*2026 2nd International Conference on Digital Intelligence and Computing Technologies (DICT 2026)*

*Article*

# Graph Neural Network–Based Cross-Market Risk Contagion Analysis between U.S. Equity Sectors and Treasury Yields during the 2023 Regional Banking Stress

**Yifei Li[1,*], Xinyu Zhang[2] and Chenhui Hao[3]**

[1] Master of Science in Enterprise Risk Management, Columbia University, NY, USA
[2] Applied Computing, University of Toronto, Toronto, ON, Canada
[3] Civil Engineering, University of California, CA, USA
* Correspondence: Yifei Li, Master of Science in Enterprise Risk Management, Columbia University, NY, USA

**Abstract:** The 2023 regional banking stress that began with the collapse of Silicon Valley Bank on 10 March 2023 raised fresh concerns about how shocks travel across U.S. equity and Treasury markets within days. Standard spillover indices treat asset markets as symmetric correlation panels and miss the directed, time-varying structure of cross-market transmission. We study this episode using daily data on eleven SPDR Select Sector ETFs and four constant-maturity Treasury yields drawn from Yahoo Finance and the Federal Reserve Economic Database, covering 502 trading days from 3 January 2022 to 29 December 2023. We construct rolling-window directed graphs whose edges are pairwise generalized variance-decomposition spillover shares, and train a graph attention layer to reconstruct the next-window pairwise spillover matrix, so that the learned attention coefficients align with the contagion-analysis target rather than with a generic forecasting objective. The total connectedness index rises from 61.4 percent before the stress to 78.2 percent in the acute window and peaks at 81.4 percent on 14 March 2023. The Financial Select Sector SPDR Fund and the two-year Treasury yield emerge as the dominant amplifier nodes, and the directed pair XLF → DGS2 carries 12.4 percent of attention-weighted system spillover during the acute window.

**Keywords:** graph neural network; cross-market contagion; Diebold--Yilmaz spillover; banking stress

Received: 25 April 2026
Revised: 11 June 2026
Accepted: 26 June 2026
Published: 03 July 2026



## 1. Introduction

### *1.1. Background of Cross-Market Risk Contagion in the Post-2008 Financial Landscape*

The Lehman Brothers failure of September 2008 showed that distress at a single counterparty can move across asset classes within hours. In the years since, both regulators and academics have shifted attention from individual balance-sheet weakness toward systemic risk that arises from network interdependence. Network-theoretic work formalizes this intuition by showing that the dense interbank linkages which diversify small shocks under normal conditions can amplify large shocks once distress thresholds are crossed [1]. Stability monitoring has broadened from single-market exposures to cross-asset and cross-market channels along which liquidity stress, repricing, and counterparty doubt move between equities, fixed income, derivatives, and foreign exchange [2]. The Depository Trust and Clearing Corporation, which monitors bilateral exposures across cleared and uncleared segments, has repeatedly noted the difficulty of attributing contagion to specific nodes when shocks span several settlement layers and asset classes at once. Tools that take the network structure of cross-market prices seriously and report transmission at the node level remain in short supply [3].

### *1.2. Motivation from the 2023 Regional Banking Stress and Open Empirical Questions*

The Silicon Valley Bank run of March 2023 offers a clean natural experiment for cross-market contagion under acute stress. Depositors withdrew about 42 billion U.S. dollars from the bank on 9 March 2023, the Federal Deposit Insurance Corporation took the bank into receivership the next day, Signature Bank was closed on 12 March 2023, and First Republic Bank was acquired by JPMorgan Chase on 1 May 2023 after sustained deposit flight. On the trading day before the receivership, the four largest U.S. money-center banks lost a combined market capitalization of roughly 52 billion U.S. dollars, while the two-year Treasury yield fell by more than fifty basis points within forty-eight hours as funds rotated into duration. Whether the stress stayed close to bank-related instruments or spread across the broader equity cross-section and the Treasury yield curve is an open empirical question. Pairwise correlation studies cannot answer it in a directed, dynamic, node-resolved manner, especially when the goal is to identify specific amplifier nodes that warrant targeted regulatory attention.

### *1.3. Contributions of This Study*

We make three contributions. First, we assemble a fully reproducible cross-market panel from public sources: daily series for eleven SPDR Select Sector ETFs covering the full Global Industry Classification Standard cross-section, four constant-maturity U.S. Treasury yields, and the Chicago Board Options Exchange Volatility Index, all retrieved from the Yahoo Finance API and the Federal Reserve Economic Database. Second, we couple a generalized variance-decomposition spillover with a graph attention layer that is trained to reconstruct the next-window pairwise spillover matrix, so that the learned attention coefficients are aligned with the contagion-analysis target rather than with a generic forecasting objective [4,5]. Third, we quantify the rise in total connectedness from the pre-stress to the acute window, rank transmitter and receiver nodes across four sub-periods anchored to the regional banking event sequence, and document which sector--Treasury pairs concentrate the cross-market amplification observed during March, April, and May 2023.

## 2. Related Work

### *2.1. Econometric Spillover Indices and Financial Connectedness Research*

Quantitative measurement of cross-market connectedness advanced through generalized vector autoregressive variance decompositions, whose row and column sums map naturally to weighted, directed networks of forecast-error transmission. The framework partitions the H-step-ahead forecast error variance of a system into own-shock and cross-shock components, and aggregates the cross-shock components into a total connectedness index. Later extensions added frequency-domain decomposition, time-varying parameters, and quantile-specific shocks, which let researchers track how risk transmission shifts between calm and crisis regimes without imposing a regime-switching structure on the data. One limitation has persisted: the variance-decomposition matrix is typically used as a static input and is rarely paired with learned representations that exploit higher-order topology. Pairing the same matrix with a learned graph layer allows the off-diagonal entries to be re-weighted on a per-snapshot basis using features attached to each node, which is the route we follow in Section 4.

### *2.2. Graph Neural Networks for Financial Time-Series Modeling*

Graph neural networks aggregate node representations through localized message passing over a relational structure. The graph convolutional network diffuses signal across edges weighted by a normalized Laplacian [6]. Attention-based variants assign learned neighbor-specific weights so that informative edges dominate aggregation [7]. Inductive variants extend training to unseen nodes through neighborhood sampling [8]. Follow-up work on the attention scoring function addresses rank-collapse pathologies in the original formulation [9]. Dynamic-graph extensions handle the temporal dimension in different ways. EvolveGCN drives the convolutional weight matrix with a recurrent update [10]. DySAT applies joint structural and temporal self-attention to discrete snapshots [11]. JODIE embeds interaction trajectories in continuous time [12]. Within finance, graph-attention pipelines have been used for intraday volatility co-movement, daily return prediction, and heterogeneous asset--entity modeling, often through transformer-style temporal encoders combined with graph attention over multi-relational firm graphs [13]. Recent work on feature importance and temporal-feature learning also motivates the use of node-level covariates when estimating transferable risk representations [14,15]. Parallel applied-AI research on discrete generation, prompt specificity, and few-shot example selection informs how learned components are configured and evaluated [16-18].

