

Minimum size ultrametric tree

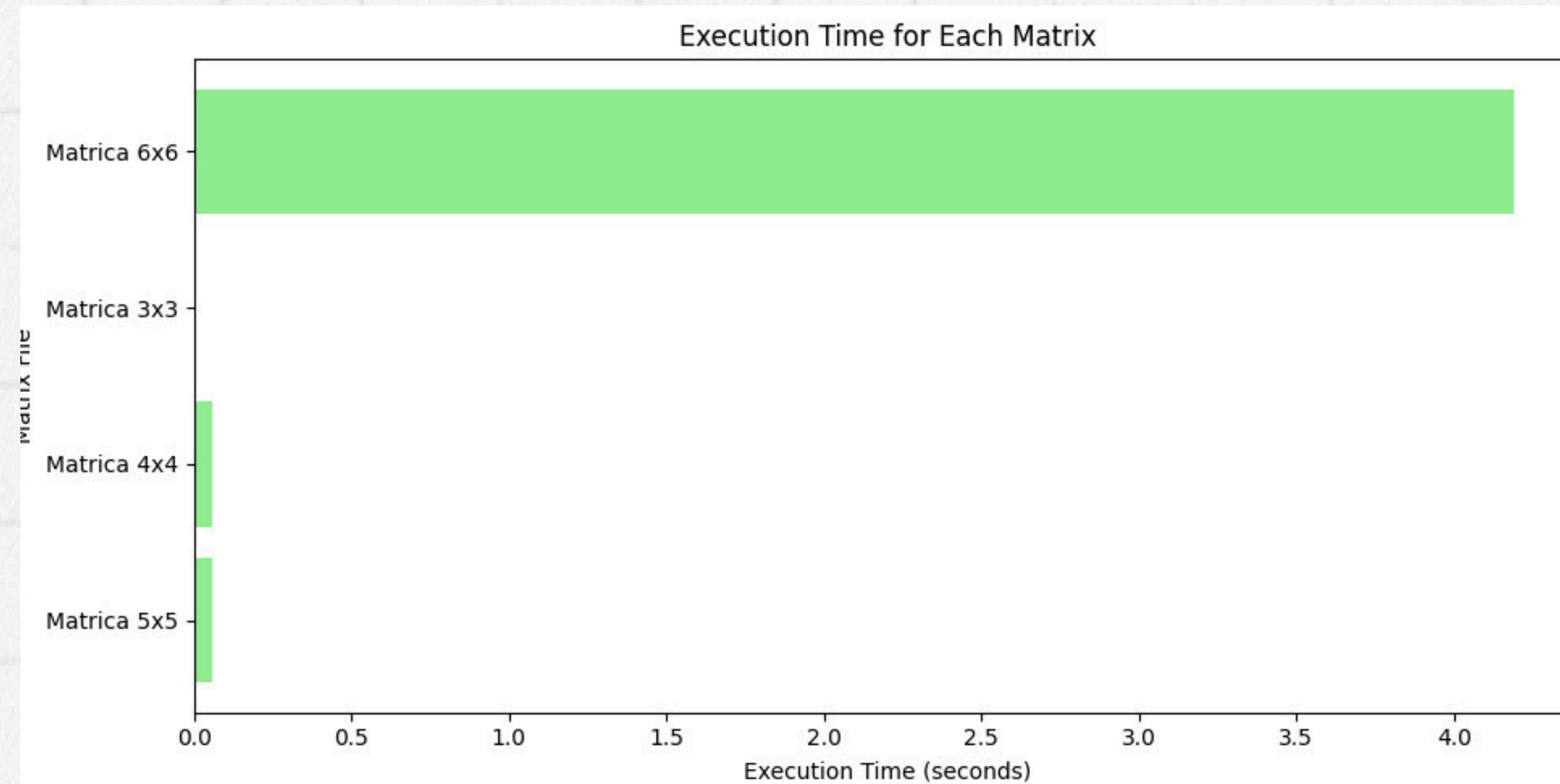
OPIS PROBLEMA

Ultrametrijsko stablo je specijalizovani oblik stabla koje se koristi u hijerarhijskom klasterovanju i srodnim oblastima. U takvim stablima, udaljenost između bilo koja dva lista preko njihovog najnižeg zajedničkog pretka ne premašuje direktnu udaljenost između njih kako je navedeno u matrici. Resenje je konstruisati stablo koje treba da bude strukturirano tako da za bilo koja dva lista i i j , zbir težina na putanji između njih u stablu bude najmanje jednaka udaljenosti $M[i,j]$ između njih u matrici.



