Adversarial Attacks in Graph Neural Networks

Graph Neural Networks (GNNs) are bringing the power of deep neural networks to the domain of graph-structured problems. This allows prediction and classification tasks on a variety of interesting problems. Like other neural architectures, they are very vulnerable to adversarial attacks. Such attacks create innocent-looking modifications to the graph — but the implications are drastic classification errors.

Consider the example in the figure. The GNN detects that the graph contains cliques. The edges that make the GNN most confident in its prediction are highlighted. Changing these edges can cause the biggest change in prediction, even though all cliques are unaffected. Clearly, removing any highlighted edge is also a “small” change: after all, it is “only” one edge. But every marked edge has the consequence of making the graph disconnected. In this thesis, we want to explore less drastic changes that still tremendously impact the performance of a trained GNN.

Requirements: Knowledge in Machine Learning and or Deep Learning is advantageous. We will have weekly meetings to discuss the intermediate progress, think together about future ideas, and tackle open questions.

Interested? Please contact us for more details!

Contact

- Lukas Faber: lfaber@ethz.ch, ETZ G60.1
- Zhao Meng: zhmeng@ethz.ch, ETZ G61.3
Detailed Project Outline

We denote the following primary tasks mandatory (on the right side you find a rough estimate for the time that we allocate to the respective task):

- Literature research (⋆)
- Familiarize yourself with the existing project (⋆)
- Select and balance one game mode for online play. (⋆)
- Adapt the game to support an online mode with multiple client devices. (★★)
- Find a way to connect clients over the Internet. (⋆)
- Allow players to form game parties. (⋆)
- Add achievements to provide challenge and motivation. (★★★)
- Add statistics and a leaderboard. (★★)
- Write a report (★★)
- Present your findings. (⋆)

Extensions

Apart from these requirements, we can think of plenty of ways to extend the project with cool features:

- Add a chat during the game setup.
- Develop an AI as an alternative to played against instead of human players.
- Create an online exchange for game mods and settings.
- Develop new game modes.
- Protect fair players from cheaters.
- Extend the game with your own ideas.

The Student’s Duties

- One meeting per week with the advisors to discuss current matters.
- Regular check-ins into the provided revision control system.
- A final report in English, presenting work and results.
- A final presentation (15 min) of the work and results obtained in the project.