
Graph-Based Entity Resolution for Real-Time Financial Sanctions Screening: Leveraging Knowledge Graphs and GNN-Enhanced Fuzzy Matching at Enterprise Scale

Suman Reddy Gaddam

Indian Institute of Science & Technology (IIST), India, sumanreddyg05@gmail.com

Abstract.

The problem of financial sanctions screening is to identify entities in financial records, which is typically a huge collection of financial records, from a wide variety of types with inconsistent identity names and confusing textual attributes. The existing rule based and string-matching methods are inadequate to recognize structural relationships between entities, which leads to high false positive rate and low efficiency in operation. The objective of this paper is to propose a framework for real-time financial sanctions screening, called Hybrid Graph-based Fuzzy Entity Resolution (HG-FER). The framework is based on knowledge graph modeling, graph neural network (GNN)-based contextual representation learning, and fuzzy similarity using adaptive fuzzy similarity in a single entity resolution pipeline. Complex relationships are captured in knowledge graphs, their structure is learned by GNNs, and textual uncertainty is manageable with fuzzy matching. Hybrid similarity mechanism: Both relational and attribute level information are used for better entity matching accuracy. Experimental results show that the proposed framework obtains a precision of 0.92, recall of 0.89 and F1-score of 0.90, and has 10 false positives which is better than conventional approaches. The findings show that HG-FER is a scalable, reliable, and effective solution for the enterprise level financial compliance and near real-time sanctions screening.

Keywords. Entity Resolution, Knowledge Graph, Graph Neural Networks, Fuzzy Matching, Financial Sanctions Screening, Fraud Detection.

1. INTRODUCTION

The nature of financial sanctions and the requirement to comply with them has become increasingly complex as financial markets and cross-border transactions have developed and become more digitized. Large data sets of customer and transaction records need to be continuously processed and compared against dynamic sanctions lists, even when the data has different naming conventions, attributes and when some data is missing or noisy. Presently, the conventional sanctions screening systems are fundamentally based on rule-based filtering and string-matching methods that tend to miss out on semantic relationships between entities and thus have high false positive and false negative rates. Such restrictions can drive up operating expenses, compliance risks, and make compliance systems less reliable.

Knowledge graphs and graph neural networks (GNNs) are introduced to represent complex relationships between heterogeneous financial entities and learn the context information

All Open Access articles are peer-reviewed, distributed under the Creative Commons Attribution-Non Commercial 4.0 International

from graph structures, which have significantly advanced the representation of entities. But most current approaches are based on one aspect of relational learning—graph-based relational learning—or one aspect of fuzzy attribute matching and don't allow for the accurate disambiguation of ambiguous entities in real-time enterprise settings.

To overcome these shortcomings, the authors introduce in this paper a new framework called Hybrid Graph-based Fuzzy Entity Resolution (HG-FER) for financial sanctions screening in real time. In contrast to traditional approaches that mainly depend on graph-based relational learning or fuzzy attribute matching, HG-FER introduces an adaptive fuzzy similarity and a unified entity resolution pipeline with knowledge graph modeling and GNN-based contextual representation learning. This integrated design not only incorporates structural relationships but also the uncertainty at the attribute level, leading to better matching accuracy, fewer false matches, and enterprise compliance scaling.

This work has the following key contributions:

- The architecture of a knowledge graph-based system to model heterogeneous financial entities and their relations;
- The incorporation of the use of graph neural networks to learn and represent complex relational dependencies and improve the prediction of entity similarity;
- The integration of the fuzzy matching methods to efficiently deal with ambiguous and inconsistent identity attributes and;
- The creation of a scalable architecture to be applied in real-time sanctions screening in enterprise environments;

As far as we know, this work is one of the first to combine fuzzy similarity with graph neural networks in a single framework in terms of real-time financial sanctions screening.

2. RELATED WORK

As financial sanctions screening becomes more complex, intelligent, data-driven approaches to entity resolution and fraud detection have come into use. Knowledge graphs have become a useful method of modelling a combination of financial entities and their semantic relationships, which has enhanced the enterprise risk analysis and decision support [1, 2]. In addition, their incorporation with graph neural networks (GNNs) further facilitates contextual representation learning by modelling structural dependencies between interconnected entities, which boosts the performance of financial analytics and knowledge reasoning [5], [7]–[10]. In the last few years, there have been several successful attempts at using graph-based learning for financial fraud detection, anti-money laundering and risk prediction. GCNs and GATs have been used to successfully uncover inner transaction patterns, detect suspicious financial activities, and enhance the predictive efficiency of large-scale financial systems [5], [7], [10], [14]–[17], [20]. These methods take advantage of relationships in the data that are not captured by traditional machine learning methods.

