FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG

OPTIMIZATION
UNDER UNCERTAINTY
& DATA ANALYSIS

Master's Thesis:

Phylogenetic trees under uncertainty

Topic

A phylogenetic tree is a branching diagram that represents the evolutionary relationships among species or taxa based on physical or genetic similarities and differences. It illustrates their shared evolutionary history and common ancestry, with all life on Earth theoretically part of a single phylogenetic tree. Computational phylogenetics uses algorithms to determine the most accurate representation of these relationships. In the language of mathematical optimization, a phylogenetic tree is a so-called Steiner tree (of degree three). Although Steiner trees have been studied well in the literature, they are difficult both in theory (NP-hard) and in practice.
In this thesis, we focus on building Steiner trees. The Steiner tree problem, named after Swiss mathematician Jakob Steiner, is a combinatorial optimization problem and a generalization of the minimum spanning tree problem. A minimum spanning tree connects all nodes in a graph by a tree with the smallest total sum of used edge lengths. In contrast, a Steiner tree may include additional nodes from a predefined set to further minimize the total network length, making it challenging to select the optimal Steiner points. For the phylogenetic trees, such Steiner points model the ancestors in the evolutionary history. As the evolutionary history is affected by uncertainty, the latter needs to be taken into account as well.
The first part of the Master's Thesis is a summary of the literature on phylogenetics, as well as Steiner trees. The main contribution is supposed to be the development and implementation of an algorithm to optimize phylogenetic trees under uncertainty by utilizing Bayesian methods. The thesis is supposed to start from existing algorithmic ideas that shall be developed further. This thesis topic has emerged from a cooperation with the GeoZentrum Nordbayern (FAU).

Your tasks:

- literature research
- creating an Algorithm for phylogenetic trees under uncertainty (for Steiner trees of degree <= 3)
- utilize well-known integer programming models for Steiner trees i.e., minimum cost trees in a graph where a given set of nodes is connected via a tree and inner nodes can be utilized
- utilize Bayesian methods in the tree building process
- implementation and evaluation on a given data set

Requirements

Prior knowledge in mathematical optimization.

Programming skills (e.g. Python).

FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG

OPTIMIZATION
UNDER UNCERTAINTY
& DATA ANALYSIS

Contact persons

Sebastian Denzler, Kevin-Martin Aigner
If you're interested, send an e-mail (including your Transcript of Records, a brief letter of motivation and your desired start date) to wima-abschlussarbeiten@lists.fau.de*, to* sebastian.denzler@fau.de *or to* kevin-martin.aigner@fau.de

Literatur

[1] T. W. ET AL., *Bioinformatics and phylogenetics*, 29 (2019).