The Analysis of the Influencing Factors on the Problems of Bike-Sharing System in China

SHIANGHAU WU1, XUXIUNAN LEI1,

1School of Business, Macau University of Science and Technology (e-mail: shwu@must.edu.mo)
2School of Business, Macau University of Science and Technology (e-mail: 1709853bj20002@student.must.edu.mo)

Corresponding author: Shianghau Wu (e-mail: shwu@must.edu.mo).

This work was supported in part by the Faculty Research Grant of Macau University of Science and Technology

ABSTRACT Bike-sharing has experienced rapid growth in China since 2016. Notwithstanding the fast expansion, or possibly because it grew too fast, some bike-sharing companies have experienced setbacks and failure in 2018. This paper aims to close the gap due to the insufficient analysis on the influencing factors of the bike-sharing system problems in China. The study first used text mining of bikes-sharing related Chinese news reports and social web discussion boards. Subsequently, the study used association rule mining to explore the relationship between keywords generated by text mining of the original data sources. The results of news reports mined keywords show that problems with shared bikes deposits were closely related to the complaints of customers. In the context of social web discussion boards, the keywords relations implies that the users concern about the possible collapse of the bike-sharing companies and related management and economy issues. The information existed before the bike-sharing companies financial failure news took place. Our results also show that the association rule mining relationship of major keywords in news reports and social media can be an early warning sign of the financial failure of sharing bikes companies.

INDEX TERMS bike-sharing, text mining, association rule mining

I. INTRODUCTION

BIKE-SHARING is one of the major applications within the sharing economy. Bike-sharing meets the requirements of the sustainability and addresses ecological concerns. Recent studies of bike-sharing discuss the implementation of the system per se, its technology and regulations. Bike-sharing in China stemmed from an ecological concern and congested traffic conditions [1].

In the past three years, the fourth-generation of dockless bike-sharing programs (DBSPs), such as Mobike and OFO, have experienced a rapid expansion in China [2]. The rise of dockless bike-sharing seems, however, to be ephemeral. The lack of supply control and unregulated DBSPs resulted in a problem-ridden sharing bikes market in China. The possible failure of bike-sharing ventures in China became hot topics in the public opinion.

In order to explore the reasons behind the failures of bike-sharing in China, the study used the text mining followed and association rule mining approach to analyse the relationship between keywords from public opinion data sourced in Chinese mass media news reports and social media discussion boards.

The study is arranged as follows: First, the study made a literature review about the development of the bike-sharing system and discussed about the insufficient parts of current literature. Second, the study introduces the methodology, followed by the research results. Third, the study analysed the research results and concluded with recommendations on the current problems of bike-sharing in China.

II. LITERATURE REVIEW

A. SHARING BIKES

According to De Maio (2009), the bike-sharing program has the four generations listed as follows [3]:

- First Generation: The first generation of bike-sharing system began in Amsterdam, Netherlands. The first generation of bike-sharing system was lack of the normal
lock sharing bikes system in China includes the following two
good solution to the urban environmental problems [6]. The
generation of sharing bikes in China was regarded as the
100 cities in four countries around the world [5]. The fourth
exported overseas to provide one billion travel services to
sharing company Mobike was founded. Intelligent dockless
rapidly in recent years. The dockless OfO bikes system was
The development of bike-sharing system in China has grown
2
VOLUME 4, 2016
Commercial Bike-sharing System: As for the commer-
Public Sharing Bikes Systems [7]: In some cities like
• Third Generation: The third generation of bike-sharing
system was launched in the U.K in 1995. The third
generation of sharing bikes were automated controlled.
It utilized the CCTV surveillance system to prevent
sharing bikes from being stolen. The third generation
of sharing bikes were regarded as the successful case
in the bike sharing history and numerous bike-sharing
programs were launched around the world.
• Fourth Generation: The fourth generation of the bike-
sharing system combined the concepts of "bikes" and
"taxi". The "BIXI" system was launched in Montreal,
Canada in 2009. The technological progress of the
fourth generation of sharing bikes was the removable
parking stations. It enabled the docking to move to
different stations easily and adapted to user patterns
and demands.

