

## Research on the impact factors of new energy automobile sales under the background of big data Era

Junting Wu<sup>1, a</sup>, Yu Guo<sup>2, b</sup> and Kewei Yang<sup>2, c</sup>

<sup>1</sup>National University of Defense Technology, Changsha 410073, China;

<sup>2</sup>National University of Defense Technology, Changsha 410073, China.

<sup>a</sup>wujunting10@163.com, <sup>b</sup>329092502@126.com, <sup>c</sup>kayyang27@nudt.edu.cn

**Keywords:** Big Data; New Energy; Automobile.

**Abstract.** At present, most research about the impact factor of new energy automobile sales is qualitative. There is controversy over different influence of each factor. To some extent hindered the market promotion of new energy automobile. This paper utilizes Internet search data to analyze the impact factors of new energy automobile sales in the era of big data. Based on the behavior process of consumers who purchased new energy automobile, we firstly use Baidu search engine data to calculate the corresponding search index. Then we use the keyword search index and related data of new energy automobile sales as a basis, to compare the domestic market sales and quantitatively analyze the influence factors. Lastly we provide some guides and suggestions targeting at the market promotion of new energy automobile

### Introduction

According to the objective situation of China's energy and environmental protection, the development of new energy automobile is not only necessary, but also a rare historic opportunity for our country to create a completely independent brand automobile. 2014 is known as the first year of China's new energy automobile business commercialization, but annual sales of new energy automobile were only 8.39 million, accounted for less than 0.0036% of the annual car sales, and most of them for the government to promote the field of public passenger cars, but sales of private new energy vehicles were less than 30 thousand[1]. Currently, despite the government departments, automobile manufacturers and related academic departments strongly support the use of new energy vehicles, however, most of the domestic research limited on China's new energy automobile marketing obstacles and qualitative analysis on preferential policies, lack of empirical research on the consumer psychology and consumer behavior of new energy automobile, only a few of articles also focused in the use of the traditional questionnaire model.

With the rapid development of network technology and the popularity of the Internet, a traditional way of access information has been completely reversed. The Internet become one of the quickest way to get materials and information, and consumers preferred to search related data and information from the Internet. According to the statistics of the China Internet network information center, about 90% of Internet users use search engines to obtain relevant information by the end of 2015[2]. Due to Internet search data recorded hundreds of millions of attention and demand, it reflects the trend and rules of main market players and provides necessary data for the research of social and economic behavior. Before buying a car new energy automobile, consumers will obtain the data and information via the related information channels. For example: consumers search data and information about the car in the Internet, and a search engine as a search information entry, when the car consumers search information at the same time the search engine will search traces recorded and saved into the database, such as the automotive consumers search the type, performance, price and description and so on. These records are not only the objective reflection of the social economy behavior of consumers, at the same time can also provide the data source to study the scientific research of social economic behavior in the whole society. The application of the new ideas and new technology makes the research process and result more objective and fair, and also makes the

empirical research method has high feasibility and rationality. Therefore, doing the quantitative research on consumers' willingness to buy new energy automobile and decision factors which influence consumers to purchase new energy vehicles under the background of big data is very necessary to solve the problem of the new energy vehicles in China's auto market weakness[3].

## **Analysis of the shopping behavior process of new energy automobile consumers in the Internet Age**

For car consumers, buying a car is a long term plan. According to the general car consumer psychological demands, during the time that between consumers having the purchase tendency and deciding to purchase, the consumers will through various channels to learn about the information of purchasing the car. For example: car brand, performance and economy, fuel consumption, after-sales service and so on. In the past, the source of information for consumers is mainly rely on newspapers, television advertising, books and other channels[4]. The emergence of the Internet, especially search engine, has broken the traditional buying habits. The search engine provides comprehensive and efficient search data and information for automotive consumers. The car consumers according to the data and information searched to support traditional buying behavior.

