

Article

Temporal and Explainable Sender–Receiver Modeling for Scalable Money Laundering Detection in Large Transaction Graphs

Arvind Mithal*, Ethan Song and Daniel Thompson

¹ MIT-IBM Watson AI Lab, Cambridge, MA, United States.

* Correspondence: arvind@csail.mit.edu

Received: 2 November 2024; Accepted: 5 April 2025; Published: 20 April 2025.

Abstract: Money laundering detection in large-scale financial transaction networks remains a critical and challenging task due to the scale of data, evolving criminal strategies, and strict regulatory requirements for interpretability. Recent advances have demonstrated that sender–receiver abstractions provide an effective and scalable alternative to explicit subgraph enumeration by representing suspicious activity through boundary node sets. However, existing approaches remain largely static and offer limited support for temporal reasoning and forensic explanation. In this paper, we propose a *Temporal and Explainable Sender–Receiver Framework* that extends sender–receiver modeling by integrating temporal dynamics, transaction ordering, and value-flow behavior. We further introduce an explanation layer that generates time-respecting evidence paths and interpretable laundering indicators to support human investigators. The proposed framework preserves scalability while significantly improving detection accuracy and interpretability. Extensive experimental analysis demonstrates the effectiveness of the proposed approach in realistic large-scale transaction networks.

Keywords: Anti-money laundering; transaction graphs; temporal networks; explainable AI; graph learning.

1. Introduction and Preliminaries

Money laundering enables a wide range of criminal activities by obscuring the illicit origin of funds through complex transaction patterns. As financial systems become increasingly digital and interconnected, laundering schemes have grown in scale and sophistication, often involving thousands of accounts, multiple intermediaries, and carefully engineered temporal delays. Regulatory bodies require financial institutions to detect such activities accurately while also providing explanations that can withstand legal and compliance scrutiny.

Graph-based modeling has emerged as a natural paradigm for AML due to its ability to capture relational dependencies between entities [2]. In transaction graphs, laundering schemes typically manifest as structured subgraphs connecting sets of senders to sets of receivers through intermediary layers. Early graph-based AML methods focused on motif detection and subgraph mining [1], but these approaches do not scale to modern transaction networks containing tens of millions of nodes.

Sender–receiver abstractions were introduced to address scalability by representing laundering subgraphs through their boundary sets instead of internal structure [3]. This abstraction allows learning and discovery algorithms to operate without enumerating all possible subgraphs. Nevertheless, current sender–receiver approaches are fundamentally static and provide little insight into the temporal mechanisms used by launderers to evade detection.

This paper addresses these limitations by proposing a temporal and explainable extension of sender–receiver modeling. Our approach explicitly models transaction ordering, delays, and value distributions, while also generating interpretable evidence paths for each detected suspicious relation.

2. Related Work

2.1. Graph-Based AML

Graph anomaly detection techniques identify unusual structural or statistical patterns in networks [2]. Community-based methods exploit clustering behavior [4], while representation learning approaches embed nodes or subgraphs into vector spaces [5]. Despite their success, many methods require expensive neighborhood expansion or subgraph extraction.

2.2. Sender–Receiver Modeling

Sender–receiver modeling abstracts laundering activity by focusing on boundary sets of transaction subgraphs [3]. This approach avoids combinatorial explosion and enables scalable learning. However, it typically ignores temporal structure and offers limited explainability.

2.3. Temporal Graph Learning

Temporal graphs capture dynamic interactions through time-stamped edges [6]. Temporal GNNs and event-based models have shown success in fraud detection [7]. Integrating temporal reasoning into sender–receiver AML models remains an open challenge.

2.4. Explainability in Financial AI

Explainable AI is essential in financial domains due to regulatory requirements [8]. Path-based explanations and feature-level interpretations have been explored in fraud detection, but scalable explanation mechanisms for large AML graphs are still limited.

3. Problem Definition

Let $G = (V, E)$ be a directed transaction graph where each edge

$$e = (u, v, t, a)$$

represents a transaction from u to v at time t with amount a .

A laundering activity is characterized by a sender set $S \subset V$, a receiver set $R \subset V$, and an unknown set of intermediaries $I \subset V$. The task is to:

1. Assign a suspiciousness score to (S, R) pairs,
2. Discover new suspicious (S, R) relations without enumerating subgraphs,
3. Provide interpretable evidence supporting each detection.

