

Xin Xiong

Email: onlybelter@outlook.com
Website: <https://onlybelter.github.io/>

EDUCATION

Shanghai Jiao Tong University

Feb. 2016 – Jan. 2021

Master of Engineering in Computer Technology, Advisor: Professor Hai Zhao.

Research interests: Machine Learning for Natural Language Processing. In particular, word embedding and its application in the representation and classification of small molecules.

Classes: The Design and analysis of Algorithms, Database Systems, Artificial Intelligence, Natural Language Understanding, Modern Algebra, Advanced Computer Architecture.

Xi'an Polytechnic University

Sep. 2006 – Jul. 2010

BE in Bioengineering

WORK EXPERIENCE

Shenzhen Institutes of Advanced Technology (SIAT), CAS

Jun 2020 – present, department of biology, Research Assistant

Southern University of Science and Technology

March 2019 – March 2020, department of biology, Research Assistant

Interdisciplinary Research Center of Biology and Chemistry (IRCBC), CAS

May 2016 – March 2019, Research Assistant

Genminix Informatics Ltd., Co.

May 2011 – April 2016, Department of R&D

I applied a patent during this period of time: Xin Xiong, Qi Li. The method of quick search and visualization of feed-forward loop. C.N. Patent 201410112193X filed on June 18, 2014 and issued on July 6, 2016.

Sangon Biotech (Shanghai) Co., Ltd.

October 2010 – April 2011, Department of Sequencing

RESEARCH PROJECTS

Gene expression profile and immune cell abundance in tumor

Jun. 2020 – present

I mainly focus on the deconvolution of bulk cell gene expression profile from primary solid

tumors and try to figure out the relation between some specific gene expression programs (such as EMT) and the abundance of immune cells in tumor microenvironment.

Allele-specific sgRNA designing in mouse-rat fusion cell Aug. 2019 – Jan. 2020

I was mainly focus on designing sgRNA library which contains 3 kind of different sgRNAs for targeting the essential genes in mouse-rat fusion cells (Li et al., 2016, Cell). The challenge of this work is that essential genes are highly conserved between mouse and rat and the sequences of each gene in these two species are very similar. So to find a good strategy to get enough allele-specific sgRNAs that can target only one species is the key point.

CCS prediction and web server development Feb. 2018 – Mar. 2019

At the beginning, I cooperated with another colleague (Zhiwei Zhou) and my main contribution was 2 user-friendly web server (MetCCS and LipidCCS) about CCS prediction and database search. In the last period of this project, I was one of the main contributors of the machine-learning-based prediction algorithm (Nat Commun, 2020). I also designed the logo for LipidCCS and MetCCS website.

Word embedding-based representation of small molecules May 2018 – Oct. 2020

This project is my master's thesis: since huge amount and structural diversity, the tasks about similarity comparison and classification for small molecular compounds become difficult. Better representation method is the basic of solving and optimizing these tasks. Inspired by natural language processing (NLP), this project use "fragment vector" and "molecular vector" to represent fragments and molecules respectively. In order to solve the problem caused by the structure of multiple branches without orientation, i.e. the relations between each two fragments are hard to represent as a linear sequence, this article shows two methods to solve this problem: TandemFragment and ParallelFragment. This article also compares different methods and hyper-parameters for training fragment vectors systematically.

Feed-forward loop in biological regulation network Aug. 2013 – Apr. 2014

To analyze regulatory relation of miRNA, lncRNA, transcript factor (TF) and target gene in an integrated way, we employed biological interaction network to represent and visualize these regulatory relations. Feed-forward loop (FFL) is a kind of small regulatory network (motif) which is comprised by 3 nodes (miRNAs/lncRNA, TF and target gene). FFL may show some specific regulatory relation in biological network (Shen-Orr SS et al., 2002). We developed an algorithm to search FFLs from big network efficiently and a tool to visualize FFL. And we applied a Chinese patent for this achievement (I am the first inventor).

PUBLICATIONS

- Zhou, Z., Luo, M., Chen, X., Yin, Y., **Xiong, X.**, Wang, R., & Zhu, Z. J. (2020). Ion mobility collision cross-section atlas for known and unknown metabolite annotation in untargeted metabolomics. Nature communications, 11(1), 1-13.
- Shen, X., Wang, R., **Xiong, X.**, Yin, Y., Cai, Y., Ma, Z., ... & Zhu, Z. J. (2019). Metabolic

reaction network-based recursive metabolite annotation for untargeted metabolomics. Nature communications, 10(1), 1-14.

- Zhou, Z., Shen, X., Chen, X., Tu, J., **Xiong, X.**, Zhu, Z. J., & Wren, J. (2018). LipidIMMS Analyzer: Integrating multi-dimensional information to support lipid identification in ion mobility-mass spectrometry based lipidomics. Bioinformatics, 1, 3.
- Zhou, Z., Tu, J., **Xiong, X.**, Shen, X., & Zhu, Z. J. (2017). LipidCCS: Prediction of Collision Cross-Section Values for Lipids with High Precision to Support Ion Mobility-Mass Spectrometry-Based Lipidomics. Analytical Chemistry, 89(17), 9559-9566.
- Zhou, Z., **Xiong, X.**, & Zhu, Z. J. (2017). MetCCS predictor: a web server for predicting collision cross-section values of metabolites in ion mobility-mass spectrometry based metabolomics. Bioinformatics, 33(14), 2235-2237.

SKILLS

Software	Python, R, TensorFlow, MySQL, Docker
Languages	Chinese (native), English (fluent)

I created sever user-friendly web servers to support the data analysis workflow using HTML/CSS/Angular on the frontend and Python/R/DjangoMySQL/Docker in the backend. I am familiar with common machine learning models (or statistics models) from linear regression to deep learning and I use these models on my daily work to analyze RNA-Seq data (bulk and single cell) or LC-MS data (identify and quantify natural small molecules). I also develop tools and algorithms by Python or R and package workflow together.

REFERENCES

Hai Zhao, Professor, Shanghai Jiao Tong University, zhaohai@cs.sjtu.edu.cn