Erfan Sayyari

Contact Information

Electrical and Computer Engineering
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093

Phone: (301) 747-5505
Website: http://esayyari.github.io/main.html
E-mail: esayyari@ucsd.edu

Research Interests

Computational Biology, Phylogenetics, Statistical learning

Education

University of California, San Diego, La Jolla, California USA

Ph.D. Candidate, Electrical and Computer Engineering (Signal and Image Processing), Fall 2017 (expected graduation date: Winter 2019)

- Dissertation Topic: "New methods to estimate and assess phylogenetic trees from genome-wide data"
- Advisor: Siavash Mirarab

M.S., Electrical and Computer Engineering (Signal and Image Processing), Fall 2016

Sharif University of Technology, Tehran, Iran

B.Sc., Electrical Engineering, June, 2013

Academic Experience

University of California San Diego, La Jolla, California USA

Research Assistant **September, 2013 - present**

My research includes developing methods on large-scale data with applications in computational biology and phylogenetics.

Teaching Assistant **Spring, 2014 - present**

Duties at various times have included holding office hours and discussion sections, grading homework, computer projects, midterm and final exams

- Undergraduate courses: Engineering Probability and Stats, Circuits and Systems, Probability and Random Processes
- Graduate courses: Algorithms for biological data analysis, Random Processes

Publications

- **Sayyari, E.**, James Whitfield, and Siavash Mirarab. "Fragmentary gene sequences negatively impact gene tree and species tree reconstruction". Molecular Biology and Evolution (in press) (2017).
- Zhang, C., **Sayyari, E.** and Mirarab, S., 2017, October. "ASTRAL-III: Increased Scalability and Impacts of Contracting Low Support Branches". In RECOMB International Workshop on Comparative Genomics (pp. 53-75). Springer, Cham.
- Mai, U., **Sayyari, E.** and Mirarab, S., 2017. 11Minimum variance rooting of phylogenetic trees and implications for species tree reconstruction". PloS one, 12(8), p.e0182238. (**Best paper award at GLBIO 2017**)
- **Sayyari, E.** and Mirarab, S., 2016. "Fast coalescent-based computation of local branch support from quartet frequencies". Molecular biology and evolution, 33(7), pp.1654-1668.
- **Sayyari, E.** and Mirarab, S., 2016. "Anchoring quartet-based phylogenetic distances and applications to species tree reconstruction". BMC genomics, 17(10), p.783.
- Weibel, N., Hwang, S.O., Rick, S., **Sayyari, E.**, Lenzen, D. and Hollan, J., 2016, January. "Hands that Speak: An Integrated Approach to Studying Complex Human Communicative

Body Movements". In System Sciences (HICSS), 2016 49th Hawaii International Conference on (pp. 610-619). IEEE

- **Sayyari, E.**, Farzi, M., Estakhrooeieh, R.R., Samiee, F. and Shamsollahi, M.B., 2012, July. "Migraine analysis through EEG signals with classification approach". In Information Science, Signal Processing and their Applications (ISSPA), 2012 11th International Conference on (pp. 859-863). IEEE.

Software (Publicly available)

ASTRAL (https://github.com/smirarab/ASTRAL)
I routinely contribute to the ASTRAL. A tool well known for estimating species trees from given gene trees.
DISTIQUE (https://github.com/esayyari/DISTIQUE)
Coalescent-based species tree estimation from gene trees
DiscoVista (https://github.com/esayyari/DiscoVista)
Command-line software package for visualizing phylogenetic discordance.
Gene tree estimation pipeline (https://github.com/esayyari/bootstrap)
Bash command-line pipeline of automatic inferring and bootstrapping gene trees

Other Projects

- **Gesture/Sign Recognition** (in C/C++ and MATLAB) - A cross modal algorithm to do gesture/sign recognition based on depth and motion information using linear dynamic systems trained by PCA and CCA.
- **Conditional Random Fields for Punctuation Prediction in Python**. Implementing a Conditional Random Fields model for prediction punctuation tags for English language text. Two different techniques are used for training the model: Gibbs sampling and Collin's Perceptron.
- **Latent Dirichlet Allocation for Document Topic Discovery in Python** Using Latent Dirichlet Allocation (LDA) to discover the underlying topics of a set of documents in an unsupervised scheme based on frequency of words in documents. Gibbs sampling is used with Harmonic mean as a measure of goodness-of-fit to train the LDA.
- **Recursive Auto Encoder (RAE)** Method of Learning Meanings for Sentences Reconstructing a recursive auto encoder neural network for recognizing sentence level sentiment using a dynamic representation for sentences.

Other Experiences

Cluster maintenance **Fall 2016 - present**
Includes updating and maintenance of lab clusters, training students, and setting up parallel job running tools like *pylaucnher*

Workshops and Tutorials **Summer 2016**
Phylogenomics Symposium and Software School on ASTRAL, PASTA, and ASTRID

Summer researcher (University of Tehran) **Summer 2012**
Finding Mid-surface 3D mesh objects Using Marching Cube algorithm to find mid-surface of convoluted objects like brain cortex.

Computer Skills

- Statistical Packages: R
- Languages: C++, Python, Unix shell scripts, MatLab, some use of Java
- Applications: InkScape, LaTeX, Markdown, Git, and presentation software
- Operating Systems: Unix/Linux, Windows.
- Language: Persion, English