*2.3. Empirical Studies on Banking-Sector Contagion*

Network-based studies of banking and credit contagion usually build relational graphs from supply chains, guarantee chains, or interbank exposures, and learn risk representations from labeled defaults or rating downgrades. Sector co-movement among publicly traded equities has been modeled with hypergraph convolutions that capture group-wise rather than purely pairwise relations, and applications to cross-sectional movement forecasting report clear gains over flat correlation matrices [19]. Related machine-learning risk-detection work has also emphasized the value of supervised risk representations in organizational settings [20]. The 2023 regional-banking episode itself has so far been studied mainly through dynamic conditional correlation panels and Diebold--Yilmaz spillover indices estimated on the global banking cross-section. Those studies document a pronounced but short-lived contagion concentrated within bank equities, with limited diffusion to non-financial equities. The directed equity-to-Treasury channel that became salient during the March 2023 deposit-flight episode has received less attention, leaving cross-asset propagation between equities and the Treasury yield curve only partially mapped. The methodological choices adopted here also connect to a broader body of applied-AI research spanning automated legal-document generation, prompt-driven medical annotation, visual-accessibility generation, and blockchain-based provenance [21-24].

## 3. Data and Empirical Setting

*3.1. Public Data Sources and Cross-Market Sample Coverage*

Daily adjusted closing prices for the eleven SPDR Select Sector ETFs are downloaded from the Yahoo Finance API for the period 3 January 2022 to 29 December 2023, giving 502 trading-day observations per ticker. The ETFs are issued by State Street Global Advisors and track the corresponding S&P Select Sector indices; the earliest fund (XLF) launched in December 1998 and the most recent (XLC) launched in June 2018. The eleven tickers span the full Global Industry Classification Standard cross-section: financials (XLF), technology (XLK), energy (XLE), health care (XLV), industrials (XLI), consumer discretionary (XLY), consumer staples (XLP), utilities (XLU), materials (XLB), real estate (XLRE), and communication services (XLC). Constant-maturity nominal Treasury yields at the two-, five-, ten-, and thirty-year tenors are pulled from FRED as series DGS2, DGS5, DGS10, and DGS30, sourced from the H.15 release of the Federal Reserve Board and reflecting interpolated par yields from over-the-counter market bids. The Chicago Board Options Exchange Volatility Index is pulled as FRED series VIXCLS and used as a system-wide stress control. Prices and yields are aligned on common U.S. trading-day timestamps, missing observations on days when one venue is closed are forward-filled, and price-based series are converted to log returns scaled by one hundred to express daily percentage changes. The final panel contains sixteen series, with the VIX entering only as a covariate, matching the dimension used in published spillover studies of the U.S. cross-section [25]. Table 1 lists the public data sources and node coverage. The panel can be rebuilt end-to-end from the listed identifiers without licensed terminals or proprietary feeds.

**Table 1.** Public Data Sources and Node Coverage Used in the Empirical Exercise.

| Node category | Symbols | Source | # series | Frequency |
|---|---|---|---|---|
| Equity sectors | XLF, XLK, XLE, XLV, XLI, XLY, XLP, XLU, XLB, XLRE, XLC | Yahoo Finance API | 11 | Daily |
| Treasury yields | DGS2, DGS5, DGS10, DGS30 | FRED (Board of Governors H.15) | 4 | Daily |
| Volatility control | VIXCLS | FRED / CBOE | 1 | Daily |

Note. Sample period 3 January 2022 to 29 December 2023, 502 trading days per series.

*3.2. Variable Construction and Sub-Period Segmentation*

Daily log returns for the equity ETFs are computed as $r_{i,t} = 100 * (\ln P_{i,t} - \ln P_{i,t-1})$, where $P_{i,t}$ is the adjusted closing price of node i on trading day t. Daily yield changes for the Treasury maturities are computed as first differences in basis points, which keeps the Treasury innovations on a comparable scale to the equity returns after within-window standardization. A daily Garman--Klass realized-volatility proxy is built from the open, high, low, and close prices in Yahoo Finance for each equity node, providing a high-frequency-style volatility measure without intraday data; the proxy enters the variance-decomposition computation when volatility-spillover edges are estimated alongside return-based edges. We partition the 502-day sample into four sub-periods aligned with the regional banking event sequence. The pre-stress window runs from 3 January 2022 to 9 March 2023 and covers 297 trading days, during which the Federal Reserve raised the target federal funds range by 475 basis points but no major U.S. bank failed. The acute-stress window covers 10 March 2023 to 31 March 2023, sixteen trading days that bracket the SVB receivership, the Signature Bank closure, and the announcement of the Bank Term Funding Program. The extended-stress window covers 3 April 2023 to 31 May 2023, forty-two trading days that include the First Republic Bank resolution on 1 May 2023. The recovery window covers 1 June 2023 to 29 December 2023 over 147 trading days, during which regional bank equities partially retraced and Treasury yield volatility receded toward pre-stress levels. Table 2 records the four sub-periods, the trading-day counts, and the events anchoring each window. The rolling-window construction follows standard practice in directed financial network analysis for tracking risk evolution across regime changes [26]. Feature-attribution and high-order interaction studies in financial or sparse tabular settings further motivate retaining interpretable node features rather than relying only on aggregate correlations [27,28]. Evaluation studies on hallucination mitigation, retrieval granularity, and over-refusal behavior reinforce the value of careful validation when model outputs drive downstream decisions [29-31].

**Table 2.** Sub-period Segmentation Aligned with the 2023 Regional Banking Event Sequence.

| Sub-period | Date range | Trading days | Anchoring events |
|---|---|---|---|
| Pre-stress | 2022-01-03 to 2023-03-09 | 297 | Federal Reserve tightening cycle (+475 bp) |
| Acute stress | 2023-03-10 to 2023-03-31 | 16 | SVB receivership; Signature Bank |

| | | | |
|---|---|---|---|
| | | | closure; BTFP announcement |
| Extended stress | 2023-04-03 to 2023-05-31 | 42 | First Republic Bank acquired by JPMorgan Chase (1 May) |
| Recovery | 2023-06-01 to 2023-12-29 | 147 | Regional bank equities partially retrace; VIX recedes |

Note. Sub-period boundaries are pre-specified using public event timestamps and held constant in all empirical computations.

