Data Science and Analytics Thrust, Information Hub
The Hong Kong University of Science and Technology (Guangzhou)

香港科技大学(广州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

AI-DRIVEN BLOCKCHAIN-BASED APPROACH FOR CARBON MARKET MECHANISM

Graph Neural Networks for Spatial-Temporal Analysis and Prediction of Carbon Emissions in China's Provinces

Prime Supervisor: Prof. Lei CHEN
Co-Supervisor: Prof. Jia LI
Project Mentor: Dr. Jenny, Beijinni LI

Presented by Mingze Gong on March 31, 2023

OUTLINE

香港科技大学(广州)
THE HONG KONG
UNIVERSITY OF SCIENCE AND
TECHNOLOGY (GUANGZHOU)

PROJECT OVERVIEW

香港科技大学(广州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

Bridging Individual and Group Efforts

- My role: Spatial-temporal analysis and prediction of carbon emissions.
- Targets: Accurate and efficient emission predictions to inform the market mechanism.
- Contributions: Enhance decision-making and optimize carbon trading strategies.
- Relationship: Provide data-driven insights for the group project

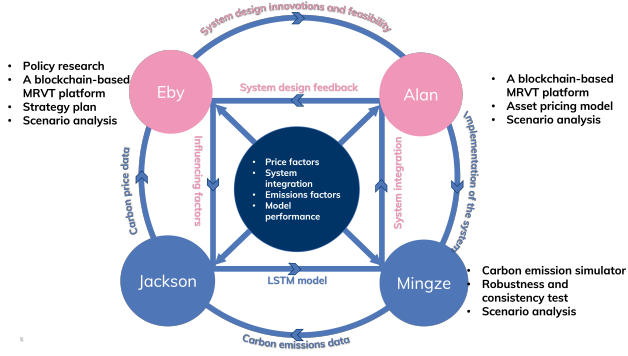

Figure 1: Team Collaboration Chart

INTRODUCTION TO INDIVIDUAL PROJECT

香港科技大学(广州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

Carbon Emissions Prediction

Background

- Disrupted balance between "growth and consumption"
- The necessity of limiting global warming to 1.5℃ above pre-industrial levels
- China's determination to achieve sustainable development

Significance

- Valuable insights for various stakeholders
- Climate change modelling
- Governments' and public's awareness

OUTLINE

香港科技大学(广州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

LITERATURE REVIEW

香港科技大学(广州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

Various purposes

- China's transportation carbon emissions and influencing factors [1]
- The relationship between land use and carbon emissions [2]
- Spatial pattern of carbon emissions in Nanjing [3]
- The relationship between economic development and carbon emissions [4]
- ...

LITERATURE REVIEW

香港科技大学(广州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

Plenty of Models

Statistical Methods

- *STIRPAT model* for emissions in Qingdao [5]
- *Grey forecasting model* for small sample data [6]
- A *dynamic time-delay discrete grey forecasting model* for time-lag effects [7]
- *ARIMA model* for emissions in four representative provinces and cities [8]
- *A novel fractional grey Riccati model* for emission prediction in US, China and Japan[9]

Machine Learning Methods

- *SVM-ELM model* to forecast emissions in specific regions [10]
- A *BP neural network model* for emissions in Beijing [11]
- *PCA and regularized extreme learning machine* in a hybrid prediction model for China's emission [12]
- *ENN model* for emissions in China [13]
- *A hybrid Fast Learning Network–Chicken Swarm Optimization model* for emissions in Guangdong [14]

LITERATURE REVIEW

香港科技大学(广州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

Graph Neural Networks (GNNs)

