



BEHAVIOR MONITORING METHODS FOR TRADE-BASED MONEY LAUNDERING INTEGRATING MACRO AND MICRO PRUDENTIAL REGULATION: A CASE FROM CHINA

Xiangrui CHAO¹, Gang KOU^{2*}, Yi PENG¹, Fawaz E. ALSAADI³

¹*School of Management and Economics, University of Electronic Science and Technology of China, No. 2006, Xiyuan Ave, West Hi-Tech Zone, Chengdu, 611731, China*

²*School of Business Administration, Southwestern University of Finance and Economics, No. 555, Liutai Ave, Wenjiang Zone, Chengdu, 611130, China*

³*Department of Information Technology, Faculty of Computing and IT, King Abdulaziz University, Jeddah, Saudi Arabia*

Received 11 June 2018; accepted 16 December 2018

Abstract. Trade-based Money Laundering, a new form of money laundering using international trade as a signboard, always appears along with speculative capital movement which has been accepted as the most concerned and consensus incentive giving rise to the collapse of the financial market. Unfortunately, preventing money laundering is very difficult since money laundering always has a plausible trade characterization. To reach this goal, supervision for regulator and financial institutions aims to effectively monitor micro entities' behavior in financial markets. The main purpose of this paper is to establish a monitoring method including accurate recognition and classified supervision for Trade-based Money Laundering by means of knowledge-driven multi-class classification algorithms associated with macro and micro prudential regulation, such that the model can forecast the predicted class from the concerned management areas. Based on empirical data from China, we demonstrate the application and explain how the monitor method can help to improve management efficiency in the financial market.

Keywords: financial risk monitor, Trade-based Money Laundering, macro- and micro- prudential regulation, machine learning, Reg-tech.

JEL Classification: C80, G18.

Introduction

Cross-border capital frequent flow, which is used as arbitrage in financial markets, is most harmful to the financial market and economics of emerging markets and developing economies (EMDEs). Many accelerating global capital flows and financial market failures require financial market regulation for reasons of export reduced, unemployment and economic stagnation in developing countries.

*Corresponding author. E-mail: kou.gang@qq.com

To cope with speculative capital, short-term capital control remains an effective tool to protect financial markets in EMDEs from disorderly shock. Therefore, cross-border money transferring accompanied by trade background has become one of the capital flow channels, which will cause cross-border capital flow and result in foreign exchange market turbulence (Naheem, 2016a, 2016b; Yousefi et al., 2018; Asongu, Akpan, & Isihak, 2018). Transferring capital with trade background has been declared Trade-based Money Laundering (TBML) by the Financial Action Task Force on Money Laundering (FATF, 2006), which means the illegal or speculative capital transferring along with international trades by means of over- and under- true value of goods and services. The TBML is seriously concerned by emerging markets and developing economics (EMDEs) and has turned into a dominating issue because of laundering funds with higher frequent cross-border capital fluctuations, which have been acknowledged as being the cause of the Southeast Asia financial crisis in 1998.

To detect TBML behavior, several methods have been developed to recognize abnormal capital mixed with good payments. Firstly, a statistical approach aims to discover the average price of goods and trace customs declaration with higher or lower prices. Secondly, the econometric method is used to find isolated points considered to be TBML in a regression process of price and payments for all trade markets. The main insight of these methods is established in the comparison model between individual and total features.

Many rule-based strategic tools including macro- and micro-prudential have been implemented to respond to capital shock from TBML. Macro-prudential tools aim to mitigate the positive impact of capital by means of counter-cyclical supervision, such as concern regulation of inflow capital in a growth period. Micro-prudential tools, such as ongoing and ex-post regulation, self-discipline mechanism against operation risk, executive interviews and currency intervention, etc. are also widely implemented in EMDEs.

However, the existing methods also have many shortages when applied to fit the current trend of TBML management. First, these methods focus on statistical properties of uni-source data such as the customs declaration price, and bank transaction data and detect outlier or deviation to average value. The methods cannot capture new trends of TBML behavior such as unmatched money transferring along with true trade background, which haven't always had false trade behavior such as using fake invoices or shipping merchandise lower than the true value. Secondly, the detection needs to integrate multi-source data to test micro market behavior. Thus, the TBML behavior is characterized by high dimensionality of features, often with high frequency, wholesale and combined monitor index of current and capital accounts. Therefore, developing a machine learning method to discover abnormal behavior in international trade is urgently required. Lastly, targeted and effective methods for macro- and micro- prudential regulation in the financial market in order to decrease human resources and increase management efficiency must also be urgently developed. Recently, Reg-tech, which simply means financial regulation using modern intelligent technologies, has become a popular tool to detect financial risk and is partly employed to regulate (Larsen & Gilani, 2017; Pishdar, Ghasemzadeh, Antucheviciene, & Saparauskas, 2018). Thus, it can provide an effective approach based on information technology, which is associated with macro- and micro- policy instruments responding to TBML, although it is very difficult to be implemented due to the lack of clue detection and intelligent technologies.

This paper aims to develop a framework for managing TBML based on data mining which begins by constructing data feature selection, and multi-class labels which are designed by means of knowledge-driven experiences under macro and micro-prudential management. A multi-class classification algorithm is employed to build a pre-warning model and forecast entities' behavior, as well as an evaluation index to assess such an algorithm's performance.

The possible contributions of this paper are to fill the gap of detecting Trade-based Money Laundering behavior using machine learning (ML) methods, and build a novel framework for TBML regulations with Reg-tech. The data used in this paper has been extracted from real-life practice and management cases. Throughout our research, a good theoretical basis will be set up for financial market management responding TBML in EMDEs.

