Baixiang Huang Personal Statement

Born and raised on the Roof of the World, the Tibetan Plateau, the spirit of adventure runs in my blood. Besides expeditions to the high peaks of the Himalayas, intellectual explorations have also been an indispensable part of my life. Science's unsolved questions have long fascinated me, far more than the close-ended questions I encountered in my coursework.

In previous research endeavors, I have explored latency reduction in information-centric networking, data collection using vehicular ad-hoc networks, and the use of Graph Neural Networks (GNNs) in traffic accident prediction. My research participation has enhanced my skills in data mining, computational modeling, and deep learning. From finding research topics to publishing papers, I have demonstrated a profound affinity for academic research.

My first experience with scientific research was as an undergraduate research assistant. I studied the mechanism of information-centric networking and devised two algorithms to improve the routing and placement of data packets. This work could improve network performance in terms of delay, cache hits, and network traffic. The second project I worked on involved simulating the process of data collection using taxis. Based on an analysis of the real-world trajectory of cabs, my supervisor and I then proposed three schemes to improve the deployment of data centers to increase the efficiency of data collection. This research proved the feasibility of utilizing taxis to collect useful data that, if applied in practice, has the potential to facilitate the maintenance of infrastructure.

Upon completion of my bachelor's degree, I studied abroad at the National University of Singapore. In my master's degree thesis, I built four graph-based traffic accident datasets using real-world geospatial data. My supervisor and I also proposed a novel GNN framework to capture angular and directional information from road networks. I then comprehensively evaluated ten state-of-the-art machine learning approaches using the created datasets. We demonstrated that the proposed framework consistently outperforms the baselines. Following that, I built and released a new set of traffic accident benchmark datasets for one thousand US cities, added four GNN baselines, and tested fourteen models on all of the released datasets. This work was presented as a contributed talk at The Workshop on Graph Learning Benchmarks.

During my GNN project, I realized that models cannot be fully trusted without reasoning the underlying mechanisms behind the predictions. For applications like traffic accident prediction, we need human-intelligible explanations to inform the design of future road networks. Therefore, it is my hope to work on explainability in GNNs. What is of particular attraction to me is Dr. Aleksandar Bojchevski's expertise in trustworthy machine learning. Working with Dr. Bojchevski could provide me with valuable inspiration for improving the trustworthiness of GNNs.

Being a first-generation college student, I am highly grateful for all the teachers and mentors who have helped me along the way. My rural background has not stymied but instead only deepened my aspiration for knowledge. Deciding to pursue a Ph.D. is a bold choice, yet bold decisions are what have brought me this far. Hence, I strongly aspire to join your research group so that I can embark upon one of the most challenging paths.

Baixiang Huang

Google Scholar GitHub huang.baixiang@u.nus.edu baixianghuang.github.io/huangbx-site

EDUCATION

National University of Singapore
Master of Computing
Singapore
Aug 2020 – Feb 2022

Central South University
B. Eng., Computer Science and Technology

PUBLICATIONS

- Baixiang Huang, Bryan Hooi. Traffic Accident Prediction using Graph Neural Networks: New Datasets and the TRAVEL Model, *The ACM Web Conference Workshop on Graph Learning Benchmarks*, 2022. Contributed talk. [PDF] [Data]
- **Baixiang Huang**, Wei Liu, Tian Wang, et al. Deployment Optimization of Data Centers in Vehicular Networks. *IEEE Access*, 2019. [PDF]
- **Baixiang Huang**, Anfeng Liu, Chengyuan Zhang, et al. Caching Joint Shortcut Routing to Improve Quality of Service for Information-Centric Networking. *Sensors*, 2018. [PDF]

RESEARCH EXPERIENCE

National University of Singapore, supervisor: Professor Bryan Hooi

Jan 2021 – Apr 2022

China

Sep 2015 - Jun 2019

- Built and released one thousand graph-based traffic accident benchmark datasets for accident prediction problems.
- Proposed a novel graph neural network architecture to capture angular and directional information from road networks.
- Evaluated the proposed framework against state-of-the-art machine learning approaches.
- Achieved the best overall performance on the released datasets.

Central South University, supervisor: Professor Anfeng Liu

Mar 2018 - Apr 2019

- Simulated the process of collecting the condition data of infrastructure using taxis.
- Proposed three schemes to improve the efficiency of data collection.
- Analyzed and tested the proposed schemes on real-world trajectory data of taxis.
- Proved the feasibility of using cabs to collect the condition data of infrastructure.

Central South University, supervisor: Professor Anfeng Liu

Feb 2017 - Mar 2018

- Simulated a scenario of information-centric networking.
- Developed two algorithms to improve the routing and replacement of data packets.
- Achieved improvements in terms of delay, cache hits, and network traffic.

