Graph Mining HW

TA: 許竣翔
jordan.hsu@bioproweb.com

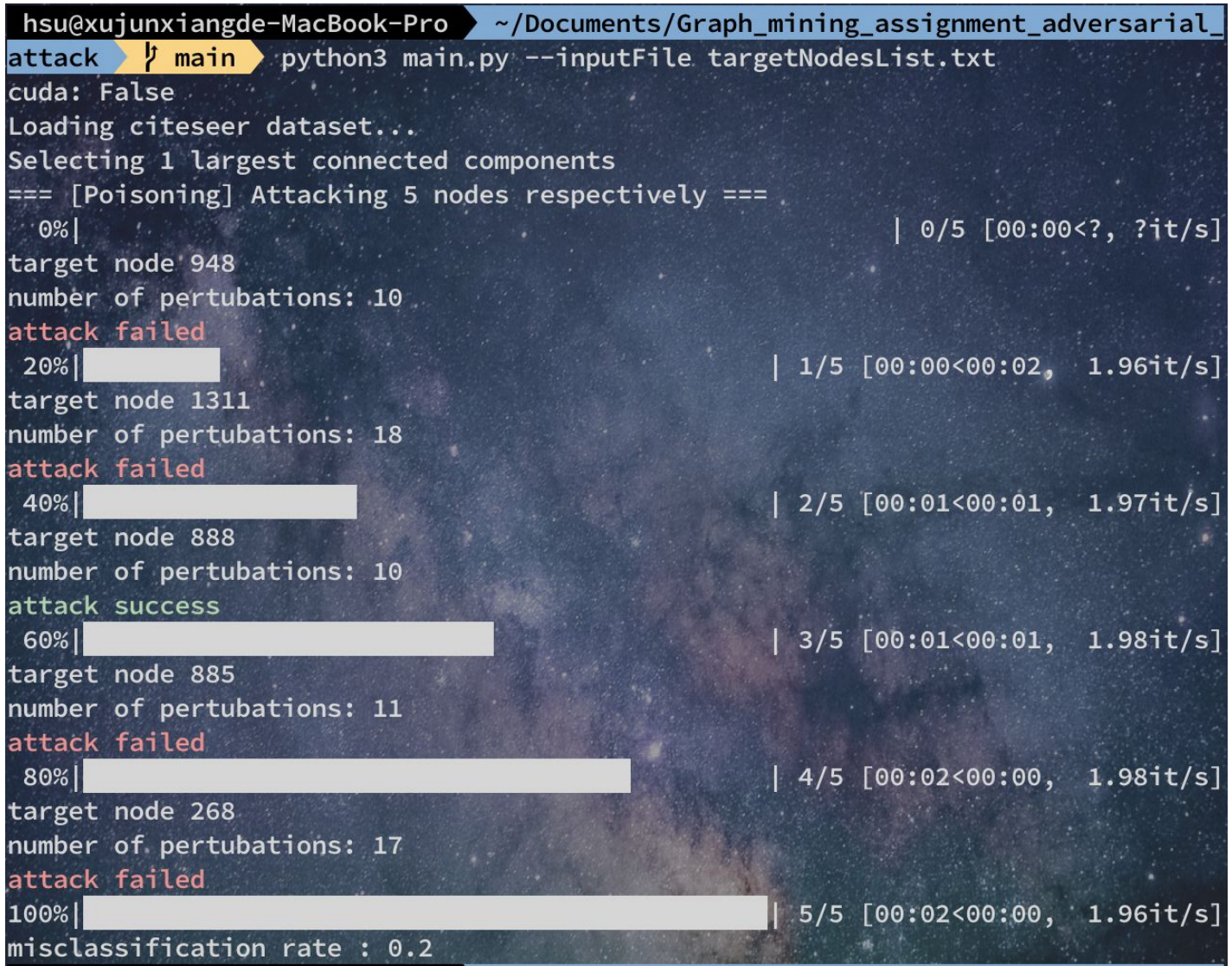
目標

- 實作Nettack：針對一target node，在符合攻擊budget constraint下，生成一modified graph，目標是讓GCN retrain在modified graph上後，target node的分類是錯誤的
- 實作direct & poisoning attack即可
- 不用考慮degree distribution constraint & feature co-occurrence constraint

要求

- main程式檔名取名為你的學號_main.py
- 執行程式時需在後面輸入--inputFile [輸入檔名]
  - 如：python3 你的學號_main.py --inputFile targetNodesList.txt
- 每個target node的攻擊budget = degree of target node + 2 (same as Nettack paper)
- 攻擊edges或攻擊features可以擇一做，也可以兩個都做，但每改一edge或一feature，都算進budget內
  - p.s. 通常攻擊structure效果會比較顯著
- 會限制總執行時間（助教的code執行時間 * 5）
  - 註：python數值計算加速可參考numba python package

範例程式成功跑起來的樣子



作業需要做的事

- attacker.py內目前是一個隨機的攻擊器class，你需要把它改編成Nettack

評分標準

- 共10筆測資（10個target nodes），五筆公開五筆隱藏，五筆公開測資全攻擊成功能得到85分，隱藏測資每成功一個加三分，五筆隱藏測資都攻擊成功即100分
- 如五筆公開測資未能全攻擊成功，但在報告中能**用自己的文字**分析為什麼攻擊失敗，最多可得30分
- 如code內使用了超出規定的攻擊budget數量，則一律以0分計
- Dataset = Citeseer’s largest connected component

繳交和執行環境

- 繳交下列東西（壓成一個zip，交到ilms）:
  - Your code, can be .py or .ipynb (100%)
  - A report, can be PDF or docx (0%，但一定要交，沒做出來時才佔30%)
- 環境
  - OS: Ubuntu 18.04
  - Python version: python 3.7
  - Cpu: Intel i9-9900
  - 顯卡: GeForce RTX 2060
  - Cuda: 11.1
  - 其他用到的Python套件請見requirements.txt

報告

- **不含圖一頁以內，中英皆可**
- You have to describe following things in your report
  - How do you implement the algorithm **by your own words** (don’t copy paper)
  - Anything you have tried

- **If you succeed, providing a screenshot like page 4**
- If you fail, analyzing the possible reasons (but not you are busy)

備註

- 範例code為pytorch-based，Gitlab: https://gitlab.com/warren30815/Graph_mining_assignment_adversarial_attack
- Paper: https://arxiv.org/abs/1805.07984
- Python加速庫numba：https://numba.pydata.org/
- **作業上有問題請發在ilms討論區**
- **Deadline: 6/5 半夜12.15前**