

Research Article

Cluster Analysis of Automobile Innovative Users Based on Interactive Innovation Value

Haowen Lin,¹ Yaping Chen,² and Yushu Yang ³

¹Undergraduate Student, Computer Science Department, University of Southern California, Los Angeles, CA 90007, USA ²Postgraduate Student, School of Economics and Management, Tongji University, Shanghai 200092, China ³Doctoral Student, School of Economics and Management, Tongji University, Shanghai 200092, China

Correspondence should be addressed to Yushu Yang; yangyushu21@163.com

Received 11 August 2018; Accepted 5 November 2018; Published 25 November 2018

Academic Editor: Marcello Pellicciari

Copyright © 2018 Haowen Lin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As a new model of product creation, "interactive innovation" can effectively improve the success rate of enterprise product innovation. Because of the wide application of the Internet and the pervasiveness of social media such as forums and blogs, there is a rapid growth of product reviews online. These social media platforms provide an effective channel for interactive innovation between enterprises and users. This paper is based on the innovation application of automobile products. The purpose of this paper is to identify and classify the innovative users in automobile forums and analyze the characteristics of different user groups. First, we summarize six typical characteristics of innovative users and quantify these six characteristics. Second, we make a cluster analysis of user data and divide innovative users into three classes according to the innovation value. And then, based on the different characteristics of three types of users, we present different interactive innovation methods for three types of users. Finally, we construct a pyramid model of innovative user to show the distribution status of various users.

1. Introduction

Product innovation is an essential factor in enterprise development. Constant technology innovation and new products development are essential for enterprises to survive in the increasingly fierce market competition [1]. However, product innovation is a very risky game. A large number of cases demonstrate that many product innovations fail or fail to recover the investment in the end. Therefore, it is important to ensure the success rate of product innovation.

As a new model of product innovation, "interactive innovation" enables enterprises to understand the overall needs of users comprehensively and effectively, thereby ensuring the success of enterprise product innovation. "Interactive innovation" refers to the innovative design and improvement process of products through the interaction and cooperation between the production enterprises and users [2]. As Prahalad et al. [3] pointed out that, in the 21st century, users would actively participate in the design and production of enterprise products and services, and users would be the co-creators of product value, thus, products designed by creative users are easily accepted by the market and companies would have a high success rate of product innovation.

With the popularization of Internet and product professional social media, such as blogs and forums, many users share their opinions and ideas, evaluate product performance and features, reflect product defects and problems, and analyze the advantages and disadvantages of different products on such platforms. A large corpus of data about product reviews has been accumulated and those data help enterprises to excavate users' needs and potential needs. Thus, it provides an effective channel and platform for the interactive innovation of enterprises and users. If enterprises can effectively utilize the convenience provided by the Internet and work closely with users to make co-creation innovations, it will be critical for enterprises to adapt to market change and win the competition. In this paper, we regard the users of automobile forum under the Internet as research objects and study the interactive innovation model between enterprises and users.

The key to carry out the interactive innovation activity for enterprises is to identify the user community of innovation value effectively [4]. Companies then can take different approaches to cooperate with users, thereby maximizing the users' innovation value. However, every user has a different contribution to the enterprise product innovation. In other words, the user value degree to the product innovation is a continuous distribution; thus users cannot simply be divided into lead users and nonlead users. Therefore, we classified different innovative users based on the innovation value and then took different methods of interactive innovation for different user groups, so that the innovation value of users can be maximized in the product innovation activities of enterprises.

2. Related Work

Von Hippel [5] from Massachusetts Institute of Technology proposed the concept of "innovation source" and divided innovation resource into three classes: user innovation, manufacturer innovation, and supplier innovation. Based on the research on innovation source, innovators who were initially successful in business were classified as high-valued innovative users [6, 7]. These innovative users are a small part of the user community [2]. Thus, the key to interactive innovation is to identify and select the right users [8, 9].

Professor von Hippel [2] firstly proposed the concept of lead user. However, this method, which simply divided users into lead users and nonlead users, neglected the contributions of nonlead users to enterprise product innovation. In order to provide new methods for enterprises to cooperate with users in product innovation, in this paper, innovative users are used to refer to the major participants of user innovation activities, and the user groups are grouped into three classes according to the different value of innovation: high-valued innovative users, medium-valued innovative users, and lowvalued innovative users.

