

Posudek práce

předložené na Matematicko-fyzikální fakultě
Univerzity Karlovy v Praze

posudek vedoucího X posudek oponenta
X bakalářské práce diplomové práce

Autor/ka: Michal Ciesla
Název práce: Identifikace meteorických rojů
Studijní program a obor: Astronomie a astrofyzika
Rok odevzdání: 2024

Jméno a tituly vedoucího/opponenta: doc. Mgr. Miroslav Brož, Ph.D.
Pracoviště: Astronomický ústav MFF UK
Kontaktní e-mail: mira@sirrah.troja.mff.cuni.cz

Odborná úroveň práce:

vynikající velmi dobrá X průměrná podprůměrná nevyhovující

Věcné chyby:

téměř žádné X vzhledem k rozsahu přiměřený počet méně podstatné četné závažné

Výsledky:

originální původní i převzaté netriviální kompilace X citované z literatury opsané

Rozsah práce:

více než veliký veliký standardní dostatečný X nedostatečný

Grafická, jazyková a formální úroveň:

vynikající X velmi dobrá průměrná podprůměrná nevyhovující

Tiskové chyby:

téměř žádné X vzhledem k rozsahu a tématu přiměřený počet četné

Celková úroveň práce:

vynikající velmi dobrá X průměrná podprůměrná nevyhovující

Slovní vyjádření, komentáře a připomínky vedoucího/oponenta:

Předložená bakalářská práce se zabývá asociacemi drah pozorovaných meteorů a rojů pomocí metriky (D-kritéria). Práce je psána česky, má 55 stran (vč. příloh), podává přehled o fyzice meteorů, obsahuje počítačový program; z formálního hlediska se zdá, že splňuje požadavky. Při bližším pohledu se ale ukáže, že má zásadní nedostatek (viz kap. 5.2 níže).

Kap. 1 popisující určování drah čerpá z Ceplechovy práce. Kap. 2 popisuje poměrně jednoduchá kritéria, v podstatě kvadráty rozdílů elementů. Hraniční hodnoty jsou zde arbitrární, např. pro D_SH je uváděno 0.2, 0.09, 0.12. Z textu není jasné, jak se tyto hodnoty určily¹, jaký je vztah k pravděpodobnosti, nebo kolik rojů lze takto odlišit od sebe.

--

¹Tj. vysvětleno až mnohem později v kap. 5.

Kap. 3 by se mohla spojit s předchozí; kritéria nejsou nijak zásadně odlišná. V kap. 4 se využívá Tisserandovo kritérium. Chybí vysvětlení, proč je časově stabilní, resp. stabilnější.

Vlastní práce studenta je v kap. 5.2 "Diskuse úspěšnosti kritérií". Jde pouze o 2 strany textu, což je naprosto nedostatečné. Není zde ani žádný obrázek, ani žádná tabulka, která by vstupní data a výstupní data jakkoliv popisovala. Není provedena žádná analýza, žádná interpretace, apod.

Jako vlastní práce by se snad mohla započítat i kap. 6, kde je popis programu pro výpočet jednotlivých kritérií jedná se však o manuál, nikoliv o vědeckou práci. Sice si cením použití pokročilejších datových struktur, použití tříd/objektů, použití vláken atd. Neměli bychom ovšem zapomínat, že program je ve své podstatě velmi jednoduchý: tzn. výpočet kvadrátu (metriky) ve dvou for-cyklech.²

--

² Pro škálování na 10^6 orbit by navíc bylo nutné použít algoritmus složitosti $N \log N$, založený např. na k-d stromu.

Proto jako interní oponent navrhuji, aby práce byla doplněna, opravena (viz též jednotlivé připomínky níže) a odevzdána znovu v červencovém termínu. Nemělo by tím vzniknout žádné zdržení; student by pokračoval ve studiu od září. Jedná se o standardní proceduru, obdobnou jako u odborných článků. Zároveň bych externím školitelům doporučil, aby zadání bylo mnohem podrobnější (nikoliv 3 řádky) a aby studentovi věnovali více času při konzultacích.

Případné otázky při obhajobě a náměty do diskuze:

Jak by vypadal algoritmus k-d stromu v Pythonu?

Za jakých podmínek by Tisserandovo kritérium neplatilo?

Jaká kritéria se používají pro rodiny asteroidů (Nesvorný et al. 2015)?

Práci

doporučuji

nedoporučuji

uznat jako diplomovou/bakalářskou.

