ALI GHAFFAARI

0171 9499741 ali@ghaffaari.me cartoonist alighaffaari ali.ghaffaari.me

EDUCATION

HHU DÜSSELDORF | *PhD in Computer Science* Apr 2015 – Mar 2021
- *expected to defend in 2025*
- **Thesis** | Indexing Schemes for Short-Read Mapping to Pangenome Graphs
- **Supervisor** | Prof. Dr. Tobias Marschall

UNIVERSITY OF TEHRAN | *M.Sc. in Computer Engineering* Sep 2011 – Sep 2014
- *GPA: 18.80/20* ≡ *4/4* (ranked top 5)
- **Thesis** | A Molecular Dynamic Approach Based on Knowledge-based Force Function for Protein Structure Prediction
- **Supervisors** | Prof. Dr. Babak Nadjar Araabi & Prof. Dr. Mahdi Sadeghi

UNIVERSITY OF TEHRAN | *B.Sc. in Computer Engineering* Sep 2006 – Sep 2011
- *GPA: 14.6/20* ≡ *2.9/4*
- **Thesis** | Functional Classification of Beta-lactamases Based on Physicochemical Features
- **Supervisors** | Prof. Dr. Ahmad Khonsari & Dr. Hamzeh Rahimi

SELECTED PROJECTS

PAIRED-END SHORT READ MAPPING TO SEQUENCE GRAPHS | `GraphAligner` Dec 2024
- Extending GraphAligner for paired-end, short-read mapping

FAST AND SCALABLE DISTANCE VERIFICATION IN SEQUENCE GRAPHS | `diverg` Sep 2024
- Scalable index construction on CUDA and OpenMP capable of handling very large graphs
- Orders of magnitude smaller in index size using a novel sparse matrix representation
- Faster construction time, and 2.5–4x speed-up in query time

FULLY-SENSITIVE SEED FINDING INDEX FOR SEQUENCE GRAPHS | `psi` Jan 2019
- A fully sensitive method for fixed-length query searches (seed finding) in sequence graphs
- Outperformed state-of-the-art methods in index size, query time, and sensitivity

KSEQ++: FAST FASTA/Q PARSER AND WRITER | `kseqpp` Jul 2018
- C++11 re-implementation of `kseq` library
- Employs RAII design for resource management and supports asynchronous writing

PYTHON LIBRARY FOR STREAMING PROTOBUF MESSAGES | `pysteam-protobuf` Jul 2016
- With asynchronous IO support

EXPERIENCE

BIELEFELD UNIVERSITY | *Research Fellow* Bielefeld
Genome Data Science Group Apr 2021 – Dec 2024
- **Project Manager** | MSCA-ITN Algorithms for Pangenome Computational Analysis (ALPACA) Coordinating the EU-funded project aiming for training a new generation of researchers in the field of computational pangenomics
- Research in Computational Pangenomics

- "Scalable Distance Index for Validation of Paired-End Alignments in Sequence Graphs"
  - Poster presentation at International Genome Graph Symposium (IGGSY) | *Jul 2024*
- Proceedings external reviewer at
  - International Conference on Research in Comp. Mol. Bio. (RECOMB) 2022, 2024, 2025
  - Intelligent Systems for Molecular Biology (ISMB) 2022 and (ISMB/ECCB) 2023
  - Workshop on Algorithms in Bioinformatics (WABI) 2022

Heinrich Heine University Düsseldorf | *Research Fellow* Düsseldorf
Institute for Medical Biometry and Bioinformatics Apr 2020 – Mar 2021
- Research in Computational Pangenomics
- Proceedings external reviewer at Workshop on Algorithms in Bioinformatics (WABI) 2021

Max Planck Institute for Informatics | *Research Fellow* Saarbrücken
Department of Algorithms for Computational Genomics Apr 2015 – Mar 2020
- Research in Computational Pangenomics
- "Fully-sensitive Seed Finding in Sequence Graphs Using a Hybrid Index"
  - presented at ISMB/ECCB | *Jul 2019*
  - presented at 5th Workshop on Data Structures in Bioinformatics (DSB) | *Feb 2019*
- Proceedings external reviewer at
  - International Conference on Research in Comp. Mol. Bio. (RECOMB) 2016, 2017
  - ISMB/ECCB 2017
  - German Conference on Bioinformatics (GCB) 2016

Institute for Research in Fundamental Sciences (IPM) | *Student Researcher* Tehran
Department of Computer Science, HPC Laboratory Sep 2009 – Dec 2013
- Scientific computing using multi-core and many-core architectures
- **Project** | Protein Feature Calculator | Project Manager
  A web-based, scalable, and integrated toolbox for protein feature calculation
- **Project** | Protein Structure Alignment | Student Researcher
  Computing pairwise protein structure alignment on NVIDIA GPUs
- **Project** | Protein Function Prediction using ML | Student Researcher
  Functional classification of beta-lactamases enzymes using SVM classifiers (B.Sc. thesis)

Saarland University | *Part-time HPC Cluster Administrator* Saarbrücken
Department of Spoken Language Systems 2016 – 2017
- Fine-tuning, configuring, and maintaining the cluster system (SGE)