The remainder of this paper is organized as follows: in Section 1, we overview the main methods for TBML monitoring and cross border capital flow control. Feature extraction and algorithm frameworks will be proposed in Section 2. The experiments demonstrate the effectiveness of ML algorithms using real-world data in Section 3. We conclude this article in last section.

1. Literature reviews

Capital transferring has many financial variables and other institutional factors including financial markets and capital control (Giovanni, 2005; Aglietta & Scialom, 2010). In EMDEs, capital control still has a response policy to protect from financial crisis (Johnson & Mitton, 2001). However, global investment and trade liberalization is still an irreversible trend. Thus, managing global capital flows requires global and national efforts to maintain their stability and growth (Gallagher, 2012). In sum, under the background of the orderly opening of capital markets in EMDEs, capital transferring hidden in trade is widely accepted as one of the most dangerous behaviors for capital flow and the most important area of financial regulation. The amount of economic and financial damage caused by international unlawful acts in recent years leaves little doubt that the financial institutions and multinational corporations (McSkimming, 2010), even customs (Xue & Zhang, 2016) and noncompliant banks (McIntosh, 2016), are becoming severely afflicted areas in global business activities. To summarize, this paper conducts empirical analysis on TBML detection using intelligent algorithm in China and theoretical arguments on feature engineering and algorithm conversion in the Reg-tech areas.

Many questions regarding current forms of money laundering have sprung up and today TBML are becoming more and more popular (Unger & den Hertog, 2012; Ferwerda et al., 2013; Delston & Walls, 2009; Liao & Acharya, 2011; Thanasegaran & Shanmugam, 2007). For this reason, regulation is becoming more important (Deng, Joseph, Sudjianto, & Wu, 2012). However, the challenge will bring a new trend for money-laundering from the traditional banks sector to the financial market along with electronic payments, trade and real estate. There are existing methods for monitoring TBML behavior with different insight and technologies.

First, the usual method used to manage TBML is implemented by the comparison of a customs declaration price and the average price of the goods, which is built from cus-

tom merchandises data. Zdanowicz (2009) illustrated many forms of Trade-based Money Laundering and proposed a statistical method to detect this behavior which is based on price deviation between the mean price of the goods and single trade. His research built an average price database of export goods based on the U.S. Merchandise trade database. By means of his study, the abnormal trades are those prices which are sharply higher than the average price. The same method is also used in China Customs (Cassara, 2011). Gao and Weng (2006) also proposed an artificial transfer pricing paradigm to test money laundering. Moreover, capital flight and tax evasion were added to the model and the impact on the money laundering pool was tested.

Second, many econometric methods in economics attempt to measure TBML behavior. Ferwerda *et al.* (2013) claimed that when using traditional methods in economic literature TBML was difficult to access since the TBML is hidden in licit trade and thus proposed an updated gravity model, which means a regression of exports and Gross Domestic Product (GDP), population, border dummy, common language, colonial background and distance of two countries, as opposed to the Walker (1995) and Unger (2007) prototype model. Their basic idea of the normal value of exports matches the inherent scale of two countries. Thus, they detected TBML flow to be an existing deviation between money transferring and the normal scale of trade based on national GDP and other trade factors. Their results verified that TBML can escape the stricter anti-money laundering regulations of financial markets. Melvin (2014) asserted that the TBML definition of FATF is seriously flawed since trade data used to detect TBML is not valid. Thus, their viewpoint on the context of TBML has always been established with the true value, quality and quantity of the shipment. Therefore, their concerns regarding cash payment should be considered when detecting TBML in trade. From the above literature, there is no doubt that trade and payment are two factors for TBML detection regardless of the type of TBML detection.

Lastly, Reg-tech technologies which try to use machine learning methods in money laundering are also a new trend to detect money laundering behavior. For example, Outlier Detection is regarded as abnormal behavior in financial transaction data including clustering and regression algorithms (Rohit & Patel, 2015; Kannan & Somasundaram, 2015). Kannan and Somasundaram (2017) employed the autoregressive-based outlier algorithm to detect money laundering behavior and their model aimed to compute the mean value of financial transaction data and then obtain deviation between individual data and mean data. Based on the proposed distance, abnormal data can be detected in terms of pre-set distance parameters.

However, we must point out that the methods proposed in current papers are based on price comparison and statistical methods under the assumption that the TBML have fake trade which is not the case for the new trend of TBML (Melvin, 2014), or the data mining using outlier algorithms in order to detect abnormal data. The money laundering regulation should adopt integrated macro- and micro-policy and global regulation mixed local regulation to protect from systemic risk caused by speculative local capital flow under international trade in financial markets (Borio, 2003, 2011; Wymeersch, 2010; Kashif, S. F. Iftikhar, & K. Iftikhar, 2016; Song, Wang, & Zhu, 2018; Chen, He, & Li, 2017; Zavadskas et al., 2018). Macro-prudential regulation can be simply divided into two main folds: on the one hand, a countercyclical operation used to alleviate cyclical capital flow which causes harmful economic shock. On the other hand, 'too big to fail' institutions should be discerned and

managed in order to respond to the risk spread (Allen, Goldstein, Jagtiani, & Lang, 2016; Arnold, Borio, Ellis, & Moshirian, 2012; Milne, 2009; Arthur, 2017; Ahmed, Socci, Severini, Yasser, & Pretaroli, 2018; Hamdi, Hakimi, & Zaghoudi, 2017). Micro-prudential regulation evaluates asset and risk response for micro-companies. Therefore, an effective Reg-tech for TBML should consider data feature engineering based on micro-prudential, knowledge-driven based macro-prudential regulation. That is, the classification labels can be assigned with different periods in terms of the countercyclical direction with the current and further trend. In addition, classification features can be designed on basis of the metrics of asset-debt and risk exposures of micro-companies.