Navrhuji hodnocení stupněm:

výborně velmi dobře dobře neprospěl/a

Místo, datum a podpis vedoucího/oponenta: Praha, 24. 5. 2024

--

Jednotlivé připomínky jsou následující (není nutné odpovídat na všechny).

s. iii

V abstraktu nejsou zmíněné žádné výsledky.

s. 3

1. věta úvodu nezmiňuje meteority, jen v poznámce pod čarou, a to jako "matoucí termín", což je nevhodné. 1. věta je zásadní větou, měla by být do-ko-na-lá.

tzv. termíny

namísto "ohromným způsobem" psát o kolik

Pozn. S větou: "Jelikož ale strojové učení funguje v principu do velké míry jako „černá skříňka,“ není prakticky příliš možné popisovat vlastnosti či fyzikální pozadí těchto metod." naprosto zásadně nesouhlasím! Samozřejmě je možné popisovat vlastnosti, proces učení, vstupní data, výstupní data, ověřování, interpretaci, vč. fyzikální interpretace, a to obzvláště tehdy, když vstupní data obsahují předpočítané fyzikální modely...

předložky na koncích řádků

s. 5

"odstranit" -> "popsat"

hvězdy jsou stacionární pouze je-li přístroj na montáži

"tečkovanou" -> "čárkovanou"

s. 6

celkové množství světla je dáno průměrem, nikoliv ohniskem

tzv. "zesilovač" sice pomáhá, když je kvantová účinnost detektoru nízká, protože z každého fotonu vznikne vícero fotonů, avšak rozhodně nezvyšuje množství světla "zvenku", tudíž ani nemůže snížit Poissonův šum

"druživ" -> "družic"

s. 7

v obecnějším pojetí je třeba parametrů 6 (vč. polohy v dráze)

"okraj elipsy"?

s. 10

Slunce bych příště prosil bílé, matné, malé.

s. 14

"... s jejíž pomocí odstraňujeme nepřesnosti pozorování."

Toto nelze! Lze s nimi počítat.

s. 15

"Průměrnou rychlost v nalezneme integrací."

Čeho? Rychlost je derivací polohy. Rychlost je integrálem zrychlení.

s. 27

geocentrická rychlost nemůže být zároveň krátkodobě oscilující
a zároveň prakticky konstantní v řádech tisíců let

Jak souvisí Tisserandův parametr s problémem tří těles?

s. 33

Tab. 1 je převzatá z práce Galligana?! Čili se naprogramovaným programem nepočítalo vůbec nic...

Problémy pro $a \rightarrow \infty$, $e > 1$ jsou triviální.

Nikde se nic nepíše o poskytnutých datech. Dle osobní konzultace šlo o 3 dráhy meteorů, což je zcela nedostatečné.

Nikde se nic nepíše o seznamu rojů.

s. 34

Co jsou "nejkonzistentnější výsledky"?

Co je "dobrá úspěšnost"?

Co jsou "nejhorší výsledky"?

Celá kap. 5.2 je psána tímto vágním jazykem. Namísto toho by měly být provedené simulace se syntetickými populacemi, výpočty pravděpodobnosti pro chybné přiřazení a především by se měla studovat observační data, např. hledat nové meteorické roje.

s. 46

1. odstavec závěru zní, jako by se počítaly dráhy, avšak se nepočítaly.

V závěru je opět vágní je hodnocení jednotlivých kritérií:

"svolné"

"slibné"

"naprosto obstojné"

"v praktické použitelnosti pokulhává"

Co si pod tím má člověk představit?

Očekával bych spíše syntetické populace a procenta meteorů, které byly identifikované chybně.

Pro fyziky je srozumitelný též... Fortran.

Ted' vážně, čitelnost je dána něčím jiným než programovacím jazykem.

Za prvé by složitost programu měla odpovídat složitosti problému!

Za druhé je na pováženou, když většina kódu je načítání vstupu:

61 hlavní program

436 obsluha vláken

447 vstup/výstup

413 konfigurace

124 D-kritéria

Za třetí je nejasné, proč nebyla použita knihovna numpy, která je pro rychlé výpočty kritická. Alternativně lze kritické části kódu psát pomocí vlastních knihoven ve Fortranu. Viz např.

<https://github.com/scraptechguy/tvet/blob/main/src/Main.py>

```
Shadowing = fmodpy.fimport("src/Shadowing/shadowing.f90")
```

Proč byl program testován, cituji, "pouze na relativně malém vzorku dat, který nám byl pro jeho vývoj poskytnut"?

Proč nebyl školitelem poskytnut větší vzorek?

s. 47

Proč se cituje:

<https://www.unicode.org/emoji/charts/full-emoji-list.html>

zatímco se cituje pouze 8 odborných prací z let 1963-2001?

odkaz nefungoval (opraveno):

<https://github.com/Akimayo/MeteorShowerIdentification>