Graphing the Insider: Innovative Applications of GNNs in Insider Threat Detection

KARTIKEYA SHARMA
SENIOR ASSOCIATE INFORMATION SECURITY ENGINEER @ EQUINIX

About me

PNW Connection

- Currently live and work in Seattle
- Completed my Master's degree in CS from UO

My Work

- Work on the intersection of Data Analytics, AI and Cybersecurity

Hobbies

- Hiking
- Drinking overpriced coffee
- Embracing flannel as a personality

Agenda

1. **Understanding Insider Threats**
2. **Introduction to Graph Neural Networks (GNNs)**
3. **GNN Approaches for Insider Threat Detection**
4. **Challenges and Future Directions**

Understanding Insider Threats

What is an Insider Threat?

Insider threats are cybersecurity threats that originate with authorized users, such as employees, contractors and business partners, who intentionally or accidentally misuse their legitimate access, or have their accounts hijacked by cybercriminals.

SOURCE: IBM

A notable example

In 2023, two ex-Tesla workers leaked sensitive data of 75,735 employees to a German newspaper, discovered by Tesla on May 10 via Handelsblatt.

SOURCE: THE VERGE

Cost of Insider Threats

- **$16.2M – Average annual cost per organization.**
- **86 days – Average time to contain an insider incident.**
- **3 Types of Insider Threats:**
  - **Negligent/Mistaken (55%) – $505K per incident**
  - **Malicious (25%) – $701K per incident**
  - **Outsmarted (Credential Theft, 20%) – $680K**

SOURCE: PONEMON INSTITUTE

Types of insiders

- **Malicious**: Discontented employees deliberately misusing access
- **Contractor**: Temporary workers with potential security risks
- **Inadvertent**: Unintentionally compromising security due to lack of awareness
- **Negligent**: Failing to meet security standards due to carelessness

Insider threat activities

- <u>IT Sabotage</u>: Disrupting or destroying IT infrastructure
- <u>Intellectual Property (IP) Theft</u>: Stealing sensitive information
- <u>Fraud</u>: Unauthorized data manipulation for personal gain
- <u>Espionage</u>: Spying for competitive advantage or national security purposes

Levels of insider threats

- Low-Level: Unintentional mistakes or careless actions
  - lack of awareness, human error, social engineering attacks
- Medium-Level: Some malicious intent with limited goals
  - disgruntled employees seeking revenge
  - individuals looking for personal gain or pressured by external forces
- High-Level: Significant risk from highly skilled insiders or foreign spies
  - skilled insiders with privileged access,
  - foreign spies

Strategies for Insider Threat Detection

- **Behavior-Based Detection**
  - Focuses on deviations from typical user behavior.
  - Example: *Tracking unusual login times or large file transfers.*
- **Rule-Based Detection**
  - Uses predefined rules and policies for threat identification.
  - Example: *Detecting unauthorized access using predefined access rules in Snort.*
- **Anomaly Detection**
  - Detects unexpected patterns and deviations from norms.
  - Example: *Identifying unusual spikes in data access using clustering algorithms.*

Strategies for Insider Threat Detection

- **Signature-Based Detection**
  - **Matches behaviors against known malicious patterns.**
  - **<u>Example</u>: *Recognizing previously documented malicious commands in network traffic.***
- **Heuristic-Based Detection**
  - **Leverages heuristics from best practices to flag suspicious activities.**
  - **<u>Example</u>: *Flagging unauthorized privilege escalation based on pre-defined rules.***

Traditional Techniques Used for Insider Threat Detection

- **Statistical Techniques**
  - Time-series analysis, clustering, and sequence pattern mining to detect deviations from normal behavior.
- **Machine Learning Techniques**
  - SVM, Random Forests, and K-Means clustering for classifying normal and anomalous behavior.
- **Matching Techniques**
  - Using regular expressions to search for suspicious patterns in log files or network traffic.
- **Deep Learning Techniques**
  - LSTM and autoencoders to capture temporal dependencies and detect subtle insider threats

Limitations of the Traditional Techniques

- Limited Non-linear Detection
- Inability to Detect Novel Threats
- High False Positive Rates
- Overfitting and Limited Generalization
- Vulnerability to Evasion

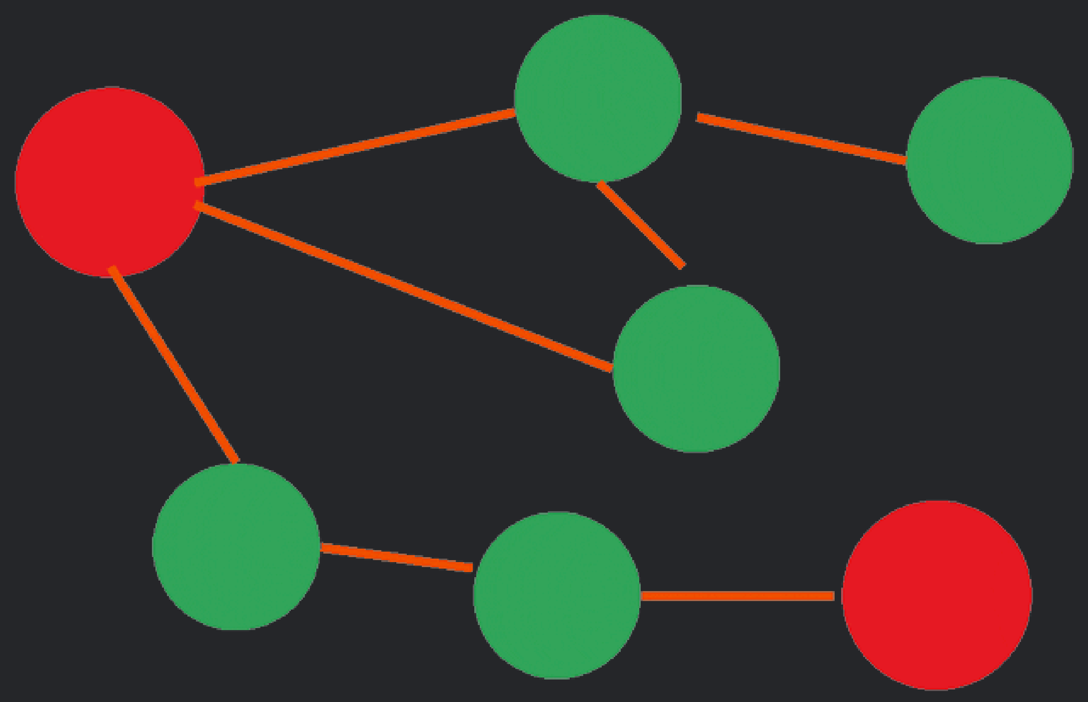
How Graph Neural Networks (GNNs) Can Help?