Much research has been done to explore the identification of innovative users. Poetz et al. [10] proposed screening. This method used questionnaires or telephone interviews to collect information of targeted users, thereby finding out the right users presenting desired characteristics. Von Hippel et al. [11] proposed pyramiding. Pyramiding builds link chain through users' recommendation and follows the link chain to find out users with the maximum of innovative characteristics. Jeppesen et al. [12] proposed broadcasting. Public details of problems are broadcasted to be solved, and this invites capable users to submit possible solution. Belz et al. [13] proposed netnography. Netnography is the combination of network and demography, and it uses demography method to analyze people in network communities. Spann et al. [14] proposed virtual stock market. This method summons targeted users through Internet, and participators trade in the virtual stock market based on their own knowledge and judgement. Finally, the users with good performance are regarded as lead users.

The above studies are based on traditional market research and customer surveys to find or identify highvalued innovative users. These methods are different only in selecting samples and identifying criteria according to the area of the enterprise [15]. The main contributions of this paper are shown as follows. Firstly, we regard the large amount of user information as the research object and obtain the user's innovation value through text mining. Secondly, we employ cluster analysis algorithm to divide the innovative users into three classes and then analyze the characteristics of different user groups with different innovation value. Finally, we proposed corresponding suggestions for different user groups on how to carry out interactive innovation with enterprise, so as to maximize the innovation value of users and ensure the success rate of product innovation.

3. Characteristic Analysis of Automobile Forum Innovative Users

3.1. Characteristic Description of Innovative Users. Research on innovative users [16] has shown that some users have more motivation and potential than the others. Other studies [17] have also emphasized the characteristics of Internet users. Therefore, we divide the typical characteristic of innovative users in the automobile forum into two categories: characteristics in innovativeness and characteristics in network.

3.1.1. Characteristics in Innovativeness. Based on the existing research about innovative users [16], we summarized the characteristics in innovativeness in four aspects: Ahead of market trend, willingness to innovate, rich domain knowledge, and rich user experience.

(1) Ahead of Market Trend. Innovative users are capable of detecting the underlying demand in the market. On the one hand, the unsatisfied needs of innovative users reflect the future mainstream market demand. Franke et al. [18, 19] proposed that the innovative users should have the ability to perceive potential market demand in advance and expect to benefit from innovation. On the other hand, the dissatisfaction with the existing production is another manifestation of demand leading. Dissatisfied users are more likely to discuss the shortcomings of the products, give suggestions for improvement, and even improve products in reality.

(2) Willingness to Innovate. Innovative users are willing to innovate. According to the researches about innovative users [4, 20], the willingness to innovate is the premise for innovative users to take actions to improve the products. Therefore, innovative users tend to be more willing to innovate than normal users. Since no products existing in the market can satisfy their special needs, innovative users are likely to create product prototypes, which can meet their needs better [21]. Compared with other users, they show a stronger motivation to transform existing products.

(3) Rich Domain Knowledge. Domain knowledge reflects user's familiarity with the products and serves as the source of inspiration to satisfy their special needs through innovation [22]. Research on innovative users shows that, as users with high innovativeness, lead users have more internal skills and external resources than other normal users [23]. Under the same conditions, users with richer expertise would cost less to innovate, and they are more likely to innovate [22].

(4) Rich User Experience. Innovative users' new ideas and solutions to current problems are inspired by their user experiences, so rich user experience leads to high potential to innovate [24]. Since they have been using the product for a long time, lead users have accumulated a rich user experience and have the ability to many drawbacks and improvements to the existing products or services. Thus, the rich user experience is another important characteristic of innovative users.

3.1.2. Characteristics in the Network. Compared with traditional user aggregation community, automobile forums take advantage of the cyberspace to connect users distributed throughout the world and share information and knowledge. Because all users' activities are online, innovative users in the automobile forum would display different characteristics in the network.

(1) Community Activity. Community activity measures the frequency of user's online behaviors, which includes both posting and replying to share expertise, ideas, insights, and solutions in the community. Knowledge sharing is one of the most important functions of a virtual community. Researches have shown that innovative users tend to be active members in virtual communities [25]. Olson et al. [26] have also found that innovative users are greatly affected by the willingness to share and cooperate in the brand community. Thus, user's community activity is a valuable indicator to measure the user's innovation value.