Therefore, in our paper, the monitoring direction of capital flow is determined according to macro-prudential regulation tasks, and a payments-goods based TBML framework will be constructed and knowledge-driven classification methods will be used to detect proposed behavior based on Chinese data.

2. Framework, feature selection and machine learning algorithms

In this section, a monitoring framework will be presented according to the following Figure 1.

2.1. Identifying TBML behavior

The rule of TBML studied in our paper is the aforementioned payment-trade principle¹ which provides a new insight for TBML based on the true value, quality and quantity of the shipment (Melvin, 2014). The TBML should identify the flow of money corresponding to trade and the deviation of the value between money and goods, even frequency and amount. The behavior should be described by many features using machine learning tools. The detail can be found in Subsection 3.3.

2.2. Monitoring process and data resource

The monitoring model is constructed on the basis of payments and goods of entities (individuality and enterprises) in the market. The class label will be assigned referring to countercyclical and management cases. At the bottom of Figure 1 are the machine learning methods used for behavior monitoring. In this paper, a classification method will be presented to illustrate the detection of TBML.

2.3. Feature extraction

The features are constructed as a combination of time frequency, rate of flow and deviation of trade and money (detail can be found in the Appendix 1). Table 1 shows the main features in monitoring models and the index is computed within a period of 12 months using dynamic recursion calculation. Parts of the index are collected from the State Administration of Foreign Exchange (SAFE) in China.

¹ Payment-trade principle means the detection of TBML is established at a money transferring and goods value match.

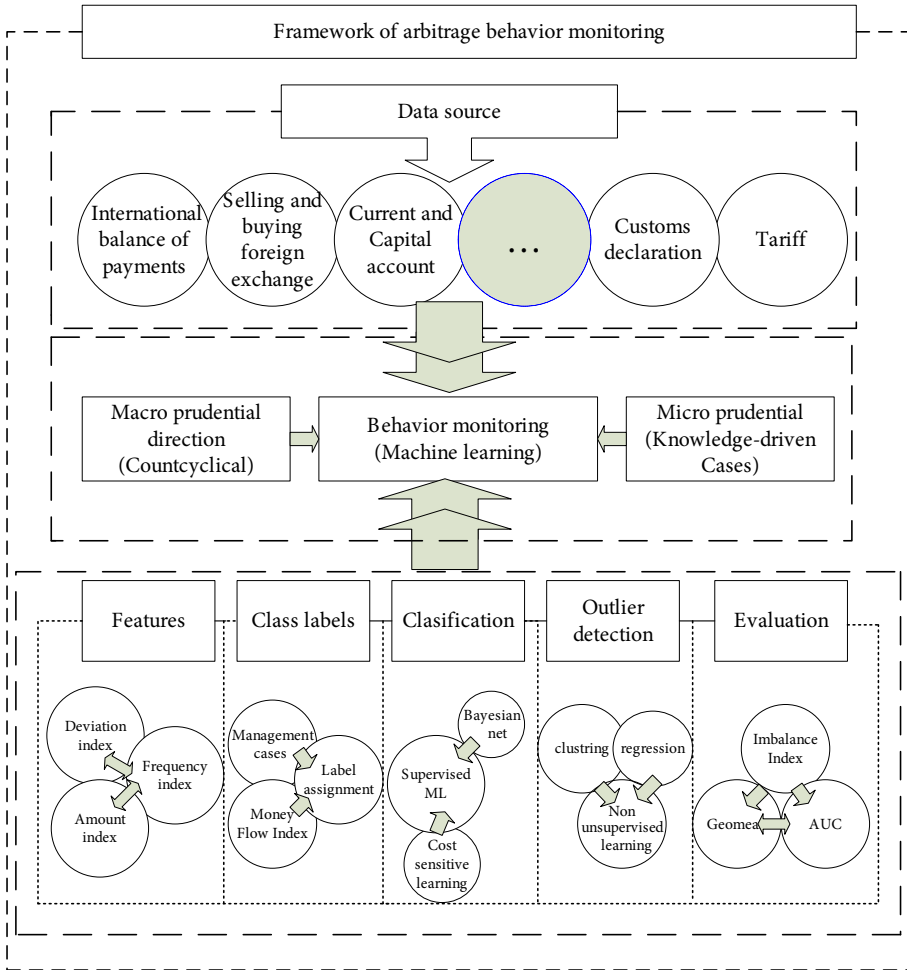


Figure 1. Framework of behavior monitoring

2.4. Machine learning algorithms

Machine learning (ML) algorithms used in our model are based on knowledge-driven multi-class classification technology. In the real world, the distribution of different class label data is different, and the numbers also have an unbalanced ratio in the production and collection of data, especially it is the case in financial regulation. The ratio of normal and abnormal institutions becomes larger since only small group arbitrage institutions are forecasted and detected. Cost sensitive learning is one of the representative algorithms for imbalanced data (Lomax & Vadera, 2013; Masnadi-Shiraziet, Vasconcelos, & Iranmehr, 2015; Huang & Kou, 2014; Huang, Kou, & Peng, 2017; Kou, Peng, & Wang, 2014; Kou, Lu, Peng, & Shi, 2012; Chao, Kou, Li, & Peng, 2018). In some specific data sets, the accuracy rate proved to be higher than standard classifiers (Liu & Zhou, 2006). Recently Chao and Peng (2017) have developed a new cost-sensitive classification algorithm.

