Exploring Users' Temporal Characteristics of "Purchase and Comment" Behaviors Based on Human Dynamics

Liangqiang Li[1], Bo Lv[2], Liang Wu[2(✉)], and Hongwei Lin[3]

[1] Business School, Sichuan Agricultural University, Dujiangyan, Sichuan, China
lilq@sicau.edu.cn
[2] School of Economics and Management,
Guizhou Normal University, Guiyang, Guizhou, China
wuliangmail@yeah.net
[3] School of Public Health and Management, Hubei University of Medicine, Shiyan, China

Abstract. This paper carried out an empirical study on the formation rules underlying online user comments. Using the human dynamics method, which is used for describing human behavior, this paper explored two typical consumer human behaviors related to "purchase-comment" behaviors within an electronic commerce environment. Using massive real data crawled by Python web crawler, this paper analyzes three types of temporal characteristics (time interval, burstiness index, and memory index) of post-purchase comments by users, explores when consumers choose to disclose their shopping experiences on the Internet after purchase, and studies the formation rules of online user comments. The study also verified that the temporal characteristics of consumers' "purchase-comment" behaviors in the e-commerce environment also conformed to power-law distribution principles, thus providing new empirical insights for the study of online human behavior dynamics.

Keywords: Online comments · Human dynamics · Temporal characteristics · Purchase-comment time interval

1 Introduction

Online user comments often post online consumer experiences, opinions, and emotional tendencies regarding products, services, brands, which consumers share with other consumers through Web 2.0 technology in the form of texts, videos, symbols, and so on [1–3]. It has also become a habit for many consumers to view online comments before making purchase decisions and to share post-purchase experiences. Massive amounts of online shopping-related comments and user generated content (UGC) with regard to

This research is support by Major Research Project of Innovative Groups in Guizhou Provincial Education Department (No. Qian Jiao He KY Zi [2017]034).

H. Song and D. Jiang (Eds.): SIMUtools 2020, LNICST 370, pp. 420–438, 2021.
https://doi.org/10.1007/978-3-030-72795-6_35

sharing experiences through social media constitute the main source of big data, and these have gradually begun to dominate the generation of website content.

Online user comments benefit increase interaction between enterprise platforms and customers; furthermore, they can create a good reputation and gain profits for enterprises. Online user comments have also played an important role in promoting the development of e-commerce. On one hand, online user comments influence consumers by eliminating the uncertainty of product information, thus affecting their online shopping decisions. In the case of new products or experiences, online user comments are an important information disclosure channel that help to eliminate potential consumers' uncertainty regarding the product. It is also used to educate potential customers. Being introduced to the experiences of other consumers can help a potential consumer use the product well. On the other hand, online user comments can influence the adjustments made to e-commerce enterprise platforms' operation strategies and the product perfection and service improvements undertaken by manufacturers.

1.1 Importance of Research on Temporal Characteristics of Users' "purchase and Comment" Behaviors

Scholars have conducted extensive research on online user comments from multiple perspectives. Previous studies have mainly focused on consumers' sentiment, usefulness [1, 3], influencing factors [4, 5], and the application of online comment corpus mining with regard to usefulness, emotional analysis, recommendation systems, and other aspects [7, 8]. However, little research has been conducted regarding the formation rules underlying the timing of online user comments (i.e. at what point users post comments after purchasing products or services). Such formation rules have great significance for theory as well as practice.

Online comments are an important source of product information, and they can affect consumers' product attitudes and purchasing behaviors [9]. Making online user comments is a voluntary contribution by consumers; undertaking this activity requires a certain amount of time and energy, and it has both cost and benefit effects. The point at which users choose to post comments on the Internet is restricted by various motivations and certain constraints. Through experimental research, McGuire and Kable found that human made decisions respond according to certain statistical time clues and that people adjust their persistence level according to the distribution of delay time intervals they encounter [10]. Therefore, consumers also tend to adopt certain strategies in order to obtain the corresponding benefits when deciding when to release a comment [11].

For potential consumers, online user comments disclose user experience-related information that is not offered by e-commerce platforms' product introductions or advertisements. The disclosure of such information plays an important role in easing any uncertainty regarding potential consumers' purchase decisions. The points of time (time points) at which any existing consumer comments are published will affect the decision-making time points of potential consumers. Some studies show that the time between the "reading time" of potential users and the "release time" of comments has a positive impact on the consumer awareness of usefulness regarding the comments [12]; the release time points of consumers' comments also have different influences on potential consumers' purchasing decisions. Recently published comments have a greater influence

on purchasing decisions to be made in the near future. Older comments are considered more important for making long-term purchase decisions. Purchase-related comments that are released after a longer consumer experience period can also make people feel that the comments are more credible [13]. If an e-commerce platform, which has mastered the general temporal rules underlying users' comments, chooses appropriate time points to motivate consumers and promotes the formation of online user comments, it can persuade consumers to generate effective comments within a specified time range. Therefore, studying the temporal characteristics of the users' "purchase-comment" behaviors and the rules underlying comment formation is important.

1.2 Using Human Behavior Dynamics to Characterize the Temporal Characteristics of User Comment Behaviors

Users purchase goods and release comments online are two typical online behavior modes. In order to understand the activity rules underlying these two behaviors, it may prove useful to discuss them using the dynamic mechanisms underlying human behaviors. The Poisson process framework is widely used to quantitatively describe the behavior rules and influences underlying human behavior activities [14]; that is, within a given time interval $[t, t+\tau]$, the time interval τ between two successive behaviors follows exponential distribution: $p(\tau) = \lambda e^{-\lambda\tau}$. The probability of occurrence of these two behaviors is not related to time t but only to their interval τ.

With the development of computer and Internet technologies, extensive data related to various human activities can be collected and stored. Thus, it is now possible to use real data to illustrate human activities. In a 2005 Nature article, Barabási dscribed an empirical counting of the time intervals between users' sending and replying behaviors regarding e-mails and ordinary mails and found that this human behavior follows a power-law distribution: $p(\tau) = \tau^{-\nu}$; this finding has opened up a new direction for modern "human dynamics" research [15]. This study found that the distribution of human behavior rules was uneven and deviated from the Poisson process in terms of classical assumptions. This uneven distribution of human behavior rules is characterized by both "short-term outbreak" and "long-term silence."

