MD HASIN ABRAR

hasinabrar1405@gmail.com • https://github.com/hasin-abrar • hasin-abrar.github.io

EDUCATION

Pennsylvania State University, State College, PA, USA
Ph.D. in Computer Science and Engineering   Aug 2021 – Present
- Research Area: Learned Data Structures, Computational Biology
- Supervisor: Paul Medvedev

Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh
B.Sc. in Computer Science and Engineering   Feb 2015 – Apr 2019
- Thesis: An Alignment-free Method for Phylogeny Estimation using Maximum Likelihood
- Supervisor: Atif Rahman
- Focus: Bioinformatics and Computational Biology, Phylogenetic tree construction
- CGPA: 3.83 / 4.00 (**Final 2 year CGPA: 3.9125**)

WORK EXPERIENCE

Pennsylvania State University, State College, PA, USA
Research Assistant, Department of Computer Science and Engineering   Jun 2022 – Present
- Supervisor: Paul Medvedev

Pennsylvania State University, State College, PA, USA
Teaching Assistant, Department of Computer Science and Engineering   Aug 2021 – May 2022
- Course: CMPSC 465: Data Structures and Algorithms

Ahsanullah University of Science and Technology (AUST), Dhaka, Bangladesh
Lecturer, Department of Computer Science and Engineering   Jul 2019 – Jul 2021
- Selective Courses: Numerical Methods and Computer Programming, Pattern Recognition Lab, Software Development

Enosis Solutions, Dhaka, Bangladesh
Software Engineer   May 2019 – Jun 2019
- Web Development

RESEARCH INTEREST
- Learned Data Structures
- Bioinformatics & Computational Biology
- Data Science

RESEARCH EXPERIENCE

Md Hasin Abrar, Paul Medvedev. **PLA-index: A k-mer Index Exploiting Rank Curve Linearity**
24th International Workshop on Algorithms in Bioinformatics, WABI (Link)   2024

We use piece-wise linear approximation to create an efficient index data structure in terms of time and space. This data structure has improvements on top of our previous work. We also showcase multiple practical bioinformatics applications where this index is very useful.

Md Hasin Abrar, Paul Medvedev. **PLA-complexity of k-mer multisets**
Overlay, RECOMB-seq (Link)   2024

We describe, implement, and evaluate the efficient index, PLA-index, which utilizes a piece-wise linear approximation of the k-mer rank function. We also study a measure of a k-mer multiset complexity, PLA-complexity, that shows interesting patterns in the hundreds of genomes compared. Finally, we show how PLA-Index can be used to improve downstream applications like searching in suffix array and doing read alignment.

Tasfia Zahin*, **Md Hasin Abrar***, Mizanur Rahman, Tahrina Tasnim, Md. Shamsuzzoha Bayzid, Atif Rahman. **An Alignment-free Method for Phylogeny Estimation using Maximum Likelihood**
BioRxiv (Link)
(*Authors contributed equally)   2019

We use *k-mer* presence information to construct binary matrix for a range of k values. Then, one specific k is chosen by evaluating the entropy values from the constructed matrices. We use this matrix to construct an alignment-free phylogenetic tree using maximum likelihood.

POSTERS

PLA-complexity of k-mer multisets
RECOMB (Link)   2024

SmartPark: A Car Parking Solution for Dhaka City Utilizing Available Parking Spaces
Smart City Innovation Hub (Link)   2017

A Dynamic Architecture for Real Time Surveillance in Smart City
Smart City Innovation Hub (Link)   2017

AWARDS & SCHOLARSHIPS	▪ Dean's Honor List, All academic years, BUET	2015 – 2019
	▪ General Scholarship, Higher Secondary School Certificate	2014
	▪ Junior Merit Scholarship, Dhaka Education Board	2008

SELECTIVE PROJECTS

Movie Recommendation System   2019
- ▪ Users, movies and the corresponding rating data are used for training
- ▪ Uses user-item utility matrix to build an alternating least square (ALS) based recommendation system

Dimensionality Reduction using PCA and Clustering using EM Algorithm   2019
- ▪ Implements two of the most widely used unsupervised methods in machine learning: Principal Component Analysis (PCA) and Expectation-Maximization (EM)
- ▪ Uses PCA for dimensionality reduction and apply the EM algorithm for Gaussian mixture model to cluster the data with dimensionality reduced.

Template Matching in Videos   2019
- ▪ Tracks a reference object from a video file using different search techniques like exhaustive, 2D logarithmic, and hierarchical approach
- ▪ Compares the performances

N-Puzzle Solver   2018
- ▪ Popular N-puzzle game(A sliding blocks game that takes place on a k * k grid with ((k * k) - 1) tiles each numbered from 1 to N) solver.
- ▪ Uses 5 different heuristics to solve the puzzle.

Cycle Finder   2017
- ▪ Hardware project which uses both arduino and an android application
- ▪ Users are sent a notification after their cycle is moved from a pre-defined position
- ▪ Features real time video as well as GPS tracking of the cycle

GRE Word Practice (like Magoosh)   2016
- ▪ Magoosh like windows application with flashcards to practice GRE words
- ▪ Uses information from user's expertise to select which word to show

(*Find more projects from my GitHub repositories*)

SKILLS

Programming Languages
- ▪ Python, C, C++, Assembly (Intel 8086), PL/SQL, PHP, HTML, Java, JavaScript, C#

Machine Learning Libraries
- ▪ PyTorch, scikit-learn

Analysis tools
- ▪ Flex, Bison, Proteus, CISCO Packet Tracer

Scripting Language
- ▪ LaTeX, Bash

Framework
- ▪ Android, ASP.NET

Database
- ▪ Oracle, Firebase, SQLite, SQL Server

SYNERGISTIC ACTIVITIES

Bioinformatics Method Developers Community Day, Pennsylvania State University   2024
- ▪ Organizing committee

Workshop on Emerging Methods for Sequence Analysis, Pennsylvania State University   2023, 2024
- ▪ Volunteer, Session Chair (2024)

Bangladesh Student Association, Pennsylvania State University   2022 - 2023

- Web Coordinator

International Conference on Computer and Information Technology (ICCIT)   2020
- Organizing Subcommittee

AUST Student Programming Development Program   2020 - 2021
- Instructor

OTHER EXPERIENCES
- Judged AUST Intra Department Project Show
- Member of the organization committee of BUET CSE Fest 2019