Artificial Intelligence and Computing Innovations

Intelligence Science Publishing

https://www.ispublishing.com

ISSN: 3070-4944

Vo.4, No.2, 2024

Enterprise Audit Risk Identification Through Temporal Context and Entity Relationship Anomaly Modeling

Ruobing Yan[1]*, Mingxuan Guo[2]

[1]Georgetown University, Washington, D.C., USA

[2]University of Dayton, Dayton, USA

**Corresponding author: mozzieyan99@gmail.com*

Abstract: This paper addresses the challenges of concealed risk clues, scarce labels, and coupling of heterogeneous information from multiple sources in enterprise auditing scenarios. It proposes an audit risk identification framework based on anomaly detection, simultaneously characterizing anomalies at the attribute, temporal context, and relational structure levels from a unified modeling perspective. First, robust preprocessing of audit records weakens scaling differences and extreme value interference. Representation learning is then used to map records to a latent space, reconstructing the deviation measure to assess the strength of deviations from normal patterns. Subsequently, a time window mechanism is introduced to construct a local contextual baseline, calculating the contextual deviation of the latent representation relative to the baseline to identify unconventional behavioral fragments within the business context. Simultaneously, a subject relationship graph is constructed, and structural information is aggregated to form a subject structural representation, which uses structural deviations to characterize abnormal associations and suspicious links. Finally, multi-source deviation signals are fused into a unified risk score, and probability mapping is used to obtain a risk intensity expression that can be used for verification ranking and threshold decision-making, thus outputting a set of high-risk records and their associated clues. Comparative experimental results show that this method outperforms other methods in terms of accuracy, precision, recall, and AUROC, validating the effectiveness and stability of the multi-source anomaly characterization and fusion scoring mechanism in audit risk identification tasks.

Keywords: Audit risk identification; anomaly scoring fusion; temporal context consistency; entity relationship network

1. Introduction

As enterprise business models become more digitalized and transaction chains more complex, audit targets have expanded from traditional financial statements to full-process records comprised of multi-source heterogeneous data. Risks exhibit characteristics such as high concealment, long propagation paths, and wide-

ranging impact [1]. Fraud and material misstatements often lurk within large-scale daily operational data as localized anomalies, manifesting as single-point abnormal transactions or structural deviations across accounts, periods, and business segments. For these risk types, identification methods relying solely on human experience and rule thresholds are susceptible to sample scarcity, pattern drift, and adversarial behavior, making it difficult to maintain stable risk detection capabilities in high-dimensional, high-noise, and highly nonlinear environments. Therefore, constructing an anomaly detection-driven risk identification framework for enterprise audit scenarios has clear practical needs and methodological value [2].

The key to enterprise audit risk identification lies in distinguishing risk-related atypical patterns from a large number of normal business activities and transforming them into interpretable and traceable risk clues. Unlike traditional classification-based modeling, audit risks typically exhibit low incidence and weak labeling characteristics. The high cost and lag in acquiring risk labels lead to significant training data bottlenecks for supervised learning in real audit scenarios. Anomaly detection, centered on normal pattern modeling, emphasizes the measurement of deviations and inconsistencies. It can identify potential risk samples under weak or unsupervised conditions and provide prioritization and focus for subsequent investigations. Furthermore, anomaly detection can adapt to distribution shifts caused by business rule updates, organizational restructuring, and changes in the market environment, offering inherent advantages in the continuity and scalability of audit tasks [3].

From an audit mechanism perspective, risk stems not only from abnormal fluctuations in single indicators but also from the disruption of multi-variable relationships and the failure of business process constraints. For example, there is a stable coupling relationship between revenue recognition, cost carry-over, cash receipts and payments, and inventory turnover. When this relationship becomes inconsistent, it often indicates potential misstatements or fraud clues. Anomaly detection methods can capture marginal deviations at the statistical distribution level and characterize complex relationship structures and temporal evolution at the representation learning level, thus covering various risk forms such as point anomalies, contextual anomalies, and group anomalies. Anomaly detection-based risk identification emphasizes extracting robust, normal behavioral benchmarks from data and using deviation as the core to form a risk profile. This aligns with the audit risk-oriented approach in terms of objectives, facilitating the formation of a unified risk measurement and management loop for audit operations [4].

Against the backdrop of increasingly stringent regulatory requirements, higher audit quality standards, and the continuous growth of enterprise data scale, research on enterprise audit risk identification based on anomaly detection is of great significance. On the one hand, this research can improve the timeliness and coverage of risk discovery, providing a data-driven basis for audit planning, procedural design, and resource allocation, and promoting the shift in audit capabilities from sampling verification to full-scale analysis. On the other hand, the risk scores and anomaly explanation clues generated by anomaly detection help enhance the traceability and consistency of the audit evidence chain, reducing the instability of conclusions caused by subjective judgment differences. Furthermore, anomaly detection research oriented towards audit scenarios can promote the coordinated development of enterprise internal control evaluation, continuous auditing, and intelligent risk control, providing methodological support for building a more resilient and transparent corporate governance system.