With the emergence of Web 2.0, the concept of user participation has become increasingly popular. Online life has become an important part of modern life. People's behavior motivation patterns on the Internet have also attracted more and more scholars' strong attention and research interest. Scholars have studied the various dynamic modes underlying people's online behaviors such as web browsing [16], online movies watching [17], online game and music [18], online instant messaging (QQ) [19], behaviors in online games [20], editing entries on Wikipedia [21, 22], login behavior of information system users [23], topic-related reply behaviors on social networks [24], and customer behaviors in e-commerce transactions [25].

Empirical data analysis of these studies shows that the time intervals involved in online human behaviors follow a power-law distribution with an index of 1–3. Vázquez et al. [26] divided the distribution of human behavior rules into two universal classes with indices of -1 and -1.5, respectively, but several empirical research results have gradually expanded beyond the scope of these two universal classes [17].

Therefore, as an important part of online human behaviors, the two typical online human behavior laws underlying consumer purchase and comment behaviors within an e-commerce environment can also be characterized by using the human behavior dynamic mechanism method as a reference. A large amount of data on consumers' purchase and review behavior accumulated on e-commerce platforms; these data are useful for studying the behavioral dynamics underlying people's buying-related comments. This article also used the time interval characteristics underlying users' purchase behaviors and comment behaviors within e-commerce environments to explore whether they also conformed to the general rules of online human behaviors.

The rest of this paper is organized as follows: Sect. 2 introduces related works; Sect. 3 proposes three kinds of indicators for analyzing the temporal characteristics of "purchase-comment" behaviors based on human behavior dynamics and provides the whole experiment's framework; Sect. 4 reports the experiment results with real data; and finally, Sect. 5 concludes this paper.

2 Relevant Works

This paper studies the temporal characteristics of users' "purchase-comments" within e-commerce environments is mainly based on the theories and methods that describe the mechanisms underlying online human behavior dynamics. One of the most commonly used distribution theories with regard to online human behavior dynamics is the power-law distribution theory (also called long tail theory).

Pareto [27] and Zipf [28] have conducted pioneering work on power-law distribution. If the probability distribution or density function of a random variable is in the form of $p(k) = ck^r$, where c and r are constants, the random variable can be said to conform to the power-law distribution. Most real-world distributions conform to the negative exponent of the power law distribution—that is, $r < 0$. Power-law distribution has an important characteristic—scale-invariance; that is, $f(cx) = \alpha(cx)^k = c^k f(x) \propto f(x)$. By considering logarithms on both sides of the power-law distribution function, $ln\, p(k) = r\, ln\, k + ln\, c$ is obtained. This function is represented as a straight line in the double logarithmic coordinates, and this graphical feature is utilized for judging whether the random variable conforms to the power-law [29]. Under power-law distribution, the frequency of most events is small, while the frequency of a few events can be large. When k increases, $P(k)$ does not rapidly approach 0 in exponential form but gradually approaches 0 in a relatively gentle power-law form, thus showing "long tail" and "wide tail" properties [30].

Power-law distribution takes various forms in various professional fields. In different subject areas, scholars have also identified corresponding characteristic rules, thus gradually setting off an upsurge of empirical research on the power-law phenomenon. For example, scholars have discussed the occurrence laws underlying extreme events such as the distribution of lunar crater diameters [31], the distribution of earthquake levels and frequencies [32], and the distribution of war scales and frequencies [33]. Simultaneously, other distributions, such as surname distribution [34], the distribution of numbers of papers published [35], distribution of numbers of web pages in websites [36] and the time interval distribution of online human behaviors can also conform to power-law distribution and thus show a typical "heavy tail" phenomenon.

3 Analysis of the Temporal Characteristics of "Purchase and Comment" Behaviors Based on Human Behavior Dynamics

This section uses a method based on the dynamic mechanisms underlying human behaviors to analyze the temporal characteristics of the two typical behaviors underlying users' "purchase-comment" behaviors within e-commerce environments. First, the paper discusses the correlation indexes (time interval, burstiness index, and memory index) of time characteristic analysis. Second, the main method used for describing time characteristic rules is introduced, and the stage division method of time interval series is provided. Finally, the paper provides an experimental scheme for analyzing "purchase-comment" behaviors from a time characteristic perspective based on human behavior dynamics.

3.1 Relevant Indicators for Analyzing the Temporal Characteristics of "Purchase and Comment" Behaviors

Utilizing the human dynamics method for depicting human behavior rules, this paper mainly uses three related indicators in its analysis of the temporal characteristics underlying users' "purchase-comment" behaviors within e-commerce environments: (1) Time interval; (2) burstiness index; and (3) memory index.

3.1.1 Time Interval

"Time interval" represents the time difference between the occurrences of two consecutive events within a certain time range. For example, if a consumer buys a product on "January 2, 2016" and makes a comment on "January 8, 2016" after receiving the product or experiencing it, the time interval between the consumer's two consecutive behaviors (i.e., buying and commenting) is 6 days. If this product receives n comments on the Internet, this indicates that there are n "buy-evaluate" time intervals. In general, the e-commerce environment includes three kinds of time intervals related to consumers' purchase and comment behaviors.

Time Interval Between Adjacent Purchases
The time interval between adjacent purchases is the interval between two purchases made by a consumer. The consumer's purchase cycle can be perceived from the perspective of the individual. At the group level, it is used to depict the time difference between the purchase times for a product among different consumers; this reflects the sales situation of the product.

In order to calculate the time intervals between the adjacent purchasing behaviors of consumers, all purchasing behaviors are first sorted based on time to form a purchasing behavior occurrence time series; second, the purchase time interval series is obtained by subtracting between two purchase time series. Finally, the interval sequence is calculated. Similarly, for the time series analysis regarding purchase behaviors, the distribution in different time cycles can be considered.

Time Interval Between Adjacent Comments
The interval between the release times for adjacent comments can also be viewed from two perspectives; on the one hand, it is aimed at all comments made by a consumer,

thus describing the distribution law for individual consumers at the time point of the comments; on the other hand, it reflects the public praise of all consumers for a given product.

