

Curriculum Vitae

Dr. Shibiao WAN

Postdoctoral Research Associate

Princeton University

Address: Room B-221, Engineering Quadrangle, Princeton NJ 08544, USA

Phone: +1 (609) 583-0989 Fax: +1 (609) 258-3745

Website: <http://scholar.princeton.edu/wan>

Email: shibiao.wan@princeton.edu



PRINCETON
UNIVERSITY

ACADEMIC EXPERIENCE

- 2016 – present *Postdoctoral Research Associate*, [Princeton University](#), Princeton, NJ, USA
Advisor: Prof. Sun-Yuan Kung
- 2014 – 2016 *Postdoctoral Fellow*, [The Hong Kong Polytechnic University](#), Hong Kong SAR
Advisor: Dr. Man-Wai Mak
- 2014 *Research Associate*, [The Hong Kong Polytechnic University](#), Hong Kong SAR
Advisor: Dr. Man-Wai Mak
- 2013 *Visiting Scholar*, [Johns Hopkins University School of Medicine](#), MD, USA
and [CBIL lab](#) of [Virginia Tech](#), VA, USA. Advisor: Prof. Yue Wang

EDUCATION

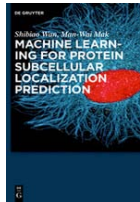
- 2010 – 2014 *Ph.D.*, [The Hong Kong Polytechnic University](#), Hong Kong SAR
Advisor: Dr. Man-Wai Mak
- 2006 – 2010 *B.Eng.*, [Wuhan University](#), Wuhan, China

RESEARCH INTERESTS

- **Bioinformatics and Computational Biology:**
 - Sequence Analysis, Protein Subcellular Localization Prediction, Database Mining;
 - Proteomics Data Analysis, Protein Type Prediction, Protein-Protein Interaction.
- **Machine Learning:**
 - Big Data Analysis, Data and Text Mining, Feature Selection, Dimension Reduction;
 - Classification/Prediction, Multi-label Classification, Regression Analysis, Optimization;
 - Semi-Supervised Learning, Transductive Learning.
- **Data Privacy:**
 - Genomic Privacy, Sequence Obfuscation;
 - Compressive Privacy, Privacy-Preserving Utility-Maximization.

Curriculum Vitae

PUBLICATIONS



BOOK

1. **S. Wan** and M. W. Mak, "[*Machine Learning for Protein Subcellular Localization Prediction*](#)", *De Gruyter*, ISBN 978-1-5015-0150-0, 2015, Germany.

JOURNALS (Impact Factor from Journal Citation Reports (JCR))

2. J. Q. Wang, C. C. Zhang, **S. Wan** and G. Peng. "Is Congenital Amusia a Connectome Disorder?: A Diffusion MRI Study Combining Tract- and Network-Based Analysis", *Scientific Reports*, 2016, submitted. (2015 Impact Factor: 5.228; JCR Q1)
3. **S. Wan**, M. W. Mak, and S. Y. Kung, "Gram-LocEN: Interpretable Prediction of Multi-Location Gram-Positive and Gram-Negative Bacterial Protein Subcellular Localization ", *Chemometrics and Intelligent Laboratory Systems*, 2017. (accepted) (2015 Impact Factor: 2.217; JCR Q1)
4. **S. Wan**, M. W. Mak, and S. Y. Kung, "FUEL-mLoc: Feature-Unified Prediction and Explanation of Multi-Localization of Cellular Proteins in Multiple Organisms", *Bioinformatics*, 2016. (accepted) (2015 Impact Factor: 5.766; JCR Q1)
5. **S. Wan**, M. W. Mak, and S. Y. Kung, "Ensemble Linear Neighbourhood Propagation for Predicting Subchloroplast Localization of Multi-Location Proteins", *Journal of Proteome Research*, 2016, vol. 15, pp. 4755–4762. (2015 Impact Factor: 4.173; JCR Q1)
6. **S. Wan**, M. W. Mak, and S. Y. Kung, "Benchmark Data for Identifying Multi-Functional Types of Membrane Proteins", *Data in Brief*, 2016, vol. 8, pp. 105–107.
7. **S. Wan**, M. W. Mak, and S. Y. Kung, "Mem-ADSVM: A Two-Layer Multi-Label Predictor for Identifying Multi-Functional Types of Membrane Proteins", *Journal of Theoretical Biology*, 2016, vol. 398, pp. 32–42. (2013 Impact Factor: 2.303; JCR Q1)
8. **S. Wan**, M. W. Mak, and S. Y. Kung, "Sparse Regressions for Predicting and Interpreting Subcellular Localization of Multi-Label Proteins", *BMC Bioinformatics*, 2016, 17:97. (2013 Impact Factor: 2.672; JCR Q1)

