ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Impact Factor: 8.206 **Volume 8, Issue 3, March 2025**

© 2025 IJMRSET | Volume 8, Issue 3, March 2025| DOI: 10.15680/IJMRSET.2025.0803257

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Enhanced Fraud Detection in Online Payment System using Fine Grained Co-Occurrence Analysis

Dr. D.J. Samatha Naidu, K. Venkata Ramya, A. Shobha

Professor, Dept. of MCA, APGCCS, New Boyanapalli, Rajampet, A.P., India

Assistant Professor, Dept. of MCA, APGCCS, New Boyanapalli, Rajampet, A.P., India

PG Student, Dept. of MCA, APGCCS, New Boyanapalli, Rajampet, A.P., India

ABSTRACT: The vigorous development of e-commerce breeds cybercrime. Online payment fraud detection, a challenge faced by online service, plays an important role in rapidly evolving e-commerce. Behaviour-based methods are recognized as a promising method for online payment fraud detection. However, it is a big challenge to build high-resolution behavioural models by using low-quality behavioural data. In this work, we mainly address this problem from data enhancement for behavioural modelling. We extract fine-grained co-occurrence relationships of transactional attributes by using a knowledge graph. Furthermore, we adopt the heterogeneous network embedding to learn and improve representing comprehensive relationships. Particularly, we explore customized network embedding schemes for different types of behavioural models, such as the population-level models, individual-level models, and generalized-agent based models. The performance gain of our method is validated by the experiments over the real dataset from a commercial bank. It can help representative behavioural models improve significantly the performance of online banking payment fraud detection. To the best of our knowledge, this is the first work to realize data enhancement for diversified behaviour models by implementing network embedding algorithms on attribute-level co-occurrence relationships

I. INTRODUCTION

Machine learning is an important component of the growing field of data science. With statistical methods, different type of algorithms is trained to make classifications or predictions, and to uncover key insights in this project. These insights subsequently drive decision making within applications and businesses, ideally impacting key growth metrics. Machine learning algorithms build a model based on this project data, known as training data, to make predictions or decisions without being explicitly programmed to do so. Machine learning algorithms are used in a wide variety of datasets, where it is difficult or unfeasible to develop conventional algorithms to perform the needed tasks.

© 2025 IJMRSET | Volume 8, Issue 3, March 2025| DOI: 10.15680/IJMRSET.2025.0803257

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

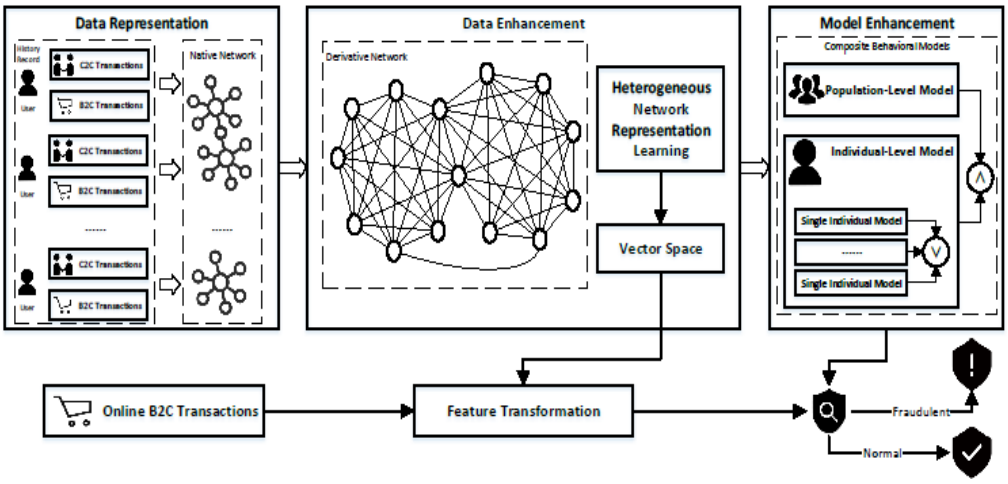

Figure 1: System Architecture.

