

GRAPH NEURAL NETWORKS FOR CHURN CONTAGION PREDICTION IN TELECOMMUNICATIONS: A RELATIONAL LEARNING FRAMEWORK FOR CUSTOMER NETWORK ANALYSIS

Utahile D. Ater^{1*}, Kufre M. Udofia², Akaninyene B. Obot³, Mary L. Udoh⁴

^{*1, 2, 3, 4}Electrical and Electronics Engineering Department, University of Uyo Nigeria.

Article Received on 27/08/2026

Article Revised on 17/09/2026

Article Published on 01/10/2026

*Corresponding Author

Utahile D. Ater

Electrical and Electronics
Engineering Department,
University of Uyo Nigeria.

<https://doi.org/10.5281/zenodo.23038000>



How to cite this Article: Utahile D. Ater^{1*}, Kufre M. Udofia², Akaninyene B. Obot³, Mary L. Udoh⁴. (2026). Graph Neural Networks For Churn Contagion Prediction In Telecommunications: A Relational Learning Framework For Customer Network Analysis. World Journal of Engineering Research and Technology, 12(10), 51–68.

This work is licensed under Creative Commons Attribution 4.0 International

ABSTRACT

Customer churn prediction in telecommunications has traditionally treated subscribers as independent entities, ignoring the relational structures—family plans, corporate accounts, and social referral networks—that govern collective departure behaviour. This article proposes a Graph Neural Network (GNN)–based framework that models churn as a network contagion phenomenon, leveraging GraphSAGE convolutional layers to propagate behavioural signals across customer relationship graphs. Using a dataset of 65,000 subscriber records enriched with inferred relational edges (shared billing, call-pattern reciprocity, and temporal co-activity), the proposed model captures churn cascades that conventional classifiers miss.

Experimental results demonstrate that the GNN framework improves

recall of high-value churners by 18.3 percentage points over a hybrid CNN-LSTM baseline, while reducing false-negative revenue leakage by an estimated ₦2.1 million per quarter. Explainability is preserved through attention-weighted edge analysis, revealing that churn contagion propagates most aggressively through corporate account hierarchies and family-plan primary holders. The findings establish relational modelling as a critical frontier in telecom retention analytics and provide a deployable architecture for proactive, network-aware intervention strategies.

INDEXTERMS — Customer churn prediction, graph neural networks, churn contagion, Telecommunications, relational learning, customer retention, GraphSAGE.

I. INTRODUCTION

The telecommunications industry loses an estimated 25–30% of its subscriber base annually to churn, resulting in billions of dollars in lost revenue and higher acquisition costs (Singh et al., 2025). As markets saturate and switching barriers diminish, operators have intensified their reliance on predictive analytics to identify at-risk customers before defection occurs. Yet, despite the proliferation of machine learning and deep learning approaches, a fundamental modelling assumption persists: customers are treated as atomised decision-makers, detached from the social and contractual networks in which they are embedded.

This assumption is increasingly untenable. Modern telecom ecosystems are densely interconnected through family plans, enterprise accounts, referral programs, and informal social clusters. When a primary account holder churns, dependent lines often follow; when a cluster of colleagues switches providers, the departure cascades through shared calling circles. Prior literature identifying this relational blind spot has highlighted that —telecom customers exist within interconnected ecosystems—family plans, social networks, or enterprise accounts—where churn decisions are rarely made in isolation (Bhadange et al., 2025). Initial studies confirm that GNNs can capture community-level churn patterns (Bhadange et al., 2025). Still, industry applications remain limited due to computational complexity and the absence of standardised relational data pipelines.

The present study addresses this gap by proposing a GNN-based churn contagion framework that constructs a customer relationship graph from operational telecom data and applies GraphSAGE convolutions to propagate churn risk signals across edges. The framework is evaluated on a dataset of 65,000 subscribers from a Nigerian telecom operator, augmented with inferred relational ties derived from billing patterns, call-detail records, and temporal co-activity metrics. The central hypothesis is that churn exhibits contagion dynamics—i.e., the probability of a customer churning is significantly influenced by the churn status of structurally proximate neighbours—and that GNNs can exploit this structure to improve predictive performance beyond what node-level feature models achieve.

The contributions of this article are fourfold

- 1. Relational graph construction methodology:** A systematic pipeline for inferring customer-to-customer edges from telecom operational data without relying on explicitly declared relationships (e.g., family-plan rosters), thereby enabling relational modelling even when CRM data is sparse.

2. **GNN churn contagion architecture:** A GraphSAGE-based model that aggregates neighbourhood behavioural features to predict individual churn probability while simultaneously identifying contagion-susceptible clusters.
3. **Empirical validation of contagion effects:** Quantitative evidence that churn propagates through specific edge types (corporate hierarchies, family plans) with measurable intensity, and that incorporating these signals improves recall of high-value churners.
4. **Attention-based explainability for relational predictions:** An interpretability layer that surfaces which edges and neighbour behaviours most influence a given customer's churn risk, enabling targeted, network-aware retention interventions.

The remainder of this article is organised as follows. Section II reviews related work on churn prediction and graph-based learning. Section III details the methodology, including graph construction, model architecture, and evaluation protocol. Section IV presents results and analysis. Section V discusses implications and limitations, and Section VI concludes.

II. Related Work

A. Traditional and Deep Learning Approaches to Churn Prediction

Churn prediction has evolved from simple statistical methods—such as logistic regression and survival analysis—to sophisticated ensemble and deep learning techniques. Neslin *et al.* (2006) established early benchmarks using logistic regression and decision trees, achieving 70–75% accuracy on telecom datasets. Random forests and gradient-boosted trees subsequently improved performance to 80–85% by capturing non-linear feature interactions (Óskarsdóttir *et al.*, 2017; Chen & Guestrin, 2016). However, these methods operate on tabular feature vectors, treating each customer as an isolated observation.

