INFLUENCING FACTORS OF MICRO-BLOGGING MARKETING OF TRAVEL SERVICE PROVIDER - AN EMPIRICAL RESEARCH IN CHINA

Wanyou Hou, College of Management, Harbin Institute of Technology, Harbin, China, houwanyou@gmail.com

Xiangbin Yan, College of Management, Harbin Institute of Technology, Harbin, China, xbyan@hit.edu.cn

Jian Sun, College of Management, Harbin Institute of Technology, Harbin, China, s.jianjian@163.com

Abstract

Micro-blogging has emerged as a powerful social medium in spotlight since Twitter launched in the US in 2006, it has already changed the transmission mode of heterogeneous information, and even the interaction patterns both in human behavior and thoughts. More and more companies have launched micro-blogging marketing to communicate with customers, to achieve brand promotion and to form a gorgeous corporate image. For the purpose of providing companies with theoretical basis and reasonable proposals to effectively practice micro-blogging marketing, this paper conducted an empirical research on two largest travel service providers in China by collecting and analyzing the data from their Sina Weibo accounts, and took the text content of microbloggs into consideration to study influencing factors of micro-blogging marketing. We selected reposts as the main measure for micro-blogging marketing effect and studied the impact of conventional micro-blogging factors on reposts. We found that followers’ amount had no obvious impact on followers growth or attracting reposts, while the text content was an effective trigger, moreover, reposts presented conspicuous temporal characteristics.

Keywords: micro-blogging marketing; text analysis; Negative Binomial Regression analysis
1 INTRODUCTION

Micro-blogging has emerged as a powerful social media in spotlight since Twitter launched in the US in 2006. People can use micro-blogging to talk about any topics within the 140-character limit and can follow any other users they interested in to achieve real-time dissemination of information and interpersonal communication. Micro-blogging has already changed the transmission mode of heterogeneous information, and even the interaction patterns both in human behaviour and thoughts. In the world of micro-blogging, everyone can become an information publisher. It will be possible for a news item, an advertisement, or an individual experience to get rapid exposure, shared and spread. Because of these unique characteristics, micro-blog platform has become an effective marketing channel that many enterprises employed to better realize their business objectives. Previous researches have shown that the majority of corporate executives and marketers approved that marketing and promotion through micro-blogging would bring great values to their companies, and many companies were trying to step into the field of micro-blogging marketing to promote their businesses. For example, in a study about Fortune 500 companies, Rybalko and Seltzer (2010) pointed that Twitter, the representative of micro-blogging, had become an important marketing channel to communicate with their customers, which was an important aspect for a company to retain its customers.

Meanwhile, since its emerge, Micro-blogging has caused the attentions of many research scholars. Generally, the current researches on micro-blogging can be classified into two categories: the individual level researches, including analysis of micro-blogging followers, personal use of micro-blogging, the categories of user intentions, etc., and the organizational level researches, which focused on enterprises’ behaviours on micro-blogging (Xing Yan, 2011). Compared with the large number of researches on personal level, researches about micro-blogging behaviours on enterprise level, especially the commercial values of micro-blogging are relatively scarcer. Studies of the enterprise level researches can be roughly classified into two aspects. One is about the influence factors of microblogs, which constitute a large proportion, and the other is about the users’ behaviours on micro-blogging. When talking about the influencing factors of micro-blogging, most scholars had identical ideas that some of the elements of micro-blogging, such as indegree, retweets and user mentions, were very important for measuring the influence of a microblog. Amount of followers of the official account had been regarded as a significant indicator by many enterprises to measure whether or not their microblogs was successful. Meeyoung Cha et al. (2010) found that the follower’s amount might not be a good measure of an official account’s influence. As for the importance and strategies of micro-blogging marketing, there had been a lot of professionals proposing recommendations to enterprises about how to run their microblogs on the internet, but these suggestions usually did not rise to the theoretical level, which were qualitative views from individuals or some organizations.

When it comes to ways to maximize the commercial values of micro-blogging and achieve the marketing purposes of enterprises using micro-blogging, some researches had also put forward many constructive suggestions. For instance, in the study of movie sales, Huaxia Rui et al. (2011) found that chatting on micro-blogging about a movie had positive impact on increasing movie sales. They also indicated that the effect of publicity from influential users was far better than word of mouth of ordinary users. In addition, other researchers also found that unique style of a microblog (Yinbin Liu et al., 2012) which could catch customers’ eyes, such as micro-blogging massages with high quality, interesting or creative themed events, and active interactions with customers (Hengchong Zhang, 2011). In particular, many findings have shown that good micro-blogging contents played a significant role in expending the influence of micro-blogging. As the context of microblogs determines whether people will follow this microblog (Yinbin Liu et al., 2012), content is the most essential influence factor to the enterprise microblogs.

