A Combined Time Series Model for the Prediction of Social Network Popularity and Content Evolution

Zhiyuan Zhang, Zemin Bao

School of Electronic and Information Engineering, Key Laboratory of Communication and Information Systems, Beijing Municipal Commission of Education, Beijing Jiaotong University, China
zhangzhiyuan@bjtu.edu.cn, baozemin@bjtu.edu.cn

Abstract

The randomness, dynamism and uncertainty of social networks pose a challenge for revealing the mechanism of content popularity growth. Evolution of content popularity is characterized by strong heterogeneity. It is difficult for a single time series prediction model to capture all kinds of content popularity dynamic evolution patterns at the same time. Using single model to predict complex social network content popularity will lead to poor prediction ability and limited application scenarios. This paper attempts to establish a combined predicting model that integrates the predicting capabilities of multiple traditional time series models. By applying multi-class regression and analyzing the historical prediction performance of each sub-model, the combined weights of the predicted values of each sub-model are generated. The model can learn to adjust the combination weights according to the real-time prediction performance of each sub-model, so as to adapt to the dynamic changes of the evolution model and to enhance the performance of predicting content popularity. The evaluation results of real social network datasets show that the performance of the proposed model is better than that of the existing single models.

Keywords: Social network popularity, Time series, Multi-class regression

1 Introduction

In a social network, users establish a user relationship network by following each other. They share and spread information, express emotions and exchange opinions through this network. The information transmitted in social network is generated by users, including text, pictures, video and other content modes, which are collectively referred as user-generated content(UGC). Compared with traditional media, social network has stronger information propagation ability. The scale-free and small-world characteristics of network users' topology make information propagate faster and have wider influence. With its convenient way of information publishing, unique incentive mechanism of sharing and reposting, as well as the interaction of mobile internet, social network attracts more and more users to participate in the process of proposing information. Then comes the massive UGC.

Popularity of social network is the popularity of UGC, which describes the scale and depth of macro-diffusion of social network content. It is usually measured by the total number of interactions between users and UGC (such as the total number of video clicks, microblog forwarding times during a period of time.). It is the embodiment of user clustering behavior. The target of this kind of research is to predict the content popularity through analyzing useful UGC. According to Herbert Simon [1], Excess information leads to scarcity of attention, that is, the attention of users is limited, they only to pay attention to some of the interesting or popular content. The competition of different content for limited users' attention and the topological structure of social networks will lead to great heterogeneity in the popularity of UGC. A few content gets a lot of attention, while most content is obscure. This also leads to different characteristics and modes of popularity propagation of different types of content. For example, news contents have short timeliness characteristics, contents popularity gathers for a period of time, and users always focuses on hot news in a shot period of time; while music contents has long timeliness characteristics, because users' preferences are different, for different users, the types of songs with high popularity and popularity propagation characteristics are different.

Prediction of content popularity in social networks includes UGC content modality prediction [2-5], time series prediction [6-10] and domain prediction, etc. In this paper, we focus on the time series prediction of social network. The goal of the mission is to predict future popularity based on the early history of information dissemination after content release. Time Series Prediction States that "Because Message X gets more attention than other messages today, Message X

*Corresponding Author: Zemin Bao; E-mail: baozemin@bjtu.edu.cn
DOI: 10.3966/160792642020072104029

will be more popular in the future". The prediction based on time series after the content begins to spread is better than that based on the features obtained in advance. Hofman et al. [11-12] also proves this point. Martin of the University of Michigan and Benjamin Shulman of Cornell University [12-13] all show that time domain characteristics play a leading role in predicting content popularity. Both platform managers and third-party research institutes can easily crawl and obtain such time domain information. Scholars use machine learning algorithm to establish mapping relationship between content popularity correlation features and future popularity. According to the continuity of prediction objectives, prediction problems can be formalized as regression or classification problems respectively. The popularity prediction in classification problem [14-21] refers to whether the predicted information can obtain a wide range of popularity, whether the popularity of the information reaches the corresponding threshold, or whether the predicted information ultimately belongs to the range of popularity. Regression problem [22-24] predicts the future popularity value of target information, that is, the specific number of reposts or clicks the content will get at a certain time in the future. Application scenarios are slightly different, so the previous method is not universal still has limitations.

Single model can not deal with the above complex problems and locate more application scenarios. Therefore, on the basis of studying the advantages of different time series propagation models, this paper proposes a combination time series prediction model, which can automatically adjust the income function of each model and automatically learn the weights in different application scenarios, so that the model has higher popularity prediction performance and more extensive scenario applicability.

This paper is organized as follows. In Section 2, we present the proposed COMFITS model motivated by some traditional models. Some useful evaluations are analyzed in Section 3. Finally, in Section 4 we present some concluding remarks.

