

# RESEARCH ON DIFFUSION OF WECHAT BASED ON BASS - BP COMBINATION MODEL

Yaping Jiang

Shanghai Normal University Tianhua College, P.R.China

## Abstract

WeChat diffusion has the nonlinear variation characteristics and the law which can not effectively described by signal prediction algorithm. Therefore, this paper presents a Bass-BP diffusion model. Firstly, use the classical Bass model to preliminary forecast the WeChat's data, and then use the BP neural network to further non-linear approach the prediction results of Bass model. The results show that compared with Bass classic model, the Bass-BP combination model has better fitting effect and higher precision.

**Key words:** Bass model; BP neural network; WeChat diffusion

**JEL code:** P43

## 1. Introduction

WeChat is a free APP launched by Tencent on January 21, 2011 to provide intelligent terminals with instant messaging service, it is made by Tencent Guangzhou R & D center product team led by Zhang Xiaolong. WeChat supports cross-communication operators, cross-operating system platforms through the network to quickly send free (only need a small amount of network flow) voice messages, videos, pictures and text, at the same time, it can use the datum shared by streaming medias, and social plug-in based on location, like "Shake", "drift bottle", "friend circle", "public platform", "voice notepad" and other service plug-ins. By the first quarter of 2017, the number of WeChat's consolidated monthly active accounts reached 938 million, an increase of 23% over the same period last year.

The strong development of WeChat has attracted the attention of researchers. Pan Chen (2012) analyzes the adoption behavior of WeChat users from five dimensions: network externality, social influence, and motivation, etc. The results show that most users contact with WeChat due to friends' recommendation. Public praises have great impacts on users whether using WeChat or not. Tian Tian (2014) proposed the eight influencing factors that affect the spread of WeChat: communication channel, superiority, ease of use, observability, testability, personal innovation, social environment and WeChat compatibility. Baoguo Zhao (2014) for the first time use the Bass model to study the market spread of WeChat, and forecast the future adoption of potential users. The results show that the diffusion of WeChat is the successful diffusion of innovation. Xiaohui Deng (2015) analyzes with SPSS from the aspect of users' objective feeling to the using of product, get the conclusion that the factors that will influence users willing to use are that: cost, ease of use, social systems, diffusion channels and personal innovation. Yan Mao (2016) divided the users participating in the dissemination of WeChat into five categories, they are namely people don't know it, people still wait and see, disseminators, people are immune to it (who has forwarded it), people have withdrew from it, and she get the conclusion that user's attraction has the most important influence on the effect of dissemination. Yaping Jiang (2017) Based on the Bass classical model, to modify the model through introducing negative factors. The fitting results show that the prediction of the modified model is more accurate than the classical Bass model.

The study of the diffusion of WeChat is mainly focused on the prediction of a single model. The process of WeChat's diffusion is influenced by a number of factors, and it will show a certain degree of disorder. Therefore, it is difficult to determine that it is a single linear or non-linear system. Among many influencing factors, some are linear, some are non-linear, and single prediction models are difficult to predict accurately. This paper firstly through the combination of Bass model and BP neural network to analyze the diffusion of WeChat, so as to better predict the diffusion process of WeChat and to provide reference for the diffusion of mobile Internet products.

## 2. Introduction of Model

### 2.1 BASS model

The first presentation of Bass model was in 1968. At that time, Bass proposed a new model which consider both Fourt & Woodlock model and Mansfield model. According to it, a part of potential users of innovative products, who were influenced externally by mass media, were called innovator. Moreover, the other part of potential users who were influenced internally by word of mouth, were defined to be imitator. Based on this theory, Bass model which combined external and internal aspects was proposed. The model below is the basic equation about Bass model.

$$n(t) = \frac{dN(t)}{dt} = p * [m - N(t)] + q * \frac{N(t)}{m} * [m - N(t)] \quad (1)$$

In the equation,  $n(t)$  stands for the number of the new users at the time  $t$ .  $N(t)$  means the cumulative number of the new users at the time of  $t$ ,  $m$  is for the number of potential users.  $p$  stands for innovation coefficient which is about the possibility of consumers who will purchase the products due to the influence from mass media.  $q$  means the imitation coefficient which is about the possibility of consumers who will purchase the products due to the reference of existing users.