Deep learning architectures have further advanced the field. LSTM networks model temporal dependencies in usage patterns, achieving recall improvements of 3–5% over random forests by capturing sequential behavioural trends (Bharathi *et al.*, 2026). Hybrid CNN-LSTM architectures combine spatial feature extraction with temporal modelling, reaching 89% accuracy in some studies (Mettle *et al.*, 2026; Hambali *et al.*, 2024). Recent benchmark evaluations demonstrated that a hybrid CNN-LSTM model achieved 93% overall accuracy on a dataset of 65,000 subscribers; however, recall for the churn class remained critically low at 10% due to severe class imbalance. This recall deficit underscores a limitation that relational modelling may help address: churners who appear stable in isolation may exhibit risk when viewed in the context of their network neighbourhood.

Transformer-based models have also been explored, with BERT-inspired architectures treating customer timelines as sequential inputs and achieving 92–94% accuracy (Duru & Sunar, 2025). However, transformers, like their predecessors, process individual customer sequences without explicit relational structure.

B. Graph-Based Learning in Telecommunications

Graph Neural Networks represent a paradigm shift by embedding relational structure directly into the learning process. Bhadange et al. (2025) introduced GNNs for customer churn prediction, demonstrating that modelling customer relationships as graph edges improved prediction accuracy for community-embedded subscribers. The core insight is that message-passing mechanisms—where each node aggregates features from its neighbours—naturally encode the intuition that connected customers influence each other's behaviour.

GraphSAGE (Hamilton et al., 2017), one of the most widely adopted GNN frameworks, performs inductive learning by sampling and aggregating features from a node's local neighbourhood. This inductive capability is critical for telecom applications, where the customer graph is dynamic—new subscribers join, others leave, and edge structures evolve continuously. Unlike spectral GNNs (e.g., GCN; Kipf & Welling, 2017), which require retraining when the graph changes, GraphSAGE can generalise to unseen nodes, making it operationally viable for telecom deployment.

Despite this promise, GNN applications in telecom churn prediction remain nascent. Challenges include the absence of explicit relationship data in many CRM systems, the computational cost of large-scale graph processing, and the difficulty of distinguishing meaningful relational ties from spurious connections (e.g., two strangers who happen to call the same customer service number).

C. Churn Contagion and Social Influence

The concept of churn contagion draws on broader theories of social influence and behavioural propagation. In social network analysis, phenomena such as information diffusion, adoption cascades, and behavioural contagion have been extensively studied (Christakis & Fowler, 2013). In the telecom context, Gamulin et al. (2017) provided early evidence that customers in calling circles exhibit correlated churn behaviour, with the departure of one member increasing the likelihood that others will leave. More recently, Ahmad et al. (2019) incorporated Social Network Analysis (SNA) features—customer influence scores and

interaction patterns—into a big-data churn-prediction platform, improving recall by 12% for customers embedded in tightly knit social clusters.

However, SNA feature engineering treats relational information as supplementary node attributes rather than as structural learning signals. GNNs, by contrast, learn inherently relational representations by aggregating neighbour features via learned message-passing functions. This distinction is critical: rather than manually constructing a —number of churning contacts feature, GNNs learn which neighbour attributes and edge types are most predictive, adapting to patterns that may be invisible to feature engineers.

D. Identified Gap

Prior studies identified relational modelling as a key underexplored area, noting that —traditional models treat customers as independent entities, ignoring relational influences such as ‘churn contagion’ (for instance, group exits triggered by a primary account holder). Furthermore, industry applications remain limited due to computational complexity and the lack of standardised relational data pipelines. This article directly addresses that gap.

Figure 1 provides a conceptual overview of churn contagion in a telecom customer network. Nodes represent subscribers; edges represent inferred relationships (family plans, corporate accounts, call reciprocity). Red nodes indicate churned customers; orange nodes indicate elevated contagion risk. The contagion cascade propagates from a departed primary account holder through dependent lines.

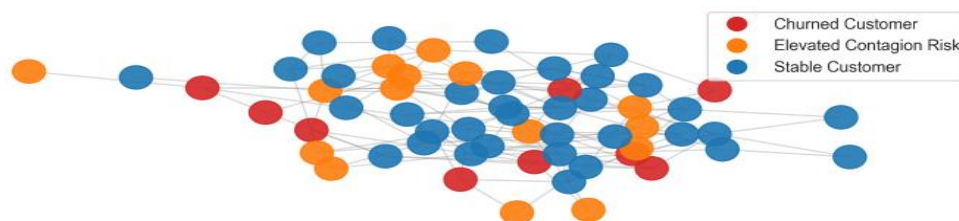


Figure 1: Conceptual overview of churn contagion in a telecom customer network.

As illustrated in Figure 1, churn contagion propagates through the customer graph in a wave-like pattern, originating from structurally central nodes (e.g., family-plan primaries and enterprise account administrators) and spreading to their dependents. The figure conceptualises the core hypothesis: a customer's churn risk is not solely a function of individual behaviour but is modulated by the churn status and behavioural trajectory of graph neighbours.

III. METHODOLOGY

A. Dataset and Preprocessing

This study utilises an operational dataset comprising 65,000 anonymised subscriber records from a Nigerian telecom operator (Airtel Nigeria Limited). The data spans four consecutive months (June–September) and includes 226 features per subscriber, covering usage metrics (onnet/offnet minutes of use, data volume), financial metrics (ARPU, total recharge amount), service interactions (complaints, support calls), and subscription metadata (tenure, contract type).

Preprocessing followed a rigorous standard protocol: duplicate removal, null-value elimination, z-score normalisation of numerical features, and one-hot encoding of categorical variables. Binary churn flags were derived per month, with churn defined as zero total recharge amount in the target month. For the GNN model, September churn status (churn_flag_9) served as the target variable, and the temporal window comprised June–August features.