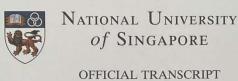
PROJECT EXPERIENCE

Online Course Platform [Code]

- Built an online course website with Spring Cloud.
- Applied Apache FreeMarker to accelerate development by generating code for similar services.
- Implemented Apsara Video VOD APIs to encrypt and authorize course videos.
- Designed and developed web pages using Vue and Bootstrap.

SKILLS

- Programming Languages: Java, Python, HTML/CSS
- Frameworks: PyTorch, PyG, MyBatis, Spring Boot, Vue
- Miscellaneous: Linux, MySQL, Hadoop, Git, Maven



HUANG BAIXIANG STUDENT NO .: DATE OF BIRTH: 30/08/1996 DATE ISSUED: 28/02/2022 DEGREE CONFERRED: MAST
CUMULATIVE AVERAGE POINT: 3.67
CUMULATIVE MODULAR CREDITS: 40.00 MASTER OF COMPUTING GRADE CREDITS MODULE MASTER OF COMPUTING CUMULATIVE AVERAGE POINT: 3.67 03/08/2020 28/02/2022 COMPUTER SCIENCE ADMISSION DATE: CONFERMENT DATE: SPECIALISATION: **END OF TRANSCRIP GRADE CREDITS MODULE ACADEMIC YEAR 2020/2021 SEMESTER 1

KNOWLEDGE DISCOVERY AND DATA MINING ADVANCED COMPUTER NETWORKS THEORETICAL FOUNDATIONS IN MULTIMEDIA

MASTER OF COMPUTING CUMULATIVE AVERAGE POINT: 3.33 ACADEMIC YEAR 2020/2021 SEMESTER 2

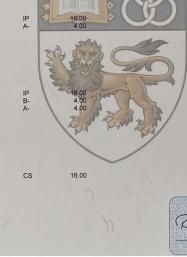
MCOMP DISSERTATION INFORMATION VISUALISATION MASTER OF COMPUTING CUMULATIVE AVERAGE POINT: 3.63

ACADEMIC YEAR 2021/2022 SEMESTER 1 MCOMP DISSERTATION COMPUTER SYSTEM PERFORMANCE ANALYSIS NEURAL NETWORKS AND DEEP LEARNING

MASTER OF COMPUTING CUMULATIVE AVERAGE POINT: 3.67 ACADEMIC YEAR 2021/2022 SEMESTER 2

CP5101 MCOMP DISSERTATION

Name: Baixiang Huang





Student ID: 1502150321

打印日期: 2019年6月25

Print Date: June 25, 2019

Category Credit Score

2

2

2

2

0.5

2

2

2

EC

EC

87

B

92

A

82

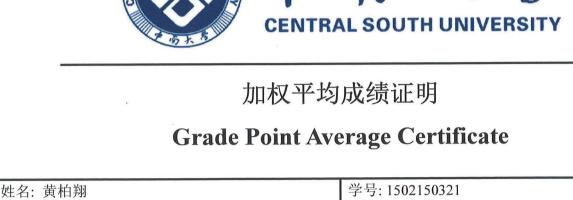
74

80

95

92

PRINTED BY: REGMB PAGE 1 OF 1



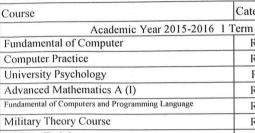
专业: 计算机科学与技术 学院: 计算机学院 Major: Computer Science and Technology College: School of Computer Science and Engineering 学制: 4年 入学日期: 2015年9月 Enroll Date: September, 2015 Years of Study: 4 总学分: 197.5 总加权平均成绩: 84.71/100 Overall Grade Point Average: 84.71/100 Overall Credit: 197.5 2015-2016 学年 Academic Year 2015-2016 加权平均成绩 78.94/100 Grade Point Average 2016-2017 学年 Academic Year 2016-2017 加权平均成绩 84.31/100 Grade Point Average 2017-2018 学年 Academic Year 2017-2018 加权平均成绩 88.71/100 Grade Point Average 2018-2019 学年 Academic Year 2018-2019 加权平均成绩 89.11/100 Grade Point Average

Enroll Date: September, 2015 Gender: male Name: HuangBaiXiang Student ID: 1502150321 Major: Computer Science and Technology Years of Study: 4 College: School of computer Science and ********************

Course

Category Credit Score

地址:中国湖南省长沙市麓山南路 932 号中南大学 Address: Central South University, 932 South Lushan Road, Changsha, Hunan, P. R. China, 410083



Computer Network Course Design RC 2.5 80 Software Architecture EC RC 1 В PE 2 80 Data Structure Course Design RC Digital Image Processing 5 76 EC RC Physical-Fitness Test (I) RC RC 2.5 80 Network Engineering RC 1 82 EC