Equation 1 can be transformed as follows:

$$N(t) = y = m * \frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p} * e^{-(p+q)t}} \quad (2)$$

### 2.2 BP neural network

BP neural network is a multi-layer feedforward network to propagate training inversely according to the errors, and is one of the most widely used neural network models. Through training, BP network can learn the intrinsic relationship between input and output, establish nonlinear equations, and theoretically can approximate arbitrary equations at any degree of accuracy.

In the BP neural network, as shown in Fig.1, the information starts from the input layer, and the individual neurons of each layer are connected with all the cells in the next layer, but there is no connection between the same layers and the amount of information of next neurons is actually the sum of the weighted average of the input of the upper neuron information and the threshold it has. In the output layer, the difference between the actual output and the expected output forms the error value. The gradient descent theory is applied in inputting errors information with reverse trend and adjust the weight and threshold between neurons to minimize the error.

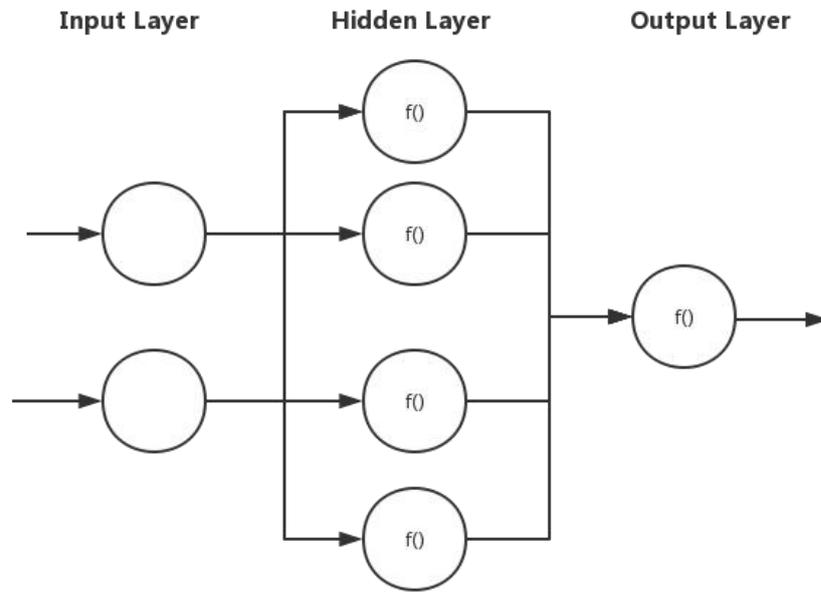


Figure 1. BP neural network

Suppose that BP neural network have  $N$  input neurons  $M$  hidden layer neurons,  $L$  output neurons, the actual output can be expressed as follows :

$$Z_k = f\left(\sum_{j=1}^M w_{jk} * o_j + b_k\right) = f(\text{net}_k), k = 1, 2, 3, \dots, L \quad (3)$$

$$o_j = f\left(\sum_{i=1}^N v_{ij} * x_i + b_j\right) = f(\text{net}_j), j = 1, 2, 3, \dots, M \quad (4)$$

In the formula:  $w_{jk}$  is the weight matrix from the hidden layer to the output layer;  $v_{ij}$  is the weight matrix of the input layer to the hidden layer, and the output of each layer is the input of the next layer. In addition,  $f(x)$  is a neuron transform function, and sigmoid or other transform functions can be used.  $B_k$ ,  $b_j$ , respectively, refers to the threshold of neurons in the hidden layer and the threshold of the output layer neurons.

The specific steps for predicting based on BP neural network are as follows:

- (1) The data acquisition of raw data and preprocessing
- (2) Design the network structure, determine the number of layers of network layer and the number of layers of neurons
- (3) Construct the training sample pair with the original data
- (4) Design learning algorithm to train neural network
- (5) For training the BP network meeting requirements, input a set of known data to get the corresponding predicated value.