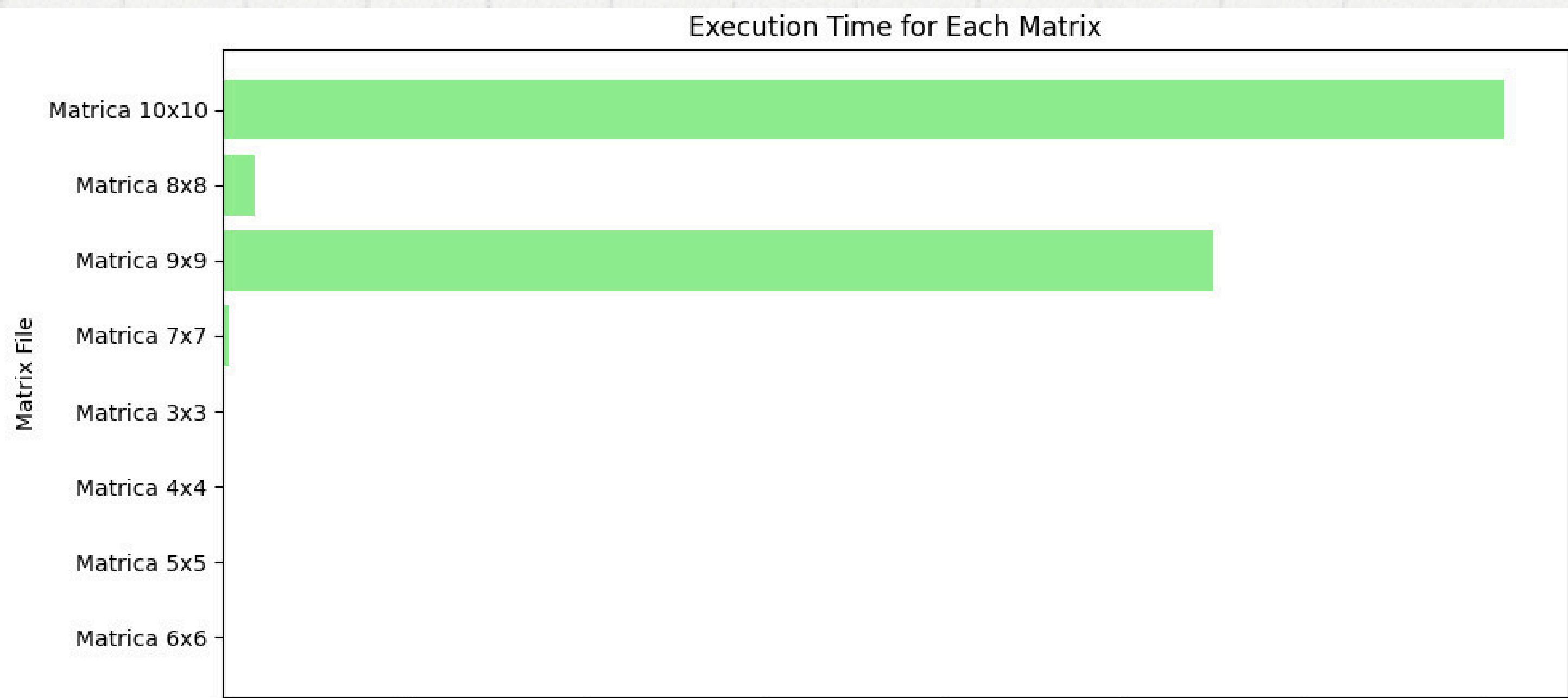
Brute force algoritam

Kreira se početno stablo dodajući sve moguće grane na osnovu matrice rastojanja M, zatim generiše sva jedinstvena ultrametrična podstabla i pronalazi ono sa minimalnom težinom. Vraća optimalno stablo i njegovu ukupnu težinu.