The growing sharing bikes has aroused some reflections. Hall
and Ince (2017) argue that the need to verify whether the
booming sharing bikes imply the post-capitalism reinvent
itself and these formulations intersect and overlap with other
forms of valuing, exchanging, producing and consuming
[11], while McLaren and Agyeman (2015) lauded the bike-
sharing system as the sustained transformation shift toward a
pro-social sharing paradigm [12].

B. PREVIOUS STUDIES OF SHARING BIKES IN CHINA

Previous studies of Sharing bikes system in China relate to
the following issues:

• Investigation on Sharing bikes Users’ Trust and Perceived Value: Ma et. al. (2018) investigates the relationship between sharing bikes users’ trust, perceived value and well-being [13]. Wu et. al. (2018) utilizes the PLS-SEM model to analyze the sharing bikes users’ customer loyalty evaluation model [14]. Yin and Qian (2018) conducts the survey on 755 public bicycle-sharing scheme (PBSS) users and finds the significance of collectivism, man–nature orientation, materialism, and face-consciousness as key determinants of the adoption of PBSS [15]. Hazen et. al. (2015) applies to the Technology Acceptance Model to analyze the 421 pub-
lic sharing bikes participants in Beijing and finds that reasonable value of use positively affects users’ intentions to adopt public bicycle sharing systems [16].

- Discussions on the Bike-sharing Problems in China: Chang et al. (2018) points out the faulty bike-sharing recycling (FBSR) problem on the dockless sharing bikes system in China and posits a framework based on the optimization method to solve the FBSR problem in order to minimize the total recycling costs by taking the route optimization and loading capacity ratio as constraints [17]. Gao et al. (2018) mentions that the problem of “difficulty picking and parking” is the most important problem of public bicycle system and affects the user’s confidence of usage. On the one hand, some of the leased-point bicycles were quickly borrowed, resulting in the increase of other users’ waiting time. They also found that the lock-up parking spaces at some rental points were soon parked by bicycles, resulting in the returning user waiting for an empty space or finding another empty lease to return [18]. Gu et al. (2019) indicates that dockless bike-sharing in China has experienced “free growth”, “regulated” and “limited” phases in a short time. They find the central government initially held a “neutral-positive” policy towards this new system, the rapid expansion of dockless sharing bikes soon exceeded cities’ limits and resulted in local government policies changing from “neutral-positive” to “neutral-negative”, and from August 2017, powerful limiting regulations have been implemented [19].

The study found that previous literature of sharing bikes in China focus less on the problems of its operations. In late October of 2018, OFO bikes were reported in debt-ridden crisis and some rumours like bankruptcy have mounted [20]. However, previous literature about the bike-sharing system problem in China seems insufficient about the analysis of the current problems of bike-sharing system. Besides, as far as the research methodology is concerned, current researches on this topic involve in the consumer demand analysis and optimization analysis. The data mining method utilized in this study would be the new attempt in this research field. The study therefore generalized our research problems: (1) What are the influencing factors of the management problems of the bikes-sharing system in China? (2) When does the management problems of bikes-sharing system begin to take shape?

III. METHODOLOGY

The study attempted to utilize the text mining method to obtain the main keywords of sharing bikes related Chinese news reports and the text contents on the social media. We also calculated the frequencies of the main keywords in each text data. The study utilized the main keywords frequency data to construct the model in order to find the relationship among the keywords in order to comprehend the problems of the bike-sharing system in China in different time frames.

![Flowchart of the Research Process](image)

**FIGURE 1. Flowchart of the Research Process**

**A. RESEARCH PROCESS**

In the beginning, the study utilized the text mining method to obtain the major keywords of sharing bikes related Chinese news reports and Chinese posts on the social media discussion board. After obtaining the major keywords, the study used the association rule mining method to explore the relationship among keywords. The research process flowchart was shown on the Figure 1. According to Figure 1, the study analyzed two data streams, including news reports data and social media data. We obtained major keywords by text mining process and utilized the TF-IDF value of each keyword to do the association rule mining in order to comprehend the relations of keywords in different time.