### 2.3 Bass-BP combination model

There are many constraints in the classical Bass model, and the prediction results are difficult to approximate the nonlinear part. Therefore, this paper proposes a Bass-BP combination model. On the basis of classic Bass model, further use BP neural network to predict the residual, get the final forecast value, the specific forecast process shown in Fig. 2:

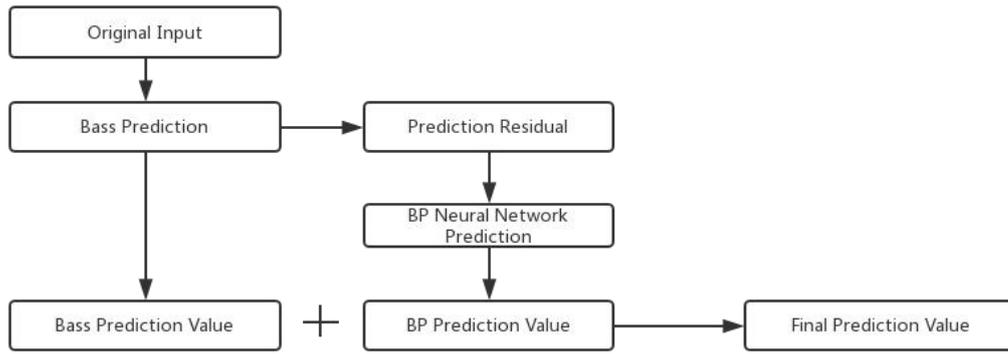


Figure 2. Bass-BP combination forecasting process

Take the data  $y(t)$  a set time series as the combination of linear and non-linear two parts. As shown in Equation 5,  $N(t)$  represents the  $t$ -time prediction result by using the classical Bass model,  $e(t)$  represents the prediction residual, and  $y(t)$  is the final prediction result.

$$y(t) = N(t) + e(t) \tag{5}$$

It is necessary to further use the BP neural network to train and predict the  $e(t)$  sequence. Since the residual data is only one-dimensional time series, the data must be reconstructed step by step. Only turn the one-dimensional sequence into a multi-dimensional structure, can use BP neural network to predict. Assume that the diffusion data is obtained quarterly, then use order 4 (four quarters of the year) to reconstruct sequence input as shown in Equation 6:

$$y = \begin{bmatrix} y_1, y_2, y_3, y_4 \\ y_2, y_3, y_4, y_5 \\ \dots \\ y_n, y_{n+1}, \dots, y_{n+3} \end{bmatrix} \tag{6}$$

### 3. Diffusion analysis of WeChat

#### 3.1 Data Sources

To test the effectiveness of the Bass-BP model, WeChat is chosen to be analyzed as the example of the diffusion of instant-messaging software. The first version of WeChat came out at the beginning of 2011. After operating for over 6 years, there are 938 millions of users in the first quarter of 2017, which diffuses successfully. Table 1 below shows the quarterly data of users of WeChat from 2011 to 2017.

Table 1. The number of WeChat users

| Year (quarter) | Actual Users (10millions) | Year (quarter) | Actual Users (10millions) | Year (quarter) | Actual Users (10millions) |
|----------------|---------------------------|----------------|---------------------------|----------------|---------------------------|
| 2011(Q2)       | 0.06                      | 2013(Q2)       | 2.35                      | 2015(Q2)       | 6.00                      |
| 2011(Q3)       | 0.15                      | 2013(Q3)       | 2.65                      | 2015(Q3)       | 6.5                       |
| 2011(Q4)       | 0.24                      | 2013(Q4)       | 3.3                       | 2015(Q4)       | 6.97                      |
| 2012(Q1)       | 0.5                       | 2014(Q1)       | 3.7                       | 2016(Q1)       | 7.62                      |
| 2012(Q2)       | 0.93                      | 2014(Q2)       | 4.1                       | 2016(Q2)       | 8.06                      |
| 2012(Q3)       | 1.18                      | 2014(Q3)       | 4.55                      | 2016(Q3)       | 8.46                      |
| 2012(Q4)       | 1.43                      | 2014(Q4)       | 5.13                      | 2016 (Q4)      | 8.89                      |
| 2013(Q1)       | 1.87                      | 2015(Q1)       | 5.49                      | 2017(Q1)       | 9.38                      |

### 3.2 Prediction on Bass Model

It is necessary to estimate the values of parameter  $m$ ,  $p$  and  $q$  before testing the model. The typical ways of parameter estimation are least square method, maximum likelihood method, nonlinear least square method. According to the comparison analysis from Fandong Meng, least square method may give higher or lower estimation. As for nonlinear least square method, its estimation has more accurate results.