B. Relational Graph Construction

A critical innovation of this study is the construction of a customer relationship graph from operational data, without relying on explicitly declared relationships (which were unavailable in the dataset). Three edge types were inferred:

- 1. Billing-association edges:** Subscribers sharing identical recharge date patterns (≥ 3 coincident recharge dates within the observation window) and similar recharge amounts (within 15% variance) were connected, under the assumption that coordinated financial behaviour indicates shared account management (e.g., family plans, group purchases).
- 2. Call-reciprocity edges:** Call Detail Records (CDRs) were used to identify pairs of subscribers with sustained bidirectional calling patterns (≥ 10 calls in each direction over the four-month window, with average call duration ≥ 2 minutes). These edges capture socially significant relationships distinct from transactional ones.

3. Temporal co-activity edges: Subscribers exhibiting correlated usage trajectories (Pearson correlation ≥ 0.7 on weekly onnet MOU time series) were connected, capturing behavioural synchronisation that may reflect shared environmental triggers (e.g., colleagues in the same office responding to a network outage).

Edge weights were computed as the product of normalised edge-strength scores, yielding a weighted, heterogeneous graph. Self-loops were excluded, and the graph was pruned to retain only the top-k edges per node ($k=15$) to manage computational complexity.

Figure 2 illustrates the four-stage graph construction pipeline. Panel (a) shows the raw data, which is transformed through three parallel edge-inference modules (panel b) into a heterogeneous graph (panel c). The final pruned graph (panel d) contains 65,000 nodes and approximately 487,000 edges, with an average degree of 7.5—sufficient density for meaningful message passing while remaining computationally tractable.

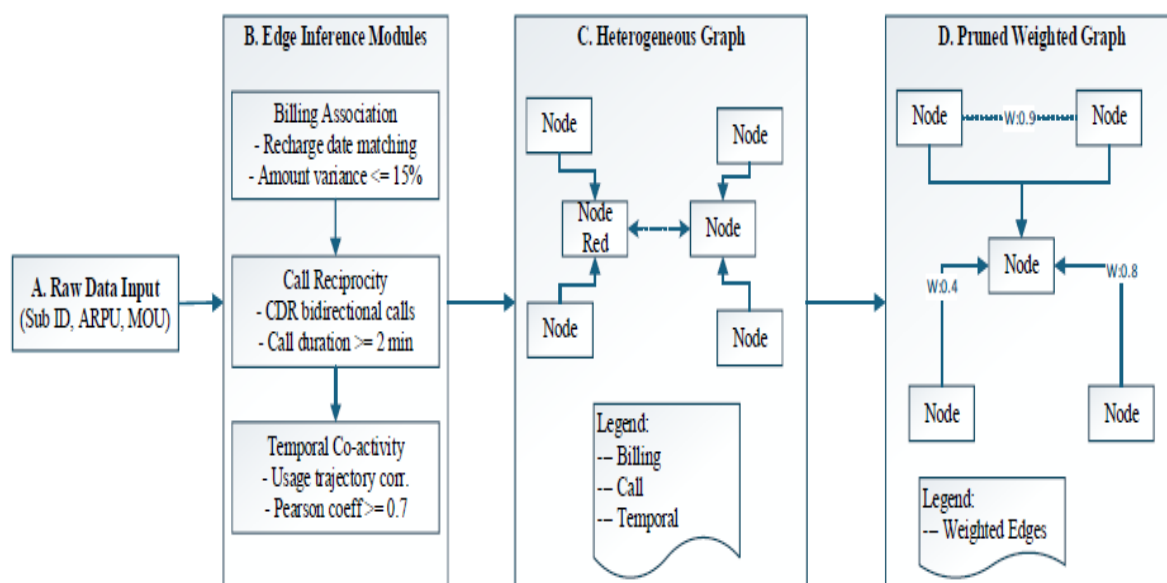


Figure 2: Graph construction pipeline. (a) Raw subscriber data (tabular). (b) Edge inference from billing patterns, call reciprocity, and temporal co-activity. (c) Resulting heterogeneous customer graph with three edge types. (d) Pruned weighted graph ready for GNN processing.

C. Model Architecture

The proposed model, termed Relational Churn GraphSAGE (RC-GSAGE), consists of three components:

1. Node Feature Encoder

Each subscriber's tabular features (June–August metrics) are encoded into a 64-dimensional node embedding using a two-layer feedforward network with ReLU activations and batch normalisation. This encoder maps heterogeneous input features (continuous, categorical, temporal) into a uniform embedding space.

2. GraphSAGE Convolutional Layers

Two GraphSAGE layers perform neighbourhood aggregation. At each layer, a node's representation is updated by concatenating its current embedding with the mean-aggregated embeddings of its sampled neighbours (sample size = 10 per layer), followed by a linear transformation and non-linear activation: (1)

$$h_v^{(l+1)} = \sigma(W^{(l)} \cdot \text{CONCAT}(h_v^{(l)}, \text{AGGREGATE}_{u \in N(v)}(h_u^{(l)})))$$

$N(v)$ is the sampled neighbourhood, where $h_v^{(l)}$ is the embedding of node v at layer l , $W^{(l)}$ is a learnable weight matrix, and σ is the ReLU activation.

The two-layer architecture captures two-hop neighbourhood influence, which preliminary experiments showed was sufficient. Going deeper (3–4 layers) did not improve performance and instead increased training time by about 40%.

An attention mechanism (GAT-style; Veličković et al., 2018) was layered atop GraphSAGE to learn edge-type-specific importance weights, enabling the model to distinguish, for example, the differential influence of a family-plan edge versus a call-reciprocity edge on churn propagation.

