Project 3, Understanding Network Mining: Attacking

PROPOSAL

Long Ma
longma2
University of Illinois Urbana-Champaign,
MCS student at Department of Computer Science
longma2@illinois.edu

Haoyu Su
haoyus2
University of Illinois Urbana-Champaign,
MCS student at Department of Computer Science
haoyus2@illinois.edu

1 PROBLEM INTRODUCTION

Deep learning models for graphs have a wide application today, ranging from the analysis of social and shopping networks, biology information collection, to marketing research and management. In recent years, researchers have found that, in contrast to its extensive application ability, there is little research content on the robustness of graph networks, which leads to severe security problems.

In graph networks, relationships between nodes are highly considered. For example, in many classic node classification tasks, some nodes that have higher credit or more connected nodes play a more important role in the network, but also bear a higher risk at the same time.

In this project, we aim to work out and improve an efficient graph adversarial attack for the classical models, in order to improve the original traditional model with higher robustness by studying the methods and models of these attacks in future research. We will mainly focus on classification models for better targeting and efficiency reasons, but we also want to verify that our solution can have a high efficiency on other similar models.

2 PRELIMINARY PLAN

Out of our current understanding of the project and the issues, we have simply divided the project into three stages, the specific ones corresponding to each of them will be divided in more detail in the next report.

Stage 1: $Sep.13^{th} - Oct.13^{th}$

We have three months from $Sep.13^{th}$ to $Dec.12^{ed}$ in total to finish the project. Cause our team is not familiar with graph data, we need sometime to study the basic contents of models based on graph data. We plan to use two weeks to get familiar with the datasets and load these datasets to our projects so we can use them to build our own models. Then, use another two weeks to learn how to build models based on graph data and try to replicate the models attacked in the introductory materials. Thus, we can settle down Step 1: $Sep.13^{th} - Oct.13^{th}$, Systematically learn the knowledge of model based on graph data and read part of the introductory papers, replicate the models attacked in papers.

Stage 2: $Oct.14^{th} - Nov.14^{th}$

Then, we already have a good reservoir of knowledge of graph data. Therefore we should move on to the subject of this project: Attacking. We want to deeply understand the various types of attacks described in the introductory papers. Thus, we need three more weeks to read those papers carefully and try to apply those attacking methods on our own models. Then try to merge the algorithms in the paper and. This is just a preliminary design of our own attack methods, which just need one week. Now we can settle down Step 2:$Oct.14^{th} - Nov.14^{th}$, Learn the attacking methods introduced in the introductory papers and apply them to our own models, try to combine those methods and come up with our own innovations.

Stage 3: $Nov.14^{th} - Dec.12^{ed}$

Finally, we want to use the last month to systematically design our own attack ideas, and implement these ideas and apply them to the models. And compare the performance of the models after being attacked by our own ideas with the original models and the models after being attacked by the methods mentioned in the papers. Summarize the strengths and weaknesses of our own attack methods and try to improve them. If time allows, we can also try to give the corresponding defense mechanism. So, Step 3: $Nov.14^{th} - Dec.12^{ed}$, Design and implement our own attacking methods and analyze their performance, summarize their strengths and weaknesses of them. Try to give the corresponding defense mechanism.

3 DIVISION OF WORK

Stage 1

Our team will learn the knowledge of model based on graph data together and read different papers separately, then organize sessions to share the content and personal understanding of the paper with other teammates. And each person is responsible for replicating the model in the paper he or she is responsible for.

Stage 2

Also, each person is responsible for learning the attacking methods in the paper he or she is responsible for, and apply on over own models. Then, we will discuss these methods together and try to combine them or come up with our own innovations.

Stage 3

In this stage, each of us will brainstorm and come up with our own ideas, then get a final solution through discussion and implement it. And analyze its performance together.

4 TEMPORARY PAPER LIST

Hanjun Dai, Hui Li, Tian Tian, Xin Huang, Lin Wang, Jun Zhu, and Le Song. 2018. Adversarial attack on graph structured data. ICML 2018.

Daniel Zügner, Amir Akbarnejad, and Stephan Günnemann. 2018. Adversarial Attacks on Neural Networks for Graph Data. KDD 2018.

Qinghai Zhou, Liangyue Li, Nan Cao, Lei Ying, and Hanghang Tong: Admiring: Adversarial Multi-Network Mining. ICDM 2019

Scarselli, Franco, Gori, Marco, Tsoi, Ah Chung, Hagenbuchner, Markus, and Monfardini, Gabriele. The graph neural network model. Neural Networks, IEEE Transactions on, 20(1):61–80, 2009.

Su, Jiawei, Vargas, Danilo Vasconcellos, and Kouichi, Sakurai. Onepixelattackforfoolingdeepneuralnetworks. arXiv preprint arXiv:1710.08864, 2017.

Zügner, Daniel, Akbarnejad, Amir, and Günnemann, Stephan. Adversarial attacks on neural networks for graph data. In KDD,2018.