Fraud Detection for O2O Commerce with Offline Users' Behavior and Relation Embedding

Abstract

Transaction fraud has become one of the most serious threats to O2O commerce recently. In this paper, we present RIC (Risk of Intelligence Center), a novel deep-learning based transaction fraud detection system designed by our team and deployed at Koubei.com, one of the largest O2O platforms in China with over 330 million active users and 2.5 million active local shops. RIC captures detailed information on offline users' behavior events using neuron-network-based embedding, and models relations among them with the graph-embedding techniques. Furthermore, RIC provides application-specific optimizations including imbalanced learning, realtime detection, and incremental model update. With large scale of real production data for several months, we show that RIC achieves over 2x improvement over the existing fraud detection approaches.

1 INTRODUCTION

Fraud in e-commerce is commonly defined as transaction with the intention to achieve financial gain on false ground by an illegal means[Alexopoulos et al., 2007]. This kind of fraud can be internal when an staff commits frauds against the organization,[Phua et al., 2010] or external which involves a variety of members including vendors, customers, shop owner or thefts from third party[Song and Gangopadhyay, 2013]. These frauds are significantly detrimental to the organizations or companies, especially in the form of loss of subsidy. In 2017, Internet Crime Complaint Centre (IC3) had received about 300,000 complaints, which caused directly financial loss of over 1.4 billion USD[ICR2017, 2017].

Online-to-offline (O2O) commerce is a business strategy that draws potential customers from online channels to make purchases in physical stores. For a quick definition, one might say O2O is anything digital which brings people to shop offline, in real-world stores. With the rapid development of O2O commerce, many companies such as Koubei, Inc. in this field are spending more money than ever on sales promotion by providing subsidies to sellers and buyers, which leads to various frauds aiming at illegal gain. Unlike the fraud of e-commerce online, which has the whole control about the transaction process, the fraud detection for O2O commerce has to face a new problem of losing control about the offline process. With the subsidies of sales promotion, offline business owners may make a fake transaction but also enjoy the related subsidy from the O2O platform. To achieve the illegal gain, some owners of local stores may ask their staffs or relatives, even hired professional fakers to act as common customers and join the transaction. Most of these deceptive users are spatially disperse and usually make fake visits at the local stores to finish transactions. These are the mainly illegal means in O2O transaction frauds, from which we can find the impressive importance of relations and users' behaviors that will be proven by the following data analysis.

However, identifying frauds from the offline massive behavior and relation data is labor-intensive and time-consuming. It is necessary to develop automatic and assistant approaches to facilitate real-time fraud tracking and debunking with the offline users logs.

In this article, we present the analysis on the problem of O2O frauds and provide an effective approach to model offline information using deep-learning and graph embedding based models. We first give an analysis on the impact of users' behavior and relations on fraudulent transaction with the offline users' logs in Koubei,Inc. and then introduce the architecture about our detection system. In the following two parts, user behavior model using bidirectional RNNs and relation model

with graph embedding are described respectively. At last, we prove that our approach is more effective than the common solutions for online transaction fraud detection using real production data. And we also compare the performance of our methods with the state-of-art baseline[Li et al., 2015] ,which first introduce the offline users' spatial features for fraud detection.

2 RELATED WORK

In the field of fraud detection, there are increasing works to tackle this challenge, including unsupervised learning methods, supervised/semi-supervised learning methods, deep learning based approaches and network/graph based solutions. Also, several industrial fraud detection system has been developed for e-commerce platforms. [Weng et al., 2018] achieved an efficient and scalable AnTi-Fraud system (ATF) to detect e-commerce frauds for large-scale e-commerce platform. To our knowledge, it's the first efficient fraud detection system for large-scale e-commerce platforms.

Supervised/Semi-supervised/Unsupervised learning methods play an important role in transaction fraud detection. [Kumar and Gupta, 2018] provides an empirical study and analysis of supervised learning techniques on a benchmark credit card transaction dataset. [Teh et al., 2018] presented an fraud detection technique in the form of a rudimentary fraud detection system that utilizes consumer spending behavior by building a consumer profile with transaction time, amount, and geographical location. [Robinson and Aria, 2018] automatically creates, updates, and compares hidden Markov models (HMM) of merchant terminals to detect prepaid cards fraud. [Ram and Gray, 2018] proposed the use of a variant of density estimation trees (DETs) for anomaly detection using distributional properties of the data. Other fraud detection approaches includes both supervised and unsupervised methods like decision trees[Kokkinaki, 1997], clustering techniques[Song and Gangopadhyay, 2013], and genetic algorithms[Mendes et al., 2001].