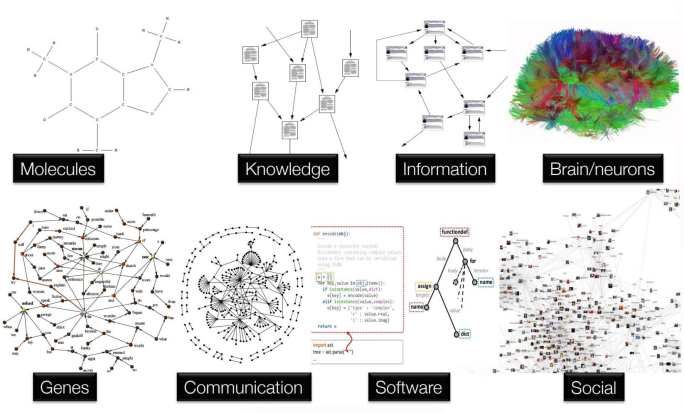

Figure 2: Knowledge from many fields of science and industry can be expressed as graphs [15]

Two types of application scenarios [16]

- Structural data
  - Graph matching, clustering etc.
  - Traffic, knowledge graph etc.
- Non-structural data
  - Image
  - Text

OUTLINE

香港科技大学(广州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

RESEARCH OBJECTIVES AND QUESTIONS

香港科技大学(广州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

Main Objectives

1. To construct a comprehensive carbon emission dataset for training and testing the GNN model.
2. To investigate the potential of GNNs in predicting carbon emissions.
3. To identify the key factors and variables that influence carbon emissions and can be effectively incorporated into the GNN model.
4. To develop a robust GNN model that accurately predicts carbon emissions and outperforms traditional forecasting statistical methods and other machine learning methods.
5. To provide insights and recommendations for policy-makers and stakeholders based on the GNN model's predictions to facilitate effective carbon emission reduction strategies.

RESEARCH OBJECTIVES AND QUESTIONS

香港科技大学(广州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

Research Questions

1. What are the most influential variables and features that should be incorporated into the GNN model to enhance its prediction accuracy?
2. How can GNNs be applied to predict carbon emissions effectively, considering the non-linearity characteristics of carbon emission data and related features?
3. How does the performance of the GNN model compare with traditional forecasting methods in predicting carbon emissions, and what are the advantages and limitations of using GNNs in this context?
4. How can the predictions generated by the GNN model be used to inform policy-makers and stakeholders in designing and implementing effective carbon emission reduction strategies?
5. Are there any potential challenges or limitations in using GNNs for carbon emission prediction, and how can these be addressed to improve the model's performance and reliability?

OUTLINE

香港科技大学(广州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

SUMMARY OF ACCOMPLISHED WORK

香港科技大学(广州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

A Comprehensive Literature Review

Methods		A feedforward neural network (FLN)	NNEnsemble	A feedforward neural network
Paper		Carbon emission forecasting and scenario analysis in Guangdong Province based on optimized Fast Learning Network	Modeling and spatio-temporal analysis of city-level carbon emissions based on nighttime light satellite imagery	Electricity production based forecasting of greenhouse gas emissions in Turkey with deep learning, support vector machine and artificial neural network algorithms
Date & Journal & Impact		2021, Journal of Cleaner Production, 9.297 (2020)	2020, Applied Energy, 9.746 (2020)	2021, Journal of Cleaner Production, 9.297 (2020)
Challenges	Application Scenario	1.Very few carbon emission-related research in Guangdong province focused on trend prediction and scenario analysis of carbon emission. 2.The search for carbon peak and neutrality is rare.	1. The utilization of night-time light satellite imagery data in urban-level carbon emissions was rarely considered. 2. A general framework and machine-learning based software is highly demanded for obtaining the relationship between carbon emissions and NSL data in various regions to explore the dynamic characteristics of carbon emissions at urban scales.	1. The number of studies related to forecasting Greenhouse Gases for Turkey is very-limited and only a few studies predicts it based on electricity production. 2. Carbon emissions gets more focus from researchers but other greenhouse gases also have higher global warming potential than that of carbon emissions.
	Data	1. Existing research mostly focuses on energy-related carbon emissions without considering carbon emissions from cement production and forest carbon sinks. 2. Carbon emissions data needs to be calculated first.	1. Carbon emission data at urban scales is difficult to estimate because of missing values of statistical variables. 2. Carbon emissions data needs to be calculated first based on IPCC formula.	
	Models	1. Compared to hybrid models, single prediction model is always ineffective in its prediction accuracy due to the nonlinear characteristics of carbon emissions.	1. Simple regression methods cannot accurately quantify the relationship between carbon emissions and night-time light satellite data.	1. DL and SVM algorithms were rarely used in Turkish greenhouse gases prediction.
Based on challenges above	What's wrong there	1. Carbon emissions trend prediction and senario analysis are limited by the number of factors introduced. 2. The reliability of carbon emissions estiamted in the paper could be improved. 3. Time-lag effect between carbon emissions and its input factors such as foreign investments is not considered.	1. The differences and influence in lighting features between urban and rural areas were ignored. 2. The reliability of carbon emissions estiamted in the paper could be improved.	1. The fact that Turkey has turned its energy to renewable might make the usage of electricity production for Green House Gases emissions prediction fail to train the model.
		1. **Investigate more influencing factors and transform them via PCA or other methods**. The percentage of errors will definitely increase as more factors are included but it is ok if our purpose is more on trend		

