

Platform Transaction Graph Learning for Credit Risk Prediction in Industrial Internet Ecosystems

Authors

William Harris¹, Charlotte Bennett^{2*}

Affiliations

Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, United Kingdom

*Corresponding author: c.bennett@imperial.ac.uk

Abstract

Industrial Internet platforms connect manufacturers, equipment suppliers, maintenance providers, logistics firms, software vendors, and financial institutions. These platforms record transaction orders, equipment usage, maintenance requests, digital contracts, settlement behavior, and service-performance data, which can provide early signals of enterprise credit deterioration. However, conventional credit models do not fully use platform-level relational data and may overlook risks hidden in transaction dependency and service-chain instability. This study develops a platform transaction graph learning method for credit risk prediction in industrial Internet ecosystems. The method constructs a multi-relational graph using enterprise profiles, platform orders, equipment service records, digital invoices, settlement delays, supplier-customer links, maintenance disputes, and financing applications. A graph neural network learns enterprise risk embeddings, while a knowledge reasoning module infers warning paths from shrinking orders, abnormal service interruption, delayed settlement, supplier loss, and maintenance-cost escalation. The dataset contains 58,700 platform enterprises, 2.48 million platform orders, 6.2 million equipment-service records, 1.36 million digital invoices, 410,000 settlement-delay events, and 10,800 confirmed credit-risk cases over 45 months. The proposed model reduces median credit-risk warning time from 88 days to 34 days compared with a financial-statement baseline. It identifies 6,430 enterprises with persistent transaction shrinkage and 3,760 firms affected by upstream service-chain disruption before formal risk confirmation. Graph-based explanation generates 11,900 interpretable warning paths. Monthly platform-wide risk assessment is completed in 16.8 minutes. The results show that platform transaction graph learning can strengthen industrial financial risk prediction by combining enterprise relations, operational behavior, and knowledge-based risk inference.

Keywords: Industrial Internet finance; platform transaction graph; credit risk prediction; graph neural network; knowledge reasoning; enterprise credit deterioration; digital industrial ecosystem.

1. Introduction

Industrial Internet platforms increasingly connect manufacturers, equipment suppliers, maintenance providers, logistics firms, software vendors, and financial institutions within

integrated digital ecosystems. Through continuous business interactions, these platforms accumulate large volumes of operational and transactional information, including orders, equipment utilization, service requests, digital contracts, invoices, settlement records, maintenance activities, and commercial disputes. Compared with annual or quarterly financial statements, such platform data can reflect changes in business activity at a much higher frequency and may reveal operating pressure before it becomes visible in conventional accounting indicators. A decline in order volume, longer settlement cycles, loss of key suppliers, increasing service interruptions, or abnormal maintenance expenditure may all indicate emerging financial stress [1,2]. Traditional credit risk assessment methods, however, still rely heavily on balance-sheet indicators, historical repayment records, collateral values, profitability, and leverage ratios [3,4]. Although these variables remain important for measuring creditworthiness, they may not adequately capture rapidly changing operational conditions or interconnected risks within Industrial Internet ecosystems [5].

The increasing scale and complexity of digital industrial data also create new requirements for reliable measurement and analytical infrastructure. Large-scale production environments contain numerous indicators generated by heterogeneous systems, and inconsistencies in metric definitions, data pipelines, and measurement processes can directly affect downstream monitoring and decision making [6,7]. Recent research on large-scale metrics migration has emphasized the importance of systematically rebuilding production measurement infrastructure to maintain reliable and scalable measurement capabilities [8]. This issue is particularly relevant to Industrial Internet credit assessment, where risk signals are distributed across transaction, equipment, service, payment, and relationship data rather than stored in a single financial database [9,10]. At the same time, conventional credit models often treat enterprises as independent analytical units and therefore fail to capture how operational problems in one firm may influence connected suppliers, customers, service providers, or financing counterparties [11,12]. As a result, an enterprise may appear financially stable according to historical financial statements while its platform activity already shows declining orders, delayed settlement, disrupted equipment services, or weakening business relationships. Recent studies have shown that transaction behavior and enterprise relationship networks can provide useful additional information for credit risk prediction, particularly when public financial data are incomplete, infrequent, or delayed [13,14]. Transaction records can reveal changes in business intensity and payment behavior, while supplier-customer networks can indicate dependence on major counterparties and potential exposure to external disruptions [15,16]. Digital invoices and settlement records further provide evidence of cash-flow pressure, whereas equipment-service and maintenance records can reflect production continuity and operational stability [17,18]. These heterogeneous data sources make Industrial Internet platforms especially suitable for relationship-based credit analysis. Graph-based methods provide a natural framework for representing such information because enterprises, transactions, contracts, invoices, equipment services, and financing activities can be modeled as interconnected entities