For calculating the release time intervals for adjacent comments, first, all users' comments on a certain product $C = \{r_1, r_2, ..., r_n\}$ are sorted according to their release time in order to form a comment release time series (from the earliest release time to the latest release time) $\{t_1, t_2, ...t_n\}$; second, two adjacent release times are subtracted to form a time interval series for releasing comments $\{\tau_1, \tau_2, ...\tau_n\}$; and finally, the time interval series is used for calculating probability distribution and ensuring the power-law fitting of the main part of time interval sequence.

Time Intervals Between Consumers' "Purchase-Comment" Behaviors
The comment page not only provides the consumers' comment time for a given product but also provides the consumers' purchase time for that product. To some extent, the calculation of the time interval between the two behaviors ("purchase-comment") reflects the motivation underlying the consumer's comment release and the factors that influence online comment generation. Therefore, by capturing all the consumer comments on the comment page of the target product C, the "purchase-comment" time interval for consumers can be calculated for each comment c_i; this method is used to measure how long consumers will post comments after making a purchase. The calculation formula is: $x_i = reviewing\_time - purchasing\_time$, where *reviewing_time* is the time point at which the consumer releases comments, and *purchasing_time* is the time point at which the consumer purchases the product.

After calculating the time interval for consumer behaviors, it is necessary to describe the rules governing these intervals. The main description method is as follows. By calculating the frequency $y_i = frequency(x_i)$ of the time interval x_i, the data series for the "purchase-comment" time interval of consumers is obtained: $\mathrm{T} = \{(x_i, y_i)\}_{i=1,...,|C|}$. This data series is used to depict the time distribution for consumers' purchases and comments ("purchase-comment" behavior).

3.1.2 Burstiness Index

Intermittence is a physical statistic that characterizes the frequency degree of human activities within a short time period and the degree of silence within a long time perod [15, 26]. Goh and Barabasi define intermittence as follows [37]:

$$B = \frac{(\sigma_\tau / m_\tau - 1)}{(\sigma_\tau / m_\tau + 1)} = \frac{\sigma_\tau - m_\tau}{\sigma_\tau + m_\tau} \tag{1}$$

Where σ_τ represents the standard deviation of the event interval sequence τ, and m_τ represents the average value of the time interval series τ. If the standard deviation σ_τ and the mean value m_τ are equal, the statistical value of intermittence is $B = 0$. For example, if the mean and standard deviation in the exponential distribution are equal, the intermittence value is $B = 0$. For the "heavy tail" distribution, the standard deviation is much larger than the mean, and the final intermittence value approaches [23].

The definition of burstiness index can also be used to describe the frequency of human behavior—that is, whether there will be many purchases and comments within

a short period of time. For example, it can be used for predicting whether people will generate a large number of purchasing activities in the Double Eleven Shopping Canival of China; it can also be used for focusing on comments made for purchased goods after receiving the goods, while there is no purchasing behavior and comment behavior for a long period during which there is no promotion activity.

3.1.3 Memory Index

The memory index is used to describe the correlation of adjacent subsequences in a given time interval series [14, 25]. The specific process for calculating memory index is as follows.

First, the time series is formed according to the time series of the occurrence of events $\{t_1, t_2, \ldots, t_{n+1}\}$, and every two adjacent time points are subtracted to obtain a time interval series $\{\tau_1, \tau_2, \ldots \tau_n\}$, with n_τ representing the total number of time interval series; a time interval series is divided into two time interval subsequences: $\{\tau_1, \tau_2, \ldots \tau_{n-1}\}$ and $\{\tau_2, \tau_3, \ldots \tau_n\}$. σ_1, σ_2 are the standard deviations of the two subsequences, and m_1, m_2 are the mean values of these two subsequences. The calculation formula is as follows [27]:

$$M = \frac{1}{n_\tau - 1} \sum_{i=1}^{n_\tau - 1} \frac{(\tau_i - m_1)(\tau_{i+1} - m_2)}{\sigma_1 \sigma_2} \tag{2}$$

The memory index is used to describe the laws underlying consumers' purchase and comment behavior in e-commerce environments. It mainly reflects whether consumers will experience a longer or shorter time interval after purchasing or commenting for a longer or shorter time interval. This indicator can help e-commerce platforms master the principles of consumer purchasing behavior and online comments generation; this will aid them in selecting appropriate time points to stimulate purchases and generate comments.

3.2 Characterization Methods for Time Characteristic Rules

In the world of online human behavior dynamics, the distribution rules and characteristics of the time interval between two consecutive behaviors are generally used to describe the activity rules governing human behavior. The most common characterization method for time interval distribution rules is to fit their distribution with power-law distribution and use statistical tests to verify whether they conform to power-law distribution.

The fitting of power-law distribution mainly involves the fitting of the power-law index. Clauset et al. gave a detailed derivation and explanation for the power-law distribution fitting [38]. This paper takes Clauset's method as a reference, mainly using linear regression and least square method to complete the curve fitting of the power-law function, thus obtaining the fitting straight line slope for the main data, namely the power index [37, 39].

The specific method for determining the power-law distribution fitting is as follows:

The distribution curve of the general power-law function is $y = ax^{-b}$, where $a > 0$; in this paper, y represents the probability of occurrence for the time interval, and x represents the time interval between consumer purchase and comment.

- Taking the logarithm of both sides of $y = ax^{-b}$, $ln\, y = (-b)\, ln\, x + ln\, a$ can be obtained;
- Certain substitutions are performed ($u = ln\, x$,$v = ln\, y$, $A = ln\, a$), and the curve equation of the power-law distribution function becomes a straight line equation: $v = A + (-b)u$.
- Using the empirical data value (x_i, y_i), we obtain (u_i, v_i) and $i = 1, 2, \ldots, n$.
- Using the least square regression straight line to calculate u and v, the following equation can be obtained:

$$\hat{b} = -\frac{\sum_{i=1}^{n} u_i v_i - \left(\sum_{i=1}^{n} u_i\right)\left(\sum_{i=1}^{n} v_i\right)}{n\sum_{i=1}^{n} u_i^2 - \left(\sum_{i=1}^{n} u_i\right)^2}$$

- $A = \overline{v} + \hat{b}\overline{u}$.
- Because $a = e^{\hat{A}}$, the value of a is obtained. Therefore, $y = ax^{-b}$.