Table 1. Features selection in monitoring model

Total Transactions of purchasing currency	Export	Actual domestic forex loan more than 90 days	Total payment
Whether or not inspect	Income	Balance of domestic forex loan more than 90 days	Purchasing/payment
Licensing status	Expense	Whether or not top 30 purchasing	Purchasing and payment at remote bank
Management level	Purchasing currency	Mean days of purchasing	Payment/import
Local sample	Exchange Settlement	Numbers of top 30 purchasing	Credit trade/total trade
Special assignment of SAFE	Account purchasing	Mean amount of purchasing	Payment of credit payment
Total deviation of money and goods	Account settlement	Whether or not individual large-amount	Settlement + income – purchasing – expense
Credit trade	Cash purchasing	Total amount of purchasing	Import + income – export – expense
Import	Cash settlement	Single payment	Settlement + import – purchasing – export
(Import + income – export – expense)/(Import + income + export + expense)			

Using our methods, Bayes net (BN) based on Bayesian network is employed as a base classifier and used to handle imbalanced data in financial markets. The main ideas of this algorithm are shown as follows:

First, BN is a set including a directed acyclic graph G and a conditional probability table. Let $G = (I, E)$ represent nodes set I which is random variable and edges set E which is conditional probability.

The joint probability can be obtained using the product of local conditional probability as follows:

$$p(X = x_1, x_2, \dots, x_n) = \prod_i p(x_i | x_{i+1}, x_{i+2}, \dots, x_n);$$

$$p(x_i | x_j) = \frac{p(x_j | x_i) p(x_j)}{p(x_j)}.$$

For a node (random variable) x , the joint probability can be denoted as follows:

$$p(x) = \prod_i p(x | Parents(x_i)),$$

where $Parents(x_i)$ is set of i -th farther node of x .

In fact, the building of Bayesian net is a solution process of parameters of network structure and probability distribution on the basis of the given data. The probability distribution of root nodes has always determined prior probability. The Bayesian network is used as a classifier if the output node is a class label and the feature is a node. The network and conditional probability will be computed in terms of training data and used to forecast the model.

The main step in BN classification is building the network structure and finding the conditional probability tables of the BN.

First, network structure always uses a heuristic-search algorithm for maximum conditional probability from parent nodes to child nodes. The representative algorithm is K2 established by Cooper and Herskovits (1992). The process can be employed to design max parent nodes, initially a naïve bays network, and then use a hill climbing algorithm restricted by variable ordering.

Second, an estimator is used for parameter estimation for the conditional probability tables of the BN, based on Bayesian conditional probability. The common estimator includes Maximum likelihood estimation, Maximum a posteriori estimation and Bayes estimator, etc. (Cooper & Herskovits, 1992).

Finally, measures including Minimum Description Length, Akaike Information Criterion, and Entropy, etc. are employed to detect quality of established net.

The advantages of Bayesian net are that the network does not require mutual independence and relatively higher accuracy can be acquired by dynamic adjustment along with posterior probability. In our paper, we use K2 and Bayes estimator without adding the ADTree structure.

3. Empirical experiments

Gathering information about entities from the balance of payments and international trade data of Qinghai province, Ningbo city and Shenzhen city gives data which is representative of western, eastern and southern China. The TBML detection is implemented by local government and bank sectors and the different data will be trained as an adaptive forecast model suited to different markets. Therefore, it is stated that the model is not fixed and will not be used in all areas, rather, a different model will be set to test various data under the proposed framework and algorithm.

Data ranges from March 2016 to Feb 2017 and there are 8837 instances with 37 features. The panel data was selected with a sample period of 12 months for two reasons. Firstly, TBML should be detected in a reasonable amount of time so as to promptly identify abnormal behavior rules. A longer time period impacts management time-effectiveness and shorter time periods decrease detection accuracy. Secondly, in management practice data features such as computing and implementation recirculate every month based on past 12 months.

In this section, experiments are implemented using WEKA 3.7.8 (Hall et al., 2009) which is a popular computer software tool widely used in data mining in scientific research.

3.1. Multi-class classification: knowledge-driven abnormal behavior detection

3.1.1. Evaluation index

When testing classification performances, TP-rate, FP-rate, AUC, Geomean, F-measure and ROC are used to evaluate the forecasting model:

$$TP_{rate} = \frac{TP}{TP + FN}; \quad TN_{rate} = \frac{TN}{TN + FP}; \quad FP_{rate} = \frac{FP}{TN + FP}; \quad FN_{rate} = \frac{FN}{TP + FN};$$

$$AUC = \frac{1 + TP_{rate} - FP_{rate}}{2};$$

$$\text{Geomean (GM)} = \sqrt{\frac{TN}{FP + TN} \times \frac{TP}{TP + FN}};$$

$$\text{F-measure } F = \frac{1}{\frac{TP + FP}{TP} + \frac{P}{TP}},$$

where *P* is number of positive instances.

The ROC curve is drawn having TP-rate and FP-rate values that compares proposed criterions with various thresholds.