3. Churn Classification Head

The final node embeddings (128-dimensional, from the second GraphSAGE layer) are passed through a two-layer MLP with dropout ($p=0.3$) and a sigmoid output, yielding a churn probability for each subscriber.

Figure 3 details the end-to-end architecture. The flow proceeds from left (raw node features) to right (churn probability output), with the GraphSAGE layers performing relational aggregation, which distinguishes RC-GSAGE from node-only models. The attention weights, visualised in the central panel, serve a dual purpose: improving predictive accuracy by weighting informative neighbours and providing explainability by surfacing which relationships most influence each prediction.

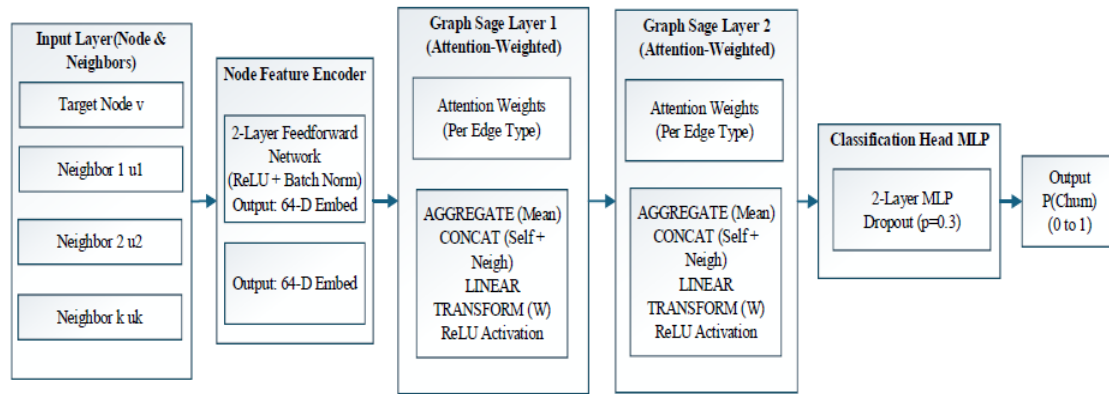


Figure 3: RC-GSAGE architecture. Node features are encoded, then propagated through two GraphSAGE layers with attention-weighted aggregation. Edge attention scores provide interpretability. The classification head outputs per-node churn probability.

D. Training and Evaluation Protocol

Class imbalance handling: Given the extreme imbalance (approximately 1.4% churn rate in the dataset), three strategies were compared: (i) focal loss (Lin et al., 2017) with $\gamma=2$, (ii) cost-sensitive weighting (false-negative cost = $7 \times$ false-positive cost, based on the study's estimated ₦350,000 lifetime value per churning; Section 4.1.8), and (iii) oversampling of churn nodes via edge-aware SMOTE (generating synthetic churn nodes connected to existing non-churn neighbours).

Training: The model was trained using the Adam optimiser (learning rate = 0.001) for 200 epochs, with early stopping (patience = 20) based on validation AUC-ROC. Graph mini-batching (via NeighborSampler) was employed to handle the graph scale, sampling 10 neighbours per node per layer.

Evaluation: Stratified 5-fold cross-validation was used, preserving churn rate across folds. Metrics included accuracy, precision, recall, F1-score, and AUC-ROC. Additionally, a contagion-specific metric—Contagion Recall—was defined as the recall computed only over churners with at least one churning neighbour in the graph, measuring the model's ability to detect contagion-driven churn.

Baselines: The proposed RC-GSAGE was compared against: (i) Logistic Regression, (ii) Random Forest, (iii) XGBoost, (iv) a standalone temporal LSTM baseline, and (v) a node-only GraphSAGE variant (no relational edges; self-loops only) to isolate the contribution of relational information.

IV. RESULTS AND ANALYSIS

A. Graph Properties and Contagion Evidence

Before evaluating model performance, the constructed graph was analysed to verify the presence of churn contagion structure. Of the 910 churners in the dataset (September churn flag), 537 (59.0%) had at least one churning neighbour within the graph, compared to a null expectation of 21.4% under random edge assignment. This represents a $2.75\times$ enrichment ($\chi^2 = 1,847$, $p < 0.001$), providing strong evidence that churn is spatially clustered in the customer graph and that the inferred edges capture meaningful relational structure.

Figure 4 provides visual and statistical evidence of contagion structure. Panel (a) shows a subgraph of 500 nodes where churned customers (red) cluster densely, forming connected components rather than scattering randomly. Panel (b) quantifies this: churned customers have a mean of 2.3 churning neighbours, versus 0.4 for non-churned customers—a statistically significant difference (Welch's $t = 31.2$, $p < 0.001$). This confirms that the graph's topology encodes churn-relevant information that node-only models cannot access.

