

Department of Mathematics and Statistics,  
University of Victoria,  
PO BOX 1700 STN CSC,  
Victoria, BC V8W 2Y2,  
Canada

yangguo@uvic.ca  
sgyzetrov.github.io

## EDUCATION

### University of Victoria, Victoria, BC, Canada

Master of Science in Statistics September 2019 – July 2021

- Thesis: Learning COVID-19 network from literature databases using core decomposition
- Supervisors: Dr. [Xuekui Zhang](#) and Dr. [Li Xing](#)
- Activity: Trainee of the [Visual and Automated Disease Analytics \(VADA\) training program](#), received training on translating complex health data into insights that can be used to improve the health of populations and support health professional decision making.

### Huazhong Agricultural University, Wuhan, Hubei, China

B.Eng. in Computer Science and Technology August 2015 – June 2019

- Bachelor Thesis: Study on Epistasis Mining Approach Based on Genetic Tabu Algorithm and Bayesian Network
- Supervisor: Dr. [Jianxiao Liu](#)

## PUBLICATIONS

### Journal Articles

- **Integrative COVID-19 biological network inference with probabilistic core decomposition** [[Link](#)]  
*Yang Guo, Fatemeh Esfahani, Xiaojian Shao, Venkatesh Srinivasan, Alex Thomo, Li Xing, Xuekui Zhang\**  
Briefings in Bioinformatics, 2021 (in press)
- **Multi-Stage Graph Peeling Algorithm for Probabilistic Core Decomposition** [[Link](#)]  
*Yang Guo, Xuekui Zhang, Fatemeh Esfahani, Venkatesh Srinivasan, Alex Thomo, Li Xing*  
Proceedings of the 2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2021)
- **Epi-GTBN: An Approach of Epistasis Mining Based on Genetic Tabu Algorithm and Bayesian Network** [[Link](#)]  
*Yang Guo<sup>#</sup>, Zhiman Zhong<sup>#</sup> (joint first authors), Chen Yang, Jiangfeng Hu, Yaling Jiang, Zizhen Liang, Hui Gao, Jianxiao Liu\**  
BMC Bioinformatics, 2019, 20(1): 444:1-444:18
- **An Improved Constraint-Based Bayesian Network Learning Method Using Gaussian Kernel Probability Density Estimator** [[Link](#)]  
*Yaling Jiang, Zizhen Liang, Hui Gao, Yang Guo, Zhiman Zhong, Chen Yang, Jianxiao Liu\**  
Expert Systems with Applications, 2018, 113: 544-554

### Books

- **Mathematical Modeling Methods 2<sup>nd</sup> Edition (textbook, in Chinese)**  
*Editorial Board, Invited Author, Chapter 9 - Selected Topics in Intelligent Algorithms*  
In Preparation.

## Patents

- **An Approach of Epistasis Mining Based on Genetic Tabu Algorithm and Bayesian Network**

Jianxiao Liu, Yang Guo, Zhiman Zhong, Chen Yang, Jiangfeng Hu, Yaling Jiang, Zizhen Liang, Hui Gao

Chinese Patent, Application ID: 201811287261.0, October 2018

## RESEARCH EXPERIENCES & PROJECTS

- **Efficient Computation of Probabilistic Core Decomposition of Biological Network Graphs**

Advised by Dr. Xuekui Zhang at UVic and Dr. Li Xing at USask, September 2019 – July 2021

- Designed computing method as well as filtering criteria for vertices'  $\eta$ -degree in order to properly screen out low incidence nodes and focusing on high/dense cores.
- Designed algorithm to effectively reduce large graph into small sub-graphs then perform further core decomposition efficiently.
- Applied the method to the [Biomine](#) database and experimentally validated SARS-CoV-2-human protein-protein interaction (PPI) networks, provided inference to the COVID-19 biological network and discovered novel knowledge relevant to COVID-19 and the SARS-CoV-2 virus.
- First-authored two papers, one is published in Briefings in Bioinformatics (IF=11.622), the other is published in the proceedings of the ASONAM 2021 conference.

- **The New York Stock Exchange (NYSE) Daily TAQ High-frequency Data Analysis and Prediction**

Advised by Dr. Xuekui Zhang and Dr. Ke Xu at UVic, September 2019 – November 2019

- Performed data section and data cleaning on data with orders of magnitude larger than 3.8 TiB.
- Designed pipeline to extract designated entries and dispose unwanted data noise. For instance, if during a period of time a specific stock was not traded in any of a total number of 16's US regulated stock exchanges, then we see that stock as unwanted.
- Utilized shell script to achieve efficient extraction of data.