II. LITERATURE REVIEW

Due to the immense growth of e-commerce and increased online based payment possibilities, credit card fraud has become deeply relevant global issue. Recently, there has been major interest for applying machine learning algorithms as data mining technique for credit card fraud detection. However, number of challenges appear, such as lack of publicly available data sets, highly imbalanced class sizes, variant fraudulent behaviour etc. In this paper we compare performance of three machine learning algorithms: Random Forest, Support Vector Machine and Logistic Regression in detecting fraud on real-life data containing credit card transactions. To mitigate imbalanced class sizes, we use SMOTE sampling method. The problem of ever-changing fraud patterns is considered with employing incremental learning of selected ML algorithms in experiments. The performance of the techniques is evaluated based on commonly accepted metric: precision and recall.

III. METHODOLOGY OF PROPOSED SURVEY

The Software Development Life Cycle (SDLC) is a series of stages that provide a structured approach to the software development process. It encompasses understanding the business requirements, eliciting needs, converting concepts into functionalities and features, and ultimately delivering a product that meets business needs. A proficient software developer should possess adequate knowledge to select the appropriate SDLC model based on project context and business requirements. Therefore, it is essential to select the right SDLC model tailored to the specific concerns and requirements of the project to ensure its success. To explore more about choosing the right SDLC model, you can follow this link for additional information. Furthermore, to delve deeper into software lifecycle testing and SDLC stages, follow the highlighted links here. The exploration will cover various types of SDLC models, their benefits, disadvantages, and when to use them. SDLC models can be viewed as tools to enhance product delivery. Therefore, understanding each model, its advantages, disadvantages, and the appropriate usage is crucial to determine which one suits the project context.

Types of Software developing life cycles (SDLC)

- Waterfall Model
- V-Shaped Model
- Evolutionary Prototyping Model
- Spiral Method (SDM)
- Iterative and Incremental Method
- Agile development

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

IV. CONCLUSION AND FUTURE WORK

For behavioural models in online payment fraud detection, we propose an effective data enhancement scheme by modelling co-occurrence relationships of transactional attributes. Accordingly, we design customized co-occurrence relation networks and introduce the technique of heterogeneous network embedding to represent online transaction data for different types of behavioural models, e.g., the individual-level and population-level models. The methods are validated by the implementation on a real-world dataset. They outperform the state of- the-art classifiers with lightweight feature engineering methods. Therefore, our methods can also serve as a feasible paradigm of automatic feature engineering. There are some interesting issues left to study. An interesting future work is to extend the data enhancement scheme into other types of behavioural models, e.g., the group-level models and generalized-agent-based models, except the population-level and individual-level models studied in this work. It would be interesting to investigate the dedicated enhancement schemes for more advanced individual-level models, since the adopted naive individual-level model does not fully capture the advantages of the proposed data representation scheme based on the techniques of heterogeneous network embedding. It is anticipated to demonstrate the generality of the proposed method by applying it to different real life application scenarios

REFERENCES

[1] B. Cao, M. Mao, S. Viidu, and P. S. Yu, "Hitfraud: A broadlearning approach for collective fraud detection in heterogeneousinformation networks," in Proc. IEEE ICDM 2017, New Orleans,LA, USA, November 18-21, 2017, pp. 769–774.

[2] M. A. Ali, B. Arief, M. Emms, and A. P. A. van Moorsel, "Doesthe online card payment landscape unwittingly facilitate fraud?"IEEE Security & Privacy, vol. 15, no. 2, pp. 78–86, 2017.

[3] X. Ruan, Z. Wu, H. Wang, and S. Jajodia, "Profiling onlinesocial behaviors for compromised account detection," IEEE Trans.Information Forensics and Security, vol. 11, no. 1, pp. 176–187, 2016.

[4] H. Yin, Z. Hu, X. Zhou, H. Wang, K. Zheng, N. Q. V. Hung, andS. W. Sadiq, "Discovering interpretable geo-social communitiesfor user behavior prediction," in Proc. IEEE ICDE 2016, Helsinki,Finland, May 16-20, 2016, pp. 942–953.

[5] Y.-A. De Montjoye, L. Radaelli, V. K. Singh et al., "Unique in the shopping mall: On the reidentifiability of credit card metadata,"Science, vol. 347, no. 6221, pp. 536–539, 2015.

[6] A. Khodadadi, S. A. Hosseini, E. Tavakoli, and H. R. Rabiee,"Continuous-time user modeling in presence of badges: A probabilistic approach," ACM Trans. Knowledge Discovery from Data, vol. 12, no. 3, pp. 37:1–37:30, 2018.