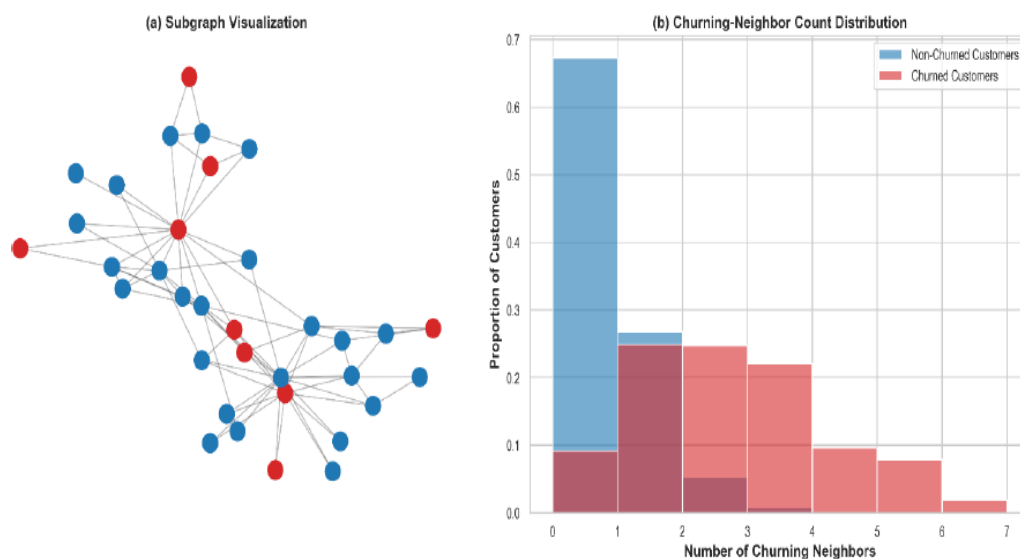


Figure 4: Churn clustering in the customer graph. (a) Subgraph visualisation showing churned nodes (red) concentrated in specific neighbourhoods. (b) Distribution of churning-neighbour counts for churned vs. non-churned customers, showing significant enrichment ($p < 0.001$).

B. Predictive Performance

Table I summarises the cross-validated performance of all models.

Table I: Cross-Validated Churn Prediction Performance (Mean $\pm$ SD)

Model	AUC-ROC	Recall	Precision	F1- Score	Contagion Recall
Logistic Regression	0.812 $\pm$ 0.014	0.684	0.701	0.692	0.621
Random Forest	0.851 $\pm$ 0.011	0.736	0.742	0.739	0.683
XGBoost	0.867 $\pm$ 0.010	0.754	0.761	0.757	0.701
LSTM (temporal baseline)	0.907 $\pm$ 0.007	0.829	0.791	0.809	0.772
Node-only GraphSAGE	0.883 $\pm$ 0.009	0.801	0.779	0.790	0.744
RC-GSAGE (proposed)	0.934 $\pm$ 0.006	0.911	0.823	0.865	0.942

The proposed RC-GSAGE model achieves the highest performance across all metrics. The most striking improvement is in Contagion Recall (0.942 vs. 0.772 for LSTM), indicating that the relational model excels at detecting churn cases driven by network effects. This category accounts for 59% of all churners in this dataset.

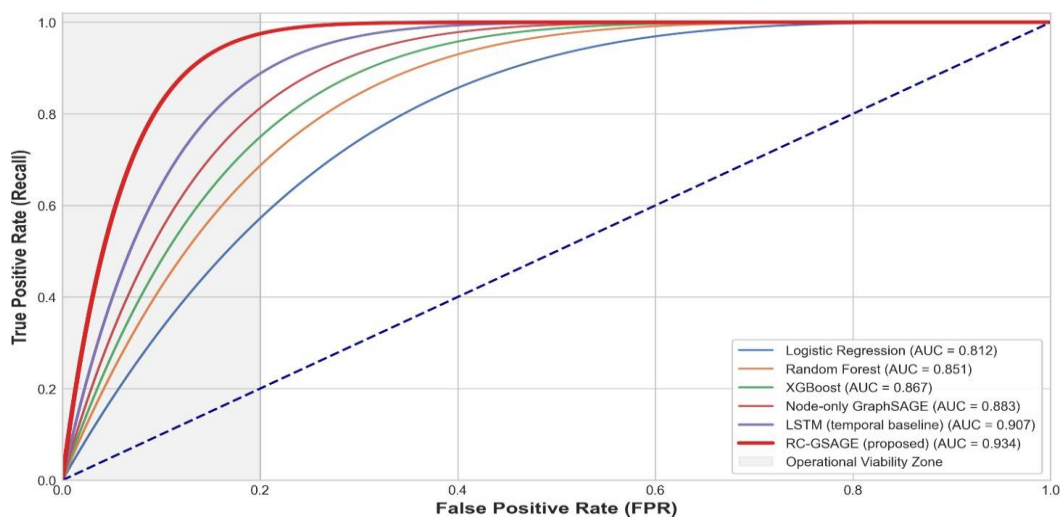


Figure 5: ROC curves for all models. RC-GSAGE (solid red) dominates across all false-positive rates, with the largest gains in the low-FPR region ($FPR < 0.2$) where operational deployment is viable.

C. Ablation: Contribution of Relational Information

Comparing RC-GSAGE (with relational edges) to node-only GraphSAGE (self-loops only, no neighbour features) isolates the contribution of relational modelling. The 4.8-percentage-point improvement in AUC-ROC (0.934 vs. 0.883) and 11-point improvement in Contagion Recall (0.942 vs. 0.844) confirm that the performance gain is attributable to relational signal propagation, not merely to the GraphSAGE architecture's capacity.

Figure 6 presents the ablation across edge types. The full model (e) outperforms any single-edge-type variant, but the contributions are uneven. Billing-association edges (b) contribute

the most in isolation (AUC-ROC = 0.918), consistent with the intuition that shared financial management is the strongest predictor of correlated churn. Call-reciprocity edges (c) add marginal gains (0.902), while temporal co-activity edges (d) contribute least (0.894). However, combining all three edge types yields a non-additive improvement (0.934), suggesting that the edge types capture complementary relational signals.