3.1.2. Classification and forecasting model

Classification is a supervised data mining algorithm, and the class labels should be assigned a training set in advance. Then, the trained model will be used to handle the testing dataset. Therefore, non-label data items will be labeled by the trained classification model. When monitoring TBML in our model, the labels are ascertained by macro-and micro prudential regulation. First, direction of capital flow is a central factor in assuring that the model can be used to counter-cycle cross-border capital flow. For instance, net outflow entities should be focused on capital flowing to other countries. Second, we should regard micro-behavior considering administrative penalty, low credit, money laundering concern, business cases and arbitrage institution, etc. to obtain general results so that the model can be extensively applied to all areas. However, abnormal entities always exist in small amounts in financial markets, for example, the ratio of normal and abnormal entities is less than 10000:1, thus, imbalance learning should be used to classify the forecasting model. We adopt the Relief feature selection to retain 40 features from the Appendix 1 in order to decrease the correction of two indexes and increase performance of the algorithm. The training data will be assigned a different cost in the sampling process for higher classification accuracy.

3.1.3. Label assignment

We divided data into five classes: net inflow, small outflow, offshore re-export trade and export without income, respectively. The detail can be seen in Table 2.

Table 2. Class labels

Label	Numbers	Ratio	Remark
Net inflow	7224	81.7%	Annotation in capital outflow stage
Small net outflow	330	3.7%	within the tolerance range in capital outflow stage
Concerned net outflow	1188	13.4%	Local samples
Offshore re-export trade	36	0.4%	Administrative penalty
Export without income	59	0.7%	Concerned by SAFE and China Customs

3.1.4. Results

Table 3 shows the confusion matrix of classification using BN. The minority classes such as offshore re-export trade can be classed correctly, and export without income can almost be correctly recognized. This demonstrates that classification can be used to distinguish ab-

normal entities in financial markets. However, small net outflow and concerned net outflow have confusion in two areas which include small net outflow being classed into concerned net outflow and concerned net outflow into net inflow. The former has no impact on actual management and the latter should be handled for more management targets. The offshore re-export trade and export without income will be regarded as TBML and further inspected by the local government and bank sector.

Table 3. Confusion matrix

		Hypothesis				
		Net inflow	Small net outflow	Concerned net outflow	Offshore re-export trade	Export without income
Ture	Net inflow	7159	0	0	36	29
	Small net outflow	18	312	0	0	0
	Concerned net Outflow	115	0	1031	0	42
	Offshore re-export trade	0	0	0	36	0
	Export without income	7	0	0	0	52

Therefore, the two classes should be recognized as the same class on application in order to improve the accuracy of the model, or net flow should adjust many delicate classes so as to fit to different situations.

Table 4 shows the results of classification performance. It is highlighted that our model aims to separate offshore re-export trade and export without income so that the model can be used to detect abnormal behavior predefined by management application.

Table 4. Performance of classification

	TP	FP	AUC	GM	F	ROC
Net inflow	0.991	0.087	0.952	0.951	0.986	0.995
Small net outflow	0.945	0.000	0.972	0.972	0.972	1.000
Concerned net outflow	0.868	0.000	0.934	0.932	0.929	1.000
Offshore re-export trade	1.000	0.004	0.998	0.998	0.667	1.000
Export without income	0.881	0.008	0.937	0.935	0.571	0.926

3.1.5. Discussion

Classification algorithms can be used to monitor abnormal behavior. The classification algorithm can classify the data sets using class labels which are based on knowledge which comes from management results and is directly labeled for monitoring classification, so as to effectively improve the maximum utilization of monitoring conclusions in different stages.

Based on the results of the experiment, three main conclusions can be summarized for further application of the proposed method.

In the first place, data feature engineering includes supervision information, trade credit and industry information, and other structural index and frequency indexes. The proposed

method draws the data from trade and money transfer, and as a composite factor, goods and payments, rather than the existing methods which are tested by means of goods or price, macro data, or simple traction data.

Next, the knowledge-driven classification can be used to effectively forecast TBML behavior and it is a systemic research tool for the Reg-tech with machine learning method as compared to the existing detection method. The algorithm can obtain better accuracy for small class data, which are regarded as having abnormal TBML behavior in international trade.

Finally, this study aims to build a new framework associated with macro-prudential and micro-prudential regulation. The concerned labeling process of knowledge-driven classification should consider the cyclicity of capital flow and the micro regulation index, which adapts to the financial market and foreign exchange transactions in EMDEs.

Remark: It is stressed that no comparison with other TBML test techniques is shown in this paper because, as far as we know, this is the first time a data-driven methodology based on payment-goods detection associated with macro-prudential and micro-prudential regulation has been implemented. The existing econometric models were used to test the existence of TBML in the financial market rather than micro-markets. Moreover, statistical methods were compared with static price which assumes the current customs declaration price to be false (Zdanowicz, 2009), and our results are also based on true trade background (Melvin, 2014). Moreover, existing data mining methods were based on features collection from bank transactions such as transaction date, transaction amount and the mode of transaction, etc. (for instance, Kannan & Somasundaram, 2017). Therefore, the comparison is unpractical due to the different principles and data features.

Conclusions

In this paper, we constructed a new insight for TBML associated with macro-prudential and micro-prudential regulation based on payments and goods under the assumption of TBML having true value in shipment. Meanwhile, feature engineering used for machine learning is established, and a data mining algorithm is employed to monitor abnormal behavior. The Empirical experiments were implemented using data from management applications from China, and the results show the effective action based on data mining. The abnormal behavior can be detected by proposed methods, although the number of entities is small compared to normal micro-entities. Furthermore, the methods can improve management efficiency and prevent abnormal cross-border capital flow for EMDEs. This framework can also be extensively implemented in EMDEs to prevent arbitrage cross-border capital flow by means of true international trade background.