- **Modeling Complex Interactions:**
  - GNNs effectively capture the relationships between users, devices, and events in graph structures.
- **Adaptive to Evolving Threats:**
  - GNNs learn new patterns dynamically, improving detection of previously unseen attacks.
- **Contextual Insights:**
  - GNNs leverage graph context, enhancing the detection of subtle, insider activities.
- **Reduced False Positives:**
  - GNNs improve accuracy by understanding connections between events, filtering out benign anomalies.

Introduction to Graph Neural Networks (GNNs)

What are Graph Neural Networks?

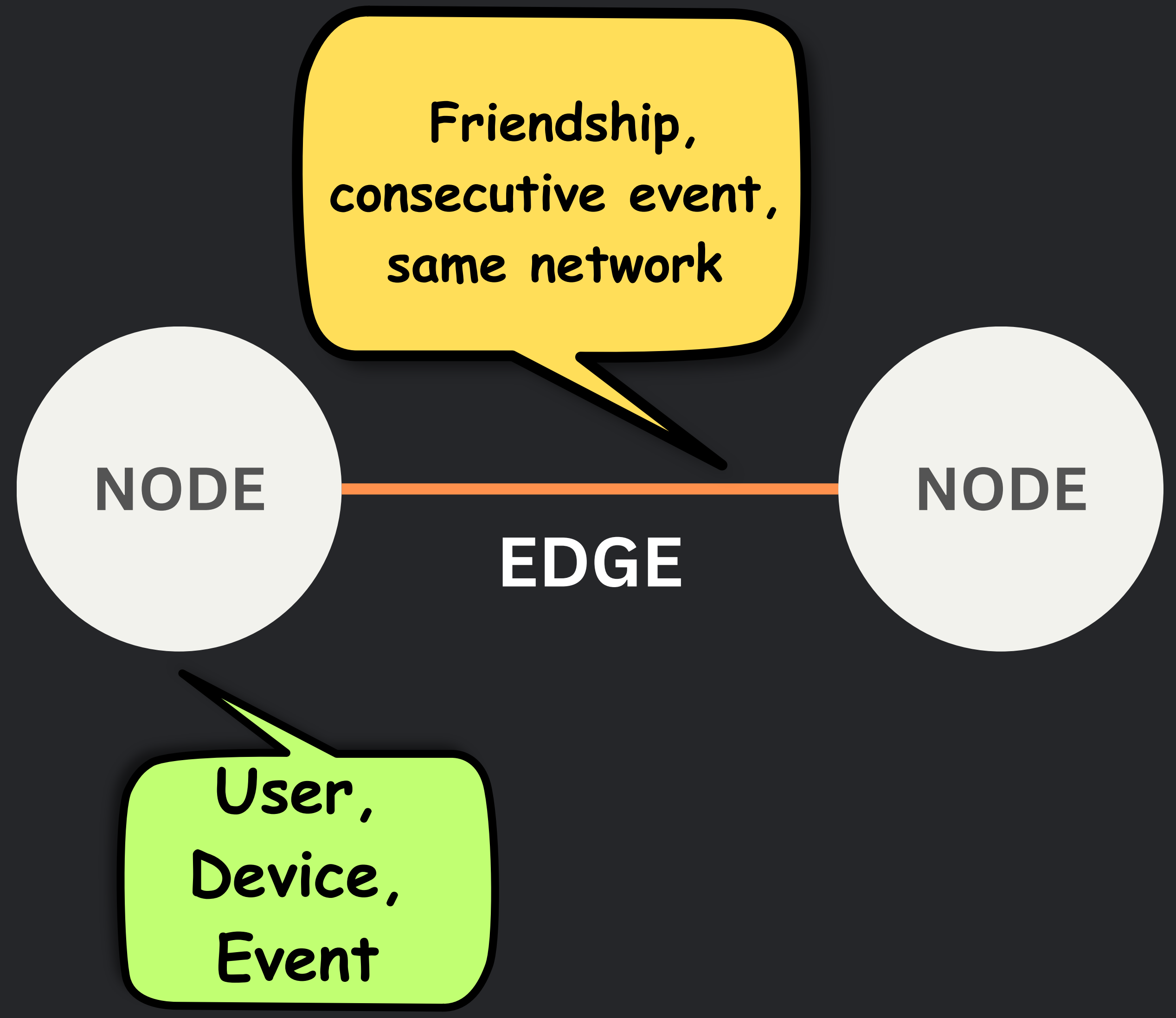
Graph Neural Networks are powerful AI tools that learn from connected data, helping us uncover hidden patterns in complex networks.



What are Graph Neural Networks?