However, one of the primary challenges in financial compliance is entity resolution, given that many real-world records will have inconsistent, incomplete or ambiguous identity attributes. To overcome these shortcomings, recent work focused on the robustness of matching under heterogeneous data conditions has introduced fuzzy matching [11]–[13] and graph representation learning and hybrid entity resolution [19] techniques. Meanwhile, several research works have focused on cyber threat intelligence and reinforcement learning, revealing the advantages of combining multiple learning paradigms in complex risk detection and adaptive decision-making for cyber threats [3], [4], [6] and the increasing relevance of hybrid intelligent frameworks. Current methods typically partition graph-based relational learning and fuzzy similarity matching, however, and fail to work well for real-

time enterprise-level sanctions screening. Furthermore, there are numerous approaches that do not scale well and do not support the processing of heterogeneous financial information in a computationally efficient and regulatory compliant manner. To overcome these drawbacks, this work presents the Hybrid Graph-based Fuzzy Entity Resolution (HG-FER) framework that combines knowledge graph modeling, GNN-based contextual representation learning, and adaptive fuzzy similarity in a single pipeline for effective and efficient real-time financial sanctions screening.

3. PROPOSED SYSTEM ARCHITECTURE

The suggested system presents a single-system to screen financial sanctions in real-time, combining the knowledge graph construction, graph neural network (GNN)-based relational learning, and fuzzy-enhanced entity similarity matching. The aim is to solve entities in heterogeneous sources of financial data correctly and with scalability and robustness in the large-scale enterprise setting. The system starts with the consumption of multi-source financial data, such as customer data, transaction data, sanctions data, and supporting identity data. Preprocessed and normalized, any entity then is the node in the knowledge graph, denoted $G = (V, E)$, where V is the set of entities and E is the set of relationships between them. Every node $v_i \in V$ is corresponded to a feature vector $x_i \in \mathbb{R}^d$, which represented features like name encodings, address encodings, and transactional features.

A graph neural network is used on the constructed graph to learn relational dependencies. The embedding of each node is iteratively updated using neighborhood aggregation:

$$h_i^{(k+1)} = \sigma \left(\sum_{j \in \mathcal{N}(i)} \frac{1}{|\mathcal{N}(i)|} W^{(k)} h_j^{(k)} \right) \quad (1)$$

$\mathcal{N}(i)$ represents the neighboring nodes of v_i and $W^{(k)}$ is the learnable parameters in layer k . The last embedding, z_i encodes both attribute-level and structural information, which allows the use of the context to represent each entity.

Simultaneously, the inconsistencies in the textual attributes are addressed by means of a fuzzy similarity mechanism. Similarity of two entities v_i and v_j computed as:

$$S_f(v_i, v_j) = \frac{1}{m} \sum_{k=1}^m \text{sim}(a_i^{(k)}, a_j^{(k)}) \quad (2)$$

In which m denotes the number of attributes and $\text{sim}(\cdot)$ is an approximate string similarity function. This allows the system to handle variations effectively like abbreviations, misspellings and phonetic variations.

A graph-based similarity score based on learned embeddings is calculated to include structural context:

$$S_g(v_i, v_j) = \frac{z_i \cdot z_j}{\|z_i\| \|z_j\|} \quad (3)$$

A graph-based similarity score is calculated with the learned embeddings to include structural context:

$$S(v_i, v_j) = \alpha S_f(v_i, v_j) + (1 - \alpha) S_g(v_i, v_j) \quad (4)$$

$\alpha \in [0,1]$ is used to tune the contribution of fuzzy matching and graph-based learning. This mixed-formulation guarantees that the textual ambiguity and relational dependencies are both taken into account to resolve entities.

