End-to-end learning and optimization on graphs

CZ/CE4071 Network Science

Mid-Progress Report:	[Google Doc] CZ/CE4071 Project 2 Mid-Progress Report
Paper Reviewed (arxiv version):	[1905.13732] End to end learning and optimization on graphs

Overview

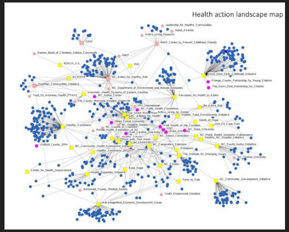
- ❏ Problem Studied in This Paper
- ❏ Most Significant Contributions
- ❏ Connections with CE/CZ4071 Network Science
- ❏ Discussion of Experiment

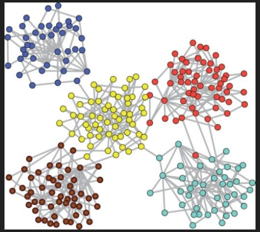
Problem Studied in This Paper

Graph Optimization (GOpt) problems is becoming popular due to:

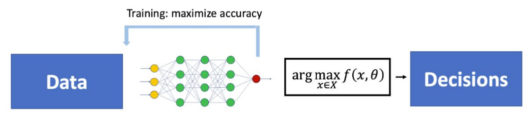
1. Abundance of open-source network datasets
2. Emergence of operational research & deep learning
3. Insufficiency of existed traditional approaches

Example Problems or Tasks:
Cluster Detection, Community Analysis



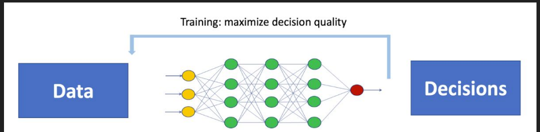


Empirically speaking, real practices of graph problems can combine learning and optimization process. There are usually two approaches (Fig 1.1 & 1.2).



(Fig. 1.1) Two-stage Approach

Learn the mapping through straightforward loss functions, which still needs optimization



(Fig. 1.2) End-to-end Approach

Learn the task output which is the ultimate goal that user wants to accomplish

In this paper, the author proposed an innovative approach to integrate a differentiable proxy for the GOpt clustering layer into the training process of deep learning. This approach is targeted at tasks including Link Prediction and Cluster Detection.

Most Significant Contributions

- Introduced a general framework for graph learning and optimization integration.
    - Utilizes a simpler problem of optimization in continuous space to replace complicated, discrete problem.
    - Simplify the optimization component instead of updating the entire algorithm.
- Proved that clustering layer is differentiable.
    - Allows the solution to be applied in deep learning systems.
    - E.g. The cluster assignments treated as a solution to the discrete problem.
- Showed experimental improvements over both *two-stage baselines* and *end-to-end approaches*.

Connections with Network Science

a) Can the paper's algorithm be used to solve problems in CE/CZ4071?

Problems in CE/CZ4071	More efficient technique mentioned in paper
Cluster Assignment Problem	Cluster nodes into K components select nodes with embedding is close to center of each cluster
Graph Partitioning (E.g. Spectral Clustering)	Graph embedding using the GCN allows for a flexible format to perform graph partitioning

b) More suitable techniques from CE/CZ4071 to solve the paper's problem?

Problem with the paper's technique	Techniques from CE/CZ4071
Does not work on uncertain graphs	Add differentiable Monte Carlo Sampling layer before the GCN node embedding layer.
Disregards feature extraction	Have a deep differentiable random forest layer as the last layer which takes into account results graph pattern mining algorithms (E.g. Pattern growth approach/Frequent Pattern Mining Approach)

Model Structure: ClusterNet

2 layer GCN with dropout followed by differentiable K means clustering

GCN: embed data structures in non-Euclidean space into continuous space

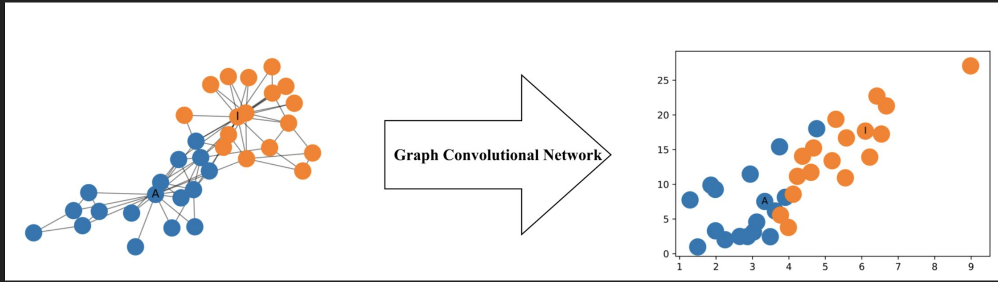
Clustering: soft assignment

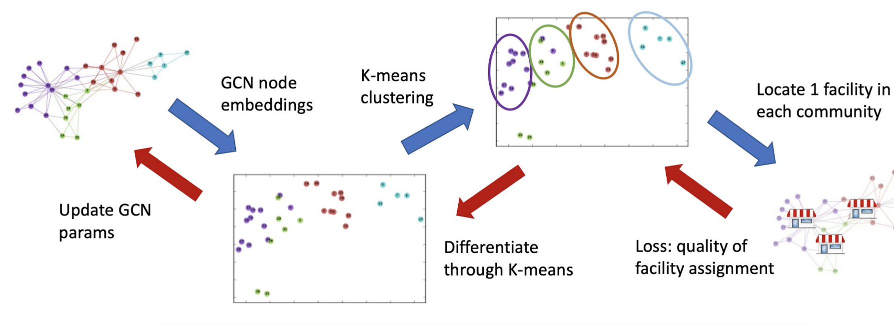
Loss function: quality of clustering

End-to-end learning and optimization

Result: communities and facilities of them

Reduce error





Discussion of Experiment

- Goal: Compare the performance of ClusterNet and other three baseline models on community detection
  - GCN Two-stage Model
  - No link prediction Model
  - GCN End-to-end Model
- Dataset: **WebKB**
  - This dataset contains 877 scientific publications in total.
  - The citation network consists of 1608 links. (WebKB.cites)
  - Each publication in the dataset is described by a 0/1-valued word vector indicating the absence/presence of the corresponding word from the dictionary(WebKB.content)

Download link: *https://linqs-data.soe.ucsc.edu/public/lbc/WebKB.tg*