Based on the analysis above, we use the following two aspects as ideological basis of paper's concepts framework of theory. In the first place, in order to find relevant information to facilitate subsequent purchase, auto consumers use search engine as the most appropriate approach to get data and information. The Internet is the best way for automotive consumers to access to relevant information[5]. For example, consumers will search for related information, generally like the brand of the car, model, pictures, cost performance, address of auto 4s shops as well as the evaluation of other consumers after they intend to buy one. The search engine is car of the most commonly used to obtain information from the Internet information retrieval tools. Therefore, auto consumers tend to search for related information at first, and then select the most appropriate 4s shop to test drive and purchase. This has formed an intangible relation and correlation between car sales volume and search data, which is the determination of the logic prediction approach we discussed later within the paper. On the other hand, most of auto consumers will upload information and data related to their car in the BBS, blog, space and other network expression after they finished purchasing[6]. These will be the source of search engine and provide reference information for subsequent car consumers.

To sum up, consumers should have consuming intentions at first, that is consume requirement. Then they could search for related information, such as internet search, talk with friends or other users, then go to convenient auto 4s shops to test, after that then purchase, share car's information, at last, this information become information source, cyclic behavior process already contains a complete information chain.

## **Research methods**

### **Introduction of Baidu index**

Baidu index, also known as big data sharing and exploration platform, refers to that user attention of a keyword and network exposure rate within 30 days, also a description and reflect on key words change development trend, it is based on the mass data, from two dimensions which are "degree of user attention" and "degree of media attention " to describe the change trend of keywords. In general, Baidu Index has a updating cycle of 24 hours, and according to the different time periods to provide different data search volume for users. For example: "within one month", "within one quarter" and all of the data search volume. In addition, the Baidu index is just a reference value, which is based on the "degree of user attention" and "degree of media attention "this two dimensions in the last 24 hours for the comprehensive reflection of keywords.

## Introduction of search index aggregation

### Process of search index aggregation

The aggregation process of search index is mainly divided into two aspects, determine the correlation coefficient and aggregate search index, the following two aspects are introduced.

#### (1) calculation of correlation coefficient

This paper applies the time difference correlation analysis method to calculate the leading order number and the correlation, the time difference correlation analysis method is an important and sensitive method, which reflect current economic index of economic activities as the benchmark index. This method uses the correlation coefficient to validate economic time series in advance, consistent or lag relationship. The mathematical principle of the time difference correlation coefficient is as follows:

Set  $Y = \{Y_1, Y_2, Y_3, \dots, Y_n\}$  as research data,  $X = \{X_1, X_2, X_3, \dots, X_n\}$  as selected indicators as,  $\gamma_t$  as time difference correlation coefficient, then:

$$\gamma_t = \frac{\sum_{i=1}^n (X_{i-t} - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_{i-t} - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

$$t = \pm 1, \pm 2, \pm 3, \dots, \pm T$$

In this research, The meaning of the variables represented in the above formula is :  $\gamma_t$  represents the time correlation coefficient, Y represents the monthly sales data of new energy automobile,  $\bar{Y}$  as means; X indicate monthly search volume data of Keyword search ,  $\bar{X}$  as means; t is the largest time difference correlation for X, this coefficient represents the correlation between the two.

#### (2) network search index aggregation

The content of network search index synthesis mainly consists of the calculation of the time difference correlation coefficient, Determine the leading order and aggregation step by step.

The aggregation steps are as follows:

a. First, determine the data lag period, and then calculate the time difference correlation coefficient between each keyword and the research data in order to calculate the correlation coefficient in the leading orders;

b. To align each keyword dislocated according to Lead orders and research data, determine its weight according to the size of the time difference correlation coefficient. And select the best of the time difference correlation coefficient as the reference index;

c. according to the sequence keywords to aggregation, which this sequence dislocate alignment by leading order, calculate the time difference of previous step correlation coefficient between the benchmark index and the research data, marked as  $W_1$ , and on the dislocation alignment for the original keyword sequence add the corresponding benchmark index respectively, then calculate the time difference correlation coefficient between the benchmark index and the research data respectively, from them elect the biggest difference correlation coefficient, marked as  $W_2$ ;

d. judge whether  $\Delta W = W_2 - W_1$  is equal to a value zero, if  $\Delta W$  greater than zero, let  $w_2$  replace  $W_1$  as new benchmark index  $W_1$ , then repeat the previous work, if  $\Delta W$  is equal to a zero, this indicates that the addition of new keywords data no longer increases the fitting degree between key words and research data, network search index aggregation process is over.

