Novel Model of E-commerce Marketing Based on Big Data Analysis and Processing

Hongsheng Xu*
Henan key Laboratory for Big Data Processing &
Analytics of Electronic Commerce
Luoyang Normal University
Henan LuoYang, China
E-mail: 85660190@qq.com
*The corresponding author

Ganglong Fan
Henan key Laboratory for Big Data Processing &
Analytics of Electronic Commerce
Luoyang Normal University
Henan LuoYang, China
E-mail: 85660190@qq.com

Ke Li
Henan key Laboratory for Big Data Processing &
Analytics of Electronic Commerce
Luoyang Normal University
Henan LuoYang, China
E-mail: 85660190@qq.com

Abstract—**The amount of information on the Internet is getting larger and larger, and the energy of the consumer and the ability to deal with information is limited. Electricity supplier enterprises in the development process to do are to use big data for personalized shopping guide. This paper analyzes the situation of the development of e-commerce industry in the background of big data, and puts forward the improvement method. The paper presents novel model of e-commerce marketing based on big data analysis and processing.**

Keywords-Big data; E-commerce; Network marketing; O2O; Massive data

I. INTRODUCTION

Big data may be heterogeneous data obtained from different physical address on the data source, and or with multi modal characteristics, is the typical representative of social media data are widely distributed on the Internet. Such as micro-blog data contains micro-blog users of gender age occupation and social networking, natural text and image, such as video and audio. For such data, even if the investigation under the same attribute of different physical space objects of the same class, cannot expect its has the same distribution characteristics. The symbols, values, time series, image, text and social network structure of a single mode, equivalence relation, symbolic and numerical characteristics of structural similarity the relationship or order relation based on object classification, can form a covering or nested grain structure.

The era of big data first determined is by data richness. The rise of social networks, a large number of UGC (Internet terminology, meaning the meaning of user generated content), audio, text information, video, pictures and other unstructured data appeared. In addition, the greater the amount of information on the Internet of things, coupled with the mobile Internet can more accurately and quickly collect user information, such as location, life information and other

data. From the amount of data, has entered the era of big data, but now the hardware has been unable to keep up with the pace of data development.

Big data is more in the national economy and the use of marketing. With the development of mobile Internet and intelligent cloud technology, big data has begun to affect more and more business. That is to be able to more effectively on the supply chain, product development, online drainage guide, thereby enhancing the efficiency of the operation of the platform. Big data provides the possibility of e-commerce, and has begun to more and more applications [1]. Key to the success of e-commerce upgrade that is necessary to highlight the traditional doors and windows retail service and experience, but also the ideological front of the Internet into the brand concept, in this context, e-commerce O2O thinking emerged, big data has become a key O2O loop.

Internet era, user habits are changing, and it is only a full understanding of the user to create products in line with user expectations. Compared with traditional retail, e-commerce, the biggest advantage is that everything can be monitored and improved through the data. Through the data can be seen from where the user, how to organize the product can achieve a good conversion rate, the efficiency of advertising and so on. Every bit of change based on data analysis is a little bit more profitable.

The data is a relatively abstract concept, especially when faced with massive data easily confusing, traditional data analysis is more to show, in a simple chart or PPT form is not intuitive, since 2010 the data information map rise, provides visual effects and understanding is very good for data analysis and output, he used a combination of simple graphics will be converted to a single chart connotation rich results, greatly stimulate people's senses, the boring data and more vivid information graph is a manifestation of the further development of data visualization, the era of big data will lead to a lot of a similar method.

Electronic commerce is the first in the network business, is using the Internet data to understand each customer's needs and tendency to provide personalized products and services for them, and then quickly and easily realize the transaction and delivery of products and services to the user as possible, personalized service features, the automation of the relationship for commercial organizations to increase the income and cost, to establish and strengthen relationship with customers.

In the big data era, e-commerce in the economic activities of the operating mode from traditional management into data management mode, performance management process and economic activities of electronic commerce business enterprise data activities, data operation development of the concept of penetration in the enterprise raw material procurement, product manufacturing and product the whole process of marketing activities. The use of electricity in the process by using professional data analysis technology will be able to carry out comprehensive and prediction of certain induction on consumer habits and consumer psychology, and clever use of distribution adjustment of product market supply and demand, and it is in order to satisfy consumer demand of consumers, reduce the cost of production and sales of products. The paper presents novel model of e-commerce marketing based on big data analysis and processing.

