



# Posudek oponenta závěrečné práce

Oponent práce: Ing. Magda Friedjungová, Ph.D.  
Student: Matyáš Turek  
Název práce: Analýza míry skupinové centrality reálných sociálních sítí  
Obor / specializace: Znalostní inženýrství  
Vytvořeno dne: 9. června 2023

## Hodnotící kritéria

### 1. Splnění zadání

- ▶ [1] zadání splněno
- [2] zadání splněno s menšími výhradami
- [3] zadání splněno s většími výhradami
- [4] zadání nesplněno

Zadání práce bylo splněno.

### 2. Písemná část práce

80 /100 (B)

Práce je psána v češtině a je logicky členěna. Teoretická část práce se zabývá potřebnými definicemi a pojmy. V praktické části jsou pak představeny použité datasety. Zde bych jako poznámku pod čarou uvítala odkazy na jednotlivé webové služby a v tabulce 3.1 informaci o počtu hran po vzorkování. V sekci 3.3 se objevuje slovní spojení "predikovat různé míry". Z dosavadního textu není jasné, zda má jít o nějaké sofistikované modelování nebo zda tím autor myslí pouze naivní odhad míry, který porovná s vypočítanou hodnotou. Pokud se jedná o naivní odhad, pak mi není jasná jeho přidaná hodnota. Obrázek 3.1 mi s danou barevnou škálou nepřijde čitelný, resp. nemám pocit, "že nejpočetnější skupinou bude nejpíše skupina tmavě modrých vrcholů", jelikož se zde vedle sebe i na sobě vyskytují vrcholy různě intenzivní modré a fialové barvy. V tabulkách s měřením centralit by bylo vhodné uvést alespoň tučně nejlepší hodnotu pro každou centralitu, takto je to trochu nepřehledné. Podle komentářů k tabulkám je predikcí opravdu nejspíš pouze odhad "top" skupin podle vizualizací (metoda "kouknu a vidím"). Můžeme na to pohlížet jako na velmi stručnou interpretaci vizualizovaného grafu, ale nelíbí se mi označení "predikce" a vůbec jakékoliv zdůrazňování ve smyslu "srovnáme naměřené hodnoty s našimi predikcemi". Srovnání by mělo smysl, pokud by se student více zabíral vizuální prezentací grafů, tzn. řešil použitý algoritmus pro layout grafu (např. minimalizace průniků hran), obarvení apod. V tabulce 3.7 student předkládá korelaci naměřených centralit, ale není z popisu jasné, jaké přesně hodnoty jsou na vstupu. V práci by se hodilo také zmínit časovou a výpočetní náročnost jednotlivých výpočtů.

Práce je věcná a rozumného rozsahu vůči řešené problematice. Obsahuje ale poměrně dost překlepů a několik gramatických chyb. U popisu tabulek je ponecháno anglické označení "Table". Student dobře pracuje s referencemi, místy se však uchyluje k výrokům, kde reference chybí (např. str. 7 "existují mnohé přístupy" - zde by se hodil výpis několika referencí s danými přístupy). Některé části by mohly být uvedeny jako poznámky pod čarou, např. vysvětlení NP-těžkého problému.

### **3. Nepísemná část, přílohy**

80 /100 (B)

V nepísemné části student odevzdal jeden jupyter notebook, který obsahuje celé řešení, včetně vzorkování dat. Další přílohou jsou data, která jsou potřeba ke spuštění práce. Odevzdaný kód je komentovaný, dobře strukturovaný a znovu spustitelný.

### **4. Hodnocení výsledků, jejich využitelnost**

75 /100 (C)

Student implementoval 3 skupinové míry centrality, které otestoval na 5 reálných datasetech. Dané výsledky na grafech také vizualizoval. Práce je to dobrá, ale myslím, že je zde ještě prostor pro invenci. Student mohl své výsledky porovnat s již existujícími implementacemi, případně si více pohrát s vizualizací jednotlivých sítí či s analýzou dat, aby si více procvičil znalosti z oboru BI-ZI. Nicméně tyto připomínky jsou nad rámec zadání. Jak student zmiňuje v závěru, na práci lze dále navázat implementací efektivnějších algoritmů.

## **Celkové hodnocení**

75 /100 (C)

Předloženou práci navrhuji hodnotit klasifikačním stupněm C. Navržené zadání považuji za méně obtížné, student dané problematice porozuměl a odvedl dobrou práci.

## **Otázky k obhajobě**

Nemám žádné otázky.

## **Instrukce**

### **Splnění zadání**

Posudte, zda předložená ZP dostatečně a v souladu se zadáním obsahově vymezuje cíle, správně je formuluje a v dostatečné kvalitě naplňuje. V komentáři uveďte body zadání, které nebyly splněny, posudte závažnost, dopady a případně i příčiny jednotlivých nedostatků. Pokud zadání svou náročností vybočuje ze standardů pro daný typ práce nebo student případně vypracoval ZP nad rámec zadání, popište, jak se to projevilo na požadované kvalitě splnění zadání a jakým způsobem toto ovlivnilo výsledné hodnocení.

### **Písemná část práce**

Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posudte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti.

Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře. Posudte správnost používání formálních zápisů obsažených v práci. Posudte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 52/2021, článek 3.

Posudte, zda student využil a správně citoval relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami. Zhodnoťte, zda převzatý software a jiná autorská díla, byly v ZP použity v souladu s licenčními podmínkami.

### **Nepísemná část, přílohy**

Dle charakteru práce se případně vyjádřete k nepísemné části ZP. Například: SW dílo – kvalita vytvořeného programu a vhodnost a přiměřenost technologií, které byly využité od vývoje až po nasazení. HW – funkční vzorek – použité technologie a nástroje, Výzkumná a experimentální práce – opakovatelnost experimentů.

### **Hodnocení výsledků, jejich využitelnost**

Dle charakteru práce zhodnoťte možnosti nasazení výsledků práce v praxi nebo uveďte, zda výsledky ZP rozšiřují již publikované známé výsledky nebo přinášející zcela nové poznatky.

### **Celkové hodnocení**

Shrňte stránky ZP, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení nemusí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích. Obecně platí, že bezvadně splněné zadání je hodnoceno klasifikačním stupněm A.