Many current studies will now be further developed to obtain more detailed results. Firstly, the empirical research on the global financial market can be studied in order to investigate correlations of capital flow and trade scale. Secondly, mixed individuals including banks, companies and non-bank institutions should be considered for further comprehensive monitoring systems. Finally, how to employ big data will be researched in the near future and multi-subjects will also be studied.

Acknowledgements

This research has been partially supported by grants from the National Natural Science Foundation of China (#71725001, #71874023, #71433001, and #U1811462).

References

- Aglietta, M., & Scialom, L. (2010). A systemic approach to financial regulation: a European perspective. *International Economics*, 123, 31-65. [https://doi.org/10.1016/S2110-7017\(13\)60013-X](https://doi.org/10.1016/S2110-7017(13)60013-X)
- Ahmed, I., Succi, C., Severini, F., Yasser, Q. R., & Pretaroli, R. (2018). Forecasting investment and consumption behavior of economic agents through dynamic computable general equilibrium model. *Financial Innovation*, 4(1), 7. <https://doi.org/10.1186/s40854-018-0091-3>
- Allen, F., Goldstein I., Jagtiani J., & Lang, W. W. (2016). Enhancing prudential standards in financial regulations. *Journal of Financial Services Research*, 49(2-3), 133-149. <https://doi.org/10.1007/s10693-016-0253-2>
- Arnold, B., Borio, C., Ellis, L., & Moshirian, F. (2012). Systemic risk, macroprudential policy frameworks, monitoring financial systems and the evolution of capital adequacy. *Journal of Banking & Finance*, 36, 3125-3132. <https://doi.org/10.1016/j.jbankfin.2012.07.023>
- Arthur, K. N. A. (2017). Financial innovation and its governance: Cases of two major innovations in the financial sector. *Financial Innovation*, 3(1), 10. <https://doi.org/10.1186/s40854-017-0060-2>
- Asongu, S., Akpan, U. S., & Isihak, S. R. (2018). Determinants of foreign direct investment in fast-growing economies: evidence from the BRICS and MINT countries. *Financial Innovation*, 4(1), 26. <https://doi.org/10.1186/s40854-018-0114-0>
- Borio, C. (2003, February). *Towards a macroprudential framework for financial supervision and regulation?* (BIS Working Papers No. 128). Retrieved from <https://www.bis.org/publ/work128.pdf>
- Borio, C. (2011). Implementing a macroprudential framework: blending boldness and realism. *Capitalism and Society*, 6(1), 1-21. <https://doi.org/10.2202/1932-0213.1083>
- Cassara, J. A. (2011). Trade-based money Laundering. *China Customs*, 5, 226-230.
- Chao, X. R., & Peng, Y. (2017). A cost-sensitive multi-criteria quadratic programming model for imbalanced data. *Journal of the Operational Research Society*, 69(4), 500-516. <https://doi.org/10.1057/s41274-017-0233-4>
- Chao, X., Kou, G., Li, T., & Peng, Y. (2018). Jie Ke versus AlphaGo: A ranking approach using decision making method for large-scale data with incomplete information. *European Journal of Operational Research*, 265(1), 239-247. <https://doi.org/10.1016/j.ejor.2017.07.030>
- Chen, T., He, J., & Li, X. (2017). An evolving network model of credit risk contagion in the financial market. *Technological and Economic Development of Economy*, 23(1), 22-37. <https://doi.org/10.3846/20294913.2015.1095808>
- Cooper, G. F., & Herskovits, E. (1992). A bayesian method for the induction of probabilistic networks from data. *Machine Learning*, 9(4), 309-347. <https://doi.org/10.1007/BF00994110>
- McIntosh, D. (2016). The costs of anti-money laundering enforcements to noncompliant banks. *Journal of Finance and Bank Management*, 4(1), 1-14.
- Delston, R. S., & Walls, S. C. (2009). Reaching beyond banks: how to target trade-based money laundering and terrorist financing outside the financial sector. *Case Western Reserve Journal of International Law*, 41, 85-97.
- Deng, X., Joseph, V. R., Sudjianto, A., & Wu, C. J. (2009). Active learning through sequential design, with applications to detection of money laundering. *Journal of the American Statistical Association*, 104(487), 969-981. <https://doi.org/10.1198/jasa.2009.ap07625>

