Unsupervised Fraud Transaction Detection on Dynamic Attributed Networks

Yangyang Hou, Daixin Wang, Binbin Hu, Ruoyu Zhuang, Zhiqiang Zhang,,
Jun Zhou, Feng Zhao, Yulin Kang, and Zhanwen Qiao

Ant Group, China
{yangyanghou.hyy, daixin.wdx, bin.hbb, zhuangruoyu.zry, lingyao.zzq,
jun.zhoujun, zhaofeng.zhf, yulin.kyl, zhanwen.qiaozw}@antgroup.com

Abstract. Fraud transaction detection is a pressing need in industrial applications, aiming to detect the fraud for a transaction involving the buyer and the seller. Due to the prohibitive cost of accessing appropriate labels for the task in a supervised fashion, unsupervised anomaly detection has become an alternative solution. However, previous methods mainly handcraft some features to detect the fraud on a single entity, which neglects the dynamic and topological nature between the buyer and the seller within the transaction. In this paper, we propose a novel Temporal Structure Augmented Gaussian Mixture Model (**TSAGMM**) for unsupervised fraud transaction detection on dynamic attributed networks. Specifically, we propose a time-encoded graph autoencoder to utilize both the topological structure and temporal information within the dynamic transaction graph to reconstruct the node attributes and graph topology. The learned latent representations as well as reconstruction errors are combined and fed into a density-based model for unsupervised fraud detection. Experimental results on the real-world transaction dataset from Alipay show the superiority of our proposed method among the state-of-the-art methods.

1 Introduction

In recent years, convenience facilitates the explosive growth of e-commerce and the booming of e-payment, while the underlying issue of the fraud transaction is not negligible. Indeed, the health development of online financial service is greatly threatened by various kinds of fraud transactions, ranging from cash-out fraud transaction [9] to malicious default fraud. In order to alleviate the negative impacts (i.e., incalculable risk-related damages and losses) on individuals and enterprises, *fraud transaction detection* has been an increasingly emerging topic in industrial applications, aiming at safeguarding the capital security in the face of fraudulent behaviors.

As the core component of ensuring a healthy environment of online financial services, recent years have witnessed a fruitful line of research in the field of the fraud transaction detection, and attain considerable success[17, 3, 1, 18]. Earlier works mainly focus on the exploration and exploitation of numerous rules summarized by fraud analysts. Unfortunately, the rapid change of fraud patterns

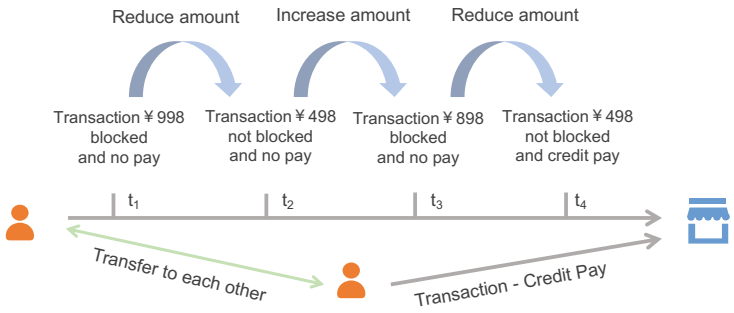

Fig. 1. A real showcase of fraud transactions for aiming at cash out. The two cash-out buyers have created multiple transaction interactions with particular suspicious merchant in a very close period. And the two buyers also have transfer relations before. The two buyers and the seller are in a group aiming to do cash-out.

is inherently difficult to be fitted with pure rules, hindering the effectiveness. Subsequently, attentions for the fraud transaction detection have been gradually shifted towards machine learning based methods, which could be roughly categorized as tree-based models, deep learning based models or graph-based models [1, 10, 18, 12]. Notwithstanding the competitive performance for automatically uncovering fraud patterns from the data, they still face the following two unresolved limitations:

- **Individual-level detection with static structure.** Previous works mainly perform the transaction fraud detection in the individual level (i.e a buyer or a seller), which commonly ignore the fact that transactions involve both buyers and sellers. On the other hand, fraud transactions intuitively reveals abnormal patterns in the temporal perspective (e.g., transactions with high frequencies), as shown in Fig. 1. Therefore, the temporal structure associated with a transaction could be a highly discriminative signal for suspicious behaviors
- **Supervised detection paradigm.** Most of current methods follows the supervised learning paradigm, whose success greatly hinges on large amounts of labeled data. However, in practical scenarios, fraud labels are usually difficult to obtain and insufficient training data also result in a serious data noise problem. Such an inevitable dilemma severely restricts the performance of detection methods with the supervised paradigm.