2 Model

2.1 Problem Definition

Assume that the popularity of the target UGC i is sampled at intervals after publication, we define the time series of i's popularity as $\phi_i = \{v_1^i, v_2^i, \ldots, v_n^i\}$. In which v_k^i is the popularity of content i in the kth sampling interval, the total number of sampling period is n. It should be noted that the sampling interval can be hourly, daily, weekly and so on, depending on the requirements of prediction task or platform propagation characteristics. The goal of popularity prediction is to predict the popularity v_{n+1}^i of content i in the next time interval, that is, we need to predict the next moment's popularity based on the current content popularity time series data. Then the problem can be defined as:

$$v_{n+1}^i = \mathcal{M}(v_1^i, v_2^i, \ldots, v_n^i) \quad \textbf{(1)}$$

In this paper, we try to find a proper model M and learn to predict content i's popularity after analyzing existing time series popularity data.

2.2 Traditional Models and Motivations

ARIMA(Auto Regressive Integrated Moving Average): ARIMA is a common time series prediction model based on mathematical statistics, which can be directly used to predict social network popularity 24]. It consists of Auto Regressive (AR) model and Moving Average (MA) model. The autoregressive model takes itself as a regression variable. Let B^k be a lag operator, $B^k v_n = v_{n-k}$, then the p-order autoregressive model can be expressed as:

$$v_{n+1} = \alpha_1 B^1 v_{n+1} + \alpha_2 B^2 v_{n+1} + \ldots + \alpha_p B^P v_{n+1} + e_{n+1} \quad \textbf{(2)}$$

where, e_{n+1} is the white noise sequence, which follow the normal distribution with mean 0 and variance σ^2. $\alpha_1, \alpha_2, \ldots, \alpha_p$ are the autoregressive coefficient. v_{n+1} can be regarded as a linear combination of self-secent p-order delay term and white noise. MA represents v_{n+1} as a linear combination of current random errors and q-order random errors. The definition details are as follows:

$$v_{n+1} = e_{n+1} - \beta_1 e_n - \beta_2 e_{n-1} - \ldots - \beta_q e_{n-q+1} \quad \textbf{(3)}$$

$\beta_1, \beta_2, \ldots, \beta_q$ are the moving average coefficients, e_{n+1} is the white noise sequence. Considering both the autoregressive part and the moving average part, the autoregressive moving average model (ARMA) can be obtained as:

$$\begin{aligned} v_{n+1} &= \alpha_1 B^1 v_{n+1} + \ldots + \alpha_p B^P v_{n+1} \\ &+ e_{n+1} - \beta_1 e_n - \ldots - \beta_q e_{n-q+1} \end{aligned} \quad \textbf{(4)}$$

where p and Q are autoregressive order and moving average order respectively, then ARMA model can be represented as ARMA(p, q). Considering that the sequence ARMA model is constrained by the sequence stationarity, stationary processing should be utilized to handle this non-stationary time series. Therefore, through d-order difference operation for non-stationary time series, ARMA(p, q) model can be updated as ARIMA(p, d, q). As a widely used time series prediction method, ARIMA has achieved good results in predicting time series with significant trend (such as continuous rise) and periodicity (such as seasonal fluctuation). Under the social network content popularity prediction scenario, the time-domain information of information diffusion and dissemination is easy to obtain, such as the time series of the change of information popularity with time, which is a universal form of time-domain information expression for various social network platforms. Therefore, ARIMA can be partly utilized to complete the task of social network popularity prediction. In order to simplify the calculation, we set d=1 and $w_{n+1} = v_{n+1} - v_n$, then the ARIMA model we need to use is defined as:

$$w_{n+1} = \alpha_1 w_n \ldots + \alpha_p v_{n-p+1} + e_{n+1} - \beta_1 e_n - \ldots - \beta_q e_{n-q+1} \quad \textbf{(5)}$$

Please check [24] for more details about ARIMA.

SVR(Support Vector Regression) [25]: SVR is a generalization of SVM in solving regression problem an implementation of Structural Risk Minimization (SRM). SVR maps input data to high-dimensional feature space by kernel function, and establishes the relationship between input data and output data in high-dimensional feature space. From this point of view, we use the SVR model to predict the trend of social popularity as follows：

$$v_{n+1} = \sum_{l=1}^{L} \alpha_l \cdot K(X_{v,n}, X_{l,n}) + b \quad \textbf{(6)}$$

where $X_{v,n}$ is the eigenvectors of i obtained at time n. and can be represented by generating epidemic time series $\phi_i = \{v_1^i, v_2^i, \ldots, v_n^i\}$. $\{X_{l,n}\}_{l=1}^{L}$ is the a set of Support Vector returned by training, in which $\{\alpha_l\}_{l=1}^{L}$ and b are parameters. $K(x, y)$ is the kernel function, which describes hypothesis space structures and characteristics. The commonly used kernel functions include polynomial kernel function, Gaussian Radial Basis Function and Sigmoid kernel function. The approximation properties of Gaussian Radial Basis Function can not only realize the non-linear mapping of input data to high-dimensional feature space, but also be suitable for dealing with non-linear problems and easy to implement. Combining the goals of our work and the evolving characteristics of social network popularity. In this paper, Gauss Radial Basis Function is chosen as the kernel function of SVR, which is

$$K(x, y) = exp(-\| x - y \|^2 / 2\sigma^2) \quad \textbf{(7)}$$

where σ width coefficient of kernel function. As a widely used time series prediction method, SVR has obvious advantages in capturing and characterizing the non-linear relationship in time series. That's motives us that SVR can be assumed as a part of the mix model in order to capture a part of non-linear relations predicting social network purpality prediction. For more details about SVR, please see the reference [25].