## **Empirical research**

### **Data source**

The new energy car sales data adopted by the Paper are derived from the auto channel SoHu website (<http://db.auto.sohu.com/cxdata/>), and the time range is between January 2007 and October 2007 (in months).

The web search data adopted by the Paper are derived from Baidu index, and the time range is January 1, 2007 to October 31, 2015 (in days). Because Baidu index is based on super huge amounts of data, from the "customer focus" and "media attention" two dimensions to describe the change trend of keywords. In general, Baidu index in 24 hours for an update cycle, and according to the different period of time to provide customers with different data searches. For example: "within one month", "1 quarter", and all the data searches. In order to be able to better with new energy vehicles in sales data for fitting, we need to aggregate the Baidu keyword search data by the day, and then we get the search data monthly.

### **Network search data selection**

Through the analysis of the consumer shopping model and the automobile consumers shopping process, we know the information search has played a vital role in the process of consumer shopping. So, the basis and key of this study are relevant keyword search and selection, keywords search and selection in paper are divided into the following three steps.

First of all, according to the different environment, we divided into macroscopic environment and microcosmic environment here. On the macro environment, under the national macroeconomic regulation and control policy, economic level and the status quo and future development direction of new energy automobile market as the guidance, we artificially select relevant keywords, such as: macroeconomic regulation and control policy, new energy automotive news, economic index, the index of residents' consumption, etc.; Look from the microcosmic environment, the initial keywords are acquired according to the information need of brand, model number, price and description, performance, fuel consumption, after-sales service and consumer evaluation of new energy vehicles

Secondly, on the basis of the initial keywords, we integrated automatic recommendation technology of Baidu index and artificial selection, got 4289 keywords, and the comprehensive keywords are 124.

Finally, according to the actual needs of paper research, we cut out the keywords with small amount data and the keywords repeated, left keywords according to the properties can be divided into: the safety protective, after-sales service, use convenience, quality, reliability, energy costs, sales prices, government related preferential policies, the highest speed and limited driving distance, auto makers brand influence, exterior trim and the surrounding crowd influence on their buying these eleven categories, to form the keywords library of new energy vehicles.

### **Network search index aggregation**

The second part elaborates the process and method of network search index aggregation, with the key word database, this paper applies stepwise dislocation aggregation method to calculate the time difference correlation coefficient between each keyword in keyword database and new energy car sales. As we all know, purchasing a car is an important economic decision, there is a certain lag existing between information search and car purchase. Similarly, in order to guarantee the feasibility of the research results and the accuracy and stability of data, in this paper, the correlation coefficient of 0-11 month is calculated for the longest lag period of 12 months, maximum correlation is obtained at the lag period of seven months. Therefore, we delete the data between 7 to 11 months during the lag period and keep the data from 0 to 6 months of the lag period.

According to the fit of the new energy automobile sales, all the keywords by categories to carry out statistical operations, correlation coefficient of various keywords' statistical calculations and lag order number range shown in the following table 1.



Table 1. The calculation results.