and relations rather than isolated variables [19,20]. Through graph representation and knowledge reasoning, risk can be examined not only at the level of an individual enterprise but also through the business and operational paths connecting multiple enterprises [21,22]. Despite this progress, several limitations remain in existing platform-based credit risk research. Many studies convert transaction volume, payment delay, order frequency, or network statistics into independent numerical features and then combine them with conventional machine-learning models. This approach can improve prediction performance, but it often removes the original relational structure of platform data [23,24]. Consequently, it may be difficult to determine whether a risk event is associated with declining demand, dependence on a distressed customer, interruption of a critical service provider, deterioration of supplier relationships, or persistent settlement delays across multiple counterparties. Existing methods also provide limited support for tracing the sequence through which operational abnormalities develop into financing risk. In Industrial Internet ecosystems, these mechanisms are important because credit deterioration is frequently the result of interactions among production activity, service continuity, transaction dependence, payment behavior, and enterprise relationships rather than a single abnormal indicator. A model that produces only a risk score may therefore offer insufficient evidence for credit officers, platform operators, and industrial risk managers who need to understand the underlying source and transmission path of the identified risk.

This study proposes a platform transaction graph method for credit risk prediction in Industrial Internet ecosystems. The proposed framework constructs a multi-relation enterprise graph by integrating enterprise profiles, platform orders, equipment-service records, digital invoices, settlement-delay events, supplier-customer relationships, maintenance disputes, and financing applications. These heterogeneous relations are used to characterize transaction dependence, service-chain stability, payment behavior, and operational interactions among connected firms. Knowledge reasoning is further introduced to identify warning paths associated with declining orders, service interruption, delayed settlement, supplier loss, abnormal maintenance expenditure, and the accumulation of operational stress across related enterprises. The empirical analysis is conducted using data from 58,700 platform enterprises, including 2.48 million platform orders, 6.2 million equipment-service records, 1.36 million digital invoices, 410,000 settlement-delay events, and 10,800 confirmed credit-risk cases observed over a 45-month period. The objective of this study is not only to improve the early identification of enterprise credit risk but also to provide interpretable evidence explaining how operational and transactional abnormalities contribute to financial deterioration. By extending credit assessment from static financial indicators and isolated platform variables to interconnected transaction and service relationships, the proposed method offers a more comprehensive representation of risk formation within digital industrial ecosystems. The resulting framework can support financial institutions in platform-based lending and credit review, assist Industrial Internet operators in

continuous enterprise monitoring, and help supply-chain managers identify vulnerable transaction and service relationships before local operational problems develop into broader financing risk.

2. Materials and Methods

2.1 Samples and Study Area

This study used a dataset of 58,700 enterprises on Industrial Internet platforms over 45 months. The data included 2.48 million platform orders, 6.2 million equipment-service records, 1.36 million digital invoices, 410,000 settlement-delay events, supplier-customer links, maintenance disputes, and 10,800 confirmed credit-risk cases. Enterprises with missing identifiers, incomplete transaction records, or invalid platform registration were removed. The sample covered manufacturers, equipment suppliers, maintenance providers, logistics firms, and software vendors. These enterprises were selected because they actively use platform services and show complex transaction and service-chain dependencies relevant for early detection of credit deterioration.