Deep learning based approaches have significant advantages over the traditional feature expression and learning. [Zhang et al., 2018b] proposed a fraud detection model based on the CNN which constructs an input feature sequencing layer that implements the reorganization of raw transaction features to form different convolutional patterns. A sandwich-structured sequence learning architecture has been proposed by stacking an ensemble model, a deep sequential learning model and another top-layer ensemble classifier in proper order [Lp et al., 2018]. [Zheng et al., 2018b] proposed a new generative adversarial network (GAN) based model that establishes a minimax game between a discriminator and a generator to accurately discriminate between positive samples and negative samples in the data distribution. [Cai et al., 2018] and [Zhao et al., 2018] handle the fraud transactions detection by reinforcement learning improving the platform's impression allocation mechanism to maximize its profit and reduce the sellers' fraudulent behaviors simultaneously. Rather than relying on a snapshot of behaviors, [Guo et al., 2018] detect frauds by considering a complete behavioral sequence. [Jurgovsky et al., 2018] phrased the fraud detection problem as a sequence classification task and employ LSTM networks to incorporate transaction sequences.

Graph-based algorithms are also studied to identify transaction frauds. [Zheng et al., 2018a] proposed logical graph of behavior profiles to characterize the diversity of transaction behaviors and then use them to verify if an incoming transaction is a fraud. [Cao et al., 2018] proposed a novel approach leveraging a graph based perspective to uncover relationships among suspicious transactions.

The class imbalance problem may lower the performance of some fraud detection models. [Zhang et al., 2018a] proposed a comprehensive model called clustering tree that can handle the class imbalance problem in the fraud transaction detection. [Dal Pozzolo et al., 2018] designed a novel learning strategy that effectively addresses class imbalance, concept drift, and verification latency in the fraud tranaction detection. [Carta et al., 2019] proposed a novel data intelligence technique based on a Prudential Multiple Consensus model to maximize the effectiveness of the model in detecting fraudulent transactions regardless the presence of any data imbalance.

Previous works, which talk about frauds of transactions, mostly focus on the transactions from online commerce systems[Abdallah et al., 2016]. And the fraud detecting application contains credit card, tax system[de Roux et al., 2018], phone call [Ying et al., 2018], fee [Edwards et al., 2017], healthcare insurance , etc. These works mainly focus on modeling the online users logs and use the models for fraud detecting. However, there are less works about the offline information modeling, which is very important for fraud identifying in O2O commerce.

3 O2O TRANSACTION FRAUDS ANALYSIS

In order to figure out how to identify transaction frauds in O2O commerce, we sample thousands of instances from frauds found by human risk experts in Koubei, Inc.

These frauds are randomly sampled and evaluated through offline double check by at least two risk experts from the business team to make sure that the precision is at least 99%. In the fraud-verifying process , we find that there are mainly three kinds of abnormal: user behavior abnormal, relation abnormal, and finlink abnormal.

User behavior abnormal means that the offline buyers who were supposed to visit the local store to complete a transaction but they did not, and instead they made a series of deceptive spatiotemporal events to complete the transactions. Relation abnormal is a broad concept which close relations were in the process of transaction where they should not be, such as the staff, family, etc of business owner or an ID sharing common profile with the owners.

Cash flow process denotes the path of cash transferring and is the common variable of any transaction including online and O2O commerce. And the finlink is the abnormal cash flow process which is also not supposed to be in the process of trade. With these assumptions of abnormalities, we classify the frauds into three groups, as showed in Figure 1. Obviously, with the sum of ratio above 70%, offline user behavior and relations, as the new variables brought by O2O commerce, are dominating causes of frauds.

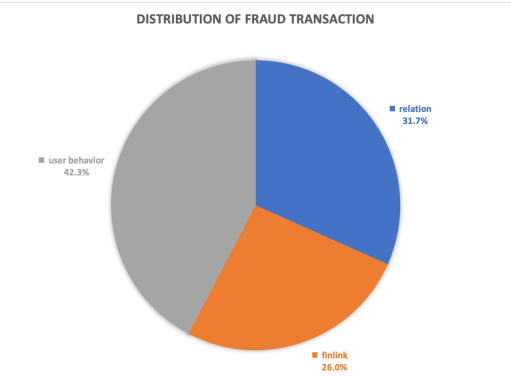

Figure 1: analysis of frauds sampled.

As for the influence of user behavior to frauds, we take out several days of users' logs in Koubei, Inc. According to the strategy of O2O commerce, customers who bought services from the online platform are supposed to reach the local stores for verifying and enjoying them. Therefore, the transaction frauds using the means of making deceptive visit to the stores should leave inevitable information about the users' behavior. With this assumption, we roughly compute the distance from users to the serving stores when the users enjoy the services. As showed in Figure 2, we can conclude that the users' behavior, especially for spatial-temporal type, really works for the identification of frauds related to local service.