From the existing researches, we found that although the marketing values of micro-blogging had been acknowledged by many enterprises (Thomas Thompson, 2011) and a large number of enterprises had engaged in launching marketing campaigns, most enterprises still remained puzzled over how to make the best use of micro-blogging to connect with their customers (Denise Sevick Bortree et al., 2009). Although micro-blogging is a platform presenting tremendous commercial values, and
enterprises can make good use of it, however, target users’ needs are complex, and finding an effective marketing strategy in micro-blogging is not easy. Many enterprises have been exploring on the road that how to gain attention and expand influence through micro-blogging marketing, but with little achievement.

Based on existing researches, this paper took reposts as the main measure of the effect produced by micro-blogging marketing, and proposed hypotheses to check the impact of conventional micro-blogging factors on reposts. Then, we verified these hypotheses, with descriptive statistics and Negative Binomial Regression model, using diachronic data crawled from the official Sina Weibo accounts of Ctrip and eLong, moreover, the impact of the content text was studied in particular, which had rarely been involved in. Several interesting results were found in the end, which would provide reasonable proposals for enterprises to effectively conduct micro-blogging marketing. The remainder of this paper is organized as follows. Section 2 describes our research focus and hypotheses. Section 3 details the research framework. Section 4 presents research results of text clustering and hypotheses verification. Finally, section 5 concludes this paper.

2 FOCUS AND HYPOTHESES

The objective of micro-blogging marketing for an enterprise is to establish an internet platform to create and increase its influence. The enterprise can initiate or participate in a variety of activities on micro-blog platform, such as positively communicate with customers to establish and maintain the intimate relationships with customers, and achieve brand promotion to obtain higher visibility, which will do a great help to form a gorgeous corporate image. To better achieve this objective, the enterprise official micro-blogging account needs to try its best to attract more followers, and to make its posted content reposted by more users. In general, microblog users can be treated enterprises’ potential target customers. Enterprises pass their information to their followers through their microblogs to attract more followers and keep their sustained attentions. In addition, a follower itself is playing the role of disseminating information to its followers when reposting the enterprises’ information. In this information spreading process, with followers’ reposting in the manner of "snowball", a message can be transmitted thousands of times within a few minutes or even seconds.

In this paper, we took reposts, a common practice on Twitter is “retweeting”, or rebroadcasting someone else's messages to one's followers(Louis Yu et al., 2011), as the main measure of micro-blogging marketing effect, we studied impacts factors of micro-blogging, including followers’ amount, being a follower means that the user can receive automatically pushed micro-blogs posted by the followed users(Haewoon Kwak et al., 2010); content text, users can post messages containing text, pictures, videos and links(Louis Yu et al., 2011); posting time, indicating the time a micro-blog is posted. Figure 1 shows the interrelationship model of micro-blogging factors followed with corresponding hypotheses.

![Figure 1: interrelationship model of micro-blogging factors](image)

In general, accounts of microblog can be divided into four categories: organizations, celebrities, talents and grassroots. The accounts of organizations, celebrities and talents usually have a large number of followers and have a great influence to their followers, which means that attracting the attention of these accounts is equivalent to transmit the messages to the large amount of their followers since the followers of these accounts are likely to become direct followers of an enterprise
just because the person or the organization they worship and trust have paid attention to the enterprise. Compared to these accounts, the followers amount and influence of grassroots would be relatively underprivileged. While, since there are many grassroots users, and they usually can form like-minded crowds, by attracting their following, an enterprise may also have chances to get more followers from these grassroots because of the high degree of homogeneity. This leads to the following hypothesis.

**Hypothesis 1**: the more followers amount a micro-blog count keeps, the faster its followers amount growth would be, which would result in an exponential followers’ growth mode.

