Jiaxuan Li

Master student, Department of Computer Science
University of Harbin Institute of Technology, Shenzhen, China

Email: jiaxuanliniki@gmail.com
Webpage: https://ycy3675001.github.io

Education

Harbin Institute of Technology，Shenzhen, China

M.S., computer science
GPA: 3.09 / 4.0
Supervisor: Dr. Philippe Fournier-Viger

Sep. 2017 to Jan. 2020

Hai Nan University, Hainan, China

B.S., computer science
GPA: 3.67 / 4.0 (88.2 / 100), Top 3%
Supervisor: Dr. Yucong Duan

Sep. 2013 to Jul. 2017

Publications

- A Survey of Pattern Mining in Dynamic Graphs. Philippe Fournier-Viger*, Ganghuan He, Chao Cheng, **Jiaxuan Li**, Jerry Chun-Wei Lin, Unil Yun. *WIREs Data Mining and Knowledge Discovery, Wiley*, 2019. **Submitted.**

- Efficiently Extracting Cost-Effective patterns from Sequential Event Log. Philippe Fournier-Viger*, **Jiaxuan Li**[+], Jerry Chun-Wei Lin, Tin Truong Chi, R. Uday Kiran. *Knowledge-Based Systems (KBS), Elsevier*, 2019. **Accepted.**

- Discovering and Visualizing Patterns in Utility Sequences. Philippe Fournier-Viger*, **Jiaxuan Li**[+], Jerry Chun-Wei Lin, Tin Truong Chi. *Proc. 21st Intern. Conf. on Data Warehousing and Knowledge Discovery (DAWAK), Springer*, 2019.

- Discovering low-cost high utility patterns. **Jiaxuan Li**[+], Philippe Fournier-Viger*, Lin, Jerry Chun-Wei Lin, Tin Truong Chi. *1st International Workshop on Utility-Driven Mining (UDM), in conjunction with the KDD 2018 conference, ACM press*, 2018. **Oral presentation.**

  *Academic supervisor, + main student contributor.

Research & Industry Experience

- **Harbin Institute of Technology, Shenzhen**   Mar. 2018 to Now
  Mentor: Dr. Philippe Fournier-Viger
  - **Cost-effective pattern mining**   Mar. 2018 to Jul. 2019
    - Mining cost-effective pattern from event logs in E-learning to provide insights about how to utilize the learning materials.
    - Combined a cost model into high utility sequential pattern mining considering the resources, effort, time or cost required to apply the patterns for getting a desirable utility.
    - Designed statistical measures to assess the correlation between utility and cost for the needs of different applications in terms of the type of utility (binary or numeric).
    - Integrated buffer structure into Prefixspan algorithm and designed pruning strategy to improve algorithms' performance in terms of memory usage and execution time.
  - **Cost-effective pattern mining from heterogeneous data source**   Jul. 2019 to Now

- Mining guidance patterns in e-learning from heterogeneous data source for specific group of users.
- Combining users' attributes, such as personal information and educational background, with their learning activities.
- Clustering users based on their attributes and mining cost-effective patterns from their sequences of activities respectively to assist different group of users use materials efficiently.
- Representing the concatenated features using vector and utilizing a statistical measure to evaluate the correlation between the feature and utility.
- Currently designing the model and searching potential datasets for testing.

- **Noah'Ark Lab, Huawei Technologies**   Aug. 2019 to Now
   Mentor: Dr. Min Zhou
   - Spatial-temporal sequence pattern mining in telecommunication network to compress alarm records, identify important alarms, and locate root-cause alarms.
   - Using dynamic attributes graph as an approach to mine important sequences of alarms that have a higher priority to be responded from various network equipment, meanwhile keeping the topology of the network.
   - Designing correlation measure and generating potential correlated sequence rules of alarms to identify the root-cause alarms.
   - Currently pre-processing data and analyzing data, including the information of alarms' occurrence time, domain, name, source and network topology information.

- **2012 Lab, Huawei Technologies**   Jun. 2019 to Aug. 2019
   Mentor: Dr. Zixian Zhang
   - Designed, implemented and tested a function for automatically extracting and checking the CAD drawings' content about servers to improve the manual inspection's accuracy.
   - Analyzed about 5 types of drawings. Extracted their components' data structure using ActiveX and mined different components' crucial features, respectively.
   - Based on those features, implemented algorithms to structurally extract contents in specified areas, compared those information with the official documents, and finally generated a detailed verification report.
   - Checked the operating specifications of the drawings, such as the intersection between texts and lines, missing arrows, and manual errors, such as the absence of a component's description or missing a component that should be contained in the drawing.
   - Tested about 100 drawings, and now this feature was integrated into their production system.