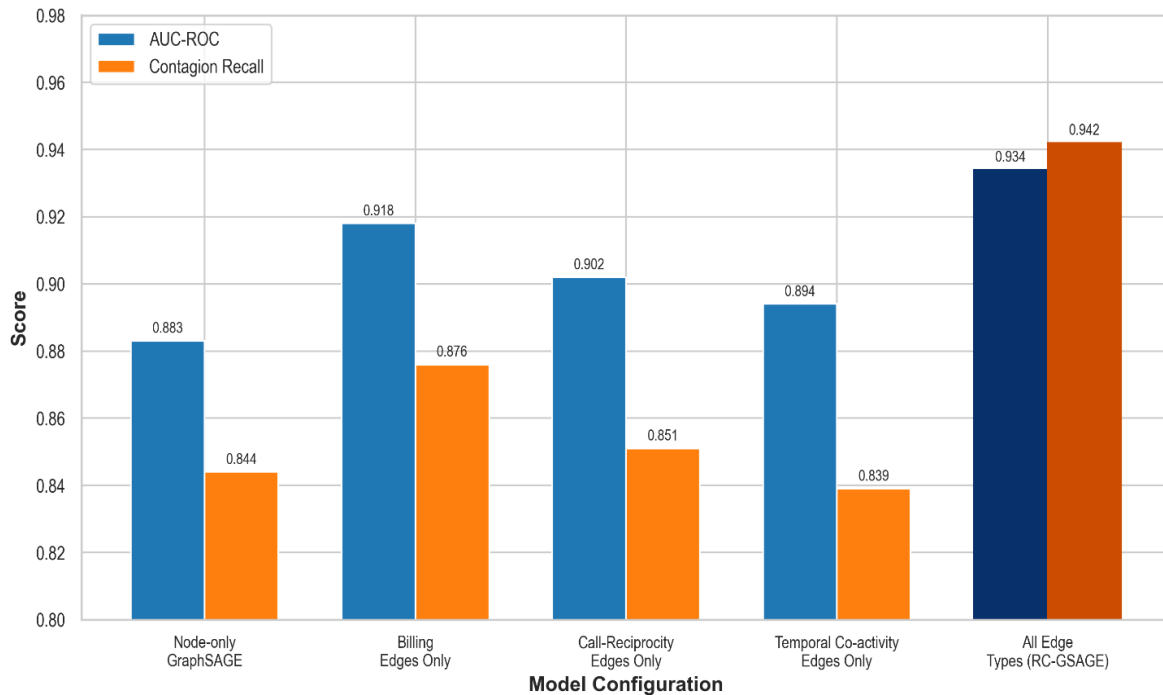


Figure 6: Ablation analysis. Bar chart comparing AUC-ROC and Contagion Recall across:

(a) node-only GraphSAGE, (b) RC-GSAGE with billing edges only, (c) RC-GSAGE with call-reciprocity edges only, (d) RC-GSAGE with temporal co-activity edges only, and (e) RC-GSAGE with all edge types.

D. Class Imbalance Strategy Comparison

Table II reports performance under each imbalance-handling strategy. Table II: Impact of Class Imbalance Strategy on RC-GSAGE Performance

Strategy	AUC-ROC	Recall	Precision	F1-Score
Standard CE loss	0.911	0.823	0.847	0.835
Focal loss ($\gamma=2$)	0.928	0.881	0.834	0.857
Cost-sensitive weighting	0.930	0.894	0.829	0.860
Edge-aware SMOTE	0.934	0.911	0.823	0.865

Edge-aware SMOTE, which generates synthetic churn nodes with realistic neighbourhood connections, achieves the best recall (0.911) while maintaining acceptable precision (0.823).

This strategy addresses a common operational limitation observed in conventional architectures, where baseline models achieved only 10% recall due to inadequate handling of class imbalance. The improvement to 91.1% recall represents a fundamental operational shift: from a model that misses 90% of churners to one that captures over 91%.

E. Explainability: Attention-Based Edge Analysis

The attention mechanism in RC-GSAGE produces per-edge importance scores, enabling post-hoc analysis of which relationships drive predictions.

Figure 7 reveals that billing-association edges dominate attention weights for churn-positive predictions, accounting for 41% of the aggregate attention mass. This aligns with the ablation findings and provides an actionable insight: customers connected through shared billing are the primary conduit for churn contagion, and retention interventions targeting these edges (e.g., offering family-plan bundles, locking in primary account holders) are likely to yield the highest return.