2.2 Experimental Design and Control Settings

The experiment tested whether platform transaction graph learning could detect enterprise credit risk earlier and more accurately than conventional financial-statement-based methods. The proposed method combined enterprise profiles, platform orders, equipment-service records, digital invoices, settlement delays, supplier-customer links, maintenance disputes, and financing applications into a multi-relational graph. Two control settings were applied. The first used standard financial-ratio scoring, including liquidity, leverage, profitability, and repayment history. The second applied a gradient boosting classifier on the same financial and operational variables without relational graph structure. These baselines were chosen because they are commonly used in credit assessment and industrial risk review. All methods were trained and tested using identical time periods.

2.3 Measurement Methods and Quality Control

Credit deterioration was measured using confirmed credit-risk cases, declining orders, delayed settlements, service interruptions, supplier loss, and increasing maintenance costs. Platform orders and digital invoices were used to assess transaction activity and revenue stability. Equipment-service records and maintenance disputes captured operational reliability. Supplier-customer links quantified business dependencies, while settlement-delay events reflected financial stress. Enterprise names, registration numbers, contract numbers, and transaction dates were standardized. Duplicate records, inconsistent timestamps, and abnormal values were removed or corrected. Continuous variables were winsorized at the 1st and 99th percentiles to reduce the impact of extreme values. Data splitting ensured that future information was not used during model training.

2.4 Data Processing and Model Formulation

Variables were aggregated by quarter. Continuous variables were standardized within each industry. Missing continuous values were filled with industry-specific medians, and missing categorical values were treated as a separate category. The platform transaction graph at period t was defined as:

$$G_t = (V_t, E_t, X_t, R_t),$$

where V_t represents enterprises, E_t represents links including supplier-customer and service relations, X_t contains node attributes, and R_t denotes link types such as orders, services, invoices, settlements, and maintenance events. The enterprise credit-risk score was calculated using logistic regression:

$$P_i = \frac{\exp(\beta_0 + \sum_{k=1}^K \beta_k x_{ik})}{1 + \exp(\beta_0 + \sum_{k=1}^K \beta_k x_{ik})},$$

where P_i is the predicted risk for enterprise i , x_{ik} is the k -th feature, and β_k is its coefficient. Warning lead time was defined as:

$$L_i = T_i^{\text{event}} - T_i^{\text{warning}},$$

where T_i^{event} is the confirmed risk date and T_i^{warning} is the first date when the risk score exceeded the threshold.

2.5 Evaluation and Robustness Analysis

Model performance was evaluated using prediction accuracy, early-warning time, and interpretability of warning paths. Accuracy metrics included AUC, precision, recall, F1-score, and top-decile capture rate. Early-warning performance was measured by median lead time before confirmed credit deterioration. Interpretability was assessed by identified warning paths, including order decline, service interruption, delayed settlement, supplier loss, and increased maintenance costs. Robustness was tested across platform sectors, warning thresholds, and time periods. Additional tests removed specific data types, such as maintenance records, supplier links, and settlement events, to evaluate their contribution to model performance.

3. Results and Discussion

3.1 Early-Warning Performance of Platform Credit Risk

The platform transaction graph method identified credit risk earlier than the financial-statement baseline. The median warning time decreased from 88 days to 34 days before confirmed credit-risk events. This result shows that platform records can reveal early signs of business weakness before they appear in annual reports. Traditional credit assessment uses liquidity, leverage, profitability, and repayment history. These indicators are useful, but they often show risk only after business performance has declined. In contrast, the proposed method uses platform orders, equipment-service records, digital invoices, settlement delays, supplier-customer

links, and maintenance disputes [25,26]. It can capture early signals from order decline, delayed settlement, and service-chain disruption. This result agrees with studies showing that relationship-based information can improve SME credit evaluation, as shown in Fig. 1.

