



**HAL**  
open science

# Research on the Micro-blog User Behavior Model Based on Behavior Matrix

Zhongbao Liu, Changfeng Fu, Chia-Cheng Hu

► **To cite this version:**

Zhongbao Liu, Changfeng Fu, Chia-Cheng Hu. Research on the Micro-blog User Behavior Model Based on Behavior Matrix. 2nd International Conference on Intelligence Science (ICIS), Nov 2018, Beijing, China. pp.373-377, 10.1007/978-3-030-01313-4\_40 . hal-02118827

**HAL Id: hal-02118827**

**<https://inria.hal.science/hal-02118827v1>**

Submitted on 3 May 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Research on the Micro-blog User Behavior Model based on Behavior Matrix

Liu Zhongbao\*, Fu Changfeng, Hu Chia-Cheng

Quanzhou University of Information Engineering, Quanzhou 362000, China  
liu\_zhongbao@hotmail.com

**Abstract.** The micro-blog user behavior model based on behavior matrix is proposed based on the analysis of the current related researches. In the proposed method, two behavior matrices named original behavior matrix and individual behavior matrix are constructed; And then, the analysis method of behavior matrix is utilized to mine the representative behaviors, which can reflect the disciplines of the user; Finally, the experiments on the 2043 blogs of 9 users from June 15, 2016 to August 1, 2016 reflect the individual behavior disciplines.

**Keywords:** micro-blog user, behavior model, behavior matrix

## 1 Introduction

With the development of information technology, micro-blog has attracted much more attention because of its instantaneity and interactivity. We can use micro-blog freely to express our own opinions and easily obtain interested information. More and more micro-blog platforms have appeared since 2016, such as Twitter, Sina, Tencent, Netease. The number of micro-blog users has grown exponentially and its number has arrived at 200 million.

The user behavior in micro-blog has attracted many researchers' interest, and a lot of achievements related the above research have been made. Kshay et al draws the conclusions that the purpose of the micro-blog user is to discuss their activities and obtain their interested information. Brian points out that Twitter plays an important role in the communication between users in the disaster relief. Kaye analyzes the relationship between motivation, leadership and social media from the perspective of the user. Schwab et al obtained the user's interests by analyzing the keywords in micro-blog. Maloof et al discusses the problem of user interest drift in micro-blog with the help of forgetting mechanism. Xia discusses the features of user's behavior based on analyzing the topics, comments and other information the user visits in the micro-blog. Ping et al analyzes the relationship of micro-blog network based on the topology relationship between micro-blog users. Wang et al analyzes the number of concerns and fans, and the relationship between the number of blog articles based on the analysis of the structure and propagation of micro-blog. Zhao et al conducts statistical analysis on the number of concerns and blog articles.

It can be seen that the researches on micro-blog users are gradually popular, and the achievements greatly improve the micro-blog service. However, the problem of micro-blog authenticity has emerged with the popularity of micro-blog. In view of this, the micro-blog user behavior model based on behavior matrix is proposed in this paper. In the proposed method, two behavior matrices named original behavior matrix and individual behavior matrix are constructed; And then, the behavior vector space is generated by user behavior matrix, and the most representative characteristic behavior is mined by Principal Component Analysis (PCA), which reflected the disciplines of user behavior; Finally, simulation experiment results on the Sina micro-blog show the effectiveness of the proposed method.

## **2 Micro-blog user behavior Model based on behavior matrix**

We try to establish the micro-blog user behavior model by analysis of blog articles and the user behavior on the micro-blog. The micro-blog user behavior refers to the micro-blog user's behaviors, such as release, transmit and praise blog articles, concerns users, login and withdraw, and make comments, when they communicate with each others on the micro-blog platform in a certain period of time.

### **2.1 Behavior matrix model**

In order to accurately describe the user behavior, two behavior matrices, the original behavior matrix and the individual behavior matrix, are introduced. The original behavior matrix is used to store user behavior records, which describes the user behavior disciplines as the user activities in a certain period of time. The individual behavior matrix is used to analyze the behavior disciplines of different individuals, which aims to record the fragmented and irregular micro-blog behaviors, and represent the micro-blog user behavior with appropriate rules and algorithms.

#### **(1) Original behavior matrix**

In order to describe the user behavior and analyze the behavior disciplines, the original behavior matrix is proposed. The basic idea of the original behavior matrix is to describe the user behavior disciplines as the user activities in a period of time. The original behavior matrix describes the behavior of user  $k$  in  $n$  time steps within  $m$  observation days. The original behavior matrix is used to store user behavior records, which is the fundamental model of a series of behavioral matrices.

#### **(2) Individual behavior matrix**

The individual behavior matrix is used to analyze the behavior disciplines of different individuals. In order to further analyze the user behavior, the individual behavior matrix is proposed based on original behavior matrix. The individual behavior matrix describes the behavior of user  $k$  in  $n$  time steps, and the behavior of each time step is the summarization of  $m$  observation days.

## 2.2 The Analysis Method of Behavior Matrix

It can be seen from the statistical analysis, the observation samples of user behaviors reflect the stochastic characteristics, however, their behavior disciplines can be still identified. The main idea of the analysis method of behavior matrix is to mine the representative behaviors based on PCA, which can reflect the user disciplines, by constructing the user behavior vector space model based on user behavior matrices.