The decision mechanism is used to decide whether two entities share the same real-world identity, using a threshold:

$$\hat{y}_{ij} = \begin{cases} 1, & \text{if } S(v_i, v_j) \geq \tau \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

τ is a constant limit. The model can be optimized with a binary cross-entropy objective to enable additional learning:

$$\mathcal{L} = - \sum_{(i,j)} [y_{ij} \log \hat{y}_{ij} + (1 - y_{ij}) \log(1 - \hat{y}_{ij})] \quad (6)$$

This expression allows the system to generalize about the best similarity bars using labeled data. Each time new data is received, the framework uses an incremental update strategy to update only the parts of the graph that are affected. This has a substantial decrease in the computation overhead and guarantees scalability to financial systems at the enterprise scale.

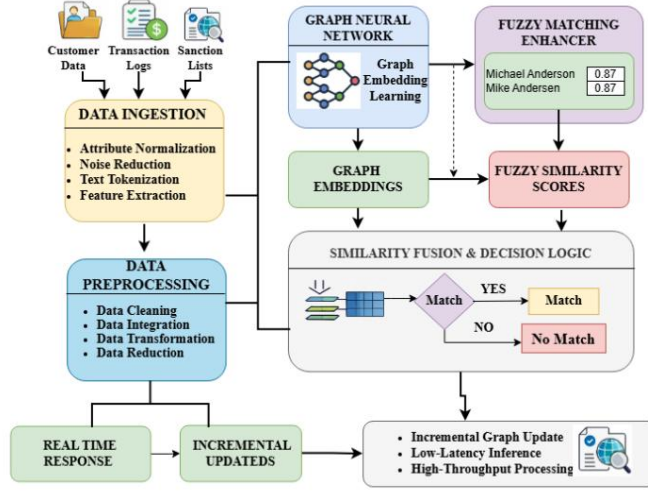


Figure 2.1. HG-FER Framework: Hybrid Graph-based Fuzzy Entity Resolution Architecture of Real-time financial sanctions screening.

The suggested architecture conceptualizes the data ingestion, preprocessing, knowledge graph building, GNN based representation learning, fuzzy similarity computation, similarity fusion and decision making in an integrated pipeline of data entity resolution. The proposed overall architecture of the HG-FER framework for real-time financial sanctions screening is presented in Fig. 2.1. A mix of real-world and synthetic datasets were used to evaluate the proposed framework for financial sanctions screening scenarios. Realistic transaction networks were generated using AMLSim and OFAC sanctions lists were used for the list of sanctioned entities and their attributes. Inconsistencies in the data were simulated by the inclusion of synthetic customer records with variations in names, addresses and identifiers. The integrated dataset offers a heterogeneous setting to test the effectiveness of the proposed HG-FER framework.

4. RESULTS AND DISCUSSION

The suggested HG-FER was tested experimentally to look into its effectiveness in real-time financial sanctions screening. Precision, recall, F1-score and false positive were used to assess performance, and visualization-based analyses of hybrid similarity patterns and entity interactions were performed. Decision thresholds and the framework's ability to correctly identify matched and non-matched entities in different financial data scenario's were also analysed.

TABLE I. PERFORMANCE COMPARISON OF ENTITY RESOLUTION METHODS

Method	Precision	Recall	F1-Score	False Positives
Rule-Based	0.72	0.85	0.78	45

Fuzzy Matching	0.80	0.83	0.81	32
GNN-Based	0.87	0.84	0.85	20
HG-FER (Proposed)	0.92	0.89	0.90	10

The proposed HG-FER framework is compared to the rule-based, fuzzy matching and GNN-based methods in Table I. HG-FER has the highest overall performance by improving the accuracy of matching, while also drastically reducing false positives. The outcomes show that the proposed entity resolution solution for enterprise financial sanctions screening is accurate and scalable, which can achieve the combination of knowledge graph-based relational learning and adaptive fuzzy matching.

TABLE II. EFFECT OF THRESHOLD ON PERFORMANCE

Threshold (τ)	Precision	Recall
0.40	0.85	0.92
0.50	0.90	0.89
0.60	0.92	0.87

The effect of the Decision threshold on the model performance is shown in Table II. A lower threshold value improves recall, and a higher threshold value increases precision. A threshold of 0.50 is a good compromise between precision and recall, and could be used for screening in practice for financial sanctions.



Figure 4.2. Entity risk interaction matrix of hybrid similarity scores.