Several researchers in Palo Alto Research Centre studied the retweet on Twitter, and they believed that the average times of retweet was related to two factors, the followers’ number of the message poster and the twitter content. They just took simple statistical methods to discuss the impact of twitter content, one of which was counting the ratio of tweets including URL among all the tweets with high retweet times, and found that the ratio of URL included among all tweets was 21.1%, while the ration in tweets with high retweet times was 28.4%. Although these two digits seemed prosaic, it may indicate that the content had influence on times of retweets to some extent. This leads to the following hypotheses.

**Hypothesis 2.1**: the followers’ amount would be a factor affecting the reposts.

**Hypothesis 2.2**: the content of a microblog would have an impact on the reposts.

Currently, a number of studies have shown that the behaviour of Internet users have obvious temporal characteristics. For example, statistical analysis conducted by Baidu, the world’s largest Chinese search engine, showed that in the year 2012, 19 to 22 o’clock was the most active time period of a whole day, the Internet users were more likely to express, share and interact with others in this period. Similarly, for Internet users, there may exist obvious temporal characteristics in using microblogs which will have influences on reposting of microblog content. This leads to the following hypothesis.

**Hypothesis 3**: microblogs with different posting time period would perform distinctively in reposting times.

### 3 METHODOLOGY

To verify proposed hypotheses, we crawled the raw data of Ctrip and eLong accounts from Sina Weibo, and then, we clustered the content text of their microblogs to conduct the subsequent analysis, finally, two successive mathematical methods were selected to complete the hypotheses verification. The research framework is shown in figure 2.
3.1 Data collection

Sina Weibo, one of the earliest Chinese micro-blogging platform, began to provide its micro-blogging service in August 2009. The number of users of Sina Weibo had exceeded 350 million by the end of 2012, which had attracted a large number of companies register their official accounts to carry out micro-blogging marketing. We used two largest travel agency account, eLong and Ctrip, as objectives to collect the data on Sina Weibo platform. eLong established its Sina Weibo account in December 2009, while Ctrip established its Sina Weibo account in March 2010, both of them were verified users. The dataset used in this paper consisted of the whole data about those two accounts.

3.2 Data pre-processing

Since the text content of a microblog is one of the important drivers to earn attention and obtain reposts, this paper paid considerable endeavor on textual content analysis. Both Ctrip and eLong had posted more than 7,000 microblogs in a decentralized fashion in their official accounts (Zhiyuan Liu et al., 2012), covering rich content information. To identify the impact of micro-blogging content on attracting attention, it was the priori work to divide them into different categories according to their content. Firstly, we conducted word segmentation process to convert the micro-blogging text content into a model that computer can recognize. This paper selected ICTCLAS (Institute of Computing Technology, Chinese Lexical Analysis System) as a word segmentation tool, which is provided by Chinese Academy of Sciences and considered as the best word segmentation tool for Chinese characters. After this process, we found out the word composition and part of speech. Secondly, we excluded words with lower frequency and selected the remaining as dimension features of the vector space model to represent micro-blogging text. Finally, we converted the processed data into the file format supported by Weka (Waikato Environment for Knowledge Analysis), a java-based software for machine learning and data mining, and undertook the text clustering with the X-Means algorithm. X-Means algorithm is K-Means algorithm extended with efficient estimation of the number of clusters (Dan Pelleg et al., 2000) by an Improve-Structure part. In this part of the algorithm, the centres are attempted to be split in its region, and the decision between the children of each centre and itself is done by comparing the BIC-values of the two structures.

3.3 Hypotheses Verification

We employed two successive means to verify the proposed hypotheses. First, in order to discover intuitionistic findings, we analysed the original and processed data by using discription statistics, through which a few charts were provided to reflect prominent features. Then, we conducted the statistical verification with Negative Binomial Regression model. For the reason that the unconditional mean of the reposts, our dependent variable, was much lower than its variance, and Negative Binomial Regression model is suitable for overdispersed count variables (Noriszura Ismail et al., 2007). Negative Binomial Regression model is the generalization for Poisson Regression model, in which, the mean, \( \lambda_i \) is assumed to be constant or homogeneous within the classes. The heterogeneity within classes is allowed in negative binomial regression by assuming \( \lambda_i \) to be a Gamma with mean \( E(\lambda_i) = \mu_i \) and variance \( Var(\lambda_i) = \mu_i^2v_1^{-1} \), and \( Y_i|\lambda_i \) to be a Poisson with conditional mean \( E(Y_i|\lambda_i) = \lambda_i \). It can be shown that the marginal distribution of \( Y_i \) follows a Negative Binomial distribution with probability density function,