Branch-and-bound

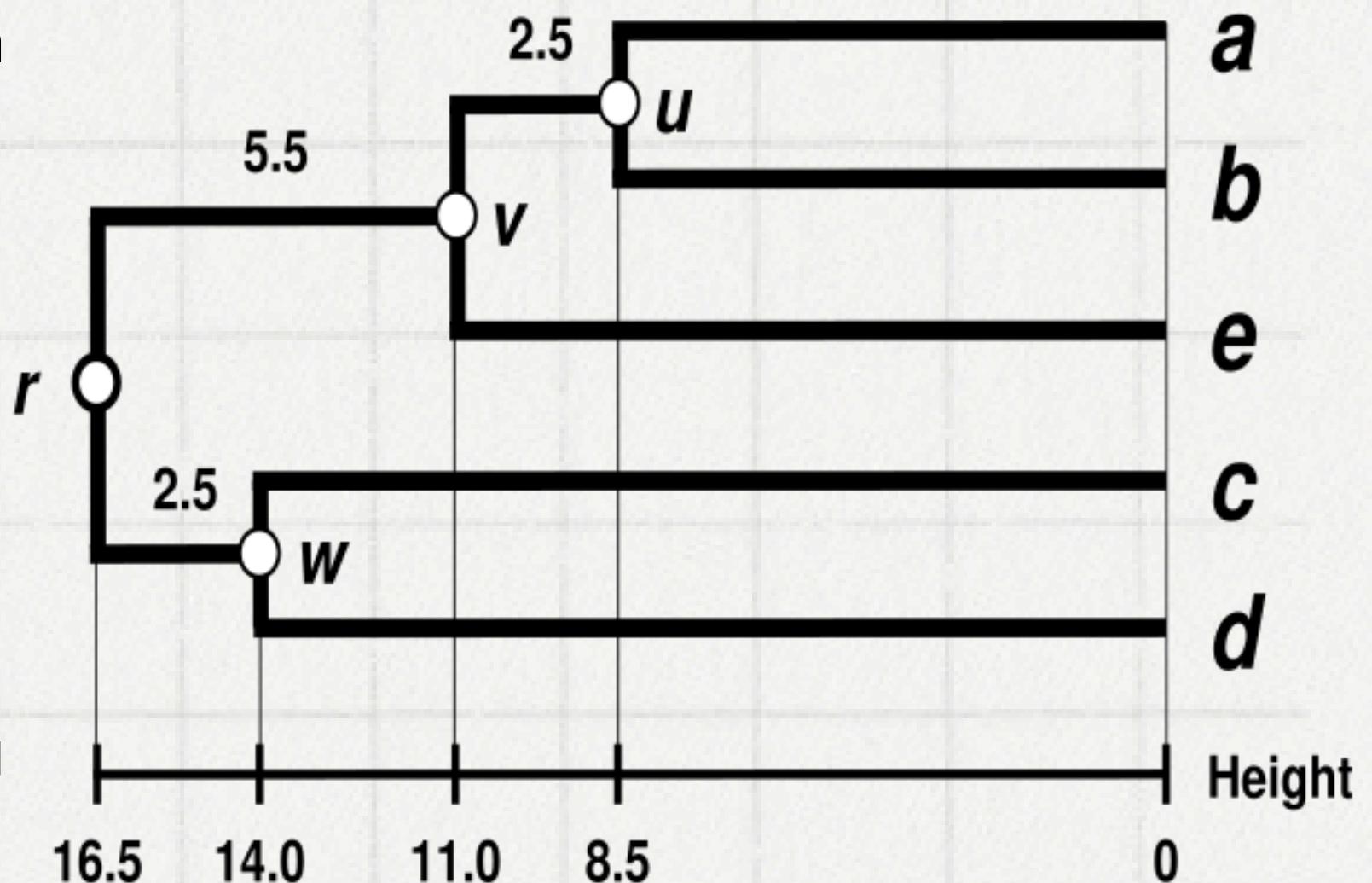
- Maxmin permutacija
- UPGMM(modifikacija UPGMA)
- Compute_lower_bound