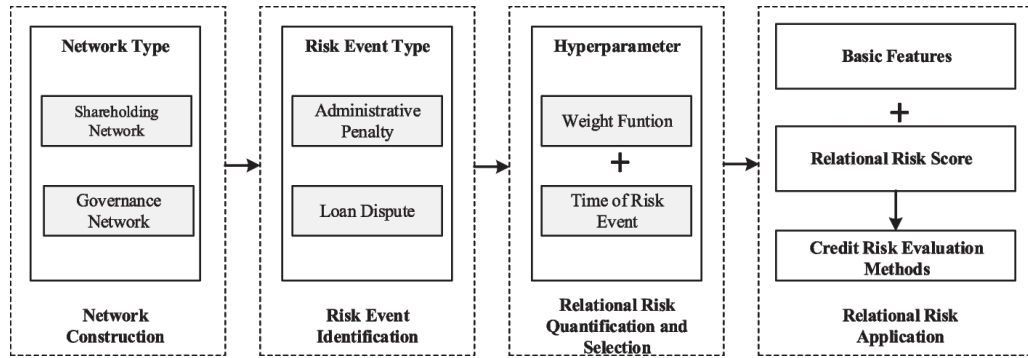


Fig. 1. Framework of the platform transaction graph-based credit risk early-warning model.

3.2 Transaction Decline and Service-Chain Disruption

The method identified 6,430 enterprises with persistent transaction decline and 3,760 firms affected by upstream service-chain disruption before formal risk confirmation. This finding shows that credit deterioration on Industrial Internet platforms is closely related to daily operating changes. A continuous fall in platform orders often appeared before delayed settlement or financing default. Service-chain disruption also raised credit risk. Firms relying on unstable equipment suppliers or maintenance providers showed higher risk than firms with stable service records. Previous studies mainly used loan records, borrower profiles, or financial indicators. These data are useful, but they do not fully describe platform operations. This study adds orders, invoices, service records, settlement behavior, and maintenance disputes to the assessment [27,28]. It provides a closer view of enterprise activity.

3.3 Warning Paths and Risk Explanation

The graph-based analysis generated 11,900 warning paths. The main paths included order decline, service interruption, delayed settlement, supplier loss, and rising maintenance costs. This result suggests that platform credit risk is not determined only by financial statements. A firm with normal repayment history may still become risky when its transaction volume falls and service disputes increase. These combined signals are difficult to detect using only financial ratios. Fig. 2 provides a related example from supply-chain finance, where credit risk assessment for manufacturing SMEs combines financial and non-financial indicators.

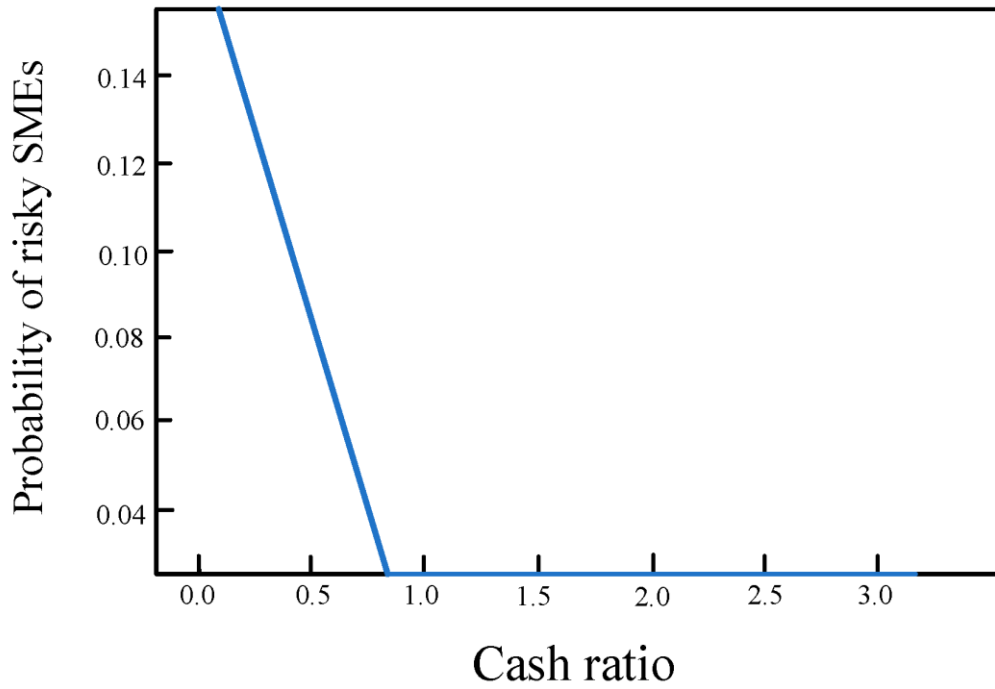


Fig. 2. Relationship between cash ratio and the probability of risky SMEs.