3.3 Division Method of Time Interval Series

In order to explore when and for what reasons users choose to release online comments after making a purchase, this paper divides the "purchase-comment" time interval series for consumers into stages and explores the behavior characteristics of users at different stages [40, 41].

The division of the time interval stages is implemented on the interval sequence $T = \{(x_i, y_i)\}_{i=1,\ldots,|C|}$, and the specific division method implementation rules are as follows.

First, the starting point $p_1 = (x_1, y_1)$ and the ending point $p_{|C|} = \left(x_{|C|}, y_{|C|}\right)$ are selected as the initial points of division.

Next, the distance between them and the starting point and the ending point $d_i = \frac{|(p_{|C|}-p_1)\times(p_1-p_i)|}{|p_{|C|}-p_1|}$ is calculated for other points $p_i = (x_i, y_i)$ and lines $\overline{p_1 p_{|C|}}$ in the interval sequence (excluding the starting point and the ending point) [42, 43].

The dividing point $i^* = \arg_{i\in\{2,\ldots,|C|-1\}} \max\{d_i\}$ is determined according to a predefined threshold value, and the number of stages are determined based on stage division, where the value of d_{i^*} is greater than the predefined distance value d_0; this process is repeated until all points with distances greater than d_0 are found.

An example of the time series stage divisions is shown in Fig. 1.

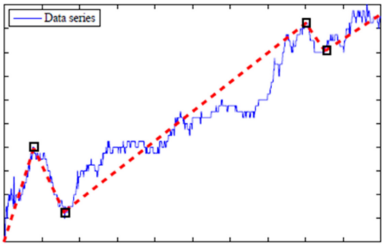


Fig. 1. Example of stage division of data series

3.4 Experimental Plan for Analyzing the Temporal Characteristics of "Purchase and Comment" Behaviors

The overall experimental framework of this paper is shown in Fig. 2. Which is mainly composed of three task stages. The first stage is the online user comment data acquisition stage. The second stage is the comment data preprocessing stage. The third stage is the time series data analysis stage.

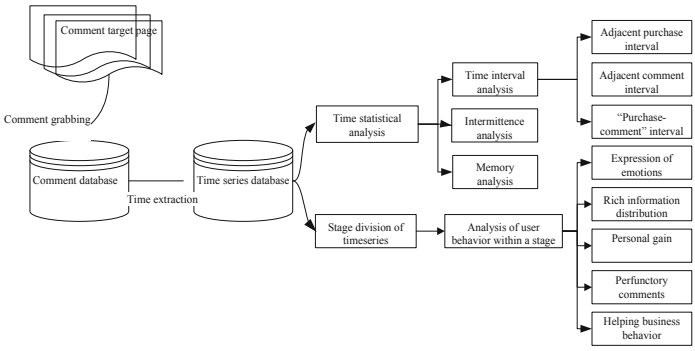


Fig. 2. The analysis framework for users' "purchase-comment" behavior dynamics

The main tasks for data acquisition follow 5 stages: (1) select target website and target comment webpage containing the e-commerce online user comments; (2) analysis the source code of webpage; (3) the determination of the presentation rules for online user comments on the selected webpage; (4) a combination of Python's web crawler toolkit and offline web crawler software (Locoy Spider) for customizing the regular expressions for web crawling rules; and (5) utilization of the web crawler for completing the grabbing of comment page data and forming the initial comment database [44].

The specific preprocessing work for comment data is as follows. First, the initial database is cleaned up to form an effective comment data set. The main operations include clearing duplicate comment records in the initial database and deleting blank data records that may not have been successfully captured due to inconsistent rules. Second, time data related to "purchase and comment" behavior are extracted from the effective data set to form a time data set. Table 1 shows an example of the time interval data set for the two important consumer behaviors (purchase and comment); the units used for recording time are days or minutes.

Two major blocks are involved in the experimental tasks for the data analysis stage. The first is the statistical analysis of time intervals. The main work is as follows. First, time intervals are calculated based on time series data sets in order to form time interval series data sets; second, probability statistics are carried out using the time intervals; and third, the event interval distribution for adjacent purchasing behaviors, the time interval distribution for adjacent comments, and the time interval distribution for "purchase-comment" behaviors are provided. The Kolmogorov-Smirnov (KS) test is carried out on each distribution to verify whether it conforms to the power-law distribution. The intermittence and memory of interval sequences are calculated again. The second major

Table 1. Example of time data set

ID	Purchasing time	Commenting time	Time interval (days)
ID1	6-30-2011	7-1-2011	1
ID2	08-01-2012	08-08-2012…	7
…	…	…	…

task is to divide the stages of the "purchase-comment" time interval series and analyze the user comment behavior at each stage.

4 Experiment Results

4.1 Data Description

According to the product classification used at the target website, comments from popular categories that contained numerous comments were selected as crawling objects. Considering representativeness, snowball sampling method was used to extracting the comment data of representative individual products in different categories. At the same time, product type like search-type and experience-type was also considered. Due to privacy protection proxy, after conducting this experiment, the comment page displayed in the online mall had been revised and that it no longer provided purchase time information. Data selected for this experiment were the comment data captured before the revision, and the time span of the selected products was relatively long.

The selected data were divided into three types. The first type was for personal historical purchase records and comment records; the second type involved selecting data sets for six major product types; and the third type included search-type and experience-type products, though it was the comment data set for each individual product.

The data set for personal historical purchases and comments used a random sampling method to capture 11,081 purchase records from 152 users of the selected online mall (name: JD) from December 1, 2008 to April 14, 2014.

For popular categories selection, this study selected six categories (office electronic products, notebook computers, routers and network cards, mobile phones, decorations, and cosmetics) for the comments. Product type, product category, number of products included in the category, and the total number of comments for each category are shown in Table 2. For this category, "purchase-comment" interval was mainly considered.

For single product selection, only the product names that were used frequently in the four experiments are listed here. The product name, product type, time span, and number of comments for these four individual products are shown in Table 3. In addition, comment data for another 55 individual products were randomly captured, but their detailed description information is not repeated here due to space reasons. For individual products, not only was the time interval for their "purchase-comment" behaviors considered in the data, but their comments were also sorted; furthermore, the time interval distribution between adjacent comments was explored.