To address the challenges discussed above, we strive to frame the fraud transaction detection in the setting of unsupervised anomaly detection problem with dynamic attributed graphs. In particular, we propose a Temporal Structure Augmented Gaussian Mixture Model (**TSAGMM** for short) to comprehensively extract the temporal and structural nature of the dynamic transaction graph to detect the transaction-level frauds. In detail, we build TSAGMM upon the general encoder-decoder framework, where a graph neural network encoder with

a temporal component explicitly characterizes temporal and topological structure, while a graph reconstruction decoder further aims at the reconstruction of the both topological structure and node attributes. Subsequently, we fuse obtained representations of sellers and buyers, coupled with reconstruction errors into an unified transaction-level representation, and then feed it into a density estimation model for unsupervised fraud detection.

In summary, we highlight the main contributions of this paper as follows:

- **Problem**: As far as we know, we are the first to investigate unsupervised fraud transaction detection in transaction level under dynamic attributed graphs.
- **Model**: We propose TSAGMM, a novel unsupervised graph model to detect the complex fraud patterns based on the temporal and topological transaction behaviors..
- **Evaluation**: Experimental results on a real-world transaction dataset in Alipay prove the effectiveness of our proposed model.

2 Related Work

2.1 Fraud Transaction Detection

In the domain of fraud transaction detection, previous methods usually treat each transaction independently and train supervised models like support vector machines, random forests, etc. [1, 18]. Structural information in these methods if applicable, are usually incorporated as hand-crafted features, which are difficult to model subtle interaction information effectively. [12] proposes to use the transaction-intention network to capture the information over transactions and intentions with additional user behaviour sequence data. However, the method is in a supervised fashion to detect some specific patterns of fraud. Furthermore, it performs fraud detection on buyers or sellers, which overlook their coupling effects within the transactions.

2.2 Unsupervised Anomaly Detection

Anomaly detection is one of the common anti-fraud approaches in data science. Tremendous effort has been devoted to unsupervised anomaly detection [3] for tabular data, such as statistical techniques, density-based methods, clustering based methods and so on. Popular used techniques are local outlier factor[2], isolation forest [13], one-class support vector machine [4] etc. Recently, deep learning approaches [16] usually outperform traditional methods for multivariate and high dimensional data. These methods typically can be categorized as a family of encoder-decoder models. The representative is DAGMM [25]. However, all of these methods do not consider the complex interaction patterns for the transaction scenario between the buyers and sellers.

2.3 Graph based Anomaly Detection

Recent years have seen significant developments in graph neural networks (GNNs) and GNN-based methods are applied to the anomaly detection field [14]. Most of these methods focus on node fraud detection [5, 24, 22]. Only a few methods focus on edge fraud detection. For example, [15, 6, 22] focus on the edge fraud detection on static networks. [21, 23] are supervised anomaly edge detection on dynamic networks. In our setting, we treat transaction-level fraud detection as an anomalous edge detection problem without any supervision in the dynamic attributed graphs, which is rarely explored before.

3 Preliminary

We first define the dynamic attributed network in the following ways:

Definition 1. *A dynamic attributed network* $\mathbf{G} = (\mathbf{V}, \mathbf{E}, \mathbf{X}, \mathbf{H})$ *consists of: (1) the set of nodes* $\mathbf{V} = \{v_i\}_{i=1}^{N}$ *including the buyers and sellers; (2) the set of edges* $\mathbf{E} = \{e_{ij}\}$ *denoting the relation between node* i *and node* j *with the timestamp* $t = t_{ij}$*. Here the relation in our problem contains the transfer relation between buyers and trade relation between the buyer and the seller; (3) the node feature matrix* $\mathbf{X}$ *where the* i^{th} *row vector* $\mathbf{X}_{i\cdot}$ *denotes the attribute information for the* i^{th} *node; and (4) the edge feature matrix* $\mathbf{H}$*, where each element* $\mathbf{H}_{ij}$ *denotes the features of the edge* e_{ij}*.*

It is worth noting that there may be multiple edges between two nodes in dynamic attributed networks, indicating there are multiple transaction or transfer events occurring between nodes. Then the topological structure of dynamic attributed network $\mathbf{G}$ can be represented by an adjacency matrix $\mathbf{A}$, where $\mathbf{A}_{ij} = k$ if there are k edge events occurring between node v_i and node v_j. Otherwise $\mathbf{A}_{ij} = 0$ if there is no edge between node v_i and v_j.

