CS 292F.300 Final project proposal

Sassan Hashemi

TOTAL POINTS

1 / 1

QUESTION 1

1 Proposal submitted 1 / 1

✓ **- 0 pts** **Correct**

This sounds great. I look forward to seeing the results.

gradescope

Final Project Proposal

Info

Name: Sassan Hashemi

Class: Graph Laplacians by John Gilbert

Summary

Both machine learning and blockchains are emerging technologies that show promise in their future impact on our society. Bringing the two fields together enables us to leverage the tools developed by machine learning researchers to gain insight on the information that exists blockchains. CoinBLAS, a platform that facilitates queries on the bitcoin blockchain (with the possibility of more blockchains in the future), along with GraphBLAS, a platform that allows for graph and matrix algorithms, as well as Tensorflow/Pytorch, allow for the analysis of the bitcoin network using ML techniques. By thinking of private keys as nodes and transactions as edges, we are able to represent the blockchain as a graph, enabling the use of GNNs and other tools for analysis.

Technologies

Below are a set of APIs that I can use as a part of my project/thesis. While I may use more than what I've listed, these will form the basis for my project.

GraphBLAS: an API that implements dense and sparse graph algorithms as linear algebraic equations (matrix multiplication, addition, ...).

CoinBLAS: an API that allows for the querying of the bitcoin blockchain, for information such as block information, transaction information, etc. This acts as the dataset for the project/thesis.

Tensorflow/Pytorch: an API that supports many machine learning models and algorithms, specifically those that can be run on graphs such as GNN, GCN, embedding nodes and graphs into vector spaces, etc.

Project

Comparison of methods for community detection on the bitcoin network

Importance: One example of the applications of community detection on the bitcoin network would allow for the development of layer 2 applications (such as the lightning network) that can decide which transactions to put on chain and which transactions to let aggregate off chain in order to optimize for network fees, transaction throughput, or any other metric.

Relevant work:

- One non-ML methods for community detection involves using the permutation required to sort the Fiedler vector of the Laplacian of the graph to permute the Laplacian matrix to put clustered nodes together.
- Another non-ML method includes recursively cutting the graph in a way that minimizes the isoperimetric ratio or the conductance. While there are definitely more methods that I can learn,

these are the one's I'm currently familiar with [John and Ambuj's courses].

Task: Compare the performance of these methods based on different graph structure representation of the blockchain network as well as different metrics. Graph structures could include a bipartite graph with one side being public keys and the other being transactions, or a graph with nodes as public keys and edges as transactions. Metrics could include isoperametric ratios or time/transaction fee savings.

1 Proposal submitted 1 / 1

✓ **- 0 pts Correct**

This sounds great. I look forward to seeing the results.

gradescope