System & Application of Linux

Introduction to Parallel Computing

Molecular Biology

中南大學本科生成绩单

Undergraduate Academic Transcript Of Central South University

Military Training RC 1.5 В Moral Education and Foundation of law RC 3 77 Physical Education (I) 88 RC 1 Introductory Course For Freshmen RC 1 Α English Reading, Writing and Translating (I) RC 2 83 English Viewing, Listening and Speaking (I) 89 RC 2 Academic Year 2015-2016 2 Term Mental Health Education 2 82 RC University Physics C RC 4.5 72 Circuit and Analog Electronic Technology RC 4 79 Advanced Mathematics A (II) RC 5 67 Advanced Programming Practice (C++) RC 2 В Arts Skills of Management Communication PE 2 94 Object-Oriented Programming (C++) RC 3 84 Physical Education (II) RC 1 89 Linear Algebra A 2 72 RC Writing PE 2 88 English Reading, Writing and Translating (II) RC 2 81 English Viewing, Listening and Speaking (II) 2 92 RC Academic Year 2016-2017 1 Term Java Language and System Design EC 3 82 Health Education PE 2 98 Electrical and Electronic Experiment RC 1 В Probability and Statistics A 3.5 RC 68 Elective Series of Advanced English 2 96 EC Discrete Mathematics RC 3 70 Cognition Practice Data Structure RC 3.5 65 Digital Electronic Technology A RC 3.5 87 Physical Education (III) RC 84 1 English Viewing, Listening and Speaking (III) RC 2 96 Professional Introduction RC В 1 Academic Year 2016-2017 2 Term Principle of Operating System 3 84 Course Design of Electronic Technology В RC Computer Network RC 3 87 Computer Principle and Assembly Language RC 4 91 Course Design of Computer Principle and Assembly 2 RC Artificial Intelligence 90 2 EC Database Principle 78 3 RC Algorithm Analysis and Design 3 91 EC Physical Education (IV) 92 RC 1 Situation and Policy RC 96 Academic Year 2017-2018 1 Term Course Design of Java 2 A RFID and Smart Card Techniques EC 3 96 90 Large Database Technology 3 EC S/N: ZNDX IHGF HEKB FDDF AIMD Web: http://print.csu.edu.cn/valid.jsp Page 1 of 1 中南大學本科生成绩单 Undergraduate Academic Transcript Of Central South University

Operating System Course Design 2 RC В Introduction to Innovation and Entrepreneurship EC 2 96 Multimedia Principle and System Design 2 87 EC Distributed Systems 2 98 EC Software Testing EC 2 90 Software Engineering EC 3 75 Integrated Course Design of Database RC 2 В Physical-Fitness Test (II) RC 0.5 83 Modern Chinese History RC 88 Academic Year 2018-2019 1 Term
Electronic-Commerce and Electronic-Government 89 2 EC Computer Simulation and Modeling 1.5 81 EC Visualization Technology EC 2 85 Basic Theory of Marxism RC 3 86 Mao Zedong Thought and theoretical system of socialism with Chinese characteristics RC 5 80 Production Practice RC 3 В 2 **Bioinformatics** EC 88 Introduction to Data Science Big Data 2 RC 88 Physical-Fitness Test (III) RC 0.5 85 Subject I Computer Control Technology EC 1 93 Subject II - The Development of Modern Computer Technology EC 1 88 Subject III - Information Processing and Fusion Technology Academic Year 2018-2019 2 Term Graduation Practice and Project RC 16 Skill Assessment ES P 3.0 ES 4.0 P Academic Achievements P Social Practice ES 1.0 CET-6: 609 CET-4:571 Elective Course: 50.5 Credits: Required Course: 131 Public Elective Course: 8 Extracurricular studies: 8 Remarks: Resit or retaken course results are marked with "*". ********Blank Below**

. Date: June 25,2019

入学年月: 2015年09月

必修

必修

选修

必修

选修

必修

选修

选修

选修

必修

选修

选修

选修

选修

选修

必修

必修

87

良

92

优

82

74

80

95

92

良

96

87

98

90

75

良

83

88

89

81

85

2

2

2

2

2

2

2

2

2

2

2

2

3

2

2

1.5

2

学制: 四年

2017-2018 学年 第2学期

Academic Year 2017-2018 2 Term



大学生心理健康教育

电路与模拟电子技术

大学物理C

体育(二)

线性代数A

学号: 1502150321 学院: 计算机学院

大学计算机基础

大学心理学 高等数学A(一)

大学计算机基础实践

课程/环节

高等数学A(二) 高级程序设计实践(C++) 必修 2 良 中国近现代史纲要 必修 管理沟通艺术与技巧 任选 2 94 2018-2019 学年 第1学期 电子商务与电子政务 选修 面向对象编程(C++) 必修 3 84 选修 计算机仿真与建模 必修 1 89 选修 2 可视化技术 必修 72

