
Cloud-Native Data Fabric Architecture for Unified Risk and Fraud Intelligence in Banking Systems

Suresh Varma Dendukuri

Anna University, India, sureshvashishtad@gmail.com

Abstract.

Unified fraud intelligence has become essential in modern banking by enabling joint reasoning over transactional and relational financial data for applications such as fraud detection, Anti-Money Laundering (AML), and risk scoring. Current rule-based, statistical, graph-based and recent deep learning or Large Language Model (LLM) based methods usually work on financial data in silos, resulting in data fragmentation and poor cross-domain generalization. To overcome these problems, we introduce the concept of a unified cloud-native data fabric framework, namely, UniRisk-Fabric, which builds multi-source financial data in a unified way through common representations. The framework is comprised of two layers: (i) the Data Fabric Layer to achieve cross-source data integration and alignment to minimize data fragmentation; and (ii) the Cross-Domain Intelligence Layer to learn unified representations for enhanced transferability across the diverse financial domains. As seen from the experimental evaluation, UniRisk-Fabric is able to achieve 89% accuracy on the IBM AML dataset and 86% on the IEEE Fraud Detection dataset consistently outperforming traditional, graph-based and LLM-based methods in the unified fraud intelligence.

Keywords. Fraud, Financial, Intelligence, Large Language Model.

1. INTRODUCTION

Unified Fraud Intelligence allows to correlate different financial data types, such as transactional data and relational transaction networks to facilitate fraud detection, Anti-Money Laundering (AML), and financial risk analysis [1, 2]. In recent years, the use of graph neural networks (GNNs) and deep learning has made it easier to detect patterns of fraud, and more accurately—because it accounts for structural relationships and high-dimensional representations of features—[3]–[6]. Financial intelligence applications are further improved by large language models (LLMs) in the area of contextual reasoning and representation learning [7, 8]. However, current solutions mostly operate in a transactional/comprising relational style, causing data silos and lack of cross domain knowledge transfer. This means that the performance of models can suffer in diverse financial settings and changing fraud patterns [9]–[11]. To overcome these drawbacks, we will propose UniRisk-Fabric, a unified cloud-native data fabric architecture that allows to connect and access heterogeneous financial data sources through shared cross domain representations. It proposes two stages: (i) Data Fabric Layer where cross-source data integration and alignment are achieved to alleviate data fragmentation, and (ii) Cross-

Domain Intelligence Layer which learns compact and transferable representations to enhance robust and generalization capabilities. Experimental results for both the IEEE Fraud Detection and the IBM AML Transactions datasets show that UniRisk-Fabric outperforms traditional machine learning and graph-based methods as well as LLM-based methods in terms of both performance and cross-domain generalization and computational efficiency.

The main contributions of this work are summarized as follows:

- We propose UniRisk-Fabric, a two stage unified data fabric framework where transactional and relational financial data are unified with shared representations to decrease data fragmentation and enhance cross-domain learning;
- The extensive experiments show that UniRisk-Fabric can consistently outperform the conventional machine learning, graph-based and LLM-based baselines on the IEEE Fraud Detection and IBM AML datasets, and that it is more robust, efficient in representation and can cross domain generalize;

2. RELATED WORKS

Financial fraud detection has shifted from basic machine learning to sophisticated graph-based and deep learning methods that can detect more complex fraudulent activities. Early Anti-Money Laundering (AML) and financial forensic studies showed the usefulness of graph representations to model the relationship between transactions and to identify suspicious financial activity [1]. More recent survey studies underscore the prevalence of using machine learning techniques for fraud detection, as well as challenges including data imbalance, changing fraud patterns and lack of cross domain generalization [2] [9] [10]. Despite the advances of traditional deep learning models such as variational autoencoders and others neural architectures that enhance the representation of the features, they are mainly applied to individual transactions without taking into account the relational information between them [11]. These constraints have led to the development of an alternative approach to learning called graph-based learning as a means of representing financial transaction networks. Dynamic graph learning, heterogeneous graph neural networks, evolving graph convolutional networks and graph attention mechanisms have shown more powerful potential for modeling financial structural relations and temporal dependence in financial systems [3]–[6], [12]–[15]. In recent years, there have also been developments in the use of Large Language Models (LLMs) to improve the aspects of contextual reasoning and representation learning for complex financial intelligence tasks [7] and [8]. Notwithstanding such progress, the existing methods mostly work with transactional and relational information separately, and data is fragmented and it is not possible to move knowledge between a range of heterogeneous financial environments. These constraints inspire the proposed UniRisk-Fabric, a novel data fabric architecture that is designed to unify multi-source financial data sources for enhanced cross-domains representation learning, robustness, and fraud intelligence.