Table 1: A partial screenshot of the summary based on comprehensive literature review regarding emission prediction

SUMMARY OF ACCOMPLISHED WORK

香港科技大学(广州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

Initial Thoughts

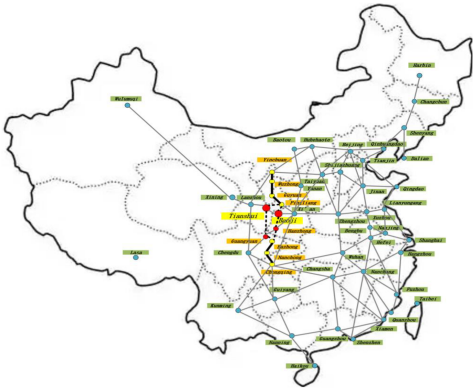

Figure 3: A figure showing the connection between each province in China [17]

Two approaches for similarities measurement

- Features/Factors-based Approach
  - Node features
  - Edge features
- Index-based Approach
  - Euclidean distance
  - Jaccard index

SUMMARY OF ACCOMPLISHED WORK

香港科技大学(广州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

Hands-on Experience

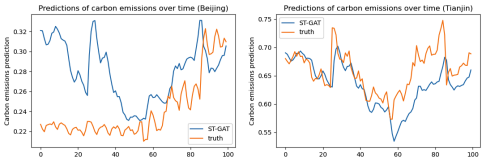

Figure 4: STGAT Baseline Model : Partial Performance (Beijing and Tianjin) in Predicting Carbon Emissions.

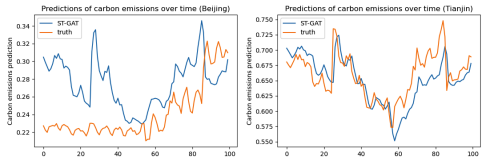

Figure 5: STGAT Model with GDP Feature: Partial Performance (Beijing and Tianjin) in Predicting Carbon Emissions.

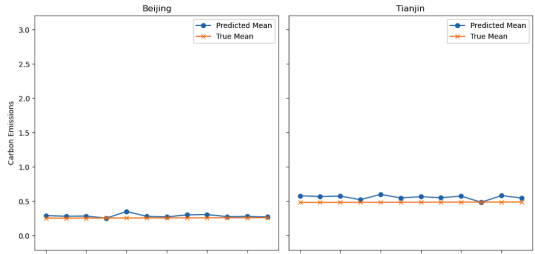

Figure 6: AGCRN Model: Partial Performance (Beijing and Tianjin) in Predicting Carbon Emissions.