**B. TEXT MINING**

Text mining is one of the data mining techniques in the knowledge discovery process. It is applied in the information extraction in the unstructured data with the Natural Language Processing (NLP) [21]. In these years, the text mining techniques are improving and most research and development efforts have centered on data mining efforts using structured data [22]. The main Text Mining applications are most often used in the following sectors, such as the publishing and media, telecommunications, energy, services industries, IT industry, financial industry, political science, legal documents, pharmaceutical industry and healthcare [23]. The Text mining process includes two steps: text preprocessing and knowledge extraction. Text preprocessing step converts unstructured data into a document-term matrix, and knowledge extraction step involves data mining. The preprocessing step includes tokenization, stop words removal, stemming and other processes followed by the formulation of a document-term matrix [24]. The document-term matrix is formed ac-
Association rule mining is one of the important data mining techniques which aims at exploring the correlations, frequent patterns and causal structures of the transaction database. It searches for the all rules in the database that satisfy some minimum support and minimum confidence constraints. The target of association rules mining is not predetermined [29]. The study uses the association rule mining to explore the relationship among keywords. It is because the association rules mining can determine the major keywords relations. The association mining is based on two basic measures: supports and confidence. Supports are calculated as follows [30]:

\[
\text{Support}(XY) = \frac{\text{Support Count of XY}}{\text{Total Number of Transactions in D}} \quad (1)
\]

In equation(1), X means antecedent and Y means consequent. D is a database with transactions T. Support count of XY means the number of coincidence occurrence of X and Y. "Support" of an item is a statistical significance of an association rule. The association rule mining also used the confidence measure. "Confidence" is the percentage of the number of transactions containing \(X \cup Y\) to the total number of transactions including X. It means that when the percentage exceeds the threshold of confidence, the relationship \(X \Rightarrow Y\) can be generated. The confidence measure can be listed as equation 2:

\[
\text{Confidence}(X|Y) = \frac{\text{support}(XY)}{\text{support}(X)} \quad (2)
\]

We mainly use "lift" as a measure of the performance of a targeting model. Lift is a better measure to compare the confidence of the rule with a benchmark value, in which we assume that the occurrence of the itemset in a transaction is independent of the occurrence of the antecedent for each rule. The benchmark confidence value for a rule can be calculated as equation 3:

\[
\text{Benchmark Confidence} = \frac{\text{no. transactions with Y itemset}}{\text{no. transactions in database}} \quad (3)
\]

The lift ratio is the confidence of the rule divided by the confidence, assuming independence of consequence from antecedent [31]

\[
\text{lift ratio} = \frac{\text{confidence}}{\text{benchmark confidence}} \quad (4)
\]

The association rules mining in this study utilized the support, confidence and lift measures. It includes the following concepts:

1) An association rule is an expression of the relationship of two sets of items, such as \(X \Rightarrow Y\). Given a database D with transactions, suppose each transaction is \(T \in D\). In this case, \(X \Rightarrow Y\) means that whenever a transaction T contains X than T probably contains Y also. The probability or rule of confidence means that the percentage of all transactions including X and Y in regard to all transactions including X, which can be represented as \(p(Y \subseteq T | X \subseteq T)\) [32].

2) Association rules mining could be constrained to the problem of finding large itemsets, where a large itemset is a collection of items existing in a database transactions equal to or greater than the support threshold [33].

3) The common algorithms of associate rule mining includes a priori algorithm and two-pass algorithm. The former is a multi-pass algorithm and needs up to \(a+1\) database scans, where a is the number of items (attributes). The latter has only two database scan passes. Association rules mining is applicable in numerous database communities. It has large applications in the retail industry to improve market basket analysis [34].

The study applied to the association rule mining method and used the support, confidence and lift measures to determine the relationship among keywords. Support, confidence and lift ratio are utilized as the threshold when initially calculating and finding rules of interests. The advantages of the support, confidence and lift ratio hinge on the easy comprehension and operation. The shortcoming of the support and confidence ratios do not allow fully assessing the quality of the association rule under consideration and often encounter redundant general rules [35].
D. DATA DESCRIPTION