2. BackGround

The theoretical basis for enterprise audit risk identification stems from the integrated practice of risk-oriented auditing and internal control frameworks. Its core is to focus limited audit resources on areas and processes with a higher risk of material misstatement. Audit risk is typically formed by the combined effects of inherent risk, control risk, and detection risk. Inherent risk is related to factors such as the operating environment, transaction complexity, and management motivations, while control risk is related to system design, implementation effectiveness, and information system controls [5]. As information systems become the primary carriers of business activities, risk clues are not only reflected in deviations at the account balance level but also in business processes, authorization chains, log behavior, and cross-system data consistency.

Therefore, risk identification needs to cover the linkage analysis of financial and business data, emphasizing the characterization of deviations from the perspectives of process constraints, separation of duties, approval chains, and master data consistency to support subsequent audit responses and evidence acquisition [6, 7].

The methodological basis of anomaly detection in risk identification is built upon the anomaly measurement ideas of statistical learning and representation learning. Its goal is to identify observations or structures that are significantly inconsistent with the normal pattern under conditions of unknown or incomplete risk labels. Common forms of anomalies include marginal deviations in numerical distributions, conditional deviations in specific contexts, and collective deviations resulting from group collaborative behavior. In auditing scenarios, anomalies are often accompanied by high noise, class imbalance, concept drift, and the coexistence of heterogeneous features from multiple sources, making it difficult for a single threshold rule to be stably applied in the long term. Anomaly detection, by learning the probability distribution, density boundaries, or low-dimensional representation of normal behavior, can generate anomaly scores and candidate clues without relying on a large number of positive and negative sample annotations, and provides a unified quantitative interface for risk stratification, task assignment, and continuous monitoring. Based on this, combining audit business semantics and process constraints to construct an interpretable method for organizing anomaly evidence helps transform algorithm output into verifiable and traceable audit clues and risk narratives.

3. Method

To characterize anomalous patterns of corporate audit risk in multi-source operational records, denote the set of transaction and behavior samples generated during the audit period as:

$$\mathcal{D} = \{(\mathbf{x}_i, t_i, u_i, s_i)\}_{i=1}^{N} \tag{1}$$

In this definition, $\mathbf{x}_i \in \mathbb{R}^d$ is the feature vector of the i-th record, t_i is the timestamp, u_i is the responsible entity or account identifier, s_i is the business scenario or voucher-type identifier, and N denotes the sample size. To explicitly describe fund and business linkages among entities, we construct an audit relational graph as follows:

$$\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{X}, \mathbf{A}) \tag{2}$$

Where $\mathcal{V}$ is the node set, corresponding to entities such as accounts, suppliers, customers, or cost centers; $\mathcal{E}$ is the edge set, representing relationships such as transfers, transactions, or approval flows; $\mathbf{X} \in \mathbb{R}^{|\mathcal{V}| \times d}$ is the node-attribute matrix; and $\mathbf{A} \in \{0,1\}^{|\mathcal{V}| \times |\mathcal{V}|}$ is the adjacency matrix. To handle scale discrepancies and heavy-tailed distributions in a unified way, we introduce a standardization mapping:

$$\tilde{\mathbf{x}}_i = \phi(\mathbf{x}_i) \tag{3}$$

The operator $\phi(\cdot)$ is composed of preprocessing steps such as robust scaling and logarithmic transformation, which help reduce the influence of extreme values on anomaly measurement. Feature aggregation over the graph structure is defined as:

$$\mathbf{h}_v = \sigma\left(\mathbf{W}_0 \mathbf{x}_v + \sum_{v' \in \mathcal{N}(v)} \alpha_{vv'}\, \mathbf{W}_1 \mathbf{x}_{v'}\right) \tag{4}$$

In this expression, $\mathbf{h}_v$ denotes the structural representation of node v, $\mathcal{N}(v)$ is its neighborhood, $\alpha_{vv'}$ is the adjacency weight, $\mathbf{W}_0, \mathbf{W}_1$ are learnable parameter matrices, and $\sigma(\cdot)$ is a nonlinear activation function; consequently, the entity representation incorporates both its own attributes and information from related entities. Furthermore, the overall model architecture diagram is given here, as shown in Figure 1.