M-L (Multivariate Linear Regression): M-L model is an extension of S-H model, S-H is a linear regression model based on logarithmic popularity, which can be used to capture the point process information (some "event granularity" information) through establishing regression relationship of logarithmic prevalence, for more details, please see reference [26]. Compared with S-H, W-L model increased dimension of input eigenvector and further analysis of information from the perspective of time series. Specifically, feature vectors are extracted from the time series of item i popularity can be set as:

$$v_n^i = (v_1, v_2, \ldots, v_n) \quad \textbf{(8)}$$

Then the predicted popularity of item i in the next time interval is obtained by the following formula:

$$\hat{v}_{n+1} = \Theta_n \cdot v_n^i \quad \textbf{(9)}$$

where $\Theta_n = (\theta_1, \theta_2, \ldots, \theta_n)$ are the parameter vectors of the model, they will be trained in the training set c, and the objective function is:

$$argmin \frac{1}{|c|} \sum_{k \in c} (\frac{\Theta_n \cdot v_n^i}{v_{n+1}} - 1)^2 \quad \textbf{(10)}$$

2.3 The Proposed Combined Model for Popularity Prediction

We assume that all the models mentioned above can be used to predict future popularity and combined into a mix model. We hope each model plays its own role in the mixed model to ensure its best performance in the popularity prediction and to achieve twice the result with half the effort. For example, the ability of ARIMA model to smooth non-stationary time series. Structural risk minimization criteria for SVR, etc. Therefore, the key point of this part is to learn the combined weights of different prediction models. By evaluating the historical prediction performance of the mentioned models and giving them different combination weights, the models with higher prediction accuracy can obtain higher combination weights, that is to say, they play a more important role in the combined model. Inspired by Mixture of Experts, in this section, we call the combined model as Combined Forecasting Model Based on Time Series, COMFITS. The basic model of COMFITS system is shown in Figure 1. The design of each sub-model (ARIMA, SVR, M-L) has been given in the previous part.

We set up performance evaluation matrix $\boldsymbol{P}$ (which should be learned in the future computing) for each model

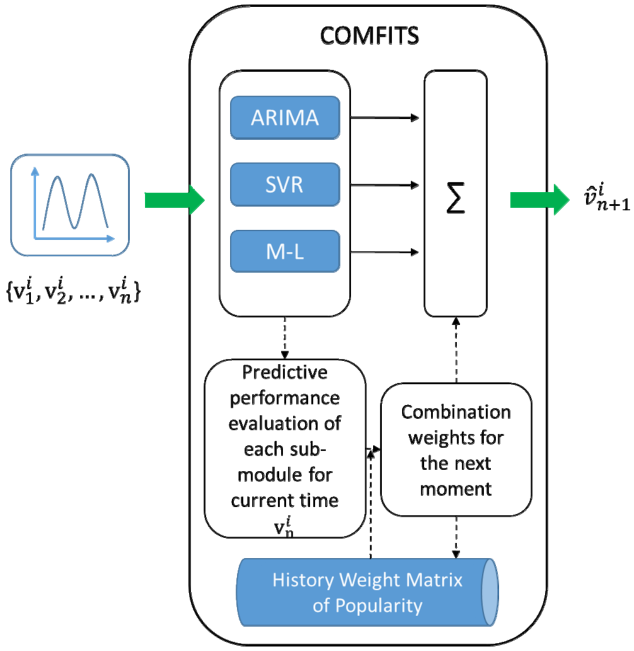

Figure 1. Architecture of our proposed model: COMFITS

$$\boldsymbol{P} = \begin{bmatrix} PE_{1,1} & PE_{1,2} & \cdots & PE_{1,m} \\ PE_{2,1} & PE_{2,2} & \cdots & PE_{2,m} \\ PE_{3,1} & PE_{3,2} & \cdots & PE_{3,m} \end{bmatrix} \tag{11}$$

where $PE_{i,j}$ is accuracy of model i for predicting popularity within time interval j, which will be defined as:

$$PE_{i,j} = (\hat{v}_{i,j} - v_{i,j})^2 \tag{12}$$

where $PE_{i,j}$ is the predictive value of model i for popularity within time interval j. $v_{i,j}$ is the true value of the popularity of the model i within the time interval j. Obviously, the smaller the value of $PE_{i,j}$, the better the performance of the model i. Considering the different popularity bases in different time intervals, it is not convenient to analyze the performance of each model in different time intervals. Here, the prediction accuracy PE is normalized to get the model weight ω:

$$\omega_{i,j} = \frac{exp(-PE_{i,j})}{\sum_{i=1}^{3} exp(-PE_{i,j})} \tag{13}$$

As can be seen from the above formula, The smaller $PE_{i,j}$ is, the smaller the prediction error of model i is and the higher the accuracy of model i is, the larger the model weight ω_i is. Meanwhile, $0 < \omega_{i,j} < 1$ can be understood as the probability of model i as the optimal prediction model when predicting the popularity value in the interval j. Then, historic weight matrix $\boldsymbol{W}$ for prediction of popularity are set as:

$$\boldsymbol{W} = \begin{bmatrix} \omega_{1,1} & \omega_{1,2} & \cdots & \omega_{1,m} \\ \omega_{2,1} & \omega_{2,2} & \cdots & \omega_{2,m} \\ \omega_{3,1} & \omega_{3,2} & \cdots & \omega_{3,m} \end{bmatrix} \tag{14}$$

In the formula, m denotes the number of nodes that need to be considered for the historical prediction performance of the model when inferring future combinatorial weights. Figure 2 shows how the historic weight matrix $\boldsymbol{W}$ of popularity prediction varies with prediction time. W_n represents the historical weight matrix for predicting popularity within time interval n+1. The corresponding training set size l is also shown in the Figure 2.

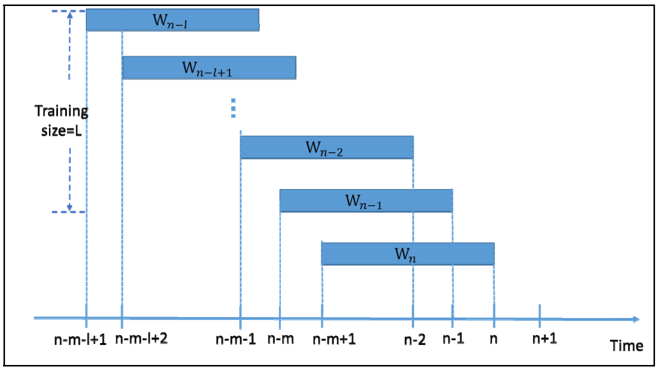

Figure 2. Temporal coverage of weight matrix W

In order to obtain the best prediction in time interval n+1, the weight vector $\hat{\boldsymbol{w}}^T = [\omega_1, \omega_2, \omega_3]$ for time interval n+1 needs to be generated. Then the problem is transformed into a multi-class regression problem, and the support vector machine algorithm is applied to establish the relationship between the weight vector $\hat{\boldsymbol{w}}^T$ and the historical weight matrix W_n. Relevant parameters of support vector machine are learned in the training set, and the size of the training set is l, then the training set is $\{\boldsymbol{W}_{n-l}, \boldsymbol{W}_{n-l+1}, \dots, \boldsymbol{W}_{n-2}, \boldsymbol{W}_{n-1}\}$, The training set covers the historical prediction weights of each model in the time interval. In the training process, the input of support vector machine is $\boldsymbol{W}_u$ and the output is the optimal weight combination vector $\hat{\boldsymbol{w}}_{u+1}^T$ at the next moment. Eventually, using the trained support vector machine, the predicted value of $\hat{\boldsymbol{w}}_{u+1}^T$ can be obtained by inputting $\boldsymbol{W}_u$.

Thus, the predicted popularity value $\hat{v}_{n+1}$ within the time interval n+1 can be obtained by the following formula:

$$\hat{v}_{n+1} = \frac{\sum_{i=1}^{3} \omega_i \times \hat{v}_{i,n+1}}{\sum_{i=1}^{3} \omega_i} \tag{15}$$

3 Simualtion

In order to verify the performance of COMFITS model, four different social network data sets are selected in this study, which cover music, movies, videos, books and other content modes. Different scenarios can verify the universality of the model and avoid the problem that the model is limited to a particular scenario.