(2) Network Connectivity. Network connectivity measures how central the user is located in the social network formed in the virtual community. Schreier et al. [27] found that the users with higher awareness of innovation tend to be the opinion leaders in the community. The trait of user's opinion leader is reflected through its connectivity in the social network, namely, the strength of network connection. The higher the user's strength of network connection in the social network, the greater its role in innovation communication and delivery, in other words, the higher the user's innovation value. Thus, user's network connectivity is an important indicator to measure the user's innovation value.

3.2. Quantitative Model of Innovative User's Characteristics. The quantitative model of innovative user's characteristics includes two parts: characteristics in innovativeness can be indicated by the content of users' posts, which can be gained using data mining and text analysis technologies; characteristics in the network can be reflected by users' demographic data and network activity data.

3.2.1. Quantitative Model of Characteristics in Innovativeness. The user's innovativeness is mainly reflected by users' post. In order to quantify the characteristics of innovativeness, we need to categorize the posts and use text classification technology to classify the posts.

The level of innovativeness can be quantified based on the text analysis of user's corpus. In summary, there

THEFT IS THE CHOOMERCHON CHIMMEN TO FOUR	TABLE 1:	The	classification	standard	for	posts
--	----------	-----	----------------	----------	-----	-------

Characteristic in Innovativeness	Post Theme
Ahead of market trend	Dissatisfaction with existing products Improvement to existing products
Willingness to innovate	Car modification Component DIY
Rich domain knowledge	Knowledge share Resource share
Rich user experience	Purchasing experience Using experience Self-driving travel
Others	Product resale Forum management Question and consultation

exists a matchup of the post theme and the corresponding characteristic in innovativeness reflected in the post. The classification standard for posts is shown in Table 1.

Using the text classification technology to classify the posts, for a specific post theme, we can get the number of posts (N), the number of essential posts (N_e), and the number of nonessential posts (N_n). The essential posts are posts pinned by forum administrators. When users think that a post is valuable, they will send a request to the administrators to highlight the posts. Generally, such posts are rich in content and have high reading value. They can be replied and the author can modify the original post. Thus, the essential posts contain high-valued information and should be differentiated from the nonessential posts. Therefore, we can get the following equation:

$$N = N_e + N_n \tag{1}$$

For the *i*th post under the corresponding theme post (L_i) , the level of the quantified characteristic in innovativeness presented in the post can be obtained by the following equation:

$$L_i = w_r N_{ir} + w_v N_{iv} \tag{2}$$

where N_{ir} represents the number of replies of the *i*th post and N_{iv} represents the number of views. w_r and w_v represent the weights given to the corresponding attributes, respectively.

Consequently, the level of the quantified characteristic in innovativeness (L) can be calculated by the following formula:

$$L = w_e \sum_{i=1}^{N_e} \left(w_r N_{ir} + w_v N_{iv} \right) + w_n \sum_{i=1}^{N_n} \left(w_r N_{ir} + w_v N_{iv} \right) \quad (3)$$

where w_e and w_n are constants, representing the weights given to essential posts and nonessential posts, respectively.

In the same way, we can quantify the level of the four characteristics in innovativeness, i.e., ahead of market trend (L_{mr}) , willingness to innovate (L_{in}) , rich domain knowledge (L_{ex}) , and rich user experience (L_{ux}) , as follows.

