Haojun Li

1234 Steelwood Road, Columbus, OH 43212
(267) 678-8162 | li.14118@buckeyemail.osu.edu

EDUCATION

The Ohio State University, Columbus, OH Aug. 2023-Present
Current Ph.D. in Computer Science | GPA 4.0/4.0 (Advisor: John Paparrizos)
Columbia University, New York, NY Sept. 2021-Dec. 2022
Master of Science in Computer Science | GPA 3.838/4.0
Pennsylvania State University, State College, PA Sept. 2017-May 2021
Bachelor of Science in Computer Science | Minor in Mathematics Applications | GPA 4.0/4.0 | Degree Honors: Summa Cum Laude (Top 2% of a college's graduating class) |

Honors and Awards

- Dean's List 2017-2020
- Undergraduate Research Award: Excellence in Info. Literacy (Honored 100 students from 19 PSU campuses) May 2021
- The Evan Pugh Scholar Senior Award (The upper 0.5 percent of the respective class) Apr. 2021
- The Evan Pugh Scholar Junior Award (The upper 0.5 percent of the respective class) Apr. 2020
- Ayoub Mathematics Achievement Award ($2,799; the only one recipient that year) Apr. 2019
- The President Sparks Award (Presented to sophomores who earned GPA: 4.0) Apr. 2019
- The President's Freshman Award (Presented to freshmen who earned GPA: 4.0) Apr. 2018

PUBLICATIONS

"Modeling policy interventions for slowing the spread of artemisinin-resistant pfkelch R561H mutations in Rwanda."

Robert J Zupko, Tran Dang Nguyen, J Claude S Ngabonziza, Michee Kabera, **Haojun Li**, Thu Nguyen-Anh Tran, Kien Trung Tran, Aline Uwimana, Maciej F Boni

Nature medicine. 2023 Nov;29(11):2775-84.

"Bridging the gap: A decade review of time-series clustering methods."

John Paparrizos, Fan Yang, **Haojun Li**

arXiv preprint arXiv:2412.20582. 2024 Dec 29.

"A survey on time-series distance measures."

John Paparrizos, **Haojun Li**, Fan Yang, Kaize Wu, Jens E d'Hondt, Odysseas Papapetrou

arXiv preprint arXiv:2412.20574. 2024 Dec 29.

"A Structured Study of Multivariate Time-Series Distance Measures."

Jens E d'Hondt*, **Haojun Li***, Fan Yang*, Odysseas Papapetrou, John Paparrizos (*: co–first author)

Proceedings of the ACM on Management of Data. 2025 Jun 18 (ACM SIGMOD 2025);3(3):1-29.

PROJECT EXPERIENCE

Pe'er Lab, Columbia University **Research Assistant** New York, NY

Supervisor: Philippe Chlenski, Computer Science PhD student in Pe'er Lab, Former researcher at Argonne National Lab; Itsik Pe'er, Associate Prof. of Computer Science & Systems Biology, Chair of Center for Health Analytics, Data Science Institute

Project: Hypotheses of Applying Phylogenetic Tree to Reduce Interaction Matrix's Parameter Jan. 2022 - May 2022

- Selected and analyzed 3 research papers and applied the results as references to support the hypotheses.
- Proposed 4 hypotheses of applying the phylogenetic tree to reduce interaction matrix's parameters in the Generalized Lotka-Volterra model, based on the references.
- Implemented hundreds of lines of Python code for hypotheses and applied generated interaction matrices to perform simulations on the microbiome data simulator.
- Analyzed plots generated for each simulation and recorded unreasonable parts in each plot based on suggested criteria.
- Gave presentation on entirety of research to lab members and planned future experiments for each hypothesis.

Project: Phylogenetic Signal Measurement of Microbial Dynamics Sept. 2021 - Dec. 2021

- Developed open-source Python code; estimated phylogenetic signal from tree-structured data; generated values and plots.
- Implemented Python code to clean and process data; applied Brownian motion model and Pagel's lambda on processed data to perform sanity check and on processed project data to measure phylogenetic signal.
- Analyzed research results and proposed plans to optimize research.

- Exhibited entirety of research in detail to lab members via PowerPoint presentation.

Boni Lab, Pennsylvania State University **Research Assistant** State College, PA

Supervisor: Robert Zupko, Boni Lab's postdoctoral scholar, Computational Scientist at PSU, Center for Infectious Disease Dynamics; Maciej F. Boni, Associate Professor of Biology, Recipient of Wellcome Trust Sir Henry Dale Fellowship, 2012

Project: Beta Calibration Methods Comparisons Jan. 2021 - Aug. 2021

- Performed research on three different Rwanda beta calibration methods: exponential curve fitting with different step sizes, second-degree polynomial curve fitting with different step sizes, and Epsilon Reduction.
- Analyzed at least 30 research papers and selected 16 to obtain simulation parameters.
- Executed, compared, and analyzed three beta calibration methods; applied most suitable method based on lab criteria.
- Wrote 20-page technical report and presented to senior faculties and researchers in Boni lab; sorted data, and supplied materials and references for Boni Lab's Rwanda research paper (Our abstract has been accepted for oral presentation at American Society of Tropical Medicine and Hygiene).
- Gave presentation in Pennsylvania State University 2021 Undergraduate Exhibition; won Undergraduate Research Award; collected by Pennsylvania State University on ScholarSphere.