3.1 The Proposed Combined Model for Popularity Prediction

We use the following data sets to evaluate the model.
Last.fm: This data set can be publicly accessed by the website. Last.fm, as the largest social music platform in the world, provides personalized recommendation through analyzing users' music listening history. Users can search, play and comment on their favorite music. The data set records listening events from January 2013 to August 2014. In order to facilitate the calibration of the model and avoid very uneven user distribution, in this section, the number of listening events for a single track is limited to 20000. For each listening event, the data set includes user ID, listening time, track ID and album artist. In this study, the number of listening events (listening music) was used as a measure of popularity.
Movie Lens-20M: This data set is publicly available on the website. MovieLens, as a movie rating and recommendation system, can recommend movies to users by using collaborative filtering technology and user's ratings. The data set contains rating events from January 1995 to March 2015. Each movie rating event includes user, movie, rating and timestamp information. In this study, the number of ratings of movies obtained from users is used as a measure of their popularity. Time series of popularity can be obtained for each movie.
Flickr: As one of the most popular photo sharing websites in the world, registered users can establish friendship with each other. Users can collect pictures they are interested in and share them with other users at the same time. The data set selected in this study contains 104 days of photo collection events between 2006 and 2007 (November 2-December 3, 2006, February 3-May 18, 2007) . Each collection event data set includes a collection timestamp, which can then generate a time series of picture collections, i.e. picture popularity time series.
Amazon-book: This data set is publicly available on the website. As the world's largest online retailer, Amazon started with online bookstores. This study focused on the popularity of books on Amazon's website, and selected the number of reviews a book received as the measure of its popularity, that is, a book with more reviews enjoy a higher online popularity. The data set contains reviewing events from May 1996 to July 2017. For each reviewing event, the data set includes reviewer ID, ASIN number of commodity (i.e. book), rating, rating timestamp and other relevant information such as the content of the comment.

3.2 Evaluation Indicators and Contrast Models

In this section, the mean absolute percentage error (MAPE) is used as an evaluation index for predicting the popularity of a message.

$$\varepsilon_v = \frac{1}{K}\sum_{k=1}^{K}\left|\frac{\hat{v}_k - v_k}{v_k}\right| \tag{16}$$

where $\hat{v}_k$ and v_k represent the predicted value sequence and the real value sequence of the test data part of a given message, respectively. K is the length of the test data. In addition, ARIMA, SVR and M-L, which are described in the previous section, are used as benchmark comparison models.

4 Results and Analysis

4.1 Evaluation of Effectiveness

It is necessary to verify whether COMFITS model can provide unbiased estimates for the prevalence series of different social platforms. Taking Last.fm data set as an example, 100 tracks are selected randomly, and the real value sequence and predicted value sequence of the test data part are obtained for each track. In order to facilitate comparative analysis, the sequence was normalized with reference to Eq. 13.

As shown in Figure 3, red dots represent the predicted results of COMFITS model, horizontal axis corresponds to the predicted values of the time series, and vertical axis corresponds to the true values of the time series. When the model provides unbiased estimates, the predicted values and the true values under ideal conditions are the same, that is, the red dots fall on the blue diagonal line in the graph. The results of data sets MovieLens-20M, Flickr and Amazon-book are also shown in Figure 3. It can be seen that the prediction results of COMFITS are scattered around the blue diagonal, that is, COMFITS can improve unbiased prediction values in different social network platforms.

Figure 4 and Figure 5 further illustrate the effectiveness of COMFITS in predicting popularity evolution. We select representative content popularity sequences from Amazon-book and MovieLens-20M datasets to show the performance of COMFITS in predicting the dynamic evolution process of popularity. Take Rain Man (Rain Man) as an example. The film was released in 1989. The first rating on MovieLens appeared on July 4, 1998. Half a year was chosen as

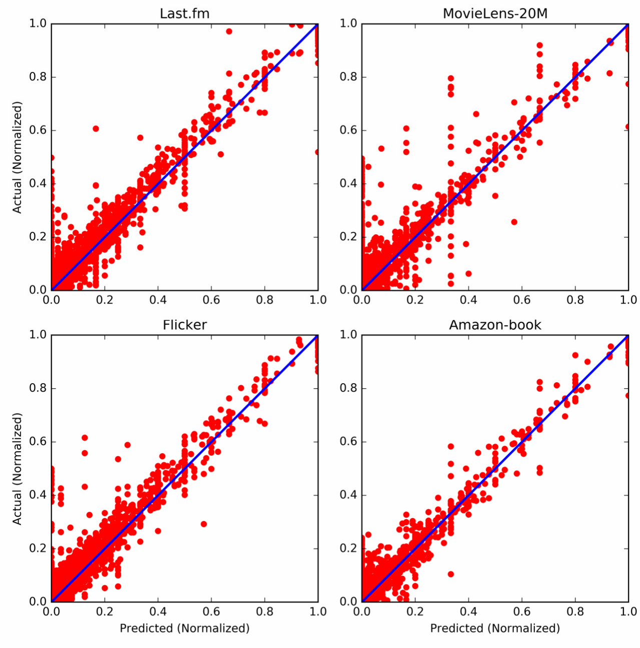

Figure 3. Comparison of predicted values and ground-truth values for different dataset

the time interval. Half a year's popularity of the film is the rating numbers obtained in half a year's time, thus we can obtain its the final popularity time series. The popularity time data from 1998 to 2002 were used to train COMFITS, and then the dynamic evolution of popularity after 2003 was predicted. As shown in the right-hand figure of Figure 5, the horizontal axis represents the time (half a year) and the vertical axis represents the popularity of Rain Man in each year. Among them, Blue solid line is the real time series of movie popularity, and COMFITS prediction results are represented by red dotted line.