\[
\Pr(Y_i = y_i) = \int \Pr(Y_i = y_i|\lambda_i)f(\lambda_i)d\lambda_i = \frac{\Gamma(y_i + \nu_i)}{\Gamma(y_i + 1)\Gamma(\nu_i)} \left( \frac{\nu_i}{\nu_i + \mu_i} \right)^\nu_i \left( \frac{\mu_i}{\nu_i + \mu_i} \right)^{y_i} (1)
\]

where the mean is \( E(Y_i) = \mu_i \), and the variance is \( Var(Y_i) = \mu_i + \mu_i^2v_1^{-1} \).

We used SPSS to test this model with the maximum likelihood estimates as the default estimating method, the likelihood equation is written as,
\[ \ell(\beta,a) = \sum_i \left( \sum_{r=1}^{y_i-1} \log(1+a r) \right) - y_i \log(1+a) - \log(1+a) - (y_i + a^{-1}) \log(1+a) \] 

The maximum likelihood estimates is conducted by maximizing \( \ell(\beta,a) \) with respect to \( \beta \) and \( a \), and the related equations are,

\[ \frac{\partial \ell(\beta,a)}{\partial \beta_j} = \sum_j (y_j - \mu_j) x_{ij} \frac{1}{1+\alpha \mu_i} = 0, \quad j=1,2,\ldots, p, \] 

and,

\[ \frac{\partial \ell(\beta,a)}{\partial a} = \sum_i \left( \sum_{r=1}^{y_i-1} \frac{r}{(1+ar)} \right) + a^{-2} \log(1+a) - \frac{(y_i + a^{-1}) \mu_i}{(1+\alpha \mu_i)} = 0. \] 

## 4 EMPIRICAL RESULTS AND RESEARCH FINDINGS

### 4.1 Data

Each Sina Weibo user keeps a brief profile, including the user’s ID, the full name, the number of micro-blogging messages produced, the number of users followed and the number of followers. By the end of October 2012, Ctrip had produced 9,095 micro-blogging messages and gained 1,391,023 followers. eLong, as a comparison, had produced 7,611 micro-blogging messages and gained 1,407,877 followers. We crawled all micro-blogging messages of these two accounts, including the time it was posted, the repost times and number of comments. We also got their daily number of followers. After data pre-processing, we got 4 categories of micro-blogging text content both for Ctrip and eLong. They had topics in common, however, some of the topics were in different types of expressions.

For the 9,095 microblogs of Ctrip, the 4 clusters were as follows. Cluster 0: Presentation of tourist attractions, hotels, local cuisine and custom, mostly talking about information for travelling, and there are discussions of timely events as well; Cluster 1: Awards-winning interactive content, launching activities to attract followers, reposts and reviews; Cluster 2: Sharing of traveling experience and UGC, Ctrip always posts manuscripts written by professional editors or content like photos, dairies generated by common users; Cluster 3: Special offers broadcast, attracting followers to visit Ctrip official website to book discount tickets and hotels.

For the 7,611 microblogs of eLong, the 4 clusters were as follows. Cluster 0: The topic #Awards-winning reposts #, often claimed followers to RT @ one or more friends; Cluster 1: Special offers, knowledge, recommendation and guides about tour, mainly talking about travel-related information, which is usually worth reading for followers in need; Cluster 2: Introductory text, introduce interviews and activities, especially include a classic topic of eLong # eLong Daily Scene #, also used to make introduction; Cluster 3: Wonderful phrases, regularly posted some phrases, most of which were said by celebrities or selected from classic movie lines. More numeric and textual information are given in table 1 and table 2 accordingly.

<table>
<thead>
<tr>
<th>Clusters</th>
<th>Quantity (Proportion)</th>
<th>Typical text content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presentation of tourist attractions, hotels, local cuisine and custom</td>
<td>5,540(60.91%)</td>
<td>里约热内卢的天梯教堂，其圆锥形造型不同于欧洲教堂古色典雅的风格，教堂四周色彩斑斓的玻璃窗从地面延伸到75米高的屋顶，教堂内部光线穿过彩绘玻璃是很美的视觉享受，当阳光照在玻璃带上，人站在圆锥体内观赏会觉得美丽异常，非常独特。（Ladder Church in Rio de Janeiro, its conical shape is different from the ancient and elegant style of European church, the church was surrounded by colourful glass windows extending from the ground to the 75-meter-</td>
</tr>
</tbody>
</table>
high roof, the interior of the church to light through the stained glass is beautiful visual enjoyment, when the sun according to the glass ribbon, the standing conical body viewing will feel beautiful and unusual, very unique.