OUTLINE

香港科技大学(广州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

RESEARCH PLAN

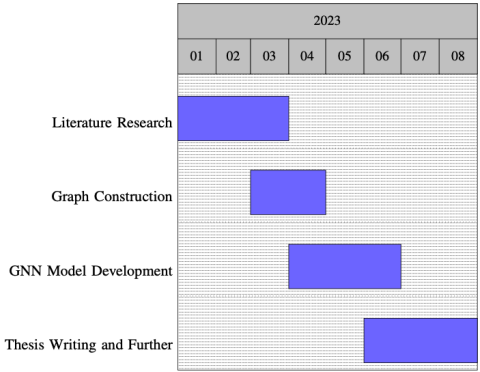

Figure 7: Gantt chart for my research plan.

- 2023.01-2023.03
  - Literature Research
  - Data Preprocessing
- 2023.03-2023.04
  - Codes Testing and Feature Selection
  - Graph Construction
- 2023.04-2023.06
  - GNN Model Development, Training, Evaluation
  - Models Comparison
- 2023.06-2023.08
  - Thesis Writing and Further

OUTLINE

香港科技大学(广州)
THE HONG KONG
UNIVERSITY OF SCIENCE AND
TECHNOLOGY (GUANGZHOU)

POTENTIAL CHALLENGES AND SOLUTIONS

香港科技大学(广州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

Anticipated Issues and Potential Solutions

Anticipated Issues

- Limited existing research on GNNs for carbon emissions prediction.
- Scarce high-dimensional GNN models built upon multiple edge features.
- Data quality and availability.
- Computational complexity.

Potential Solutions

- An extensive literature review on related fields where GNNs have been applied.
- Thorough data preprocessing and cleaning to ensure the consistency and accuracy.
- GNN model optimization for computational complexity reduction.

References

[1] J. Cai, s. ma, h. ji, w. jiang, and z. bai, "Spatial-temporal characteristics and decoupling effects of China's transportation carbon emissions," In Review, Preprint, 2022. DOI: 10.21203/rs.3.rs-1632059/v1. (visited on 02/05/2023).

[2] H. Huang and J. Zhou, "Study on the Spatial and Temporal Differentiation Pattern of Carbon Emission and Carbon Compensation in China's Provincial Areas," *Sustainability*, vol. 14, no. 13, p. 7627, 2022, ISSN: 2071-1050. DOI: 10.3390/su14137627. (visited on 03/18/2023).

[3] Z. Wu, L. Zhou, and Y. Wang, "Prediction of the Spatial Pattern of Carbon Emissions Based on Simulation of Land Use Change under Different Scenarios," *Land*, vol. 11, no. 10, p. 1788, 2022, ISSN: 2073-445X. DOI: 10.3390/land11101788. (visited on 02/05/2023).

[4] S. Yin, C. Bai, and W. Xiao, "Spatial–temporal evolution of the allometric growth between economic scale and carbon emissions in China's cities and its driving mechanism," *Environment, Development and Sustainability*, 2022, ISSN: 1387-585X, 1573-2975. DOI: 10.1007/s10668-022-02708-6. (visited on 02/05/2023).

[5] C. B. Wu, G. H. Huang, B. G. Xin, and J. K. Chen, “Scenario analysis of carbon emissions’ anti-driving effect on Qingdao’s energy structure adjustment with an optimization model, Part i: Carbon emissions peak value prediction,” *Journal of Cleaner Production*, vol. 172, pp. 466–474, 2018, ISSN: 0959-6526. DOI: 10.1016/j.jclepro.2017.10.216. (visited on 03/17/2023).

[6] D. Ju-Long, “Control problems of grey systems,” *Systems & Control Letters*, vol. 1, no. 5, pp. 288–294, 1982, ISSN: 0167-6911. DOI: 10.1016/S0167-6911(82)80025-X. (visited on 03/17/2023).

[7] L. Ye, D. Yang, Y. Dang, and J. Wang, “An enhanced multivariable dynamic time-delay discrete grey forecasting model for predicting China’s carbon emissions,” *Energy*, vol. 249, p. 123 681, 2022, ISSN: 03605442. DOI: 10.1016/j.energy.2022.123681. (visited on 02/11/2023).