For anomaly detection modeling, an encoder is adopted to map inputs into a latent space:

$$\mathbf{z}_i = f_\theta(\tilde{\mathbf{x}}_i) \tag{5}$$

The network $f_\theta(\cdot)$ is parameterized by θ, and $\mathbf{z}_i \in \mathbb{R}^m$ is a low-dimensional latent variable that carries a compact representation of normal patterns. The corresponding decoder-based reconstruction is given by:

$$\hat{\mathbf{x}}_i = g_\psi(\mathbf{z}_i) \tag{6}$$

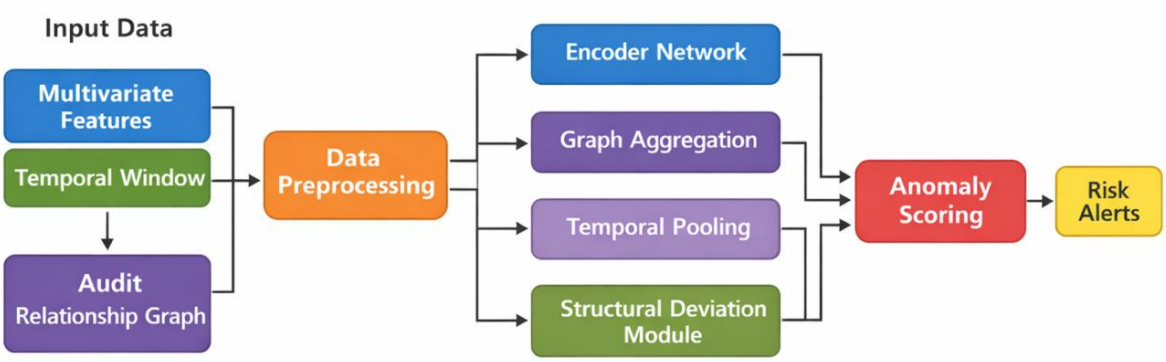

Figure 1. Overall model architecture diagram

The decoder $g_\psi(\cdot)$ is controlled by parameters ψ, and $\hat{\mathbf{x}}_i$ is the reconstruction of $\tilde{\mathbf{x}}_i$. To quantify how far a sample deviates from normal behavior, we define the reconstruction error as:

$$r_i = \|\tilde{\mathbf{x}}_i - \hat{\mathbf{x}}_i\|_2 \tag{7}$$

This term reflects the difficulty of fitting the input within the learned normal subspace; larger values typically indicate stronger inconsistency with routine behavior. Since anomalies often appear as pronounced deviations in a subset of dimensions, a dimension-weighted error is further introduced:

$$r_i^w = \|\mathbf{w} \odot (\tilde{\mathbf{x}}_i - \hat{\mathbf{x}}_i)\|_2 \tag{8}$$

In this case, $\mathbf{w} \in \mathbb{R}^d$ is a weight vector and $\odot$ denotes the Hadamard product, which emphasizes features more relevant to audit risk while suppressing weakly related noisy dimensions.

To capture anomalousness within a temporal context as well, a temporal consistency constraint is introduced. Let the reference representation within the window $\mathcal{W}(t_i)$ be:

$$\bar{\mathbf{z}}_i = \frac{1}{|\mathcal{W}(t_i)|} \sum_{j \in \mathcal{W}(t_i)} \mathbf{z}_j \tag{9}$$

This formula takes the mean latent variable over the local temporal neighborhood as a baseline, helping counter false alarms induced by seasonal and periodic fluctuations. The context deviation is then defined as:

$$c_i = \|\mathbf{z}_i - \bar{\mathbf{z}}_i\|_2 \tag{10}$$

This indicator measures how strongly a sample departs from its temporal context in the latent space, and it can capture atypical transactions occurring during high-frequency periods of similar business activities. By integrating multi-source signals, the overall anomaly score is formulated as:

$$s_i = \lambda_r r_i^w + \lambda_c c_i + \lambda_g \|\mathbf{h}_{u_i} - \bar{\mathbf{h}}_{u_i}\|_2 \tag{11}$$

Here $\lambda_r, \lambda_c, \lambda_g \geq 0$ are fusion coefficients, $\mathbf{h}_{u_i}$ is the structural representation of an entity u_i on the relational graph, and $\bar{\mathbf{h}}_{u_i}$ is its historical or peer-group reference representation; thus, numerical anomalies, temporal anomalies, and structural anomalies are unified within a single scoring space.