Awards-winning interactive contend 1,566(17.22%) #关注有礼#微博粉粉们看过来！想赢北京豪华酒店，现在仅需一个关注！3月22日-3月27日，关注@携程旅行网，就有机会赢取价值5400元北京兴基铂尔曼饭店行政大床间两晚人住房券（门市价：2700元/晚）。老北京，新玩法，Two night in Beijing！你留下关注，我送出酒店！(#follows to win# Look over here. Want to live freely in a luxury hotel in Beijing, and now only need to follow! March 22-March 27, 2011, follow @Ctrip, you will have the chance to win Beijing Pullman Hotel for two nights in the Executive King Room into a housing voucher (Retail price: 2700 yuan/night)! Old Beijing, new play, Two nights in Beijing! You follow me; I give out of the hotel!)

Sharing of traveling experience and UGC 1,101(12.11%) #象鼻山#或许提起桂林，所有的人都会在脑海中想到这座山，因为太过著名，我就不多做介绍了，选张网友的照片分享给大家。(图片来自：携程网友开心的人生) (#the Elephant Trunk Hill# Perhaps when talking about Guilin, this hill will come to everyone’s mind since it is so famous. Instead of introducing, I will share a photo shot by a Ctrip friend. (Picture from: Ctrip friends’ happy life))

Special offers broadcast 888(9.76%) #秒杀机票# 携程“88抢票”今天推出河北航空专场 - 全场88张机票全部只要88元。今天晚上8:08准时开抢，只此一场，不要错过哦。http://t.cn/hr2wk (#Grab flight tickets# Ctrip today launch Hebei Airlines special - 88 tickets all sold at 88 yuan. Tonight 8:08 to open on time to grab do not miss it. http://t.cn/hr2wk)

Table 1 Clusters of Ctrip micro-blogging text content

<table>
<thead>
<tr>
<th>Clusters</th>
<th>Quantity (Proportion)</th>
<th>Typical text content</th>
</tr>
</thead>
<tbody>
<tr>
<td>The topic #Awards-winning reposts #</td>
<td>423(5.56%)</td>
<td>#五星酒店免费住# 深圳瑞吉酒店，坐落于蔡屋围金融中心内。参与方式：关注@艺龙旅行网并转发此微博。活动结束后将有2名幸运儿，分别获豪华城景房1晚免费住宿。订酒店，用艺龙！<a href="http://t.cn/zlsaML">http://t.cn/zlsaML</a> (#Star hotel for free# St. Regis is located in Caiweiwu financial centre, Shenzhen. A participatory manner: follow @eLong and forward this microblog, at the end of this activity, there will be two winners, awarded one night living for free. Book hotels, with eLong! <a href="http://t.cn/zlsaML">http://t.cn/zlsaML</a>)</td>
</tr>
<tr>
<td>Special offers, knowledge, recommendation and raiders about tour</td>
<td>5,912(77.68%)</td>
<td>【如何应对高原反应终极版】如果能够按照以下的方法去做，高原反应就不会成为困扰高原之行的绊脚石。给要去西藏的同学参考。(Ultimate Edition for how to deal with altitude sickness) altitude sickness will not become a stumbling block for troubled plateau trip if able to do it in accordance with the following method. classmates to go to Tibet for reference.)</td>
</tr>
<tr>
<td>Introductory text Introduce interviews and other activities The topic #Presentation of Tourist Attractions#</td>
<td>923(12.12%)</td>
<td>#艺龙每日一景# 济州岛—东方夏威夷 济州岛位于朝鲜半岛南端，是韩国最大的岛屿。这里气候温和、风光秀美，有“东方夏威夷”、“远东巴厘岛”的美誉。作为韩国最知名的观光度假胜地，曾有《冬季恋歌》、《人鱼小姐》等多部热门韩剧集在此取景。(#eLong Daily Scene# Jeju Island - Oriental Hawaii Jeju Island is located in the southern portion of the Korean Peninsula, South Korea's largest island. Mild climate, beautiful scenery, with the reputation of &quot;Oriental Hawaii&quot; and “Far East Bali”. As the most well-known tourist resort in South Korea, has been shot by &quot;Winter Sonata&quot;, &quot;Miss Mermaid&quot; and many other popular Korean dramas)</td>
</tr>
</tbody>
</table>
4.2 Descriptive Verification and Findings

First, we observed the changes of the amounts of Ctrip and eLong’s followers from February 29, 2012 to October 31, 2012, which was shown in figure 3. We found that the number of the two companies’ followers were generally in a growing trend in this period, but their growth rates and growth modes were quite different.