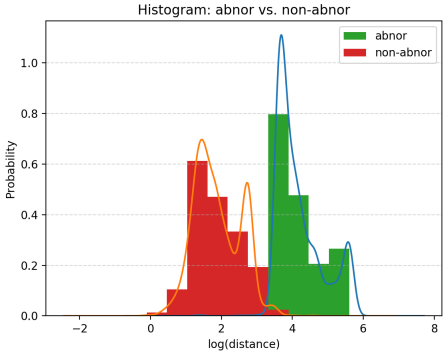

Figure 2: spatial sensitivity of frauds

In Table 1 , we present the analysis on statistics of relations in transaction frauds, which has been summarized by risk experts as relational abnormality. The proportion of each relation reflects the importance of identifying frauds in corresponding field. We can see that frauds due to the relation of employee is so impressive that we must find a model to detect them.

Table 1: abnormal relations of frauds

#employee	#device	#same person	#other
70.3%	22%	6.7%	1%

4 PROBLEMS AND OUR APPROACH

4.1 Problem formulation

Following the analysis above, there are two problems to be addressed, namely modeling users' 1) behaviors and 2) relations, both are definitely important for fraud detection in O2O commerce. In this section, we mainly give the formulation of the problems and some notations, which could be used in the following parts. As Eq.1 shows, the loss of identifying transaction frauds can be noted as a classifier with common loss function as usual:

$$Loss(W) = \frac{1}{N}\sum_{i=1}^{N} L_i(f(x_i, W), y_i) + \lambda R(W) \quad (1)$$

where $L_i = -logP(Y = y_i|X = x_i)$, $x_i = (x_{1i}, x_{2i}, x_{3i}, ..., x_{mi})$, N is the number of transactions, x_i is m dimensional feature-vector for the ith trade, $y_i \in \{0, 1\}$ denotes whether the ith trade is fraudulent(1) or not(0), $f(x_i, W)$ is some function combining information from x_i and W. The $R(W)$ is the regularization for suitable parameters W.

In the application of frauds detecting, lots of attention is paid to events representing, which is somewhat similar to other machine learning problems. In this work, we want to represent the users' spatial-temporal preference, such as velocity and so on, to latent and dense vectors, which may produce more effective ability to identify whether the users reach the local stores or not.

$$Loss_b(W) = \frac{1}{N}\sum_{i=1}^{N} L_i(f(s_i, W), y_i) + \lambda R(W) \quad (2)$$

where $L_i = -logP(Y = y_i|X = s_i)$, $s_i = (l_1, l_2, l_3, ..., l_k, l_{k+1}, ..., l_n)$, $|t_k - t_i| \leq \varepsilon$. s_i is the ith instance of users' moving behavior sequence. l_i is the ith spatial-temporal event. t_i is the time of user reaching local stores. t_k is the time we get spatial temporal information closest to t_i. $Loss_b$ denotes the loss function about whether users' instance is spatial-temporal abnormal or not.

$$Loss_r(W) = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} L_{ij}(f(g_{ij}, W), y_{ij}) + \lambda R(W) \quad (3)$$

where $L_{ij} = -logP(Y = y_{ij}|X = g_{ij})$, $g_{ij} =< x_i, x_j, w_{ij} >$ is the graph of relations, which derive from the real application problem. $E = \{x_i, 1 <= i <= n\}$ denotes the entities in application, and w_{ij} is the weight computed from the information of x_i and x_j. And $Loss_r$ is the loss function to evaluate whether the two nodes has a relation.

4.2 Bidirectional RNNs based User Behavior Modeling

Given a user's offline behavior instance with moving sequence denoted by s_i^k, where k is the indice of event which happens nearby the time t_i he make a visit to local store. Following the Eqs 2, the representation of users' behavior events plays an important role before we optimize the loss function. As to the sequence $s_i^k = (l_1, l_2, l_3, ..., l_{k-1}, l_k, l_{k+1}, ..., l_n)$, we can see there are two parts of events including before and after in time, which corresponds to the architecture of Bidirectional RNNs. The mobile apps such as Koubei or Alipay, can collect the location events of users, who use GPS carrying longitude and latitude. We denote a local user's event as $l_k = (l_k^x, l_k^y, t_k)$, where l_k^x is longitude, l_k^y is the latitude and t_k is the time of data collected. And the neural architecture of user behavior modeling, which we call DSTM for deep learning based spatial-temporal modeling, is depicted as Figure 3. L_1 is the input-layer for l in the form of fully linear connection. And L_2 is the Bidirectional RNNs layer, using LSTM or GRU as the core net, which gives us the ability of modeling local behavior information by end-to-end training. In order to combining the long term information with the short term one, we usually use the LSTM 4 in application. The layer L_3 take the embedding representation of behavior information and user-wide features as the final input vector for loss optimization.