Selected Awards

- Outstanding graduates of Hainan University, 2017.
- Mathematical Contest in Modeling Certificate of Achievement, Honorable Mention, 2016.

Essay

I want to pursue a Ph.D. degree in Computer Science, and my primary research interest is in the intersection area of data mining with natural language processing, machine learning or deep learning. My initial passion and interest in data mining was inspired by an undergraduate course.

My undergraduate study at Hainan University was full and tense. In my sophomore year, I transferred my major to computer science due to my interest in programming. Then I finished all the required courses within two years and kept my GPA in the top 3% ranking. Among all the courses, "data warehouse and data mining" attracted me the most. I was always wondering how people could get valuable guidance from a series of seemingly messy numbers. This curiosity persuaded me to participate in the American Mathematical Contest in Modeling. As the captain of a team of three girls, I mainly charged with analyzing data, building mathematical models, and setting paper's writing logic. We used prediction and evaluation models to determine whether an economical opportunity exists to remove the space debris and provided a recommended solution to remove the debris. After a month of training and four days of tough contests, we won the second prize. This experience greatly improved my confidence and inspired me to further explore interesting problems in data mining.

With the honor of waiving the entrance exam, I was admitted to the Department of Computer Science in Harbin Institute of Technology, Shenzhen, which is famous for its artificial intelligence research in China. Following my passion, I joined Professor Philippe Fournier-Viger's group to continue my study in data mining. My research projects focus on high utility pattern mining, which is about mining patterns that have great importance in the database. Observing the limitation that existing models often ignore the cost (e.g., effort, resources) required for obtaining the high utility, I added a cost model into the utility model to mine the patterns that can generate a high utility but consume a relatively small cost. First, I designed innovative pruning measures and a buffer structure to improve the algorithm's efficiency under the huge search space. The experiments showed that the execution time is improved by up to 10 times and the memory usage is reduced by up to 20%. Second, two reasonable statistical measures were proposed to rank the discovered patterns in different scenarios. The proposed models were applied to e-learning datasets to provide insights about time-efficient ways of utilizing learning materials to obtain good performance.

Last August, I presented my work at the Utility-Driven Mining and Learning Workshop, held in conjunction with the 24th KDD conference in London. It was an excellent experience. In the conference, I learned about many interesting research topics and exciting applications in the area of natural language processing, machine learning, and deep learning. It also gave me an opportunity to discuss my research problems with several pioneer researchers in data mining. Their inspiring responses encouraged me to further investigate one of the problems. In my current project, I try to

provide guidance for people with different characteristics by implementing mining methods to properly combine heterogeneous data sources. In summary, I need to integrate people's profile and their sequence of activities to learn and mine the best ways of utilizing resources for an individual or a group.

Conducting research at school cannot satisfy my curiosity on understanding data mining's power in solving real-world problems. I was eager to work with researchers in industry and apply my knowledge on data mining to solve challenging problems with real data. In the summer of 2019, I joined the 2012 Lab of Huawei Technology as an intern. During my internship, I designed and implemented a framework to discover erroneous patterns from the CAD drawing data. This system was deployed in production and greatly improved the CAD design team's efficiency by reducing many manual efforts. This successful internship experience brought me another opportunity with Huawei. Now I am participating in a project of alarm compression and causal root location in the telecommunication network with Noah' Ark Lab. The main focuses of the project are using dynamic attributes graph as an approach to mine important sequences of alarms that have high priorities and generating potential correlated sequence rules of alarms to identify the root-cause alarms.

Renowned for its prominent faculty and state-of-the-art research facilities, the graduate program at The University of Texas at Dallas is an ideal platform for me to advance my pursuit in data mining research. In particular, there are several professors whose projects are especially appealing to me: Prof. Jessica Ouyang (natural language processing on summarization, paraphrasing, and style transfer); Prof. Vincent Ng (unsupervised and semi-supervised machine learning for natural language processing); Prof. Feng Chen (data mining in complex and heterogeneous network data). Reading those professors' project website and seeing the researches' broad applications have convinced me that pursuing Ph.D. degree in your university is a great opportunity to fulfill my long-term professional objectives: keep learning advanced technics, make data mining technology a positive impact on human life and devote myself to build a more intelligent and convenient world. I hope that I have convinced you of my passion and ability, and I sincerely wish to be admitted to your program.