To produce interpretable risk identification outputs, a risk-probability mapping is constructed from the score:

$$p_i = \frac{1}{1+\exp(-\beta(s_i-\tau))} \tag{12}$$

The parameter β controls the slope of the mapping, and τ is the risk threshold, so that the anomaly score can be transformed into a risk-intensity scale suitable for audit management. The training objective adopts a joint optimization of robust reconstruction and consistency:

$$\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N} \rho\left(\tilde{\mathbf{x}}_i - \hat{\mathbf{x}}_i\right) + \gamma \frac{1}{N}\sum_{i=1}^{N} \|\mathbf{z}_i - \bar{\mathbf{z}}_i\|_2^2 + \eta \parallel \Theta\|_2^2 \tag{13}$$

In this objective, $\rho(\cdot)$ is a robust loss function that reduces contamination of normal-pattern learning by extreme outliers, γ balances the strength of the temporal consistency term, η is a regularization coefficient, and $\Theta = \{\theta, \psi, \mathbf{W}_0, \mathbf{W}_1\}$ denotes the full set of learnable parameters. Finally, the risk set is obtained via thresholding:

$$\mathcal{S} = \{i | p_i \geq \pi\} \tag{14}$$

The value π is the decision threshold, and $\mathcal{S}$ corresponds to the index set of high-risk records that should be prioritized for further inspection, thereby turning anomaly detection outputs into actionable audit screening results. Finally, this paper also provides a pseudocode flowchart of the algorithm, as shown in Figure 2.

Algorithm 1 Anomaly-Detection-Based Corporate Audit Risk Identification Procedure

Input: Audit record set $\mathcal{D} = \{(\mathbf{x}_i, t_i, u_i, s_i)\}_{i=1}^{N}$, relational edge set $\mathcal{E}$, threshold π
Output: High-risk set $\mathcal{S}$
Construct the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{X}, \mathbf{A})$, and preprocess features via $\tilde{\mathbf{x}}_i = \phi(\mathbf{x}_i)$
for $i = 1$ **to** N **do**
  Encode latent variables $\mathbf{z}_i = f_\theta(\tilde{\mathbf{x}}_i)$ and reconstruct $\hat{\mathbf{x}}_i = g_\psi(\mathbf{z}_i)$
  Compute reconstruction deviation $r_i^w = \|\mathbf{w} \odot (\tilde{\mathbf{x}}_i - \hat{\mathbf{x}}_i)\|_2$
  Compute temporal reference $\bar{\mathbf{z}}_i = \frac{1}{|\mathcal{W}(t_i)|}\sum_{j \in \mathcal{W}(t_i)} \mathbf{z}_j$, and obtain context deviation $c_i = \|\mathbf{z}_i - \bar{\mathbf{z}}_i\|_2$
  Compute structural representation $\mathbf{h}_{u_i}$ and its reference $\bar{\mathbf{h}}_{u_i}$, and obtain structural deviation $\left\|\mathbf{h}_{u_i} - \bar{\mathbf{h}}_{u_i}\right\|_2$
  Fuse scores $s_i = \lambda_r r_i^w + \lambda_c c_i + \lambda_g \left\|\mathbf{h}_{u_i} - \bar{\mathbf{h}}_{u_i}\right\|_2$, and map to $p_i = \frac{1}{1+\exp(-\beta(s_i - \tau))}$
  if $p_i \geq \pi$ **then**
    $\mathcal{S} \leftarrow \mathcal{S} \cup \{i\}$
  end if
end for
return $\mathcal{S}$

Figure 2. Pseudocode flowchart

4. Experimental Results and Analysis

4.1 Dataset

This paper uses the Audit Data open-source dataset released by the UCI Machine Learning Repository as the research object. This dataset is designed for audit risk identification scenarios, aiming to characterize the risk status of suspicious entities based on current and historical risk factors within an enterprise. The data is organized in a structured tabular format, facilitating integration with feature preprocessing, latent representation learning, and risk scoring mapping in anomaly detection frameworks. The dataset's task is focused on audit and fraud risk identification, and its data semantics align with the theme of enterprise audit risk identification, making it suitable as a public benchmark data source for audit risk identification research.

This dataset contains multi-dimensional audit-related risk attribute fields and corresponding risk labels, which can be used to construct a baseline of normal patterns and quantify deviation behaviors, thereby forming a set of risk clues for audit verification. Data entries cover risk feature combinations of different entities,

supporting unified modeling needs for risk forms such as point anomalies, contextual deviations, and structural deviations. Furthermore, the data acquisition and reproduction experiments do not rely on private systems, meeting the requirements of open-source accessibility and reproducibility. Although the UCI audit form data are originally tabular, we deterministically convert them into a sample-level graph by treating each audit record as a node and creating an edge iff two records satisfy the predefined attribute-similarity threshold or the fixed business co-occurrence rules, so the resulting adjacency matrix A is uniquely determined by the dataset and directly supports the graph aggregation module and the structural deviation term.