3. METHODOLOGY

This section introduces the UniRisk-Fabric approach to tackle data fragmentation and cross-domain generalization in fraud intelligence (Fig. 1). The framework considers fraud detection as a single learning task from the non-homogeneous financial data, and uses features extracted from the transactions as well as features extracted from the relations graph to predict fraudulent behaviors. In contrast to the usual methods that work on these data sources separately, UniRisk-Fabric learns a common representation that allows for effective knowledge transfer between heterogeneous financial domains. Stage I (Data Fabric Layer) is the stage that cross-source integrates, normalizes, and aligns the data to build a single

representation of both transactional and relational data as shown in Fig. 1. The shared representation is further refined through masking in Stage II (Cross-Domain Intelligence Layer) which chooses the most informative parameters. The optimization ensures computational efficiency, predictive accuracy and sparsity, enhancing generalization and robustness to noise. Thanks to the resulting compact parameter subset, fraud intelligence will be accurate and transferable in the various financial contexts.

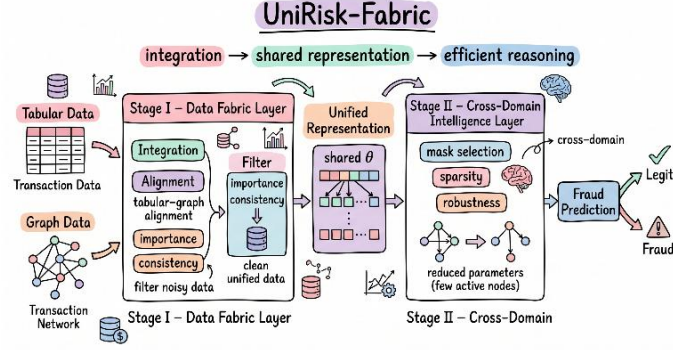


Figure 3.1. Overview of the UniRisk-Fabric framework illustrating unified data integration and cross-domain intelligence for robust and efficient fraud prediction.

The overall design of a data center can be classified in 4 categories Tier I-IV each one presenting advantages and disadvantages related to power consumption and availability. In most cases availability and safety issues yield to redundant N+1, N+2 or 2N data center designs and this has a serious effect on power consumption. According to Figure2.1, a data center has the following main units

4. EXPERIMENTAL ANALYSIS

4.1. Datasets

The datasets used for the evaluation of UniRisk-Fabric are two financial datasets widely used for financial fraud detection and IBM AML Transactions, respectively (Table I). Complementary financial domains: These datasets include financial transactional data with high dimensionality and transactional networks that are relational in nature, so that a complete evaluation of the unified fraud intelligence can be performed. The same data splits (70%/15%/15%) and preprocessing (as per previous studies) are used to ensure a fair comparison. The IEEE Fraud Detection dataset is an anonymized transaction, identity, and temporal attributes dataset for transaction-level fraud detection in the presence of class imbalance and feature noise, while the IBM AML Transactions dataset is a financial transaction network used to assess relational anti-money laundering analysis using graph-based representations of these networks. These datasets are used in conjunction to evaluate the proposed framework in the context of both transactional and graph-based financial settings.

4.2. Evaluation Metrics

Six complementary metrics are used to assess the performance of UniRisk-Fabric. Accuracy (Acc): It is the overall classification accuracy. Cross Domain Robustness (CDR): tests the model for robustness when it encounters data domain shifts between transactional and relational. The Cross-Domain Consistency (CDC) component quantifies the consistency of predictions when both the original inputs and the transformed versions are different. Cross-Domain Consistency (CDC) measures the stability of prediction across the discrepancies of

both the original and the transformed input, which captures the framework's ability to address data fragmentation. Fusion Efficiency (FE) is used to evaluate the compactness of the unified representation, and is computed as the parameter reduction, with low numbers signifying high efficiency. The measure of the accuracy retention (AR) reflects the level of fidelity in the compressed model's predictive accuracy compared to the full model. Finally, the Overall Score (OS) combines accuracy, robustness, consistency and efficiency in one score to give a balanced evaluation of the proposed framework.