Project: Simulations of 561H Frequency with Fast and Slow Mutation Rates in Rwanda Aug. 2020 - Dec. 2020

- Studied at least 35 research papers and selected 19 to collect simulation arguments and data files; cleaned and processed data files via ArcMap, expedited by instructions on helper files composed by me in previous semester.
- Tested thousands of lines of code that generate spatial maps of beta parameters to identify errors; iterated Beta Calibration on ACI computing cluster to reduce epsilon value to an acceptable level and generate spatial maps of beta parameters.
- Performed model validation against reference data; compared simulated $PfPR_{2\text{ to }10}$ and reference $PfPR_{2\text{ to }10}$.
- Simulated 561H frequency with fast and slow mutation rates in Rwanda for future twenty years; ran 50 replicates for each fast and slow mutation rate; generated heatmaps of 561H frequency and plotted the trend of 561H frequency in different provinces and districts in Rwanda.
- Created 3-page references for Boni Lab members and external audiences to perform experiments.
- Analyzed research results and exhibited this research via slide presentation to faculty and researchers in Boni lab; drafted 17 pages technical report and provided materials and references for Boni Lab's Rwanda research paper (Our abstract has been accepted for oral presentation at American Society of Tropical Medicine and Hygiene).

Project: Cambodia Research Data Management and Helper Files Collation Feb. 2020 - May 2020

- Dived deep into 15 official malaria websites, like WHO and Malaria Atlas Project and their related research papers, and collected shapefiles, raster files, and numerical data for malaria in Cambodia.
- Changed projection coordinates of files; aggregated, aligned, and edited files via ArcMap for future experiment use.
- Recorded 3-page helper files to provide processing research data preparation for new lab members.

INTERNSHIP & PROFESSIONAL EXPERIENCE

Allset Home LLC **Programmer** May 2022-Aug. 2022

- Analyzed 12-page project requirements and designed project top to bottom.
- Reviewed Django; documented 32 pages Django tutorials for project members; attached corresponding codes to every section in the tutorial.
- Designed databases for company's website by using Entity-Relationship (ER) Diagram.
- Mapped designed ER diagram into a relational schema in SQL; refined and normalized SQL schema; implemented and tuned physical database design; performed front-end and back-end implementation.
- Tested applications for websites and troubleshot website problems; composed help documents to make future code maintenance and updates easier.

Boni Lab, PSU **Programmer** June 2020-Aug. 2020

- Applied Python, HTML, CSS, JavaScript, Django, psycopg2, SQL, and PostgreSQL to program database management website for Boni Lab.
- Displayed data in tables and plotted data in charts;
- Incorporated and achieved twelve features below: navigation bar; real-time data refresh; dynamic tables display; data charts display; clickable table headers for data sorting; dynamic page-flipping bar; databases managements; clone databases; special 404 error page; special loading pages; statistics summary; databases switch.
- Reduced probability of databases misuse, prevented malicious attacks, and provided straight views of data in tables and charts to improve lab members' research efficiencies by at least 15%.

Abington Campus, PSU **Programmer** May 2019-June 2019/July 2018-Aug. 2018

- Developed a Graphical User Interface (GUI) via PyQt and Python; managed to take images and videos as inputs and demonstrate OpenCV image-processing results immediately; enabled users to save results in different formats.

- Coded different applications by using Python and OpenCV and connected them with GUI.
- Programmed robotic application via Python and Arduino allowing users to control robot car by using direction keys on the keyboard, and attached this functionality to GUI.
- Constructed tens of pages of helper files for future users; accelerated co-workers' research efficiencies by at least 10%.
- Achieved real-time robot car trace tracking, and enabled users to define moving areas for robot cars by applying OpenCV, NumPy, and Python.
- Enabled robot car to change moving directions or stop based on its position related to boundary and target point by building signal connections between Python code and Arduino code.
- Explored a new method in robot navigation and obstacle avoidance; created and implemented a new algorithm called "ellipse obstacle avoidance and navigation algorithm" which calculated shortest path to reach target point while avoiding obstacles; enabled programs to precisely fit irregular obstacles into proper ellipses to maximize moving area for robot car.

TECHNICAL SKILLS

- Programming: Python, Java, C, C++, HTML, CSS, SQL, Verilog, MATLAB
- Other: Linux operating system, GitHub, Visual Studio Code, NetBeans, PyCharm, NI Multisim circuit simulation, Arduino microcontroller, OpenCV, NumPy, PyQt, MySQL, PostgreSQL, Xilinx Vivado, scikit-learn, pandas, PyTorch, Django, and ArcMap