Názov výsledku: Vplyv asortativity siete na výsledky metód detekcie komunít v grafových dátach reálneho sveta.
Anglický názov výsledku: Effects of network assortativity on results of community detection methods on real world networks.
Číslo a typ projektu: VEGA 2/0185/13, VENIS FP7-284984
Autor: Marek Ciglan
Zhlukovanie v grafoch, často označované aj ako detekcia komunít, je prominentnou úlohou v doméne dolovania grafových dát, čoho dôkazom sú aj desiatky algoritmov navrhnutých pre túto úlohu v posledných rokoch. V tejto práci sme sa zamerali na niekoľko populárnych algoritmov pre detekciu komunít s nízkou časovou náročnosťou a s výbornými výsledkami na umelo generovaných sieťach pre testovanie metód detekcie komunít a študovali sme ich správanie na rozsiahlych sieťach reálneho sveta s explicitne definovanými komunitami.
Pre niektoré zo skúmaných sietí reálneho sveta použité algoritmy doručili pomerne kvalitné rozdelenie vrcholov grafov do komunít, pre niektoré však doručené rozdelenie obsahovalo väčšinu uzlov v zopár obrovských komunitách, pričom takéto rozdelenie malo kvalitu porovnateľnú s náhodným. Motivovaní pozorovaním, že existuje trieda sietí na ktorých algoritmy na detekciu komunít zlyhávajú v odhaľovaní zhlukovej štruktúry, skúmali sme vlastnosti použitých sietí. Štúdia ukázala koreláciu medzi koeficientom asortativity siete a detegovanou komunitnou štruktúrou s majoritou uzlov v niekoľko málo zhlukoch. Koeficient asortativity vyjadruje tendenciu uzlov siete byť spojených hranami s uzlami podobného stupňa (kde stupeň uzla je počet hrán incidentých s daným uzlom). Ďalej sme ukázali, že koeficient asortativity siete môže byť značne odlišný od asortativity štruktúry komunít siete. Následne sme skúmali možnosť zmeny využiť spomínané pozorovanie pre zlepšenie detekcie komunít v sieťach s asortatívnou štruktúrou komunít pomocou váhovania hrán siete. Vyhodnotenie ukázalo, že navrhovaný spôsob váhovania sietí môže výrazne vylepšiť výsledky metód na detekciu komunít v sieťach s asortatívnou štruktúrou komunít.
Práca bola publikovaná na konferencii KDD 2013 (19th ACM SIGKDD Conference on Knowledge Discovery and Data Mining). KDD je prestížnou svetovou konferenciu v oblasti dolovania dát, čo potvrdzuje aj dlhodobo nízky pomer prijatých a podaných príspevkov (v roku 2013 bolo priatych na publikovanie 17% zo 726 podaných príspevkov).

Hlavné scientometrické výstupy:

[1] CIGLAN, M. - LACLAVÍK, M. - NØRVÅG, K.: On Community Detection in Real-World Networks and the Importance of Degree Assortativity, 19th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, KDD 2013], **SCOPUS**