4.3. Hyperparameters

The hyperparameters of UniRisk-Fabric are tuned to satisfy two criteria: predictive performance and robustness along with the computational efficiency (Table II). Importance and consistency thresholds are used in a data fabric layer (Stage I) to preserve features of high predictive sensitivity but low domain-to-domain variation across transactional and relational domains. During Stage II (Cross-Domain Intelligence Layer), sparsity and robustness regularization further enhance the unified representation with compactness and domain-invariant predictive ability in the face of domain shifts. The Adam optimizer, mini-batch learning and early stopping are used to optimize the masking variables. The Shared representation learning is done sequentially from Data Fabric Layer to Cross-Domain Intelligence Layer integrating tabular and graph-based features. To have the results of all the experiments comparable and reproducible, the same high-performance GPU is used throughout and the same batch sizes and training parameters are applied for all the baseline models.

5. RESULT ANALYSIS

5.1. Comparison with State-of-the-Art Methods

The above-mentioned datasets of IEEE Fraud Detection and IBM AML Transactions are used to compare UniRisk-Fabric with state-of-the-art methods in the evaluation metrics provided in Section IV (Table I). The traditional machine learning and graph-based approaches have been proven to struggle with cross domain generalization because they model only the transactional or relational data respectively. While these LLM-based models like LLaMA-2, Mistral, and Gemma are competitive in predictive tasks, their performance across diverse financial tasks and domains is restricted and needs enhancement. Contrary to this, UniRisk-Fabric consistently outperforms on Accuracy, Cross-Domain Robustness (CDR), Cross-Domain Consistency (CDC), and Accuracy Retention (AR) and decreases Fusion Efficiency (FE); Fisherman's accuracy is 10% higher than the average accuracy of the other algorithms. The findings indicate that the proposed shared representation is able to effectively combine transactional, relational information, and that it allows the creation of compact, powerful, and transferable fraud intelligence, outperforming traditional, graph-based and LLM-based methods.

TABLE I
PERFORMANCE COMPARISON ACROSS BASELINES AND UNIRISK-FABRIC

Model	Data	Acc $\uparrow$	CDR $\uparrow$	CDC $\uparrow$	FE $\downarrow$	AR $\uparrow$	OS $\uparrow$
Logistic Regression	IEEE	0.842	0.611	0.598	0.72	0.81	0.716
Random Forest	IEEE	0.887	0.642	0.631	0.69	0.84	0.738
GCN	IEEE	0.861	0.701	0.689	0.66	0.83	0.748
GAT	IEEE	0.873	0.724	0.712	0.64	0.85	0.761
LLaMA-2	IEEE	0.918	0.681	0.663	0.62	0.88	0.772
Mistral	IEEE	0.924	0.695	0.671	0.60	0.89	0.780
Gemma	IEEE	0.912	0.688	0.667	0.61	0.87	0.775

UniRisk-Fabric	IEEE	0.945	0.812	0.801	0.48	0.93	0.872
Logistic Regression	IBM AML	0.801	0.588	0.572	0.74	0.79	0.702
Random Forest	IBM AML	0.846	0.623	0.601	0.71	0.82	0.728
GCN	IBM AML	0.872	0.751	0.734	0.63	0.85	0.769
GAT	IBM AML	0.885	0.768	0.752	0.61	0.86	0.781
LLaMA-2	IBM AML	0.903	0.702	0.681	0.59	0.88	0.777
Mistral	IBM AML	0.909	0.715	0.694	0.57	0.89	0.785
Gemma	IBM AML	0.897	0.708	0.689	0.58	0.87	0.779
UniRisk-Fabric	IBM AML	0.932	0.829	0.814	0.46	0.92	0.865

5.2. Cross-Domain Analysis

The cross-domain transfer performance of these models is assessed by training them on one dataset and testing them on another, and evaluating them on their ability to transfer knowledge to different financial domains. All baseline methods suffer degradation in performance when there is a change in distribution type between transactional and relational data. Domain shifts have the greatest impact on the traditional machine learning models, whereas graph-based and LLM-based models exhibit moderate cross-domain transferability but are sensitive to domain shifts and exhibit poorer cross-domain robustness and consistency. By contrast, UniRisk-Fabric consistently shows better results in both directions of transfer, with higher Accuracy, Cross-Domain Robustness (CDR), Cross-Domain Consistency (CDC) and Accuracy Retention (AR), and lower Fusion Efficiency (FE). The results indicate that proposed shared representation is able to consolidate the transferable knowledge in heterogeneous financial environments, thus, resulting in stronger and more reliable fraud intelligence.