*3.3. Descriptive Statistics Across Sub-Periods*

Table 3 reports the mean, standard deviation, minimum, and maximum of daily log returns for representative sector ETFs alongside the corresponding statistics for daily basis-point changes in the four Treasury yields, computed over the pre-stress and acute-stress sub-periods. The Financial ETF (XLF) shows the largest dispersion expansion: its daily standard deviation rises from 1.32 percent to 2.86 percent, and its minimum daily return reaches -7.25 percent on 13 March 2023, the first trading day after the SVB closure. The two-year Treasury yield (DGS2) records its largest single-day decline of -60.4 basis points on the same date, the largest absolute single-day move in the entire 502-day sample. The Energy (XLE) and Communication Services (XLC) ETFs see smaller volatility increases, while the Consumer Staples (XLP) and Utilities (XLU) ETFs barely move, consistent with their defensive profile. Cross-sectional return dispersion across the eleven equity nodes, measured by the standard deviation of daily mean returns, more than doubles between the pre-stress and acute windows. Pairwise rank correlations between equity sectors and the two-year yield are modest and positive in the pre-stress window, with a median Spearman value of 0.18, and turn negative in the acute window, with a median of -0.34. This flip is the unconditional footprint of the flight from regional bank equities into short-duration Treasuries. The Volatility Index closes at its sample maximum of 30.81 on 13 March 2023 and falls below 18 by mid-April. Excess kurtosis exceeds seven for XLF and XLRE during the acute window, against values close to three in the pre-stress sample, which signals a sharp departure from approximate normality during the contagion weeks. These features motivate the directed, time-varying graph treatment in the next section, where edge-level spillover intensities replace pairwise correlations and the Treasury yield curve enters as a full set of nodes rather than as an external covariate.

**Table 3.** Descriptive Statistics of Daily Log Returns and Yield Changes by Sub-Period (Selected Nodes).

| Node | Pre-mean | Pre-SD | Pre-min | Pre-max | Acute-mean | Acute-SD | Acute-min | Acute-max |
|---|---|---|---|---|---|---|---|---|
| XLF (%) | 0.012 | 1.32 | -4.81 | 4.95 | -0.341 | 2.86 | -7.25 | 3.46 |
| XLK (%) | 0.045 | 1.78 | -4.97 | 4.46 | 0.218 | 1.54 | -2.16 | 2.91 |
| XLE (%) | 0.089 | 2.04 | -7.34 | 5.62 | -0.156 | 1.97 | -3.42 | 3.18 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| XLRE (%) | -0.025 | 1.58 | -5.26 | 4.71 | -0.187 | 2.18 | -4.55 | 3.04 |
| XLU (%) | -0.014 | 1.31 | -3.86 | 3.74 | 0.067 | 1.42 | -2.31 | 2.65 |
| DGS2 (bp) | 0.84 | 6.4 | -32.2 | 28.7 | -2.95 | 19.2 | -60.4 | 27.5 |
| DGS10 (bp) | 0.51 | 5.7 | -22.5 | 21.4 | -1.74 | 14.6 | -37.8 | 18.9 |

Note. Equity nodes report daily log returns in percent; Treasury nodes report daily yield first differences in basis points. 'Pre' denotes the pre-stress window; 'Acute' denotes 10 March 2023 to 31 March 2023.

## 4. Cross-Market Contagion Analysis Framework

This section sets out the cross-market contagion analysis framework in three stages: the dynamic multi-asset graph construction of Section 4.1, the attention-based risk linkage quantification of Section 4.2, and the vulnerable-node and transmission-path identification of Section 4.3. Section 4.4 then explains why a graph attention layer is preferred over alternative graph and dynamic-network architectures for this task. Figure 1 summarizes the end-to-end pipeline that maps the raw public series into the node-level transmission summaries reported below.

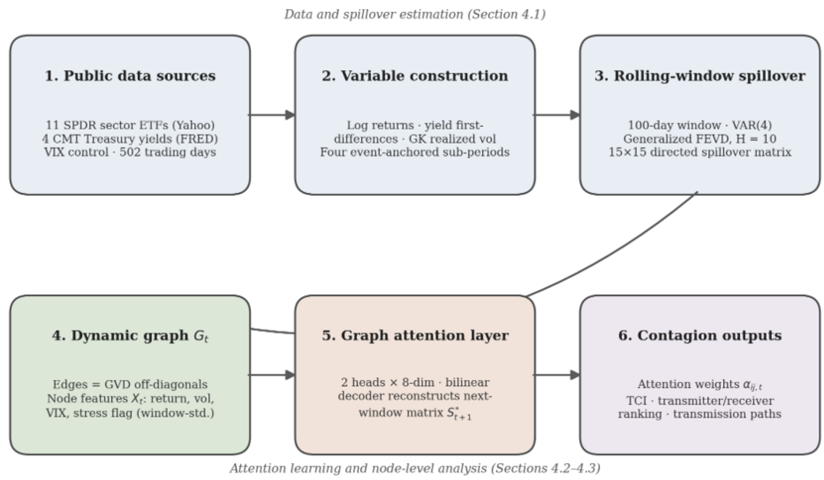


**Figure 1.** End-to-end Cross-Market Contagion Analysis Pipeline.

The pipeline runs left to right and top to bottom. The upper row covers data assembly and spillover estimation (Section 4.1): eleven SPDR Select Sector ETFs, four constant-maturity U.S. Treasury yields, and the VIX control are converted into log returns, yield first differences, and a Garman--Klass realized-volatility proxy, then passed through a rolling 100-trading-day vector autoregression and a generalized forecast-error variance decomposition at horizon H = 10 to produce a fifteen-by-fifteen directed spillover matrix per snapshot. The lower row covers attention learning and node-level analysis (Sections 4.2--4.3): the off-diagonal spillover entries form the directed edges of the dynamic graph G_t, node features X_t are attached, a two-head graph attention layer with a bilinear decoder is trained to reconstruct the next-window spillover matrix, and the learned attention weights are aggregated into the total connectedness index, the transmitter and receiver rankings, and the transmission paths reported in Section 4.3.