### (1) Behavior vector space model

The main idea of the behavior vector space model is to characterize the behavior disciplines of different individuals. The behavior disciplines are reflected by the distribution of the behavior vectors in a period of time. The vector representing the characteristic of user behaviors is defined as the behavior vector, and the  $n$ -dimensional space whose behavior vector belongs to is named as behavior vector space. The behavior discipline similarity of different individuals can be obtained by calculating the correlation coefficient of their behavior vectors.

### (2) Characteristic behavior analysis method

The main idea of the characteristic behavior analysis method is to mine the representative behaviors in the behavior vector by PCA. The eigenvectors (characteristic behavior) of correlation coefficient matrix can be obtained by the characteristic behavior analysis method, and let the disciplines of high-frequency behavior with high weights. The correlation coefficient matrix is used to describe the relationship between variables of the behavior matrix. The process of characteristic behavior analysis method is as follows.

Input: the user behavior matrix

Output: eigenvectors and their eigenvalues

Step1: calculate the correlation coefficient matrix  $R$  of the user behavior matrix;

Step2: calculate the eigenvectors of  $R$  and its eigenvalues;

Step3: descending sort according to the eigenvalues;

Step4: calculate the proportion and the cumulative proportion of each eigenvalue;

Step5: obtain the eigenvectors and eigenvalues of the behavior matrix and the proportional and cumulative proportions of the eigenvalues.

## 3 Experimental analysis

The 2043 micro-blogs released by 9 Sina micro-blog users from June 15, 2016 to August 1, 2016 were selected in our experiment, which was obtained by Sina micro-blog Application Programming Interface (API). In our experiment, a day was equally divided into 24 intervals and each time step was 60 minutes. The behavior matrix is analyzed by the characteristic behavior analysis method. The experiment results are recorded in Table 1, in which  $U_i$  ( $i=1,2,9$ ) stands for the micro-blog users.

**Table 1.** The correlation coefficients of 9 micro-blog users' individual behavior matrix

	U <sub>1</sub>	U <sub>2</sub>	U <sub>3</sub>	U <sub>4</sub>	U <sub>5</sub>	U <sub>6</sub>	U <sub>7</sub>	U <sub>8</sub>	U <sub>9</sub>
U <sub>1</sub>	1								
U <sub>2</sub>	0.24	1							
U <sub>3</sub>	0.52	0.51	1						
U <sub>4</sub>	0.82	0.54	0.70	1					
U <sub>5</sub>	0.64	0.64	0.79	0.70	1				
U <sub>6</sub>	0.44	0.28	0.43	0.78	0.34	1			
U <sub>7</sub>	0.77	0.45	0.72	0.77	0.81	0.70	1		
U <sub>8</sub>	0.02	0.03	0.02	0.01	0.15	0.29	0.20	1	
U <sub>9</sub>	0.67	0.30	0.73	0.71	0.62	0.66	0.75	0.03	1

It can be seen from Table 1 that the correlation coefficient between U<sub>8</sub> and other users is quite small, which indicates his activity disciplines are irrelevance to other users, and it can be inferred that his micro-blog behaviors is quite different from other users. The correlation coefficient U<sub>1</sub> and U<sub>4</sub> is largest, which indicates these two users have similar activity disciplines, and it can be inferred that their work environment and habits and customs maybe the same.

## 4 Conclusions

With the popularity of micro-blog, the number of micro-blog user grows exponentially. The problem of authenticity restricts the improvement of micro-blog services. In view of this, the micro-blog user behavior model based on behavior matrix is proposed based on the analysis of micro-blog user behaviors. Two behavior matrices named the original behavior matrix and the individual behavior matrix are introduced in our method, the former matrix is used to store the user original behavior records, and the latter is used to analyze the behavior disciplines of different individuals. The characteristic behavior analysis method is to mine the characteristic behaviors in the behavior vector by PCA. The simulation experiment results on Sina micro-blog show the effectiveness of the proposed method.

## References

1. Java A, Song X, Finin T, *et al.*, "Why We Twitter: Understanding Micro-blogging Usage and Communities", *Proceedings of the 9th Web KDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis*, pp.56-65 (2007)
2. Brian G. Smith., "Socially distributing public relations: Twitter, Haiti, and interactivity in social media", *Public Relations Review*, vol.36, no. 4, pp. 329-335 (2010)

3. Kaye D. Sweetser & Tom Kelleher, "A survey of social media use, motivation and leadership among public relations practitioners", *Public Relations Review*, vol.37,no.4, pp. 425-428 (2011)
4. Schwab I, Pohl W., "Learning User Profiles from Positive Examples", *Proceedings of the Acai'99 Workshop on Machine Learning in User Modeling* (1970)
5. Maloof M, Michalski S., "Selecting examples for partial memory learning", *Machine Learning*, no.41, pp. 27-52 (2000)
6. Xia Y.H., "The Structure and Mechanism of Microblog Interaction-Based on the Empirical Study of Sina Weibo", *Journalism & Communication*,no.4, pp. 60-69 (2010)
7. Ping L, Zong Y. L. "Research on Microblog Information Dissemination Based on SNA Centrality Analysis-A Case Study with Sina Microblog", *Document, Information & Knowledge*,no.6, pp. 92-97 (2010)
8. Wang X.G., "Empirical Analysis on Behavior Characteristics and Relation Characteristics of Micro-blogging Users-Take "Sina Micro-blog" for Example", *Library and Information Service*, vol.54, no.14, pp. 66-70 (2010)
9. Zhao W. B, Zhu Q. H, Wu K.W, *et al.* "Analysis of Micro-blogging User Character and Motivation-Take Micro-blogging of Hexun.com as an Example", *New Technology of Library and Information Service*, no.2, pp. 69-75 (2011)