Keywords category	correlation coefficient				lag order number
	means	minimum	maximum	standard deviation	
Safety protection	0.9642	0.8735	0.9842	0.4106	1
After sales service	0.9247	0.8425	0.9802	0.4957	0
Use convenience	0.8977	0.8203	0.9751	0.6310	3
Quality reliability	0.8619	0.7632	0.9312	0.7362	2
The cost of energy consumption	0.8459	0.7225	0.9321	0.7570	4
Sales price	0.8150	0.7127	0.8730	0.7702	1
Government related preferential policies	0.8080	0.7365	0.8722	0.7942	6
Maximum speed and mileage	0.7996	0.6983	0.8563	0.7649	3
Car manufacturers brand impact	0.7864	0.7069	0.8235	0.8206	2
exterior interior	0.7767	0.6532	0.8427	0.7789	3
surrounding crowd influence on their buying	0.7603	0.6327	0.8245	0.7750	6

## Research conclusions and Countermeasures

### Research conclusions

The average value of the new energy automobile aftermarket service correlation coefficient is 0.9247, second only to the average value of the new energy automobile safety factor, which is a very important factor in the decision of consumers to buy new energy automobiles; The average value of the impact score that car exterior interior and manufacturers brand respectively are 0.77667 and 0.7864, just above the average of surrounding crowd influence on their buying, which is not considered to be an important factor in consumer purchase decisions.

The correlation coefficient of new energy automobile sales price, the highest speed and the driving range is relatively average, the former score is 0.8150, the latter score of 0.7996. It can be seen from the consumer attention to the sales price, the price is a very important variable that decides the consumer whether to buy traditional car or buy new energy automobile, the survey showed 27.4% and 27.7% of respondents were only subjected to new energy automobile sales price with the grade of the traditional automobile higher within 1 to 3 million and 3 million to 5 million, otherwise they will choose to purchase conventional cars.

New energy vehicle safety protection, convenience and quality of the reliability of the coefficient is larger, with the average score of three points up to 0.9642, 0.8977 and 0.8619. Among them, the new energy vehicle safety protection is the highest score in the whole survey variable, which is the most important deciding factors in the purchase of new energy automobile. In addition, the use of convenience and quality reliability are also an important variable in consumer purchasing decisions. From the scores of these three variables, we can illustrate that the quality of new energy automobile and supporting facilities play an important part in purchase decision-making process, because its quality is directly related to new energy automobile user after the user satisfaction. Poor quality of new energy vehicles may lead not to sell, even if being sold out, the car with bad quality will be not useful, becoming a "problem car" and "trouble car".

Using energy consumption cost correlation coefficient is larger, new energy automobile use energy costs mean score for the 0.8459, indicating that at present, in the condition of China's gasoline and diesel prices stubbornly high, new energy automobile in the low energy consumption is one of the most important factors to attract consumers to buy. In addition, the government introduced preferential policies for consumers to buy the use of new energy automobile will also reduce the cost of new energy automobile to a certain extent.

The impact of the crowd is relatively small to its purchase, the average score is 3.4035, indicating that when consumers making the decision of buying a new energy automobile, the speech of their relatives, friends and colleagues has a certain influence on their purchase.

### **Countermeasures and suggestions**

#### **Cooperation with the media to increase publicity efforts**

From the empirical results of this paper, we can know that safety concerns are the main factors that influence the purchase of new energy automobile. Therefore, new energy automobile production enterprises should strengthen with all kinds of media, especially with the cooperation of the new media, in the promotion of new energy v automobile, not only should they promote automotive quality, but also stress that the excellent safety performance and green environmental protection concept of life, to strengthen the consumer's positive recognition to the new energy automotive.

#### **Improve the after-sale service quality of new energy automobile**

From the empirical results of this paper, it can be known that the second factor that influences consumers to buy new energy automobile is 'after-sales service'. There has a huge technical differences between electric vehicles and traditional cars, but for the reason of it belongs to a new type of products, so the temporary production enterprises have not established a perfect after-sales service mechanism. However, after-sales service is a very important and very complicated section, it contains a series of related content after the new energy automobile sales section include the claim, quality assurance, maintenance service, technical advisory guidance, market information feedback and so on. After-sales service as an important part of the new energy vehicle sales, reflects the business to the customer's humanistic care and emotion, and is the important link between manufacturers and customers. After-sales service mechanism's imperfect, not only will make the actual consumers feel troubled, but also let the potential consumers be terrified by the sight electric cars. Therefore improve the after-service quality of new energy vehicles can promote consumers to buy new energy vehicles.