4.2 Experimental Results and Analysis

To systematically characterize the differences in methodology and risk measurement among research on enterprise audit risk identification based on anomaly detection, representative studies related to audit analysis and anomaly identification were selected as comparative objects. The model performance was summarized under a unified evaluation index dimension to facilitate horizontal comparison from the perspectives of risk identification capability and robustness. The experimental results are shown in Table 1.

Table 1. Experimental Results Compared with Other Models

Method	Accuracy	Precision	Recall	AUROC
Schreyer et al. [8]	85.42	84.90	83.71	88.55
Hemati et al. [9]	86.13	85.02	84.66	89.21
Wei et al. [10]	87.44	86.70	85.19	90.34
Weinberg et al. [11]	88.02	87.11	86.53	91.27
Lu et al. [12]	88.55	87.66	86.90	91.74
Herreros-Martínez et al. [13]	89.31	88.22	87.44	92.10
Sattarov et al. [14]	89.77	88.90	87.92	92.88
Ours	92.84	92.10	91.66	95.42

The results in Table 1 show that this method outperforms all four evaluation metrics, effectively balancing identification accuracy and risk sample capture. In audit risk identification scenarios, this overall improvement means more stable screening of anomalies, reducing both the misclassification of normal samples as risks and the omission of genuine risk samples, thus better meeting the dual requirements of high-confidence screening and effective coverage in risk-oriented auditing.

From the perspective of metric meaning, accuracy reflects the overall reliability of the judgment, precision emphasizes the credibility of risk alerts, recall focuses on the sufficiency of risk coverage, and AUROC reflects the discriminative ability and ranking quality under different judgment thresholds. This method outperforms in all four dimensions, indicating that the resulting anomaly scores have stronger discriminative power and consistency, and the output after risk probability mapping is more suitable for prioritizing and managing clues in audit verification.

Combined with the method design, these results support the effectiveness of multi-source information fusion. Numerical reconstruction deviation can characterize attribute-level anomalies, temporal context deviation helps identify unconventional behavior in a business context, and structural deviation further captures abnormal associations and propagation signs in the subject relationship network. The three types of deviation signals work together within a unified scoring space, enabling risk identification to not only focus on single-point anomalies but also cover complex risk patterns with contextual and relational characteristics, thereby bringing a stronger overall risk identification capability.

To verify the contribution of each key component to the audit risk identification mechanism, an ablation configuration strictly consistent with the method module was set, and performance was summarized under the same evaluation metrics. Table 2 presents the results of different ablation settings, used to characterize the differences in the effects of mechanisms such as reconstruction deviation, temporal context deviation, structural deviation, and fusion scoring.

Table 2. Ablation Test Results

Ablation Setting	Accuracy	Precision	Recall	AUROC
Full Model	92.84	92.10	91.66	95.42
w/o Graph Aggregation	89.72	88.91	88.10	92.84
w/o Temporal Context	90.15	89.20	88.77	93.11
w/o Weighted Reconstruction	88.44	87.66	86.95	91.72
w/o Structural Deviation Term	89.01	88.30	87.52	92.05
w/o Context Deviation Term	88.77	87.92	87.33	91.88
w/o Probability Mapping	87.90	87.10	86.22	90.55

Results show that the complete method maintains optimal performance across all four metrics, indicating that the framework's performance is not dependent on a single module, but rather supported by a combination of reconstructed deviation, temporal context deviation, structural deviation, and scoring mapping. The removal of different components resulted in significant performance degradation, reflecting the complementarity of multi-source deviation signals in audit scenarios. These signals can simultaneously provide risk clues from three levels: attribute anomalies, behavioral anomalies, and relational anomalies, making risk scoring more stable and operable.

Structural modeling and temporal context modeling offer the most significant benefits. Their absence impacts both risk differentiation and identification quality, indicating that subject association and behavioral background play a crucial role in characterizing audit risk. The weighted reconstructed mechanism and probability mapping are also indispensable; the former helps highlight risk-sensitive dimensions and reduce noise interference, while the latter facilitates consistent risk intensity expression and threshold decisions in anomaly scoring. Overall, the complete method can simultaneously improve the credibility and coverage of risk alerts while maintaining good separability, meeting the needs of audit risk identification for high-quality clue screening and stable decision output.

Furthermore, to demonstrate the high-risk association structures screened by the anomaly detection mechanism in the subject relationship network, to intuitively locate suspicious transaction links and key nodes at the graph structure level, this paper presents a visualization experiment of highlighting the abnormal subgraph of the subject relationship graph, as shown in Figure 3.