Table 2. Data description of six product categories

Product type	Product category	Number of products	Number of comments
Search type	Office electronic products	10	15 928
	Notebook computers	12	34 504
	Routers and network cards	9	285 535
	Mobile phones	71	487 817
Experience type	Decorations	10	3450
	Cosmetics	11	42 545

Table 3. Data set description for four individual products

Product type	Product name	Time span	Number of comments
Search-type	D90 Single-lens reflex digital camera	November 6, 2008 to December 31, 2011	5460
	Double Happiness badminton racket	May 22, 2010 to December 8, 2013	13 733
Experience-type	Books: One Hundred Years of Solitude	June 1, 2011 to December 8, 2013	49 422
	Red wine	February 14, 2011 to July 25, 2012	4246

4.2 Analysis Results for the Temporal Characteristics of Purchase Behavior

4.2.1 Data Description

For dynamic analysis of users' purchase behavior, this study used a data set of 152 users' purchase history, with a total of 11 081 purchase records and 73.87 purchase records per person on average; these included at least 2 purchase records and at most 1063 purchase records per person.

4.2.2 Time Interval Characteristics of Adjacent Purchasing Behaviors

In order to verify the temporal characteristics shared by users' adjacent purchasing behaviors, the purchase records of a consumer (ID: Kom_isun) from a purchase database of 152 consumers (with a total of 857 records) were selected, and a time span from May 19, 2009 to October 25, 2013 was chosen. Figure 3 shows the online purchase sequence for the consumer with the ID Kom_isun. The horizontal axis of the series indicates the interval (in days) between the first purchase and the current purchase, and the vertical axis indicates that the purchase occurred on the same day.

Figure 4 gives the scatter plot for the time interval of Kom_isun's adjacent purchase behavior and the power-law distribution fitting for the main part data. The exponent of power-law fitting is -1.372, and the goodness of fit is 0.853.

Fig. 3. Online shopping series (Kom_isun)

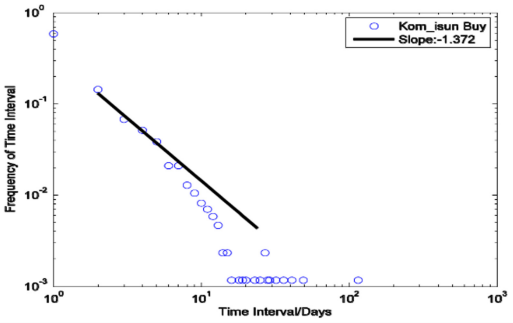


Fig. 4. Statistical double logarithmic chart for adjacent purchase time interval (Kom_isun).

Using Sect. 3.1 to introduce the calculation formula for intermittence and memory of purchase behavior, it was found that the burstiness index value for the adjacent purchasing behaviors of Kom_isun was 0.50128, thus showing intensive purchase behavior for a short time and non-purchase behavior for a long time. Statistics regarding its purchasing behaviors showed that 58.52% of all the purchasing records for the user had no interval (i.e., 0 day) between adjacent purchasing behaviors. The proportion within 3 days was 84.8%; the longest interval between two adjacent purchases was 114 days. The M value calculated by depicting the memory of purchasing behavior was 0.1264, which indicates that the purchasing behaviors of this user were characterized based on memory. That is, after adjacent purchasing behaviors were conducted within a short time period, some adjacent purchasing behaviors continued within a short time period.

4.3 Analysis Results for Temporal Characteristics of Comment Behaviors

According to the temporal characteristics' statistics for comment release through online user comment behaviors, the time intervals for pairs of adjacent comments released by consumers in the product comments category were depicted based on the data of three separate products (D90 single-lens reflex digital camera, red wine, and badminton racket). The shortest time interval, longest time interval, standard deviation of time interval, and average value for the three groups of data are shown in Table 4.

Figures 5, 6, and 7 show the probability distributions for the time intervals of the three products under the double logarithmic coordinate and the power-law fitting of the main part's curve, respectively. The circle in the figure represents the original curve of the release time interval between two adjacent comments, and the black straight line represents the fitting curve. The power-law indexes of the three data groups were −1.092, −0.935, and −0.878, respectively, and the goodness of fit was above 85%. The Kolmogorov-Smirnov (KS) test results showed that the release time intervals for adjacent

Table 4. Basic description of adjacent comment time intervals of three products

Product type	Minimum interval (minutes)	Maximum interval (minutes)	Standard deviation	Mean value
D90 digital camera	0	65 406	1115.10	271.27
Red wine	0	9 800	351.16	178.36
Badminton racket	0	15 236	303.38	128.68

comments conformed to the power-law distribution and that the end of the distribution had obvious "fat tail" characteristics.

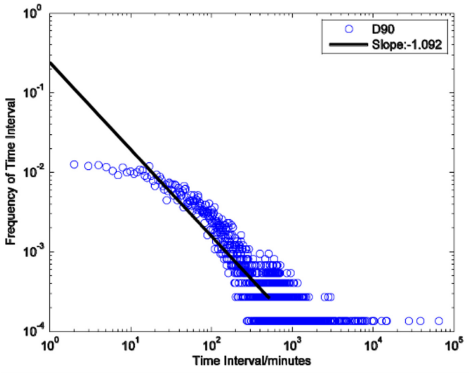


Fig. 5. Time interval distribution for adjacent comments regarding D90 digital camera

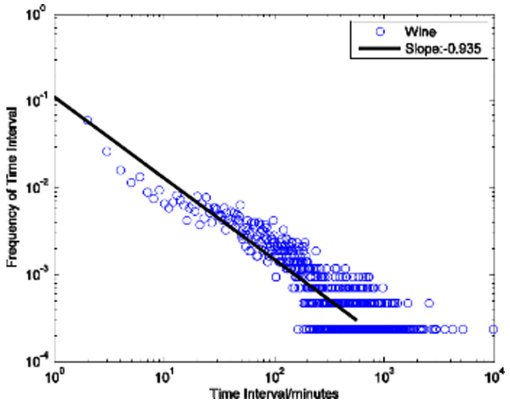


Fig. 6. Time interval distribution for adjacent comments regarding red wine.

This study calculated the intermittence and memory indexes for the time interval between two adjacent consumers' comments for three online products by using the previous formulas (3–1) and (3–2). The calculation results are shown in Table 5.