#### **Popularization construction of complete supporting facilities**

New energy automobile, energy-saving and environmental protection, use much lower cost compared to the traditional fuel vehicle in the late time, take the F3DM of BYD electric cars, if the models take fast charge mode to charge in professional charging, so full of battery capacity of 50% of the electricity need only 10 minutes, but the general household charging equipment with the same power requires about 10 hours. It can also be seen from the results the empirical study of this paper that the convenience of supporting facilities is one of the main factors that restrict consumers to buy new energy vehicles.

#### **All kinds of ways to reduce the purchase cost of new energy vehicles**

Although the country has implemented the purchase of new energy vehicles subsidy policy, but on the view of the current preferential strength, it is more expensive than ordinary fuel vehicles 5-6 million in the end. Although the new energy vehicle is lower than normal fuel vehicle in the late fees, but paying tens of thousands or even hundreds of thousands yuan at one time will greatly weaken the consumer desire to buy. Therefore, reducing the purchase cost of new energy vehicles through the following ways will promote the popularization and use of new energy vehicles.

- (1) Increase government support for new energy vehicles.
- (2) Develop technology to reduce production cost of enterprises.
- (3) Provide financial support from related financial institutions.

#### **Focus on product research and development, improve the user experience**

From the conclusion of this empirical study, we can see that, in addition to the price of the car, the quality and comfort of the car as well as the car's battery life and charging time is also the focus of attention. The empirical analysis of this paper shows that the comprehensive performance of vehicles is an important factor affecting the behavior of consumers purchase a car. Therefore, it is necessary to solve the problems that pure electric vehicles with short endurance mileage and long charging time, and improve the quality and comfort of the car.

### **Clear the new energy vehicles' market positioning**

The results of this paper show that the brand and appearance factors are also the factors that can infect the consumers to buy new energy vehicles, shows that in addition to the car price and performance requirements, consumers are also valued at the car's appearance and reputation. For Chinese independent automobile brand, the improvement of manufacturing process and technology needs a process, and design a car that it's appearance can attractive other's attentions can rapidly enhance their products market competitiveness. At present, the automobile consumption subject is becoming younger and younger, we should take 25 to 35 years old consumers as the main target customers, when we design models we should consider more about the characteristics of this age group. Because the age group belongs to young adults, we can combine with fashionable and dynamic elements when designing the shape of the car, to attract consumers which belongs to this age groups form the exterior and interior of the car.

### **References**

- [1] Schmidt S V T. A Monthly Consumption Indicator for Germany Based on Internet Search Query Data[J]. *Applied Economics Letters*, 2010, 19(7):683-687.
- [2] Ocampo A J, Chunara R, Brownstein J S. Using search queries for malaria surveillance, Thailand[J]. *Malaria Journal*, 2013, 12(1):23-30.
- [3] Tierney H L R, Pan B. A poisson regression examination of the relationship between website traffic and search engine queries[J]. *Netnomics Economic Research & Electronic Networking*, 2012, 13(3):155-189.
- [4] Cho S, Sohn C H, Jo M W, et al. Correction: Correlation between National Influenza Surveillance Data and Google Trends in South Korea[J]. *Plos One*, 2014, 9(12):1254-1258.
- [5] Chan E H, Sahai V, Conrad C, et al. Using Web Search Query Data to Monitor Dengue Epidemics: A New Model for Neglected Tropical Disease Surveillance[J]. *Plos Neglected Tropical Diseases*, 2011, 5(5):e1206-e1206.
- [6] Kostkova P, Szomszor M, Luis C S. The Experimental Studies on the Shock Wave/Turbulent Boundary Layer Interaction Region Induced by the Blunt Protuberance.[J]. *Journal of the Japan Society for Aeronautical & Spaceences*, 1996, 44(515):735-740.