Then our problem can be defined here:

Definition 2. *Unsupervised fraud transaction detection on dynamic attributed networks: Given the dynamic attributed network* $\mathbf{G}$*, a transaction involving a buyer, denoted as* $\mathcal{B}$*, and a seller denoted as* $\mathcal{S}$*, unsupervised fraud transaction detection aim to predict the fraud score* $s_{\mathcal{B},\mathcal{S}}(\mathbf{G})$ *only based on* $\mathbf{G}$*.*

4 The Proposed Model

In this section, we introduce our proposed method **TSAGMM** in detail. As shown in Fig 2, our model consists of three components: (1) The time-encoded graph encoder to model the temporal-structural information within the dynamic attributed graph. (2) The graph reconstruction part to restore the node attributes and graph structure for unsupervised graph learning and (3) The gaussian mixture model to do density-based fraud detection. Since the learning process of graph autoencoders for buyers and sellers are quite similar, we then mainly introduce buyers' as an illustration for space saving.

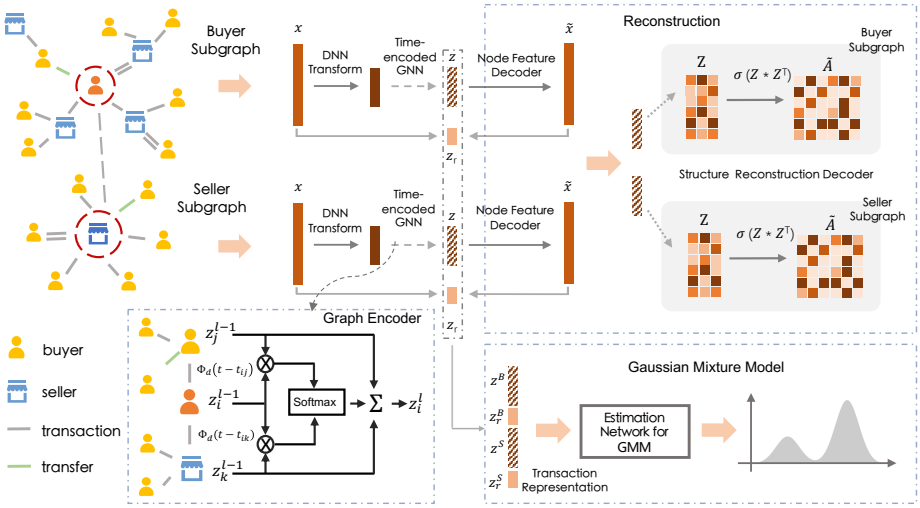

Fig. 2. The overall architecture of the Temporal Structure Augmented Gaussian Mixture Model (TSAGMM) for unsupervised fraud transaction detection.

4.1 Time-encoded Graph Encoder

We propose a time-encoded graph encoder with the attention mechanism to combine the temporal and structural information to learn node representations in a GNN manner. Supposing v_i is the target node at time t, it will aggregate the information from neighboring nodes via the following forms:

$$\mathbf{z}_i^l(t) = \sum_{j \in \mathcal{N}_i(t)} \gamma_{i,j}(t) \mathbf{z}_j^{l-1}(t), \tag{1}$$

where $\mathbf{z}_i^l(t)$ denotes the node embedding at layer l for node v_i, and $\mathbf{z}_i^0 = f(\mathbf{X}_{i\cdot})$, where $f(\cdot)$ denotes a DNN-based model to compress the original node features. $\mathcal{N}_i(t)$ denotes the neighbors of node v_i, whose interaction with v_i takes place at time prior to t. $\gamma_{i,j}(t)$ denotes the attention value to aggregate the representations from neighbor v_j to v_i, whose calculation process will be introduced later.