The study at first used the text mining method to obtain the major keywords TF-IDF data. Major keywords means the keywords with the largest occurrence in the text data. The Chinese text resource was from two origins: sharing bikes related Chinese news reports and sharing bikes related posts on the discussion boards of the social media. The Chinese news reports text materials were assorted from "Wise News" (http://wisesearch.wisers.net) via the portal of Macau University of Science and Technology Library. It provides the access to more than 350 titles of newspapers, magazines, journals and newswires published in the Mainland China, Hong Kong SAR, Macau SAR, and Taiwan. The Chinese text materials of the social media discussion boards were assorted from "Tianya" (http://www.tianya.cn/). "Tianya" is referred to as China's most popular public forum. With approximately 270,000 visitors per day, "Tianya" has generated the most controversial and influential online debates and, in turn, has attracted massive media exposure and public responses [36]. The time frame of Chinese news text started from December 2016 to early May 2018. The Tianya web discussion boards text materials ranged from February 2017 to May 2018. Both materials were assorted on the monthly basis. For the Chinese social media discussion board text materials, the time frame ranged from February 2017 to early May 2018 and the text materials were also assorted on the monthly basis. The number of Chinese news texts was listed on Figure 2. In Figure 2, there are 10 or 11 related Chinese news texts in each month except May 2018, because the research time frame ended in early May 2018. The number of Tianya Web posts texts was listed on Figure 3:

IV. RESULTS

A. TEXT MINING RESULTS

The study utilized the text mining method to explore the major keywords of sharing bikes related Chinese news reports and posts on the discussion boards of "Tianya" website. The study obtained the major keywords of the sharing bikes related Chinese news reports on Table 1. The study also found the major keywords of sharing bikes related posts on the discussion boards of "Tianya" website on Table 2.

TABLE 1. Major Keywords of Chinese News Reports

<table>
<thead>
<tr>
<th>Keywords</th>
<th>Chinese Words or Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>increase</td>
<td>zhang</td>
</tr>
<tr>
<td>enterprise</td>
<td>gongxian</td>
</tr>
<tr>
<td>sharing</td>
<td></td>
</tr>
<tr>
<td>Hello</td>
<td></td>
</tr>
<tr>
<td>bike</td>
<td></td>
</tr>
<tr>
<td>city</td>
<td></td>
</tr>
<tr>
<td>market</td>
<td></td>
</tr>
<tr>
<td>launch</td>
<td></td>
</tr>
<tr>
<td>investment</td>
<td></td>
</tr>
<tr>
<td>deposit</td>
<td></td>
</tr>
<tr>
<td>DiDi</td>
<td></td>
</tr>
<tr>
<td>user</td>
<td></td>
</tr>
<tr>
<td>problem</td>
<td></td>
</tr>
<tr>
<td>complaint</td>
<td></td>
</tr>
<tr>
<td>financing</td>
<td></td>
</tr>
<tr>
<td>a brand name of Chinese bike-sharing company</td>
<td>danche</td>
</tr>
<tr>
<td></td>
<td>chengshi</td>
</tr>
<tr>
<td></td>
<td>shichang</td>
</tr>
<tr>
<td></td>
<td>toufang</td>
</tr>
<tr>
<td></td>
<td>touzi</td>
</tr>
<tr>
<td></td>
<td>yajin</td>
</tr>
<tr>
<td>a brand name of Chinese bike-sharing company</td>
<td>shiyongzhe</td>
</tr>
<tr>
<td></td>
<td>wenti</td>
</tr>
<tr>
<td></td>
<td>toussu</td>
</tr>
<tr>
<td></td>
<td>rongzi</td>
</tr>
</tbody>
</table>

We got the major keywords frequency data (TF-IDF) from the text mining process. The descriptive statistics of major keywords frequency data of sharing bikes related Chinese news reports was shown on Table 3.

The descriptive statistics of major keywords frequency data of "Tianya" web discussion board text materials was shown on Table 4.

B. ASSOCIATION RULE MINING RESULTS
TABLE 2. Major Keywords of Tianya Website Posts

<table>
<thead>
<tr>
<th>Keywords</th>
<th>Chinese Words or Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>mess</td>
<td>luan</td>
</tr>
<tr>
<td>enterprise</td>
<td>qiyi</td>
</tr>
<tr>
<td>use</td>
<td>tingfang</td>
</tr>
<tr>
<td>parking</td>
<td>chengshi</td>
</tr>
<tr>
<td>city</td>
<td>shichuang</td>
</tr>
<tr>
<td>market</td>
<td>yajin</td>
</tr>
<tr>
<td>deposit</td>
<td>yonghu</td>
</tr>
<tr>
<td>convenience</td>
<td>jingji</td>
</tr>
<tr>
<td>user</td>
<td>wenti</td>
</tr>
<tr>
<td>economy</td>
<td>daohi</td>
</tr>
<tr>
<td>problem</td>
<td>fazhan</td>
</tr>
<tr>
<td>collapse</td>
<td>toufang</td>
</tr>
<tr>
<td>development</td>
<td>guanli</td>
</tr>
</tbody>
</table>