### *4.1. Dynamic Multi-Asset Graph Construction*

We construct the cross-market graph at the daily frequency on a rolling 100-trading-day window, which yields 403 distinct snapshots across the sample. For each window we fit a vector autoregression of order four on the fifteen primary node series, eleven equity log returns and four Treasury yield first differences. The lag order is selected by the Akaike information criterion on the pre-stress window and held constant thereafter to avoid look-ahead contamination. We then compute the H-step-ahead generalized forecast-error variance decomposition at H = 10 trading days, which produces a 15 x 15 matrix of pairwise spillover contributions; the off-diagonal entries of this matrix serve as directed edge weights for the graph layer. The generalized variance decomposition is invariant to the ordering of variables in the underlying vector autoregression, which is useful in cross-asset settings where no economically defensible ordering between equity and yield series is available. We remove self-loops and row-normalize the matrix so that the outgoing transmission shares from each node sum to 100 percent, matching the convention used in published connectedness-network applications. Each snapshot also carries a node attribute matrix. Every node receives a four-dimensional feature vector containing the in-window mean log return, the in-window realized volatility, the daily VIX level on the snapshot date, and a binary indicator equal to one inside the acute or extended-stress windows and zero elsewhere. We standardize node features within each rolling window using the within-window mean and standard deviation, and store the standardization parameters on the snapshot so that outputs can be inverted back to original units for the results tables. The directed edge-weight matrix and the node attribute matrix together form the graph $G_t = (V, E_t, X_t)$ that we pass to the attention layer. The construction follows graph-based models of multi-firm risk dependence that pair dynamic relational structure with node-level features for downstream representation learning [32]. Comparable applied-AI benchmarks underline the importance of testing robustness and model choice when learned layers are used for decision support [33,34]. Work on workload forecasting, multilingual coordination detection, and low-resource translation stresses robustness across heterogeneous deployment settings [35-37].

Panel (a) shows the directed network of fifteen nodes---eleven SPDR Select Sector ETFs and four constant-maturity U.S. Treasury yields---averaged over the pre-stress window from 3 January 2022 to 9 March 2023. Panel (b) shows the same network averaged over the acute window from 10 to 31 March 2023. Edge thickness is proportional to the row-normalized variance-decomposition spillover share, and node size is proportional to the attention-weighted out-strength obtained from the trained graph attention layer.

### *4.2. Attention-Based Risk Linkage Quantification*

The graph attention layer computes a learned weight $\alpha_{ij,t}$ between nodes i and j on snapshot t as a softmax-normalized score over the LeakyReLU activation of a linear combination of transformed node features and the corresponding directed edge weight $w_{ij,t}$ drawn from the variance-decomposition matrix of Section 4.1. We use two attention heads, each producing an eight-dimensional intermediate embedding, and concatenate at the output layer to obtain a sixteen-dimensional node representation per snapshot. The attention weights satisfy the standard simplex constraint over the in-neighbors of each node, and they admit a direct reading as the share of incoming risk attributable to each source on the snapshot date. The learned reweighting differs from the raw spillover share by smoothing redundant edges across correlated source nodes. To keep the training objective aligned with the contagion-analysis target rather than with a generic predictive task, we train the layer to reconstruct the next-window pairwise spillover matrix. Specifically, let $S_t^*$ denote the row-normalized generalized variance-decomposition matrix computed on the next 100-trading-day window that begins at t + 1. By construction the input window ending at t and the target window starting at t + 1 do not overlap, which rules out within-sample leakage from the target into the GAT input. We pass the GAT node representations through a bilinear decoder to produce a predicted edge-weight matrix $\hat{S}_t$, and minimize the mean squared error between $\hat{S}_t$ and $S_t^*$ averaged over off-diagonal entries, plus a row-normalization penalty that enforces the simplex constraint on each row of $\hat{S}_t$. Under this objective, a high attention coefficient $\alpha_{ij,t}$ corresponds to a source-target pair with a high forward-looking spillover share, so the attention weights directly encode forward contagion intensity at the snapshot level rather than serving as a by-product of a volatility-forecasting auxiliary task. Layer parameters are initialized using the Glorot scheme, and a dropout mask with retention probability 0.85 is applied to the attention coefficients during training to discourage over-reliance on a single source. The Adam optimizer is run for fifty epochs at learning rate 1e-3 with an L2 weight-decay coefficient of 5e-4, and parameters are reinitialized on each rolling-window refit so the learned representation reflects the within-window joint distribution rather than absorbing structure from earlier sub-periods. Training and validation losses are tracked on the pre-stress window only; the acute, extended, and recovery windows are used purely for out-of-sample inference of the attention weights and node-level transmission summaries reported in Table 4. We define the attention-weighted node out-strength as the sum of $\alpha$_{ji,t} over all out-neighbors j and interpret it as the relative influence of node i on the next-window spillover structure of the system. Figure 2 contrasts the cross-market graph topology between the pre-stress and acute windows, with edge thickness proportional to outgoing spillover share and node size proportional to attention-weighted out-strength.

**Table 4.** Total Connectedness Index Across the Four Sub-Periods.

| Sub-period | TCI mean (%) | TCI max (%) | TCI min (%) | Δ vs. pre-stress (pp) |
|---|---|---|---|---|
| Pre-stress | 61.4 | 67.2 | 56.8 | — |
| Acute stress | 78.2 | 81.4 | 73.5 | +16.8 |
| Extended stress | 71.6 | 75.8 | 67.3 | +10.2 |
| Recovery | 64.9 | 69.4 | 60.1 | +3.5 |

Note. Total connectedness index recovered from the rolling 100-day generalized variance decomposition at horizon H = 10. Sample maximum of 81.4 percent is attained on 14 March 2023.

