**Portfolio Construction using Graph Sampling**

**Group members**: Yinghao Li, Michael Chu, Junyu Wang

**Name and contact information of project mentor**: Chris Romano <[romano.chris@gmail.com](mailto:romano.chris@gmail.com)>

**Problem statement:**

The traditional Markowitz model, when given too large a dataset, faces practical difficulties, such as numerical instability and sensitivity to perturbations/noisy data. We attempt to solve this problem by curating and filtering through a superset of assets using graph sampling methods.

**Approaches:**

Using publicly available equity data as the superset, we will:

* Construct stock network using similarity measures (e.g. correlation) for the superset
* Down-sample the stock network defined by such similarity scores
* Further sift through the down-sampled network using graph clustering algorithms, such as k-spanning tree, highly connected subgraph, maximal clique and kernel k-means clustering.
* Construct portfolios using the down-sampled set of stocks, and compare its performance with the market portfolio

**Criteria of success:**

* Gain insight into the equity market, with the help of graph clustering algorithms. For instance, the centrality of the equities, so we can reduce correlation and reduce diversity.
* Down-sample a subset of stocks that captures the salient graph/network properties from our predefined “stock universe”
* Construct a sparse portfolio with low tracking error and information ratio on par with a common index, e.g. S&P500.

**List of references:**

Tse Chi-Kong, Liu Jing, Lau Chung-Ming: A Network Perspective of the Stock Market

Gustavo Peralt, Abalfazl Zareei: A Network A Network Approach to Portfolio Selection

​​F. Pozzi, T. Di Matteo, T. Aste: Spread of risk across financial markets: better to invest in the peripheries

Jure Leskovec, Christos Faloutosos: Sampling from Large Graphs

Yu Wang, Bortik Bandyopadhyay, Aniket Chakrabarti, David Sivakoff, Srinivasan Parthasarathy: Spread Sampling for Graphs: Theory and Applications

Olivier Ledoit, Michael Wolf: Honey, I Shrunk the Sample Covariance Matrix