

# HIDETOSHI SHIMODAIRA

SEP. 26, 2019

Professor, Graduate School of Informatics, Kyoto University  
Team Leader, RIKEN Center for Advanced Intelligence Project

@ shimo@i.kyoto-u.ac.jp    📍 Kyoto, Japan    🌐 <http://stat.sys.i.kyoto-u.ac.jp/>    in <https://www.linkedin.com/in/shimo/>  
📄 <https://orcid.org/0000-0002-3371-7724>    G <https://scholar.google.com/citations?user=LvolalsAAAAJ>

## EXPERIENCE

---

Professor, Graduate School of Informatics, Kyoto University

📅 Apr 2017 – Present    📍 Kyoto, Japan

Team Leader, RIKEN Center for Advanced Intelligence Project

📅 Sep 2016 – Present    📍 Tokyo, Japan

Professor, Graduate School of Engineering Science, Osaka University

📅 Apr 2012 – Mar 2017    📍 Osaka, Japan

Associate Professor, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology

📅 May 2005 - Mar 2012    📍 Tokyo, Japan

Lecturer, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology

📅 Jun 2002 - Apr 2005    📍 Tokyo, Japan

Visiting Scholar, Department of Statistics, Stanford University

📅 Jul 1999 - Jan 2001    📍 Stanford, CA

Assistant Professor, Department of Prediction and Control, Institute of Statistical Mathematics

📅 Jul 1996 - May 2002    📍 Tokyo, Japan

Visiting Scholar, Department of Genetics, University of Washington

📅 Oct 1995 - Mar 1996    📍 Seattle, WA

Research Fellow of JSPS, University of Tokyo

📅 Apr 1995 - Jun 1996    📍 Tokyo, Japan

## EDUCATION

---

Ph.D., Mathematical Engineering and Information Physics, The University of Tokyo

📅 1995    📍 Tokyo, Japan

M.S., Mathematical Engineering and Information Physics, The University of Tokyo

📅 1992    📍 Tokyo, Japan

B.S., Mathematical Engineering and Information Physics, The University of Tokyo

📅 1990    📍 Tokyo, Japan

## RESEARCH

---

### Multiscale Bootstrap and Post-Selection Inference

- A statistical theory and method for computing confidence measure of hypotheses via “scaling law” of bootstrap probabilities (i.e., the change of confidence value when changing the sample size)
- The method has been widely used in life sciences
- **Paper citation is about 9300 for our first four papers**
- Higher-order asymptotic theory has been developed in later research
- Now, the method is extended to “selective inference” (aka post-selection inference) for hypotheses selected after looking at data

### Transfer Learning via Covariate Shift

- The seminal paper (Shimodaira 2000) has been cited nearly 1000
- This paper developed the framework of transfer learning with density ratio, and first coined the terminology “covariate shift”

### Multi-view Representation Learning

- Theory and methods for graph embedding with neural networks
- Applications to natural language processing and image recognition

## SOFTWARE

---

### pvclust

- An R package for hierarchical clustering with p-values
- Suzuki and Shimodaira (2002) [1549 citations]
- <https://cran.r-project.org/web/packages/pvclust/index.html>

### CONSEL

- CONSEL: for assessing the confidence of phylogenetic tree selection
- Shimodaira and Hasegawa (2001) [1930 citations]
- <http://stat.sys.i.kyoto-u.ac.jp/prog/consel/>

### scaleboot

- scaleboot: Approximately Unbiased P-values via Multiscale Bootstrap
- Shimodaira (2008)
- <https://cran.r-project.org/web/packages/scaleboot/index.html>

### PAFit

- PAFit: Generative Mechanism Estimation in Temporal Complex Networks
- T. Pham, P. Sheridan and H. Shimodaira (2015)
- <https://cran.r-project.org/web/packages/PAFit/index.html>