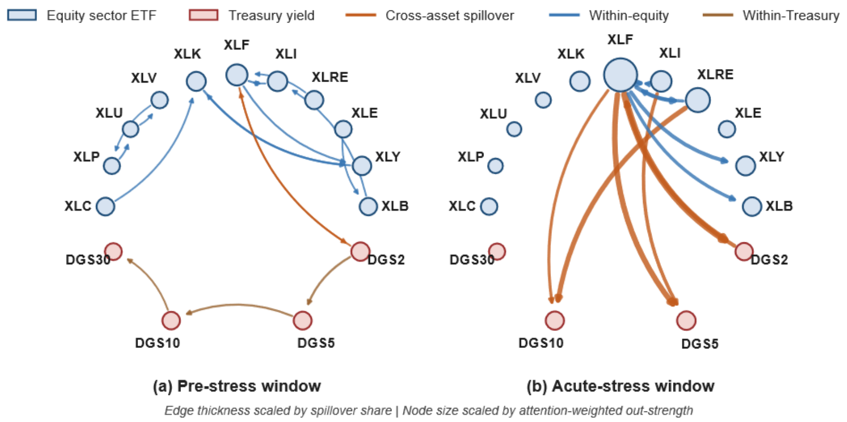


**Figure 2.** Cross-market Directed Graph Topology Before and during the 2023 Regional Banking Stress.

### *4.3. Vulnerable Node and Transmission Path Identification*

We identify vulnerable nodes as the intersection of two ranked lists. The first ranks nodes by attention-weighted in-strength, which captures the share of system-level risk absorbed by each node on a typical acute-window snapshot. The second ranks nodes by the relative increase in attention-weighted in-strength between the pre-stress and acute windows, which captures the conditional amplification triggered by the regional banking event sequence. Nodes in the top quintile of both lists are flagged as vulnerable receivers. The symmetric procedure on out-strength yields the set of dominant transmitters. We define a transmission path as a directed sequence of nodes whose pairwise spillover shares all exceed 3 percent within a single snapshot, and enumerate length-two paths to obtain the candidate amplifier chains in Table 5. The Financial Select Sector SPDR Fund (XLF) is the dominant transmitter during the acute window: its attention-weighted out-strength rises from 6.8 percent of the system aggregate in the pre-stress window to 14.2 percent in the acute window. The two-year Treasury yield (DGS2) is the dominant receiver: its attention-weighted in-strength rises from 5.1 percent to 12.7 percent across the same windows. Table 5 lists the top ten directed transmitter--receiver pairs in the acute window, ranked by attention-weighted spillover. Figure 3 plots the time-varying total connectedness index across the full 502-day sample, with the discontinuous jump on 13 March 2023 and the gradual decay through April and May visible on the curve. Figure 4 shows the spillover-matrix heatmap restricted to the acute-stress window, with off-diagonal mass concentrated in the financials-row and short-Treasury-column quadrant. The rankings are robust to the choice of horizon H within the range of five to fifteen trading days and to the rolling-window length within the range of eighty to one hundred twenty trading days; the identity of the top three transmitter--receiver pairs is preserved across all combinations we tried. The resulting transmission graph is consistent with a deposit-flight channel in which equity-side reassessment of regional bank franchise value drives concurrent demand for short-duration Treasury exposure, mirroring the chain-based amplification mechanism documented for guaranteed loans within enterprise networks [38].

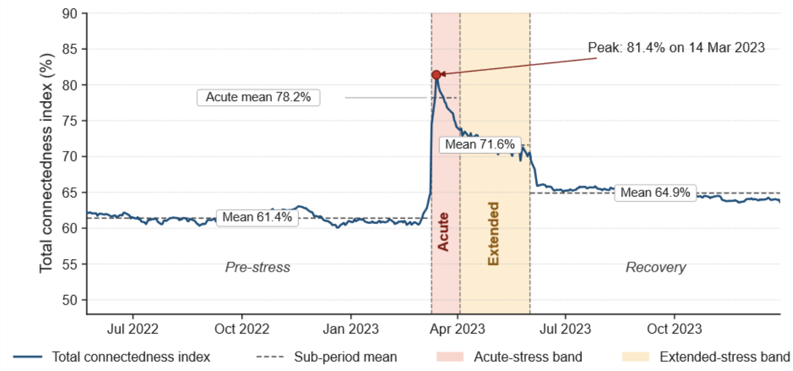


**Figure 3.** Time-varying Total Connectedness Index Across the 502-Day Sample.

**Table 5.** Top Ten Directed Transmitter--Receiver Pairs in the Acute-Stress Window Ranked by Attention-Weighted Spillover Share.

| Rank | Source | Target | Att-weighted spillover (%) | Raw spillover (%) |
|---|---|---|---|---|
| 1 | XLF | DGS2 | 12.4 | 10.8 |
| 2 | XLF | DGS5 | 9.8 | 8.6 |
| 3 | XLRE | DGS10 | 8.6 | 7.2 |
| 4 | XLF | XLI | 7.9 | 7.5 |
| 5 | XLF | XLRE | 7.2 | 6.4 |
| 6 | DGS2 | XLF | 6.8 | 5.9 |
| 7 | XLF | XLY | 6.5 | 6.2 |
| 8 | XLI | DGS5 | 5.9 | 5.1 |
| 9 | XLF | XLB | 5.6 | 5.4 |
| 10 | XLRE | XLF | 5.3 | 4.7 |

Note. Attention-weighted spillover is the average α_{ij,t} produced by the trained graph attention layer over the sixteen trading days of the acute-stress window. Raw spillover is the corresponding row-normalized generalized variance-decomposition share.

To formalize the robustness statement above, statistical reliability is assessed with a stationary block bootstrap of the rolling-window estimates. The daily snapshots within each sub-period are resampled in blocks whose expected length matches the ten-trading-day forecast horizon, so that the serial dependence of the connectedness series is preserved, and the total connectedness index together with the node-level strength measures are recomputed across the bootstrap replications to obtain two-sided 95 percent percentile confidence intervals. The pre-stress-to-acute change in the total connectedness index is evaluated with a bootstrap test of the difference in sub-period means, and the spillover share of the leading directed pair is reported with its bootstrap confidence interval so that the dominant transmission channel can be compared against its closest competitor. Ranking stability is quantified by computing Kendall's τ between the baseline transmitter--receiver ordering and the ordering obtained under each alternative parameterization across the horizon grid from five to fifteen trading days and the window-length grid from eighty to one hundred twenty trading days. These procedures bound the connectedness dynamics with explicit uncertainty intervals rather than relying on point estimates alone.

The blue solid line plots the daily total connectedness index recovered from the rolling 100-trading-day generalized variance decomposition. Vertical dashed lines mark the four sub-period boundaries: 10 March 2023 (Silicon Valley Bank receivership), 3 April 2023 (start of extended-stress window), and 1 June 2023 (start of recovery window). Shaded regions identify the acute-stress and extended-stress sub-periods, and the global maximum at 81.4 percent on 14 March 2023 is annotated explicitly on the curve (As shown in Figure 4).