- FATF. (2006). *Trade based money laundering*. Financial Action Task Force (FATF). Retrieved from <http://www.fatf-gafi.org/media/fatf/documents/reports/TradeBasedMoneyLaundering.pdf>
- Ferwerda, J., Kattenberg, M., Chang, H. H., Unger, B., Groot, L., & Bikker, J. A. (2013). Gravity models of trade-based money laundering. *Applied Economics*, 11(22), 3170-3182. <https://doi.org/10.1080/00036846.2012.699190>
- Gallagher, K. P. (2012). Regaining control? Capital controls and the global financial crisis. In *The consequences of the global financial crisis: the rhetoric of reform and regulation*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199641987.003.0007>
- Gao, Z. A., & Weng, L. F. (2006). Transfer price-based money laundering in international trade. *International Conference on Management Science and Engineering*, 1128-1132.
- Giovanni, J. D. (2005). What drives capital flows? The case of cross-border M&A activity and financial deepening. *Journal of International Economics*, 65(1), 127-149. <https://doi.org/10.1016/j.jinteco.2003.11.007>
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. (2009). The WEKA data mining software: an update. *SIGKDD Explorations*, 11, 10-18. <https://doi.org/10.1145/1656274.1656278>
- Hamdi, H., Hakimi, A., & Zaghoudi, K. (2017). Diversification, bank performance and risk: have Tunisian banks adopted the new business model? *Financial Innovation*, 3(1), 22. <https://doi.org/10.1186/s40854-017-0069-6>
- Huang, Y., & Kou, G. (2014). A kernel entropy manifold learning approach for financial data analysis. *Decision Support Systems*, 64, 31-42. <https://doi.org/10.1016/j.dss.2014.04.004>
- Huang, Y., Kou, G., & Peng, Y. (2017). Nonlinear manifold learning for early warnings in financial markets. *European Journal of Operational Research*, 258(2), 692-702. <https://doi.org/10.1016/j.ejor.2016.08.058>
- Johnson, S., & Mitton, T. (2001). Cronyism and capital control: evidence from Malaysia. *Journal of Financial Economics*, 67(2), 351-382. [https://doi.org/10.1016/S0304-405X\(02\)00255-6](https://doi.org/10.1016/S0304-405X(02)00255-6)
- Kannan, S., & Somasundaram, K. (2015). A review of outlier prediction techniques in data mining. *Research Journal of Applied Sciences Engineering & Technology*, 10(9), 1021-1028. <https://doi.org/10.19026/rjaset.10.1869>
- Kannan, S., & Somasundaram, K. (2017). Autoregressive-based outlier algorithm to detect money laundering activities. *Journal of Money Laundering Control*, 20(2), 190-202. <https://doi.org/10.1108/JMLC-07-2016-0031>
- Kashif, M., Iftikhar, S. F., & Iftikhar, K. (2016). Loan growth and bank solvency: evidence from the Pakistani banking sector. *Financial Innovation*, 2(1), 22. <https://doi.org/10.1186/s40854-016-0043-8>
- Kou, G., Peng, Y., & Wang, G. (2014). Evaluation of clustering algorithms for financial risk analysis using MCDM methods. *Information Sciences*, 275, 1-12. <https://doi.org/10.1016/j.ins.2014.02.137>
- Kou, G., Lu, Y., Peng, Y., & Shi, Y. (2012). Evaluation of classification algorithms using MCDM and rank correlation. *International Journal of Information Technology & Decision Making*, 11(01), 197-225. <https://doi.org/10.1142/S0219622012500095>
- Larsen, K., & Gilani, S. (2017). Regtech is the new black – the growth of RegTech demand and investment. *Journal of Financial Transformation*, 45, 22-29.
- Liao, J., & Acharya, A. (2011). Transshipment and trade-based money laundering. *Journal of Money Laundering Control*, 14(1), 79-92. <https://doi.org/10.1108/13685201111098897>
- Liu, X. Y., & Zhou, Z. H. (2006). Training cost-sensitive neural networks with methods addressing the class imbalance problem. *IEEE Transactions on Knowledge and Data Engineering*, 18, 66-37.
- Lomax, S., & Vadera, S. (2013). A survey of cost-sensitive decision tree induction algorithms. *ACM Computing Surveys*, 45, 1-35. <https://doi.org/10.1145/2431211.2431215>

- Masnadi-Shirazi, H., Vasconcelos, N., & Iranmehr, A. (2015). Cost-sensitive Support Vector Machines. *Journal of Machine Learning Research* (arXiv:1212.0975V2).
- Mckimming, S. (2010). Trade-based money laundering: responding to an emerging threat. *Deakin Law Review*, 15(1), 37-63. <https://doi.org/10.21153/dlr2010vol15no1art116>
- Milne, A. (2009). Macroprudential policy: what can it achieve? *Oxford Review of Economic Policy*, 25(4), 608-629. <https://doi.org/10.1093/oxrep/grp036>
- Melvin, R. J. S. (2014). A critical approach to trade-based money laundering. *Journal of Money Laundering Control*, 17(2), 230-242. <https://doi.org/10.1108/JMLC-01-2013-0001>
- Naheem, M. A. (2016a). Trade based money laundering: a primer for banking staff. *International Journal of Disclosure & Governance*, 14(2), 95-117. <https://doi.org/10.1057/ijdg.2015.21>
- Naheem, M. A. (2016b). Risk of money laundering in the US: HSBC case study. *Journal of Money Laundering Control*, 19(3), 225-237. <https://doi.org/10.1108/JMLC-01-2015-0003>
- Nitsch, V. (2010). *The dynamics of illicit flows from developing countries*. Paper presented at the World Bank conference, 14-15 September. Washington.
- Pishdar, M., Ghasemzadeh, F., Antucheviciene, J., & Šaparauskas, J. (2018). Internet of things and its challenges in supply chain management; a rough strength-relation analysis method. *E & M Ekonomie a Management*, 21(2), 208-222. <https://doi.org/10.15240/tul/001/2018-2-014>
- Rohit, K. D., & Patel, D. B. (2015). Review on detection of suspicious transaction in anti-money laundering using data mining framework. *International Journal for Innovative Research in Science and Technology*, 1, 129-133.
- Song, Y., Wang, H., & Zhu, M. (2018). Sustainable strategy for corporate governance based on the sentiment analysis of financial reports with CSR. *Financial Innovation*, 4(1), 2. <https://doi.org/10.1186/s40854-018-0086-0>
- Thanasegaran, H., & Shanmugam, B. (2007). International trade-based money laundering: the malaysian perspective. *Journal of Money Laundering Control*, 10(4), 429-437. <https://doi.org/10.1108/13685200710830916>
- Unger, B. (2007). *The scale and impacts of money laundering*. Cheltenham, UK: Edward Elgar. <https://doi.org/10.4337/9781781007624>
- Unger, B., & den Hertog, J. (2012). Water always finds its way: Identifying new forms of money laundering. *Crime Law & Society Change*, 57(3), 287-304. <https://doi.org/10.1007/s10611-011-9352-z>
- Walker, J. (1995). *Estimates of the extent of money laundering in and throughout Australia*. The Australian Financial Intelligence Unit AUSTRAC.
- Wymeersch, E. (2010). Global and regional financial regulation: the viewpoint of a European securities regulator. *Global Policy*, 1(2), 201-208. <https://doi.org/10.1111/j.1758-5899.2010.00031.x>
- Xue, Y. W., & Zhang, Y. H. (2016). Research on money laundering risk assessment of customers – based on the empirical research of China. *Journal of Money Laundering Control*, 19(3), 249-263. <https://doi.org/10.1108/JMLC-01-2015-0004>
- Yousefi, V., Haji Yakhchali, S., Šaparauskas, J., & Kiani, S. (2018). The impact made on project portfolio optimisation by the selection of various risk measures. *Inzinerine Ekonomika-Engineering Economics*, 29(2), 168-175. <https://doi.org/10.5755/j01.ee.29.2.17405>
- Zavadskas, E., Šaparauskas, J., & Antucheviciene, J. (2018). Sustainability in construction engineering. *Sustainability*, 10(7), 2236. <https://doi.org/10.3390/su10072236>
- Zdanowicz, J. S. (2009). Trade-based money laundering and terrorist financing. *Review of Law & Economics*, 5(2), 855-878. <https://doi.org/10.2202/1555-5879.1419>