性别: 男

专业: 计算机科学与技术 (大数据方班级: 计算机(大数据)1501

课程/环节

分子生物学

软件体系结构

数字图像处理

网络工程

计算机网络课程设计

数据结构课程设计

体育课外测试 (一)

Linux系统及应用

操作系统课程设计

多媒体原理与系统设计

数据库综合课程设计

体育课外测试 (二)

并行计算导论

创新创业导论

分布式系统

软件测试

软件工程

姓名: 黄柏翔

性质

必修

必修

任选

必修

学分

2.5

1

2

5

2.5

1

1.5

3

1

2

2

2

4.5

4

5

成绩

80

良

80

76

80

82

良

77

88

优

83

89

82

72

79

67

2015-2016 学年 第1学期

2015-2016 学年 第2学期

写作		日	壬选	2	88	马克思主义基本原	理	必何	多 3	86
英语读写证	羊 (二)	业	必修	2	81	毛泽东思想和中国特色	色社会主义理论体系概	论 必	多 5	80
英语视听说		业	必修	2	92	生产实习	1	必何	多 3	良
		7 学年 第1学期				生物信息学		选作	多 2	88
Java 语言	与系统设计	选	先修	3	82	数据科学与大数据	技术导论	必何	多 2	88
大学生健康	 表教育	白	壬选	2	98	体育课外测试 (三)	必有	多 0.5	85
电工电子实	实验	业	必修	1	良	专题I计算机控制	技术	选值	多 1	93
概率论与数		业	公修	3. 5	68	专题 II -现代计算标	几技术发展	选作	多 1	88
高级英语系	系列选修	选	先修	2	96	专题III-信息处理与	 	选值	多 1	86
离散数学		业	公修	3	70		2018-2019 学年	第2学期		
认识实习		业	公修	2	良	毕业实习与设计		必任	多 16	优
数据结构		业	公修	3. 5	65	技能考试		课名	小 3.0	通过
数字电子打	支术A	业	公修	3. 5	87	论文成果		课名	小 4.0	通过
体育(三)		业	公修	1	84	社会实践		课名	· 1.0	通过
英语视听说		业	公修	2	96	******	******	*****	*****	*****
专业导论		业	公修	1	良	英语四级: 571	英语六:	级: 609		
	2016-201	7 学年 第2学期					**************************************	: ::+*******	*****	*****
操作系统原			先修	3	84	获得学分: 必修:		先修: 50.5		
电子技术设			公修	2	良	任选:		果外: 8		
计算机网织		业	公修	3	87			<u> ለ</u>		
计算机原理		业	公修	4	91	备注:重修或补考成	《须杯识 *** 亏 ********以下空白	******	*****	*****
	里与汇编课程设计		公修	2	优		TITLE			
人工智能			先修	2	90					
数据库原理	Ψ.	业	公修	3	78					
算法分析与	ラ设计	逆	先修	3	91				, i	
体育(四)		业	必修	1	92					
形势与政策		业	必修	1	96					
		8 学年 第1学期						Xa >	1	
Java课程证			必修	2	优			B62 -	1. 1	
RFID与智能		逆	先修	3	96			A		1
大型数据风		逆	先修	3	90				.12	
	验证网址: http:/	M NLCI IILK NLCD		id. js	; 第 1	页共1页	ACI LE	节百期 ,201 9年	00 AZ	
	百分制、五级设	生成绩记载说明: 平分制与绩点的:				70.70	60.60	/50		
				89-8	30	79-70	69-60 及格	≪59		×

绩点 4

e one-hundr	ed-point score can	be converted in	to grade and GPA	as follows:		
Score	100-90	89-80	79-70	69-60	≤59	
Grade	Excellent (A)	Good (B)	Average (C)	Pass (D)	Fail (F)	
GPA	4	3	2	1	0	

2

1

0

3

Description of Transcript for student of Central South University:

Abstract of Baixiang Huang's Master's Thesis

Traffic accident prediction is crucial for reducing and mitigating road traffic accidents. Existing machine learning approaches aim to predict the number of traffic accidents in each cell of a discretized grid, based on features from each grid cell, without considering the underlying graph structure of road networks. To allow us to incorporate road network information, graph-based approaches such as Graph Neural Networks (GNNs) are a natural choice, but are made challenging by a lack of suitable graph-based traffic accident prediction datasets. To overcome this problem, we first construct a new set of four graph-based traffic accident datasets, along with two benchmark tasks (accident occurrence prediction and accident severity prediction). We then comprehensively evaluate a variety of state-of-the-art GNN variants using the created datasets. Furthermore, we propose our novel Traffic Accident Vulnerability Estimation via Linkage (TRAVEL) model, which is designed to capture angular and directional information from road networks. We demonstrate that our proposed model consistently outperforms the baselines by a large margin.