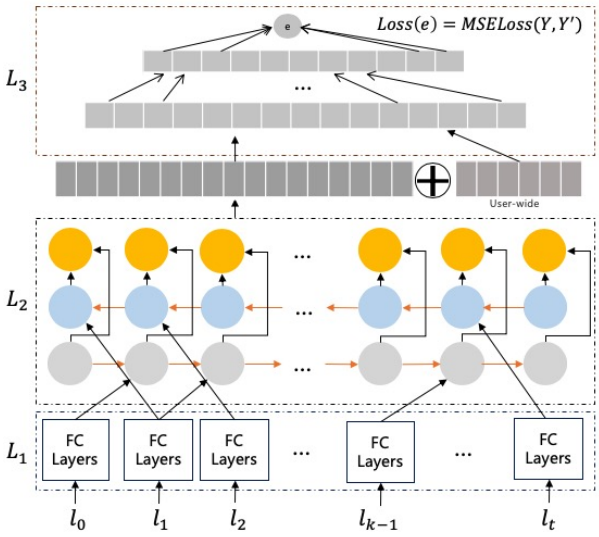

Figure 3: neural architecture of the deep spatial-temporal model (DSTM)

$$\begin{aligned} z_t &= \sigma(W_z \cdot [h_{t-1}, x_t]) \\ r_t &= \sigma(W_r \cdot [h_{t-1}, x_t]) \\ \tilde{h}_t &= tanh(W \cdot [r_t * h_{t-1}, x_t]) \\ h_t &= (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \end{aligned} \quad (4)$$

4.3 Context-aware Graph Embedding based Relation Modeling (CGRM)

In this section, we use the following notation to introduce the methods:U denotes the set of users, and $u_i \in U$ is the ith user, a customer in our context. S is the set of all local stores, and s_i is the ith one of S. E is the set of some relational entities such as wifi or device, and x_i is the ith one of it. e_i denotes the embedding vector of x_i in latent space, which is the output of our method. According to Eq.3, the structure of the data $G = \{g_{ij} =< x_i, x_j, w_{ij} >, 1 \leq i, j \leq n\}$ is a complex network or graph. To learn a vector representation of user associated with some elements in E, we construct a graph based on entities' relations among them.

Under the context of a specific problem, the graph we

build should be fully convinced by its' intent. The relation of employee modeling with wifi graph embedding is illustrated in the figure 4, which shows the procedure of our method. we define the wifi ID set by E,

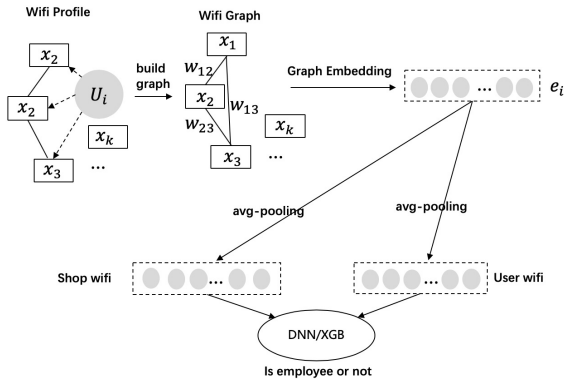

Figure 4: architecture of employee relation modeling using wifi graph embedding

and $E_i = \{x_k, k \in \Omega_i\}$ denotes the user u_i wifi profile, where Ω_i is the indices set of wifi IDs used by u_i. The information of a user i using wifi j can be stated as $< u_i, x_j, t_{ij} >$, where t_{ij} is the duration time. It is obviously that the frequency and duration time of using wifi behavior contains enough information about employee relationship. With these consideration, we want to find the representation of users and shops in wifi latent space as showed in Figure 5.

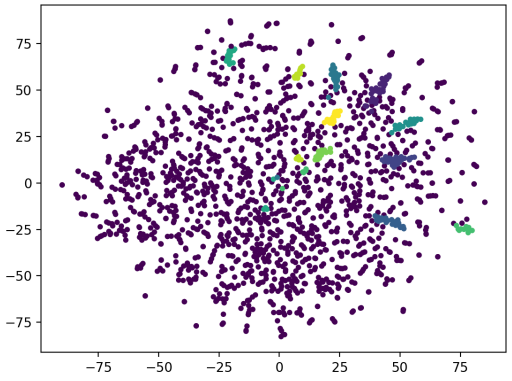

Figure 5: latent space of relations embedding

The procedure of building the graph with nodes $E = \{x_i\}$ and edges as follow:

$$w_{ij} = \frac{1}{|U|} \sum_{k \in U} \frac{h_i^k * c(t_{ki}) + h_j^k * c(t_{kj})}{h_i^k + h_j^k} \tag{5}$$

where $h_i^k = \# < u_k, x_i, t_{ki} >$ denotes the frequency of user k connects wifi i. The $c(t)$ is a context function defined for specific application problem.