Therefore, nonlinear least square method(OLS) is chosen in this article. Make the initial value  $m(0) = 5$ ,  $p(0) = 0.1$ ,  $q(0) = 0.1$ . Then after calculation, the estimated values of the parameters are showing in the table below.

Table 2. The estimations of parameters

| parameters | estimated values |
|------------|------------------|
| $m$        | 12.096           |
| $p$        | 0.011            |
| $q$        | 0.150            |
| $R^2$      | 0.998            |

Several points are gotten from the results above.

1.  $p < q$ . WeChat has diffused successfully since 2011.
2.  $R^2$  is close to 1. The estimated parameters have good fitting result and are able to reflect the actual situations of the users.
3.  $m = 12.096$ . There are around 130 millions population in China and 120.96 millions WeChat users, which indicates that WeChat gets a broad range of market shares. Besides, the software also satisfies the needs of customers in different ages.

We can further calculate the number of WeChat users, as shown in the Table3.

Table 3. The estimated numbers of WeChat users based on Bass Model

| Year(quarter ) | Estimated Numbers of the Users (10millions) | Year (quarter) | Estimated Numbers of the Users (10millions) | Year (quarter) | Estimated Numbers of the Users (10millions) |
|----------------|---|----------------|---|----------------|---|
| 2011(Q2)       | 0.14  | 2013(Q2)       | 2.2   | 2015(Q2)       | 6.01  |
| 2011(Q3)       | 0.31  | 2013(Q3)       | 2.6   | 2015(Q3)       | 6.52  |
| 2011(Q4)       | 0.49  | 2013(Q4)       | 3.02  | 2015(Q4)       | 7.03  |
| 2012(Q1)       | 0.70  | 2014(Q1)       | 3.48  | 2016(Q1)       | 7.52  |
| 2012(Q2)       | 0.94  | 2014(Q2)       | 3.95  | 2016(Q2)       | 7.98  |
| 2012(Q3)       | 1.21  | 2014(Q3)       | 4.45  | 2016(Q3)       | 8.42  |
| 2012(Q4)       | 1.51  | 2014(Q4)       | 4.96  | 2016(Q4)       | 8.83  |
| 2013(Q1)       | 1.84  | 2015(Q1)       | 5.49  | 2017(Q1)       | 9.21  |

To make curve fitting of the estimated number of users and the actual number of users, as shown in Fig.3, the estimated number of users is in line with the actual number of users, and it can be used as a Preliminary prediction results.