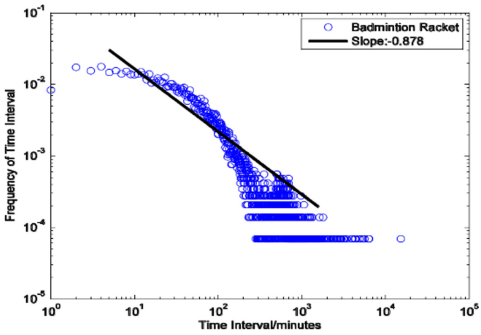


Fig. 7. Time interval distribution of adjacent comments for badminton racket.

Table 5. Burstness and memory for time intervals of consumer comments for three products

Products	Number of comments	Power exponent	Burstness	Memory
D90 digital camera	5460	−1.092	0.6087	2.206e−22
Badminton racket	13 733	−0.878	0.4044	2.755e−22
Red wine	4246	−0.935	0.3263	1.432e−22

Burstness index to comments released for the three products sold online were 0.6087, 0.4044, and 0.3263, respectively; this indicated that the time interval for the adjacent consumer comment released for each product had a strong intermittence. However, the memory values for the three products were almost zero. In other words, the adjacent comment behaviors for online products showed a strong intermittence and weak memory.

Based on the actual situation, the possible reason for this phenomenon is that, during the specific sales activities for the products or when the e-commerce platform was urging consumers to evaluate products that had not been evaluated, consumers tended to release comments within a short time period, and this resulted in a short interval between adjacent comments made within a certain time period, thus resulting in intense intermittence. On the e-commerce platform, each consumers' "buying and commenting"-related behavior was independent. Consumers' tend to comment on a product according to their own behavioral preferences or rules. The time point at which a consumer releases a comment has nothing to do with the time point at which the persons before him/her released their comments; therefore, the weak memory can be explained well.

4.4 Analysis Results for Temporal Characteristics of "Purchase and Comment" Behaviors

In order to verify the time interval distribution of consumers' "purchase-comments" macroscopically, the experimental data sets of six major categories were used for verification. The basic description of the "purchase-comment" time intervals for the consumers in the six product categories is shown in Table 6. Among them, the shortest time interval for "purchase-comment" was 0 (day)—that is, purchase and comment on the same

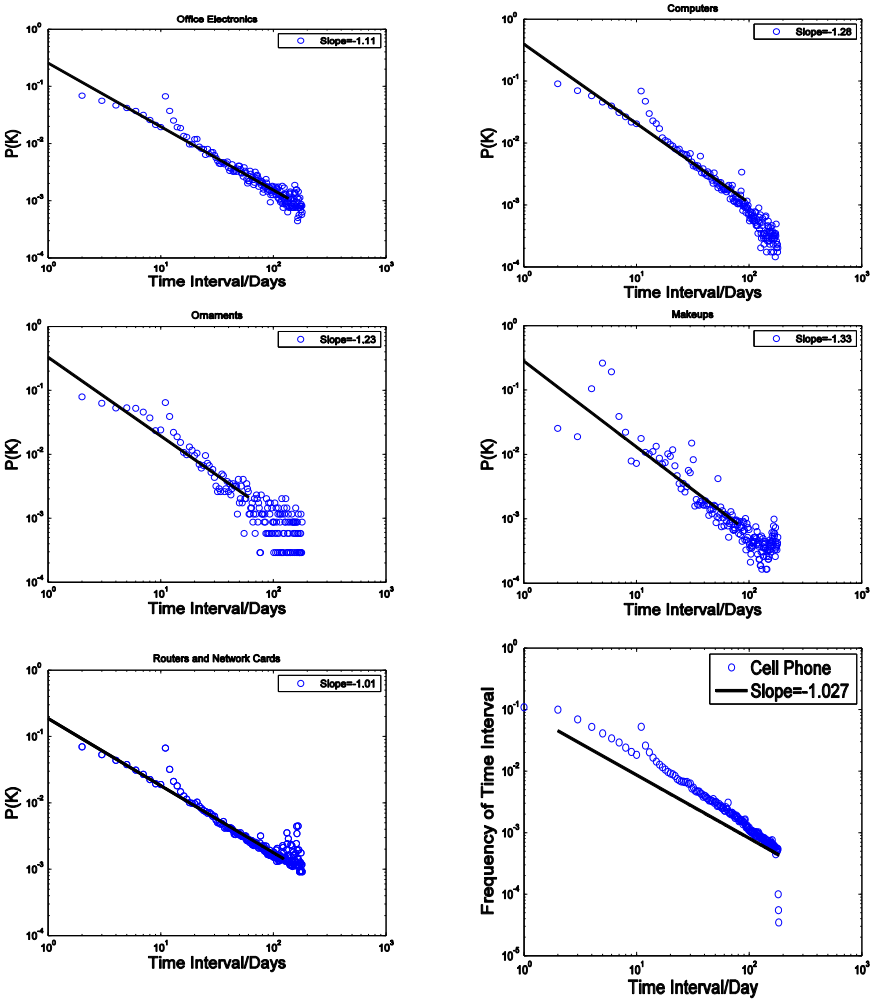


Fig. 8. Distribution of "purchase and comment" time intervals for six types of products.

day. The longest time interval was 183 days, and it was related to the platform's comment policy. The JD online mall only allowed consumers to post comments on products purchased within half a year.

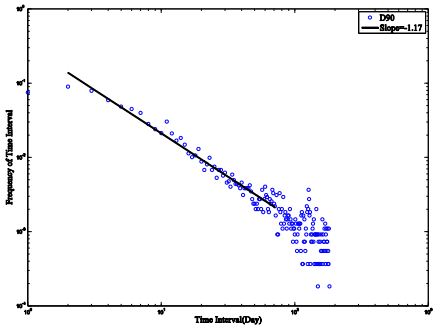
The double logarithmic scatter plot for the "purchase-comment" time interval distribution of the six product categories and the curve fitting of the main part are shown in Fig. 8. The power-law indexes of these six product categories were as follows: Office electronic products: −1.11; notebook computers: −1.28; decorations: −1.23; cosmetics: −1.33; routers and network cards: −1.01; and mobile phones: −1.027.