University of Tehran | *Team Member* Tehran
School of Electrical and Computer Engineering Feb 2014 - Apr 2014
- **Project** | Advanced Metering Infrastructure (AMI) System
  Designing a REST API as a part of Smart Grid Project in order to improve the efficiency and reliability of the production and distribution of electricity
- **Supervisor** | Dr. Ashkan Rahimi-Kian

GynapSys Inc. | *Part-time Remote Developer* Redwood City, CA
May 2013 – Aug 2013
- System-level implementation of algorithms for DNA base calling with real-time constraints
- Integrate and automate developed tools and services on Google Cloud

TEACHING

HEINRICH HEINE UNIVERSITY DÜSSELDORF | *Teacher Assistant* — Düsseldorf
Algorithms for Sequence Analysis — SS 2020
- **Instructor** | Prof. Dr. Tobias Marschall
- Conducting tutorial sessions and exams, as well as designing and grading homework

SAARLAND UNIVERSITY | *Teacher Assistant* — Saarbrücken
Algorithms for Sequence Analysis — SS 2016, SS 2019
- **Instructor** | Prof. Dr. Tobias Marschall
- Teaching two sessions on "Sequence Similarity Search using Locality Sensitive Hashing"
- Conducting tutorial sessions and exams, as well as designing and grading homework

INSTITUTE FOR RESEARCH IN FUNDAMENTAL SCIENCES (IPM) | *Speaker, Co-organiser* Tehran
Two Days Workshop on Cloud Computing — Mar 2012
- **Talk** | "Introduction to Amazon S3 with hands on exercises"

One Day Workshop on IBM Cell/B.E. Programming — Apr 2010
- **Talk** | "How to set up an IBM Cell/B.E. cluster using Sony PS3"
- **Talk** | "An Introduction to IBM Cell Programming"

PASTEUR INSTITUTE OF IRAN | *Instructor* — Tehran
An introduction to Linux for bioinformaticians — Feb 2011

UNIVERSITY OF TEHRAN | *Speaker, Co-organiser* — Tehran
Introduction to GNU/Linux systems — 2009, 2010, 2011

UNIVERSITY OF TEHRAN | *Teaching Assistant* — Tehran
Computer Network Laboratory — WS 2010/2011
- **Instructor** | Prof. Dr. Ahmad Khonsari

Operating Systems Laboratory — SS 2010
- **Instructor** | Prof. Dr. Nasser Yazdani

Operating Systems — SS 2010
- **Instructor** | Prof. Dr. Mahdi Kargahi

HOMA EDUCATION COMPLEX | *Instructor* — Tehran
Introduction to Statistics — 2007, 2009
Models of Computational — 2008

SERVICES

UNIVERSITY OF TEHRAN | *Graduate Student Advisor* — Tehran
ACM Student Chapter — WS 2012/2013
- Assisting core members in the planning, organisation, and execution of workshops

UNIVERSITY OF TEHRAN | *Vice Chair (2008), Secretary (2007)* — Tehran
Student Council — 2008, 2007
- Promoting student welfare and addressing issues affecting the community
- Facilitating communication between the students and the university administration
- Organising and coordinating events and exam schedules

HONOURS AND AWARDS

Top 0.1% (525 out of ca. 500,000 participants) The Nationwide University Entrance Exam 2006

Outstanding Student Award Tehran Scientific Fair, Tehran Central Dept. of Education 2005

Ranked 8th Tehran RoboCup Contest, 2nd Round Sep 2004
earning entry into the qualification phase for the Osaka RoboCup 2D Soccer Simulation League 2005

Ranked 1th Tehran RoboCup Contest, 1st Round Jul 2004

SKILLS

Programming Experienced in C, C++, Python | Familiar with Rust, Elisp, R

Technical GNU/Linux development tools (Autoconf, Automake), CMake

Workflow Snakemake

Language Fluent in English, Basic German, Native in Persian

Laboratory Basic wet-lab experience

Interests Character and digital graphic design, IoT, and system-level programming

PUBLICATIONS

[1] **A. Ghaffaari**, T. Marschall, and A. Schoenhuth. "DiVerG: Scalable Distance Index for Validation of Paired-End Alignments in Sequence Graphs". 2024.

[2] J. M. Eizenga, A. M. Novak, J. A. Sibbesen, S. Heumos, **A. Ghaffaari**, et al. "Pangenome Graphs". In: *Annual Review of Genomics and Human Genetics* 21.1 (May 2020). DOI: 10.1146/annurev-genom-120219-080406.

[3] **A. Ghaffaari** and T. Marschall. "Fully-sensitive seed finding in sequence graphs using a hybrid index". In: *Bioinformatics (Proceedings of ISMB)* 35.14 (July 2019), pp. i81–i89. DOI: 10.1093/bioinformatics/btz341.

[4] T. Marschall, M. Marz, T. Abeel, L. Dijkstra, B. E. Dutilh, **A. Ghaffaari**, et al. "Computational Pan-Genomics: Status, Promises and Challenges". In: *Briefings in Bioinformatics* (2016). DOI: 10.1101/043430.