Hainan University Student's Academic Record

Student ID No:20130801310013   Name: LiJiaXuan   Gender: Female

School:School of Computer Science & Cyberspace Security

Major:Computer Science and Technology

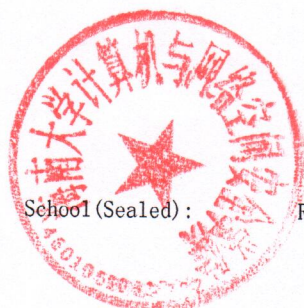Date of Enrollment:20130902   Length of Program: 4   Years

Course Title	Type	Credits	Score	Point
2013-2014-1				
Physical Education I	Required	1.0	98	4.0
Military Theory	Required	1.0	88	3.8
College English I	Required	4.0	86	3.6
Situation and Policy I	Required	0.0	90	4.0
Ideological and Moral Cultivation and Legal Basis	Required	3.0	85	3.5
Military Abilities Training	Required	1.0	87	3.7
2013-2014-2				
Physical Education II	Required	1.0	93	4.0
College English II	Required	4.0	83	3.3
Situation and Policy II	Required	0.5	90	4.0
College Chinese	Required	3.0	80	3.0
Outline of Modern and Contemporary Chinese History	Required	3.0	85	3.5
2014-2015-1				
Physical Education III	Required	1.0	95	4.0
JAVA Programming	Required	4.0	95	4.0
Java Programing Experiment	Required	1.0	87	3.7
Social Practice I	Required	1.0	83	3.3
Advanced mathmatics	Required	5.0	88	3.8
Linear Algebra A1	Required	3.0	85	3.5
Probability and Statistics A1	Required	3.0	80	3.0
English for International Communication II	Elective	3.0	71	2.1
Situation and Policy III	Required	0.0	85	3.5
Introduction to Basic Principles of Marxism	Required	3.0	80	3.0
Discrete mathematics	Required	3.5	94	4.0
Principle of Computer Composition	Required	4.0	85	3.5
Data Structure	Required	3.5	91	4.0
Experiment of Data Structure	Required	1.0	90	4.0
Course Project of Data Structure	Required	1.5	Excellent	4.0
Course Design in JAVA Programming	Required	2.0	72	2.2
Mathematical Thinking	Elective	3.0	86	
Evolutionary Biology	Elective	2.0	90	
2014-2015-2				
Physical Education IV	Required	1.0	97	4.0
C++ Programming	Required	3.5	94	4.0
Fundamentals of Compiling	Required	4.0	95	4.0
Experiment of C++ Programming	Required	1.0	96	4.0
Information Retrieval and Utilization	Required	1.0	89	3.9
Higher Mathematics A1(volume two)	Required	6.0	90	4.0
Advanced English Reading	Elective	3.0	85	3.5
Situation and Policy IV	Required	0.5	80	3.0
Introduction to Mao Zedong Thought and theoretical system of socialism with Chinese characteristics	Required	5.0	93	4.0
Computer Networks	Required	3.0	87	3.7

