

Posudek bakalářské práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce Daniel Trlifaj
Název práce Graphlets in Complex Networks
Rok odevzdání 2023
Studijní program Informatika
Studijní obor Obecná informatika

Autor posudku Martin Černý Oponent
Pracoviště Katedra aplikované matematiky

K celé práci

lepší OK horší nevyhovuje

	lepší	OK	horší	nevyhovuje
Obtížnost zadání		X		
Splnění zadání	X			
Rozsah práce <i>... textová i implementační část, zohlednění náročnosti</i>	X			
<p>Práce je věnována tzv. grafetům (indukovaným podgrafům zakořeněným ve vrcholu) a jejich roli v analýze a porovnávání grafů a komplexních sítí. Práce začíná úvodem do komplexních sítí. Je představen jak historický kontext, tak metody pro popis a modelování komplexních sítí. Druhá část práce je věnována samotným grafetům; autor uvádí základní definice, které ilustruje na příkladech, shrnuje dosavadní výsledky týkající se jejich vlastností a následně představuje řadu vlastních pozorování a tvrzení o vlastnostech grafetů. Třetí a poslední část práce se zabývá rolí grafetů v komplexních sítích. Zde autor představuje metodu na porovnávání komplexních sítí za pomoci grafetů, aplikuje ji na reálných příkladech, které porovnává s dříve představenými modely. Na základě těchto výsledků vyplývá, že tzv. AB model se zdá nejvíce vhodným pro modelování komplexních sítí. Na konci obou kapitol autor shrnuje dosažené výsledky a navrhuje několik směrů vhodných pro budoucí výzkum v této oblasti.</p> <p>Práce splňuje v plné míře zadání a nad rámec zadání odvozuje autor výsledky o vlastnostech grafetů. Celkově je práce velmi čtivá a slouží jako dobrý výchozí bod pro studium jak grafetů, tak také komplexních sítí. Obvláště dobře sepsaný je úvod do komplexních sítí, kde se autorovi podařilo najít rovnováhu mezi širším představením tématu a mírou detailů, do kterých zabíhá.</p>				

Textová část práce

lepší OK horší nevyhovuje

	lepší	OK	horší	nevyhovuje
Formální úprava <i>... jazyková úroveň, typografická úroveň, citace</i>	X			
Struktura textu <i>... kontext, cíle, analýza, návrh, vyhodnocení, úroveň detailu</i>	X			
Analýza	X			

Úprava práce je na velmi dobré úrovni (z hlediska typografie, struktury, vyjadřování atp.). Práce je psána v angličtině a obsahuje minimum překlepů, či drobných gramatických nedostatků (např. chybějící člen, spelling některých slov, ...). V kapitole 2, která se zabývá zavedením a analýzou grafletů, je u základních definic a značení prostor pro vylepšení (zjednodušení, vyjasnění). Zde je třeba zdůraznit, že na vině není samotný autor, neboť definice a značení jsou v rámci existující literatury nekonzistentní a vedou k mnoha nedorozuměním i mezi odbornou veřejností. V návaznosti na tuto těžkopádnost jsou důkazy některých tvrzení a vět hůře čitelné, nicméně myšlenky, na kterých důkazy stojí jsou korektní. Vezmeme-li v úvahu, že se jedná o studenta bakalářského studia, vypořádal se s těmito problémy i tak velice dobře. S přihlédnutím k výše zmíněnému hodnocení bych studentovi doporučoval výsledky publikovat i v odborném žurnálu. K práci mám dále několik jmenovitých poznámek a otázek:

- str. 15, druhý odstavec po Def. 1: ...and g_i with $i \leq 72$... - g_i definováno až v Def. 2.
- str. 15, Def. 3: Navrhoval bych pojmenovat *Graphlet touching* nebo *Touch*. Také mi není jasné, proč se ve značení objevuje \sim pro $g_j(G)\tilde{v}$?
- str. 15, Def. 3: ...we can denote the relation of touching without G_i as $g_j(G)\tilde{v}$ - navrhoval bych přeformulovat.
- str. 16, Def. 4: Nebyli mi až do pozdější části práce jasné, proč *Graphlet degree vector* značíme zkratkou *gdd* (*graphlet degree distribution*). Značení $\{p, \dots, q\} - gdd_G(v)$ mi připadá těžkopádné. Potřebujeme nutně množinovou podobu, když uvádíme rozsah od p do q ? Nestáčílo by $(p, q) - gdd_G(v)$? Případně zjednodušit na $gdd_G(v, p, q)$ nebo něco podobného. Také bych zvážil jiné značení pro $n_i(G_j)_v$ a $g_i(G)\tilde{v}$.
- str. 18, Fig. 2.4: Téměř identický s Fig. 2.2. Využil bych Fig. 2.2 a Fig 2.4 smazal.
- str. 19, výraz (2.4): Nejedná se o $2n_3(G)_v$ namísto $2n_2(G)_v$?
- str. 19, poslední odstavec: Pro odvození výrazu (2.6), kde se nachází obecné G se analyzuje nejprve menší graf v Fig. 2.5., který je nicméně také značen G , když se odvodí vztah $n_9(G)_x + n_{12}(G)_x = c(y, z)$. Při prvním čtení byla tato shoda jmen matoucí. Navrhoval bych oba případy rozlišit.
- str. 20: ... G^4 *isomprhic*... $\rightarrow G_4$ *isomorphic*
- str. 21, Theorem 1: Je matoucí, že *Reconstruction conjecture* je značená jako Theorem.
- str. 23, Lemma 2: Není mi úplně jasná první část argumentu, tedy proč maximální počet grafletů dostáváme pro úplné grafy. Chtělo by více vyjasnit argument.
- str. 23, Lemma 3: ...*a path that have nonzero entries will...* \rightarrow ...*nonzero entries in graphlet degree vector...*
- str. 24: ... $G\mathcal{G}_\gamma^n$... - mělo být $G \in \mathcal{G}_\gamma^n$
- str. 24: ... $\{2\} - gdd_G = \{2\} - gdd(H)$... - sjednotit značení
- str. 27, Theorem 8: Nemá být ...*have the same* $\{2, \dots, \mu - 1\} - gdd?$
- str. 38, Theorem 12: Ekvivalentní s Theorem 1. Stejný problém jako u Theorem 1.
- str. 45, výraz (3.2): Na pravé straně rovnosti chybí uvést, přes co se počítá suma.

Implementační část práce

lepší OK horší nevyhovuje

Nebyla součástí práce.

Celkové hodnocení Výborně
Práci navrhuji na zvláštní ocenění Ano

Datum 20. srpna 2023

Podpis