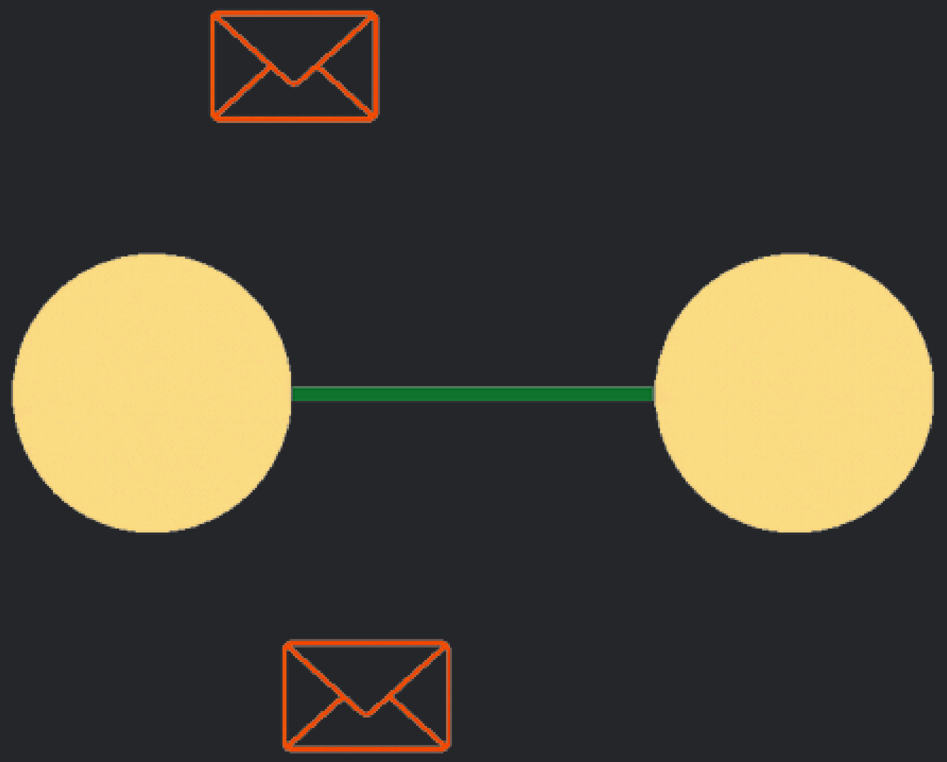
<u>Nodes</u> (also known as vertices) represent entities or objects in a graph.

<u>Edges</u> represent the relationships or connections between nodes.



What are Graph Neural Networks?

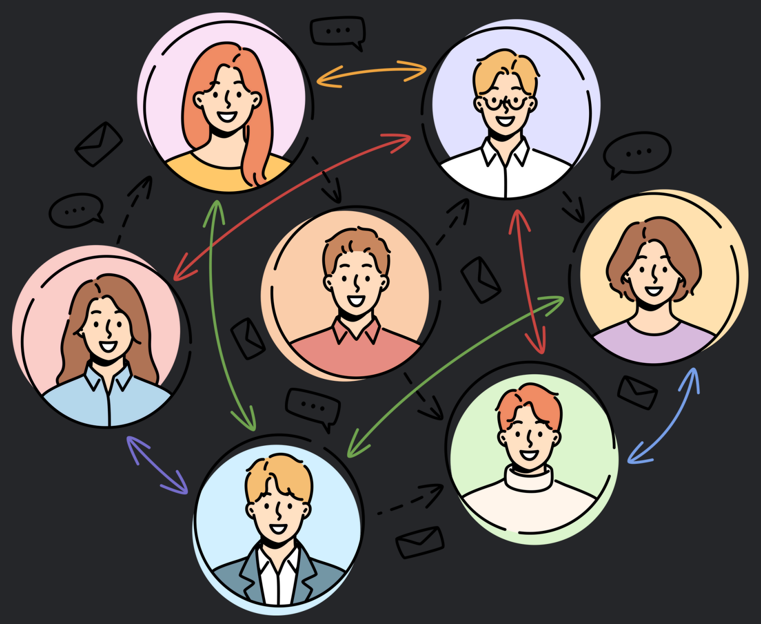
GNNs learn rich node representations, called embeddings using Message Passing.