II. ANALYSIS ON THE DEVELOPMENT OF E-COMMERCE INDUSTRY UNDER THE BACKGROUND OF BIG DATA

Big data marketing refers to the behavior of a large amount of data collected through the Internet, first of all to help advertisers find the target audience, the advertising content, time and form of anticipation and allocation, and ultimately complete the process of marketing advertising.

Big data can pass the verification and evaluation of massive data, increasing the risk of controllable line and management, timely find and solve the possible risk points, have accurate grasp for the regularity of the risk, will drive demand for more in-depth and thorough analysis of the data of financial institutions. Support business refinement management [2]. Although the bank has a lot of data to pay for water, but the Department is not cross, the data can not be integrated, big data banking model prompted the bank began to effectively use the data deposited. Big data will promote the innovation of financial institutions and service brand, do fine service, are customized to the customer, the use of data analysis and prediction of the new development model, realize the analysis of customer consumption patterns to improve customer conversion rate. Big data is bound to bring more opportunities for financial companies to update the data based business and internal management optimization.

The security and integrity of data is the core elements of the enterprise must consider the elements of e-commerce sites in accordance with the three levels to do this, the first level: the basic hardware. In data storage, data backup, most companies can do. The second level is the key data management, who can see these data. This level is the largest investment in the company, as is shown by equation (1), where p is in the management of data, a lot of data leaks may

be internal reasons. There is some internal staff to see the data he should not see, or get the data he should not get [3].

$$p_0(t) = \frac{u}{\lambda+u} + \frac{\lambda}{\lambda+u} e^{-(\lambda+u)t} \tag{1}$$

This is an era of information explosion, in the face of a large number of consumer choice, the same marketing approach has been difficult to attract consumer's loyal consumption. Therefore, today's retail marketing must come up with the content and form that can really touch the hearts of consumers. We believe that to do this, we must rely on the collection of electricity supplier and store sales data as one of the large database. Through the analysis of these data, we can not only accurately understand each member's age, gender, consumer consumption, consumption frequency, also can analyze their shopping habits and preferences, in shopping malls throughout the residence time and order from.

Found that grain size model (fusion) is a natural big data super large-scale, multi modality, hybrid features, internal logic requirements for solving complex problems and granular computing under the framework of big data. After granulation, the particle pattern of each homogeneous data sample from the perspective of problems can be solved by the existing methods though however, global data on pattern discovery needs to carry on the fusion strategy under the guidance of the heterogeneous data. The particle (corresponding to different feature subspaces of the abstract concept level, and according to the different modes under the framework of data granulation results) on the issue of pattern discovery, as shown in equation (2) [4].

$$u_c(t) = \begin{cases} 0, 0 \leq t \leq c, \\ \frac{1}{4}(t-c)^2, t \geq c. \end{cases} \tag{2}$$

The total time of feature vector B (N)=23P*P(O). Because the N P so when we choose namely the above namely the above, a feature vector of B K-means clustering algorithm, the total time running by (3n)O to P (3O), the efficiency of the algorithm is improved obviously. This low complexity algorithm is well suited for large data applications.

Spark is an efficient distributed computing system, compared to Hadoop, it is 100 times higher than the performance of Hadoop. Spark provides a higher level of API than Hadoop, the same algorithm implemented in Spark is often only Hadoop of the length of 1/10 or 1/100. Shark similar to the SQL on Spark is a data warehouse in Spark implementation, as is shown by equation(3), where f(t) is in the case of compatibility with Hive, the highest performance can reach one hundred times Hive.

$$f^1(t) = \sum_{i=i_1}^{i_2} b_i^1(t) k_i^1(t) x_i^1(t)$$

$$(3)$$

Big data market for new technologies, new products, new services, and new formats will continue to emerge. In the field of hardware and integrated equipment, big data will have an important impact on the chip, the storage industry, will also generate integrated data storage processing server, memory computing and other markets. In the field of software and services, big data will lead to the rapid processing of data analysis, data mining technology and the development of software products.

Big data not only refers to the massive data, also contains data segments, almost all of the links within the enterprise will be in the form of data to show, for example, efficiency optimization derived time node of the business aspects of the Amazon, in this area has been developed greatly, there will be a report and data processing operations based on a large number of every day operation strategy, marketing strategy, the change is mainly to see the data, automatic replenishment model it defined is the principle of time series and based on extreme value formed, effectively solve completely rely on the order, the artificial replenishment model to improve the efficiency of inventory management.