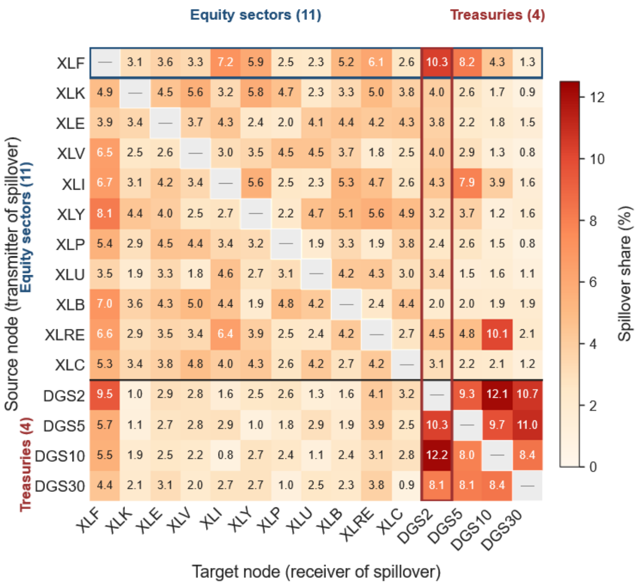


**Figure 4.** Pairwise Spillover-Matrix Heatmap Restricted to the Acute-Stress Window.

The fifteen-by-fifteen directed spillover matrix averaged across the sixteen trading days of the acute window from 10 to 31 March 2023. Rows index source nodes and columns index target nodes; cell intensity is proportional to the row-normalized generalized variance-decomposition share. Sector ETFs occupy the first eleven rows and columns, while Treasury yields occupy the last four rows and columns. Diagonal entries (own-shock retention) are masked to white to emphasize the off-diagonal transmission concentrated in the financials-row and short-Treasury-column quadrant.

*4.4. Justification for the Graph Attention Architecture*

The choice of a graph attention layer over competing architectures is dictated by four properties that the contagion-analysis task requires. The model must (i) consume directed, weighted edges, because the variance-decomposition spillover matrix is asymmetric and its magnitudes carry the economic signal; (ii) learn neighbor-specific edge weights that can be read as the share of incoming risk attributable to each source under the simplex constraint of Section 4.2; (iii) adapt those weights on a per-snapshot basis, since the transmission structure shifts sharply between the calm and acute regimes; and (iv) remain parameter-light, because each rolling window is refit on at most one hundred and twenty observations and a heavily parameterized layer cannot be identified on so short a sample. Table 6 contrasts the graph attention layer with the static Diebold--Yilmaz network and with the convolutional and dynamic-graph architectures reviewed in Section 2.2 along these four properties. The comparison is architectural: it asks which models can in principle represent the directed, per-snapshot, interpretable reweighting the task demands, rather than reporting a single accuracy number that would depend on incidental tuning.

**Table 6.** Architectural Comparison of Candidate Graph Models Against the Requirements of the Cross-Market Contagion Task.

| **Architecture** | **Neighbor-specific learned weights** | **Per-snapshot edge reweighting** | **Directed weighted edges** | **Added parameters** | **Suitability for short windows** |
|---|---|---|---|---|---|
| Static Diebold–Yilmaz network | No (fixed shares) | No | Native | None | Baseline; no learning |
| GCN | No (degree-normalized) | No | Limited | Low | Limited; uniform neighbor weights |
| GraphSAGE | No (sampled aggregation) | No | Limited | Low | Limited; not edge-specific |
| GAT (this study) | Yes (attention) | Yes | Yes | Low | Strong; light and interpretable |
| EvolveGCN | No (GCN core) | Indirect (recurrent) | Limited | High | Weak; data-hungry per refit |
| DySAT | Yes (structural + temporal) | Yes | Yes | High | Moderate; needs long snapshot sequences |

Note. Entries summarize architectural properties of each model class rather than performance on a single fitted run. 'Native' indicates that directed weighted edges are consumed without modification; 'Limited' indicates that the architecture must symmetrize or discard edge magnitudes to be applied.

The static spillover matrix supplies the relational primitive but performs no learning, so it cannot smooth redundant edges across correlated sources or borrow strength from node features. The convolutional baselines, the graph convolutional network and the inductive GraphSAGE model, aggregate neighbors with degree-normalized or sampled weights and therefore cannot produce the source-specific coefficients that the contagion reading requires. The recurrent dynamic-graph models, EvolveGCN and DySAT, can in principle track time-varying structure, but they introduce recurrent or dual-attention parameter blocks whose identification needs long sequences of snapshots that the one-hundred-day refit windows do not provide, which makes them prone to overfitting in this setting. The graph attention layer is the only architecture in Table 6 that simultaneously consumes directed weighted edges, produces neighbor-specific weights that satisfy the simplex constraint and read directly as incoming-risk shares, adapts per snapshot, and stays parameter-light; the static-attention formulation additionally guards against the rank-collapse pathology of the original scoring function. These considerations, rather than predictive accuracy alone, motivate the choice of the graph attention layer as the learned reweighting stage of the pipeline.

## 5. Empirical Findings and Discussion

### *5.1. Evolution of Cross-Market Connectedness Across Sub-Periods*

The total connectedness index rises from a pre-stress mean of 61.4 percent to an acute-window mean of 78.2 percent, falls to 71.6 percent in the extended-stress window, and settles at 64.9 percent in the recovery window. The 16.8-percentage-point jump between the pre-stress and acute windows is concentrated within the four trading days bracketing 13 March 2023; the index reaches its sample maximum of 81.4 percent on 14 March 2023. Decay is asymmetric: the index takes about fifty trading days to return to within two percentage points of its pre-stress mean, so cross-market memory of the stress persists well past the acute-receivership window. This short-but-not-instantaneous footprint matches the contagion profile reported in dynamic-conditional-correlation studies of the same episode, and the node-resolved view here pins down the specific transmission pairs in Table 5 and Figure 4 that drive the aggregate.

### *5.2. Key Risk Amplifiers Identified Across Equity-Treasury Markets*

XLF is the dominant transmitter and DGS2 the dominant receiver during the acute window, with the directed pair XLF → DGS2 carrying 12.4 percent of attention-weighted system spillover. The Real Estate ETF (XLRE) acts as a secondary amplifier, transmitting 8.6 percent of system spillover during the same window. This pattern is consistent with the duration-sensitive balance-sheet exposure of listed real estate to the same rate-shock channel that destabilized regional bank securities portfolios. Defensive sectors, Consumer Staples (XLP) and Utilities (XLU), record the smallest changes in attention-weighted strength between sub-periods, in line with the cross-sectional heterogeneity in transmission roles familiar from daily-frequency network studies. The Treasury curve responds in a tenor-dependent way: the two-year and five-year nodes absorb most of the equity-originated stress, while the thirty-year node sees only a modest in-strength increase. The contagion channel is concentrated at the short end, where money-market and deposit-substitution dynamics dominate price discovery during the acute window of mid-March 2023.