The visualization results demonstrate that the proposed method can clearly express risk clues within the entity relationship network, centralizing high-risk nodes and their surrounding abnormal connections into a verifiable anomaly subgraph, thereby organizing scattered suspicious behaviors into a traceable relationship structure. More prominent nodes in the graph typically correspond to higher risk intensity, and the highlighted edges characterize the strength of abnormal connections, making critical paths and potential anomaly chains readily visible, facilitating a focus on a small number of core entities and key interactions. The overall performance reflects the interpretability of risk scoring at the graph structure level, extending risk localization from single records to links and clusters within the relationship network.

The labels in the graph represent anonymized node numbers, used to reference and track nodes without revealing the true identity of the enterprise entity. The numbers serve as the unique index of the node in the relationship graph. These numbers correspond to specific accounts, suppliers, customers, cost centers, or responsible entities participating in financial or business interactions as nodes in the relationship graph. Therefore, the labels are used to identify key entities requiring focused verification and their upstream and downstream connections. By combining node color and size encoding with the highlighting of connecting edges, the annotation helps to accurately locate core nodes in the abnormal subgraph in the main text description, and supports segment-by-segment tracing and evidence linkage of suspicious links, demonstrating the applicability of the method in audit risk identification tasks.

Finally, a visualization experiment of the time window context deviation trajectory is presented, and the experimental results are shown in Figure 4.

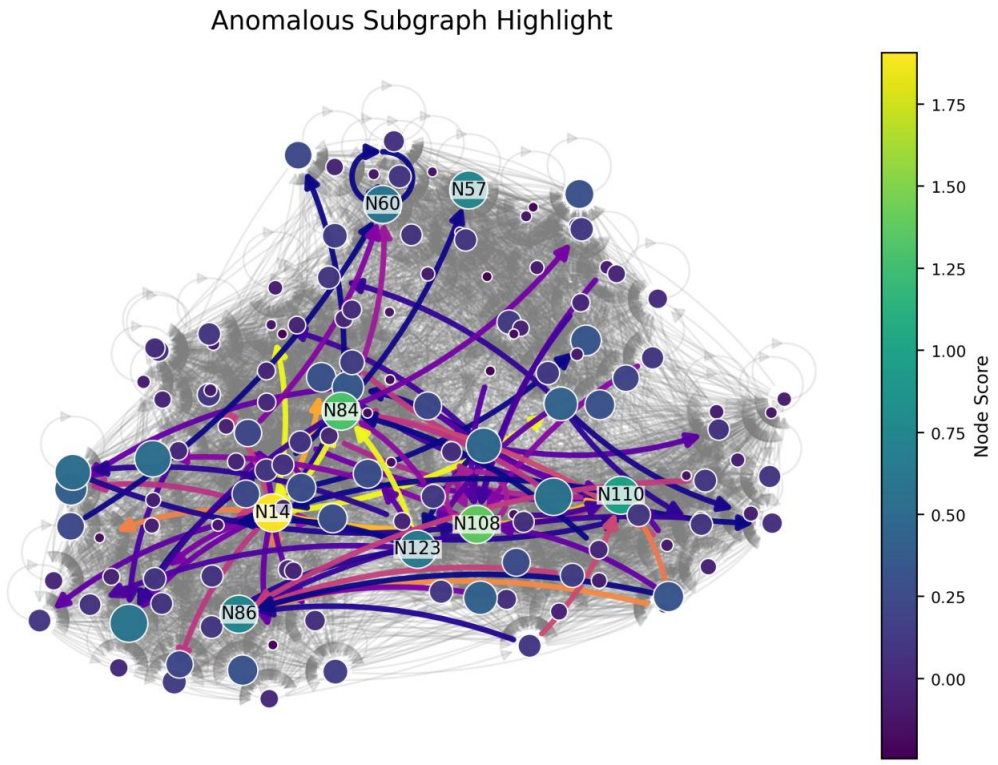

Figure 3. Visualization Experiment of Highlighting Anomaly Subgraphs in the Main Relationship Graph