In particular, our proposed aggregation process not only considers the neighbors' features and edge features as common GNN models do, but more importantly considers the temporal information on each edge. To achieve the idea, we define a functional time encoding technique to represent the time as a combination of several periodic functions [20]:

$$\Phi_d(t) = \sqrt{\frac{1}{d}}[\cos(\omega_1 t), \sin(\omega_1 t), \ldots, \cos(\omega_d t), \sin(\omega_d t)] \tag{2}$$

parameterized by the frequency set $\{\omega_1, \ldots, \omega_d\}$.

Then the attention coefficient between target node i and neighbor node j is computed as follows:

$$\begin{aligned}\alpha_{i,j}(t) &= \text{attn}(\mathbf{q}_i(t), \mathbf{k}_j(t)) = \sigma(\mathbf{a}^T[\mathbf{W}_s\mathbf{q}_i(t) + \mathbf{W}_d\mathbf{k}_j(t)]),\\ \mathbf{q}_i(t) &= [\mathbf{z}_i^{l-1}(t) \parallel \mathbf{0} \parallel \Phi_d(0)],\\ \mathbf{k}_j(t) &= [\mathbf{z}_j^{l-1}(t) \parallel \mathbf{H}_{ij} \parallel \Phi_d(t - t_{ij})]\end{aligned} \tag{3}$$

where $\parallel$ denotes the concatenate operation, $\sigma(\cdot)$ is the tanh activation function, $\text{attn}(\cdot)$ denotes the attention function parameterized by $\mathbf{a}$, $\mathbf{W}_s$ and $\mathbf{W}_d$. Then, the attention value $\gamma_{i,j}$ can be achieved through the softmax function:

$$\gamma_{i,j}(t) = \frac{\exp(\alpha_{i,j}(t))}{\sum_{k\in\mathcal{N}_i(t)}\exp(\alpha_{i,k}(t))} \tag{4}$$

In our model, we use a two-layer time-encoded graph encoder to obtain the node embedding $\mathbf{Z} = \{\mathbf{z}_i^2(t_i)\}_{i=1}^N$, which aggregates neighbors' information within two hops.

4.2 Graph Reconstruction

Then we use the learned node embedding $\mathbf{Z}$ to reconstruct both the original node features and the adjacency matrix, aiming to make the model preserve the node attribute information and the graph topology.

In detail, we first use a DNN transform $g(\cdot)$ to project the node embedding $\mathbf{Z}$ into the feature reconstruction space $\tilde{\mathbf{X}} = g(\mathbf{Z}; \theta_g)$. Then given original node feature $\mathbf{X}$ and reconstruction node feature $\tilde{\mathbf{X}}$, we define the feature reconstruction error as: $\mathbf{\Delta}_f = [d_e, d_c]$ where d_e and d_c are the Euclidean distance and cosine similarity respectively between $\mathbf{X}$ and $\tilde{\mathbf{X}}$.

Another part aims to reconstruct the original network topology. The reconstruction adjacency matrix is calculated as the inner product between two node embeddings $\tilde{\mathbf{A}} = \sigma(\mathbf{Z}\mathbf{Z}^\mathbf{T})$, where $\sigma(\cdot)$ is the sigmoid function. Then the structure reconstruction error is given by $d_s = \|\bar{\mathbf{A}}_{\mathbf{i}\cdot} - \tilde{\mathbf{A}}_{i\cdot}\|_2$, where $\bar{\mathbf{A}}$ is the row-normalization of $\mathbf{A}$.

The overall loss function for graph reconstruction is given by:

$$\mathcal{L}_{recon} = \frac{1}{N}\sum_{i=1}^{N}\left(d_e\left(\mathbf{X}_{i\cdot}, \tilde{\mathbf{X}}_{i\cdot}\right) + d_c\left(\mathbf{X}_{i\cdot}, \tilde{\mathbf{X}}_{i\cdot}\right) + d_s\left(\bar{\mathbf{A}}_{i\cdot}, \tilde{\mathbf{A}}_{i\cdot}\right)\right) \tag{5}$$

4.3 Density Estimation

Fraud transactions often consist of abnormal information regarding its buyer and the seller. Therefore, to detect fraud transaction, we obtain the transaction representation by combining the obtained node embedding and reconstruction errors from both the buyer and the seller: $\mathbf{z}^\mathcal{T} = [\mathbf{z}^\mathcal{B}, \mathbf{\Delta}_f^\mathcal{B}, \mathbf{z}^\mathcal{S}, \mathbf{\Delta}_f^\mathcal{S}]$. It is worth noting that feature reconstruction errors for the buyer and seller can characterize

their own anomaly scores, because a good reconstruction-based model will focus on reconstructing the normal patterns, resulting in a larger reconstruction errors for anomaly data. Therefore, we also combine the feature reconstruction errors with the node embedding for the buyer and seller for fraud detection.