# SELECTED PUBLICATIONS

---

## Multi-view Representation Learning and Natural Language Processing

- A. Okuno and H. Shimodaira, Hyperlink Regression via Bregman Divergence, arXiv 2019.
- G. Kim, A. Okuno, K. Fukui and H. Shimodaira, Representation Learning with Weighted Inner Product for Universal Approximation of General Similarities, Proceedings of 28th International Joint Conference on Artificial Intelligence (IJCAI-19), 5031–5038, 2019. (acceptance rate 17.8%)
- G. Kim, K. Fukui and H. Shimodaira, Segmentation-free compositional  $n$ -gram embedding, Proceedings of 2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT), 3207–3215, 2019. (acceptance rate long+short approx. 22%)
- A. Okuno and H. Shimodaira, Robust Graph Embedding with Noisy Link Weights, Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics (AISTATS), PMLR 89, 664–673, 2019. (acceptance rate 32.4%)
- A. Okuno, G. Kim and H. Shimodaira, Graph Embedding with Shifted Inner Product Similarity and Its Improved Approximation Capability, Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics (AISTATS), PMLR 89, 644–653, 2019. (acceptance rate 32.4 %)
- A. Okuno, T. Hada and H. Shimodaira, A probabilistic framework for multi-view feature learning with many-to-many associations via neural networks, Proceedings of the 35th International Conference on Machine Learning (ICML), PMLR 80, 3888–3897, 2018. (acceptance rate 25.1%)
- K. Fukui, A. Okuno and H. Shimodaira, Image and tag retrieval by leveraging image-group links with multi-domain graph embedding, 2016 IEEE International Conference on Image Processing (ICIP), 221–225, 2016.
- T. Oshikiri, K. Fukui and H. Shimodaira, Cross-Lingual Word Representations via Spectral Graph Embeddings, Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL), 493–498, 2016.
- H. Shimodaira, Cross-validation of matching correlation analysis by resampling matching weights, Neural Networks, 75, 126–140, 2016.

## Network Science

- T. Pham, P. Sheridan and H. Shimodaira, PAFit: an R Package for the Non-Parametric Estimation of Preferential Attachment and Node Fitness in Temporal Complex Networks, Journal of Statistical Software, accepted, 2019.
- M. Inoue, T. Pham and H. Shimodaira, Transitivity vs Preferential Attachment: Determining the Driving Force Behind the Evolution of Scientific Co-Authorship Networks, Unifying Themes in Complex Systems IX. ICCS 2018, Springer Proceedings in Complexity, 262–271, 2018
- T. Pham, P. Sheridan and H. Shimodaira, Joint estimation of preferential attachment and node fitness in growing complex networks, Scientific Reports, 6, 32558, 2016.
- T. Pham, P. Sheridan and H. Shimodaira, PAFit: A Statistical Method for Measuring Preferential Attachment in Temporal Complex Networks, PLoS ONE, 10, e0137796, 2015.

## Multiscale Bootstrap and Post-Selection Inference

- Y. Terada and H. Shimodaira, Selective inference after variable selection via multiscale bootstrap, arXiv 2019.
- H. Shimodaira and Y. Terada, Selective Inference for Testing Trees and Edges in Phylogenetics, Frontiers in Ecology and Evolution, 10.3389/fevo.2019.00174, 2019.
- Y. Terada and H. Shimodaira, Selective inference for the problem of regions via multiscale bootstrap, arXiv 2017.
- H. Shimodaira, Higher-order accuracy of multiscale-double bootstrap for testing regions, Journal of Multivariate Analysis, 130, 208–223, 2014.
- H. Shimodaira, Testing regions with nonsmooth boundaries via multiscale bootstrap, Journal of Statistical Planning and Inference, 138, 1227–1241, 2008.
- R. Suzuki and H. Shimodaira, Pvcust: an R package for assessing the uncertainty in hierarchical clustering, Bioinformatics, 22, 1540–1542, 2006. (Google Scholar Citations 1549)
- H. Shimodaira, Approximately unbiased tests of regions using multistep-multiscale bootstrap resampling, The Annals of Statistics, 32, 2616–2641, 2004. (Google Scholar Citations 304)
- H. Shimodaira, An approximately unbiased test of phylogenetic tree selection, Systematic biology, 51, 492–508, 2002. (Google Scholar Citations 2015)
- H. Shimodaira and M. Hasegawa, CONSEL: for assessing the confidence of phylogenetic tree selection, Bioinformatics, 17, 1246–1247, 2001. (Google Scholar Citations 1930)
- H. Shimodaira, Multiple comparisons of log-likelihoods and combining nonnested models with applications to phylogenetic tree selection, Communications in Statistics-Theory and Methods, 30, 1751–1772, 2001.
- H. Shimodaira and M. Hasegawa, Multiple comparisons of log-likelihoods with applications to phylogenetic inference, Molecular Biology and Evolution, 16, 1114–1116, 1999. (Google Scholar Citations 3840)

## Covariate Shift and Information Criterion

- Imori and H. Shimodaira, An Information Criterion for Auxiliary Variable Selection in Incomplete Data Analysis, Entropy, 21, 281, 2019.
- H. Shimodaira and H. Maeda, An information criterion for model selection with missing data via complete-data divergence, Annals of the Institute of Statistical Mathematics, 70, 421–438, 2018.
- H. Shimodaira, Improving predictive inference under covariate shift by weighting the log-likelihood function, Journal of statistical planning and inference, 90, 227–244, 2000. (Google Scholar Citations 981)