Electricity supplier is born with a big data aura. Compared to the traditional retail and channel, B2C platform can obtain consumer behavior data, shopping preferences, status, contact information, etc., for the user to accurately identify and locate [5]. On the platform of the third party service providers, such as logistics companies, payment companies, but also contributed to the operating conditions, product service records, consumer reviews, including important data. Do not have the advantage of big data, traditional businesses, on the one hand, self built electronic business platform, with the Internet and cooperation, to create their own customer management or membership system, on the other hand, with the opening of the O2O, such as the use of WIFI technology to create indoor interactive system, thus completing the data collection.

For large data in e-commerce applications can be an example, if the customer wants to buy a Chinese style doors and windows, he usually use search engines. When he entered the doors and windows, Chinese and other keywords, the search engine can display the relevant product information in the forefront of the search. For customers, large data makes him very precise access to the products they want, and Internet companies can analyze the user behavior, more accurately the needs of mining, with the use of large data analysis, a high degree of concern for the further promotion of investment products.

Individual performance, unstructured and semi-structured data, the general characteristic and basic principle is not clear, these are required by multi disciplines including mathematics, economics, sociology, computer science and management science, to study and discuss. Given a semi-structured or unstructured data, such as images, how to convert it into a multi - dimensional data table, object-oriented data model or

directly based on the image data model, as is shown by equation(4).

$$w_{i+1}^1(t+1) = \left(1 - wd_i^1(t)\right) x_i^1(t) - rs_i \alpha N^1(t)$$

$$(4)$$

The use of large data analysis to send personalized EDM. If customers have in the e-commerce website to view a product but did not buy, there are several possibilities: A. stock, B. price is not right, C. do not want to brand or not to goods, just to see if the D. in the customer view the goods out of stock is delivered immediately notify the customer; if there were goods and customers do not buy because the price is likely to be caused by, will inform the customer in the commodity price promotion; at the same time, inform the customer in the introduction and the similar goods or related commodity when warm [6]. In addition, through the excavation of the customer's periodic buying habits, in the vicinity of the customer's purchase cycle timely remind customers.

Shop operators need to analyze a lot of data, including external data, which is the industry's market share; it also reflects the website visits PV; reflect the electricity supplier website sales flow rate, order conversion rate, average order value and so on; of course the most important thing is lying on the Bank of the digital.

III. NOVEL MODEL OF E-COMMERCE MARKETING BASED ON BIG DATA ANALYSIS AND PROCESSING

Electricity supplier big data collection, integration, analysis is a professional and rigorous method system, the traditional method is difficult to do, which requires the use of sophisticated computer software systems. Take Shanghai far Fang multi-user mall, its independent research and development of electricity supplier transaction data analysis module can be a very good statistical analysis of business data.

The emergence of the Internet and the development of related technologies make it possible to collect and analyze massive data. The characteristics of the Internet has led to the spread of these data can be high speed and large capacity [7]. The Internet introduces a pattern of data generated by the user. This model is characterized by multiple sources, low cost, and timelier. Of course, the authenticity and reliability of these data need to be certified. One of the advantages of building e-commerce based on the Internet and traditional retail is the availability of data. E-commerce can get the source of the customer's visit in real time, the search, collection, purchase behavior in the website, and the relevance of the goods purchased. These data can help companies more accurate customer service.

Application of electricity supplier data broadly divided into the following steps: A. data collection, verification and filtering; B. classification and stored in the data warehouse; C. data mining is to find association rules and data of the data between the D. data model and parameter establishment; adjustment; e. application development and decision support based on data.

The module consists of three parts: system statistics, sales data analysis, operational data analysis. Companies can easily access to the day of the site's traffic, website products, news and other published articles, online members, online visitors, etc.. As well as on-line members, and it is the current online membership calculation and the proportion of the total membership. At the same time, the module has the function of historical statistics, statistics of the past traffic data, with the trend chart in the form of display. The most important point is that the system can be presented for the mall, shop sales details, commodity classification statistical analysis, as is shown by equation (5) [8].

$$\frac{f'(x_1)}{2x_1} = (b^2 + a^2)\frac{f'(x_2)}{4x_2^3} = \frac{\ln\frac{b}{a}}{b^2 - a^2}x^3 f'(x_3) \tag{5}$$

Big data is a kind of artificial nature has the hidden law, searched for scientific mode of big data will bring a general method to study the beauty of big data, although this exploration is very difficult, but if we find the unstructured and semi-structured data conversion method of structured data, data mining methods known will become a major tool for data mining.