Given the graph $G = \{< x_i, x_j, w_{ij} >\}$, we can apply conventional network embedding techniques, such as random walk or line, to get the latent representation e_i of node x_i, where i is the indice of wifi IDs. For a user $u_i \in U$ and a store $s_k \in S$, we can get the latent representation with average-pooling method in wifi embedding space:

$$\begin{aligned} r(u_i) &= \frac{1}{|\Omega_i|} \sum_{k \in \Omega_i} h_k^i * e_k \\ r(s_k) &= \frac{1}{|N_k|} \sum_{l \in N_k} e_l \end{aligned} \tag{6}$$

where $r(x)$ is the representation function of entity, N_k is the indice set for all wifi IDs, connected by users who make face-to-face transactions with the local store s_k at the same time. To figure out the relation of user u_i and store s_k in the latent space, we can perform not only the supervised methods, such as xgboost or dnn, but also just compute the similarity in a cosine way as Eq 7.

$$sim(u_i, s_k) = \frac{\|u_i * s_k\|}{\|u_i\| * \|s_k\|} \tag{7}$$

4.4 Architecture for our methods

Figure 6 shows the basic architecture of our system RIC and the flow of data processing towards the final business application in App client. Our detector system contains three layers of fraud identifiers: FLM, DSTM and CGRM, which give us the ability of combining users' behavior and relations with cash transferring events. The result of each layer give the probability of abnormal, which used as inputs for supervised classifier to identify real frauds. Due to the rareness of labels in fraud detection, conventional supervised methods may not get helpful performance, for which we adopted downsampling and PU learning methods. And rules engine, which is part of RIC, takes the responsibility of managing different frauds with enough flexibility. App client will receive the fraudulent transaction, the output of RIC, with detailed information for punishment.

5 EVALUATION

5.1 Dataset

It is well known that the details of commercial data are secrets for any company or organization because of sensitivity required. Therefore, in our experiments, the logs

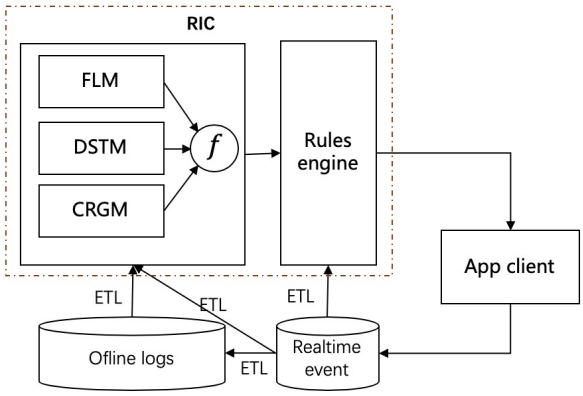

Figure 6: architecture of the RIC

of transaction used will not be described detailedly. The dataset is composed of several sources, including realtime events and offline logs from different business units, mainly Koubei, in Alibaba Group. After randomly sampling from logs for several months, we draw a portion of real production dataset, which contains 6.8 million instances with 2.5 million local stores and 3.3 billion users. Using the module of ETL, we extract the necessary offline user behavior and relational information, such as the speed of moving, stores visited, mobile phones used, cash-transferring records, etc. All of these private information like real names or telephone numbers are anonymized and thus masked.

5.2 Baseline

Cash transfering process has been mainly used as a practical method for identifying frauds in commerce, not only helpful for online e-commerce but also applicable for O2O platform. In our work, we call this method finlink model (FLM), which has been proved as a fundamental and effective way of fraud detection in large scale real production data.

By matching transaction instances with the cash transfering logs in Ant Finance, we are able to identify users who got paid by persons closely related to the shop, or even shop owner itself. Given a transaction τ and its happening time τ_t, the matching instances set T can be extracted during the specific period $|t - \tau_t| \leq \varepsilon$, where t is the time of cash transferring event and ε is an empirical value. We denote the buyer of transaction τ as $\tau(u)$ and seller as $\tau(s)$. To obtain illegally financial gain, buyers would have to get back the cost of transaction, which can be measured as follow:

$$cost(T|\tau) = \sum_{\forall \pi \in \Theta_\tau} \pi.amount \tag{8}$$

where $\Theta_\tau = \{\forall \pi \in T | \pi(s) \to \pi(u), |\pi_t - \tau_t| < \varepsilon\}$.

