) t=1/2

(b) t=1

(c) t=40

Figure 2: The time evolution at t=1/2,1,40.

As can be seen from the above figure, the evolution trend of the partial-rogue lump ripple solution to the nonlocal Kadomtsev-Petviashvili equation is closely aligned with the popular topic of the film Full River Red. Based on soliton theory, we provide a mathematical explanation for the evolution trend of Full River Red topic.

## 3.4 Application on Weibo

In this subsection, we turn back to the popular topics on Weibo. Based on the the partial-rogue lump ripple solution to the nonlocal Kadomtsev-Petviashvili equation, the following conclusions are worth to point out.

(1) The wave with the highest amplitude in Figure 2(a) represents the overall popularity of the super hit movie Full River Red, and the waves with slightly lower amplitudes correspond to the sub-hot topics. It can be seen that this partial-rogue lump ripple solution aptly simulates the status of this popular event.

(2) Figure 2(b) shows that the amplitude of the smaller waves decreases but becomes more stable. By comparison, it reflects the state of the popular event after a period of development. The reason for this is that after a certain period of time, netizens' discussions mainly focus on the core sub-hot topics, such as Event 1 to Event 5.

(3) The partial-rogue lump ripple solution to the nonlocal Kadomtsev-Petviashvili equation tends to zero when $t \to +\infty$, it is consistent with the assumption as the same as the evolution of the popular topics on Weibo. As we know, A is the variable describing the new messages per 2 days. Indeed, the new messages about the popular issue on

Weibo will be almost zero after a long time. This process is shown in Figure 2(c).

(4) The partial-rogue lump ripple solution of nonlocal Kadomtsev-Petviashvili equationn shares multiple solitons, i.e. several events could coexist in the system at the same time. The popular online public opinion topic of Full River Red reflects this point.

(5) This equation also has a solution as shown in following figures.
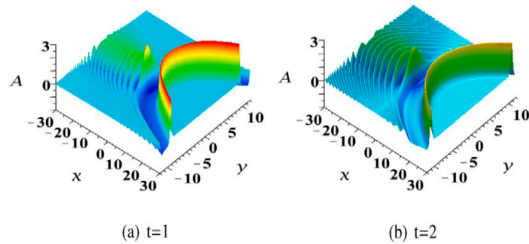


(a) t=1    (b) t=2

Figure 3: The time evolution at t=1,2.

When applied to the popular online public opinion topic, it indicates that the overall event is relatively strong and dominates the absolute popularity, while the sub-events are weaker. The image corresponding to the following solution is even more prominent.
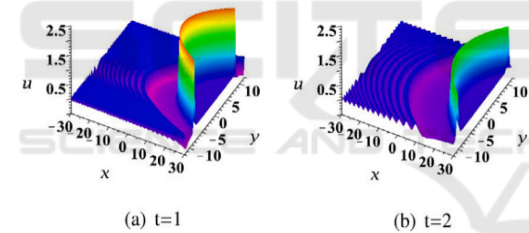


(a) t=1    (b) t=2

Figure 4: The time evolution at t=1,2.

(6)From the soliton theory, the partial-rogue lump ripple wave is once formed, and then it will continue for a period. This result tells us that the popular topics will keep be popular for several days normally.

## 4 CONCLUSIONS

The partial-rogue lump ripple solution theory of the nonlocal Kadomtsev-Petviashvili equation we introduced provides a good description and explanation of the popular topic of the film "Full River Red" in online public opinion. This offers a new perspective for understanding popular topics in online public opinion.

## REFERENCES

Yin, F., She, Y., Yang, Q., et al, 2024. Modeling and analyzing dynamic information propagation on sina weibo in a semi-directed network. *Journal of biological systems*, vol.32, no.01, pp.219-237, DOI: 10.1142/S0218339024500086.

Ma, D., Zhang, C., Zhao, L., et al, 2023. An analysis of the evolution of public sentiment and spatio-temporal dynamics regarding building collapse accidents based on Sina Weibo data. *International journal of geo-information*, vol.12, no.10, Article Number. 388, DOI: 10.3390/ijgi12100388.

Chen, X., Chen, Y., Yin, G., et al, 2023. Dynamic evolution of information diffusion networks of news agencies in emergencies: a case study of microblogs of urban fire disasters on Sina Weibo. *Multimedia tools and applications*,vol.83,no.9, pp. 25287-25319, DOI: 10.1007/s11042-023-16498-0.

Cui, H., Kertesz, J., 2023. Competition for popularity and interventions on a Chinese microblogging site. *Plos one*, vol.18, no.5, Article Number. e0286093, DOI: 10.1371/journal.pone.0286093.

Chen, Q., Sun, X., Wang, J., et al, 2023. User-based hierarchical network of Sina Weibo emotion analysis. *ACM transactions on Asian and low resource language information processing,* vol.22, no.5, Article Number.143. DOI: 10.1145/3579048.

Wu, W., Zhang, M., 2023. Exploring the motivations and obstacles of the public's garbage classification participation: evidence from Sina Weibo. *Journal of material cycles and waste management*, vol. 25, no.4, pp. 2049-2062, DOI: 10.1007/s10163-023-01659-y