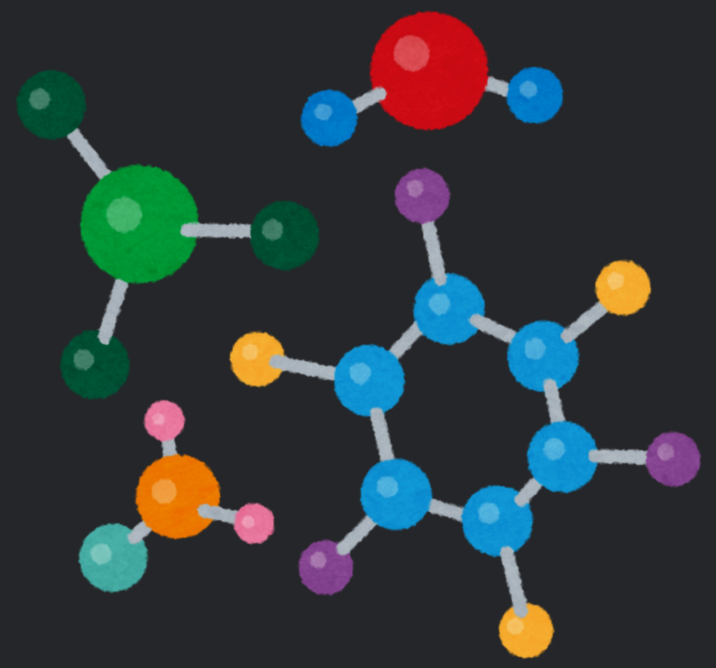


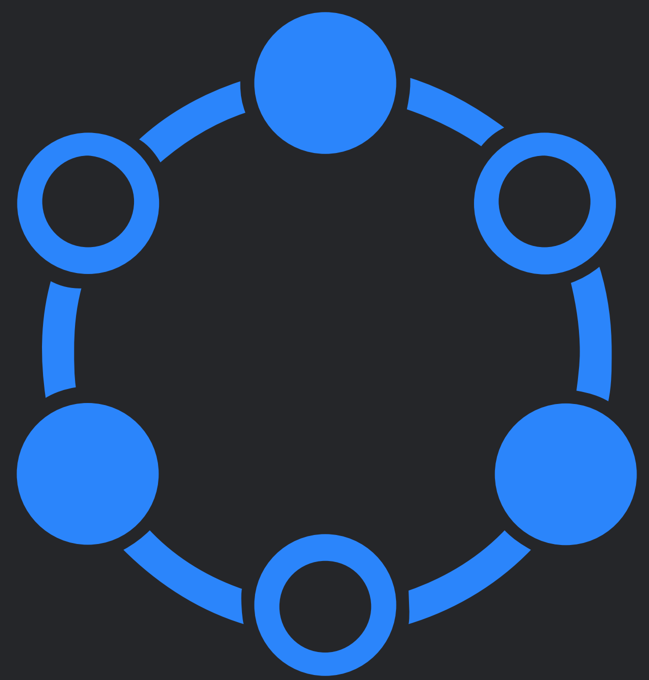
What are Graph Neural Networks?

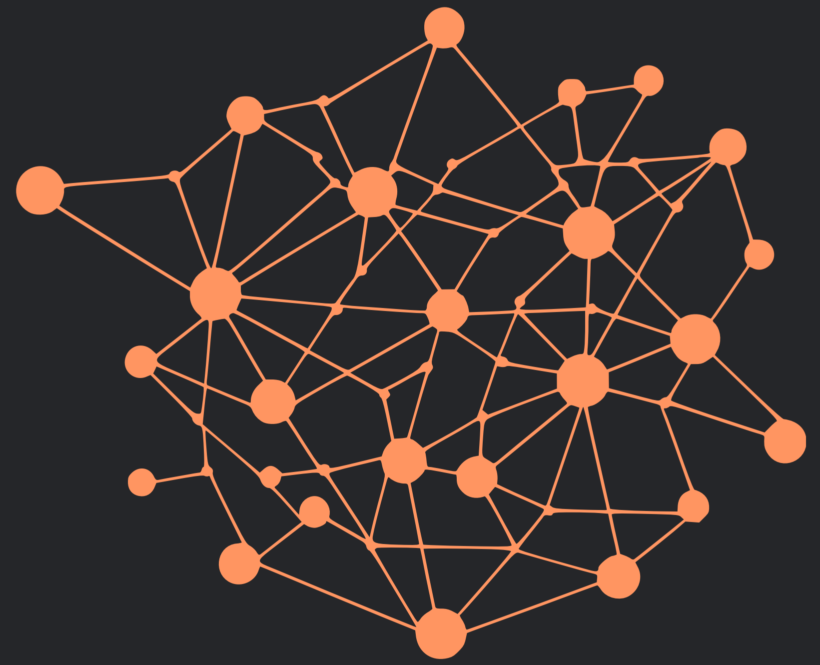
GNNs have found applications in various domains, including:

- **Social network analysis**
- **Molecular property prediction**
- **Knowledge graph completion**
- **Recommender systems**