In order to verify the distribution rules followed by consumers within the "purchase-comment" time interval for a single product, in this experiment, the time intervals of the selected products for two types (experience-type and search-type) were selected. Two

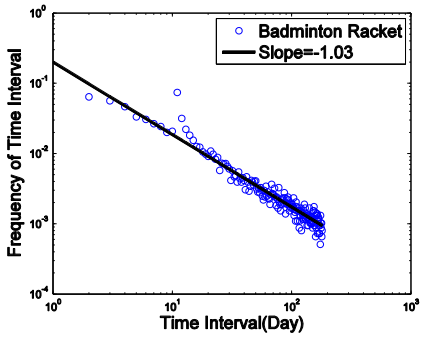
Table 6. Basic description of "purchase and comment" time intervals for six types of products

	Shortest interval	Longest interval	Standard deviation	Mean value
Office electronic products	0	180	28.94	19.83
Notebook computers	0	180	29.84	39.97
Routers and network cards	0	180	46.49	36.15
Mobile phones	0	183	35.87	24.58
Decorations	0	180	33.10	21.49
Cosmetics	0	182	28.32	16.18

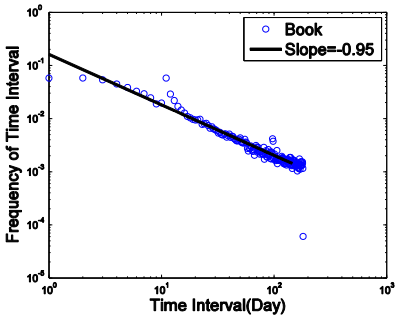
of the selected products were search-type products (D90 digital camera and badminton racket), and the other two were experience-type products (One Hundred Years of Solitude [book] and red wine). Their interval distribution is shown in Fig. 9. Their power-law indices were −1.17, −1.03, −0.95, and −1.14, respectively.



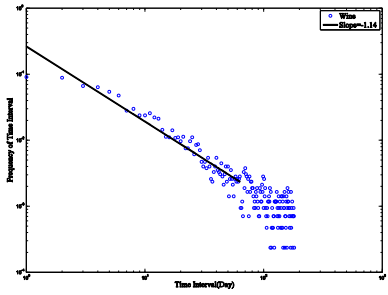
(a) "D90" Digital Camera



(b) Badminton Racket



(c) García Márquez's Book



(d) "Arabella" Wine

Fig. 9. Double logarithmic diagram of time interval distribution for four products.

5 Conclusion

This paper used the theories and methods of human behavior dynamics to describe the temporal characteristics of consumers' online purchasing and commenting behaviors. The main contents included the following. (1) By introducing the rules of online human behaviors, the form, cause, application scope, and fitting method of the power-law distribution theory utilized in this study were introduced; (2) three indexes describing the temporal rules for human behaviors were introduced (time interval, intermittence, and memory); (3) a method for dividing time interval series into stages was proposed; (4) based on online user comment behavior data, this empirical study examined three indicators used for describing the temporal rules underlying purchase behaviors, comment behaviors, and "purchase-comment" behaviors. The following conclusions were drawn.

5.1 Temporal Characteristics of Consumer Purchase Behavior

Based on the empirical results of the burstiness index and memory index tests, the temporal characteristics of consumers' purchasing behavior showed intensive purchasing behavior within a short time period, non-purchasing behavior within a long time period, and memory of purchasing behavior.

5.2 Temporal Characteristics of Consumer Comment Behavior

Based on the temporal characteristics of consumer comment behavior, the empirical results showed that the time interval between adjacent comment releases conformed to the power-law distribution and that the end of the distribution had clear "fat tail" characteristics. Furthermore, according to the results of the intermittence value and memory value tests, the adjacent comment behaviors for online products showed strong intermittence and weak memory.

5.3 Temporal Characteristics of Intervals for Consumer "Purchase-Comment" Behaviors

For the time interval characteristics of consumers' "purchase-comment" behaviors, it was found that the time interval between the two online human behaviors ("purchase-comment") followed the power-law distribution and that the power-law index distribution was in the range of -0.7 to -2.

This research on the dynamic mechanisms underlying consumer purchase and comment behavior also showed that it was an extension of the research on online human behavior dynamics in the e-commerce environment, thus providing new empirical evidences for research on online human behavior dynamics. On the other hand, mastering the temporal characteristics underlying consumer comments can provide good guidance practices for e-commerce enterprises. The main implication of the study findings was as follows. E-commerce enterprises can provide timely intervention based on the rules of delayed comments after consumers' purchase, encourage consumers to spend enough time and, at the same time, make real comments quickly, thus generating good publicity

for the selected product or service. In addition, such enterprises can allocate reasonable resources to customer service in order to better serve customers and improve customer satisfaction and loyalty.