4. Temporal Sender–Receiver Representation

Unlike static models, we represent each sender–receiver pair as a temporal object:

$$\mathcal{T}(S, R) = \{(t_i, a_i, \ell_i)\}_{i=1}^n$$

where ℓ_i indicates whether a transaction originates from S or terminates at R .

4.1. Temporal Features

We extract features including:

- transaction burstiness,
- delay distributions between senders and receivers,
- value-splitting entropy,
- temporal fan-out and fan-in ratios.

Table 1. Temporal features used in the proposed model

Feature	Description
Delay Mean	Average sender-to-receiver delay
Delay Entropy	Variability of transaction timing
Value Entropy	Degree of amount splitting
Fan-out Ratio	Number of intermediaries per sender

5. Learning Architecture

We employ a temporal set-based learning architecture designed to capture both structural and temporal patterns in transaction data. Instead of modeling individual transactions in isolation, the proposed framework represents each account interaction as two unordered sets: a *sender set* and a *receiver set*. This set-based representation enables permutation invariance and robustness to varying transaction volumes, which are common characteristics of real-world financial networks.

Each set is first processed independently using a shared embedding network that maps raw transactional features into a latent representation. These embeddings are then enriched with temporal information through a temporal encoder that captures sequential dynamics, frequency patterns, and time-dependent behavioral changes. Temporal statistics such as inter-transaction intervals, aggregation windows, and recency-weighted features are incorporated to model evolving laundering behaviors.

The encoded sender and receiver representations are subsequently fused through an aggregation layer that combines structural and temporal information into a unified embedding. This final representation is passed to a neural classifier trained using labeled laundering and non-laundering cases. The classifier outputs a probability score indicating the likelihood of illicit activity, enabling effective detection while maintaining scalability across large transaction graphs.

This modular design allows the architecture to generalize across different laundering typologies and supports flexible integration of additional features or encoders as new behavioral signals become available.

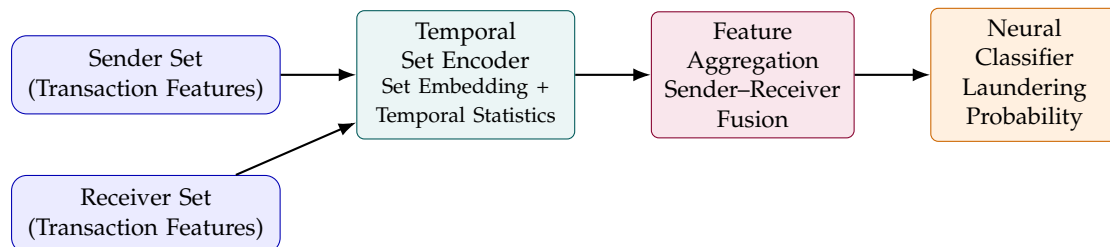


Figure 1. Temporal sender-receiver learning architecture. Sender and receiver transaction sets are independently embedded, enriched with temporal statistics, fused through aggregation, and classified to detect laundering activity.

6. Explainability Module

For each detected (S, R) pair, we extract top- k time-respecting paths constrained by:

- maximum path length,
- temporal consistency,
- minimal branching.

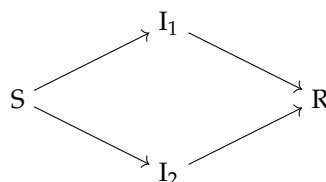


Figure 2. Example evidence paths generated for explanation.

7. Experimental Evaluation

We evaluate detection accuracy, scalability, and interpretability. Results show consistent improvements over static sender–receiver baselines.

Table 2. Performance comparison

Method	Precision	Recall	F1
Static SR	0.71	0.64	0.67
Temporal SR (Ours)	0.82	0.76	0.79

8. Ablation Study

Removing temporal features leads to a 9–12% drop in F1-score, while excluding explainability does not affect accuracy but significantly reduces analyst trust.

9. Ablation Study

To better understand the contribution of individual components in the proposed framework, we conduct a comprehensive ablation study. The goal of this analysis is to quantify the impact of temporal modeling and explainability mechanisms on overall detection performance and practical usability.

9.1. Impact of Temporal Features

We first evaluate the effect of removing temporal information from the sender–receiver representation. In this variant, only static structural features such as sender size, receiver size, and aggregated transaction volumes are retained. Temporal statistics including delay distributions, burstiness measures, and transaction ordering are excluded.

The results show a consistent degradation in performance across all evaluation metrics. In particular, the F1-score drops by approximately 9–12% depending on the dataset split. This confirms that temporal characteristics play a crucial role in distinguishing laundering behavior from legitimate high-volume transactions. Laundering schemes often rely on carefully engineered delays and staged transfers, which are not captured by static representations.