TABLE 3. Descriptive Statistics of Major Keywords TF-IDF Data of Chinese News Reports

<table>
<thead>
<tr>
<th>Keywords</th>
<th>Max.</th>
<th>Min.</th>
<th>Average</th>
<th>Std.Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>increase</td>
<td>2.35</td>
<td>0</td>
<td>1.56</td>
<td>1.14</td>
</tr>
<tr>
<td>enterprise</td>
<td>0.155</td>
<td>0</td>
<td>0.029</td>
<td>0.05</td>
</tr>
<tr>
<td>sharing</td>
<td>1.80</td>
<td>0</td>
<td>0.24</td>
<td>0.44</td>
</tr>
<tr>
<td>bike</td>
<td>1.157</td>
<td>0</td>
<td>0.26</td>
<td>0.44</td>
</tr>
<tr>
<td>Hello</td>
<td>0.21</td>
<td>0</td>
<td>0.017</td>
<td>0.05</td>
</tr>
<tr>
<td>city</td>
<td>0.17</td>
<td>0</td>
<td>0.035</td>
<td>0.06</td>
</tr>
<tr>
<td>market</td>
<td>0.19</td>
<td>0</td>
<td>0.025</td>
<td>0.06</td>
</tr>
<tr>
<td>launch</td>
<td>0.119</td>
<td>0</td>
<td>0.02</td>
<td>0.04</td>
</tr>
<tr>
<td>complaint</td>
<td>0.338</td>
<td>0</td>
<td>0.019</td>
<td>0.079</td>
</tr>
<tr>
<td>investment</td>
<td>0.17</td>
<td>0</td>
<td>0.029</td>
<td>0.059</td>
</tr>
<tr>
<td>deposit</td>
<td>0.40</td>
<td>0</td>
<td>0.046</td>
<td>0.107</td>
</tr>
<tr>
<td>DiDi</td>
<td>0.48</td>
<td>0</td>
<td>0.064</td>
<td>0.138</td>
</tr>
<tr>
<td>user</td>
<td>0.21</td>
<td>0</td>
<td>0.03</td>
<td>0.065</td>
</tr>
<tr>
<td>financing</td>
<td>0.26</td>
<td>0</td>
<td>0.047</td>
<td>0.08</td>
</tr>
<tr>
<td>problem</td>
<td>0.13</td>
<td>0</td>
<td>0.024</td>
<td>0.049</td>
</tr>
</tbody>
</table>

FIGURE 4. Item Frequency Plot of Chinese News Reports Major Keywords

1) Results of Chinese News Reports Major Keywords

In order to facilitate the association rule mining process, the study divides each keywords TF-IDF data into two parts: below average value and above average value. The author chooses the average TF-IDF value as the standard to divide the keywords into two parts. The rationale of choosing the average as the threshold hinges on the descriptive statistics of the keywords TF-IDF values. As the TF-IDF values of the major keywords have the larger divergence, the division of the keywords does not change the characteristics of the keywords TF-IDF values distribution. The study made use of the "arules" software package [37] of the R language to analyze the association rules among keywords. The item frequency plot was listed as Figure 4. According to the Figure 2, the most frequent items were "complaint (0-average value)", "deposit(0-average value)", "Hello (0-average value)", "market(0-average value)" and "launch (0-average value)".

The study employed the "Apriori" algorithm to obtain the association rules and used the support measure to search for the important relationship. In order to make sure the antecedent (LHS) and the consequent (RHS) independent in all rules, the study used the chi-squared statistic to obtain the rules. The chi-squared statistic was set at 0.05 [38]. The descriptive statistics of these rules was listed as follows:

TABLE 5. Descriptive Statistics of Association Rules

<table>
<thead>
<tr>
<th>support</th>
<th>confidence</th>
<th>lift</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min.</td>
<td>0.055</td>
<td>0.60</td>
</tr>
<tr>
<td>Max.</td>
<td>0.944</td>
<td>6.00</td>
</tr>
<tr>
<td>Mean</td>
<td>0.287</td>
<td>5.181</td>
</tr>
</tbody>
</table>

The study got top 10 rules according to the lift ratio of
The parallel coordinates plot of the top 10 association rules was listed as Figure 5. It shows the content of the major association rules.

