ARTEMIS Brief

ARTEMIS is a graph neural network (GNN) system designed to detect airdrop hunters in NFT markets. These hunters exploit airdrop mechanisms by using multiple wallet addresses to unfairly claim token rewards, threatening the decentralization and fairness of Web3 ecosystems.

ARTEMIS introduces a multi-modal graph learning framework that captures rich transactional and metadata features from NFT activity. It combines three key innovations:

1. Multimodal Feature Extraction: Utilizes transformer models to analyze both visual and textual NFT metadata, converting them into embedding features.
2. Behavior-Aware Node Aggregation: Introduces a customized GNN aggregation that models sequential transaction behavior, capturing patterns typical of hunter activity
3. Market Manipulation Features: Integrates handcrafted features inspired by financial market manipulation theory to enhance detection accuracy

In training, ARTEMIS processes a graph where nodes represent wallets, and edges represent NFT transactions with associated metadata. The model is trained using binary classification to predict whether a node is a hunter.

Training Summary:

In Run 1, ARTEMIS trained for 100 epochs with strong performance:

- Final Test Accuracy: 96.5%
- Precision: 84.1%
- Recall: 79.3%
- F1 Score: 81.7%

These metrics indicate the model generalizes well to unseen wallets and is effective at identifying hunter behavior.

Limitation Observed:

A concrete limitation of ARTEMIS is its high training time and resource demand. Run 1 took approximately 90 minutes on a Colab T4 GPU, which may be impractical for large-scale or real-time detection scenarios. Additionally, occasional recall instability warnings (e.g., zero recall on some batches) suggest sensitivity to imbalanced data and the need for careful sampling strategies.