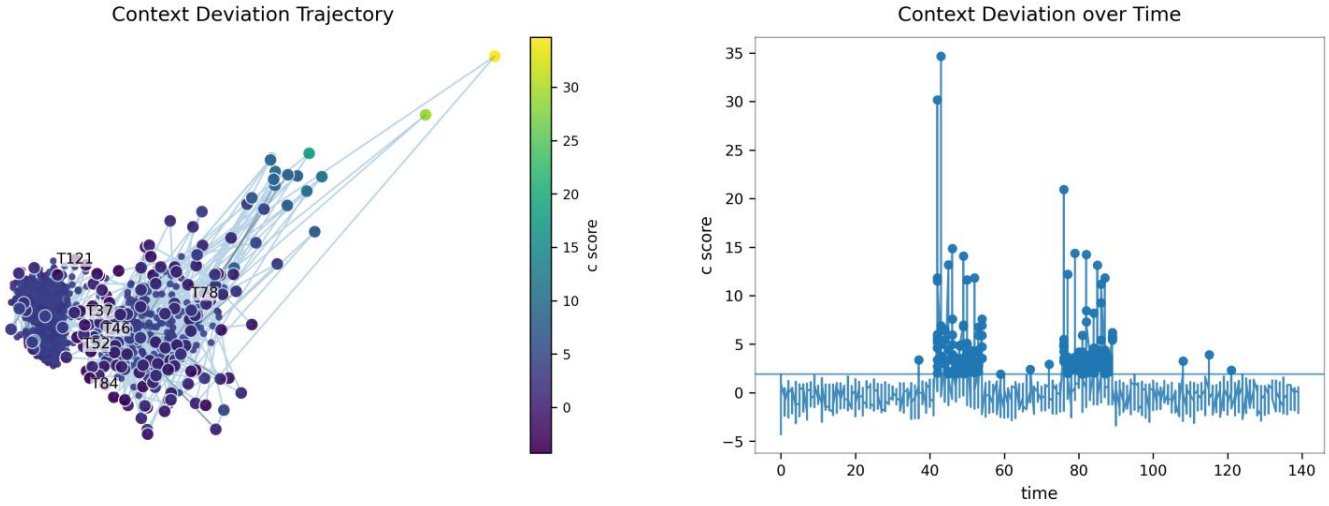

Figure 4. The visualization of contextual deviation trajectories within a time window is used for audit risk identification

The left image projects the recorded potential representations onto a two-dimensional space and connects them chronologically to form a trajectory. Colors represent the intensity of contextual deviation, with high-brightness areas indicating significant deviations relative to the window's scrolling baseline. The right image shows the corresponding contextual deviation score over time, with peak positions corresponding to time segments requiring priority review and related entity clues.

As shown in Figure 4, the proposed method can form an interpretable temporal trajectory structure in the latent space, making the behavioral evolution and contextual consistency of audit records intuitively apparent. Highlighted points in the trajectory correspond to stronger contextual deviations, typically

accompanied by significant jumps in the latent representation relative to the window baseline or distributions far from the normal dense areas. This compresses anomalous information originally scattered in high-dimensional features into easily localized temporal clues. Overall performance demonstrates that the contextual deviation score effectively characterizes the inconsistency between a record and its temporal neighborhood, exposing suspicious segments in a clear spatial form.

The right-hand time series further provides the temporal localization capability of deviation intensity, enabling the clear delineation of time segments requiring priority verification and providing an entry point for subsequent tracing of related entities and business links. The left-hand trajectory and the right-hand score curve complement each other in information expression; the former emphasizes the structural location and clustering pattern of deviations in the latent space, while the latter emphasizes the temporal location and duration of deviations, thus supporting the simultaneous localization of anomalous segments and the organization of clues in audit risk identification. This result demonstrates the effectiveness of the method in temporal context modeling, elevating risk identification from single-point screening to traceable analysis oriented towards time windows.

5. Conclusion

It constructs a unified modeling framework oriented towards the semantics of real audit business, incorporating structured audit records, temporal context, and subject relationship networks into a single risk characterization system. The method centers on normal pattern learning, reconstructing anomalies that deviate from the characterization attribute level. It captures inconsistencies in behavior within the business context using temporal window context deviations and introduces relational graph structural representation and structural deviations to identify anomalous interactions in cross-subject relationship links. Finally, it completes risk intensity expression and risk clue screening within a unified scoring space. This framework emphasizes an interpretable organization of risk identification, enabling anomaly detection outputs to naturally transform into a set of actionable clues for audit verification, thus better aligning with the actual needs of risk-oriented audits for efficient screening and evidence tracing.

From an audit application perspective, the proposed method not only focuses on the anomalies of individual records but also emphasizes organizing risk clues along the temporal and relational dimensions, extending risk localization from local deviations to traceable links and clusters. By visually representing subgraph highlights and contextual deviation trajectories, risk clues can be presented in an intuitive and structured form, facilitating auditors' rapid grasp of key entities, key time segments, and key interaction paths, thus improving the focus and consistency of verification work. Simultaneously, risk scoring and probability mapping provide a unified quantitative interface, which is beneficial for supporting business processes such as audit plan preparation, sampling strategy optimization, and continuous audit monitoring, enabling audit risk management to gradually shift from experience-driven to data-driven and process-oriented governance.