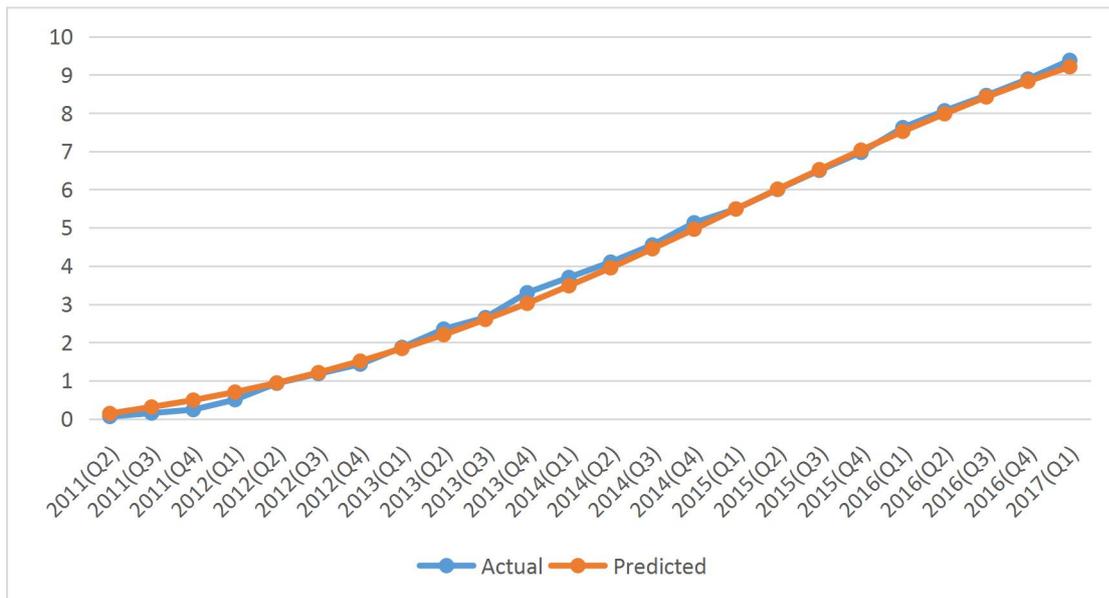


Figure 3. The Fitting curves of actual users and the estimated numbers of WeChat users

### 3.3 Prediction based on BP Neural Network

Through the number of users from the previous section which based on Bass model, we can further calculate the number of users residual, as shown in the Table 4 .

Table 4. The residual of number of WeChat predicted by Bass model

| Year (quarter) | residual (10millions) | Year (quarter) | residual (10millions) | Year (quarter) | residual (10millions) |
|----------------|-----------------------|----------------|-----------------------|----------------|-----------------------|
| 2011(Q2)       | -0.08                 | 2013(Q2)       | 0.15                  | 2015(Q2)       | -0.01                 |
| 2011(Q3)       | -0.16                 | 2013(Q3)       | 0.05                  | 2015(Q3)       | -0.02                 |
| 2011(Q4)       | -0.25                 | 2013(Q4)       | 0.28                  | 2015(Q4)       | -0.06                 |
| 2012(Q1)       | -0.20                 | 2014(Q1)       | 0.22                  | 2016(Q1)       | 0.10                  |
| 2012(Q2)       | -0.01                 | 2014(Q2)       | 0.15                  | 2016(Q2)       | 0.08                  |
| 2012(Q3)       | -0.03                 | 2014(Q3)       | 0.10                  | 2016(Q3)       | 0.04                  |
| 2012(Q4)       | -0.08                 | 2014(Q4)       | 0.17                  | 2016(Q4)       | 0.06                  |
| 2013(Q1)       | 0.03                  | 2015(Q1)       | 0.00                  | 2017(Q1)       | 0.17                  |

Through the above-mentioned topological reconstruction method, use one-year data as an iteration and remove the initial seven data, the one-dimensional data of Table 4 can be reconstructed into a  $17 * 4$  two-dimensional data set. Therefore, the predicted BP neural network input layer node is 4, the output layer node is 1, the hidden layer node is  $(2n + 1) = 9$ . In this paper, the first 11 sets of data were selected as the training set, and predict the data of later six groups. The prediction results are shown in Table 5.

Table 5. Prediction of the number of WeChat users

| Year (quarter) | Actual Value | Predicted Value On Bass Model | Predicted Value of Bass-BP combination model | Prediction error of Bass model | Prediction error of Bass-BP combination model |
|----------------|--------------|-------------------------------|--|--------------------------------|---|
| 2015(Q4)       | 6.97         | 7.03                          | 7.08   | 0.8%                           | 0.7%  |
| 2016(Q1)       | 7.62         | 7.52                          | 7.57   | 1.3%                           | 0.6%  |
| 2016(Q2)       | 8.06         | 7.98                          | 8.03   | 0.9%                           | 0.6%  |
| 2016(Q3)       | 8.46         | 8.42                          | 8.47   | 0.4%                           | 0.5%  |
| 2016(Q4)       | 8.89         | 8.83                          | 8.89   | 0.6%                           | 0.6%  |
| 2017(Q1)       | 9.38         | 9.21                          | 9.27   | 1.8%                           | 0.6%  |

Analyze the prediction results, the error of the combination prediction is less than that of the single Bass model, which indicates that the combination prediction model is more accurate for the prediction of the actual results. The comparison of prediction of two models is shown in Fig. 4.