$$L_{mr} = w_e \sum_{i=1}^{N_{nd,e}} (w_r N_{nd,ir} + w_v N_{nd,iv})$$

$$+ w_n \sum_{j=1}^{N_{nd,n}} (w_r N_{nd,ir} + w_v N_{nd,iv})$$

$$L_{in} = w_e \sum_{i=1}^{N_{in,e}} (w_r N_{in,ir} + w_v N_{in,iv})$$

$$+ w_n \sum_{j=1}^{N_{in,n}} (w_r N_{in,ir} + w_v N_{in,iv})$$

$$L_{ux} = w_e \sum_{i=1}^{N_{ux,e}} (w_r N_{ux,ir} + w_v N_{ux,iv})$$

$$+ w_n \sum_{j=1}^{N_{in,n}} (w_r N_{ux,ir} + w_v N_{ux,iv})$$

$$L_{ex} = w_e \sum_{i=1}^{N_{ex,e}} (w_r N_{ex,ir} + w_v N_{ex,iv})$$

$$+ w_n \sum_{j=1}^{N_{in,n}} (w_r N_{ex,ir} + w_v N_{ex,iv})$$

$$(7)$$

3.2.2. Quantitative Model of Characteristics in the Network

(1) Community Activity. In consideration of the scalability of data, we choose the duration of registration (T) and the number of activities (F) as metrics to measure the community activity. The duration of registration is the total number of months from the user registration to the date of data acquisition. The number of activities is the total number of the users' posting and replying activities. Therefore, the level of community connectivity L_{act} can be represented by the frequency of the user's activity frequency, as shown in formula (8).

$$L_{act} = \frac{F}{T} \tag{8}$$

(2) Network Connectivity. In this study, we choose two metrics to measure user's network connectivity: the number of times replying to others' posts (N_{out}) and the number of times being replied by others in the community (N_{in}) . Thus, the level of network connectivity can be quantified using formula (9).

$$L_{con} = w_{in}N'_{in} + w_{out}N'_{out}$$
⁽⁹⁾

where N'_{in} and N'_{out} represent the values after normalization. w_{in} and w_{out} are constants, representing the weights of N'_{in} and N'_{out} . 3.3. *Indicator System*. The indicator system for identifying innovative users in automobile forum (e.g., AutoHome) is shown in Table 2.

4. Clustering of Innovative Users

4.1. Preprocessing of User Data for Automobile Forum. In this subsection, we will introduce the data preprocessing, including data collection, text categorization, and quantitative processing of user data.

4.1.1. Data Collection. The experimental data of this paper are derived from "Tiguan Forum" (https://www.autohome.com .cn/874/). AutoHome was launched in June 2005. It is one of the biggest automobile forums around the world and the most famous automobile forums in China. As the most valuable online automobile community in China, AutoHome provides service for automobile consumers throughout the entire course of purchasing and using and influences automobile consumers in China widely.

The research object, Tiguan Forum, is one of the 49 forums in AutoHome. The paper applies web crawler to collect user data from Tiguan Forum and constructs the Tiguan forum database, which includes all user information and posts from April 20, 2009, to January 8, 2017.

4.1.2. Text Categorization. To measure user's level of innovativeness through posting in the automobile forum, we need to identify which posts reflect the poster's characteristics in innovativeness and what characteristic in innovativeness is reflected in the post first. Therefore, we need to classify the posts according to their post themes. We choose to use short text classification technology to classify collected posts so that we can identify posters' characteristics in innovativeness and characteristics in the posts.

Through text categorization processing, the posts from Tiguan Forum in AutoHome are divided into five categories: posts indicating user's ahead of market, willingness to innovate, domain knowledge, user experience, and others, respectively.

As the text categorization method is a supervised learning process, it is necessary to construct a training set in order to provide a reference for the automatic classification. In this study, we randomly selected some posts from the AutoHome Tiguan user database, labeled the posts with the 5 aforementioned tags (ahead of market, willingness to innovate, domain knowledge, user experience, and others), and constructed the test set. The training set contains 4,000 forum posts from five categories. The last tag is used to identify and filter posts that do not reflect the characteristics in innovativeness.

The categorization result is shown in Table 3. Through sample survey of the result, we labeled the sample manually and compared the results of automatic processing and manual processing. The comparison result shows that the accuracy rate is 69.03%, which is quite acceptable to the research requirement.