![Growth trends of followers](image)

**Figure 3: Growth trends of followers**

From figure 3, we can observe that Ctrip, compared with eLong, had a much greater growth rate of the amount of its followers. On February 29, 2012, the initial followers’ amount of Ctrip was 584,122, which was only half of eLong with 1,179,211 followers. However, on October 31, 2012, the terminal date we had observed, the amount of Ctrip’s followers grew to 1,391,023, very close to the amount of eLong’s followers, which climbed to 1,407,877. In the observation period, the growth rate of the amount of Ctrip’s followers was much higher, with an increase of 806,901 followers, 138% of the initial amount; while the amount of eLong’s followers only increased by 19%, with 228,656 new followers. Meanwhile, considered from the point of view of the growth mode, the followers’ quantity of Ctrip, in generally speaking, had grown exponentially; while the followers quantity of eLong had grown in a stepwise mode. Before July 31, 2012, the growth trend of Ctrip’s followers was obvious. The regression analysis showed that it fitted with the exponential model well. Although there was a slightly drop of the growth speed from August 2, 2012, the followers amount continued growing exponentially. Until late September, the Ctrip followers amount growth trend changed obviously, showing a two-stage ladder growth: after reaching 1,260,580 on September 30, the followers’ quantity grew slower than the previous period; while from October 10 to 12, the followers quantity went up sharply again, with a growth of 7,250.25 per day on average. In contrast, the amount of eLong’s followers had increased in a phased manner. The followers’ amount of eLong went through a gentle period, even with a slow decline sometimes, after every significant growth. The most representative changes occurred in the period from May to July. After grown to 1,277,038, the number of eLong’s followers had slowly dropped to 1,251,320 on June 13, a total decline of 25,718 followers, and the average daily decline was 695. Then, from June 25 to July 1, the amount of its followers appeared a jump growth, added 87,427 followers within a week, with an average daily growth of 12,490 followers.

**Finding 1:** the exponential followers’ growth mode only exists in specific stage of individual account; followers’ number is not a decisive factor to attract more followers.
Then, we studied the impact of followers. As mentioned above, no matter what mode they performed in, the followers’ amount of these two companies were always in growth, if assumption 2 is true, the reposts would also present obvious gradual upward trend. Figure 4(a) is the reposts situation of Ctrip from February 29, 2012 to October 31, 2012, compared with followers in the same period in figure 1. We found that on June 11 the reposts amount was 71,706, much higher than that of most days with more followers. In addition, September 25 and October 9 were two start point of rapid growth for followers amount, but the reposts quantity were also much larger than the one after the substantial growth of followers. Figure 4(b) is the reposts of eLong in a longer period. It obviously shows that the average reposts amount in middle stage was far greater than that in the later stage.

**Finding 2.1:** the followers’ number plays an important role in the exposure of a microblog, which may help increase the possibility for Netizen to click the repost button; but when the followers’ amount grows to a certain extent, it is not likely to have direct bearing on reposting any longer.

![Figure 4(a): the reposts over time of Ctrip](image)

We have proven that followers’ number is not the reason of the text reposts. Now we pin our hopes on the text content of microblogs. As mentioned in section 2, both Ctrip and eLong had 4 categories of text content, and the average reposts for their respective categories can be seen in Figure 5(a) and figure 5(b). We can know that the average reposts of different content clusters were imbalanced. Both Ctrip and eLong have a particularly prominent cluster, whose reposts times was much more greater than the other clusters respectively. Moreover, the two clusters pointed to a common text content type - # awards-winning interactive content#.
Finding 2.2: we believe that the text content of a microblog is the main trigger to lead to reposts, and the content with bonus does best in mobilizing enthusiasm.