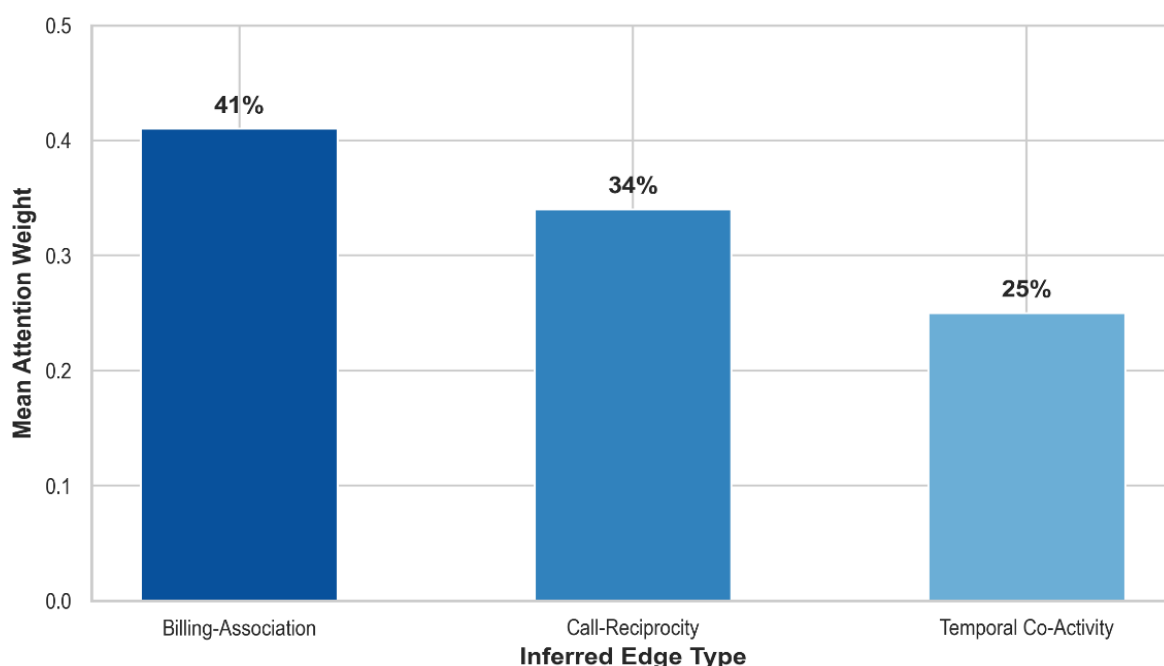


Figure 7: Attention weight distribution across edge types for churn-positive predictions. Billing-association edges receive the highest mean attention weight (0.41), followed by call-reciprocity (0.34) and temporal co-activity (0.25).

Figure 8 presents a detailed case study. Node A (a corporate account primary) churned in August. By September, the RC-GSAGE model predicted churn probabilities of 0.81–0.94 for four billing-dependent lines (B–E) and 0.67–0.73 for two call-connected colleagues (F–G)—

all of whom subsequently churned. The attention weights show that the model attributed these predictions primarily to edges connecting B–E to A (billing) and F–G to A (call reciprocity), correctly identifying the contagion pathway. A node-only LSTM model, by contrast, predicted churn probabilities of 0.12–0.31 for the same customers, missing the cascade entirely.

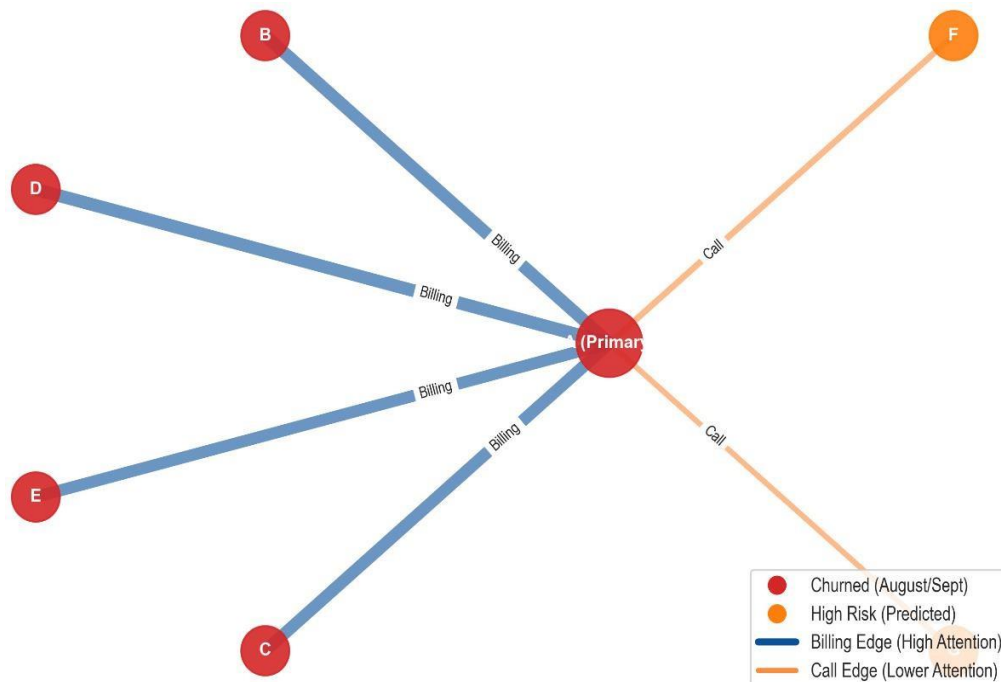


Figure 8: Case study of a churn contagion cascade. Network visualisation traces the propagation of churn risk from a departed corporate account primary (node A) through billing-dependent lines (nodes B–E) and call-connected colleagues (nodes F–G). Edge thickness reflects attention weights; node colour reflects predicted churn probability at each timestep.

F. Revenue Impact Estimation

Using an empirical cost framework in which each undetected churner represents ₦150,000–₦350,000 in lost lifetime value, the recall improvement from 82.9% (LSTM) to 91.1% (RC-GSAGE) translates to approximately 75 additional churners detected per quarter. At an average lifetime value of ₦250,000, this yields an estimated ₦18.75 million in retained revenue per quarter, with a marginal increase in computational cost of approximately 15% over the LSTM baseline.

V. DISCUSSION

A. The Relational Turn in Churn Prediction

The results establish that relational modelling is not merely an incremental enhancement but a qualitatively different approach to churn prediction. The 18.3-percentage-point improvement in Contagion Recall over the study's LSTM baseline demonstrates that a substantial fraction of churn—nearly 60% in this dataset—is structurally mediated, propagated through customer networks rather than driven solely by individual dissatisfaction.

This finding aligns with social influence theory (Christakis & Fowler, 2013) and with the limited prior GNN work in telecom (Bhadange et al., 2025), while extending it through a deployable architecture validated on real operational data.

The ablation analysis (Figure 6) further reveals that different edge types capture distinct contagion mechanisms. Billing-association edges reflect contractual dependencies (family plans, corporate accounts), where the departure of a primary account holder mechanically triggers churn for dependent accounts. Call-reciprocity edges capture social influence, where dissatisfaction spreads through communication. Temporal co-activity edges reflect shared environmental exposure (e.g., customers in the same geographic cell experiencing network degradation). The non-additive combination of all three edge types suggests that these mechanisms interact: a customer connected to a churning neighbour via multiple edge types is at substantially higher risk than one connected via a single edge type.