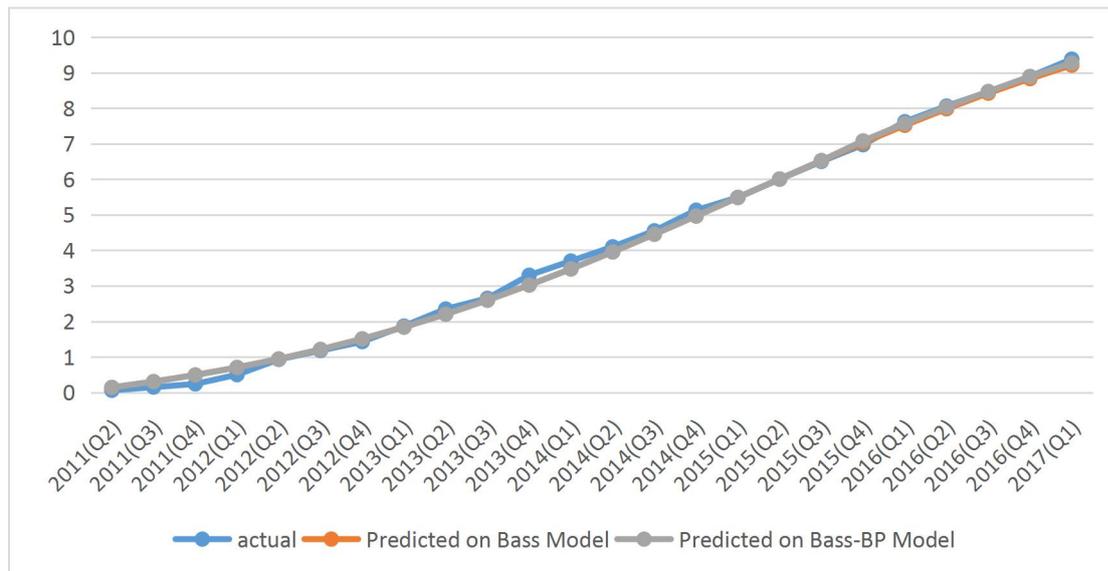


Figure 4. The prediction of the number of WeChat users

It can be seen from Fig.4 that the Bass model and the Bass-BP combination model all have good prediction characteristics. But in contrast, the results of combination prediction are more close to the actual data.

#### 4. Conclusion

Based on the Bass model and the BP neural network method, this paper proposes a new combination method, and establishes the combined prediction model, and analyzes and validates the actual data with the data of WeChat. The results show that the model proposed in this paper is effective and has good precision and good use value.

Compared with a single method, combination prediction is often able to learn from each other, and have better predictive results. The prediction work of the future services like WeChat can try to find the advantages from more forecasting models and integrate them. Select and dig deeply the historical data to reduce the forecast error and improve accuracy.

#### References

Pan Chen. The research of WeChat users' adoption based on the mobile Internet [D]. Huazhong University of Science and Technology, 2012.

Tian Tian. The WeChat diffusion under the Innovation diffusion theory and analyze the influencing factors of use [D]. Anhui University, 2014.

Baoguo Zhao, Feng Xuyan. Research on the diffusion of WeChat users based on Bass model [J]. Journal of Central University of Finance and Economics, 2014, (11): 108-112

Xiaohui Deng. Influence factors of products like Kiki [D]. Guangxi University, 2015.

Yan Mao. Research on the dissemination mechanism of WeChat information based on the theory of diffusion of innovation theory [D]. Shanghai University of Engineering and Science, 2016.

Yaping Jiang, Yueming Weng. Research on the Diffusion of Instant Communication Products Based on Bass Correction Model [J]. *China Management Informationization*, 2017, (11): 164-167.

F M Bass. A new product growth model for consumer durables [J]. *Management Science*, 1969, 15 (5): 215 - 227.

L A Fort, J W Woodlock. Early prediction of early success of new grocery products [J]. *Journal of Marketing*, 1960, 25 (2): 31-38.

E Mansfield, Technical change and the rate of imitation [J]. *Econometrical*, 1961, 29 (4): 741-766.

Fandong Meng, He Mingsheng. Research and Review of Parameter estimation method based on Bass Model [J]. *Space Control*, 2009, (01): 104-108.