Cluster 0: Presentation of tourist attractions, hotels, local cuisine and custom; Cluster 1: Awards-winning interactive contents; Cluster 2: Sharing of occupational traveling experience and UGC; Cluster 3: Special offers broadcast

Figure 5(a): Text Content Chart of Ctrip, the average reposts changes and rollup for three years

Cluster 0: The topic #Awards-winning reposts#; Cluster 1: Special offers, knowledge, recommendation and raiders about tour; Cluster 2: Introductory text of interviews and other activities, the topic #Presentation of Tourist Attractions#; Cluster 3: Wonderful phrases

Figure 5(b): Text Content Chart of eLong, the average reposts changes and rollup for three years

As mentioned above, both Ctrip and eLong had the content cluster of #awards-winning reposts#, and this category showed a strong advantage in achieving reposts. We compared the reposts situation of this category, and in this way, the impact caused by different content could be excluded. Moreover, in order to form intuitive feelings, we removed the dimensional impact, and showed the result in figure 6 and figure 7. We found that both Ctrip and eLong could receive more reposts on two periods of time: 13:00-15:00, the noon nap to the beginning of afternoon work, and 19:00-21:00, after dinner to bedtime. Considering separately, the first three time for Ctrip to be pampered followed by 14:00-
15:00, 15:00-16:00 and 16:00-17:00; while, for along, the time went to 13:00-14:00, 11:00-12:00 and 19:00-20:00.

Finding 3: the reposts amount shows temporal characteristics, however, for the specific account the characteristic would be different.

![Average Reposts of Every Two Hours](image1)

**Figure 6: Average Reposts of Every Two Hours**

![Average Reposts per Hour](image2)

**Figure 7: Average Reposts per Hour**

4.3 Statistical Verification

We set reposts as dependent variable, and hourSeg, contentClu as categorical predictors. Thereinto, the variable contentClu was a four-level nominal variable with the same identifier as the clustering results mentioned above, and the variable hourSeg was a four-level nominal variable indicating the hour segment of posting time of microblogs. Each level represented different 6 hours of a whole day - level 0 stood for 7:00 to 9:00 and 21:00 to 23:00; level 1 stood for 1:00 to 7:00; level 2 stood for 9:00 to 13:00 and 17:00 to 19:00; level 3 stood for 13:00 to 17:00 and 19:00 to 21:00. Table 3(a) and Table 3(b) show the tests of model effects. We can see that both hourSeg and contentClu are statistically significant predictors of reposts.
### Table 3(a) Tests of Model Effects, Ctrip

<table>
<thead>
<tr>
<th>Source</th>
<th>Type</th>
<th>Wald Chi-Square</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>III</td>
<td>150.234</td>
<td>1</td>
<td>.000</td>
</tr>
<tr>
<td>hourSeg</td>
<td></td>
<td>66.617</td>
<td>3</td>
<td>.000</td>
</tr>
<tr>
<td>contentClu</td>
<td></td>
<td>582.573</td>
<td>3</td>
<td>.000</td>
</tr>
</tbody>
</table>

Dependent Variable: reposts
Model: (Intercept), hourSeg, contentClu

### Table 3(b) Tests of Model Effects, eLong

<table>
<thead>
<tr>
<th>Source</th>
<th>Type</th>
<th>Wald Chi-Square</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>III</td>
<td>94.900</td>
<td>1</td>
<td>.000</td>
</tr>
<tr>
<td>hourSeg</td>
<td></td>
<td>229.829</td>
<td>3</td>
<td>.000</td>
</tr>
<tr>
<td>contentClu</td>
<td></td>
<td>165.906</td>
<td>3</td>
<td>.000</td>
</tr>
</tbody>
</table>

