

ChenyuanWu

E-mail: chenyanwu1998@163.com, Phone: +86-18800172721

Room 402, Block B, Jiayuan Student Apartment, Beijing Jiaotong University, Beijing, China

Education

Beijing Jiaotong University (BJTU), School of Electronic and Information Engineering

- B.E.in Communication Engineering Sep. 2016 – Jun. 2020(Expected)
- General **GPA:3.93/4.00, Weighted Average Mark:92.9/100, Rank:1/330**
- Key Courses: Probability Theory and Mathematical Statistics 99, Complex Functions and Integral Transformations 100, Geometry and Algebra 100, Computer Networks 99, Principles of Communication Systems 97, Digital Signal Processing 98, Signals and Systems 99, Network Content Security (Principles of Machine Learning) A, Seminars on Information Network A, C Programming A, Advanced Programming Training (Data Structure) A-
- TOEFL:108 (Best Score 110), GRE:326

Papers

- [1] Haoxian Chen, **Chenyuan Wu**, Manqiu Zhang, Yahui Li, Zhijia Chen, Boon Thau Loo, Mayur Naik, Mukund Raghothaman, *Example-Driven Network Program Synthesis*, ACM SIGCOMM, 2020. Status: Will submit.
- [2] Shaobo Wang, Hui Lyu, Jiachi Zhang, **Chenyuan Wu**, Xinyi Chen, Wenchao Zhou, Boon Thau Loo, Susan B. Davidson and Chen Chen, *Provenance for Probabilistic Logic Programs*, International Conference on Extending Database Technology (EDBT), 2020. Status: Accepted.

Awards and Honors

President's Scholarship in Beijing Jiaotong University (10 out of 15,000)	Nov.2019
The National Scholarship	Sep.2019
Weinuo Times Scholarship	Sep. 2018
The National Scholarship	Sep.2017
The Second Prize in the National College Physics Competition	Dec. 2017
The First Prize in the Mathematics Competition of Beijing Jiaotong University	Jun. 2017

Working Experience

HUAWEI TECHNOLOGIES CO., LTD (Huawei), 2012 Labs, Central Research Institute

Research intern in Network Technology Lab **Nov. 2019 – Present**

- Did literature review on multipath transport layer protocols;
- Networked the servers, constructed the test platform in Linux, conducted bandwidth tests of MPTCP in heterogeneous networks;
- Work in progress: design a mechanism to stabilize the network performance when long-distance traffic meets short-distance traffic in a bottleneck, run the simulation on NS3;

Research Experiences

Example-Driven Network Program Synthesis

July. 2019 – Sept. 2019

Research Assistant, Network and Distributed System Lab, University of Pennsylvania; *Advisor: Prof. Boon Thau Loo*

Objective: We propose an interactive programming-by-example toolkit that synthesizes network-domain specific programs from input-output examples based on numerical optimization and active learning; we are able to synthesize various SDN controller programs and distributed routing programs from example traces in three different Domain Specific Languages (Scala, Mininet Tool).

- Did literature review on Declarative Networking, Software Defined Networking such as Ethane, Flowlog, Pyretic and Frenetic;
- Implemented a compiler module of the toolkit in Scala (300 lines of code), which can translate the synthesized logical representation automatically into Flowlog rules in a line-to-line pattern, supporting various forwarding behaviors and controller state update;
- Implemented another compiler module in Scala (350 lines of code), which can translate the synthesized logical representation automatically into dynamic policies in Pyretic (essentially in Python programs) by leveraging Pandas as database operation API and properly handling policy initialization and network default behavior;
- Used the compiler to generate mac learning, stateful firewall and static routing programs in both Flowlog and Pyretic;
- Set up Pyretic runtime and Mininet tools, and validated the translated program in Mininet;
- Created a FatTree topology in Mininet with one host as HTTP server and eighteen others as client, run httpperf request on all the clients upon mac learning programs, aiming to compare the end-to-end difference between handwritten and generated Pyretic programs;
- Did Cbench tests for all three generated programs, aiming to compare the response ability of generated controller programs;
- Contributed to a SIGCOMM paper [1];