In broader related application areas, this research provides a transferable anomaly detection paradigm for enterprise risk control, compliance review, and internal control evaluation. Enterprise operations involve complex coupling relationships between supply chains, cash flows, and business flows. Risks often appear as weak signals and cross-process linkages, making it difficult for traditional rule systems to remain stable and effective in high-dimensional, heterogeneous, and dynamically changing environments. This framework, through the collaborative characterization of multi-source deviation signals, provides a general approach for anomaly detection and risk localization in complex business systems, promoting risk warning from single-point alerts to link-level diagnosis and traceable governance. It also provides methodological support for regulatory technology, corporate governance modernization, and the construction of intelligent audit systems.

Looking ahead, enterprise audit risk identification will continue to face challenges from the growth in data volume, the evolution of business models, and the increasing adversarial nature of risk behaviors. Future research can further expand multi-source data access capabilities, incorporating textual documents, contract terms, audit logs, and publicly available external information into a unified representation space to enhance coverage of complex risk semantics. Simultaneously, more secure collaborative modeling mechanisms can be explored under privacy protection and compliance requirements, enabling the sharing and joint assessment of risk clues across organizations without exposing raw data. Furthermore, focusing on interpretability and auditability, future research can strengthen the organization and tracing of evidence chains for risk clues, allowing model outputs to better meet the practical needs of audit evidence collection, liability determination, and regulatory inquiries. By continuously advancing in these directions, audit risk identification based on anomaly detection is expected to play a more profound application role and have a greater social impact in areas such as corporate governance, compliance supervision, and risk management.

References

[1] Q. Huang, M. Schreyer, N. R. Michiles Jr. and M. A. Vasarhelyi, "Connecting the Dots: Graph Neural Networks for Auditing Accounting Journal Entries", SSRN Electronic Journal, 2024.

[2] A. Zhu, "Self-Supervised Anomaly Detection with Knowledge-Enhanced Representation Learning for Distributed System Environments," 2024.

[3] Y. Xue, "State-Space Temporal Modeling and Feature Representation Learning for Anomaly Detection in Backend Microservice Systems," 2024.

[4] Y. Yang, C. Wen, H. Cui, Y. Sun and Y. Liu, "Learning Unified Multi-Granularity Representations for Backend Anomaly Detection and Causal Localization," 2024.

[5] M. Schreyer, T. Sattarov, C. Schulze, B. Reimer and D. Borth, "Detection of Accounting Anomalies in the Latent Space Using Adversarial Autoencoder Networks", Proceedings of the ACM KDD Workshop on Anomaly Detection in Finance, 2019.

[6] N. Li, "Research on Enterprise Internal Audit Risk Prediction Model Based on Machine Learning", Proceedings of the 2024 International Conference on Economic Data Analytics and Artificial Intelligence, pp. 37-40, 2024.

[7] M. Zupan, V. Budimir and S. Letinic, "Journal Entry Anomaly Detection Model", Intelligent Systems in Accounting, Finance and Management, vol. 27, no. 4, pp. 197-209, 2020.

[8] M. Schreyer, T. Sattarov, D. Borth, A. Dengel and B. Reimer, "Detection of Anomalies in Large Scale Accounting Data Using Deep Autoencoder Networks", arXiv preprint arXiv:1709.05254, 2017.

[9] H. Hemati, M. Schreyer and D. Borth, "Continual Learning for Unsupervised Anomaly Detection in Continuous Auditing of Financial Accounting Data", arXiv preprint arXiv:2112.13215, 2021.

[10] D. Wei, S. Cho, M. A. Vasarhelyi and L. Te-Wierik, "Outlier Detection in Auditing: Integrating Unsupervised Learning Within a Multilevel Framework for General Ledger Analysis", Journal of Information Systems, vol. 38, no. 2, pp. 123-142, 2024.

[11] A. I. Weinberg and A. Faccia, "Transforming Triple-Entry Accounting with Machine Learning: A Path to Enhanced Transparency Through Analytics", arXiv preprint arXiv:2411.15190, 2024.

[12] H. Gu, M. Schreyer, K. Moffitt and M. A. Vasarhelyi, "Artificial Intelligence Co-Piloted Auditing", International Journal of Accounting Information Systems, vol. 54, Art. no. 100698, 2024.

[13] A. Herreros-Martínez, R. Magdalena-Benedicto, J. Vila-Francés, A. J. Serrano-López and S. Pérez-Díaz, "Applied Machine Learning to Anomaly Detection in Enterprise Purchase Processes", arXiv preprint arXiv:2405.14754, 2024.

[14] T. Sattarov, D. Herurkar and J. Hees, "Explaining Anomalies Using Denoising Autoencoders for Financial Tabular Data", arXiv preprint arXiv:2209.10658, 2022.