The probability evaluating whether τ is fraudulent, can be computed as this:

$$P(\tau) = \frac{1}{2} + \sigma(-|\frac{cost(T|\tau)}{cost(T|\tau) + \tau.amout} - 0.5|) \tag{9}$$

where σ is the sigmoid function $\sigma(x) = \frac{1}{1+e^{-x}}$.

5.3 Experiments

In this section, several experiments are carried out to show the performance of our proposed models: a) experiments of DSTM with different setup; b) supervised models using CGRM; c) improvements for our proposed methods comparing with conventional approach in real production data.

As showed in Figure 7, we can see the different performances for several setups in the way of ROC. With the same user behaviors' sequences for one month as the input, the results show that DSTM using gru is impressive better than conventional shallow model such as logistic regression and is also slightly better than models using uni-gru. That means the performance of our approach DSTM has the best ability in the problem of user's behavior modeling. And we also combine user's wide features with embedding one, but make little difference.

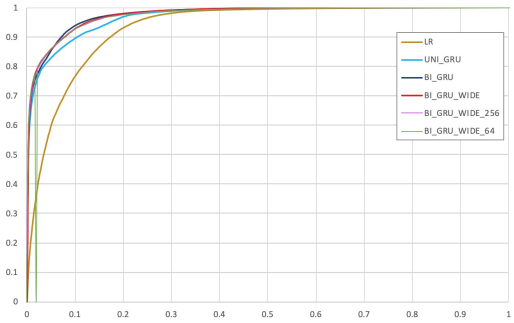

Figure 7: ROC curves of DSTM with different setup. x-axis is false positive rate, y-axis is true positive rate.

As to the proposed CGRM, experiments in Figure 8 show that supervised methods, xgboost and random forest, have a better performance with the same embedding vector as input. To model the relationship of employee between buyers and sellers and investigate influence of different context, we applied CGRM with modified context and got the experiments' results showed in Table 2. And the context, namely the function $c(t)$, is modified as follow: unique number of users, who connected the wifi once, denoted as UC; accumulated duration time of all users denoted as DT; MW means sum of duration divided by #users; WW is short for

working time and using wifi, much more related to the scenario of employee in store; MWW is mean value of WM. Considering different graph embedding techniques may play an extra effect on the performance of our experiments, we also use them combined with the mentioned context respectively. The experiments setup: a) DeepWalk,number-walks=10, walk-length=80, workers=8, window-size =10; b) Line, negative-ratio=5, order=3; c) Node2Vec number-walks=10, walk-length=80, workers=8, window-size =10, p=0.25, q=0.25; d) SDNE encoder-list=[1000, 128], alpha=1e-6, beta=5, nu1=1e-5, nu2=1e-4, batch size=200, learning rate=0.01. Both metrics of auc and f1 indicate the performance of CGRM using node2vec and the context of MMW is the best.

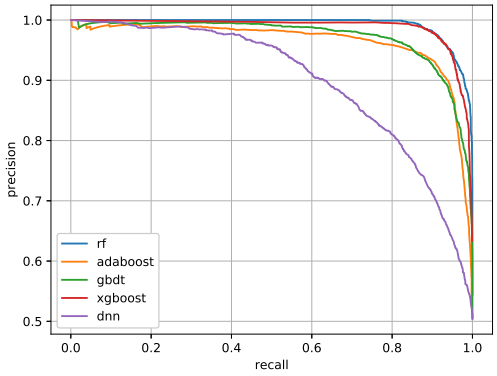

Figure 8: Precision-recall curves of different models using relation embedding by CGRM.x-axis is recall rate, y-axis is precision rate

The method [Li et al., 2015], named as DPS model in this paper, is proposed as a practical solution for fraud detection in Dianping Inc. , which is also one of the largest O2O platform in china. We use DPS as state-of-the-art method in O2O commerce, and compare the performance of our approach with it. The result, showed in Figure 9, indicates our approach produce 10x improvements over that

After our proposed models DSTM and CGRM deployed in real production environment, our system RIC shows an impressive 2.x improvement over previous baseline FLM, as showed in Table 3.

