Eurographics Conference on Visualization (EuroVis), Posters Track (2016), pp. 1–3
T. Isenberg and F. Sadlo (Guest Editors)

Visual Analytics for Fraud Detection: Focusing on Profile Analysis

Roger Almeida Leite[1], Theresia Gschwandtner[1], Silvia Miksch[1], Erich Gstrein[2] & Johannes Kuntner[2]

[1]Vienna University of Technology, Austria
[2]Erste Group IT International, Austria

Abstract

Financial institutions are always interested in ensuring security and quality for their customers. Banks, for instance, need to identify and avoid harmful transactions. In order to detect fraudulent operations, data mining techniques based on customer profile generation and verification are commonly used. However, these approaches are not supported by Visual Analytics techniques yet. We propose a Visual Analytics approach for supporting and fine-tuning profile analysis and reducing false positive alarms.

Categories and Subject Descriptors (according to ACM CCS): Visual Knowledge Discovery, Time Series Data, Business and Finance Visualization, Financial Fraud Detection.

1. Introduction

Financial institutions that provide credit and money transaction services are interested in efficiently identifying suspicious behavior. Some examples of such frauds are 'money laundry', 'straw persons', and 'fake user'. In order to adapt to new kinds of frauds and at the same time avoid too many false positive alarms, automatic systems for the identification of frauds need to be constantly supervised and refined. Bank transaction data and credit control data contain time-oriented and multivariate features, which are of complex nature [AMST11] and demand for appropriate visualization and exploration means.

Commonly, artificial Intelligence (AI) techniques and fraud detection metrics are used to identify suspicious patterns and outliers in the financial operational field [FHPS98]. Despite the wide amount of available outlier detection techniques in the literature, most solutions implement machine learning approaches [PLSG10, YWW*07]. Common shortcomings of existing approaches are outlined in Section 2. One major problem in successfully identifying fraudulent behaviour, is that these strategies are constantly changed and adapted, and are thus hard to detect. On the other hand, reporting each transaction outside the norm would increase the number of false-positive alarms, which should also be avoided [CCM*14]. To this end, we propose a Visual Analysis (VA) approach for the investigation and fine-tuning of algorithms for the automatic generation of alarms.

Based on a literature study and on a collaboration with a financial company, we could identify important tasks. Our contributions are: (1) listing the challenges of fraud detection focusing on customer profile analysis, (2) the integration of a VA loop into the profile analysis process, and (3) the prototypical implementation of a VA approach for the investigation of suspicious behaviour and fine-tuning of automatic alert systems.

2. Related Work

Kirkland, et al. [KSH*99] published one of the first works in fraud detection enhanced by visualization analysis. In this work AI, visualization, pattern recognition, and data mining were combined to support regulatory analysis, alerts (pattern detection), and knowledge discovery. WireVis's [CGK*07] main idea is to explore big amounts of transaction data using a multiple coordinated view visualizations. The multiple-coordinated views approach aims to depict relationships among accounts and keywords over time. Huang, et al. [HLN09] presented a VA framework for stock market security. In order to reduce the number of false alarms presented by traditional AI techniques, this work presents a visualization approach combining a 3D tree map for market performance analysis and a node-link diagram for network analysis. Carminati, et al. [CCM*14] present a semi-supervised online banking fraud analysis and decision support based on profiles generation and analysis. However, this is a statistic oriented approach without visual support. This approach is directly related to our approach in the sense that we are also focusing on profile analysis, yet we believe that VA methods have the potential to investigate the data and scoring in more detail and enable the analyst to better fine-tune the scoring system.

According to our literature study the following challenges can be derived: (a) false alarm reduction; (b) development of a comprehensive VA design; (c) precision enhancement on event analysis

and prediction (monitoring). (d) enhancement of the scalability for frauds analysis; (e) knowledge base construction in order to support further fraud identifications;

3. Conceptual Design

We design our interactive VA approach with respect to the data, users, and tasks [MA14].

Data: Financial transaction events.
Users: Analysts from financial institutions that monitor, investigate, and validate transactions and alert systems.
Tasks: The overall task is fraud detection by means of profile analysis. This task includes fine-tuning of automatic alert algorithms as well as managing the trade-off between the sensitivity of the approach and the reduction of false alarms.

In the field of fraud detection, purely automatic approaches tend to produce false alarms [HLN09]. There are also some VA approaches focused to improve fraud detection which mainly use node-link diagrams to represent sequences of suspicious events identified by AI and metric techniques. These aim to reduce the number of false alarms and increase the quality of the query by adding human interaction and cognition in the fraud discovery loop [LGM*15]. Yet, there are no VA approaches to aid the task of profile analysis in fraud detection. This task is usually supported by machine learning techniques used to generate the profiles (e.g., for credit card fraud detection and for online banking fraud detection [CCM*14]). To fill the current lack of VA support in profile analysis, we propose a VA approach to support the exploration of customer profiles to aid reasoning as well as the adaption and fine-tuning of the fraud detection system (i.e., the transaction evaluation system).

Automatic Transaction Evaluation. The system builds a profile for each user based on past transactions in which he/she was involved. For instance, considering financial transactions, we could construct a profile based on (1) how often the costumer executes operations, (2) how much money does he/she usually transfer, and (3) the geographic locations involved. New transactions are then evaluated against these profiles to figure out whether they are suspicious or not. This evaluation can involve different metrics depending on the dimension that is being analyzed. The output of this analysis is a set of score values that indicate how uncommon each dimension of the transaction is compared to the profiles (see Figure 2 (a)) – for instance, a high score for 'amount' indicates that the transaction transfers an uncommon amount of money. Based on this set of single score values, an overall score is computed. If this overall score exceeds a given boundary value, the transaction is classified as fraudulent. An alert prompts the analyst to check the transaction and to decide, based on his/her experience, if this is indeed a case of fraud or if this is a false alarm.