Big data is not in the computer CRM system and various forms, really big data is reflected in consumer behavior and feelings, all the bits exist in the form of Internet information, its basic features are: mass, high speed, flexibility, diversity! For example, you decide to buy is not the business of "pretty, but the evaluation of other buyers, the seller's credit rating, this is the big data; such as unknown problem, Baidu to search, today if there are a lot of people search for" how to treat colds, can predict ten days there will be a flu outbreak. This is called big data.

Using the data of the biggest weakness is the lack of relevance to grasp, once isolated data considerations, most core elements may be missing or not accurate and comprehensive expression of e-commerce internal information transfer can be transformed into data, by operating on the basis of the data association will become the basis for data analysis. Using the data of multi dimension and multi angle of view, through a core dimension of data will be gradually expanded, the reason will be a behavior and rationality through more than a dozen or more standard data to show, make it more accurate and focused.

The Internet is a National People's Congress in the age of media data, people through the computer and it together, not a historical product of large data is not because of strict environmental control and content, which fully reflects the dynamic data it is timely developed on the basis of the Internet content, big data information can be generated in any time is dynamic not only in the data collection process, data processing and data preservation technology are constantly changing, so it can be said that the facility is dynamic data processing.

According to statistics, 82% of the electricity supplier is being challenged to deal with massive amounts of information, and they spend a lot of time to study it, and 89% of the electricity supplier because of overload processing data and lost sales opportunities [9]. Just sitting on big data is not enough, the big data analysis and mining capabilities has become the core competitiveness of the electricity supplier. This shows that the key is not the number of large data in the number of raw materials, but the ability to process data, which can make the real value of big data.

Under the direction of big data products include: crowd, product features, and unique selling points (customized). Large number of products must be selected to do the STP analysis, the product must meet the following characteristics: 1) product profit space, for example, now sells stockings on Taobao, basically no profit margins. What do you do, three kinds of methods: one is to sell special products, for example, can sell sell sexy stockings at the speed, through the crowd of foreigners love China taste stockings. Two is engaged in wholesalers in Alibaba, Taobao market has been saturated, in doing this product is dead. The three is to increase product features, stockings with a special function of burning belly.

When the number of clusters is larger, that is, M is larger, the number of samples of the new training sample I NewX will be much less than the original sample number, it is easy to achieve a significant reduction in the amount of KNN calculation, improve the speed of classification purposes. However, with the increase of M, the overhead of clustering will increase, and the number of samples in the new training sample I NewX will gradually decrease, which may lead to the decrease of classification accuracy. So in order to avoid this situation, the number of clusters m need to be set in a more reasonable value, as shown in the following formula [10].

$$p_1 = \rho p_0, p_2 = \frac{\rho^2}{2!}p_0, \cdots, p_n = \frac{\rho^n}{n!}p_0, p_{n+1} = \frac{\rho^{n+1}}{n \cdot n!}p_0,$$

$$p_{n+2} = \frac{\rho^{n+2}}{n^2 \cdot n!}p_0, \cdots, p_{n+r} = \frac{\rho^{n+r}}{n^r \cdot n!}p_0, \cdots \tag{6}$$

E-commerce marketing precision, as the electronic commerce platform profitable growth of consumer data, collected by the consumers from different channels of data, timely and accurate understanding of the comprehensive information of the customer. Especially the development of mobile smart devices can whenever and wherever possible to provide consumers with related services and products, provides a location data for the mobile users, on the other hand, because there are a large number of data in the intelligent mobile phone, is unique to the individual use, so that consumers one-on-one service possible.

In the case of similar classification time, KNN classification algorithm LSC based on clustering accuracy rate was significantly higher than that of KNN algorithm of random block, and can be found from table 8, the classification accuracy is closer to the classical KNN algorithm. Based on the above analysis, we can see that the number of clusters of the same size of M should be to ensure that all sub clusters can run in memory, as small as possible.

It can not only reduce the clustering time, but also improve the speed of classification, as is shown by equation (7).

$$I(t) = 1 - R_0 - S + \frac{1}{\sigma} \ln(\frac{S}{S_0})$$
(7)

On the importance of enterprise data on the contrary, the massive data used by companies to do the absolute value ratio, add, subtract, multiply and divide method, trend, are the most frequently used methods, data is the abstraction and limitations of all a hideous mess, did not get a breakthrough; there are many reasons for this result, may be different stages of development, may is the lack of human resources, no matter what the reason is, a waste of such an important resource is a great loss to the enterprise, innovation in the field of data need to be improved.

