

# Brief Analysis on the Application of Big Data in Internet Financial Risk Control

Hu Jiabin

Accounting and Finance, Hong Kong Polytechnic University, 999077

**Keywords:** big data; Internet; financial risk control; application

**Abstract:** With the rapid development of information technology, the Internet has been applied to all walks of life. The integration of big data in Internet finance has resulted in the rapid development of Internet finance. The application of big data to Internet finance has played an important role in controlling Internet financial risks, which has improved the accuracy of Internet financial risk control systems, and improved the services of Internet finance. This paper briefly summarizes big data technology and Internet finance, analyzes the specific application of big data in Internet finance and clarifies the principle of Internet financial risk control, and the specific application of big data in Internet financial risk control is obtained finally.

## 1. Introduction

Big data technology was first proposed by Silicon Graphics, which can accurately find valuable information in a large amount of information through the analysis of data, and then the valuable information can be extracted to achieve secondary utilization that can increase the effective utilization rate of information [1]. The big data technology system can be divided into many aspects, including storage and management, privacy and security, acquisition and preprocessing, computing models and systems. Big data technology can effectively optimize data processing, and quickly find the main information points in massive data, to improve the efficiency of data processing [2]. The main characteristics of big data technology is that the amount of data is very large, resulting in the unpredictability and the wide range of the processed data, which is involved all aspects. Besides, the value density of big data technology is low and the processing speed is high, based on which a large amount of data can be used to optimizing the use of data quickly, to shorten the data processing time greatly and improve the work efficiency, satisfying the needs of modern fast-paced life.

At present, with the rapid development of information technology, the big data technology has been applied everywhere we study, live and work. For example, the cumulative lending and transaction growth rates of microfinance in the Alibaba Group are constantly rising, based on which it can be found that the use of big data for financial risk transactions not only can effectively increase the number of transactions and complete a wide range of transactions, but also can reduce transaction costs, to promote the rapid development of Internet finance [3]. However, there are certain risks and hidden dangers, and the user's personal information can be leaked in the course of the transaction, so that the user's information security cannot be protected, and the personal property will be subject to great threats once hacked. Besides, the application of the big data to the Internet finance can also play an important role in restricting the management decisions made by enterprises.

## 2. The Characteristics of Internet Finance

The transaction cost of Internet finance is much less than that of actual financial transactions. The transaction cost refers to the cost incurred by both sides of the fund supply in the process of exchange of funds. Utilizing the Internet to carry on capital transactions can greatly reduce the input costs of human and material resources to reduce the resource consumption [4]. In addition, the information issue depends on the Internet: the product information release on the Internet in place of traditional TV commercials or leaflets, and the process of pricing and trading on the Internet can effectively reduce capital investment and transaction cost. Based on the internet finance, the transaction can be completed simply and quickly through computer operation, which are not limited

by factors such as time and space, and can greatly improve the transaction efficiency, to implement efficient internet financial transaction.

The transactional data generated by Internet finance is of great value. With the popularization and ripeness of science and technology, the society has entered into the information era. With the continuous advancement of e-commerce and logistics technology, people can use the Internet to purchase what they want without leaving home. As shown in Table 1, the application of Internet finance has become more and more widespread.

Tab.1 The Transaction Amount and Year-on-Year Growth Rate of Internet Finance

Year	2010	2011	2012	2013	2014	2015	2016	2017
Amount (billion ¥)	5.1	8.4	12.4	17.2	23.3	31.2	57.9	102
year-on-year growth rate	—	64.7%	47.6%	38.7%	35.5%	41.8%	85.6%	74%

Based on the analysis of these information and data, enterprise managers can find out consumption trend, which is a kind of valuable resource for advancing the development of enterprises [5].

However, everything has its pros and cons, there exists some inevitable potential risks in the rapid development of the Internet. In recent years, there has been a variety of online payment methods including WeChat wallet, Alipay and online banking [6]. At present, the governmental supervision of Internet finance is relatively weak, and some lawbreakers raise funds by Internet finance. Besides, the phenomenon of online fraud emerges in endlessly, showing that the high risk of the Internet finance greatly affects the security of Internet finance.