As shown in the categorization result, the number of posts labeled as user experience is more than the number

Characteristics		Indicator	Data Acquire
	Ahead of market trend	Replies to related recommended posts	(1) Use text categorization algorithm to tag post titles and classify user's posts to different themes
	Willingness to innovate	Views related recommended posts	(2) Delete posts that cannot reflect characteristics in innovativeness
Characteristics in innovativeness	User experience	Replies to non-recommended posts	(3) Collect data on the four indicators for user's posts under a specific theme
	Domain knowledge	Views non-recommended posts	$L = w_e^{N_e} \sum_{i=1}^{N_e} (w_r N_{ir} + w_{\nu} N_{i\nu}) + w_n \sum_{i=1}^{N_n} (w_r N_{ir} + w_{\nu} N_{i\nu})$
	Network connectivity	Total times of replies Total times of being replied	Retrieve required data from the established database
Characteristics III network	Community activity	Duration of registration total times of activity	$L = w_e^{\sum_{i=1}^{2r_e}} (w_r N_{ir} + w_{\nu} N_{i\nu}) + w_n^{2r_n} (w_r N_{ir} + w_{\nu} N_{i\nu})$

TABLE 2: Indicator system of innovative users identification in the automobile forum.

TABLE 3: Text categorization result.

Tag	Number	Percentage
Ahead of market	776	3.97%
Willingness to innovate	2642	13.52%
User experience	8287	42.40%
Domain knowledge	3328	17.03%
Others	4513	23.09%

of posts labeled as other tags, accounting for nearly half of the total number. Posts labeled as willingness to innovate and domain knowledge account for similar percentage, 13% and 17%, respectively. Posts labeled as ahead of market are the fewest, less than 4% of the total. The categorization result is consistent with the overall trend observed manually from Tiguan Forum, which partially proves that the categorization result is reliable.

4.1.3. Quantitative Processing of User Data. Through web crawler and text categorization, the user data from Tiguan Forum in AutoHome is obtained and processed. According to the indicator system constructed in Section 3, we need to quantify the data to get the level of user's characteristics.

Through seminar and survey of experts in automobile field, the weights of each characteristic are determined. The values of these weights, including w_e , w_n , w_r , w_v , w_{in} , and w_{out} , are 0.65, 0.35, 0.65, 0.35, 0.50, and 0.50, respectively. First, we need to substitute the values of each weight into formula (1)-(9) and calculate the quantization values of corresponding innovation characteristics of each user, respectively. And then, 12624 users' information are obtained. Finally, we normalize the data results of users' innovative features and network characteristics. Partial normalized results of the data above are shown in Table 4.

4.2. Cluster Analysis of Innovative Users

4.2.1. Results of Cluster Analysis. The paper adopts cluster analysis to classify users in the automobile forum and identify users with high potential to innovate. Based on the study of clustering algorithms, we choose to combine Balanced Iterative Reducing and Clustering Using Hierarchies (BIRCH) algorithm with Agglomerative Nesting (AGENES) algorithm. First, we use BIRCH algorithm to fulfill the preclustering and construct the clustering feature tree, and then use AGNES algorithm to undertake the clustering and achieve the classification of users.

Applying the chosen two-step clustering algorithm to cluster members of Tiguan Forum in AutoHome, the final result shows the 12624 samples are divided into 3 user groups, one cluster with 379 samples, one with 1880 samples, and the third one with 10365 samples, as shown in Table 5.

The following results can be obtained from Table 5. First, users in Cluster 1 are the minority, accounting for around 3% of the total. These users are significantly higher than the other two groups of users in one or a few innovative characteristics. Thus, users in Cluster 1 are named as highvalued innovative users. Second, users in Cluster 2 are slightly more than those in Cluster 1, accounting for around 15% of the total. Users in Cluster 2 have a higher level of characteristics in innovativeness than users in Cluster 3 but lower than users in Cluster 1. Thus, users in Cluster 2 are named as mediumvalued innovative users. Finally, as the majority of users, users in cluster 3 account for 82% of the total. The level of characteristics of users in Cluster 3 is generally lower than that of the other two types of users. Thus, users in Cluster 3 are named as low-valued innovative users.

4.2.2. Cluster Validation. After clustering, we need to evaluate the validation of clustering result. First, this paper adopts Analytical Hierarchical Process (AHP) to obtain suitable weights of different characteristics of innovative users. After weighted calculation, each user has a composite score. Second, the users were sorted and classified by the users' AHP comprehensive scores, and the manually annotated clusters were obtained. By comparing the results of clustering and manual annotation, this paper employs three external criteria to evaluate the cluster validation: precision, recall, and F value.