Without labels, we consider using the density-based methods for fraud detection. We assume that normal transactions can be modelled by a mixture of gaussian models, while the fraud transactions will be far-away from the combination of these gaussian distributions. Based on the assumption, we first predict its soft mixture-component membership prediction given the number of mixture components K: $\hat{\gamma} = \text{softmax}(h(\mathbf{z}^{\mathcal{T}}; \theta_h))$, where $h(\cdot)$ is a multi-layer neural network parameterized by θ_h. By traversing all the samples, we can estimate the mixture probability $\hat{\phi}_k$, the mean value $\hat{\mu}_k$, the covariance matrix $\hat{\Sigma}_k$ for each component k in GMM respectively [25]. The sample energy can be inferred by:

$$E(\mathbf{z}^{\mathcal{T}}) = -\log\Big(\sum_{k=1}^{K} \hat{\phi}_k \frac{\exp(-\frac{1}{2}(\mathbf{z}^{\mathcal{T}} - \hat{\mu}_k)^T \hat{\Sigma}_k^{-1}(\mathbf{z}^{\mathcal{T}} - \hat{\mu}_k))}{\sqrt{|2\pi\hat{\Sigma}_k|}}\Big) \tag{6}$$

4.4 Model learning and Fraud Detection

To jointly learn the reconstruction errors as well as GMM estimation, given a batch of N transaction data, the training objective function of our proposed model can be formulated as:

$$\mathcal{L} = (\mathcal{L}^{\mathcal{B}}_{recon} + \mathcal{L}^{\mathcal{S}}_{recon}) + \frac{\lambda_1}{N}\sum_{i=1}^{N} E(\mathbf{z}_i^{\mathcal{T}}) + \lambda_2 P(\hat{\Sigma}) \tag{7}$$

This objective function includes three components: (1) the first two terms denote the graph reconstruction errors for the buyer and seller. (2) $E(\mathbf{z}_i^{\mathcal{T}})$ is the energy defined in Eq. (6). It describes how possible we could see the transaction samples in the whole training dataset. (3) To avoid trivial solutions when the diagonal entries of covariance matrices degenerate to 0, we penalize small values of the diagonals by the fourth component $P(\hat{\Sigma}) = \sum_{k=1}^{K}\sum_{j}\frac{1}{\hat{\Sigma}_{kjj}}$ as a regularizer.

In the prediction phase, the sample energy is then employed to assess the abnormality of the transaction data.

5 Experiments

To demonstrate the effectiveness of the proposed method for fraud transaction detection, we conduct comprehensive experiments and present the result.

5.1 Experiments Setup

Evaluation Dataset With the real-world transaction datasets from Alipay, we sample about 30 million transaction data completed by credit pay in one month

for training, and the next month for evaluation. The dynamic graph which contains fund transfer relations between buyers and transaction relations (including credit pays and non-credit pays) between buyers and sellers, has around 70 million interactions per day. For each transaction, we sample two-hops dynamic attributed subgraphs centered by both the buyer and the seller respectively, within 7 days before the creation time of the transaction. We extract 43-dimension features for each user, including user profile, credit history, platform behaviors.

Evaluation Metrics Since there is no direct labels for the transactions, we collect the evaluation labels from business expert and from buyer and sellers' behaviors a few months later. In all, about 0.6% of the transactions are fraudulent for the evaluation. We select Precision, Recall, F1 score and the LIFT@k% as the evaluation metric. The LIFT@k% measure means the ratio between the bad rate of the top k% transactions with the average bad rate of all transactions. We use this metric because fraud detection usually focuses more on the results ranking ahead.