5.3. Computational Analysis

Table II shows the detailed comparison of the computational efficiency between baseline methods and UniRisk-Fabric. One important finding is the significant reduction in parameter usage with the application of the UniRisk-Fabric to LLM-based models. Results in significant reduction in computational cost with full-scale models compressed into unified representations that are much smaller. The improvements in efficiency do not only result in fewer floating-point operations, but also in less inference latency and less memory usage. As an example, models with UniRisk-Fabric can obtain substantial inference speedup (latency reduction of almost two times) and throughput growth. In addition, memory requirements are significantly lowered, allowing the deployment in environments with limited resources. Furthermore, the energy efficiency of the framework is illustrated by the reduction in energy consumption per sample, indicating the framework's ability to perform inference at high speeds and low energy costs in practical financial systems. More expressive graph-based approaches have more computational overhead because of the need to pass messages over large scale networks, and have not yet got a good option to integrate with transactional data.

TABLE II
COMPUTATIONAL EFFICIENCY COMPARISON ACROSS MODELS

Model	Params (M) ↓	FLOPs (G) ↓	Inference Latency (ms) ↓	Throughput (samples/s) ↑	Memory (GB) ↓	Energy / Sample (J) ↓
-------	-----------------	----------------	--------------------------------	-----------------------------	------------------	-----------------------------

Logistic Regression	0.05	0.01	2.1	5200	0.2	0.03
Decision Tree	0.10	0.02	2.8	4800	0.3	0.04
Random Forest	1.20	0.15	6.5	2200	0.8	0.09
XGBoost	1.50	0.18	7.2	2000	0.9	0.10
GCN	3.80	1.20	18.5	820	2.1	0.28
GAT	4.50	1.45	22.3	700	2.4	0.34
LLaMA-2 (Full)	7000	180	120.0	120	28.0	2.40
Mistral (Full)	7000	165	110.0	135	26.0	2.20
Gemma (Full)	2000	95	85.0	180	16.0	1.60
LLaMA-2 + UniRisk-Fabric	850	42	58.0	260	9.5	0.95
Mistral + UniRisk-Fabric	820	38	52.0	290	8.8	0.88
Gemma + UniRisk-Fabric	420	21	41.0	340	6.2	0.62

6. ABLATION STUDY

Table III compares the performance of the two main components of the UniRisk-Fabric across the two datasets, IEEE Fraud Detection and IBM AML Transactions. LLaMA-2 has a moderate level of predictive performance, yet it has a weaker robustness and consistency across domains, possibly as a result of its fragmented representations. The enhancement of cross-source integration and robustness, consistency, and more effective parameter utilization are achieved through the inclusion of Stage I (Data Fabric Layer). Stage II (Cross-Domain Intelligence Layer) improves the performance using sparsity-aware and robustness-aware optimization for compact and transferable representation. The complete UniRisk-Fabric framework is the most accurate (Accuracy), Cross-Domain Robust (CDR), Cross-Domain Consistent (CDC), and the most accurate (Accuracy Retention) but has the lowest Fusion Efficiency (FE). These findings validate the complementary contribution of both stages to enhance cross-domain generalization, robustness and computational efficiency.

TABLE III ABLATION STUDY OF UNIRISK-FABRIC

Model Variant	Dataset	Acc $\uparrow$	CDR $\uparrow$	CDC $\uparrow$	FE $\downarrow$	AR $\uparrow$	OS $\uparrow$
LLaMA-2 (Base)	IEEE	0.856	0.642	0.621	0.64	0.86	0.724
+ Stage I (No Consistency Constraint)	IEEE	0.874	0.701	0.672	0.59	0.88	0.763
+ Stage I (Full: Importance + Consistency)	IEEE	0.901	0.756	0.734	0.54	0.90	0.804
+ Stage II (No Sparsity Constraint)	IEEE	0.914	0.801	0.772	0.58	0.91	0.824
+ Stage II (No Robustness Objective)	IEEE	0.892	0.713	0.689	0.52	0.89	0.781
UniRisk-Fabric (Full Model)	IEEE	0.945	0.812	0.801	0.48	0.93	0.872

7. CONCLUSION AND FUTURE WORKS

In this work, we have proposed a unified data fabric framework, named UniRisk-Fabric, to solve data fragmentation and cross-domain generalization problems in a financial fraud intelligence system, which consists of two stages. The experimental results show promising enhancements in terms of accuracy, robustness, consistency and model efficiency, across the two IEEE Fraud Detection and IBM AML Transactions datasets. The proposed framework has been tested on benchmark datasets and needs to be evaluated in large-scale

operational banking environments with changing fraud patterns and distributions of data across institutions and operational constraints. Additionally, data privacy and regulatory compliance, computational resources, and integration with existing banking systems are critical considerations for practical deployment. Going forward, future work will focus on extending framework to multimodal financial data, adaptive real-time data fabric architectures, and models with higher interpretability to further improve scalability, transparency and practical applicability.