The dominance of the XLF → DGS2 channel admits a clear economic reading grounded in the mechanics of the deposit-flight episode. The Federal Reserve's 475-basis-point tightening between early 2022 and the onset of the stress left regional-bank securities portfolios carrying large unrealized losses on long-duration holdings, so the equity market's reassessment of bank franchise value and the repricing of short-dated rate expectations were driven by a single underlying shock rather than by two independent ones. As uninsured depositors withdrew funds and equity investors marked down the financial sector---XLF is dominated by money-center and regional-bank constituents---the same information flowed into the front end of the Treasury curve through two reinforcing routes. First, a flight to quality pushed fleeing deposits and balance-sheet-constrained investors into the safest short-duration instruments, bidding up Treasury bills and the two-year note. Second, the market rapidly repriced the expected monetary-policy path, rotating from further hikes toward an anticipated pause and cuts and pricing in the liquidity backstop announced through the Bank Term Funding Program; because the two-year yield is the maturity most sensitive to the expected policy path over the coming cycle, this expectational channel mapped almost mechanically onto DGS2. The directed edge therefore captures equity-side distress in banks propagating into short-rate expectations, which is precisely why the absorbed share concentrates on the two- and five-year nodes and leaves the thirty-year node---governed more by term premia and long-run inflation expectations---comparatively untouched.

The secondary role of XLRE reflects the same rate-and-funding channel operating through a different exposure. Listed real estate is highly duration-sensitive and depends on bank credit for refinancing, and regional banks are the marginal lender to large parts of the commercial-real-estate market; a contraction in regional-bank lending capacity therefore raised refinancing and valuation risk for real-estate equities at the same moment that the rate-cut repricing was lowering discount rates, producing the elevated XLRE → DGS10 and XLRE → XLF linkages recorded in Table 5. By contrast, Consumer Staples (XLP) and Utilities (XLU) move the least because their cash flows are stable, their operating leverage to the policy rate is low, and they are not funded at the margin by the stressed regional-bank sector, so they neither originate nor absorb much of the cross-market shock. The appearance of the reverse edge DGS2 → XLF in the same ranking is consistent with a feedback loop rather than a contradiction: once the front-end rally signalled lower expected funding costs and a credible liquidity backstop, falling short-term yields fed back into bank-equity valuations and partially stabilized XLF, so the two nodes are bidirectionally coupled with the dominant direction running from financial-sector distress into the short end of the curve. Read together, these mechanisms explain why the cross-market amplification observed in March 2023 is best described as a directed, rate-mediated transmission from financial-sector equity to short-duration Treasuries rather than as a symmetric rise in correlations.

### *5.3. Policy Implications and Limitations*

The node-resolved cross-market view offers a candidate input to the systemic-risk monitoring frameworks maintained by central clearing infrastructure providers and prudential regulators. The identification of amplifier nodes can inform the calibration of cross-margining haircuts and the sequencing of liquidity backstops during acute episodes such as the March 2023 deposit-flight window. Several limitations should be noted. The analysis covers a single stress episode and does not establish out-of-sample predictive validity for unrelated future events. The omission of derivatives, foreign-exchange, and credit-default-swap nodes restricts the cross-market scope to two of the four asset classes typically discussed in the systemic-risk literature; adding the missing classes is straightforward once corresponding public series are obtained. The variance-decomposition edges remain conditional on the underlying vector autoregressive specification, and alternative connectedness primitives such as quantile or frequency-domain decomposition are likely to yield qualitatively similar but quantitatively distinct rankings of amplifier nodes. We leave extensions to multi-event stress panels and higher-frequency intraday measurements for future work.

## References

1. D. Acemoglu, A. Ozdaglar, and A. Tahbaz-Salehi, "Systemic risk and stability in financial networks," *American Economic Review*, vol. 105, no. 2, pp. 564–608, 2015.
2. Y. Li, "Enhancing Financial Compliance Transparency through Automated Data Governance and Intelligent Risk Reporting," *Journal of Science, Innovation & Social Impact*, vol. 2, no. 1, pp. 299–313, 2026.
3. M. Deng and S. Xu, "Temporal-Structural Propagation Graph Analysis for Coordinated Misinformation Campaign Detection and Source Attribution in Social Networks," *Journal of Advanced Computing Systems*, vol. 6, no. 5, pp. 1–11, 2026.

