

Bachelor's / Master's / Semester Project

Exploring New Algorithms and Hardware Architectures for Bioinformatic Studies

Context & background:

Bioinformatics is an interdisciplinary field aiming to understand large and complex biological data through mathematical and computational models that use computer programming. Bioinformatic studies consist of several genome analysis pipelines designed to enrich our understanding of a particular problem in genomics.

The fast advances in the computing units of computers in the past few decades have helped us with our bioinformatic studies by reducing the overall analysis time. However, even with the daily advances in the algorithms and our processing platforms (e.g., CPUs and GPUs), the timely and cost-efficient analysis of such pipelines still remains a challenge, mainly due to the sheer amount of new genomic data produced and the complex nature of the underlying algorithms.

Therefore, in this project, we aim to develop novel, fast, and accurate methods (at the algorithmic or hardware level) to analyze different stages of the genomic pipeline more quickly and efficiently. You will join our researchers to implement and explore new ideas, algorithms, and hardware architectures and evaluate them using real genomic data.

Expected effort:

This project can take shape as a semester, BSc, or MSc project.

In this project, the student will

- investigate different computational stages of the genomic pipeline,
- run and analyze the state-of-the-art, publicly-available algorithms and hardware implementations of methods in one or two stages,
- develop a simulation platform to compare related previously-proposed tools fairly,
- propose novel ideas to improve the task at hand and evaluate them compared to previous works

For example studies you may perform please see:

- Wood, et al., "[Improved metagenomic analysis with Kraken 2.](#)" Genome biology, 2019.
- Firtina, et al., "[Blend: A fast, memory-efficient, and accurate mechanism to find fuzzy seed matches.](#)", arXiv, 2021.
- M. Alser, et. al., "[SneakySnake: A Fast and Accurate Universal Genome Pre-Alignment Filter for CPUs, GPUs, and FPGAs](#)", Bioinformatics, 2020.
- Houtgast, et al., "[An FPGA-based systolic array to accelerate the BWA-MEM genomic mapping algorithm.](#)" SAMOS, 2015.

Requirements:

- Outstanding programming and simulation skills (C++, CUDA, HDL, or SPICE)
- Good knowledge of design and analysis of algorithms and data structures
- An interest in developing and evaluating new ideas, discovering why things do or do not work, making systems efficient and usable
- Strong work ethic

Plus Points:

- Background in bioinformatics and genomics

Contact information:

If you are interested, please email: Said Hamdioui via S.Hamdioui@tudelft.nl, Stephan Wong via J.S.S.M.Wong@tudelft.nl, and Michael Shahroodi via tahamichaelshahroodi@gmail.com.