Provenance for Probabilistic Logic Programs

July. 2019 – Sept. 2019

Research Assistant, Network and Distributed System Lab, University of Pennsylvania; Advisor: Prof. Boon Thau Loo

Objective: We propose P3, a novel provenance model and system for debugging probabilistic logic programs; we enable four kinds of queries: traditional explanation queries, queries for finding the set of most important derivation within an approximate error, most influential queries, and modification queries that change the output probability with fewest modifications to input data or program (C++, OpenCL Framework, Rapidnet Tool).

- Did literature review on Datalog, OpenCL and rule based Visual Question Answering;
- Implemented a parallel method for the most influential query in C++ and OpenCL (400 lines of code for parallel module together);
- Implemented another parallel method for the most influential query, where each work unit is in charge of a single round of Monte Carlo simulation;
- Set up parallel computing environment on both Nvidia GPU and Intel CPU; run experiments on large provenance graph, analyzing results from different method and on different hardware setting, boosting the performance by more than 20-fold faster than sequential query at most;
- Made a VQA case study, further enabling the system to find inappropriate input data and show the modification strategy with minimum cost in order to make prediction right;
- Contributed to an EDBT paper [2];

Prostate Image Segmentation based on Machine Learning

Sept. 2018 – Feb. 2019

Independent Project, Institute of Intelligent Sensing and Information Processing (National Teaching Base for Electrical and Electronic Engineering), Beijing Jiaotong University; Advisor: Prof. Yahui Peng

Objective: I aimed to realize the separation of different tissue structures, especially the prostate tissue by using the method of Gaussian Mixture Model and the nearest neighbor method for multimodal images (250 lines of code, Python, Scikit-Learn Framework).

- Did literature review, and learned how to process DICOM files through Python;
- Data Preprocessing: For a given slice, programmed to find the slices with the shortest distance and the same normal direction in each set of slices with different modes; and got the gray value of the points in different modal slices;
- Feature Engineering: Did sparse sampling with equal spacing on 2D plane, and the different gray values of each sample point were used as features to form a multi-dimensional vector space;

- Used GMM clustering algorithm to divide the sample points in vector space into different sets of points and outputted different labels to represent them; the remaining unclassified pixels in the 2D plane were mapped to the vector space, and the nearest neighbor algorithm was adopted to classify the remaining points; find the best parameters;
- Repeated the above process for a case of multi-slice 3D volume, completed the segmentation under 3D conditions, and used the 3D slicer software for stereo reconstruction, showing the three-dimensional structure after segmentation;
- Data analysis: compared the results with those obtained under 2D conditions;

Prediction of Hot Spots based on PageRank Method

Oct. 2018 – Jan. 2019

Hot Spots: Specific locations with large network demand and operators should give priority of network speed and quality for these places.

Independent Project, Network Science and Computation Lab, Beijing Jiaotong University; Advisor: Prof. **Yuchun Guo**

Objective: I expected to get the prediction of potential hot spots on analyzing user communication data provided by China Unicom Operator (Python, Pandas).

- Calculated the network traffic of each base station and ranked the priority of each base station according to the traffic;
- Used Pandas to integrate information in multiple CSV files to construct transfer matrix and realize PageRank algorithm by programming; the priority of each base station was ranked according to PageRank value;
- Counted the number of times that users arrived at each base station and the length of calls in one year, so as to adjust the transfer probability in the transfer matrix, realized the weighted PageRank algorithm, and finally ranked the priority of each base station according to PageRank value;
- Analyzed and compared the various ranking results;

Skills

Programming skills: Skilled in the following languages: Scala, C++, Python (Pandas, Numpy, Matplotlib), NDlog (Network Datalog), Pyretic

Tools/Frameworks: Scikit-Learn, OpenCL, Mininet, Rapidnet, Tensorflow