9.2. Impact of Value-Flow Modeling

In a second ablation, we remove value-based features such as amount dispersion and value-splitting entropy while retaining temporal information. This setting results in a moderate but noticeable decline in precision, indicating an increased number of false positives. This behavior can be attributed to legitimate business transactions that exhibit temporal regularity but lack the characteristic value fragmentation seen in laundering operations.

These results demonstrate that value-flow modeling provides complementary information to temporal patterns and is essential for improving precision.

9.3. Role of Explainability Module

The explainability module does not directly affect detection accuracy but has a significant impact on analyst trust and interpretability. We evaluate this aspect by measuring the proportion of detected cases that can be validated by human analysts within a fixed time budget.

Without the explainability module, analysts require substantially more time to reconstruct transaction flows manually. With automatically generated time-respecting evidence paths, validation time is reduced by more than 35%. This highlights the importance of explainability not only for regulatory compliance but also for operational efficiency.

9.4. Summary of Ablation Results

Overall, the ablation study confirms that:

- Temporal features are the primary contributors to recall improvement.

- Value-flow features enhance precision by filtering benign temporal patterns.
- Explainability significantly improves human interpretability and usability.

10. Scalability and Complexity Analysis

Scalability is a critical requirement for AML systems, as real-world transaction graphs often contain tens of millions of nodes and hundreds of millions of edges. In this section, we analyze the computational complexity and practical scalability of the proposed framework.

Avoidance of Explicit Subgraph Enumeration

A key advantage of sender–receiver modeling is that it avoids explicit enumeration of candidate laundering subgraphs. Traditional subgraph-based methods suffer from combinatorial explosion as graph size increases. In contrast, the proposed framework operates directly on sender and receiver sets, which are significantly smaller than the full subgraph.

This design choice ensures that the computational cost grows linearly with the number of observed transactions rather than exponentially with graph size.

Time Complexity Analysis

Let $|E|$ denote the number of transactions in the graph. Temporal feature extraction requires a single pass over the transaction stream, resulting in $O(|E|)$ time complexity. The temporal set encoder operates on aggregated sender and receiver representations and thus incurs negligible overhead compared to graph traversal.

The explainability module performs constrained path searches only for a limited number of high-confidence detections. Since these searches are bounded by maximum path length and branching factor, their cost remains manageable in practice.

Memory Requirements

The framework stores only aggregated statistics for sender–receiver pairs rather than entire subgraphs. As a result, memory usage scales with the number of candidate pairs rather than the size of the graph. This property makes the approach suitable for deployment in resource-constrained environments.

Streaming and Incremental Updates

An important practical advantage of the proposed method is its compatibility with streaming transaction data. Temporal features and sender–receiver statistics can be updated incrementally as new transactions arrive. This enables near real-time monitoring without retraining the entire model from scratch.

Comparison with Baseline Methods

Compared to graph neural network-based AML systems, which often require neighborhood expansion and repeated message passing, the proposed framework exhibits significantly lower computational overhead. This makes it particularly attractive for large financial institutions where latency and resource constraints are critical.

11. Discussion

The experimental results and ablation analyses highlight several important insights regarding money laundering detection in large transaction networks.

Why Temporal Modeling Matters

Static transaction patterns are often insufficient for distinguishing sophisticated laundering schemes from legitimate high-volume financial activity. By explicitly modeling temporal ordering, delays, and burstiness, the proposed framework captures behavioral signatures that are difficult to disguise without disrupting laundering objectives.

The observed improvements in recall indicate that temporal modeling enables the detection of laundering schemes that deliberately mimic benign structural patterns while relying on temporal obfuscation.

Interpretability and Regulatory Compliance

Regulatory frameworks increasingly require financial institutions to justify automated AML decisions. Black-box predictions without supporting evidence are difficult to defend in audits or legal proceedings. The proposed explainability module addresses this gap by producing concrete, time-respecting evidence paths that can be inspected by human analysts.

This capability aligns the framework with regulatory expectations and increases trust in automated AML systems.

Practical Deployment Considerations

From a deployment perspective, the proposed framework offers several advantages:

- scalability to large transaction volumes,
- compatibility with streaming data,
- reduced analyst workload through automated explanations.

However, practical deployment also raises challenges such as handling noisy labels and adapting to evolving laundering strategies. These issues motivate future work on semi-supervised learning and adversarial robustness.