Comparing Methods To show the performance, we compare our proposed model **TSAGMM** with three lines of unsupervised methods: popular traditional methods on anomaly detection, NN-based methods and GNN-based methods. The first line contains **LOF** (Local Outlier Factor[2]), **OC-SVM** (One-class support vector machine [4]) and **iForest** (Isolation Forest [13]). NN-based methods contains reconstruction-based method **DAE** [19] and density-based method **DAGMM** [25]. Note that we choose these five methods as representatives because according to [8], these five methods perform good and stable. To make fair comparisons, we extract topology features like degree and the average of neighbors' features and combine them with the basic node features as the node features for these five methods. For graph-based baseline methods, since there is no existing unsupervised graph-based methods for edge-level fraud detection, we combine Graph Autoencoder [11, 7] and our temporal encoder to form **TGAE**.

Since LOF and OC-SVM don't scale well, we sample 1% from training dataset for these two methods. We set n_neighbors=10 for LOF and use RBF kernel in OC-SVM. For iForest, we use 100 trees to train. For all deep autoencoding instances, the embedding dimension is set as 8 with three hidden layers (32-unit, 16-unit and 8-unit, respectively). Moreover, we set the number of GMM components is 3 and set λ_1 as 0.1 and λ_2 as 0.001 as they render desirable results. In addition, two time-encoded graph attention layers are employed for our proposed models. We use Adam algorithm to optimize the loss function with the learning rate 0.001 and set batch size to 1024.

5.2 Offline Results

We present the results for each method in Table 1. We report the best F1 score with corresponding Precision and Recall, as well as the top 1% and 2% LIFT metric. The major results are summarized: It is worth noting that because of

Table 1. The performance in the evaluation dataset of credit transactions (best F1 score with corresponding Precision and Recall, as well as LIFT metrics). We also report the relative improvement ratio of our proposed model over all baselines.

Methods	Precision	Recall	F1 score	Lift on top1%	Lift on top2%
LOF	1.26%	6.17%	2.09% (+276%)	2.49 (+318%)	2.13 (+267%)
OC-SVM	2.59%	8.45%	3.96% (+99%)	4.72 (+121%)	4.23 (+85%)
iForest	3.65%	12.02%	5.60% (+40%)	6.80 (+53%)	6.02 (+30%)
DAE	1.02%	11.71%	1.87% (+321%)	1.68 (+520%)	1.70 (+360%)
DAGMM	4.12%	**13.58%**	6.32% (+24%)	7.59 (+37%)	6.75 (+16%)
TGAE	2.70%	8.91%	4.15% (+90%)	5.28 (+96%)	4.45 (+73%)
TSAGMM	**6.31%**	10.41%	**7.86%**	**10.41**	**7.82**

high class balance, the absolute values of precision, recall and F1 are not very high. But the F1 score of our proposed method still outperforms the comparing methods by at least 24%, which demonstrates the superiority of the overall performance of our method on fraud detection. Furthermore, we further find that our method achieves at least 35% and 15% improvement on Lift@1% and Lift@2% compared with all baseline methods. It is very important for real-world scenarios because fraud detection usually cares more about the entities ranking very ahead. In addition, we note that although TGAE utilizes the graph data, it performs not good. It demonstrates that designing a sophisticated component for unsupervised fraud detection is very important. In summary, the performance improvement demonstrates our model is able to capturing temporal and structural information effectively for fraud transaction detection.

5.3 Ablation Studies

We conduct the ablation studies to explore the impact of the main components of TSAGMM. Comparing methods are shown as follows:

Table 2. Ablation studies in the evaluation dataset of credit transactions.

Methods	Precision	Recall	F_1	Lift on top1%	Lift on top2%
TSAGMM	**6.312%**	**10.410%**	**7.859%**	**10.41**	**7.82**
TSAGMM_sd	5.733%	9.457%	7.139 %	9.46	6.88
TSAGMM_te	5.610 %	9.254%	6.985%	9.25	7.23
TGAE	2.701%	8.910 %	4.145%	5.28	4.45

- **TSAGMM_sd** removes the graph structure decoder reconstruction module.
- **TSAGMM_te** removes the time encoder part in the encoder function.