Curriculum Vitae

9. **S. Wan**, M. W. Mak, and S. Y. Kung, "Transductive Learning for Multi-Label Protein Subchloroplast Localization Prediction", *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2016. (accepted) (2013 Impact Factor: 1.536; JCR Q1)
10. **S. Wan**, M. W. Mak, and S. Y. Kung, "Mem-mEN: Predicting Multi-Functional Types of Membrane Proteins by Interpretable Elastic Nets", *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2016, vol. 13, pp. 706–718. (2013 Impact Factor: 1.536; JCR Q1)
11. **S. Wan** and M. W. Mak, "Predicting Subcellular Localization of Multi-Location Proteins by Improving Support Vector Machines with Adaptive-Decision Schemes", *International Journal of Machine Learning and Cybernetics*, 2015. (accepted) (2015 Impact Factor: 1.110; JCR Q3)
12. **S. Wan**, M. W. Mak, and S. Y. Kung, "mLASSO-Hum: A LASSO-Based Interpretable Human-Protein Subcellular Localization Predictor", *Journal of Theoretical Biology*, 2015, vol. 382, pp. 223–234. (2013 Impact Factor: 2.303; JCR Q1)
13. **S. Wan**, M. W. Mak, and S. Y. Kung, "mPLR-Loc: An Adaptive-Decision Multi-Label Classifier Based on Penalized Logistic Regression for Protein Subcellular Localization Prediction", *Analytical Biochemistry*, 2015, vol. 473, pp. 14–27. (2013 Impact Factor: 2.305; JCR Q2)
14. **S. Wan**, M. W. Mak, and S.Y. Kung, "HybridGO-Loc: Mining Hybrid Features on Gene Ontology for Predicting Subcellular Localization of Multi-Location Proteins", *PLoS ONE*, 2014, 9(3): e89545. (2013 Impact Factor: 3.534; JCR Q1)
15. **S. Wan**, M. W. Mak, and S. Y. Kung, "R3P-Loc: A Compact Multi-Label Predictor Using Ridge Regression and Random Projection for Protein Subcellular Localization", *Journal of Theoretical Biology*, 2014, vol. 360, pp. 34–45. (2013 Impact Factor: 2.303; JCR Q1)
16. **S. Wan**, M. W. Mak, and S. Y. Kung, "Semantic Similarity over Gene Ontology for Multi-Label Protein Subcellular Localization", *Engineering*, 2013, vol. 5, pp. 68-72.
17. **S. Wan**, M. W. Mak, and S. Y. Kung, "GOASVM: A Subcellular Location Predictor by Incorporating Term-Frequency Gene Ontology into the General Form of Chou's Pseudo-Amino Acid Composition", *Journal of Theoretical Biology*, 2013, vol. 323, pp. 40–48. (2013 Impact Factor: 2.303; JCR Q1)

Curriculum Vitae

18. **S. Wan**, M. W. Mak, and S. Y. Kung, "mGOASVM: Multi-Label Protein Subcellular Localization Based on Gene Ontology and Support Vector Machines", *BMC Bioinformatics*, 2012, 13:290. **Highly accessed** (2013 Impact Factor: 2.672; JCR Q1)

CONFERENCE PAPERS

19. **S. Wan**, M. W. Mak and S. Y. Kung, "Protecting Genomic Privacy by a Sequence-Based Obfuscation Method", *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'17)*, submitted.
20. Mert Al, **S. Wan** and S. Y. Kung, "Ratio Utility and Cost Analysis for Privacy Preserving Subspace Projection", *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'17)*, submitted.
21. **S. Wan**, M. W. Mak, B. Zhang, Y. Wang and S. Y. Kung, "Ensemble Random Projection for Multi-Label Classification with Application to Protein Subcellular Localization", *2014 IEEE International Conference on Acoustic Speech and Signal Processing (ICASSP'14)*, Florence, Italy, May 2014, pp. 6040-6044.
22. **S. Wan**, M. W. Mak, B. Zhang, Y. Wang and S. Y. Kung, "An Ensemble Classifier with Random Projection for Predicting Multi-Label Protein Subcellular Localization", *The 2013 IEEE International Conference on Bioinformatics and Biomedicine (BIBM'2013)*, Shanghai, China, Dec. 2013, pp. 35-42.
23. **S. Wan**, M. W. Mak, and S. Y. Kung, "Adaptive Thresholding for Multi-Label SVM Classification with Application to Protein Subcellular Localization Prediction", *2013 IEEE International Conference on Acoustic Speech and Signal Processing (ICASSP'13)*, Vancouver, Canada, May 2013, pp. 3547-3551.
24. **S. Wan**, M. W. Mak, and S. Y. Kung, "GOASVM: Protein Subcellular Localization Prediction Based on Gene Ontology Annotation and SVM", *2012 IEEE International Conference on Acoustic Speech and Signal Processing (ICASSP'12)*, Kyoto, Japan, Mar. 2012, pp. 2229-2232.
25. **S. Wan**, M. W. Mak, and S. Y. Kung, "Protein Subcellular Localization Prediction Based on Profile Alignment and Gene Ontology", *2011 IEEE International Workshop on Machine Learning for Signal Processing (MLSP'11)*, Beijing, China, Sep. 2011, pp. 1-6.