The formulas for these three external criteria are as follows. Assume that the final clustering result is

$$C = \{C_1, C_2, \dots, C_n\}$$
 (10)

And the cluster structure determined by human being is

$$P = \{P_1, P_2, \dots, P_n\}$$
(11)

For every cluster P_j , assume that there is a corresponding cluster C_i , and the correspondence is unknown. In order to find the exact C_i for P_j , we need to scan the *n* clusters in *C*. Then we can calculate precision, recall, and F value.

Precision:

$$P\left(P_{j},C_{i}\right) = \frac{\left|P_{j}\cap C_{i}\right|}{\left|C_{i}\right|}$$
(12)

Recall:

$$R\left(P_{j}, C_{i}\right) = \frac{\left|P_{j} \cap C_{i}\right|}{\left|P_{j}\right|} \tag{13}$$

F value:

$$F\left(P_{j},C_{i}\right) = \frac{2 \cdot P\left(P_{j},C_{i}\right) \cdot R\left(P_{j},C_{i}\right)}{P\left(P_{j},C_{i}\right) + R\left(P_{j},C_{i}\right)}$$
(14)

Since the number of samples is enormous, we adopt the method of sampling comparative analysis. The manual annotation method is as follows. First, 10% samples were extracted from Cluster 1, Cluster 2, and Cluster 3, and the selected samples were mixed. Second, combining with AHP, we invited experts to rate the six characteristics according to the importance compared with the other characteristics. We then obtained the weights of each characteristic being 0.36,

			LABLE 4: Quantification r	normalized result.		
Host ID	Community Activity	Network Connectivity	User Experience	Domain Knowledge	Ahead of Market Trend	Willingness to Innovate
host_id	Leve <u>l a</u> ct	level_link	level_ux	level_ex	level_nd	level_in
13305157	0.3188	2.0588	0.8903	-0.1998	-0.1218	3.3609
10571765	0.2799	2.9561	0.0552	1.0753	-0.1218	3.0887
1315792	-0.1861	0.6073	-0.0771	9.2434	-0.1218	-0.0665
11251975	4.05738	18444	11671	19979	12182	1.07331
11401964	1.38303	23722	11671	-19979	12182	.19088
11503528	1.05727	2.21713	00123	19979	12182	02169
11580181	.10057	.92398	.33313	.23936	12182	.40028
10403030	28586	15804	11671	19979	12182	.23888
10405822	04203	.42255	09676	19979	12182	02749
10423821	15930	.34338	08522	.13090	12182	.102.27



FIGURE 1: Pyramid of innovative users.

TABLE 5: Clustering result overview.

Cluster	Number	Percentage
1	379	3.00%
2	1880	14.89%
3	10365	82.11%

TABLE 6: Cluster validation.

Cluster	1	2	3
Precision	0.8846	0.8131	0.9615
Recall	0.6675	0.7660	0.9802
F value	0.7609	0.7888	0.9708

0.07, 0.13, 0.19, 0.20, and 0.05, respectively. Finally, after the calculation of the composited AHP scores, users ranked as the top 3% are classified as Class 1, users ranking from 3% to 20% are classified as Class 2, and the others are classified as Class 3.

According to the comparison of cluster analysis and manually labeled testing samples, the precision, recall, and F value are shown in Table 6. The range of the three measurements above is [0, 1]. The bigger the value, the more similar the result of clustering analysis and the result of artificial classification. As shown in Table 6, the clustering result of the chosen algorithm is acceptable, and the analysis of the result is reasonable.

5. Characteristic Analysis of Cluster User Groups

The cluster analysis automatically divides the automobile forum users into three classes according to their innovation value: high-valued innovative users, medium-valued innovative users, and low-valued innovative users.

According to the analysis result, the distribution of innovative users is a typical pyramid. High-valued innovative users have the greatest ability to innovate. They are the minority with leading characteristics at the top of the pyramid. The number of medium-valued innovative users is slightly more than high-valued innovative users and less than low-valued innovative users. Low-valued innovative users have little potential to innovate, but they are the majority of the user group. The cluster analysis result shows that as the potential to innovate increases, the number of users decreases. Therefore, even for automobile industry with a large number of users, the number of high-valued innovative users will be small as well.