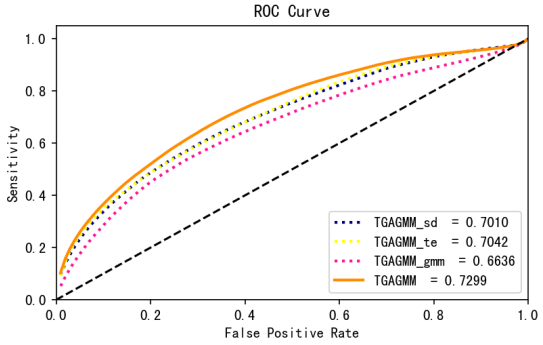

Fig. 3. ROC curves and AUC scores for ablation studies.

We report the results of ablation studies in Table 2 and Fig 3. Specifically, the performance of TSAGMM_sd and TSAGMM_te drops because the effects of the structural and temporal information are not fully exploited. For TGAE, the result shows the density estimation part is critical for fraud detection. Overall, we can clearly observe that full TSAGMM achieves the best result.

5.4 Performance in Online Environment

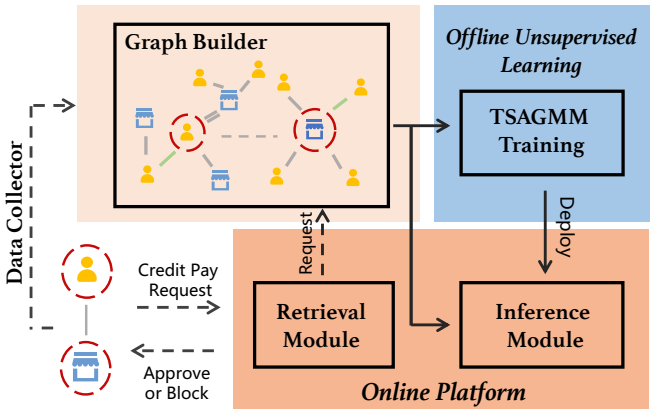

Fig. 4. The deployment of TSAGMM for the online credit transaction service in Alipay.

We deploy the proposed TSAGMM into the online environment of Alipay and report the online results. As shown in Fig 4, TSAGMM is trained offline every month and then deployed into online environment for serving. Then once a buyer issue a new credit pay request with a seller in the online serving, the

retrieval module are employed to extract node features and subgraphs associated with the request efficiently. Then TSAGMM predicts the anomaly score for the request. Together with other online strategies, the system would return the final decision, i.e. approve or block, to the transaction for the credit pay request.

Since we block high-risky transactions in the online environment, we use the repay rate, RP rate for short, as an important online metric. Once the system denies the credit pay request, risky buyers usually close the transactions, while normal buyers would continue to complete the transactions by other means of payments, e.g. debit cards. Therefore, the lower RP rate is, the more risky the transactions are. The average RP rate of online baseline strategy is 27.3%, while our proposed TSAGMM can detect additional highly risky transactions with RP rate 7.8%. Moreover, we randomly extract 100 samples from top ranked transactions and have a check by business experts. The results are as follows: 77% are abnormal (aiming to cash out or fraudulent), 10% are suspicious and 13% are misjudgements. Fig 1 shows a real abnormal transaction example detected by TSAGMM model for cashing out.

6 Conclusion

In this paper, we propose a novel time-encoded graph autoencoding gaussian mixture model for unsupervised fraud transaction detection on dynamic attributed networks. Specifically, we propose a time-encoded graph autoencoder to model the topological structure and temporal information within the dynamic transaction graph. The learned node representations and reconstruction errors are combined for density estimation to perform the fraud detection. Experimental results on the real-world transaction dataset from a credit payment service institute show the superiority of our proposed method among the state-of-the-art methods. The future work will focus on more types of nodes like devices to link users better.

References

1. Bhattacharyya, S., Jha, S., Tharakunnel, K., Westland, J.C.: Data mining for credit card fraud: A comparative study. Decision Support Systems (2011)
2. Breunig, M.M., Kriegel, H.P., Ng, R.T., Sander, J.: Lof: Identifying density-based local outliers. In: Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data (2000)
3. Chandola, V., Banerjee, A., Kumar, V.: Anomaly detection: A survey. ACM Computing Surveys (2009)
4. Chen, Y., Zhou, X.S., Huang, T.: One-class svm for learning in image retrieval. In: Proceedings 2001 International Conference on Image Processing (2001)
5. Ding, K., Zhou, Q., Tong, H., Liu, H.: Few-shot network anomaly detection via cross-network meta-learning. In: Proceedings of the Web Conference (2021)
6. Duan, D., Tong, L., Li, Y., Lu, J., Shi, L., Zhang, C.: Aane: Anomaly aware network embedding for anomalous link detection. In: 2020 IEEE International Conference on Data Mining (ICDM). pp. 1002–1007. IEEE (2020)