The hybrid similarity scores between each pair of images are shown in Figure 4.1, obtained by the proposed HG-FER framework. High similarity regions suggest that there are strong relationships between entities, while low similarity regions show weak or unrelated entities. The smooth transition of similarity scores shows that the framework is able to express structural relationships as well as attribute-level uncertainty. This analysis demonstrates the utility of HG-FER in solving complicated relational patterns needed for accurate financial entity resolution.

The false positives by the various approaches to entity resolution are compared in Figure 4.2. The proposed HG-FER framework also performs the fewest number of false positive results relative to rule-based, fuzzy matching, and GNN-based methods, proving that the hybrid approach of integrating graph learning and fuzzy similarity is an effective way to minimize unnecessary compliance investigations and increase screening reliability. In summary, experimental results demonstrate that HG-FER is able to consistently outperform traditional entity resolution approaches. The framework integrates knowledge graph-based relational learning with adaptive fuzzy matching, and the two methods plus the framework

can realize accurate, scalable, and reliable entity resolution appropriate for real-time enterprise financial sanctions screening.

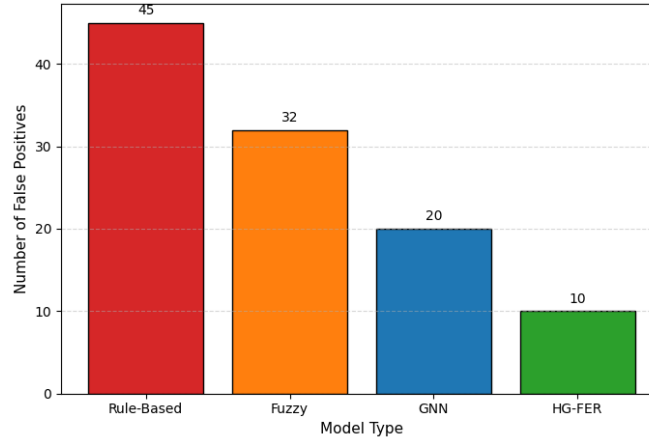


Figure 4.2. Comparisons of model false positives.

5. CONCLUSION AND FUTURE ENHANCEMENTS

Financial sanctions screening aims to accurately identify fraudulent transactions and users by matching the identity of financial transactions and users. The financial sanctions screening aims at accurately identifying the fraudulent transactions and users through matching the identity of financial transactions and users. The experimental results show that the framework consistently and effectively outperforms the traditional entity resolution methods in terms of the accuracy of matching entity and the number of false positives. Such results validate the effectiveness and scalability of structural relationship modeling when joined with uncertainty-aware similarity matching in financial environments with different types of data. One goal of future research is to generalize the model to handle streaming data on large graphs in an efficient way via more efficient graph update mechanisms. Furthermore, integrating different graph structures, financial data modalities, and explainable decision-making methods can enhance model flexibility, transparency, and adherence to regulatory shifts.

6. REFERENCES

- [1] [1]P. Li, Q. Zhao, Y. Liu, C. Zhong, J. Wang and Z. Lyu, 'Survey and Prospect for Applying Knowledge Graph in Enterprise Risk Management', *Computers, Materials & Continua*, Vol. 78, No. 3, 2024.
- [2] [2]X. Wang, M. Li, W. Chen, Y. Yao, Z. Li, Y. Liu and H. Zhang, 'Review and Application of Knowledge Graph in Crisis Management', *International Journal of Software Engineering and Knowledge Engineering*, Vol. 34, No. 03, pp. 393-425, 2024.
- [3] [3]P. Santos, R. Abreu, M. J. C. S. Reis, C. Serôdio and F. Branco, 'A systematic review of cyber threat intelligence: the effectiveness of technologies, strategies, and collaborations in combating modern threats', *Sensors*, Vol. 25, No. 14, Article 4272, 2025.
- [4] [4]H. Kheddar, D. W. Dawoud, A. I. Awad, Y. Himeur and M. K. Khan, 'Reinforcement-learning-based intrusion detection in communication networks: A review', *IEEE Communications Surveys & Tutorials*, Vol. 27, No. 4, pp. 2420-2469, 2024.