6 CONCLUSTIONS

In this paper, we present a way of modeling offline users' information using deep-learning and network embedding techniques, and proved its effectiveness in O2O fraud detection with enough experiments. The system RIC pow-

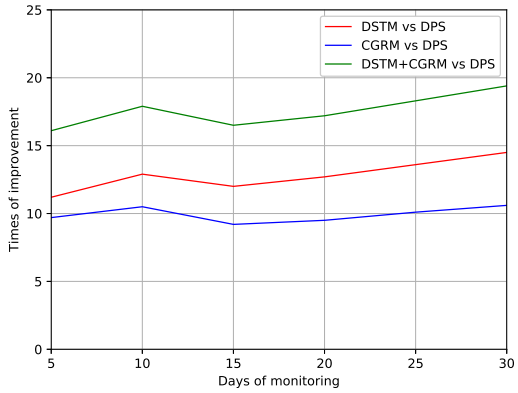

Figure 9: improvements of proposed Models vs DPS.

ered by our proposed models has been undergone extensive tests and deployed in large scale production environment. With the rapid development of O2O commerce, there will be new difficult problems for frauds detection. And the proposed methods in this work may give us some hints for solving them.

Table 2: Performance of CGRM with different context

Context	UC		DT		MW		WW		MWW	
	auc	f1	auc	f1	auc	f1	auc	f1	auc	f1
DeepWalk	0.943	0.905	0.923	0.892	0.915	0.893	**0.948**	**0.906**	0.924	0.891
Line	**0.968**	**0.928**	0.926	0.899	0.925	0.901	0.948	0.911	0.958	0.905
Node2Vec	0.915	0.864	0.943	0.903	0.926	0.913	0.965	0.908	**0.967**	**0.931**
SDNE	0.931	0.879	0.935	0.893	0.933	0.894	**0.961**	**0.918**	0.940	0.888
Mean	0.939	0.894	0.932	0.897	0.925	0.900	**0.955**	**0.911**	**0.947**	**0.904**

Table 3: Improvements of proposed Models vs FLM in large scale production data

	1 Day	3 Days	1 Week	1 Month	2 Months	1 Season
DSTM vs. FLM	112.83%	105.31%	109.38%	134.60%	147.09%	156.41%
CGRM vs. FLM	92.99%	91.57%	92.35%	98.58%	105.88%	109.08%
DSTM+CGRM vs. FLM	**160.58%**	**152.73%**	**155.18%**	**180.64%**	**198.89%**	**207.68%**

References

[Abdallah et al., 2016] Abdallah, A., Maarof, M. A., and Zainal, A. (2016). Fraud detection system: A survey. *Journal of Network & Computer Applications*, 68:90–113.

[Alexopoulos et al., 2007] Alexopoulos, P., Kafentzis, K., Benetou, X., Tagaris, T., and Georgolios, P. (2007). Towards a generic fraud ontology in e-government. In *Ice-b - International Conference on E-business*.

[Cai et al., 2018] Cai, Q., Filos-Ratsikas, A., Tang, P., and Zhang, Y. (2018). Reinforcement mechanism design for fraudulent behaviour in e-commerce. In *Thirty-Second AAAI Conference on Artificial Intelligence*.

[Cao et al., 2018] Cao, B., Mao, M., Viidu, S., and Yu, P. (2018). Collective fraud detection capturing inter-transaction dependency. In *KDD 2017 Workshop on Anomaly Detection in Finance*, pages 66–75.

[Carta et al., 2019] Carta, S., Fenu, G., Recupero, D. R., and Saia, R. (2019). Fraud detection for e-commerce transactions by employing a prudential multiple consensus model. *Journal of Information Security and Applications*, 46:13–22.

[Dal Pozzolo et al., 2018] Dal Pozzolo, A., Boracchi, G., Caelen, O., Alippi, C., and Bontempi, G. (2018). Credit card fraud detection: a realistic modeling and a novel learning strategy. *IEEE transactions on neural networks and learning systems*, 29(8):3784–3797.

[de Roux et al., 2018] de Roux, D., Perez, B., Moreno, A., Villamil, M. d. P., and Figueroa, C. (2018). Tax fraud detection for under-reporting declarations using an unsupervised machine learning approach. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 215–222. ACM.

[Edwards et al., 2017] Edwards, M., Peersman, C., and Rashid, A. (2017). Scamming the scammers: towards automatic detection of persuasion in advance fee frauds. In *Proceedings of the 26th International Conference on World Wide Web Companion*, pages 1291–1299. International World Wide Web Conferences Steering Committee.

[Guo et al., 2018] Guo, J., Liu, G., Zuo, Y., and Wu, J. (2018). Learning sequential behavior representations for fraud detection. In *2018 IEEE International Conference on Data Mining (ICDM)*, pages 127–136. IEEE.

[ICR2017, 2017] ICR2017 (2017). 2017 internet crime report.

[Jing et al., 2016] Jing, M., Wei, G., Mitra, P., Kwon, S., Jansen, B. J., Wong, K. F., and Cha, M. (2016). Detecting rumors from microblogs with recurrent neural networks. In *International Joint Conference on Artificial Intelligence*.