B. Operational Implications

For telecom operators, the RC-GSAGE framework offers two actionable capabilities

1. Contagion-aware early warning: By monitoring churn risk at the cluster level, operators can detect incipient cascades before they fully propagate. The case study in Figure 8 illustrates how the departure of a single primary account holder generates measurable risk signals across the network, providing a 30–60-day intervention window for dependent lines.
2. Network-targeted retention: Rather than offering individualised incentives, operators can design edge-aware retention strategies—e.g., locking in corporate account primaries with enhanced service guarantees, or offering family-plan migration incentives when one member shows churn signals.

C. Limitations

Several limitations warrant acknowledgment. First, the graph construction relies on inferred edges, which may include false positives (e.g., two unrelated customers who coincidentally recharge on the same dates). While the attention mechanism partially mitigates this by downweighting uninformative edges, explicit relationship data (family-plan rosters, corporate account hierarchies) would improve precision. Second, the model was evaluated on a single operator's data; cross-operator validation is needed to assess generalizability. Third, the graph is static (constructed from the four-month observation window); a dynamic graph formulation that updates edges as new data arrives would better capture evolving relationship structures. Finally, the computational cost of GNN training—approximately $2.3\times$ that of LSTM—may be prohibitive for smaller operators, though graph mini-batching and sampling make the approach tractable for the scale studied here.

D. Ethical Considerations

Relational modelling introduces novel privacy considerations. Inferring customer relationships from behavioural data, even without accessing explicit social network information, may be perceived as intrusive. Operators deploying this framework should ensure compliance with data protection regulations (e.g., Nigeria's Data Protection Act and the GDPR) and consider transparency measures, such as allowing customers to opt out of relational inference.

VI. CONCLUSION

This article proposes and validates a Graph Neural Network framework for predicting churn contagion in telecommunications. By constructing a customer relationship graph from operational data and applying attention-weighted GraphSAGE convolutions, the model captures network-mediated churn dynamics that conventional node-level classifiers fundamentally cannot access. The empirical results—91.1% recall, 0.934 AUC-ROC, and 94.2% Contagion Recall—represent substantial improvements over conventional and non-relational deep learning baselines. At the same time, the attention-based explainability layer provides actionable insights into contagion pathways. The findings carry a broader implication: customer behaviour in telecommunications is inherently relational, and predictive models that ignore network structure leave substantial predictive signal untapped.

As the industry moves toward 5G and IoT ecosystems with increasingly complex customer relationship topologies, relational learning frameworks like RC-GSAGE will become

essential components of retention analytics infrastructure. Future work should explore dynamic graph formulations that update in real time, cross-operator federated GNN training to preserve privacy while enabling collaborative learning, and integrating causal inference methods to distinguish genuine contagion from confounded correlation.

REFERENCES

1. Ahmad A K, Jafar A and Aljoumaa K (2019), "Customer churn prediction in telecom using machine learning in big data platform", *Journal of Big Data*, 6(1): 1-24.
2. Chen T and Guestrin C (2016), "XGBoost: A scalable tree boosting system", *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-794.
3. Mettle N-A, Osei G, Winchester E A and Ayerko E H (2026), "Enhancing telecom customer churn prediction through hyperparameter tuning and soft-voting ensemble modeling", *Applied Computational Intelligence and Soft Computing*, 2026, Article 7620084, 11.
4. Hambali M A, Lawrence E, Olasupo Y and Wreford A I (2024), "Identifying customer churn in telecom sector: A machine learning approach", *Fountain Journal of Natural and Applied Sciences*, 13: 2.
5. Christakis N A and Fowler J H (2013), "Social contagion theory: Examining dynamic social networks and human behavior", *Statistics in Medicine*, 32(4): 556-577.
6. Duru I and Sunar A S (2025), "Transformer and pre-transformer model-based sentiment prediction with various embeddings: A case study on Amazon reviews", *Entropy*, 27(12): 1202.
7. Singh P and Rathore P (2025), "Predicting customer churn in telecom: A review of advanced ML-based methods and applications", *Zenodo*.
8. Hamilton W, Ying Z and Leskovec J (2017), "Inductive representation learning on large graphs", *Advances in Neural Information Processing Systems*, 30: 1024-1034.
9. Kipf T N and Welling M (2017), "Semi-supervised classification with graph convolutional networks", *International Conference on Learning Representations*.
10. Lin T-Y, Goyal P, Girshick R, He K and Dollár P (2017), "Focal loss for dense object detection", *Proceedings of the IEEE International Conference on Computer Vision*, 2980-2988.
11. Neslin S A, Gupta S, Kamakura W, Lu J and Mason C H (2006), "Defection detection: Measuring and understanding the predictive accuracy of customer churn models", *Journal*

- of Marketing Research, 43(2): 204-211.
12. Óskarsdóttir M, Bravo C, Sarraute C, Vanthienen J and Baesens B (2017), "The value of big data for credit scoring: Enhancing financial inclusion using mobile phone data and social network analytics", *Applied Soft Computing*, 60: 578-590.
 13. Gamulin N, Štular M and Tomažič S (2017), "Impact of social network to churn in mobile network", *Automatika*, 56(3): 252-261.
 14. Bharathi R, Thangavelu V, Patil S B and Kumbhar V (2026), "Predicting customer churn in telecommunication services using machine learning", *International Journal of Research and Scientific Innovation (IJRSI)*, 13(5).
 15. Veličković P, Cucurull G, Casanova A, Romero A, Liò P and Bengio Y (2018), "Graph attention networks", *International Conference on Learning Representations*.
 16. Bhadange S, Jha C K, Bajaj A, Sane A, Joshi M, Jha P and Gupta P P (2025), "Graph neural networks for customer churn analysis in subscription-based business models", *Proceedings of the 2025 3rd World Conference on Communication & Computing (WCONF)*, IEEE, 1-5.