# Projects Proposal

Keyue Jiang

## 0. Outlines

- Graph Signal Processing on Dynamic graphs

- (*)GSP for balancing Node Attribute and Topological Structure

- GNN design with CSBM models

- Graph Signal Processing on Multi-relational Graphs?

- Multi-Agent AI with Graph-structured data, or Network Games?

- Graph Bandits?

## 1. Asynchronous Message Passing Networks for Dynamic Graphs

(More explaination in file 2)

While most graph machine learning methods have been targeted at static graphs, many important real-world graphs are dynamic – they change with time. For example, infectious diseases spread along with the changing graph of interpersonal contacts. Graphs of social networks and e-commerce also change as new members join, or new data becomes available.

In the project, we argue that the paradigm of most existing Graph Neural Networks (GNNs), so-called synchronous message passing process [4], is not ideal for modeling dynamic graphs. Thus, we then propose to use Asynchronous Message Passing (AMP) [3] models instead to overcome the challenges in dynamic graph machine learning. We believe the AMP models can better model the time-varying nature and solve the information redundancy problem of dynamic graphs.

## 2. Balancing the Information from Node Attribute and Topological Structure in Graph Neural networks

The message-passing framework is the foundation of many graph neural networks (GNNs). A typical message-passing framework usually consists of two parts - Transformation (T) and Propagation (P) [6], respectively responsible for capturing the node attributes information and the structural information. The node information and structural information can have different roles in the homogenous/heterogeneous graphs, while in most cases we will not have prior information about which aspect should we use for prediction. Thus, building a GNN model without a special design usually cannot achieve satisfactory performance. Graph signal processing (GSP) is a useful toolkit to infer graph structure from the graph signals (i.e. node attributes). In this project, we wish to use the methods in graph signal processing to balance the utility of node attributes and structural information. This project aims to develop a mechanism to measure the information from node attributes and topological structures, and consequently design novel GNNs based on this mechanism.

## 3. Multi-task Graph Representation Learning with Agent Random Walk

Graphs are prominent tools to model relational data in many domains, such as traffic networks, social networks and protein discovery. In many of these applications, the success of algorithms is attributed to recognizing the presence or absence of specific substructures for a specific task, e.g. atomic groups in the

case of molecule and protein functions, or cliques in social networks. However, current works based on message-passing networks are spending intensive computation to travel through all possible paths in the graphs by aggregating information from neighbours for every node. Thus, it is natural to think, whether we can avoid this redundant computation but just extract the subgraph instead.

Recent work considers this problem with a new paradigm called agent-based networks, which could directly mine the subgraph in the whole graph and learn the embeddings while walking around the graph. Inspired by this, the project aims to use a similar method to solve multi-task learning or meta-learning on graphs. Specifically, an agent will be designed to walk the graphs with task-specific guidance, so it could be capable to make different decisions while walking on the graph. The difficulty lies in how to design the strategy of walking and how can the agent efficiently extract the information from the graphs. (E.g. Goal-conditioned RL methods could be considered)

## 4. Contextual Stochastic Block based Graph neural networks

Graph Neural Networks (GNNs) have shown their competency in encoding graph-structured data. However, the explainability of GNNs is not always coherent with the model, and most of the existing models gain explainability instance-wise - they provide input-dependent explanations for each input graph. This project aims to construct a novel type of graph neural networks that can gain model-level explainability and expressiveness by combing Contextual Stochastic Block Models (CSBM) and slot attention networks into GNNs.

Contextual Stochastic Block Model (CSBM) [1, 2] is a widely used tool in the research of graph theories with applications like spectral clustering and community detection. CSBM assumes that each node $\mathcal{V}_i$ in graph $\mathcal{G}$ belongs to a block (i.e. community) $\mathbf{B}_j$, and the blocks are internally connected to each other. The idea is to use a framework similar to CSBM to construct graph neural networks to gain explainability and expressiveness by iteratively refining the graph information using CSBM. The difficult part would be how to incorporate community discovery into GNNs. A method could be using slot attention networks or capsule networks, but there are always other possible solutions.

## 5. Associating Graphs and Natural Language with Multimodal Deep Learning

Multimodal learning is a technique aiming to combine information from different resources, and is gaining more popularity in the field of image-text translation recently thanks to the giant success of large cross-modal pretrained models like CLIP, DALL-e. With this inspiration, it is natural to wonder if we can do a similar job in the community of graphs.

This project will first consider a novel cross-modal task - the translation between graphs and Natural Language. We will start with generating molecule descriptions based on molecule graphs as the dataset can be easily retrieved. And the next step could be generating graphs from description tasks, which is a much more complicated task. Finally, we will build a large pre-trained model cross the modalities between graphs and natural language to achieve better knowledge representation and better performance on several downstream tasks, like property prediction, drug discovery or protein design.

## 6. Graph Bandits

I was recently getting fascinated by online learning and bandit problems. I saw Kaige's paper [5] about graph bandits and felt its setting to be very interesting. I was just curious about whether this could be a promising direction and if there is more to do on the topic [5]?

# References


[1] Binkiewicz, N., Vogelstein, J. T., and Rohe, K. Covariate assisted spectral clustering. *CoRR abs/1411.2158* (2014).

[2] Deshpande, Y., Sen, S., Montanari, A., and Mossel, E. Contextual stochastic block models. In *NeurIPS* (2018), pp. 8590–8602.

[3] Faber, L., and Wattenhofer, R. Asynchronous neural networks for learning in graphs. *CoRR abs/2205.12245* (2022).

[4] Papp, P. A., and Wattenhofer, R. An introduction to graph neural networks from a distributed computing perspective. In *ICDCIT* (2022), vol. 13145 of *Lecture Notes in Computer Science*, Springer, pp. 26–44.

[5] Yang, K., Toni, L., and Dong, X. Laplacian-regularized graph bandits: Algorithms and theoretical analysis. In *AISTATS* (2020), vol. 108 of *Proceedings of Machine Learning Research*, PMLR, pp. 3133–3143.

[6] Zhang, W., Sheng, Z., Yin, Z., Jiang, Y., Xia, Y., Gao, J., Yang, Z., and Cui, B. Model degradation hinders deep graph neural networks. In *KDD* (2022), ACM, pp. 2493–2503.