Figure 4 and Figure 5 show only four representative predictions of content popularity. As can be seen from the figure, COMFITS can predict the evolution process of content popularity in different modes. For example, as shown in Figure 4, the popularity evolution of Ender's Game shows a single peak pattern, while that of the movie Farewell My Concubine in Figure 5 shows a periodic peak pattern. That means our proposed mix model is capable of grasping the evolving characteristics of popularity of different content by adjusting the learning weights and performance evaluation matrix of the model. Therefore, in this sense, it can be explained that the model is effective in predicting the accuracy of popularity and diversity of popularity evolution process.

4.2 Evaluation of the Performance of Predicting Popularity

Figure 3 illustrates to a certain extent that the model presented in this paper has high accuracy in predicting popularity. In this part, we will continue to analyze the performance of the model.

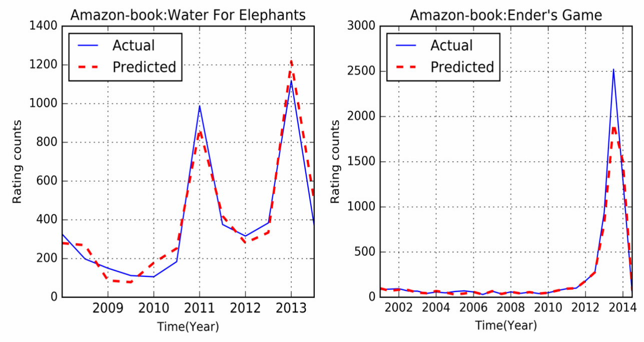

Figure 4. The prediction results of COMFITS on Amazon-book

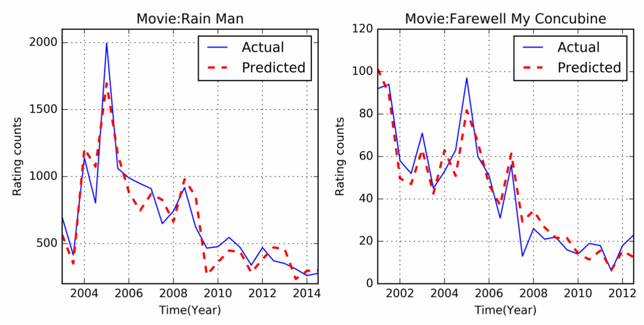

Figure 5. The prediction results of COMFITS on MovieLens-20M

Figure 6 shows the predictive performance comparison between COMFITS model and benchmark models, which will be expressed in boxplot. It can be seen that the COMFITS model always performs better than the comparative models in different prediction scenarios of different data sets, that is, for different social application platforms (Last.fm, Movie Lens, Flickr and Amazon) on content (tracks, movies, pictures, books).

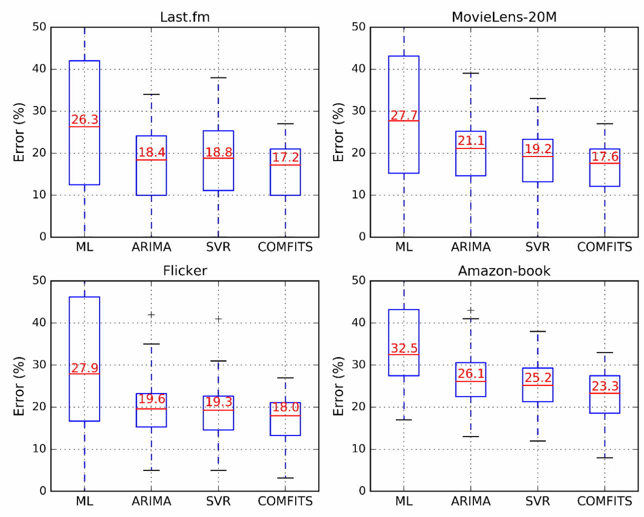

Figure 6. The performance comparison in popularity dynamic prediction

Taking Movie Lens-20M as an example, it can be seen from Figure 6 that the boxplot of COMFITS model has lower median and quartile than ARIMA, SVR and M-L. This shows that COMFITS model can provide better prediction for the popularity prediction process of the evaluation number obtained by most movies. In the Movie Lens-20M data set, the performance of SVR of the benchmark comparison model is better than that of other comparison models. The median value of prediction error of COMFITS model is nearly 8.3% better than that of SVR, and the mean value of prediction error is 16.2% better than that of SVR.

Throughout all data sets, COMFITS compared with the best performance benchmark comparison model in all data sets, the median prediction error decreases by 6.5%~9.2%, and the mean error decreases by 13.6%~23.3%. The experimental results can be explained as follows: the trend of popularity evolution of user-generated content in social networking platforms is diverse and heterogeneous. Popularity of user-generated content is affected by different factors such as content attraction, content publisher characteristics and so on. It has different communication life cycles and are manifested in different content popularity evolution patterns.

