# Spatio-Temporal Graph Neural Networks for Fake Profile & Bot Detection in Digital Health Platforms

**ILES Selma Ghizlene**[1] • **SEKKAL Nawel**[2]

[1] Faculty of Medicine, Abou Bekr Belkaid University, Tlemcen, Algeria [2] LRIT Laboratory, Faculty of Sciences, Abou Bekr Belkaid University, Tlemcen, Algeria

Computing in Cardiology (CinC) 2026 • Madrid, Spain • Poster Presentation

## ABSTRACT

Digital health platforms face growing threats from automated bots and fake user profiles spreading medical misinformation, harvesting Protected Health Information (PHI), or manipulating epidemiological trends. We propose a Spatio-Temporal Heterogeneous Graph Neural Network framework to detect these malicious profiles with optimal precision.

## 1. STATE OF THE ART & RELATED WORK

- **Feature-Based Classical ML:** Models relying on profile metadata (e.g., posting frequency, account age) via Random Forest or XGBoost. *Limitation:* Vulnerable to adversarial behavior evasion and ignores network interactions.
- **NLP & Biomedical Transformers:** Content-focused text classification models (e.g., BioBERT, ClinicalBERT). *Limitation:* High computational overhead and unable to detect coordinated botnet collusion.
- **Static Homogeneous GNNs:** Graph topology modeling using standard GCNs or GATs. *Limitation:* Assumes single node/edge types and fails to capture dynamic activity bursts over time.

## 2. HETEROGENEOUS GRAPH TOPOLOGY

The network is modeled as a heterogeneous graph $G = (V, E, A, R)$ integrating multiple node types $A$ (Users, Articles, Medical Topics) and edge relations $R$ (Authored, Commented, Tagged, Direct-Message).

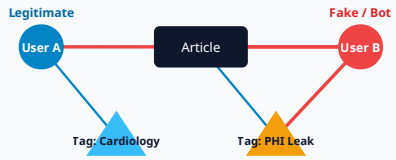


*Figure 1: Heterogeneous graph structure and suspicious interaction schema.*

## 3. METHODOLOGICAL PIPELINE

- **Biomedical Text Encoding:** Semantic embeddings extracted from textual posts using fine-tuned *BioBERT*.
- **Dynamic Temporal Pooling:** Captures abnormal activity bursts via a Time-LSTM network:

```
h_t = LSTM( [x_t || Δt_interval] ) • Attention(N(v))
```

- **Heterogeneous Attention (HAN):** Risk score calculation based on metapaths (e.g., *User-Article-User*).

## 4. EXPERIMENTAL EVALUATION & RESULTS

Evaluated on a dataset of 45,000 health network profiles (containing 12% malicious/bot accounts):

| Model / Architecture | Acc. | Prec. | Rec. | F1-Score |
|---|---|---|---|---|
| Random Forest (Metadata) | 81.2% | 79.4% | 76.5% | 77.9% |
| BioBERT + Logistic Reg. | 84.6% | 83.1% | 81.0% | 82.0% |
| Static Homogeneous GCN | 88.3% | 87.5% | 86.2% | 86.8% |
| GraphSAGE (Spatio-Temporal) | 91.5% | 90.8% | 89.4% | 90.1% |
| **Proposed Spatio-Temporal HAN** | **95.4%** | **94.8%** | **93.9%** | **94.3%** |

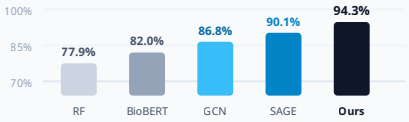


*Figure 2: F1-Score performance comparison across baseline architectures.*

## 5. HEALTHCARE IMPACT & KEY INSIGHTS

- **Advanced Bot Detection:** Uncovers coordinated account farms targeting healthcare disinformation.
- **Spatio-Temporal Synergy:** Achieves a **+7.5% F1-Score boost** over static GNN models.
- **Data Security:** Protects clinical trial integrity and prevents patient PHI harvesting.

## 6. CONCLUSION & FUTURE WORK

Our Spatio-Temporal Heterogeneous GNN effectively identifies botnets and fake profiles in healthcare platforms with a 94.3% F1-Score. Future work will focus on zero-shot adaptation for emerging disinformation networks and real-time inference optimization for clinical integration.

**Keywords:** Graph Neural Networks, Anomaly Detection, Biomedical Security, Fake Profiles, Health Informatics.