4. F. X. Diebold and K. Yilmaz, "On the network topology of variance decompositions: Measuring the connectedness of financial firms," *Journal of Econometrics*, vol. 182, no. 1, pp. 119–134, 2014.
5. Y. Li and X. Fu, "Comparative Evaluation of Graph Neural Networks for Cross-Market Risk Contagion Path Identification in Multi-Layer Financial Networks," *Journal of Sustainability, Policy, and Practice*, vol. 2, no. 3, pp. 1–14, 2026.
6. T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *Proceedings of the 5th International Conference on Learning Representations (ICLR)*, 2017.
7. P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph attention networks," in *Proceedings of the 6th International Conference on Learning Representations (ICLR)*, 2018.
8. W. L. Hamilton, R. Ying, and J. Leskovec, "Inductive representation learning on large graphs," in *Advances in Neural Information Processing Systems 30 (NeurIPS)*, 2017, pp. 1024–1034.
9. S. Brody, U. Alon, and E. Yahav, "How attentive are graph attention networks?," in *Proceedings of the 10th International Conference on Learning Representations (ICLR)*, 2022.
10. A. Pareja et al., "EvolveGCN: Evolving graph convolutional networks for dynamic graphs," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 4, 2020, pp. 5363–5370.
11. A. Sankar, Y. Wu, L. Gou, W. Zhang, and H. Yang, "DySAT: Deep neural representation learning on dynamic graphs via self-attention networks," in *Proceedings of the 13th ACM International Conference on Web Search and Data Mining (WSDM)*, 2020, pp. 519–527.
12. S. Kumar, X. Zhang, and J. Leskovec, "Predicting dynamic embedding trajectory in temporal interaction networks," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD)*, 2019, pp. 1269–1278.
13. S. Xiang, D. Cheng, C. Shang, Y. Zhang, and Y. Liang, "Temporal and heterogeneous graph neural network for financial time series prediction," in *Proceedings of the 31st ACM International Conference on Information and Knowledge Management (CIKM)*, 2022, pp. 3584–3593.
14. X. Liu, "Covariate Feature Importance and Cross-Indication Transferability Analysis for Phase III Oncology Trial Outcome Prediction," *Journal of Science, Innovation & Social Impact*, vol. 2, no. 2, pp. 117–130, 2026.
15. X. Liu, "Fairness-Constrained Temporal Feature Learning Algorithm for Cross-Population Glucose Prediction and Adherence Risk Stratification Using Continuous Glucose Monitoring Data," *Journal of Sustainability, Policy, and Practice*, vol. 2, no. 3, pp. 50–61, 2026.
16. X. Fu and F. Zhao, "An Empirical Comparison of Few-Shot Example Selection Strategies for In-Context Learning on Public Reasoning and QA Benchmarks," *Journal of Computing Innovations and Applications*, vol. 3, no. 2, pp. 119–128, 2025.
17. S. Xu, F. Zhao, and X. Wang, "An Empirical Comparison of Generation Quality and Diversity Between Discrete Diffusion and Autoregressive Text Generation," *Artificial Intelligence and Machine Learning Review*, vol. 6, no. 2, pp. 16–26, 2025.
18. M. Li, F. Zhao, and T. Tang, "How Prompt Specificity Affects Edge Case Handling in LLM-Generated Code: An Empirical Evaluation," *Artificial Intelligence and Machine Learning Review*, vol. 5, no. 4, pp. 139–149, 2024.
19. R. Sawhney, S. Agarwal, A. Wadhwa, and R. Shah, "Spatiotemporal hypergraph convolution network for stock movement forecasting," in *Proceedings of the 20th IEEE International Conference on Data Mining (ICDM)*, 2020, pp. 482–491.
20. L. Long, D. Zou, and W. Shi, "NLP-Driven Psychological Contract Risk Detection in Cross-Cultural Teams: An XGBoost Approach with Cultural Adaptation," *Artificial Intelligence and Machine Learning Review*, vol. 7, no. 2, pp. 43–53, 2026.
21. H. Zhang, "Comparative Evaluation of Automated Approaches for Legal Aid Document Generation: Template-Based, Rule-Based, and LLM-Based Methods," *Journal of Sustainability, Policy, and Practice*, vol. 2, no. 3, pp. 62–79, 2026.
22. Y. Zhang, "Language-Driven Interactive Annotation for Pulmonary Nodules in Chest CT: An LLM Prompt-Translation and Multi-Round Refinement Approach," *Journal of Science, Innovation & Social Impact*, vol. 2, no. 2, pp. 55–66, 2026.
23. J. Li, "Evaluating BLIP-2, LLaVA, and GPT-4V for Accessible Artwork Description Generation in Virtual Museums: A Comparative Study Based on WCAG-Aligned Evaluation and User Satisfaction," *Journal of Sustainability, Policy, and Practice*, vol. 2, no. 3, pp. 1–12, 2026.
24. J. Li, F. Zhang, and M. Li, "Comparative Effectiveness of Blockchain Provenance Verification on Counterfeit Reduction in Art Transactions: A Multi-Scenario Empirical Assessment," *Artificial Intelligence and Machine Learning Review*, vol. 7, no. 2, pp. 82–92, 2026.
25. Y. L. Hsu, Y. C. Tsai, and C. T. Li, "FinGAT: Financial graph attention networks for recommending top-K profitable stocks," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 1, pp. 469–481, 2021.
26. D. Cheng, Y. Tu, Z. Ma, Z. Niu, and L. Zhang, "Risk assessment for networked-guarantee loans using high-order graph attention representation," in *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*, 2019, pp. 5822–5828.
27. P. Xiao and X. Fu, "Comparative Evaluation of Post-Hoc Feature Attribution Methods on Tabular Financial Data: Faithfulness, Stability, and Computational Efficiency," *Journal of Science, Innovation & Social Impact*, vol. 2, no. 3, pp. 1–11, 2026.
28. T. Tang, X. Fu, and C. Luo, "An Empirical Comparison of High-Order Feature Interaction Operators for Conversion Rate Prediction in Sparse, High-Cardinality Message-Ads Traffic: Accuracy, Efficiency, and Offline--Online Consistency," *Journal of Science, Innovation & Social Impact*, vol. 2, no. 3, pp. 12–22, 2026.

29. S. Xu, M. Li, and F. Zhao, “Comparative Empirical Evaluation of Hallucination Mitigation Strategies in LLM-Based Text Generation,” *Journal of Sustainability, Policy, and Practice*, vol. 2, no. 3, pp. 38–49, 2026.
30. X. Wang, X. Fu, and D. Zou, “Passage, Sentence, or Proposition? An Empirical Comparison of Retrieval Granularity Effects on LLM Answer Accuracy in Retrieval-Augmented Generation,” *Journal of Global Engineering Review*, vol. 3, no. 1, pp. 81–90, 2025.
31. X. Fu and D. Zou, “A Comparative Empirical Study of Over-Refusal Behavior in Closed-Source Large Language Models on Pseudo-Harmful Prompts,” *Artificial Intelligence and Machine Learning Review*, vol. 6, no. 3, pp. 34–45, 2025.
32. S. Yang, Z. Zhang, J. Zhou, Y. Wang, W. Sun, X. Zhong, Y. Fang, Q. Yu, and Y. Qi, “Financial risk analysis for SMEs with graph-based supply chain mining,” in *Proceedings of the 29th International Joint Conference on Artificial Intelligence (IJCAI)*, 2020, pp. 4661–4667.
33. Y. Li, “Benchmarking Robustness-Efficiency Trade-offs of Camera-LiDAR Fusion Strategies for 3D Object Detection Under Environmental Corruptions,” *Journal of Science, Innovation & Social Impact*, vol. 2, no. 2, pp. 104–116, 2026.
34. J. Li, M. Liu, and M. Li, “Comparative Evaluation of Ensemble Learning Algorithms for Visitor Engagement Prediction and Content Recommendation Optimization in Virtual Museum Environments,” *Journal of Advanced Computing Systems*, vol. 6, no. 2, pp. 64–74, 2026.
35. Y. Zhang and M. Li, “Accuracy Evaluation of Multi-Factor Forecasting Methods for Customer Service Workload Prediction Under Holiday and Promotional Fluctuations: Evidence from US Service Industry Data,” *Journal of Advanced Computing Systems*, vol. 6, no. 5, pp. 12–20, 2026.
36. S. Wen and T. Tang, “A Comparative Evaluation of URL-Sharing, Content Similarity, and Temporal Synchronicity Signals for Detecting Coordinated Inauthentic Behavior in Multilingual Political Discourse,” *Journal of Global Engineering Review*, vol. 3, no. 2, pp. 69–78, 2025.
37. M. Li, X. Wang, and M. Yu, “Comparative Evaluation of Zero-Shot and Few-Shot Performance of Large Language Models in Low-Resource Language Machine Translation,” *Journal of Global Engineering Review*, vol. 3, no. 2, pp. 59–68, 2025.
38. D. Cheng, Z. Niu, and Y. Zhang, “Contagious chain risk rating for networked-guarantee loans,” in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD)*, 2020, pp. 2715–2723.