- [5] [5] G. Mo, W. Jia, C. Tan, W. Zhang and J. Rong, 'Risk Prediction of International Stock Markets with Complex Spatio-Temporal Correlations: A Spatio-Temporal Graph Convolutional Regression Model Integrating Uncertainty Quantification', *Journal of Risk and Financial Management*, Vol. 18, No. 9, Article 488, 2025.
- [6] [6] T. Qin, 'Multimodal Deep Learning Framework for Shadow Banking Risk Prediction: Dynamic Decision Optimization Integrating Knowledge Graph and Reinforcement Learning', 2025.
- [7] [7] X. Li, L. Chu, Y. Li, Z. Xing, F. Ding, J. Li and B. Ma, 'An Intelligent Financial Fraud Detection Support System Based on Three-Level Relationship Penetration', *Mathematics*, Vol. 12, No. 14, Article 2195, 2024.
- [8] S. Prakash, 'Demystifying Knowledge Graphs for AI-Enhanced Financial Decision Support with Graph Neural Networks', *Journal of Computer Science and Technology Studies*, Vol. 7, No. 6, pp. 580-590, 2025.
- [9] N. Xu, X. Chen, J. Luo, F. An, L. Wang and X. Li, 'Knowledge Graph–Large Language Model Fusion Approach for Emergency Knowledge Recommendation in Gas Tunnels', *Scientific Reports*, 2026.
- [10] J. Fan, L. K. Shar, R. Zhang, Z. Liu, W. Yang, D. Niyato and K.-Y. Lam, 'Deep Learning Approaches for Anti-Money Laundering on Mobile Transactions: Review, Framework, and Directions', *IEEE Internet of Things Journal*, 2026.
- [11] V. V. S. B. Nunna and R. Sahu, 'Entity Resolution in Distributed Systems: From Fuzzy Matching to Knowledge Graph Integration', *Journal of Computer Science and Technology Studies*, Vol. 7, No. 8, pp. 489-495, 2025.
- [12] N. T. M. Linh and T. Vo, 'An Integration of Fuzzy-Enhanced Heterophilic Graph Representation Learning for Improving Performance of Link Prediction', *Cybernetics and Systems*, pp. 1-29, 2025.
- [13] D. Yao, Y. Gu, G. Cong, H. Jin and X. Lv, 'Entity Resolution with Hierarchical Graph Attention Networks', *Proc. International Conference on Management of Data (SIGMOD)*, pp. 429-442, 2022.
- [14] O. R. Tiamiyu, 'Unveiling Hidden Money Laundering Networks: The Application of Graph Neural Networks in Financial Transaction Analysis', *Journal of Computational Analysis and Applications*, Vol. 34, No. 9, pp. 50-74, 2025.
- [15] S. Chakraborty and S. Sharov, 'Graph Based Approach on Financial Fraudulent Detection and Prediction', in *Applied Graph Data Science*, Morgan Kaufmann, pp. 25-37, 2025.
- [16] O. M. Olaniyi, I. S. Aroh, H. Onyii, O. C. Metibemu and O. I. Akinola, 'Graph Neural Networks for Multi-Layered Financial Crime Network Detection: An Explainable AI Framework for Anti-Money Laundering', *Journal of Engineering Research and Reports*, Vol. 28, No. 2, pp. 18-36, 2026.
- [17] G. Yu and Z. Luo, 'Financial Fraud Detection Using a Hybrid Deep Belief Network and Quantum Optimization Approach', *Discover Applied Sciences*, Vol. 7, No. 5, Article 454, 2025.
- [18] J. Liu, D. Cheng and C. Jiang, 'Preferential Selective-Aware Graph Neural Network for Preventing Attacks in Interbank Credit Rating', *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 36, No. 6, pp. 11414-11427, 2024.
- [19] M. Sharifi and D. Ahmadzadeh, 'Transformer-Gather, Fuzzy-Reconsider: A Scalable Hybrid Framework for Entity Resolution', *Proc. 15th International Conference on Computer and Knowledge Engineering (ICCKE)*, IEEE, pp. 1-6, 2025.
- [20] C. Wang, M. Wang, X. Wang, L. Zhang and Y. Long, 'Multi-Relational Graph Representation Learning for Financial Statement Fraud Detection', *Big Data Mining and Analytics*, Vol. 7, No. 3, pp. 920-941, 2024.