[8] L. Ning, L. Pei, and F. Li, “Forecast of China’s Carbon Emissions Based on ARIMA Method,” *Discrete Dynamics in Nature and Society*, vol. 2021, L. Xie, Ed., pp. 1–12, 2021, ISSN: 1607-887X, 1026-0226. DOI: 10.1155/2021/1441942. (visited on 02/11/2023).

[9] M. Gao, H. Yang, Q. Xiao, and M. Goh, "A novel fractional grey Riccati model for carbon emission prediction," *Journal of Cleaner Production*, vol. 282, p. 124 471, 2021, ISSN: 09596526. DOI: 10.1016/j.jclepro.2020.124471. (visited on 03/17/2023).

[10] M. Li, W. Wang, G. De, X. Ji, and Z. Tan, "Forecasting Carbon Emissions Related to Energy Consumption in Beijing-Tianjin-Hebei Region Based on Grey Prediction Theory and Extreme Learning Machine Optimized by Support Vector Machine Algorithm," *Energies*, vol. 11, no. 9, p. 2475, 2018, ISSN: 1996-1073. DOI: 10.3390/en11092475. (visited on 03/17/2023).

[11] J.-p. Liu, X.-b. Zhang, and X.-h. Song, "Regional carbon emission evolution mechanism and its prediction approach driven by carbon trading – A case study of Beijing," *Journal of Cleaner Production*, vol. 172, pp. 2793–2810, 2018, ISSN: 09596526. DOI: 10.1016/j.jclepro.2017.11.133. (visited on 03/17/2023).

[12] W. Sun and J. Sun, "Prediction of carbon dioxide emissions based on principal component analysis with regularized extreme learning machine: The case of China," *Environmental Engineering Research*, vol. 22, no. 3, pp. 302–311, 2017, ISSN: 1226-1025, 2005-968X. DOI: 10.4491/eer.2016.153. (visited on 03/17/2023).

[13] J. Zhou, X. Xu, W. Li, F. Guang, X. Yu, and B. Jin, "Forecasting CO_2 Emissions in China's Construction Industry Based on the Weighted Adaboost-ENN Model and Scenario Analysis," *Journal of Energy*, vol. 2019, pp. 1–12, 2019, ISSN: 2356-735X, 2314-615X. DOI: 10.1155/2019/8275491. (visited on 03/17/2023).

[14] F. Ren and D. Long, "Carbon emission forecasting and scenario analysis in Guangdong Province based on optimized Fast Learning Network," *Journal of Cleaner Production*, vol. 317, p. 128 408, 2021, ISSN: 0959-6526. DOI: 10.1016/j.jclepro.2021.128408. (visited on 02/11/2023).

[15] R. Merritt, *What Are Graph Neural Networks?* 2022. [Online]. Available: https://blogs.nvidia.com/blog/2022/10/24/what-are-graph-neural-networks/ (visited on 03/29/2023).

[16] J. Zhou, G. Cui, S. Hu, *et al.*, "Graph neural networks: A review of methods and applications," *AI Open*, vol. 1, pp. 57–81, 2020, ISSN: 26666510. DOI: 10.1016/j.aiopen.2021.01.001. (visited on 03/18/2023).

[17] M. Ma, Y. Zhang, Y. Li, X. Li, and Y. Liu, “Scatter-GNN: A Scatter Graph Neural Network for Prediction of High-Speed Railway Station—A Case Study of Yinchuan–Chongqing HSR,” *Applied Sciences*, vol. 13, no. 1, p. 150, 2023, ISSN: 2076-3417. DOI: 10.3390/app13010150. (visited on 03/29/2023).

香港科技大学(广州)
THE HONG KONG
UNIVERSITY OF SCIENCE AND
TECHNOLOGY (GUANGZHOU)

THANK YOU!