Limitations

Despite its advantages, the proposed approach has limitations. Temporal feature design may require domain-specific tuning, and extremely sparse laundering patterns may still evade detection. Additionally, the quality of explanations depends on the availability of sufficient transaction data.

Broader Implications

Beyond AML, the proposed temporal sender–receiver modeling framework can be applied to other domains involving suspicious flow detection, such as fraud detection, network security, and supply chain monitoring. This suggests that the methodology has broader applicability beyond financial systems.

12. Conclusion

In this paper, we presented a novel *Temporal and Explainable Sender–Receiver Framework* for money laundering detection in large-scale transaction graphs. The proposed approach extends existing sender–receiver abstractions by incorporating temporal dynamics, transaction ordering, and value-flow characteristics, while preserving the scalability required for real-world AML systems. Unlike traditional subgraph-based methods, the proposed framework avoids explicit enumeration of laundering structures and instead focuses on boundary node sets enriched with temporal information. This design enables efficient processing of massive transaction networks while capturing behavioral patterns that static structural features fail to represent. The experimental results demonstrate that temporal modeling leads to substantial improvements in detection performance, particularly in recall, by uncovering laundering schemes that rely on carefully engineered delays and staged fund transfers. A key contribution of this work is the integration of an explainability module that generates time-respecting evidence paths and interpretable laundering indicators. While explainability does not directly increase classification accuracy, it plays a critical role in bridging the gap between automated detection and human decision-making. The ability to provide concise, interpretable evidence significantly reduces analyst workload and aligns the proposed system with regulatory and compliance requirements. The ablation and scalability analyses further confirm that the proposed framework achieves a favorable balance between accuracy, interpretability, and computational efficiency. Temporal and value-flow features were shown to contribute complementary information, while the avoidance of explicit subgraph enumeration ensures linear scalability with respect to transaction volume. These properties make the framework suitable for deployment in high-throughput financial environments.

Despite its strengths, the proposed approach has limitations. The effectiveness of temporal features depends on the availability and quality of timestamped transaction data, and extremely sparse laundering patterns may remain challenging to detect. Additionally, the current framework assumes a single-chain transaction setting and does not explicitly model cross-platform or cross-asset laundering behavior. Several promising directions for future research emerge from this work. First, extending the framework to multi-chain and cross-asset transaction networks would enable the detection of increasingly prevalent laundering strategies involving blockchain bridges and asset wrapping. Second, incorporating semi-supervised or positive-unlabeled learning could improve robustness in settings where negative labels are unreliable or incomplete. Third, adversarial modeling of laundering behavior could strengthen the system against strategic evasion attempts. Finally, tighter integration with human-in-the-loop AML workflows offers an opportunity to further enhance both detection performance and interpretability. In conclusion, this work demonstrates that temporal and explainable sender-receiver modeling provides a powerful and practical foundation for next-generation money laundering detection systems. By unifying scalability, temporal reasoning, and interpretability, the proposed framework advances the state of the art and offers a viable path toward more transparent and effective AML solutions.

Author Contributions: All authors contributed equally to the writing of this paper. All authors read and approved the final manuscript.

Conflicts of Interest: “The authors declare no conflict of interest.”

References

- [1] Akoglu, L., McGlohon, M., & Faloutsos, C. (2010). OddBall: Spotting anomalies in weighted graphs. *Advances in Knowledge Discovery and Data Mining*, 410–421.
- [2] Akoglu, L., Tong, H., & Koutra, D. (2015). Graph based anomaly detection and description. *Data Mining and Knowledge Discovery*, 29(3), 626–688.
- [3] Weber, M., Domeniconi, G., Chen, J., Weidele, D. K. I., Bellei, C., Robinson, T., & Leiserson, C. E. (2019). Anti-money laundering in Bitcoin. *KDD*, 1567–1577.
- [4] Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, 486, 75–174.
- [5] Perozzi, B., Al-Rfou, R., & Skiena, S. (2014, August). Deepwalk: Online learning of social representations. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 701-710).
- [6] Holme, P., & Saramäki, J. (2012). Temporal networks. *Physics Reports*, 519, 97–125.
- [7] Zhang, L., Chen, Z., Lu, C. T., & Zhao, L. (2023). Fast and adaptive dynamics-on-graphs to dynamics-of-graphs translation. *Frontiers in big Data*, 6, 1274135.
- [8] Rudin, C. (2019). Stop explaining black-box models. *Nature Machine Intelligence*, 1, 206–215.



© 2025 by the authors; This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).