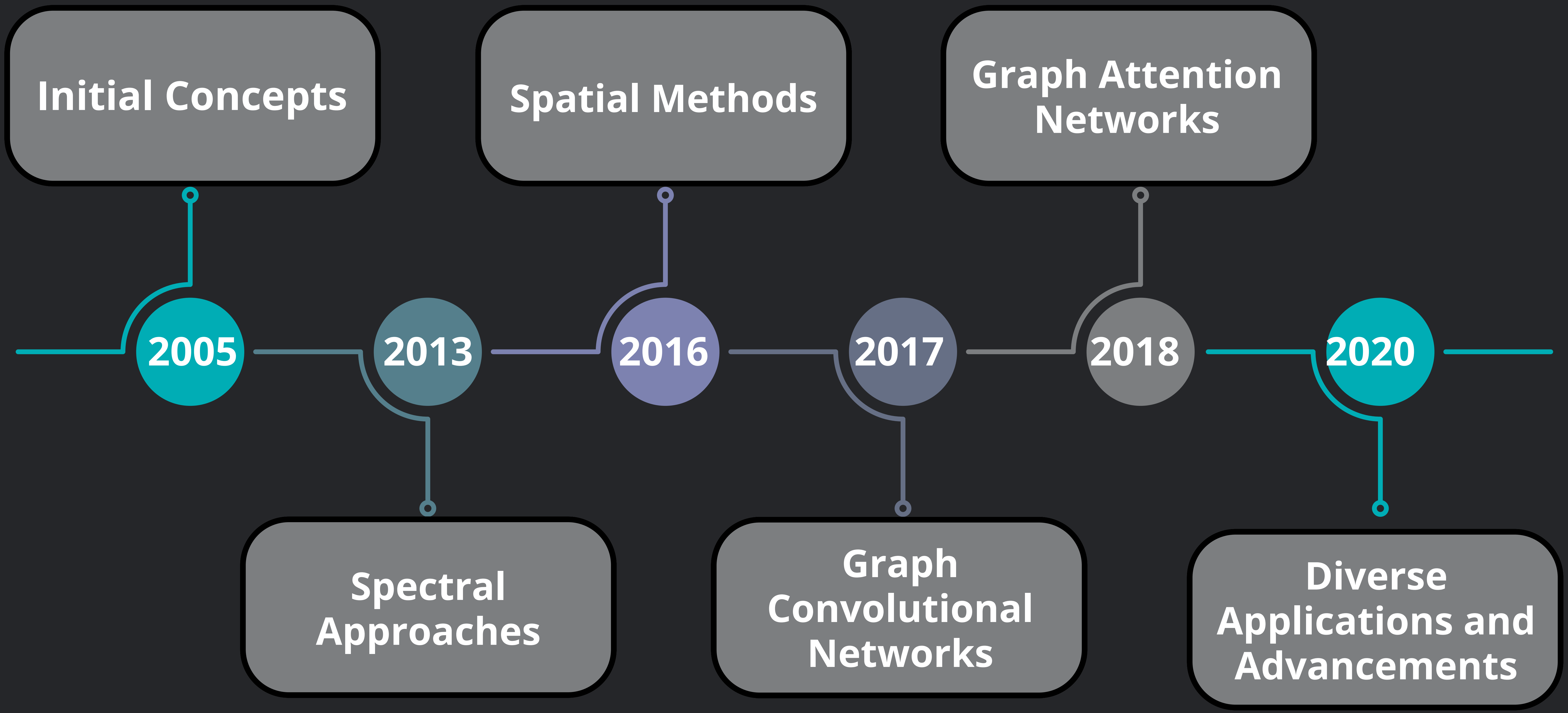




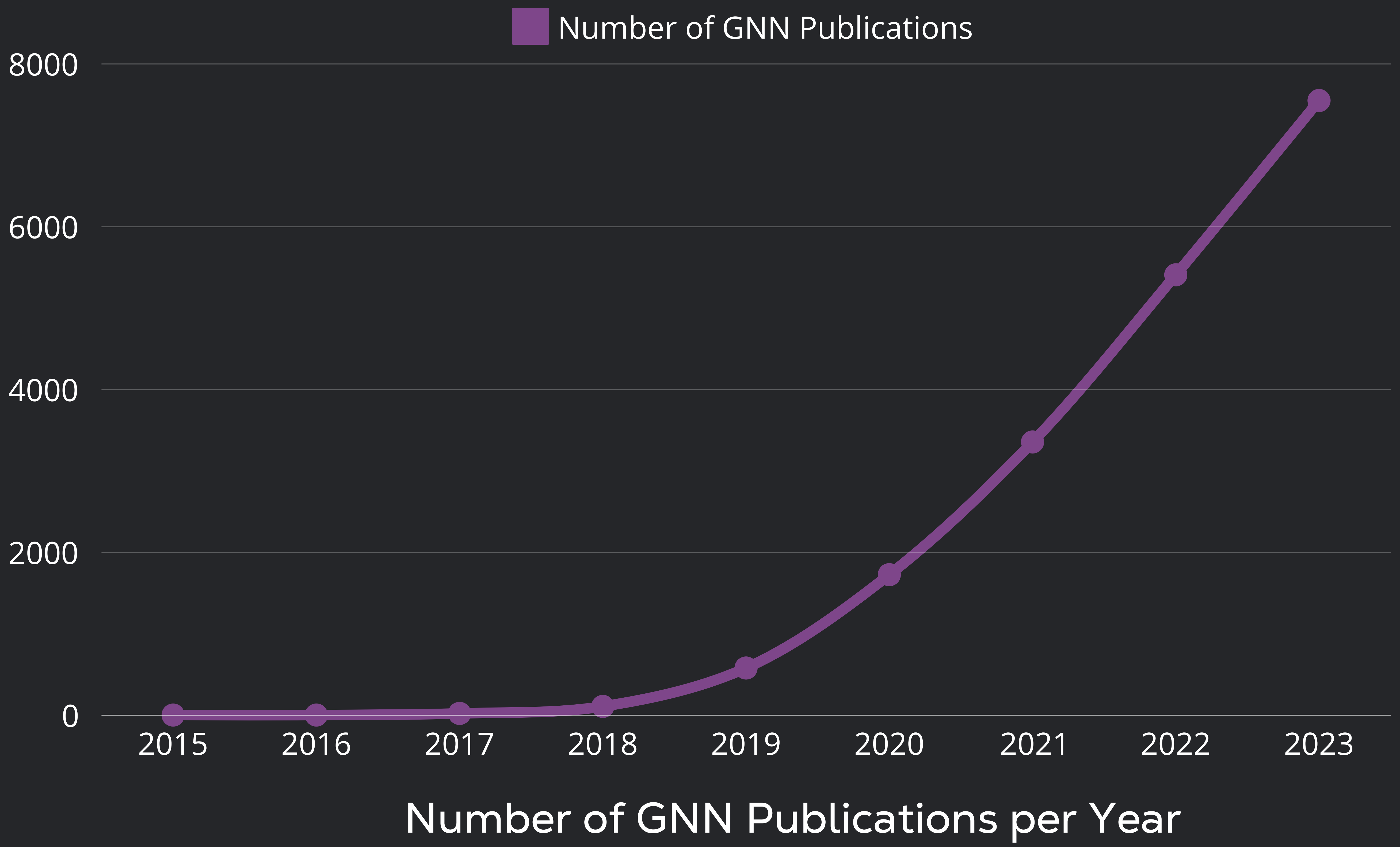
GNNs vs Traditional Neural Networks

Aspect	Graph Neural Networks	Traditional Neural Networks
Input Structure	Graphs with variable size	Fixed-size, grid-like input (e.g., images, sequences)
Relationships	Model learns from relationships between input features	Assumes independence between input features
Node-level Tasks	Node classification, node regression, node clustering	Sample-level classification, regression
Edge-level Tasks	Link prediction, edge classification	Not applicable
Graph-level Tasks	Graph classification, graph regression	Not applicable
Permutation Invariance	Inherently permutation-invariant due to message passing	Requires explicit techniques (e.g., pooling) for permutation invariance
Interpretability	Can provide insights into important nodes, edges, and subgraphs	Often difficult to interpret learned features