Rank Examinations Results

CET4	511
CET6	524

Achievement Results

Required Total Credits of the Training Program:0

Course Credits: 181.5 Quality Expansion Credits: 0 Total Credits: 181.50

GPA:3.67 Average Score: 88.2 Average Credit Grade:88.13

The grade point of make-up exams and re-take courses are all marked as 1.0

School (Sealed):   Reviewer: Dean's Office(Sealed):   Print Date: 20-06-2019

Page 1 , 2 Page in Total

Hainan University Student's Academic Record

Student ID No:20130801310013   Name: LiJiaXuan   Gender: Female

School:School of Computer Science & Cyberspace Security

Major:Computer Science and Technology

Date of Enrollment:20130902   Length of Program: 4   Years

Course Title	Type	Credits	Score	Point
Computer Network Experiment	Required	1.0	85	3.5
database system	Required	3.0	90	4.0
experiment of database system.	Required	1.0	90	4.0
Course Design in Computer Network	Required	1.5	Excellent	4.0
Course Design in Database	Required	1.5	90	4.0
Course Design in C++ Programming	Required	1.5	Excellent	4.0
Contemporary Chinese Economy	Elective	2.5	87	
Chinese Ancient Etiquette and Rituals	Elective	2.0	93	
2015-2016-1				
Operating System	Required	4.5	92	4.0
Social Practice II	Required	1.0	85	3.5
Situation and Policy V	Required	0.0	90	4.0
Introduction to Computer Science	Required	3.0	87	3.7
Software Engineering	Required	4.0	94	4.0
Course Design in Software Engineering	Required	1.5	92	4.0
Internet Information Retrieval	Elective	2.0	88	3.8
2015-2016-2				
Analysis and Design of Algorithms	Required	3.0	89	3.9
Experiment of Algorithm Analysis and Design	Required	1.0	98	4.0
Innovation and entrepreneurship	Elective	2.0	87	
Innovative business practices	Elective	1.0	85	
Situation and Policy VI	Required	0.5	93	4.0
Principle of Circuits	Required	4.0	74	2.4
Lab. of Principle of Circuits	Required	0.5	Excellent	4.0
Graph Theory and Its Application	Elective	2.0	93	4.0
Computer Assistded Design	Elective	2.0	95	4.0
Experiment of Computer Aided Design	Elective	1.0	80	3.0
Software Testing Theory and Method	Elective	4.0	90	4.0
Network Security Technology	Elective	4.0	87	3.7
2016-2017-1				
Data Warehouse and Data Mining	Required	4.0	Good	3.5
Digital Image Processing Technology	Required	2.0	85	3.5
Experiment of Digital Image Processing Technology	Required	1.0	80	3.0
Situation and Policy VII	Required	0.5	86	3.6
Analog Electronics Technique	Required	4.0	87	3.7
Experiment in Analog Electronics Technique	Required	0.5	Good	3.5
Course Design in Comprehensive Computer System	Required	1.5	87	3.7
Computer Graphics	Elective	4.0	84	3.4
2016-2017-2				
Graduation Internship	Required	5.0	87	3.7
Graduation Thesis	Required	10.0	Excellent	4.0

Rank Examinations Results

CET4	511
CET6	524

Achievement Results

Required Total Credits of the Training Program:0

Course Credits: 181.5 Quality Expansion Credits: 0 Total Credits: 181.50

GPA:3.67 Average Score: 88.2 Average Credit Grade:88.13

The grade point of make-up exams and re-take courses are all marked as 1.0

School(Sealed):   Reviewer: Dean's Office(Sealed):   Print Date: 20-06-2019

哈尔滨工业大学（深圳）
HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

学生成绩单
Student's Academic Transcript

Name:	LI Jiaxuan	**Program:**	Master	
Student ID:	17S051070	**School:**	Computer Science and Technology	
Entrance Date:	201709	**Major:**	Computer Science and Technology	

Course	Semester	Hour	Credit	Grade	Mark
Academic Activities	Spring,2018	32	2.0	PASS	85
Data Mining	Spring,2018	16	1.0	B+	79
Professional Ethics	Spring,2018	16	1.0	PASS	85
Pattern Recognition Algorithm and Progress	Fall,2017	32	2.0	B	73
Academic English Writing	Fall,2017	16	1.0	B	73
Research on Socialism with Chinese Characterics: Theory and Pratice	Fall,2017	32	2.0	B+	79
Capitalism Against Capitalism:Western Family Business Management	Fall,2017	16	1.0	A+	96
Dialectics of Nature	Fall,2017	16	1.0	A-	84
Combinatorics	Fall,2017	32	2.0	B	73
Matrix Analysis	Fall,2017	32	2.0	B+	79
Introduction to Machine Learning	Fall,2017	32	2.0	B	73
Complex Networks: Modeling and Dynamics	Fall,2017	32	2.0	A-	84
Advanced Algorithms	Fall,2017	32	2.0	B	73
Data Mining	Fall,2017	32	2.0	B	73
Introduction to Internet of Things	Fall,2017	32	2.0	A-	84
Computer Science and Technology Seminar Topics	Spring,2019	32	2.0	PASS	85
Advanced Computer Networks Modeling and Simulation	Fall,2018	32	2.0	A-	84
Advanced Database System Principle	Fall,2018	32	2.0	A+	96
Numerical Analysis	Fall,2018	32	2.0	A+	96
Thesis Proposal	Fall,2018	16	1.0	PASS	85
Information Security	Fall,2018	32	2.0	A+	96

GPA:3.09

Deputy director:Zhu Ning

(scan verification) 毕业成绩单校核人：张晓婉