### 3. Principles of Internet Financial Risk Control

As with all risk controls, the most effective way to control Internet financial risk is to diversify risk: it is difficult to obtain sufficient risk resistance ability when putting risks together, but the way of spreading a large risk into several small risks is effective. For example, if a dozen of eggs is packed in a basket, all of eggs in the whole basket will break once the basket overturned; but if the eggs are spread in different baskets, the number of broken eggs will be much less than that in the case of all eggs put in one basket, which proves the importance of "distribution" and "minority" in Internet finance.

The borrowers' academic record, age, economic level, repayment ability and personal credit are different, and so it is the area where borrowers stay. The borrowers may be in breach of contract, but it occurs independently and there is no correlation between borrowers. Therefore, the probability that these decentralized individual borrowers break contract is relatively independent, and the probability that borrowers break contract simultaneously is small. For example, if there are 100 independent individuals applying for a loan, the probability of a single loan user defaulting is 15%; if two loan users are arbitrarily selected, the probability of simultaneous defaulting is 3%; if four individuals are selected at the same time, the probability that the two loan users default at the same time is 6%.

If there is some kind of connection between borrowers, the borrower B will cancel loan once borrower A cancels loan, and then the probability of defaulting will change from 3% to 7.5% (as shown in Table 2), which is far greater than that in the case of no connection between borrowers. Therefore, in order to reduce risks in making loans, it is necessary to ensure that the individual loan entities are independent, to reduce the correlation between loan individuals

Tab.2 Default Risk in Different Cases

Number of People	Default risk (no correlation between individuals)	Default risk (there is correlation between individuals)
2	3%	7.5%
4	6%	15%
6	9%	22.5%

Small amount is aimed at avoiding small sample deviations. For example, there is a lending company with a loan of 1 billion ¥, on the basis of statistics rules, if there are 10,000 customers at the current stage, each customer can borrow 100,000 ¥; but if there are 1000 customers at this stage, then each customer only can borrow 1 million ¥ (as shown in Table 3). From this table, it can be seen that the more borrowers, the less the amount of loans each person can acquire. In other words, the greater the number of samples, it is more consistent with the law of normal distribution. Therefore, if the default rate is 2%, and then the default rate of 10,000 customers will be higher than that of 1000 customers.

Tab.3 The Relationship Between Number of Customer and Amount of Loan

Number of Customer	10000	1000	100
Amount of Loan (ten thousand ¥)	10	100	1000

Small share is easy to result in too much decentralized borrowing and excessive borrowing customers [7]. The traditional credit approval model does not meet the characteristics of small-share borrowers that the individual differences among small-share borrowers are large, and the willingness and repayment ability of many borrowers cannot be effectively measured, which can greatly increase the cost of risk control.

It is required to change the credit approval model that using P2P to promote interest-bearing liabilities (see Figure 1) and establish a risk control model, to understand customer's general characteristics and credibility by observing the behavioral characteristics of customers, which is beneficial to determine the size of the customer's default risk. According to this credit audit model, the cost of manual auditing can be greatly reduced, and some errors in the process of judging standards and manual auditing can be reduced at the same time. The core method of the risk control system is to study the possibility of default by different individual users, and to follow the principle of small share at the same time [4].

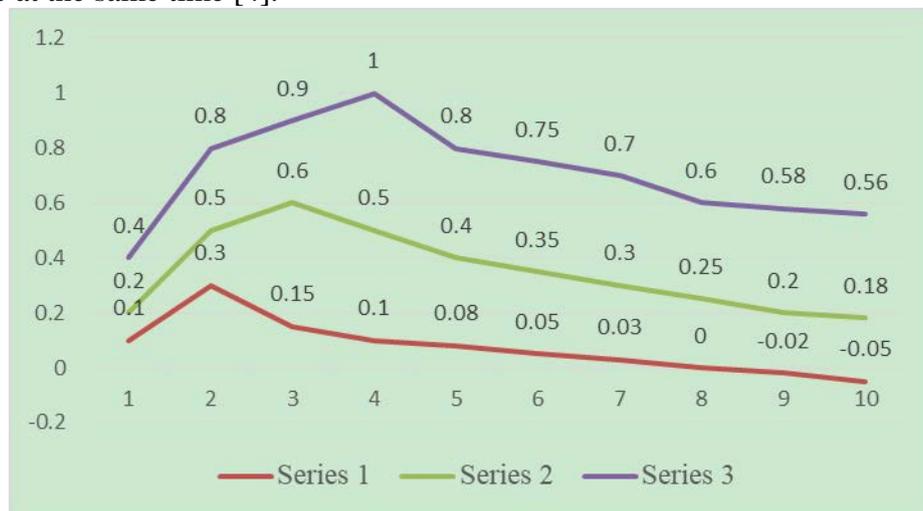


Fig.1 The Interest-Bearing Liabilities Promotion Based on P2P

#### 4. The Application of Big Data in Internet Finance

With the gradual development of big data technology, it has been gradually applied to Internet finance in recent years. On this occasion, two algorithms of high-frequency trading and algorithmic trading are currently used.