As shown in Figure 4 and Figure 5; the popularity evolution model of the same content's dissemination process is not invariable, it will be changed due to interference from external environment and other factors. For example, hotspot event can lead to a sharp increase in the attention of its relevant content on social platforms. Moreover, Different social network platforms have different network topology and popularity growth mechanism. Therefore, in the prediction of popularity evolution based on time series, it is difficult for a single model to accurately capture the evolution mode of popularity, which will future leads to the increase of prediction error. COMFITS can continuously adjust the weight of different prediction values in the combination model by evaluating the historic performance of each prediction model, and then, it will update the combination model to adapt to the change of evolution model, so as to improve the prediction accuracy.

It can also be concluded from Figure 6 that the prediction performance of SVR is superior to other benchmark comparison models (ARIMA and M-L), because it benefits from the advantages of support vector regression in non-linear feature extraction and high-dimensional pattern recognition in popularity evolution prediction. It should be noted that the median and quartile of ARIMA's prediction error of boxplot are lower than that of SVR in the evolution prediction of track popularity on Last.fm platform, that is to say, ARIMA performs better than SVR. That's because the track shows a stronger periodicity than other content in its whole propagation cycle. For example, users may play a certain kind of song regularly according to their interests or law of work and rest. As a classical time series analysis method, ARIMA is good at capturing such periodic features when dealing with such time series prediction. At the same time, on Last.fm platform, the prediction errors of the benchmark model SVR and ARIMA are also improved compared with those of other prediction scenarios(Movie Lens-20M, Flickr, Amazon-book). This shows that the prediction of track popularity evolution on Last.fm is easier than other prediction scenarios. It has a lot to do with the platform itself. Nevertheless, seasonality and periodicity of track popularity may change over time, so COMFITS's performance is still better than other benchmark models.

5 Conclusion

In view of the strong heterogeneity of content popularity evolution process in social networks, it is difficult for a single time series prediction model to accurately capture all kinds of popularity evolution models. A combined predicting model (COMFITS) that integrates multiple models (ARIMA, M-L, SVR) is established. By applying multi-class regression and analyzing the historical prediction performance of each sub-model, the combined weights of each sub-model for future time prediction are generated. COMFITS model can automatically adjust the weight of sub-prediction model according to the trend and characters of popularity evolution, and has achieved the desired results on different social network application platforms, and its performance is always better than the benchmark comparison models. The platform applicability of COMFITS has also been verified.

Acknowledgements

This work has been supported by the Fundamental Research Funds for the Central Universities (grant number: 2020JBM002), National Key Research and Development Program of China (grant number 2018YFC0831703) and China Postdoctoral Science Foundation (grant number: 2018M641172)

References

[1] H. A. Simon, Designing Organizations for an Information-rich World, in: M. Greenberger (Ed.), *Computers, Communications and the Public Interest*, The Johns Hopkins Press, 1971, pp. 37-72.

[2] N. Zarei, M. A. Ghayour, S. Hashemi, Road Traffic Prediction Using Context-Aware Random Forest Based on Volatility Nature of Traffic Flows, *5th Asian Conference on Intelligent Information and Database Systems*, Kuala Lumpur, Malaysia, 2013, pp. 196-205.

[3] H. Ma, W. Qian, F. Xia, X. He, J. Xu, A. Zhou, Towards Modeling Popularity of Microblogs, *Frontiers of Computer Science*, Vol. 7, No. 2, pp. 171-184, April, 2013.

[4] S. Kong, F. Ye, L. Feng, Predicting Future Retweet Counts in a Microblog, *Journal of Computational Information Systems*, Vol. 10, No. 4, pp. 1393-1404, February, 2014.

[5] L. Hong, O. Dan, B. D. Davison, Predicting Popular Messages in Twitter, *20th International Conference Companion on World Wide Web*, Hyderabad, India, 2011, pp. 57-58.

[6] M. Cornia, L. Baraldi, G. Serra, R. Cucchiara, Predicting Human Eye Fixations via an lstm-based Saliency Attentive Model, *IEEE Transactions on Image Processing*, Vol. 27, No. 10, pp. 5142-5154, October, 2018.

[7] M. Ahmed, S. Spagna, F. Huici, S. Niccolini, A Peek into the Future: Predicting the Evolution of Popularity in User Generated Content, *Sixth ACM International Conference on Web Search and Data Mining*, Rome, Italy, 2013, pp. 607-616.

[8] S. Jamali, H. Rangwala, Digging Digg: Comment Mining, Popularity Prediction, and Social Network Analysis, *2009 International Conference on Web Information Systems and Mining*, Shanghai, China, 2009, pp. 32-38.

[9] P. Yin, P. Luo, M. Wang, W.-C. Lee, A Straw Shows Which Way the Wind Blows: Ranking Potentially Popular Items from Early Votes, *Fifth ACM International Conference on Web Search and Data Mining*, Seattle, WA, United States, 2012, pp. 623-632.