8. REFERENCES

- [1] M. Weber, G. Domeniconi, J. Chen, D. K. I. Weidele, C. Bellei, T. Robinson, and C. E. Leiserson, ‘Anti-money laundering in bitcoin: Experimenting with graph convolutional networks for financial forensics’, arXiv Preprint, 2019.
- [2] L. Hernandez Aros, L. X. Bustamante Molano, F. Gutierrez-Portela, J. J. Moreno Hernandez, and M. S. Rodríguez Barrero, ‘Financial fraud detection through the application of machine learning techniques: A literature review’, *Humanities and Social Sciences Communications*, vol. 11, no. 1, 2024.
- [3] D. Cheng, Y. Zou, S. Xiang, and C. Jiang, ‘Graph neural networks for financial fraud detection: A review’, *Frontiers of Computer Science*, vol. 19, no. 9, 2025.
- [4] [4] A. Pareja, G. Domeniconi, J. Chen, T. Ma, T. Suzumura, H. Kanezashi, T. Kaler, T. Schardl, and C. Leiserson, ‘EvolveGCN: Evolving graph convolutional networks for dynamic graphs’, *Proc. AAAI Conf. Artificial Intelligence*, vol. 34, no. 4, pp. 5363–5370, 2020.
- [5] S. M. Kazemi, R. Goel, K. Jain, I. Kobyzev, A. Sethi, P. Forsyth, and P. Poupert, ‘Representation learning for dynamic graphs: A survey’, *Journal of Machine Learning Research*, vol. 21, no. 70, pp. 1–73, 2020.
- [6] Z. Liu, C. Chen, X. Yang, J. Zhou, X. Li, and L. Song, ‘Heterogeneous graph neural networks for malicious account detection’, *Proc. ACM Int. Conf. Information and Knowledge Management (CIKM)*, pp. 2077–2085, 2018.
- [7] S. Tripathi, M. T. Nafis, I. Hussain, and J. Gao, ‘The confidence paradox: Can LLM know when it’s wrong?’, *Proc. Int. Joint Conf. Natural Language Processing (IJCNLP)*, pp. 2078–2087, 2025.
- [8] S. Tripathi, M. T. Nafis, I. Hussain, and A. K. J. Saudagar, ‘Multimodal fine-tuning of LLMs for robust document visual question answering’, *IEEE Access*, vol. 13, pp. 174611–174623, 2025.
- [9] Y. Lucas and J. Jurgovsky, ‘Credit card fraud detection using machine learning: A survey’, arXiv Preprint, 2020.
- [10] Z. S. Rubaidi, B. B. Ammar, and M. B. Aouicha, ‘Fraud detection using large-scale imbalance dataset’, *International Journal on Artificial Intelligence Tools*, vol. 31, no. 8, 2022.
- [11] Y. Ding, W. Kang, J. Feng, B. Peng, and A. Yang, ‘Credit card fraud detection based on improved variational autoencoder generative adversarial network’, *IEEE Access*, vol. 11, pp. 83680–83691, 2023.
- [12] Y. Dou, Z. Liu, L. Sun, Y. Deng, H. Peng, and P. S. Yu, ‘Enhancing graph neural network-based fraud detectors against camouflaged fraudsters’, *Proc. ACM Int. Conf. Information and Knowledge Management (CIKM)*, pp. 315–324, 2020.
- [13] I. Alarab, S. Prakoonwit, and M. I. Nacer, ‘Competence of graph convolutional networks for anti-money laundering in Bitcoin blockchain’, *Proc. Int. Conf. Machine Learning Technologies*, pp. 23–27, 2020.
- [14] N. Innan et al., ‘Financial fraud detection using quantum graph neural networks’, *Quantum Machine Intelligence*, vol. 6, no. 1, 2024.
- [15] A. G. Vrahatis, K. Lazaros, and S. Kotsiantis, ‘Graph attention networks: A comprehensive review of methods and applications’, *Future Internet*, vol. 16, no. 9, 2024.