Dependent Variable: reposts
Model: (Intercept), hourSeg, contentClu

The table 4(a) and table 4(b) are parameter estimates containing the negative binomial regression coefficients for each of predictor variables along with their standard errors, 95% confidence intervals for the coefficients, Wald chi-square values and p-values. The dummy variables for the variable hourSeg are statistically significant. For the dataset of Ctrip, compared to level 3 of hourSeg, the expected log count of level 1 decreases by 11.57. Moreover, for the dataset of eLong, compared to level 3 of hourSeg, the expected log count of level 0, 1 and 2 respectively decreases by 0.118, 0.667 and 0.457, presenting the same results enfolded in Fig. 5 and Fig. 6, which indicated the significant impacts of posting time to the reposts. The level 1 of contentClu for Ctrip and the level 0 of contentClu for eLong both have much higher log count than any other levels, together with the investigation in last section, verified the obvious effect of micro-blogging content to the reposts.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>B</th>
<th>Std.Error</th>
<th>95% Wald Confidence Interval</th>
<th>Hypothesis Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>7.832</td>
<td>.1865</td>
<td>7.466</td>
<td>1763.847</td>
</tr>
<tr>
<td>[hourSeg=0]</td>
<td>6.181</td>
<td>2.3621</td>
<td>1.552 10.811</td>
<td>6.849</td>
</tr>
<tr>
<td>[hourSeg=2]</td>
<td>-6.49</td>
<td>.1123</td>
<td>-8.69 -4.29</td>
<td>33.378</td>
</tr>
<tr>
<td>[hourSeg=3]</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[contentClu=0]</td>
<td>.136</td>
<td>.1885</td>
<td>-.233 .506</td>
<td>.523</td>
</tr>
<tr>
<td>[contentClu=1]</td>
<td>8.251</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[contentClu=3]</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(scale)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Negative binomial)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dependent Variable: reposts
Model: (Inception), hourSeg, contentClu

a. Set to zero because this parameter is redundant
b. Fixed at the displayed value

### Table 4(a) Parameter Estimates, Ctrip
Table 4(b) Parameter Estimates, eLong

<table>
<thead>
<tr>
<th>Parameter</th>
<th>B</th>
<th>Std.Error</th>
<th>95% Wald Confidence Interval</th>
<th>Hypothesis Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>6.806</td>
<td>.0609</td>
<td>6.686, 6.925</td>
<td>12472.944</td>
</tr>
<tr>
<td>[hourSeg=0]</td>
<td>-.118</td>
<td>.0467</td>
<td>-.210, -.027</td>
<td>66.432</td>
</tr>
<tr>
<td>[hourSeg=1]</td>
<td>-.667</td>
<td>.2201</td>
<td>-1.009, -.236</td>
<td>9.190</td>
</tr>
<tr>
<td>[hourSeg=2]</td>
<td>-.457</td>
<td>.0320</td>
<td>-.519, -.394</td>
<td>203.542</td>
</tr>
<tr>
<td>[hourSeg=3]</td>
<td>0(^a)</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>[contentClu=0]</td>
<td>11.318</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>[contentClu=1]</td>
<td>.144</td>
<td>.0586</td>
<td>.029, .259</td>
<td>6.045</td>
</tr>
<tr>
<td>[contentClu=2]</td>
<td>4.543</td>
<td>.3598</td>
<td>3.829, 5.239</td>
<td>158.826</td>
</tr>
<tr>
<td>[contentClu=3]</td>
<td>0(^b)</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>(scale)</td>
<td>1(^b)</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>(Negative binomial)</td>
<td>1</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

Dependent Variable: reposts
Model: (Inception), hourSeg, contentClu
   a.Set to zero because this parameter is redundant
   b.Fixed at the displayed value

5 CONCLUSION AND DISCUSSION

Ctrip and eLong are two largest travel service providers in China, and they both have launched their micro-blogging marketing and have got great achievement with their Sina Weibo accounts. In this paper, we studied the role of followers’ amount and micro-blogging content in attracting reposts, and found several interesting results. Firstly, the followers’ amount is not as valuable as it seems to be, it does not perform well in promoting followers increase or attracting reposts. Secondly, posting microblogs frequently is not a sensible way to obtain a large amount of reposts. In the time period observed, the microblogs of ctrip were nearly 1,500 more than the quantity of eLong, but the cumulative reposts were 4,717,366 times less. It is the quality, not the quantity, which attracts the attention and affection, based on this fact, we analysed the content-clustered data and found that the text content of a microblog was the main trigger to lead to reposts, and the content with bonus did best in mobilizing enthusiasm. The enterprise micro-blogging accounts may post microblogs with exciting temptation to expand the influence dramatically. Finally, each enterprise account has its own fans crowd with specific active time to read microblogs, in order to make larger exposure and gain more attention, the enterprise’s official accounts should select the correct time to post microblogs due to the temporal characteristics of reposts.

6 ACKNOWLEDGEMENTS

This work is partly supported by the National Natural Science Foundation of PRC (No. 71171067, 70890082) and Postdoctoral Science-research Developmental Foundation of Heilongjiang province (No. LBH-Q11114).
References