According to the analysis result, we can conclude that the higher the user's value or potential in innovation, the more deeply the users should be involved in co-creation. Based on the user clustering results, we proposed the corresponding innovative interaction methods and constructed the pyramid, as is shown in Figure 1.

Different methods are adopted to co-create with different user groups. For high-valued innovative users, enterprises should cooperate with the users to carry out the product innovation, thereby excavating the innovation value of the high-valued user as the co-creator. For medium-valued innovative users, enterprises should lead and guide the users during the innovation process, thereby excavating the innovation value of the medium-valued users as information providers. For low-valued innovative users, enterprises should interact with this user group after the launching of products to acquire feedback and propagate products, thereby excavating the innovation value of the low-valued user as the product user.

The pyramid of innovative users illustrates the distribution of innovative users in the user group and methods to cocreate with different types of users. As shown in the model, the higher the user's value for innovating, the more active they are in the innovation process.

6. Conclusions

The purpose of this paper is to identify and classify the innovative users in automobile forums and research the interactive innovation method for different user groups. First, we analyze characteristics of innovative users in automobile forums and construct an innovative identification system to quantify the innovative score for each user. Second, the system acquires the user data of automobile forums and utilizes text mining to calculate the innovation value of each user. We then also conduct the cluster analysis of users according to the different innovation value. Third, we analyze the characteristics of user groups with different value of innovation and give corresponding suggestions on the cocreation innovation method for different users. At the end of this paper, we construct a pyramid model of innovative users to show the distribution status of various users.

This study contributes to innovation user identification and co-creation with users. However, there are still limitations in the study. For example, the determination of index weight inevitably adopts the method of questionnaire surveys and expert interview. Thus, the method proposed in this paper cannot completely avoid subjective influence. In conclusion, further research can continue in the following two directions. First, the application of text analysis and deep learning in innovative user identification can be further explored. Second, the application of cluster analysis in innovative user identification can also be deeply discussed.

Data Availability

The experimental data used to support the findings of this study are included within the article. The website of this data is "Tiguan Forum" (https://www.autohome.com.cn/874/).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work has been supported by National Natural Science Foundation of China (No. 71672128) and the Fundamental Research Funds for the Central Universities.

References

- T. Mikkonen, C. Lassenius, T. Männistö, M. Oivo, and J. Järvinen, "Continuous and collaborative technology transfer: Software engineering research with real-time industry impact," *Information and Software Technology*, vol. 95, pp. 34–45, 2018.
- [2] E. Von Hippel, "Lead users: a source of novel product concepts," Management Science, vol. 32, no. 7, pp. 791–805, 1986.
- [3] C. K. Prahalad and G. Hamel, "The core competence of the corporation," *Harvard Business Review*, vol. 68, no. 3, pp. 275– 292, 2006.
- [4] C. Lüthje and C. Herstatt, "The lead user method: an outline of empirical findings and issues for future research," *R&D Management*, vol. 34, no. 5, pp. 553–568, 2004.
- [5] E. Von Hippel, "The dominant role of users in the scientific instrument innovation process," *Research Policy*, vol. 5, no. 3, pp. 212–239, 1976.
- [6] C. Freeman, A. B. Robertson, P. J. Whittaker et al., "Chemical process plant: Innovation and the world market," *National Institute Economic Review*, vol. 45, no. 45, pp. 29–57, 1968.