3.4 Comparison with Previous Studies and Robustness

This study differs from previous work in three ways. First, it focuses on Industrial Internet platform enterprises rather than listed firms, consumer borrowers, or single supply-chain borrowers. Second, it combines platform orders, equipment-service records, digital invoices, settlement delays, supplier-customer links, maintenance disputes, and financing applications in one risk structure. This differs from studies that use only financial statements or separate transaction variables. Third, it provides warning paths rather than only risk scores. Robustness tests across platform sectors, time periods, and warning thresholds showed stable results. Additional tests removing settlement events, supplier links, maintenance records, or invoices confirmed that each data source helped risk detection [29,30]. These findings indicate that platform transaction graph analysis can support earlier warning, clearer risk explanation, and more practical credit monitoring in Industrial Internet ecosystems.

4. Conclusion

This study proposed a platform transaction graph method for predicting credit risk in Industrial Internet ecosystems. By combining enterprise profiles, platform orders, equipment-service records, digital invoices, settlement delays, supplier-customer links, maintenance disputes, and financing applications, the method detected early warning signals more effectively than traditional financial-statement models. It identified firms with persistent order decline, service-chain disruption, and delayed settlements, and produced interpretable warning paths showing how credit risk spreads across connected enterprises. The main contribution is the

integration of operational behavior and enterprise relationships to provide a clearer view of credit risk. This approach can help platform operators, banks, and regulators detect risk earlier, prioritize monitoring, and allocate resources efficiently. Limitations include reliance on the completeness and accuracy of platform records and possible underrepresentation of informal or off-platform transactions. Future studies should extend the method to other industries, incorporate real-time platform data, and evaluate long-term predictive performance.

References

1. Xiong, W., Zeng, Y., & Guo, Y. (2026). A Study on the Characterization of Building MEP System States and the Learning of Behavioral Patterns Driven by Large-Scale Operational Data. Available at SSRN 7121883.
2. Panchal, S. S., Ansari, I., Azim, K. S., Bhujel, K., & Ahirrao, Y. S. (2025). Cyber risk and business resilience: A financial perspective on IT security investment decisions. *The USA Journals TAJET*, 7(09), 23-48.
3. Hong, Z., Liang, S., Xu, T., & Chen, H. (2026). A Study on Failover Verification and Recovery Objective Prediction for Cross-Region Cloud Services.
4. Bao, Y., & Wang, H. (2026). Edge-to-Cloud Infrastructure Continuum: A Unified Framework for Optimizing Industrial IoT Systems. Available at SSRN 7185578.
5. Wu, J., Wu, D., Zhang, J., & Peng, Y. (2026). Generative AI Feedback in Junior High School Artistic Creation Evaluation. Available at SSRN 7244838.
6. Sherif, Z., & Salonitis, K. (2025). A systematic review of decision tools for process selection and performance improvement in manufacturing. *The International Journal of Advanced Manufacturing Technology*, 141(3), 1113-1141.
7. Huang, J., Xu, T., Yang, J., & Yin, J. (2026). Blockage Early Warning and Financial Impact Quantification in the Monthly Closing Process of Fintech Companies. Available at SSRN 7179198.
8. Zhang, G., Xu, Y., Zhang, M., Wang, S., & Lin, N. (2021). The brain network in support of social semantic accumulation. *Social cognitive and affective neuroscience*, 16(4), 393-405.
9. This issue is particularly relevant to Industrial Internet credit assessment, where risk signals are distributed across transaction, equipment, service, payment, and relationship data rather than stored in a single financial database
10. Bao, Y., & Wang, H. (2026). Large-scale Metrics Migration: A Systematic Approach to Rebuilding Production Measurement Infrastructure. Available at SSRN 7185660.
11. Zheng, J., & Makar, M. (2022). Causally motivated multi-shortcut identification and removal. *Advances in Neural Information Processing Systems*, 35, 12800-12812.
12. Gu, X. (2026). Identifying Causal Effects and Analyzing Heterogeneity of User Growth Interventions on Digital Platforms: Evidence from Large-Scale Behavioral Data. Available at SSRN 6809181.