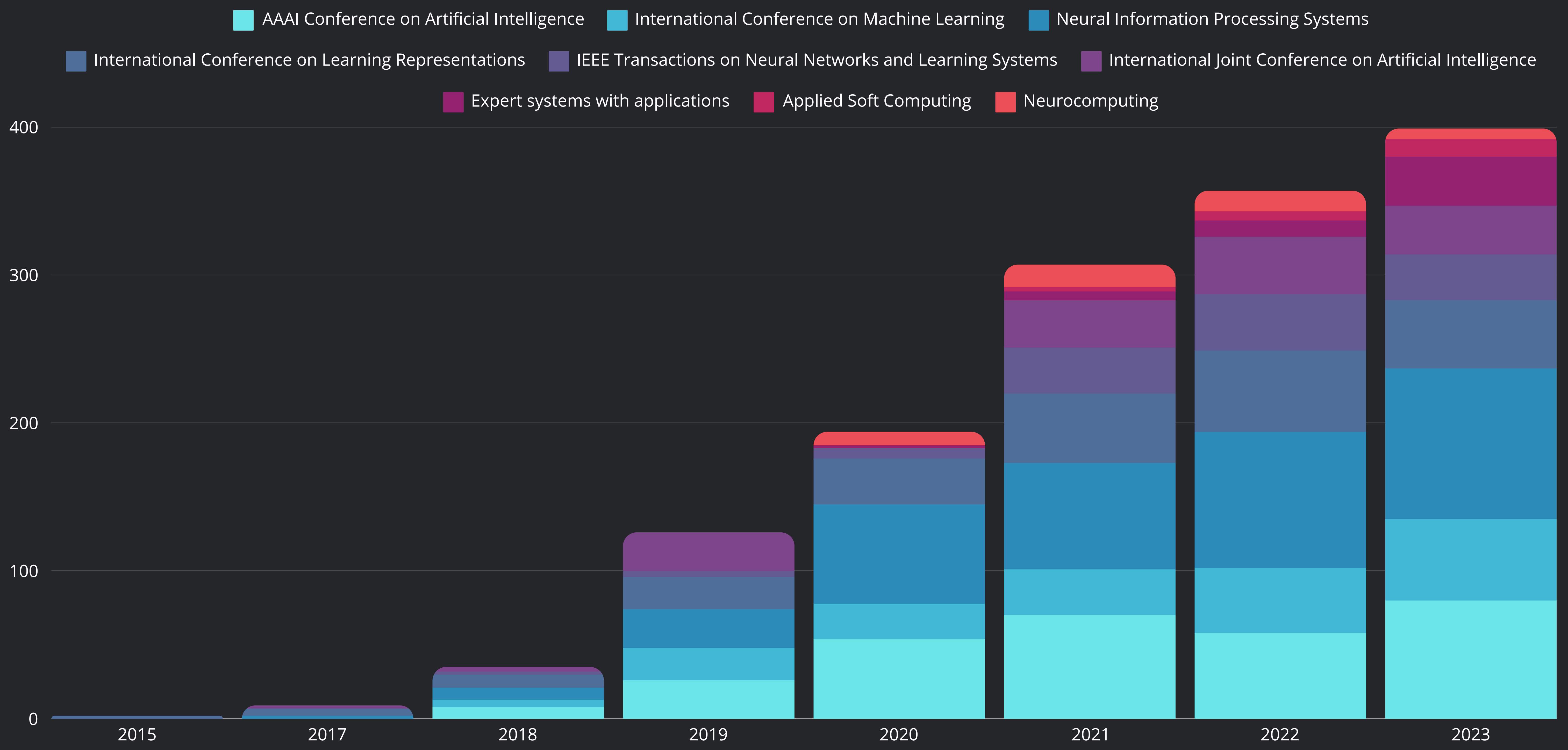
Milestones in GNN Evolution



Milestones in GNN Evolution



Milestones in GNN Evolution

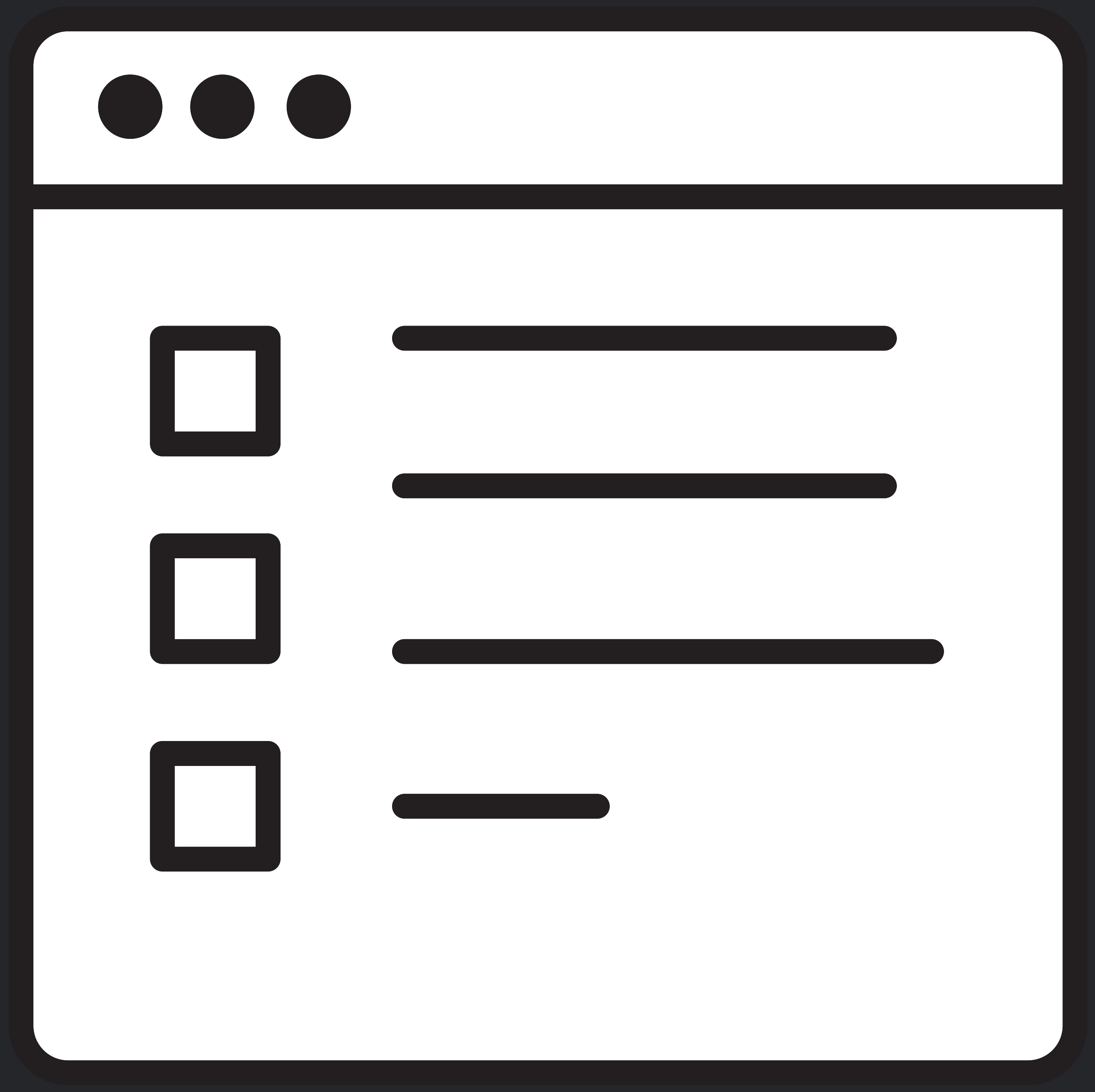


Number of GNN Publications at important conferences/journals per Year

GNN Approach for Insider Threat Detection

Detection Workflow Template

- Data Collection
- Graph Construction
- Extract the node/edge features
- Select the best Graph Neural Network Model
- Train the model
- Detect the anomaly



GCN Based Internal Threat Detection

- Data Collection
  - CMU CERT v4.2 Dataset
- Graph Construction
  - User Behavior-Interaction Graph
- Extract the node/edge features
  - User features are extracted
- Select the best Graph Neural Network Model
  - Graph Convolutional Network
