

VEDA O SIEŤACH / ALGORITMY NA SIEŤACH 2023

ZADANIE DOMÁCEJ ÚLOHY 2

Odovzdať vo forme jedného pdf s riešeniami (s názvom Priezvisko.pdf), plus priečinku obsahujúceho zdrojový kód (s názvom Priezvisko_kod).

Úloha 1. V roku 2018 vyšiel článok v časopise Science o tom, ako mravce menia sieť svojich kontaktov za účelom obmedzenia prenosu infekcie. My sa v tejto úlohe pozrieme na jeden z množstva príkladov siete kontaktov mravcov z týchto experimentov. Bude to sieť bez kontaktov mravcov bez prítomnosti infekcie a bližšie sa pozrieme na jej topológiu. Budeme sa zaujímať najmä o to, či sú mravce v kontakte prevažne s mravcami im rovnakého typu (asortativita) alebo sú kontakty medzi mravcami náhodné vzhľadom na ich typ. Tiež budeme sledovať, či sa patogén v sieti šíri ľahšie alebo ťažšie v porovnaní s náhodnou sieťou a to prostredníctvom hustoty a priemerného stupňa vrcholov siete v porovnaní s randomizovanou sieťou kontaktov.

V Moodle nájdete spomínaný článok aj so supplementom (v ktorom sa nachádzajú všetky technické informácie), ako aj súbor s menom 'dataDU.xlsx', ktorý obsahuje experimentálne dáta. Dáta boli poskytnuté so súhlasom autorky štúdie Nathalie Stroeymeyt (University of Lausanne). Tieto dáta sú v nasledujúcom formáte:

- Súbor obsahuje tri listy. V prvej liste (observed) sú pozorované kontakty medzi mravcami, kým v druhej liste sú randomizované dáta s tou istou štruktúrou.
- Každý mravec má jednoznačné číselné označenie a v každom riadku tabuľky je zaznamenaný jeden kontakt medzi dvojicou mravcov s označeniami v stĺpcoch Tag1 a Tag2. Pri kontakte nezáleží na poradí mravcov. V riadku je zaznamenané trvanie kontaktu v jednotkách 0.5 sekundy (čiže kontakt dĺžky 2 predstavuje trvanie 1s).
- V tretej liste je prehľad číselných označení jednotlivých mravcov aj s ich typom, ktorý je jednou z nasledujúcich možností: queen (kráľovná), nurse (mravce v mravenisku), forager (potravu hľadajúce mravce). Kráľovná je len jedna.

Tu sú jednotlivé úlohy, pomocou ktorých by ste mali objasniť, ako sa správa táto kolónia:

- (a) Skonstruujte sieť kontaktov medzi mravcami, ktorá bude mať rozmer $n \times n$, kde $n = 105$ je celkový počet mravcov tak, že nájdete maticu kontaktov A . Prvok $a_{ij} = a_{ji}$ tejto matice predstavuje celkový čas kontaktu medzi mravcom i a j , teda súčet trvaní všetkých kontaktov medzi touto dvojicou mravcov (nezáleží na ich poradí). Urobte to pre pozorované dáta, ako aj pre randomizované dáta. Vizualizujte obe siete (v dvoch separátnych grafoch, pre porovnanie vykreslených vedľa seba), ideálne aj farebne odlíšte jednotlivé typy mravcov.
- (b) Spočítajte a reportujte modularitu oboch sietí. Môžete použiť už implementovanú funkciu, ak takú vo vami zvolenom výpočtovom softwari nájdete. Vysvetlite, ako ste ju spočítali.
- (c) Spočítajte a reportujte hustotu oboch sietí. Môžete použiť už implementovanú funkciu, ak takú vo vami zvolenom výpočtovom softwari nájdete. Vysvetlite, ako ste ju spočítali.
- (d) Spočítajte a reportujte priemerný stupeň vrchola oboch sietí. Môžete použiť už implementovanú funkciu, ak takú vo vami zvolenom výpočtovom softwari nájdete. Vysvetlite, ako ste ho spočítali.
- (e) Porovnajete, či sa infekcia bude lepšie šíriť v pozorovanej sieti alebo v randomizovanej sieti. Použijete na to argumenty založené na vypočítaných štatistikách.
- (f) Stručne popíšte, aký software a ako ste použili. Výsledný kód, spreadsheet, atď. priložte k riešeniu domácej úlohy do separátneho priečinku.

Pri riešení úloh použijete vami preferovaný výpočtový software (Python, R, Matlab, C++, Excel, atď.). Na stránke <https://fmph.uniba.sk/osobne-stranky/bodova/katarina-bodova/teaching/> v ľavom menu môžete nájsť sadu úloh s riešeniami v Matlabe. Matlab je dostupný na školských počítačoch, ale môžete si stiahnuť aj voľne dostupnú verziu softwaru pod názvom Octave, ktorá má mierne obmedzenú funkcionality (sada funkcií na prácu s grafmi sa dá stiahnuť z GitHubu).

Úloha 2. Majme Poissonovský náhodný graf $G(n, p)$ s dostatočne veľkým n .

- (a) Aký je priemerný stupeň vrcholu ak viete, že v sieti je gigantický komponent, ktorý zahŕňa presne polovicu jej vrcholov?
- (b) Aká je v tejto sieti pravdepodobnosť, že vrchol má stupeň 5?
- (c) Ak viete, že má vrchol stupeň 5 zistíte pravdepodobnosť, že je súčasťou gigantického komponentu.
- (d) Aká je proporcia vrcholov stupňa 5 v gigantickom komponente?

Jednotlivé riešenia poriadne vysvetlite.

Úloha 3. Študujme konfiguračný model, v ktorom má každý vrchol rovnaký stupeň k

- (a) Ako vyzerá sieť pre $k = 1$? Má táto sieť gigantický komponent?
- (b) Aké je rozdelenie stupňa vrchola pre túto sieť? Nájdite vytvárajúcu funkciu pre rozdelenie stupňa vrchola $g_0(u)$ a pre rozdelenie stupňa susedného vrchola $g_1(u)$.
- (c) Aký veľký je gigantický komponent siete pre $k = 3$?

Jednotlivé riešenia poriadne vysvetlite.