VA support. Aiming to reduce the number of false alarms, we propose to add VA support to the existing transaction evaluation routine (see Figure 2 (b)). To this end, we use linked views of parallel coordinates (see Figure 1 (a)), a matrix of scatter plots (see Figure 1 (b)), and stacked bar charts (see Figure 1 (c)). Each line in the parallel coordinates and each dot in a scatter plot represents a transaction. The stacked bar chart shows the overall score value for a given transaction and its composition of single score values.

Each axis of the parallel coordinates encodes the score for a given dimension of the transaction while the last axis shows the overall score. Transactions that were classified as fraudulent by the automatic system are highlighted. Moreover, we provide interactive selection and filtering, as well as brushing & linking for the exploration of different dimensions, why transactions were classified as fraudulent, and how the overall score is composed. The scatter plot matrix shows pair wise scatter plots of the different dimensions and facilitates the detection of interesting patterns and outliers. Based on the insights from this exploration, the analyst may fine-tune the automatic classification system by adjusting the weights of different dimensions for score calculation, re-calculate the scores, and verify these changes on the basis of immediate visual feedback.

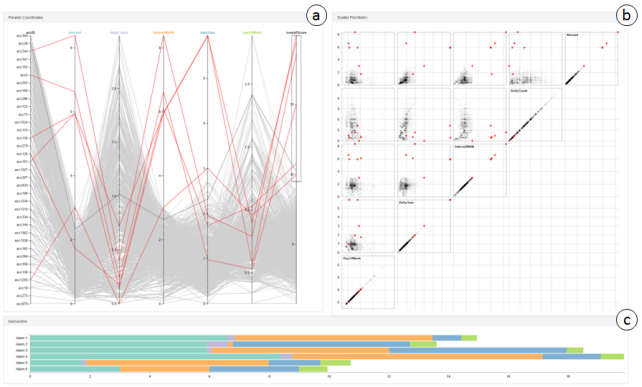

Figure 1: *Filter and selection interactions linking parallel coordinates (a), scatter plots (b), and horizontal stacked bar charts (c).*

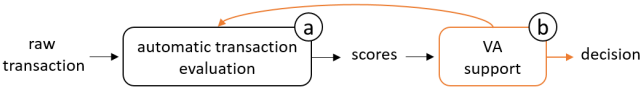

Figure 2: *The transaction evaluation system. The interactive VA approach for investigating transactions and fine-tuning the scoring system is highlighted in orange.*

4. Conclusion and Further Work

In this work we summarize the main challenges of fraud detection in the financial field. Based on these challenges we propose a VA approach for profile analysis to support fraud detection and user monitoring. We integrate this VA approach into the fraud detection process to efficiently combine AI techniques with interactive visual means. We believe our approach may as well be applicable in similar monitoring techniques and in similar domains, such as malware detection or tax usage analysis.

5. Acknowledgements

This work was supported by the Austrian Federal Ministry of Science, Research, and Economy via CVAST, a Laura Bassi Centre of Excellence (No. 822746).

References

[AMST11] AIGNER W., MIKSCH S., SCHUMANN H., TOMINSKI C.: *Visualization of time-oriented data*. Springer Science & Business Media, 2011. 1

[CCM*14] CARMINATI M., CARON R., MAGGI F., EPIFANI I., ZANERO S.: Banksealer: an online banking fraud analysis and decision support system. In *ICT Systems Security and Privacy Protection*. Springer, 2014, pp. 380–394. 1, 2

[CGK*07] CHANG R., GHONIEM M., KOSARA R., RIBARSKY W., YANG J., SUMA E., ZIEMKIEWICZ C., KERN D., SUDJIANTO A.: Wirevis: Visualization of categorical, time-varying data from financial transactions. In *Visual Analytics Science and Technology. VAST. IEEE Symposium on* (2007), IEEE, pp. 155–162. 1

[FHPS98] FAWCETT T., HAIMOWITZ I., PROVOST F., STOLFO S.: AI approaches to fraud detection and risk management. *AI Magazine 19*, 2 (1998), 107. 1

[HLN09] HUANG M. L., LIANG J., NGUYEN Q. V.: A visualization approach for frauds detection in financial market. In *Information Visualisation, 13th International Conference* (2009), IEEE, pp. 197–202. 1, 2

[KSH*99] KIRKLAND J. D., SENATOR T. E., HAYDEN J. J., DYBALA T., GOLDBERG H. G., SHYR P.: The nasd regulation advanced-detection system (ads). *AI Magazine 20*, 1 (1999), 55. 1

[LGM*15] LEITE R. A., GSCHWANDTNER T., MIKSCH S., GSTREIN E., KUNTNER J.: Visual analytics for fraud detection and monitoring. In *Visual Analytics Science and Technology (VAST), 2015 IEEE Conference on* (2015), IEEE, pp. 201–202. 2

[MA14] MIKSCH S., AIGNER W.: A matter of time: Applying a data–users–tasks design triangle to visual analytics of time-oriented data. *Computers & Graphics 38* (2014), 286–290. 2

[PLSG10] PHUA C., LEE V., SMITH K., GAYLER R.: A comprehensive survey of data mining-based fraud detection research. *arXiv preprint arXiv:1009.6119* (2010). 1

[YWW*07] YUE D., WU X., WANG Y., LI Y., CHU C.-H.: A review of data mining-based financial fraud detection research. In *Wireless Communications, Networking and Mobile Computing, 2007. WiCom 2007. International Conference on* (2007), IEEE, pp. 5519–5522. 1