Boyu Qiao

(+86) 15602005125 | qiaoboyu@iie.ac.cn | Homepage

Education

Institute of Information Engineering, Chinese Academy of Sciences - PhD in Cyber Security, Prof. [Songlin Hu] 2021.08 - Now

Hong Kong University of Science and Technology - Academic Exchange, Prof. [Céline Yunya Song] 2025.07 - 2026.07

- **Research Area**: Social Bot Detection, Rumor /Fake News Detection, Graph Representation Learning, Multi-Agent Simulation

Hebei University of Technology (211) - Bachelor of Network Engineering 2017.08-2021.06

- **GPA**: 3.84/4.00 (2/71). **Graduation Project**: News Event Matching

Research Experience

[1] Dynamic Simulation Framework for Disinformation Dissemination and Correction With Social Bots. Boyu Qiao, Kun Li, Wei Zhou, and Songlin Hu. EMNLP 2025 (CCF-B, Review Stage) [PDF] [Code]

- **Task Objective:** Simulating disinformation dynamics to assess the impact of malicious bots and the effectiveness of correction strategies from legitimate bots.
- **Technical Framework: 1.** Constructing a social propagation network based on the **Stochastic Block (SB)** and the **Barabási-Albert (BA)** Model. **2.** Utilizing **prompt engineering** and **statistical modeling** methods to generate multi-dimensional attribute features for user agents, including identity labels, interest preferences, and behavioral patterns. **3.**Establishing a **time-driven mechanism** and **tool-calling interfaces** to schedule dynamic interaction simulations among **malicious bot agents**, **legitimate bot agents**, and **regular human agents** at various stages. **4.** Designing **dynamic quantitative evaluation metrics** to assess the influence of malicious and legitimate bots on regular users.

[2] BotSim: LLM-Powered Malicious Social Botnet Simulation. Boyu Qiao, Kun Li, Wei Zhou, Shilong Li, Qianqian Lu, and Songlin Hu. AAAI 2025 (CCF-A) [PDF] [Code] [Dataset]

- **Task Objective:** Construct a social network simulation based on LLM to model social bot strategies and develop an LLM-driven social bot dataset.
- **Technical Framework: 1.** Build the social network simulation using the **LangChain framework and Prompt engineering**. **2.** Use the **PRAW library** to collect real account data from Reddit and build the raw ecological data for the simulation environment. **3.**Integrate simulation flow, bot role settings, and actions to drive Prompt-based bot agent interactions, creating a dataset with **1,907 humans and 1,000 bots**. **4.** Evaluate dataset quality using four baseline methods: **feature engineering, semantic encoding, homogeneous GNN, and heterogeneous GNN**.

[3] Identifying Bots on Social Media through Coordinated Group Perception. Boyu Qiao, Kun Li, Wei Zhou, Shilong Li, Qianqian Lu, and Songlin Hu. ICASSP 2025 (CCF-B) [PDF] [Code]

- **Task Objective:** Identify the inherent coordination behaviors of bot account groups using community detection techniques.
- **Technical Framework: 1.** Extract coordination features from posting content using **Autoencoders and K-means**. **2.** Uncover coordination features in social relationships with **GCN and differentiable modularity**. **3.** Aggregate user attributes and coordination group features based on **RGCN**.

[4] Dispelling the Fake: Social Bot Detection Based on Edge Confidence Evaluation. Boyu Qiao, Wei Zhou, Kun Li, Shilong Li, and Songlin Hu. TNNLS 2024 (SCI Q1) [PDF] [Code]

- **Task Objective:** Identify disguised interaction edges of bots in social networks to mitigate their impact on the GNN homogeneity assumption.
- **Technical Framework: 1.** Encode account metadata and text using **MLP and RoBERTa**, and fuse features with **multi-head attention**. **2.** Construct edge features based on L1-distance between user nodes, and utilize **Parametric Gaussian** to reconstruct edge representations in latent semantic space. **3.** Estimate edge confidence with the **Sigmoid function**, and remove low-confidence edges based on **Bernoulli distribution** to update the social graph. **4.** Apply six types of **homogeneous and heterogeneous GNN** methods on the updated graph for node classification.

[5] Multi-modal Social Bot Detection: Learning Homophilic and Heterophilic Connections Adaptively. Shilong Li, Boyu Qiao, Kun Li, Qianqian Lu, Meng Lin, and Wei Zhou. MM 2023 (CCF A) [PDF]

- **Task Objective:** Enhance edge relationships in the social network graph using graph augmentation strategies.
- **Technical Framework: 1. Graph Augmentation Strategy 1**: Add potential edge relationships based on **node representation similarity**. **2. Graph Augmentation Strategy 2**: Design an **edge classifier** to add homophilic and heterophilic edges. **3.** Combine the original graph with Strategies 1 and 2 to **build a more complete social graph**. **4.** Apply **heterogeneous GNN** on the updated graph to aggregate neighbor features for node classification.

[6] Social Bot Detection Based on Window Strategy. Boyu Qiao, Kun Li, Wei Zhou, Zhou Yan, and Songlin Hu. ICME 2023 (CCF B) [PDF]

- **Task Objective:** Identify differences in posting behavior motivations between social bots and humans using a sliding window strategy.

- **Technical Framework: 1.** Build initial posting embeddings using **Roberta**, and update with **Bi-LSTM** to capture past and future sequence information. **2.** Use **sliding windows** to segment user posting sequences, and apply **multi-head attention** to assess representation changes across windows. **3.** Integrate user metadata and window-based representation changes as node embeddings in the social graph. **4.** Apply **heterogeneous GNN** to aggregate neighbor features for node classification.

Competition Experience

[1] Champion of the 'Explainable Social Network Community Detection' Track at the Songshan Lake AI Algorithm Challenge. Boyu Qiao, Shiling Li, Kun Li, Wei Zhou. 2023. [PPT] [Code]

- **Task Objective:** Use modularity metric to partition users into different communities and explore the interpretability of the partitioning results.
- **Technical Framework: 1.** Construct social graphs based on user forwarding relationships, initialize node embeddings in **one-hot form**, and update embeddings using **GCN encoder**. **2.** Apply **K-means** to group nodes into clusters. **3.** Design a **differentiable modularity function** to optimize the graph neural network as the loss function. **4.**Explore the interpretability of the community divisions by **mapping them to user attribute features**.

Professional Skill

- **Language**: English, CET-6
- **Expertise**: Deep Learning, Social Network Analysis, Graph Neural Networks, Natural Language Processing
- **Skills**: PyTorch, TensorFlow
- **Reviewer**：MM, TKDE, ICASSP