[7] F. M. Naini, J. Unnikrishnan, P. Thiran, and M. Vetterli, "Whereyou are is who you are: User identification by matching statistics,"IEEE Trans. Information Forensics and Security, vol. 11, no. 2, pp.358–372, 2016.

[8] M. Egele, G. Stringhini, C. Kruegel, and G. Vigna, "Towardsdetecting compromised accounts on social networks," IEEE Trans.Dependable and Secure Computing, vol. 14, no. 4, pp. 447–460, 2017.

[9] A. Alzubaidi and J. Kalita, "Authentication of smartphone usersusing behavioral biometrics," IEEE Communications Surveys andTutorials, vol. 18, no. 3, pp. 1998–2026, 2016.

[10] H. Mazzawi, G. Dalaly, D. Rozenblatz, L. Ein-Dor, M. Ninio, andO. Lavi, "Anomaly detection in large databases using behavioralpatterning," in Proc. IEEE ICDE 2017, pp. 1140–1149.

[11] Q. Cao, X. Yang, J. Yu, and C. Palow, "Uncovering large groupsof active malicious accounts in online social networks," in Proc.ACM SIGSAC 2014, pp. 477–488.

[12] X. Zhou, X. Liang, H. Zhang, and Y. Ma, "Cross-platform identification of anonymous identical users in multiple social media networks," IEEE Trans. Knowledge and Data Engineering, vol. 28,no. 2, pp. 411–424, 2016.

[13] T. W¨ uchner, A. Cislak, M. Ochoa, and A. Pretschner, "Leveraging compression-based graph mining for behavior-based malware detection," IEEE Trans. Dependable Secure Computing, vol. 16, no. 1,pp. 99–112, 2019.

[14] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in Proc. ACM SIGKDD 2016, CA, USA, August 13-17,2016, pp. 785–794.

[15] B. Jia, C. Dong, Z. Chen, K. Chang, N. Sullivan, and G. Chen,"Pattern discovery and anomaly detection via knowledge graph,"in Proc. FUSION 2018, Cambridge, UK, July 10-13, 2018, pp. 2392–2399.

[16] P. Cui, X. Wang, J. Pei, and W. Zhu, "A survey on network embedding," IEEE Trans. Knowledge and Data Engineering, vol.31,no. 5, pp. 833–852, 2019.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

[17] M. Abouelenien, V. P´erez-Rosas, R. Mihalcea, and M. Burzo,"Detecting deceptive behavior via integration of discriminative features from multiple modalities," IEEE Trans. Information Forensics and Security, vol. 12, no. 5, pp. 1042–1055, 2017.

[18] W. Youyou, M. Kosinski, and D. Stillwell, "Computer- basedpersonality judgments are more accurate than those made by humans," PNAS, vol. 112, no. 4, pp. 1036–1040, 2015.

[19] V. Sekara, A. Stopczynski, and S. Lehmann, "Fundamental structures of dynamic social networks," PNAS, vol. 113, no. 36, pp.9977–9982, 2016.

[20] K. Rzecki, P. Plawiak, M. Niedzwiecki, T. Sosnicki, J. Leskow, and M. Ciesielski, "Person recognition based on touch screen gesturesusing computational intelligence methods," Information Science,vol. 415, pp. 70– 84, 2017.

[21] S. Lee and J. Kim, "Warningbird: Detecting suspicious urls intwitter stream," in Proc. NDSS 2012, San Diego, California, USA,February 5-8, 2012, vol. 12, pp. 1–13.

[22] G. Stringhini, P. Mourlanne, G. Jacob, M. Egele, C. Kruegel, and G. Vigna, "EVILCOHORT: detecting communities of malicious accounts on online services," in Proc. USENIX Security 2015,Washington, D.C., USA, August 12-14, 2015, pp. 563–578.

[23] Z. Meng, L. Mou, and Z. Jin, "Hierarchical RNN with staticsentence- level attention for text-based speaker change detection,"in Proc. ACM CIKM 2017, Singapore, November 06 - 10, 2017, pp.2203–2206.

[24] A. Rawat, G. Gugnani, M. Shastri, and P. Kumar, "Anomalyrecognition in online social networks," International Journal of Security and Its Applications, vol. 9, no. 7, pp. 109–118, 2015.

[25] C. VanDam, J. Tang, and P. Tan, "Understanding compromisedaccounts on twitter," in Proc. ACM WI 2017 , Leipzig, Germany,August 23-26, 2017, pp. 737–744.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com