Data audit includes two aspects, on the one hand is the integrity of the audit, as well as his right or wrong, or that he has not been unsafe to use. There may be a data produced, there will be an audit procedure to verify whether the data is correct or wrong, should not be produced at this time. E-commerce sites for big data security, but also the use of some conventional means. For example, it is the separation of internal and external networks and it is a data management process is being set up, including large data audits. When we have just discovered that a data has been tampered with, or illegally used, the audit mechanism will start the alarm.

In the development of big data technology today, more and more e-commerce companies have to use big data to seek their own business development and talent needs. Many colleges and universities are based on the creation of e-commerce curriculum corresponding social development, but in the professional training program design is often biased in favor of the popular professional, lack of electronic business enterprise and the social demand for talent data collection, analysis. The use of big data technology will be able to better provide a more accurate and reliable data basis for personnel training programs. The main data collection requirements, social classification and acquisition of electronic commerce environment in general post supply demand; electronic business enterprise of professional talents, skills and quality requirements; school teachers professional settings, and other data; student information, interest, skill biased data.

IV. SUMMARY

Big data has changed the traditional mode of management decision-making structure. The study of the impact of big data on management decision making will be an open research question. In addition, the change of the decision-making structure requires people to explore how to support the higher level of decision-making and do the two mining". No matter what kind of data heterogeneity big data, big data in the rough knowledge can still be seen as a mining category. It is very necessary to find the bridge between the heterogeneity of data and the heterogeneity of decision making by looking for the "knowledge" generated by the two

mining. How to change the decision structure under the condition of big data is equivalent to the study of how to make the decision maker's subjective knowledge in the process of decision making.

The paper presents novel model of e-commerce marketing based on big data analysis and processing. E-commerce marketing, has been changing with the spread of the way, from the traditional to the network, and then from the network to social, and then from PC to mobile, each change is profound. Every technological progress also brings the progress of e-commerce industry! The main characteristics of e-commerce network marketing in the era of big data is accurate demand forecasting, mining, interpersonal orientation, directional communication, and ultimately by the user to grasp the decision-making power and the two transmission right, rather than through its own e-commerce business. Therefore, e-commerce network will usher in the era of big data interaction.

REFERENCES

[1] Sirivimol Thanchalatudom, Namfon Assawamekin, Using Big Data Technology for Information Management in Hybrid Learning System, *RNIS*, Vol. 12, pp. 179 -182, 2013.

[2] Hyoung woo Park, Il Yeon Yeo, Haengjin Jang, Seo-Young Noh, "Simulation based Analysis on Big Data Service Bottleneck for Data Center", JNIT, Vol. 4, No. 8, (2013),pp. 185 – 189.

[3] Hongxin Wan, Yun Peng, "Clustering and Evaluation on Electronic Commerce Customers Based on Fuzzy Set", IJACT, Vol. 5, No. 3, pp. 199 ~ 206, 2013.

[4] Wei Li, Hongtu Zhang, Tingting An, "Optimal Decision-Making in E-Commerce Platform Based On Optimal Stopping Theory", JCIT, Vol. 8, No. 8, pp. 922 ~ 929, 2013.

[5] Dawei Sun, Ge Fu, Xinran Liu, Hong Zhang, "Optimizing Data Stream Graph for Big Data Stream Computing in Cloud Datacenter Environments", IJACT, Vol. 6, No. 5, pp. 53 ~ 65, 2014.

[6] Shoupu Wan, Yongjia Wang, Michael Recce, "A Case Study of the Internet Dynamics -From Big Data to Marketing Insights", IJIIP, Vol. 5, No. 2, (2014), pp. 39-52.

[7] JinHui Lei, XianFeng Yang, "Construction the E-commerce Trading Platform Based on Rough Set Data Mining Technology", JCIT, Vol. 8, No. 3, pp. 460 ~ 469, 2013.

[8] Sun Chengshuang, Shen Liyin, Gan Lin, "An Analysis on Competitiveness Enhancement by Applying E-commerce to Construction Corporations", JDCTA, Vol. 7, No. 8, pp. 772 ~ 780, 2013.

[9] Fan Yang, "The Research of E-Commerce Recommendation System Based on Cloud Computing", IJACT, Vol. 4, No. 16, pp. 256 ~ 263, 2012.

[10] Yanqing Lv, Jianmin Gao, Zhiyong Gao and Hongquan Jiang, "Multifractal information fusion based condition diagnosis for process complex", Process Mechanical Engineering, pp.1-8. (2012),