Curriculum Vitae

26. **S. Wan**, C. Yao, Y. Hu, G. Zhang, "A Method of Continuous Data Flow Embedded within Speech Signals", *The 2-nd International Conference on Signal Acquisition and Processing (ICSAP'10)*, Bangalore, India, Feb. 2010, pp. 362-365.

BIOINFORMATICS WEB-SERVERS

- **PolyU-Loc**: A Package of Web-Servers for Protein Subcellular Localization http://bioinfo.eie.polyu.edu.hk/Book_website/
- **Mem-mEN**: A Predictor for Membrane Protein Multi-Functional Type <http://bioinfo.eie.polyu.edu.hk/MemmENServer/>
- **SpaPredictor**: An Interface of Two Predictors for interpretable Protein Subcellular Localization <http://bioinfo.eie.polyu.edu.hk/SpaPredictorServer/>
- **GOASVM**: Single-label Protein Subcellular Localization Prediction (for Eukaryote and man) <http://bioinfo.eie.polyu.edu.hk/mGoaSvmServer/GOASVM.html>
- **mGOASVM**: Multi-label Protein Subcellular Localization Prediction (for Virus and Plant) <http://bioinfo.eie.polyu.edu.hk/mGoaSvmServer/mGOASVM.html>
- **HybridGO-Loc**: Mining Hybrid GO Features for Multi-label Protein Subcellular Localization Prediction (for Virus and Plant) <http://bioinfo.eie.polyu.edu.hk/HybridGoServer/>
- **R3P-Loc**: Compact Predictor for Multi-Label Protein Subcellular Localization (for Eukaryote and Plant) <http://bioinfo.eie.polyu.edu.hk/R3PLocServer/>
- **mPLR-Loc**: Probabilistic Predictor for Multi-label Protein Subcellular Localization (for Virus and Plant) <http://bioinfo.eie.polyu.edu.hk/mPLRLocServer/>
- **mLASSO-Hum**: An Interpretable Predictor for Human Protein Subcellular Localization <http://bioinfo.eie.polyu.edu.hk/mLASSOHumServer/>
- **Mem-ADSVM**: A Two-Layer Predictor for Multi-Label Membrane Protein Type <http://bioinfo.eie.polyu.edu.hk/MemADSVMServer/>
- **EnTrans-Chlo**: Transductive Learning for Protein Subchloroplast Localization <http://bioinfo.eie.polyu.edu.hk/EnTransChloServer/>
- **LNP-Chlo**: Linear Neighborhood Propagation for Protein Subchloroplast Localization Prediction <http://bioinfo.eie.polyu.edu.hk/LNPChloServer/>

Curriculum Vitae

- **FUEL-mLoc:** Feature-Unified Prediction and Explanation of Protein Multi-Localization of Cellular Proteins (for Eukaryote, Human, Plant, Gram-positive, Gram-negative and Virus) <http://bioinfo.eie.polyu.edu.hk/FUEL-mLoc/>
- **Gram-LocEN:** Interpretable Prediction of Subcellular Multi-Localization of Gram-Positive and Gram-Negative Bacterial Proteins (for Gram-positive Bacteria, Gram-negative Bacteria) <http://bioinfo.eie.polyu.edu.hk/Gram-LocEN/>

INVITED TALKS (EXCLUDING CONFERENCES)