- Train the model
- Detect the anomaly

GCN Based Internal Threat Detection

CMU CERT v4.2 Dataset

- Simulates activity logs of 1,000 users from January 2010 to May 2011
- Types of Logs Collected:
  - Logon/Logoff Activity
  - Email Communication
  - File Access
  - Web Browsing
  - Device Usage

GCN Based Internal Threat Detection

User Behavior-Interaction Graph

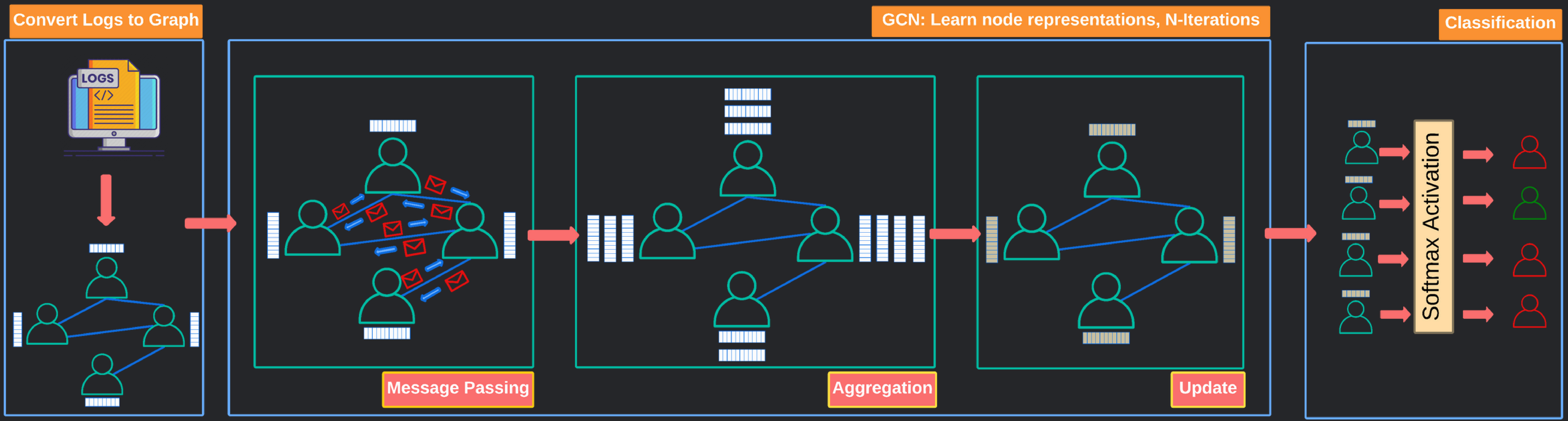
- **Nodes:**
  - Each node represents a user within the network
- **Edges:**
  - Edges represent interactions or similarities between users.
  - Two types of edges:
    - Direct connections: Email communication or shared activities.
    - Similarity-based connections: Behavioral similarity quantified using cosine similarity.