[10] B. Chang, H. Zhu, Y. Ge, E. Chen, H. Xiong, C. Tan, Predicting the Popularity of Online Serials with Autoregressive Models, *23rd ACM International Conference on Conference on Information and Knowledge Management*, Shanghai China, 2014, pp. 1339-1348.

[11] J. M. Hofman, A. Sharma, D. J. Watts, Prediction and Explanation in Social Systems, *Science*, Vol. 355, No. 6324, pp. 486-488, February, 2017.

[12] B. Shulman, A. Sharma, D. Cosley, Predictability of Popularity: Gaps Between Prediction and Understanding, *International Conference on Weblogs and Social Media*, Cologne, Germany, 2016, pp. 348-357.

[13] T. Martin, J. M. Hofman, A. Sharma, A. Anderson, D. J. Watts, Exploring Limits to Prediction in Complex Social Systems, *25th International Conference on World Wide Web*, Montréal, Québec, Canada, 2016, pp. 683-694.

[14] M. Tsagkias, W. Weerkamp, M. De Rijke, Predicting the Volume of Comments on Online News Stories, *18th ACM Conference on Information and Knowledge Management*, Hong Kong, China, 2009, pp. 1765-1768.

[15] Z. Zhang, W. Zhang, F.-H. Tseng, Satellite Mobile Edge Computing: Improving Qos of High-speed Satellite-terrestrial Networks Using Edge Computing Techniques, *IEEE Network*, Vol. 33, No. 1, pp. 70-76, January/February, 2019.

[16] A. Kupavskii, L. Ostroumova, A. Umnov, S. Usachev, P. Serdyukov, G. Gusev, A. Kustarev, Prediction of Retweet Cascade Size over Time, *21st ACM International Conference on Information and Knowledge Management*, Maui, Hawaii, USA, 2012, pp. 2335-2338.

[17] A. Kupavskii, A. Umnov, G. Gusev, P. Serdyukov, Predicting the Audience Size of a Tweet, *Seventh International AAAI Conference on Weblogs and Social Media*, Cambridge, Massachusetts, USA, 2013, pp. 693-696.

[18] Z. Ma, A. Sun, G. Cong, Will this #hashtag be Popular Tomorrow? *35th international ACM SIGIR Conference on Research and Development in Information Retrieval*, Portland, Oregon, USA, 2012, pp. 1173-1174.

[19] L. Weng, F. Menczer, Y. Ahn, Virality Prediction and Community Structure in Social Networks, *Scientific Reports*, Vol. 3, Article number: 2522, August, 2013.

[20] W. Zhang, Z. Zhang, S. Zeadally, H.-C. Chao, V. C. M. Leung, MASM: A Multiple-algorithm Service Model for Energy-delay Optimization in Edge Artificial Intelligence, *IEEE Transactions on Industrial Informatics*, Vol. 15, No. 7, pp. 4216-4224, July, 2019.

[21] Q. Kong, W. Mao, G. Chen, D. Zeng, Exploring Trends and Patterns of Popularity Stage Evolution in Social Media, *IEEE Transactions on Systems, Man, and Cybernetics*, pp. 1-11, August, 2018.

[22] G. Szabo, B. A. Huberman, Predicting the Popularity of Online Content, *Communications of The ACM*, Vol. 53, No. 8, pp. 80-88, August, 2010.

[23] H. Pinto, J. M. Almeida, M. A. Goncalves, Using Early View Patterns to Predict the Popularity of Youtube Videos, *Sixth ACM international conference on Web Search and Data Mining*, Rome Italy, 2013, pp. 365-374.

[24] D. J. Bartholomew, Reviewed Work: Time Series Analysis Forecasting and Control by G. E. P. Box, G. M. Jenkins, *Journal of the Operational Research Society*, Vol. 22, No. 2, pp. 199-201, June, 1971.

[25] C. H. Wu, G.-H. Tzeng, R.-H. Lin, A Novel Hybrid Genetic Algorithm for Kernel Function and Parameter Optimization in Support Vector Regression, *Expert Systems with Applications*, Vol. 36, No. 3, pp. 4725-4735, April, 2009.

[26] R. A. Jacobs, M. I. Jordan, S. J. Nowlan, G. E. Hinton, Adaptive Mixtures of Local Experts, *Neural Computation*, Vol. 3, No. 1, pp. 79-87, Spring, 1991.

Biographies

Zhiyuan Zhang received his Ph.D. degree in Beijing Jiaotong University. Currently, he is a teacher in the School of Electronics and Information Engineering at Beijing Jiaotong University. His current research fields include social network analysis, recommender system, machine learning, data mining and artificial intelligence.

Zemin Bao received his Ph.D. Candidate in the Department of Communication and Information Systems, Beijing Jiaotong University. Now he worked in National Computer network Emergency Response technical Team/Coordination Center of China. His research interests include social network analysis, information networks and Popularity prediction.