According to Table 6, we can find that "deposit" (above average) as the consequent account for all top 10 association rules. The study also finds that "complaint" (above average) plays the important role as the first or second element in the antecedents (LHS) on Figure 5.

2) Results of "Tianya" Web Discussion Boards Major Keywords

The study also divided each "Tianya" web discussion boards major keywords TF-IDF data into two parts: below average value and above average value. According to Table 4, the keywords TF-IDF values have the larger discrepancy. It can explain the rationale of the usage of the average as the cut-off point. The average value threshold division does not change the characteristics of keywords TF-IDF values. The item frequency plot was listed as follows.

The study used the chi-squared statistic to obtain the rules. The chi-squared statistic was set at 0.05. The descriptive statistics of association rules was listed in Table 7.

The first 10 items with larger lift ratio were listed as follows.

The parallel coordinates plot of the top 10 association rules was listed as Figure 7. It shows the content of the major association rules.

According to Table 8 and Figure 7, we can observe that "launch" (above the average) and "economy" (above the average) are the two major consequent (RHS) in all of the association rules. In top 10 association rules, use(0-average), development(0-average), collapse(above average) are major elements of precedents(LHS).
V. DISCUSSION

The study aimed at bridging the gap of the current literature of the problems of influencing factors of the bike-sharing system in China. The major findings of this study were listed as follows.

1) According to the association rule mining results of Chinese news reports major keywords on Table 6 and Figure 5, complaints (above the average) is included in all of the antecedents of the top ten association rules. Besides, deposits (above the average) is included in all of the precedents of the top ten association rules. It indicates that the deposits are the public opinion focus and the main reason of complaints in bikes-sharing related Chinese news reports. As the TF-IDF values of "complaint" reached the highest point in January 2018 and that of "deposit" reaches the second highest point and the highest point in January and March 2018 (Figure 8 and Figure 9, the orange line means the average TF-IDF value), the association rule mining results imply that the relationship between "complaint" and "deposits" exists in the early period of 2018. It also shows that the deposits problem is the focus of the complaints in the related news reports. Therefore we can conclude that the deposits problem as the news reports focus existed in the early 2018. It can be taken as the early warning sign of the sharing bikes companies crisis in December 2018. In the late December 2018, bike-sharing company OFO has encountered the financial failure, lots of the customers demanded to get their deposits refunded [39].

2) We also obtain the "Tianya" web discussion boards keywords on the basis of the association rule mining results. According to the results of Table 8 and Figure 7, the major one in top 10 association rules is "mess" (above average) and "use" (0-average) as the precedent and "launch" (above average) as the consequence because of its largest support ratio (0.1875). This major rule is also shown in thicker arrow in the middle of Figure 7. It means the launch of sharing bikes is the main public opinion focus, and the related precedents include the element of "economy", "collapse", "market", "convenient" and "use". Besides, as far as the other 6 association rules with the "economy" (above
average) as the consequent (RHS) are concerned, the precedents (LHS) include “use”, “launch”, “convenience”, “enterprise”, “city” and “collapse”. It can be inferred that the relationship between usage, launch, collapse and the economy concern of the sharing bikes is also the public opinion focus in Tianya web posts discussion.

Besides, the top 10 association rules of Tianya web posts keywords took place when the TF-IDF values of keywords “launch” and “economy” are above the average value. According to Figure 10 and Figure 11, the consequent (RHS) of the top 10 association rules, the time frame of the above average TF-IDF values (average value in orange line) took place in June, August, September, November 2017 and January 2018 for the keyword “launch”. The time frame of the above average TF-IDF value of keyword “economy” took place in June, July, December 2017 and April, May 2018. We can also obtain from Figure 12 that the keyword “collapse”, which belongs to the precedents of three of top 10 association rules, has the over average TF-IDP value (orange line) in June, July, November and December 2017 and January and April 2018. The time frame almost corresponds to that of keyword “launch” and “economy”. It indicates that the collapse crisis of bike-sharing companies has become the top issues almost one year before the financial failure and bankrupt happened in December 2018. It can be also concluded that the “collapse” of the bike-sharing companies in China relates to the launching of resources and the management issues.