GCN Based Internal Threat Detection

Key User Features

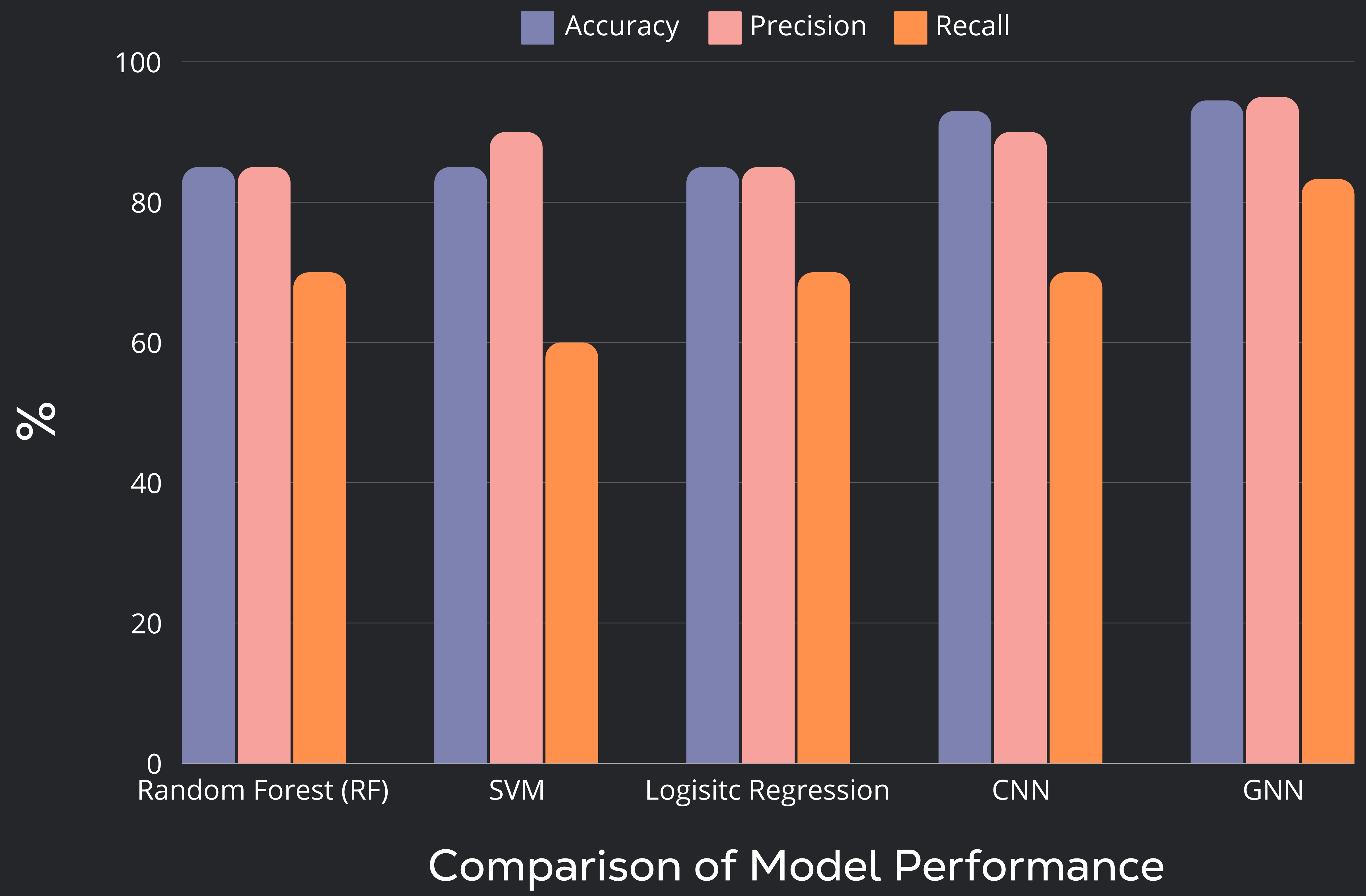
- **Logon/Logoff:**
  - Daily, off-hours logins, devices used
- **Device Usage:**
  - Connections, off-hours activity
- **File Access:**
  - Number and type of files, off-hours, devices used
- **Email:**
  - Sent emails, internal/external, size, topic, sentiment
- **Web Browsing:**
  - Pages visited, WikiLeaks, sentiment, key-logger sites

GCN Based Internal Threat Detection



Model Architecture

GCN Based Internal Threat Detection



Challenges and Future Directions

Challenges

- **Limited Malicious Data:**
  - Real-world insider incidents are rare, leading to data imbalance and insufficient labeled samples for training.
- **Evolving User Behavior:**
  - Employee behavior and interactions change dynamically over time, requiring adaptive models to capture new patterns.
- **Scalability in Large Networks:**
  - Corporate networks have millions of interactions, making GNNs computationally intensive and hard to deploy in real time.
- **Incorporating Low-Interaction Users:**
  - Many employees exhibit few or sporadic interactions, which challenges the graph structure and prediction accuracy.

Future Directions

- **Real-Time Detection Systems:**
  - Develop streaming GNN models to process dynamic, real-time data and adapt to evolving user behaviors.
- **Handling Data Imbalance with Few-Shot Learning:**
  - Integrate few-shot or self-supervised learning techniques to address the challenge of limited labeled malicious data.
- **Scalable Architectures:**
  - Explore distributed or federated GNN architectures for large-scale enterprise networks with millions of nodes and interactions.
- **Explainable AI for Insider Threat Detection:**
  - Design models that provide interpretable insights to security teams, so they understand why specific users or activities are flagged as threats.

Questions?

References

1. https://www.ibm.com/topics/insider-threats
2. https://www.theverge.com/2023/8/21/23839940/tesla-data-leak-inside-job-handelsblatt
3. https://www2.dtexsystems.com/2023ponemonreport
4. Jiang, Jianguo et al. "Anomaly Detection with Graph Convolutional Networks for Insider Threat and Fraud Detection." *MILCOM 2019 - 2019 IEEE Military Communications Conference (MILCOM)* (2019): 109-114.
5. Manoharan, Phavithra et al. "Insider Threat Detection: A Review." *2024 International Conference on Networking and Network Applications (NaNA)* (2024): 147-153.
6. Al-Mhiqani, Mohammed Nasser et al. "A Review of Insider Threat Detection: Classification, Machine Learning Techniques, Datasets, Open Challenges, and Recommendations." *Applied Sciences* (2020): n. pag.