- **Epistasis Mining Based on Genetic Tabu Algorithm and Bayesian Networks**

Advised by Prof. Jianxiao Liu at HZAU, December 2016 – June 2019

- Conducted research in swarm intelligence algorithms and Bayesian networks (BNs), improved the hill-climbing algorithm in Bayesian network structure learning by combining the algorithm with genetic tabu algorithm and mutual information entropy.
- Boosted the accuracy of the original hill-climbing algorithm from **38%** to now's **78%** (tested on the [ALARM network](#)).
- Implemented the improved algorithm into an R package: 'epiGTBN'.
- Compared the detection accuracy and efficiency of different methods ([AntEpiSeeker](#), [BEAM](#), [BOOST](#), [hill-climbing](#), [MDR](#), [SNPRuler](#)) with epiGTBN on the simulated datasets. The result obtained was found to be highly satisfactory, with epiGTBN's accuracy close to **100%**.
- Applied the improved method to the age-related macular degeneration (AMD) [dataset](#) and found interesting interactions between sensible genes in terms of disease.
- First-authored a paper published in *BMC Bioinformatics*. Co-authored a paper published in *Expert Systems with Applications*. A patent has been submitted and is currently pending.

## SOFTWARE TOOLS (by me)

- **epiGTBN**

- R package for epistasis mining based on genetic tabu algorithm and Bayesian network.
- Project page is found [here](#).

- **m3u8\_downloader**

- A tool written in Python that downloads the online m3u8 playlist and all its media segments.
- Project page is found [here](#).

## HONORS & AWARDS

For associated scholarships and honors, only the honors are listed. There are also honorary titles without any matching scholarship. The names of organizations are put in parentheses.

### 2020

- **VADA Student Research Presentation Award (VADA Summer School)**
  - Title: Performing Efficient Core Decomposition on Probabilistic Biological Network
- **UVic Graduate Award (University of Victoria)**
  - C\$8,000

### 2019

- **VADA Trainee Scholarship (NSERC CREATE)**
  - C\$16,000
- **First Class in Bachelor Thesis (Huazhong Agricultural University)**
  - Top 5%
- **Outstanding Graduates (Huazhong Agricultural University)**
  - Two consecutive wins of 'Merit Student' honor.

### 2018

- **Merit Student (Huazhong Agricultural University)**
  - Top 20%
- **Top 10 Outstanding Undergraduate Research Projects (Huazhong Agricultural University)**
  - Maize Carotenoid Gene Mining Based on Bayesian Networks

### 2017

- **National 2<sup>nd</sup> prize in Chinese Undergraduate Mathematical Contest in Modeling (China Society for Industrial and Applied Mathematics)**
  - Top 5.8% / 36,375 submitted papers.
- **Merit Contest Paper prize in 10<sup>th</sup> Central China University Students Mathematical Modeling Invitational Tournament (Hubei Society for Industrial and Applied Mathematics)**
  - Top 13% / 959 contenders.

## CONFERENCES & WORKSHOPS ATTENDED

### 2021

- [UVic Global Days](#) (Online, November 15-19)
  - Gave talk in the UVic Seeds of Change and Hope session, introduced past research experience and current career, and how they contributed to the 17 UN sustainable development goals.
- [The 2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining](#) (Online, November 8-11)
  - Presented the paper: [Multi-Stage Graph Peeling Algorithm for Probabilistic Core Decomposition](#)

- 2021 Joint Statistical Meetings (Online, August 8-12)
  - Presented the paper: [Integrative COVID-19 biological network inference with probabilistic core decomposition](#)

## 2020

- [SORA-TABA Annual Workshop - Statistical Machine Learning for Biomedical Data](#) (Online, August 20-21)
- UVic Math & Stat Biostatistics Reading Group Bi-weekly Meeting (Online, August 12)
  - Gave talk on paper: [A SARS-CoV-2 protein interaction map reveals targets for drug repurposing](#)
- [SciNet Summer Training Program on R and High Performance Computing](#) (Online, August 11-13)
- [WestGrid Research Computing Summer School](#) (Online, July 9-10)
- [Statistical Society of Canada \(SSC\) Business and Industrial Statistics Workshop: Interactive Exploratory Data Visualization in R](#) (Online, June 6)
- [SSC BioStatistics Workshop: Intro to Epidemic Modelling](#) (Online, June 5)
- [VADA Summer School](#) (Online, June 2-4)
  - Gave talk titled: [Performing Efficient Core Decomposition on Probabilistic Biological Network](#)

## 2019

- [Vancouver Machine Learning: Genomics \(VanML: Genomics\) conference](#) (Simon Fraser University, December 16)
- [WestGrid Research Computing Winter School](#) (University of Victoria, December 10-13)

## MISCELLANEA

### Standardized Test

- IELTS (Academic): 8.0 (Listening: 9, Reading: 9, Writing: 6, Speaking: 7).

### Certifications

- Huawei Certified ICT Associate in Big Data (Credential ID: 010101801193807639051437).

### Additional Skills

- Fluent in R, MATLAB, C, Java, and Python. Capable of Shell, C++, HTML, CSS, and JavaScript.
- Familiar with source management and collaboration with Git.
- Skilled in Excel, PowerPoint, XMind, Jupyter Notebook, RStudio, and typesetting with Word, Markdown as well as  $\text{\LaTeX}$ .