A Credit Scoring Model Based on Alternative Mobile Data for Financial Inclusion

Xinhai Liu（1） Ti Wang（2） Wei Ding (2) Yanjun Liu (3) Qiuyan Xu (4)

(1) The CEIBS-World Bank China Center for Inclusion Finance, Shanghai, China
(2) China Unicom Network Technology Research Institute, Beijing, China
(3) AsiaInfo Inc, Beijing, China
(4) Cignifi Company, Boston, MA, USA

Abstract: Credit is becoming more and more important for consumers in China as both Internet Economy and consumer finance are booming in the recent years. However, more than 2/3 people in China do not have credit records because they have never been to banks, thus leading to the difficulty of making credit risk evaluation for them. With big data time coming, a lot of alternative data shed light on the financial inclusion for these unbanked people. Mobile phone data is deemed to be a good type of alternative data for credit scoring. So in this research, collaborating with China Telecom and Cignifi, we develop a new credit score based on mobile phone data, which is believed to be widely used in China in the near future.

Key words: Mobile phone data, Credit Scoring, Financial Inclusion，Alternative Data;

Introduction

Credit scoring is a vital tool for consumer credit risk management. The

traditional credit scoring is based on FICO model, depending on the credit record of a consumer. However, there are nearly 3 billion consumers all around the world who are credit invisible. Especially, in the emerging country, the demand for alternative data of consumers is urgent as financial service is expanding quite fast. The alternative data (or big data) based credit scoring is becoming a hot topic. Several big consumer credit bureaus and even FICO are involved in the related R&D. For example, in Oct. 25th, 2016, FICO declared its financial Inclusion Initiative plan, including the alternative data based credit scoring.

There are more than several hundreds of high-tech companies (even in China, around one hundred of high-tech companies) are working toward this direction, using mobile data, payment data, facility data, e-commerce data and even social network data. Mobile data of consumers is believed to be a good alternative data for credit scoring, which has correlation with loan credit, from capability to willingness. Also，just a few months of mobile data for an individual can provide a sufficient sample. However, how to use mobile data in an efficient way for financial inclusion remains a challenge for China.

In this work, we focus on the mobile data as the alternative data for credit scoring. Here mobile data refers to call detail record (CDR) and billing data. We collaborate with China Unicom, China Mobile and some banks. The mobile data analytic company named Cignifi also provides

some support for this research. The real data is tested for our mobile data based credit scoring model.

Related work

There are some research about the application of mobile phone data in credit scoring overseas while it lacks in China.

Björkegren and Grissen employ behavioral signatures in mobile phone data to predict default with accuracy approaching that of credit scoring methods that rely on financial histories [1]. The method is demonstrated using call records matched to loan outcomes for a sample of borrowers in a Caribbean country. Individuals in the highest quartile of risk by our measure are 6 times more likely to default than those in the lowest quartile.

Jose San Pedro et al present MobiScore, an approach to build a model of the user's financial risk from mobile phone usage data, which previous work has shown to convey information about e.g. personality and socioeconomic status [2]. MobiScore could replace traditional credit scores when no financial history is available, providing credit access to currently excluded population sectors, or be used as a complementary source of information to improve traditional finance-based scores. They validate the proposed approach using real data from a telecommunications operator and a financial institution in a Latin American country, resulting in an accurate model of default comparable to traditional credit scoring techniques.

Speakman demonstrate how boosted decision trees (Adaboost) may be used to create credit scores (probability of repaying a low-value, short term loan) for under-banked populations, allowing them to access credit that was previously unavailable due to a lack of financial data [4].

The mobile phone data is based on Mobile Money platforms which are gaining traction across developing markets as a convenient way of sending and receiving money over mobile phones, IFC, Cignifi & Airtel applies CDR models to increase Mobile Wallet Adoption & activity [5]. Cignifi company in USA applied mobile phone data for Brazil microcredit approval [6].

Dataset

In China, most mobile phone users are pre-paid consumers. So only with mobile phone data, we could not find the strong meaning of " bad person" in term of lending credit. In order to overcome this problem, we adopt financial data in the form of credit default reports, besides the mobile data. This method could be named as cross-training, using the mobile phone data to match the credit default.

Collaborating with a commercial bank in China, we got more than 400,000 consumers, each of them have a mobile phone number, but only 20% of them belong to China Telecom company.

Methods and Results

We take the credit scoring problem as a binary classification task, which we model using logistic regression. A lot of work is for extracting the credit behavior from mobile phone data. Originally, more than 500 features are generated including some behavior tags, which are used for marketing. But less than 15 features are left in the final model.

The default of consumers is defined as the 30 days past due during 12 months performance window. The observation time window for mobile phone behavior is three months. At last, the KS value is around 42 in the evaluation test.

Conclusion and future work

In China, nearly 2/3 people do not have the traditional credit data while more than 2/3 people own at least one mobile phone. With mobile data based credit score, we can help solve the problem of financial service for unbanked people.

Regarding future, first, we will extend the credit score to more consumption scenarios; second, we will add more alternative data, such as payment data and utility data to refine the current model.

References

[1] Daniel Björkegren, Darrell Grissen, Behavior Revealed in Mobile Phone Usage Predicts Loan Repayment, Netmob 2015.

[2] Jose San Pedro, Davide Proserpio, Nuria Oliver, MobiScore: Towards Universal Credit Scoring from Mobile Phone Data, International Conference on User Modeling, Adaptation, and Personalization UMAP 2015: User Modeling, Adaptation and Personalization pp 195-207.
[3] Xinhai Liu, Credit Information Service and Big Data, CITIC Press, Beijing China, Nov. 2016.
[4] Skyler Speakman, Eric Mibuari, Isaac Markus, Felix Kwizera, Mobile –phone based Credit Scoring, NetMob2017, Milan Italy.
[5] IFC, CIgnifi & Airtel, Mobile Wallet Usage Study, 5-7 April, 2017, NetMob2017, Milan Italy.
[6] Data Science Group of Cignifi, Cignifi Risk Solutions, NetMob2015, MIT, Cambirdge,USA.