7. Fan, H., Zhang, F., Li, Z.: Anomalydae: Dual autoencoder for anomaly detection on attributed networks. In: ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 5685–5689. IEEE (2020)
8. Han, S., Hu, X., Huang, H., Jiang, M., Zhao, Y.: Adbench: Anomaly detection benchmark. arXiv preprint arXiv:2206.09426 (2022)
9. Hu, B., Zhang, Z., Shi, C., Zhou, J., Li, X., Qi, Y.: Cash-out user detection based on attributed heterogeneous information network with a hierarchical attention mechanism. In: Proceedings of the 33rd AAAI Conference on Artificial Intelligence (2019)
10. Huang, D., Mu, D., Yang, L., Cai, X.: Codetect: Financial fraud detection with anomaly feature detection. IEEE Access (2018)
11. Kipf, T.N., Welling, M.: Variational graph auto-encoders. arXiv preprint arXiv:1611.07308 (2016)
12. Liu, C., Sun, L., Ao, X., Feng, J., He, Q., Yang, H.: Intention-aware heterogeneous graph attention networks for fraud transactions detection. In: Proceedings of the 27th ACM SIGKDD Conference (2021)
13. Liu, F.T., Ting, K.M., Zhou, Z.H.: Isolation forest. In: 2008 Eighth IEEE International Conference on Data Mining (2008)
14. Ma, X., Wu, J., Xue, S., Yang, J., Zhou, C., Sheng, Q.Z., Xiong, H., Akoglu, L.: A comprehensive survey on graph anomaly detection with deep learning. IEEE Transactions on Knowledge and Data Engineering (2021)
15. Ouyang, L., Zhang, Y., Wang, Y.: Unified graph embedding-based anomalous edge detection. In: 2020 International Joint Conference on Neural Networks (2020)
16. Pang, G., Shen, C., Cao, L., Hengel, A.V.D.: Deep learning for anomaly detection: A review. ACM Computing Surveys (2021)
17. Phua, C., Lee, V., Smith, K., Gayler, R.: A comprehensive survey of data mining-based fraud detection research. arXiv preprint arXiv:1009.6119 (2010)
18. Prusti, D., Rath, S.K.: Fraudulent transaction detection in credit card by applying ensemble machine learning techniques. In: 2019 10th International Conference on Computing, Communication and Networking Technologies (2019)
19. Sakurada, M., Yairi, T.: Anomaly detection using autoencoders with nonlinear dimensionality reduction. In: Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis (2014)
20. Xu, D., Ruan, C., Korpeoglu, E., Kumar, S., Achan, K.: Inductive representation learning on temporal graphs. In: International Conference on Learning Representations (ICLR) (2020)
21. Zhang, G., Li, Z., Huang, J., Wu, J., Zhou, C., Yang, J., Gao, J.: efraudcom: An e-commerce fraud detection system via competitive graph neural networks. ACM Transactions on Information Systems (TOIS) (2022)
22. Zhang, G., Wu, J., Yang, J., Beheshti, A., Xue, S., Zhou, C., Sheng, Q.Z.: Fraudre: fraud detection dual-resistant to graph inconsistency and imbalance. In: 2021 IEEE International Conference on Data Mining (ICDM) (2021)
23. Zheng, L., Li, Z., Li, J., Li, Z., Gao, J.: Addgraph: Anomaly detection in dynamic graph using attention-based temporal gcn. In: Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (2019)
24. Zheng, P., Yuan, S., Wu, X., Li, J., Lu, A.: One-class adversarial nets for fraud detection. In: Proceedings of the AAAI Conference on Artificial Intelligence (2019)
25. Zong, B., Song, Q., Min, M.R., Cheng, W., Lumezanu, C., Cho, D., Chen, H.: Deep autoencoding gaussian mixture model for unsupervised anomaly detection. In: International conference on learning representations (2018)