3) The study concludes that the association rule mining results of the keywords from sharing bikes related Chinese news reports and “Tianya” web discussion boards focused on the commercial sharing bikes system. The influencing factors of the bike-sharing system problems in China includes the following parts:

- Deposit Problem: The association rule mining results of Chinese news reports keywords TF-IDF data indicate that deposits are the main focus of the sharing bikes related Chinese news reports.
It relates to the customers’ complaints. Such a circumstance took place earlier than the management failure really happened. It implies the deposit problem is one of the key issues of the sharing bikes companies management problems. As most of the commercial bike-sharing systems in China belong to the Free Floating Bike Sharing system (FFBS system, or dockless sharing bikes system), the functioning of this system needs the support of customers’ deposits in the beginning. However, as the rebalancing of sharing bikes to ensure bikes always available was not well controlled, the sharing bike system turned to be chaotic. The deposit mechanism initially aimed at solving the trust problem [40]. However, it can not solve the bike-sharing system problem. In late December of 2018, some bike-sharing companies fell prey to the bankruptcy and aroused the criticism of their business operations. Bike-sharing and car-sharing companies absorb the users’ deposits and do not need to pay annual interest for their users. The Chinese authorities have not promulgated the regulations of the deposits yet [41]. When the bike-sharing system encountered problems, the customers would turn to ask for the deposits to get refunded. It would aggravate the daily operation problems and make the sharing bikes companies collapse. In order to solve the deposits problem of sharing bikes, the deposits regulations need to be ratified and the authorities concerned should prevent the bikes-sharing companies from embezzling the deposits.

- Management and Sustainability Problems: According to the associate rule mining results, we observed that the public opinion on the “Tianya” web discussion boards focused on the collapse of bike-sharing companies and the related factors, such as “management” and “launch” (means providing resources for business operations in Chinese). It can be inferred that the social web media users concerned about the possibility of bike-sharing companies collapse. As far as the sustainability is concerned, it plays the role as the guidance for companies to format new institutions. Sustainability practices are shaped and consolidated owing to the need to respond to the pressures from society, government and competition [42]. We can conclude that the problem of bike-sharing companies in China was also the sustainability problem due to the breakdown of sustainability concept in the business operations. It needs the government to make the regulations about the bike-sharing industry and reasonably control the number of the sharing bikes.

In sum, the development of commercial bike-sharing system should care about the sustainability dimensions, such as environmental, social and economic. In order to reach the goal of sustainability, the government should play the role as the contributor of the institutional sustainability logic [42] and regulate the bike-sharing market in fine tune. The sharing bikes companies should dedicate to the corporate social responsibility to finely rebalance the sharing bikes.

VI. CONCLUSION

According to the association rule mining results of the Chinese news reports keywords, it shows that the sharing bikes deposits are closely related to the customers’ complaints. As for the social web discussion boards context, the keywords association rules results imply that the users worry about the collapse and the management problems of the bike-sharing companies and the economy effects. The deposits and management problem exists as the public opinion focus almost one year before the bike-sharing companies’ financial failure news takes place. It indicates that the public opinion focus can be taken as the early warning sign of the bike-sharing problems in China.

We suggest that researchers can utilize other methods to explore the relationship among keywords and assort more related text materials from other media as the resource of text mining.

ACKNOWLEDGEMENT

This research was funded by Macau University of Science and Technology Faculty Research Grant. The authors would like to show the gratitude for Macau University of Science and Technology Library as the proxy for the linkage of the Wisers Information Portal.

REFERENCES


[27] XUXIANAN LEI Xuxianan Lei completed her Bachelor’s degree in Economics from Nankai University in 2017. Then she obtained her Master’s degree in Applied Economics from Macau University of Science and Technology in 2019. Her research interests include the data mining and algorithm application in the social science research.

[28] SHIANGHAU WU Shianghau Wu received his Ph.D. degree from National Chengchi University in 2006. He serves in the School of Business at Macau University of Science and Technology as the Associate Professor. His research interests include the data mining and algorithm application in the social science research.