- [7] G. L. Urban and E. Von Hippel, "Lead user analyses for the development of new industrial products," *Management Science*, vol. 34, no. 5, pp. 569–582, 1988.
- [8] G. L. Lilien, P. D. Morrison, K. Searls, M. Sonnack, and E. Von Hippel, "Performance assessment of the lead user ideageneration process for new product development," *Management Science*, vol. 48, no. 8, pp. 1042–1059, 2002.
- [9] D. Mahr and A. Lievens, "Virtual lead user communities: Drivers of knowledge creation for innovation," *Research Policy*, vol. 41, no. 1, pp. 167–177, 2012.
- [10] M. K. Poetz and R. Prügl, "Crossing domain-specific boundaries in search of innovation exploring the potential of pyramiding," *Journal of Product Innovation Management*, vol. 27, no. 6, pp. 897–914, 2010.
- [11] E. Von Hippel, N. Franke, and R. Prügl, "Pyramiding: Efficient search for rare subjects," *Research Policy*, vol. 38, no. 9, pp. 1397– 1406, 2009.
- [12] L. B. Jeppesen and K. R. Lakhani, "Marginality and problemsolving effectiveness in broadcast search," *Organization Science*, vol. 21, no. 5, pp. 1016–1033, 2010.
- [13] F. M. Belz and W. Baumbach, "Netnography as a method of lead user identification," *Creativity and Innovation Management*, vol. 19, pp. 304–313, 2010.
- [14] M. Spann, H. Ernst, B. Skiera, and J. H. Soll, "Identification of lead users for consumer products via virtual stock markets," *Journal of Product Innovation Management*, vol. 26, no. 3, pp. 322–335, 2009.
- [15] C. S. Stockstrom, R. C. Goduscheit, C. Lüthje, and J. H. Jørgensen, "Identifying valuable users as informants for innovation processes: Comparing the search efficiency of pyramiding and screening," *Research Policy*, vol. 45, no. 2, pp. 507–516, 2016.
- [16] E. Von Hippel, "The dominant role of the user in semiconductor and electronic subassembly process innovation," *IEEE Transactions on Engineering Management*, vol. 24, no. 2, pp. 60–71, 1977.
- [17] K. Boudreau, "Open platform strategies and innovation: granting access vs. devolving control," *Management Science*, vol. 56, no. 10, pp. 1849–1872, 2010.
- [18] N. Franke and S. Shah, "How communities support innovative activities: an exploration of assistance and sharing among endusers," *Research Policy*, vol. 32, no. 1, pp. 157–178, 2003.
- [19] N. Franke and E. Von Hippel, "Satisfying heterogeneous user needs via innovation toolkits: the case of Apache security software," *Research Policy*, vol. 32, no. 7, pp. 1199–1215, 2003.
- [20] L. Aarikka-Stenroos and E. Jaakkola, "Value co-creation in knowledge intensive business services: A dyadic perspective on the joint problem solving process," *Industrial Marketing Management*, vol. 41, no. 1, pp. 15–26, 2012.
- [21] P. D. Morrison, J. H. Roberts, and E. von Hippel, "Determinants of user innovation and innovation sharing in a local market," *Management Science*, vol. 46, no. 12, pp. 1513–1527, 2000.
- [22] C. Lüthje, "Characteristics of innovating users in a consumer goods field: An empirical study of sport-related product consumers," *Technovation*, vol. 24, no. 9, pp. 683–695, 2004.
- [23] P. D. Morrison, J. H. Roberts, and D. F. Midgley, "The nature of lead users and measurement of leading edge status," *Research Policy*, vol. 33, no. 2, pp. 351–362, 2004.
- [24] A. K. Chatterji and K. R. Fabrizio, "Using users: When does external knowledge enhance corporate product innovation?" *Strategic Management Journal*, vol. 35, no. 10, pp. 1427–1445, 2014.

- [25] W. D. Hoyer, R. Chandy, M. Dorotic, M. Krafft, and S. S. Singh, "Consumer cocreation in new product development," *Journal of Service Research*, vol. 13, no. 3, pp. 283–296, 2010.
- [26] E. L. Olson and G. Bakke, "Implementing the lead user method in a high technology firm: A longitudinal study of intentions versus actions," *Journal of Product Innovation Management*, vol. 18, no. 6, pp. 388–395, 2001.
- [27] M. Schreier, S. Oberhauser, and R. Prügl, "Lead users and the adoption and diffusion of new products: Insights from two extreme sports communities," *Marketing Letters*, vol. 18, no. 1-2, pp. 15–30, 2007.





International Journal of Mathematics and Mathematical Sciences





Applied Mathematics

Hindawi

Submit your manuscripts at www.hindawi.com



The Scientific World Journal



Journal of Probability and Statistics







International Journal of Engineering Mathematics

Complex Analysis

International Journal of Stochastic Analysis



Advances in Numerical Analysis



Mathematics



Mathematical Problems in Engineering



Journal of **Function Spaces**



International Journal of **Differential Equations**



Abstract and Applied Analysis



Discrete Dynamics in Nature and Society



Advances in Mathematical Physics