References

1. Wang, T., Wang, K., Chen, H.: The impact of temporal distance on the increase of the perceived usefulness of online reviews: from the perspective of the attribution theory. J. Bus. Econ. **280**(2), 46–55 (2015)
2. Jiang, D., Huo, L., Song, H.: Rethinking behaviors and activities of base stations in mobile cellular networks based on big data analysis. IEEE Trans. Netw. Sci. Eng. **7**(1), 80–90 (2020)
3. Jiang, D., Wang, Y., Lv, Z., et al.: Big data analysis-based network behavior insight of cellular networks for industry 4.0 applications. IEEE Trans. Ind. Inform. **16**(2), 1310–1320 (2019)
4. Shen, C.C., Chiou, J.S.: The impact of perceived ease of use on Internet service adoption: the moderating effects of temporal distance and perceived risk. Comput. Hum. Behav. **26**(1), 42–50 (2010)
5. Shi, W., Gong, X., Zhang, Q., et al.: A comparative study on the first-time online reviews and appended online review. J. Manag. Sci. **29**(4), 45–58 (2016)
6. Min, Q., Qin, L., Zhang, K., Qing, F.M., Liang, Q., Keliang, Z.: Factors affecting the perceived usefulness of online reviews. Manag. Rev. **29**(10), 95–107 (2017)
7. Bauman, K., Liu, B., Tuzhilin, A.: Aspect based recommendations: recommending items with the most valuable aspects based on user reviews. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 717–725 (2017)
8. Ren, X., Lv, Y., Wang, K., et al.: Comparative document analysis for large text corpora. In: Proceedings of the Tenth ACM International Conference on Web Search and Data Mining, pp. 325–334 (2017)
9. Pang, J., Qiu, L.: Effect of online review chunking on product attitude: the moderating role of motivation to think. Int. J. Electron. Commer. **20**(3), 355–383 (2016)
10. McGuire, J.T., Kable, J.W.: Decision makers calibrate behavioral persistence on the basis of time-interval experience. Cognition **124**(2), 216–226 (2012)
11. Shen, W., Hu, Y.J., Rees, J.: Competing for attention: aAn empirical study of online reviewers' strategic behaviors. MIS Q. **39**(3), 683–696 (2015)
12. Gong, Y.P., Huang, K., Zhang, Q., Gu, H.P.: Study on the relationship between time and distance of new product forecast, consumer online comments and their purchase targets. R&D Manag. **04**, 36–44 (2015)
13. Jin, L., Hu, B., He, Y.: The recent versus the outdated: an experimental examination of the time-variant effects of online consumer reviews. J. Retail. **90**(4), 552–566 (2014)
14. Santhanam, M.S., Kantz, H.: Return interval distribution of extreme events and long-term memory. Phys. Rev. E **78**(5), 051113 (2008)
15. Barabasi, A.L.: The origin of bursts and heavy tails in human dynamics. Nature **435**(7039), 207–211 (2005)
16. Gonçalves, B., Ramasco, J.J.: Human dynamics revealed through web analytics. Phys. Rev. E **78**(2), 026123 (2008)
17. Zhou, T., Kiet, H.A.T., Kim, B.J., et al.: Role of activity in human dynamics. EPL (Europhys. Lett.) **82**(2), 28002 (2008)
18. Hu, H.B., Han, D.Y.: Empirical analysis of individual popularity and activity on an online music service system. Phys. A **387**(23), 5916–5921 (2008)
19. Chen, G., Han, X., Wang, B.: Multi-level scaling properties of instant-message communications. Phys. Procedia **3**(5), 1897–1905 (2010)

20. Henderson, T., Bhatti, S.: Modelling user behavior in networked games. In: Proceedings of the Ninth ACM International Conference on Multimedia, pp. 212–220. ACM (2001)
21. Zhao, F., Liu, J.H., Zha, Y.L., et al.: Human dynamics analysis of online collaborative writing. Acta Physica Sinica **60**(11), 118902–118902 (2011)
22. Sumi, R., Yasseri, T., Rung, A., et al.: Edit wars in Wikipedia. arXiv preprint arXiv:1107.3689 (2011)
23. Ren, J.J., Wang, N.X., Ge, S.L.: Information system user access law based on human behavior dynamics. Comput. Mod. **10**, 10–15 (2015)
24. Wu, Y., Zhou, C., Chen, M., et al.: Human comment dynamics in on-line social systems. Phys. A **389**(24), 5832–5837 (2010)
25. Wang, J., Gao, K., Li, G.: Empirical analysis of customer behaviors in Chinese e-commerce. J. Netw. **5**(10), 1177–1184 (2010)
26. Vázquez, A., Oliveira, J.G., Dezsö, Z., et al.: Modeling bursts and heavy tails in human dynamics. Phys. Rev. E **73**(3), 036127 (2006)
27. Chris, A.: The Long Tail, pp. 1–100. CITIC Press, China (2006)
28. Zipf, G.K.: Human Behavior and the Principle of Least Effort, pp. 1–573. Addisom-Wesley Press, Cambridge (1949)
29. Hu, H.B., Wang, L.: Brief research history of power-law distribution. Physics **34**(12), 889–896 (2005)
30. Newman, M.E.J.: Power laws, Pareto distributions and Zipf's law. Contemp. Phys. **46**(5), 323–351 (2005)
31. Zhang, J.Z.: Fractals, pp. 300–326. Tsinghua University Press, Beijing (1997)
32. Gutenberg, B., Richter, C.F.: Frequency of earthquakes in California. Bull. Seismol. Soc. Am. **34**, 185–188 (1944)
33. Roberts, D.C., Turcotte, D.L.: Fractality and self organized criticality of wars. Fractals **6**, 351–357 (1998)
34. Zanette, D.H., Manrubia, S.C.: Vertical transmission of culture and the distribution of family names. Phys. A **259**, 1–8 (2001)
35. Lotka, A.J.: The frequency distribution of scientific productivity. J. Wash. Acad. Sci. **16**, 317–323 (1926)
36. Crovella, M., Bestavros, A.: Self-similarity in world wide web traffic: evidence and possible causes. IEEE/ACM Trans. Netw. **5**(6), 835–846 (1997)
37. Goh, K.I., Barabási, A.L.: Burstiness and memory in complex systems. EPL (Europhys. Lett.) **81**(4), 48002 (2008)
38. Clauset, A., Shalizi, C.R., Newman, M.E.J.: Power-law distributions in empirical data. SIAM Rev. **51**(4), 661–703 (2009)
39. Jiang, D., Wang, Y., Lv, Z., et al.: An energy-efficient networking approach in cloud services for IIoT networks. IEEE J. Sel. Areas Commun. **38**(5), 928–941 (2020)
40. Jiang, D., Wang, W., Shi, L., et al.: A compressive sensing-based approach to end-to-end network traffic reconstruction. IEEE Trans. Netw. Sci. Eng. **7**(1), 507–519 (2020)
41. Jiang, D., Huo, L., Lv, Z., et al.: A joint multi-criteria utility-based network selection approach for vehicle-to-infrastructure networking. IEEE Trans. Intell. Transp. Syst. **19**(10), 3305–3319 (2018)
42. Jiang, D., Huo, L., Li, Y.: Fine-granularity inference and estimations to network traffic for SDN. PLoS ONE **3**(5), 1–23 (2018)
43. Jiang, D., Li, W., Lv, H.: An energy-efficient cooperative multicast routing in multi-hop wireless networks for smart medical applications. Neurocomputing **220**, 160–169 (2017)
44. Jiang, D., Zhang, P., Lv, Z., et al.: Energy-efficient multi-constraint routing algorithm with load balancing for smart city applications. IEEE Internet Things J. **3**(6), 1437–1447 (2016)