



**The College of Graduate Studies and the College of Information Technology  
Cordially Invite You to a  
Phd Dissertation Defense**

Entitled

*STREAM-EVOLVING BOT DETECTION FRAMEWORK USING GRAPH-BASED AND FEATURE-BASED APPROACHES FOR IDENTIFYING SOCIAL BOTS ON TWITTER*

by

Eiman Alothali

Faculty Advisor

Dr. Hany Alashwal

College of Information Technology

Date & Venue

8 Jun 2023

1:00-3:30 PM

Room 1060 , E1 Building

Abstract

This dissertation focuses on the problem of evolving social bots in online social networks, particularly Twitter. Such accounts spread misinformation and inflate social network content to mislead the masses. The main objective of this dissertation is to propose a stream-based evolving bot detection framework (SEBD), which was constructed using both graph- and feature-based models. It was built using Python, a real-time streaming engine (Apache Kafka), and our pretrained model (bot multi-view graph attention network (Bot-MGAT)). The feature-based model was used to identify predictive features for bot detection and evaluate the SEBD predictions. The graph-based model was used to facilitate multiview graph attention networks (GATs) with fellowship links to build our framework for predicting account labels from streams. A probably approximately correct learning framework was applied to confirm the accuracy and confidence levels of SEBD.

The results showed that the SEBD can effectively identify bots from streams and profile features are sufficient for detecting social bots. The pretrained Bot-MGAT model uses fellowship links to reveal hidden information that can aid in identifying bot accounts. The significant contributions of this study are the development of a stream-based bot detection framework for detecting social bots based on a given hashtag and the proposal of a hybrid approach for feature selection to identify predictive features for identifying bot accounts. Our findings indicate that Twitter has a higher percentage of active bots than humans in hashtags. The results indicated that stream-based detection is more effective than offline detection. Finally, semi-supervised learning can solve the issue of labeled data in bot detection tasks.

**Keywords:** social bots, detection, real-time, stream, machine learning, GAT, GNN, PAC, SL, Twitter