The high-frequency trading refers to the trading with high frequency and high efficiency, in which a trader can complete an intelligent process by using a series of intelligent software systems such as trading programs and software facilities to comprehensively analyze, collate, calculate,

generate, and send transaction instructions for big data. This highly efficient approach can increase the speed of the entire transaction process, increase the number of transactions, and maximize the number of transactions in a shorter period of time, to obtain high profits. Besides, this high-frequency algorithm can also be used to analyze big data, to identify and analyze the trading traces of specific participants as well as gain certain profits. The risk and technical requirement for algorithmic trading is pretty high.

The algorithm identification used for the big data under certain requirements can determine the trend of funds, and the transaction time can be judged by algorithm to obtain profits [8]. Taking the fund for example, the large data of the fund order is always generated in the last second before fund closing. At the moment, the algorithm transaction must identify the pattern algorithm of fund, to predict the funding direction of this fund in other time. At present, the Alipay transactions occupy the top position in the Internet finance (as shown in Figure 2), and the Alipay accounts for 49% of the transaction share, the Tenpay accounts for 40%, and the other payment method accounts for 11%.

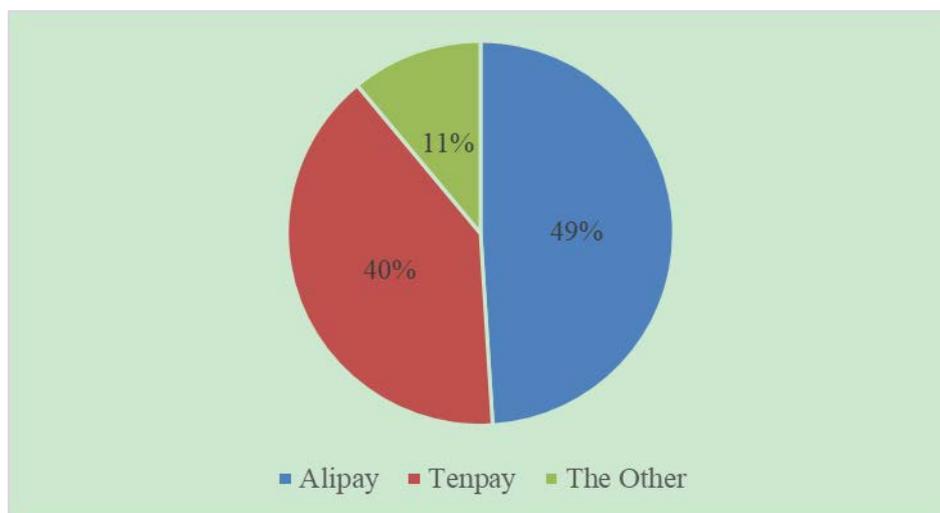


Fig.2 The Market Share of Third-Party Payment

In recent years, the hedge funds have further developed algorithmic transactions by extracting market sentiment information from various social media (Facebook, websites, blogs, etc.). This type of information can be used to make various emergency decisions in the nick of time: once the unforeseeable factors such as terrorist attacks, hacking, and natural disasters occur, the orders can be automatically generated by the information to determine the trading strategy for funds, which is aimed to minimize losses and protect corporate interests greatly. This information method was widely used many years ago, and some people are even willing to give a long price for big data information on social media such as WeChat, Weibo, and chat rooms. Through the big data information, the sentiment analysis on enterprises can be conducted, which can be used to establish user information and determine fund trading strategies further.

The aforementioned approach is not a simple data research, which can be adjusted and improved through the response of big data. According to market data, the market return rate of this approach is very high, and therefore it can be said that the market sentiment analysis is a kind of mature information analysis method.

