

# Posudek bakalářské práce

Matematicko-fyzikální fakulta Univerzity Karlovy

**Autor práce** Matúš König  
**Název práce** Detection of Influential Individuals, Communities, and Link Prediction in Social Networks  
**Rok odevzdání** 2023  
**Studijní program** Informatika  
**Specializace** Umělá inteligence

**Autor posudku** Jan Hric **Role** Oponent  
**Pracoviště** KTIML MFF UK

Prosím vyplňte hodnocení křížkem u každého kritéria. Hodnocení *OK* označuje práci, která kritérium vhodným způsobem splňuje. Hodnocení *lepší* a *horší* označují splnění nad a pod rámeček obvyklý pro bakalářskou práci, hodnocení *nevyhovuje* označuje práci, která by neměla být obhájena. Hodnocení v případě potřeby doplňte komentářem. Komentář prosím doplňte všude, kde je hodnocení jiné než *OK*.

**K celé práci** lepší O horší nevyhovuj

	lepší	O K	horší	nevyhovuj e
Obtížnost zadání	<input type="checkbox"/>	X	<input type="checkbox"/>	<input type="checkbox"/>
Splnění zadání	<input type="checkbox"/>	X	<input type="checkbox"/>	<input type="checkbox"/>
Rozsah práce ... <i>textová i implementační část, zohlednění náročnosti</i>	X	X	<input type="checkbox"/>	<input type="checkbox"/>

**Komentář** Zadání je poměrně široké. Většina práce je věnována detekci komunit. Práce je různorodá, textová i implementační část je zvládnutá a je vidět snaha o dotažení detailů. Práce má rozsáhlé přílohy.

**Textová část práce** lepší O horší nevyhovuj

	lepší	O K	horší	nevyhovuj e
Formální úprava ... <i>jazyková úroveň, typografická úroveň, citace</i>	<input type="checkbox"/>	X	<input type="checkbox"/>	<input type="checkbox"/>
Struktura textu ... <i>kontext, cíle, analýza, návrh, vyhodnocení, úroveň detailu</i>	X	X	<input type="checkbox"/>	<input type="checkbox"/>
Analýza	X	X	<input type="checkbox"/>	<input type="checkbox"/>
Vývojová dokumentace	<input type="checkbox"/>	X	<input type="checkbox"/>	<input type="checkbox"/>
Uživatelská dokumentace	<input type="checkbox"/>	X	<input type="checkbox"/>	<input type="checkbox"/>

**Komentář** Práce je psána anglicky. Na začátku je popis metrik a algoritmů zaměřený na hledání komunit, konkrétně rozkladu sítě na komunity. Algoritmy jsou i převzaté z literatury, i navrženy vlastní varianty, například dvoufázové alg. S využitím skriptů a konfiguračních souborů jsou získána data, která jsou dále statisticky a graficky zpracována. Student se snažil o automatizaci a opakovatelnost výsledků. Práce dává představu o použitelnosti a škálovatelnosti algoritmů pro hledání komunit.

**Klady:** Tab. 2.1 s přehledem složitosti algoritmů. Používání knihoven pro kompresi grafů (METIS) a s algoritmy, statistické zpracování v ipython noteboocích, grafické výstupy.

Návrh nových kombinovaných variant algoritmů, i když neuspěly. Některé alg. jsou rychlé a následně škálovatelné.

**Zápory:** Vzorec 2.4/s.19 je divný, drobné chyby okolo (indexy). Poslední odstavec 2.3.5, první věta, je asi nepravdivá, případně dvojznačná. Druhý experiment se zvětšováním počtu tříd dává základní představu o změnách komunit při zvětšování jejich počtu, ale je to kazuistika, tj. není vypovídající (např. o stabilitě komunit, kap. 4.2, 5.2). Testování je jen na

neorientovaných datových sadách. Girvan-Newman alg. měl dostat předzpracovaná data pro Enron sadu tak, že nesouvislé části grafu (jako outliersy) se odstraní. Pak by obr. 5.16 dával lepší smysl.

Další možnosti: Pro pomalé alg. navrhnout *aproximační* varianty, které by využívaly zrychlené metriky. (Např. rekurzivní Kernighan-Lin rozkládá a/nebo slučuje komponenty *hladově* na poslední vrstvě, Girvan-Newman vybírá v metrice *vzorek* ze všech cest.) Řešená úloha je rozklad sítě, chybí mi zobecnění na hledání jedné komunity (při vhodném způsobu zadání, některé metriky jsou na to použitelné). Kap. 5.1.4: je překvapující, že jen jedna ze tří metrik byla použitelná pro odhad počtu komunit.

### Implementační část práce

	lepší	O K	horš í	nevyhovuj e
Kvalita návrhu ... <i>architektura, struktury a algoritmy, použité technologie</i>	<input type="checkbox"/>	X	<input type="checkbox"/>	<input type="checkbox"/>
Kvalita zpracování ... <i>jmenné konvence, formátování, komentáře, testování</i>	X	X	<input type="checkbox"/>	<input type="checkbox"/>
Stabilita implementace	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Práce používá různorodé nástroje a knihovny vhodným způsobem. Práce je několik programů pro jednotlivé části, tj. není to monolit. Programy a konfigurace jsou dotažené, je vidět snaha o automatizaci a opakovatelnost. Knihovna METIS komprimuje grafy a následně podporuje škálovatelnost. Zpracování dat (statistiky, zobrazení) je pomocí ipython notebooků. Součástí práce je nástroj na vizualizaci komunit. Program jsem si nechal předvést.

**Celkové hodnocení** Výborně

**Práci navrhuji na zvláštní ocenění** Ne

**Datum** C22. června 2023

**Podpis**