UPGMA algoritam

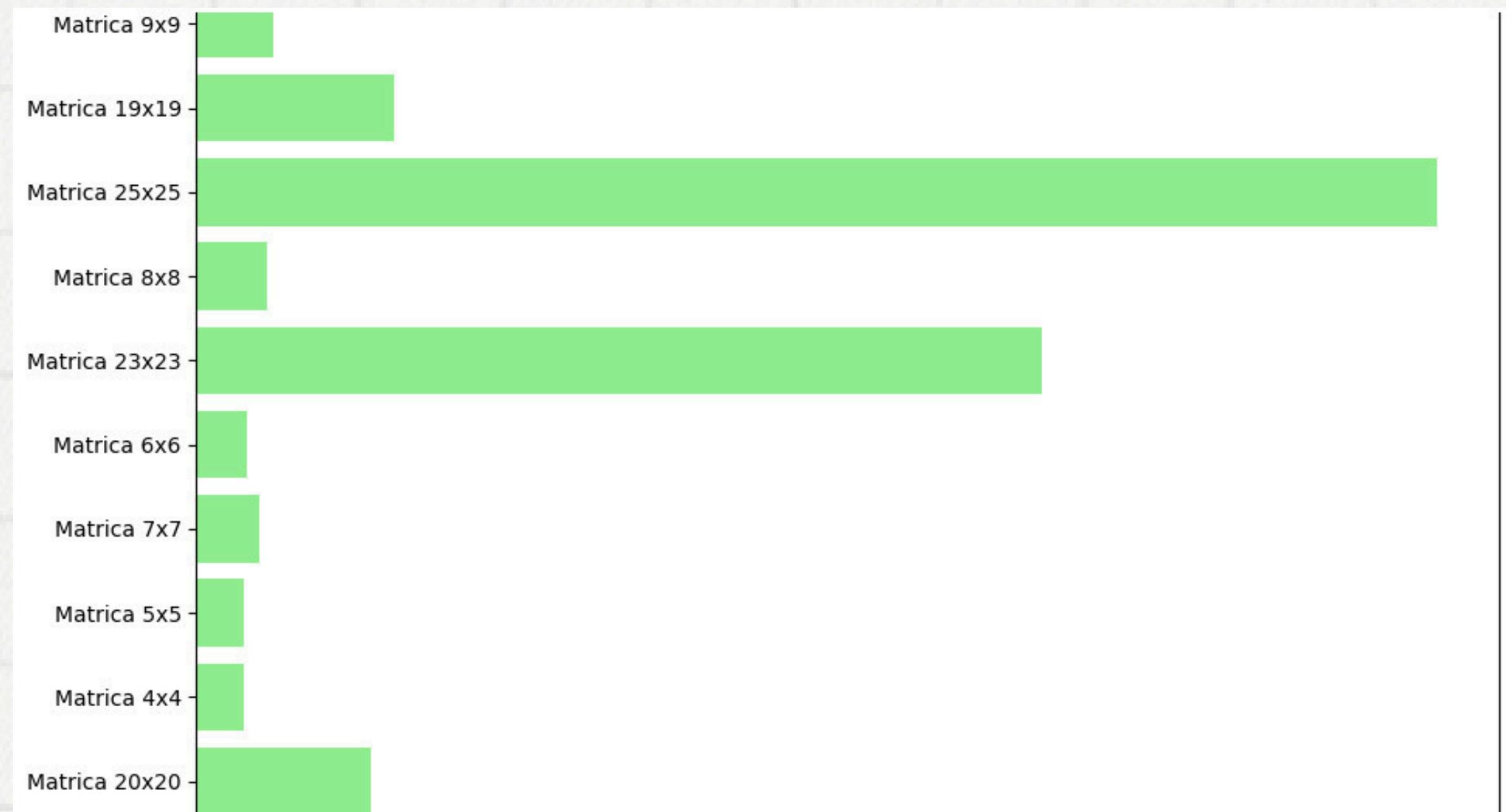


- Inicijalizacija: Svaki element (ili sekvenca) tretira se kao poseban klaster. Udaljenosti između svih parova klastera se izračunavaju i smeštaju u matricu udaljenosti.
- Spajanje: Pronalazi se par klastera sa najmanjom međusobnom udaljenošću. Ovi klasteri se spajaju u novi klaster, a njihove grane na filogenetskom stablu dobivaju dužinu proporcionalnu toj udaljenosti.
- Ažuriranje udaljenosti: Udaljenosti između novog klastera i preostalih klastera se ponovo izračunavaju kao aritmetička sredina udaljenosti između starih klastera koji su spojeni.
- Ponavljanje: Koraci 2 i 3 se ponavljaju sve dok ne ostane samo jedan klaster, koji predstavlja koren filogenetskog stabla.



Genetski algoritam

Za razliku od prethodnih algoritama metod calc_fitness je zasnovana pored tezine grana i na "kaznama" za broj ivica i balansiranje. Algoritam koristi selekciju, ukrštanje, i mutaciju da iterativno poboljšava populaciju stabala.



Hvala na paznji!

Ana Knezevic 356/2022
Pavle Ponjavic 162/2020