[Jurgovsky et al., 2018] Jurgovsky, J., Granitzer, M., Ziegler, K., Calabretto, S., Portier, P.-E., He-Guelton, L., and Caelen, O. (2018). Sequence classification for credit-card fraud detection. *Expert Systems with Applications*, 100:234–245.

[Kokkinaki, 1997] Kokkinaki, A. I. (1997). On atypical database transactions: identification of probable frauds using machine learning for user profiling.

In *IEEE Knowledge & Data Engineering Exchange Workshop*.

[Kumar and Gupta, 2018] Kumar, A. and Gupta, G. (2018). Fraud detection in online transactions using supervised learning techniques. In *Towards Extensible and Adaptable Methods in Computing*, pages 309–321. Springer.

[Li et al., 2015] Li, H., Chen, Z., Mukherjee, A., Bing, L., and Shao, J. (2015). Analyzing and detecting opinion spam on a large-scale dataset via temporal and spatial patterns.

[Lp et al., 2018] Lp, X., Yu, W., Luwang, T., Zheng, J., Qiu, X., Zhao, J., Xia, L., and Li, Y. (2018). Transaction fraud detection using gru-centered sandwich-structured model. In *2018 IEEE 22nd International Conference on Computer Supported Cooperative Work in Design ((CSCWD))*, pages 467–472. IEEE.

[Mendes et al., 2001] Mendes, R. R. F., Voznika, F. B. D., Freitas, A. A., and Nievola, J. C. (2001). Discovering fuzzy classification rules with genetic programming and co-evolution. In *European Conference on Principles of Data Mining & Knowledge Discovery*.

[Phua et al., 2010] Phua, C., Lee, V., Smith, K., and Gayler, R. (2010). A comprehensive survey of data mining-based fraud detection research. *arXiv preprint arXiv:1009.6119*.

[Ram and Gray, 2018] Ram, P. and Gray, A. G. (2018). Fraud detection with density estimation trees. In *KDD 2017 Workshop on Anomaly Detection in Finance*, pages 85–94.

[Robinson and Aria, 2018] Robinson, W. N. and Aria, A. (2018). Sequential fraud detection for prepaid cards using hidden markov model divergence. *Expert Systems With Applications*, 91:235–251.

[Song and Gangopadhyay, 2013] Song, C. and Gangopadhyay, A. (2013). A novel approach to uncover health care frauds through spectral analysis. In *IEEE International Conference on Healthcare Informatics*.

[Teh et al., 2018] Teh, B., Islam, M. B., Kumar, N., Islam, M. K., and Eaganathan, U. (2018). Statistical and spending behavior based fraud detection of card-based payment system. In *2018 International Conference on Electrical Engineering and Informatics (ICELTICs)(44501)*, pages 78–83. IEEE.

[Weng et al., 2018] Weng, H., Li, Z., Ji, S., Chu, C., Lu, H., Du, T., and He, Q. (2018). Online e-commerce fraud: a large-scale detection and analysis. In *2018 IEEE 34th International Conference on Data Engineering (ICDE)*, pages 1435–1440. IEEE.

[Ying et al., 2018] Ying, J. J.-C., Zhang, J., Huang, C.-W., Chen, K.-T., and Tseng, V. S. (2018). Fraudetector+: An incremental graph-mining approach for efficient fraudulent phone call detection. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 12(6):68.

[Zhang et al., 2018a] Zhang, Y., Liu, G., Zheng, L., Yan, C., and Jiang, C. (2018a). A novel method of processing class imbalance and its application in transaction fraud detection. In *2018 IEEE/ACM 5th International Conference on Big Data Computing Applications and Technologies (BDCAT)*, pages 152–159. IEEE.

[Zhang et al., 2018b] Zhang, Z., Zhou, X., Zhang, X., Wang, L., and Wang, P. (2018b). A model based on convolutional neural network for online transaction fraud detection. *Security and Communication Networks*, 2018.

[Zhao et al., 2018] Zhao, M., Li, Z., An, B., Lu, H., Yang, Y., and Chu, C. (2018). Impression allocation for combating fraud in e-commerce via deep reinforcement learning with action norm penalty. In *IJCAI*, pages 3940–3946.

[Zheng et al., 2018a] Zheng, L., Liu, G., Yan, C., and Jiang, C. (2018a). Transaction fraud detection based on total order relation and behavior diversity. *IEEE Transactions on Computational Social Systems*, (99):1–11.

[Zheng et al., 2018b] Zheng, Y.-J., Zhou, X.-H., Sheng, W.-G., Xue, Y., and Chen, S.-Y. (2018b). Generative adversarial network based telecom fraud detection at the receiving bank. *Neural Networks*, 102:78–86.