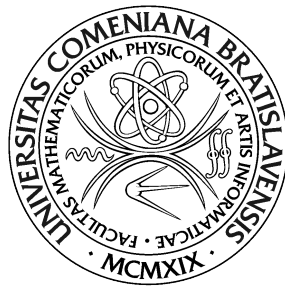


COMENIUS UNIVERSITY IN BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS AND
INFORMATICS



GRAPHLET STRUCTURE ANALYSIS OF THE REAL NETWORKS

Master's thesis

COMENIUS UNIVERSITY IN BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS AND
INFORMATICS



GRAPHLET STRUCTURE ANALYSIS
OF THE REAL NETWORKS

Master's thesis

Study programme: Applied Informatics
Field of study: 2511 Applied Informatics
Training center: Department of applied informatics
Supervisor: doc. RNDr. Mária Markošová, PhD.

Bratislava, 2020

Bc. Michal Puškel



Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Michal Puškel
Študijný program: aplikovaná informatika (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: aplikovaná informatika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: anglický
Sekundárny jazyk: slovenský

Názov: Graphlet structure analysis of the real networks
Analýza grafletovej štruktúry reálnych sietí

Anotácia: Študent vytvorí nový softvér alebo vylepší existujúci softvér na analýzu grafletovej štruktúry komplexných sietí. S pomocou tohto softvéru zanalyzuje grafletovú štruktúru sietí, reálnych dát a vypočíta všetky miery s touto štruktúrou súvisiace. Zanalyzuje tiež grafletovú štruktúru umelo vytvorených sietí a porovná s dátami.

Cieľ: Analyzovať grafletovú štruktúru reálnych sietí pomocou vylepšeného alebo vlastného softvéru.

Vedúci: doc. RNDr. Mária Markošová, PhD.
Konzultant: Mgr. Peter Náther, PhD.
Konzultant: Mgr. Andrej Jursa
Katedra: FMFI.KAI - Katedra aplikovanej informatiky
Vedúci katedry: prof. Ing. Igor Farkaš, Dr.

Spôsob prístupnenia elektronickej verzie práce:
prípustná pre vlastnú VŠ

Dátum zadania: 25.09.2017

Dátum schválenia: 09.10.2017

prof. RNDr. Roman Ďurikovič, PhD.
garant študijného programu

.....
študent

.....
vedúci práce

I hereby affirm that this Master's thesis represents my own written work and that I have used no sources and aids other than those indicated. All passages quoted from publications or paraphrased from these sources are properly cited and attributed.

Bratislava, 2020

.....

Bc. Michal Puškel

Acknowledgement

TODO podakovat strasne vela ludom

Abstract

TODO en abstrakt

Keywords: graphs, networks, degree distribution, graphlet degree distribution, orca

Abstrakt

TODO sk abstrakt

Kľúčové slová: grafy, siete, distribúcia stupňa vrcholov, distribúcia orbít
grafletov, orca

Contents

1	Introduction	1
1.1	Graph comparison using graphlet degree distribution	1
2	Goals of the thesis	3
3	Graph theory	4
4	Solution	5
4.1	The brute force, but as good as even possible	5
4.2	Orca	5
4.2.1	Custom orca implementation	6
4.3	Workflow	6
5	Results	7
6	Conclusion	8

Chapter 1

Introduction

This thesis is a direct sequel to our former thesis [Pu7]. Let's briefly refresh the problem we solve in next section 1.1.

1.1 Graph comparison using graphlet degree distribution

In this thesis we strive for finishing our ultimate goal to find out, whether attribute of brain functional network “to suffer by Alzheimer’s disease” is projected to its graph structure.

We obtained input dataset of examined brain functional networks from work of McCarthy and col. [MBF14]. Dataset consists of space standardized brain functional networks from 40 research participants. Participants were divided into 3 groups: young people, elder people with diagnosed tendency for Alzheimer’s disease and elder people without this diagnosis.

Should we be able to distinguish between two subsets of input dataset according to diagnosed disease by usage a computer program comparing graph

structure, we may assume that Alzheimer's disease is really projected to brain functional networks structure.

blab bla tymtoo stylom takmer 1:1 cely uvod prelozit podla starej prace
- motivacia - jasne lahke priklady pre pochopenie
na konci citovat [Pr7]
da sa to robit bruteforcom ako v minulej praci alebo:
spomenut aj [HD14] ze existuje aj kombinatoricky princip

Chapter 2

Goals of the thesis

TODO popisat ciel

- zmerat mozgy
- overit GDD metodu == zmerat umele siete ER GEO AB ???

Chapter 3

Graph theory

TODO teoria grafov - iba to co potrebujem bude vychadzat z predoslej prace

- lepsie vysvetlit orbity / automorfizmy

+ vysvetlit aj BA ER GEO - resp tie umele ake zmeram / co su bezskalove...

nieco z komplexnyxh sieti - iba to co sa tyka mojej prace

Chapter 4

Solution

bla bla kombinatorcke riesenie na kt som si najskor netrufol (orca) vs. chcel som svoje nieco spravit (brute force GDD konzolova aplikacia)

4.1 The brute force, but as good as even possible

TODO "STRUCNE!!!" iba popisat co som urobil zmeral v predoslej praci

- bol to distrib system
- vyhody nevyhody

4.2 Orca

TODO vysvetlit co je Orca

-aj s obrazkami podrobnejsie 4node graflety

4.2.1 Custom orca implementation

TODO trochu popisat co som sparalalelizoval v orce - ci sa zrychlila..

4.3 Workflow

TODO popisat / zhrnut co sa spravilo experimenty

- distrib system meranie (este bak)
- preteceny integer (este bak)
- dalsie meranie cez letne prazdniny
- modul na GDD agreement
- implementacia s arbitrarnou presnostou (zavisi iba od RAM) - vlastna odmocnina
- podozrive vysledky
- odhalena chyba v orbite [55] - porovnanie vsetky vysledky meranych grafov moje vs orca
- modul na spracovanie vysledkov z orcy mojim gdd agreementom - konecny experiment
- overenie GDD agreementu s nejakymi umelymi

Chapter 5

Results

TODO -distribuvany system s kopec funkciami na meranie gdd
bruteforce nie najlepší, ale z hľadiska memory resp ak máme viac slabších
PC a málo času tak lepšie riešenie ako orca
-vlastná implementácia orcy (dufam že bude rýchlejšia)
-vykonaný experiment s mozgami - výsledky - nedajú sa rozlíšiť
-experiment s umelými sieťami - GDD ich rozlišuje (dufam)

Chapter 6

Conclusion

alzheimer sa pravdepodobne neprejavuje v strukrue grafu sieti mozgu resp

GDD agreement ich nerozlisi

future work 6 vrcholove graflety ?

Bibliography

[HD14] Tomaž Hočevar and Janez Demšar. A combinatorial approach to graphlet counting. *Bioinformatics*, 30(4):559–565, 12 2014.

[MBF14] Paul McCarthy, Lubica Benusková, and Elizabeth A. Franz. The age-related posterior-anterior shift as revealed by voxelwise analysis of functional brain networks. In *Front. Aging Neurosci.*, 2014.

[Pr7] Nataša Pržulj. Biological network comparison using graphlet degree distribution. *Bioinformatics*, 23(2):e177–e183, 01 2007.

[Pu7] Michal Puškel. *Pravdepodobnostné porovnávanie funkčných sietí mozgu (Bakalárska práca)*. Univerzita Komenského v Bratislave, 2017.

List of Figures

Appendix

TODO popisat prilohy - sukromny repozitar gitlab