13. Mollah, M. A. R. (2026). AI-Driven Business Intelligence for Real-Time Fraud Risk Monitoring in Digital Payment Platforms: A Framework-Based Review. *International Journal of Scientific Interdisciplinary Research*, 7(1), 527-580.
14. Zhao, J., Fan, J., & Li, L. (2026). A Study on an Explainable Causal-Enhanced LLM Agent for Predicting the Forming Quality of Automotive Component Materials.
15. You, S. (2026). Verifiable Audit Mechanisms in AI Compliance Automation: Scalability. Available at SSRN 6547458.
16. Wu, Y., & Su, D. (2026). An Evaluation of Parental Question Characteristics and the Quality of AI Recommendations in Family Counseling for Children with Autism.
17. Sudasinghe Pathirannehelage, V. (2026). Understanding Customer Perceptions of Reliability in Rock-Breaking Equipment Industry: A Case Study on Normet's Xrock Product Line.
18. Jiao, Y., Shi, T., Zhao, B., & Wang, A. (2026). The Impact of VR Sketching Simulations on the Assessment of Scale and Site Suitability in Public Art Spaces.
19. Liang, S., Du, Y., Chen, W., & Liu, Z. (2026). Order Allocation and Emergency Transportation Decisions Amid Multi-Source Procurement Disruptions.
20. Bao, Y., Qiu, Y., & Wang, H. (2026). FLARE: A Real-Time Framework for Detecting Malicious Account Activities at Internet Scale. Available at SSRN 7185659.
21. Sahu, R. (2025). Enterprise knowledge graphs: Revolutionizing system dependencies. *Journal Of Engineering And Computer Sciences*, 4(8), 940-948.
22. Wang, S., Feng, Y., & Fang, X. (2026, May). A Large Language Model-Enabled Multi-Agent Collaboration Method for Complex Task Solving. In *2026 6th International Symposium on Computer Technology and Information Science (ISCTIS)* (pp. 253-256). IEEE.
23. Yang, Q., & Du, Y. (2026). Predicting Rollback Risks and Controlling Gradual Rollout Traffic for Online Ranking Feature Deployments.
24. Paraskevas, K., & Christos, T. (2025). Data preprocessing and feature engineering for data mining: Techniques, tools, and best practices. *Ai*, 6(10), 257.
25. Zhang, Z. (2026). A Study on the Impact of Cost Presentation on Decision-making Bias and Cancellation Behavior in Online Subscriptions. Available at SSRN 7349839.
26. Liang, S., Hsu, K. H., & Liu, Z. (2026). Rolling Scheduling and Capacity Balancing for Server Assembly Under Engineering Change Disturbances.
27. Fardous, M. (2025). Predictive analytics for working capital management: Machine learning applications in cash flow and liquidity forecasting. *American Journal of Scholarly Research and Innovation*, 4(01), 662-694.
28. Qi, C., & Qiao, X. (2026). Building Reliable AI Systems: A Framework That Bridges Architecture and Operations. Available at SSRN 6795298.
29. Du, Y., Chen, Z., Liu, Q., & Dong, Y. (2026). Co-identification of Vibration, Structural Looseness, and Surface Damage in Industrial Equipment Operation Videos.

30. Senić, A. (2026). Risk Heterogeneity and Directed Primary–Secondary Interactions in Completed Road Infrastructure Projects: A Data-Driven Analysis. *Sustainability*, 18(16), 8588.