1. “Linear and Kernelized Neighborhood Propagation: A Semi-Supervised Approach”, in *Department of Electrical Engineering of Princeton University*, Princeton, NJ, USA, December, 2016.
2. “Semi-Supervised Multi-Kernel Learning in Image Classification”, in *Department of Electrical Engineering of Princeton University*, Princeton, NJ, USA, October, 2016.
3. “Large-Scale in-Silico Prediction of Protein Subcellular Localization”, in *the 3-rd Wuhan University International Forum for Interdisciplinary Sciences and Engineering*, in *School of Electronic Information of Wuhan University*, Wuhan, China, April, 2016.
4. “Machine Learning for Multi-Label Protein Subcellular Localization”, in *Department of Computer Science of Hong Kong Baptist University*, Hong Kong SAR, China, May, 2015.
5. “Machine Learning Approaches for Protein Subcellular Localization”, in *Department of Statistics of The Chinese University of Hong Kong*, Hong Kong SAR, China, Apr., 2015.
6. “CPTAC Sub-Project: Customized Sample-Specific Protein Sequence Database”, in *Computational Bioinformatics and Bioimaging Laboratory (CBIL) of Virginia Tech*, VA, USA, May, 2013.
7. “mGOASVM: A Multi-Label Protein Subcellular Localization Predictor”, in *Computational Bioinformatics and Bioimaging Laboratory (CBIL) of Virginia Tech*, VA, USA, May, 2013.
8. “Semantic Similarity over Gene Ontology for Multi-label Protein Subcellular Localization”, in *2013 International Conference on Bioinformatics and Biomedical Engineering (iCBBE2013)*, Beijing, China, Sep. 2013.
9. “Protein Subcellular Localization Prediction Based on Profile Alignment and Gene Ontology”, in *the 6-th Beijing-Hong Kong International Doctoral Forum 2011*, Hong Kong SAR, China, Aug., 2011.
10. “Support Vector Data Description (SVDD) vs Support Vector Machine (SVM)”, in *Center for Multimedia Signal Processing of The Hong Kong Polytechnic University*, Hong Kong SAR, China, Mar., 2011.

Curriculum Vitae

AWARDS & HONORS

- “**Best Poster Award**” in *the 6-th Beijing-Hong Kong International Doctoral Forum 2011*, Hong Kong SAR, August, 2011;
- Postgraduate Scholarship, The Hong Kong Polytechnic University, Hong Kong SAR, 2010-2014;
- National Assistantship, Wuhan, China, 2008-2009;
- National Inspirational Scholarship, Wuhan, China, 2007-2008;
- Excellent Scholarship (the 3-rd prize), Wuhan, China, 2007-2008;
- The Freshmen Scholarship, Wuhan, China, 2006-2007.

TEACHING & TUTORING EXPERIENCE

- **2016**
 - *Kernel-Based Machine Learning* (ELE477), Princeton University;
- **2015**
 - *Object-Oriented Design and Programming* (EIE320/EIE3375), The Hong Kong Polytechnic University;
- **2014**
 - *Information Technology* (ENG2003), The Hong Kong Polytechnic University;
 - *Database Systems* (EIE3114), The Hong Kong Polytechnic University;
- **2013**
 - *Information Technology* (ENG2003), The Hong Kong Polytechnic University;
 - *Object-Oriented Design and Programming* (EIE320/EIE3375), The Hong Kong Polytechnic University;
 - *Database Systems* (EIE3114), The Hong Kong Polytechnic University;
 - *Distributed Systems and Network Programming* (EIE424/EIE4108), The Hong Kong Polytechnic University;
- **2012**
 - *Object-Oriented Design and Programming* (EIE320/EIE3375), The Hong Kong Polytechnic University;
 - *Distributed Systems and Network Programming* (EIE424/EIE4108), The Hong Kong Polytechnic University;
 - *Communication Fundamentals* (EIE331), The Hong Kong Polytechnic University

PROFESSIONAL ACTIVITIES

Curriculum Vitae

- Editorial Board Member for *Gene & Translational Bioinformatics (GTB)*;
- TPC Member for *The annual IEEE International Conference on Tools with Artificial Intelligence (IEEE ICTAI) 2016*;
- Reviewer for Journals
 - *IEEE Transactions on Neural Networks and Learning Systems (IEEE TNNLS)*; (1 time)
 - *BMC Bioinformatics*; (1 time)
 - *IEEE Transactions on Nanobioscience (IEEE T-NB)*; (1 time)
 - *Journal of Theoretical Biology (JTB)*; (1 time)
 - *Analytical Biochemistry (AB)*; (3 times)
 - *International Journal of Machine Learning and Cybernetics (JMLC)*; (3 times)
 - *Applied Mathematics and Computation (AMC)*; (1 time)
 - *Journal of Applied Mathematics (JAM)*; (1 time)
 - *Advances in Artificial Neural Systems (AANS)*; (2 times)
 - *International Journal of Biomedical Imaging (IJBI)*; (1 time)
 - *Computer Methods and Programs in Biomedicine (CMPB)*; (1 time)
- Reviewer for International Conferences
 - *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP2017)*; (3 times)
 - *2016 IEEE International Conference on Tools with Artificial Intelligence (IEEE ICTAI)*; (2 times)
 - *2015 IEEE International Conference on Bioinformatics and Biomedicine (IEEE BIBM 2015)–Workshop on Semantic Data Analytics and Bioinformatics (SDAB)*; (2 times)
- Registration coordinator for *The 8-th International Symposium on Chinese Spoken Language Processing (ISCSLP) 2012*;
- Registration coordinator for *The 19-th International Conference on Digital Signal Processing (DSP) 2014*.