Whether in the ancient time or the modern society, there are always certain risks associated with any transaction, and it seems there is a natural law that before obtaining revenue, the one has to know how to accept the risk. The law is no exception for any industry, especially for financial institutions.

The financial industry is filled with many uncertainties, which experiences drastic fluctuations every day, so it is necessary to learn that how to prevent risks while enjoying benefits.

In the face of risks, it is necessary to make a variety of countermeasures in advance to manage the risks strictly, to minimize losses and ensure the interests of enterprises. The risk management is mainly based on the collection and analysis of big data involved in small data users' transaction

requirements, transaction volume, transaction time and transaction methods, to prejudge various possible risk situations. Besides, the big data can be used to conduct accurate and detailed analysis for the demand, budget and cost of small users, as well as the development prospect for the entire industry, which is helpful for reducing risks.

The method of pre-managing risks timely is favorable for the business operation of SME (small medium enterprise), and the risk management ability can be improved through various innovative modes.

## **5. The Application of Big Data in Internet Financial Risk Control**

The China's financial system is still under construction and improvement, and the number of users in China is large, the users' data is inaccurate and the data source is relatively wide, which greatly increases the difficulty of applying big data to Internet finance. At this stage, various types of risk control systems are gradually formed in the Internet finance. Some large companies have established credit evaluation systems based on big data technology, and small companies can also share the information generated by these systems. The central bank's credit system incorporates the user's identity review and the bank's data display to provide user credit rating reports and credit inquiry results. However, the abovementioned system only works for banks, which not targets Internet finance companies, and the users' personal credit situation cannot be reflected in this system.

The Micro-credit provided by Alibaba Company has changed the traditional offline loan model, based on which an online loan method is established by combining loans with big computer data. The online loan method refers that the entire process from the beginning of the auditing of funds to the completion of the loan can be completed online, which can grasp the risk management of the entire process clearly and is conducive to reducing risks as well as provide guaranteed personalized micro-credit manner for vulnerable groups.

The data analysis system established by Alibaba is relatively early, and many people use the data reflected by Taobao and Tmall as the raw materials for data analysis. Then the payment and sales data of various shops can be combined to rate the Internet financial credit based on the online scoring model. In addition, the credit card websites can also provide valuable big data for Internet financial risk control. For example, the users' credit rating can be examined by the repayment status, type, limit, and the processed time of iQiyi credit card. The water, electricity and gas payment in Taobao are also important data references that reflects the user's basic information.

## **6. Conclusion**

With the development of society and the advancement of information technology, the Internet finance will be more and more consummate, and an Internet financial risk control system can be established relying on big data technology. Through this study, it can be seen that the Big Data can be used to conduct algorithmic transactions and high-frequency transactions in Internet finance, make the analysis of market sentiment and improve the risk management ability. As a result, it is demonstrated that the risk management occupies an important position in Internet finance, and it is necessary to combine the actual characteristics of small-scale users and small transaction amounts, to continuously optimize Internet finance risk control model based on the big data technology.

## **References**

- [1] Yin S, Kaynak O. Big data for modern industry: challenges and trends [point of view] [J]. *Proceedings of the IEEE*, 2015, 103(2): 143-146.
- [2] Jagadish H V, Gehrke J, Labrinidis A, et al. Big data and its technical challenges[J]. *Communications of the ACM*, 2014, 57(7): 86-94.
- [3] Wang G, Gunasekaran A, Ngai E W T, et al. Big data analytics in logistics and supply chain

management: Certain investigations for research and applications[J]. International Journal of Production Economics, 2016, 176: 98-110.

[4] Choi T M, Chan H K, Yue X. Recent development in big data analytics for business operations and risk management[J]. IEEE transactions on cybernetics, 2017, 47(1): 81-92.

[5] Lee I, Lee K. The Internet of Things (IoT): Applications, investments, and challenges for enterprises[J]. Business Horizons, 2015, 58(4): 431-440.

[6] Hashem I A T, Yaqoob I, Anuar N B, et al. The rise of “big data” on cloud computing: Review and open research issues[J]. Information Systems, 2015, 47: 98-115.

[7] Li J, Tao F, Cheng Y, et al. Big data in product lifecycle management[J]. The International Journal of Advanced Manufacturing Technology, 2015, 81(1-4): 667-684.

[8] Nalls M A, Pankratz N, Lill C M, et al. Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease[J]. Nature genetics, 2014, 46(9): 989.