APPENDIX 1

Features used in data mining

Index	Attribute	Remark
Inspection status	Logic	Whether or not inspected and concerned
Licensing status	Logic	Whether or not licensing of SAFE
Management level	Logic	Whether degraded by SAFE
National samples	Logic	Whether or not a national sample
Local samples	Logic	Whether or not a local sample
Assignment status	Logic	Whether or not a special assignment by SAFE
Special assignment of SAFE	Logic	What is a special assignment
Total deviation of money and goods	Numeric	Import + income – export – expense
Balance of Credit trade (CT)	Numeric	Balance of Credit trade
Import	Numeric	Total amount of import
Export	Numeric	Total amount of export
Expense	Numeric	Total amount of expense
Income	Numeric	Total amount of income
Import after CT	Numeric	Import after CT
Export after CT	Numeric	Export after CT
Expense after CT	Numeric	Expense after CT
Income after CT	Numeric	Income after CT
Deferred payment	Numeric	Deferred payment
Deferred receipt	Numeric	Deferred receipt
Payment in advance	Numeric	Payment in advance
Advances on sales	Numeric	Advances on sales
Balance of deferred payment	Numeric	Balance of deferred payment
Balance of deferred receipt	Numeric	Balance of deferred receipt
Balance of payment in advance	Numeric	Balance of payment in advance
Balance of advances on sales	Numeric	Balance of advances on sales
Total amount of purchasing currency	Numeric	Total amount of purchasing currency
Total Transactions of purchasing currency	Numeric	Total Transactions of purchasing currency
Currency settlement	Numeric	Currency settlement
Account purchasing	Numeric	Account purchasing
Account settlement	Numeric	Account settlement
Cash purchasing	Numeric	Cash purchasing

End of Appendix 1

Index	Attribute	Remark
Cash settlement	Numeric	Cash settlement
Usance credit more than 90 days	Numeric	Usance credit more than 90 days
Balance of usance credit more than 90 days	Numeric	Balance of usance credit more than 90 days
Overseas payment more than 90 days	Numeric	Overseas payment more than 90 days
Balance of overseas payment more than 90 days	Numeric	Balance of overseas payment more than 90 days
Actual domestic forex loan more than 90 days	Numeric	Actual domestic forex loan more than 90 days
Balance of domestic forex loan more than 90 days	Numeric	Balance of domestic forex loan more than 90 days
Top 30 purchasing	Numeric	Top 30 purchasing
Mean days of purchasing	Numeric	Mean days of purchasing
Numbers of top 30 purchasing	Numeric	Numbers of top 30 purchasing
Mean amount of purchasing	Numeric	Mean amount of purchasing
Whether or not individual large-amount	Numeric	Whether or not individual large-amount
Mean payments	Numeric	Mean payments
Single large purchasing currency	Numeric	Single large purchasing currency
Purchasing/payment	Numeric	Purchasing/payment
Purchasing and payment at remote bank	Numeric	Purchasing and payment at remote bank
Payment/import	Numeric	Payment/import
Credit trade/total trade	Numeric	Credit trade/total trade
Net flow of credit trade	Numeric	Net flow of credit trade
Relative deviations	Numeric	$(\text{Import} + \text{income} - \text{export} - \text{expense}) / (\text{Import} + \text{income} + \text{export} + \text{expense})$
Exchange gap	Numeric	Settlement + income - purchasing - expense
Income and expense gap	Numeric	Import + income - export - expense
Absolute deviation	Numeric	Settlement + import - purchasing - export