Network Management with Graph Machine Learning: Challenges and Solutions

Yu Wang (Vanderbilt University)  Ramakrishnan Durairajan (University of Oregon)

The exponential growth of network traffic and attacks, combined with operator scarcity, underscores the crucial role of machine learning (ML) for network management (e.g., detecting the onset of a DDoS attack, identifying network congestion, etc.). The application of ML for network performance and security tasks is henceforth referred to as ML4Nets. Today, ML4Nets is applied for a wide variety of network management tasks, including traffic classification, anomaly detection, and beyond [1, 2, 3, 4]. While great progress has been made in solving some of the key management tasks in networking, we posit that all these ML4Nets efforts focus only on the network’s forwarding behavior (e.g., traffic patterns).

Ideally, networks can be conceptualized as graphs, with nodes representing devices (e.g., switches, routers, etc.) and edges denoting links (e.g., fiber-optic cables) or traffic paths [5, 6]. Such a graph-based representation not only captures the forwarding behavior (i.e., intricate traffic interactions) of a network but also the network connectivity (i.e., topological behavior), opening a new dimension for network management. For example, optical topology-aware traffic engineering has been shown to be effective against fiber cuts and flash crowds [7]. This makes graph machine learning (GML) an ideal tool for analyzing traffic patterns, attacks, and automating decision-making processes.

GML learns the node and edge embeddings by iteratively applying message-passing to capture topological patterns and neural transformations to extract task-specific insights from attributes [8, 9, 10]. These learned embeddings can then be applied for various tasks such as cyber-attack detection [11, 12] and traffic restoration [13, 14]. Despite these advancements, two key challenges persist which will be the focus of this talk.

First, a comprehensive exploration of graph construction methods and datasets for communication networks is lacking. Effective GML relies on appropriately constructed graphs tailored to specific applications. We address this gap by reviewing three prevalent graph construction techniques: the original network structure, the line graph capturing link homophily, and the byte-level correlation graph for identifying similar traffic patterns [5, 6, 11, 13, 15]. We analyze their advantages, disadvantages, and applications such as network disintegration, encrypted flow classification, and intrusion detection.

Second, the reliability and explainability of GML predictions require further investigation to achieve network operator buy-in. We review current research on explaining GML predictions [16] and quantifying the outcome uncertainty [17, 18]. Additionally, we apply a graph explanation framework, GNNExplainer [16], to understand traffic classification through flow subgraph and attribute mask analysis.

By addressing these challenges, our vision is to equip the network operators and researcher communities with improved graph construction techniques and a deeper understanding of graph-structured datasets for more effective GML4Nets. Furthermore, by exploring network explanation and uncertainty quantification methods, we strive to enhance the trustworthiness of GML in network management tasks.
References


