

Poročilo o uresničitvi predloženega programa dela na raziskovalnem projektu

Izhodišče projekta predstavlja povečana pozornost, ki jo raziskovalci namenjajo analizi bolj kompleksnih vrst omrežij, kot so tako imenovana povezana omrežja, ki vključujejo tudi več-nivojska in časovna omrežja (omrežja, merjena v več časovnih točkah) ter pomanjkanje metod za razvrščanje takih omrežij. Razvili smo sicer že metode za posplošeno bločno modeliranje (BM) povezanih omrežij, ki pa so zaradi računske zahtevnosti manj primerne za večja omrežja (z več kot nekaj sto enotami v vsakem omrežju). Zato je bil glavni cilj projekta razvoj metod za BM povezanih (predvsem več-nivojskih) omrežij, primernih za večja (več 100 enot) omrežja. V prijavi projekta so bile aktivnosti za doseganje tega in podpornih ciljev podrobneje razdeljene v 11 delovnih paketov (DP), od katerih nekateri predstavljajo podporne ali pripravljalne aktivnosti. Zaradi preglednosti poročilo sledi spodaj navedenim ciljem projekta. Pri navedbi vsakega cilja navajamo tudi v okviru katerih DP so bili naslovljeni (razen DP11 - Upravljanje projekta).

Glavna cilja sta bila (I. – DP1 + DP3-5) pregled stanja in izboljšanje optimizacijske metode za posplošeno BM za uporabo pri BM povezanih omrežjih in (II. – DP6-8 + DP9) prilagoditev hitrejših pristopov za BM povezanih omrežij. Dodatna cilja projekta sta bila tudi (III. – DP10) uporaba razvitih metod za BM empiričnih omrežij sodelovanj v znanosti in (IV. – DP2) spoznavanje lokalnih mehanizmov, ki vodijo v določeno globalno (bločno strukturo) oz. spremembo globalne strukture in ki omogočajo generiranje omrežij z določeno bločno strukturo.

Zaradi polovičnega financiranja projekta smo se, kot je bilo navedeno že pri prijavi, pri vseh ciljih razen pri zadnjem večinoma omejili le na eno vrsto povezanih omrežij, torej na večnivojska omrežja, smo pa nekatere metode preizkusili tudi na časovnih in časovnih večnivojskih omrežjih.

Cilji:

I. Izboljšanje optimizacijske metode za posplošeno bločno modeliranje za uporabo pri bločnem modeliranju povezanih omrežij (DP1 + DP3-5)

V projektu smo precej pomembnosti namenili pregledu literature oz. trenutnega stanja na področju razvrščanja v omrežjih. Pomemben del tega je razviden iz knjige *Advances in network clustering and blockmodeling* [1], ki so jo uredili in v veliki meri (s soavtorji) napisali sodelavci pri projektu. Izpostavljamo štiri poglavja. V 2. [2] je predstavljen pregled literature, v 3. [3] teoretični pregled osnov razvrščanja in BM v omrežjih ter v 6. [4] pregled BM omrežij z vrednostmi na povezavah. Za ta cilj je še posebej pomembno 10. poglavje [5], kjer je predstavljeno (posplošeno) BM povezanih omrežij.

V sklopu tega cilja smo želeli predvsem izboljšati optimizacijo pri posplošenem BM. Glavna ugotovitev je, da samo z manjšimi posegi v optimizacijo ne bo mogoče zagotoviti ustreznega napredka, ter da so predlogi za izboljšanje optimizacije, navedeni v prijavi projekta, primerni le za določeno vrsto povezanih omrežij. Zato smo več pozornosti namenili II. cilju in raziskovali možnosti, kako bi za pohitritev izkoristili dejstvo, da so velika omrežja običajno redka.

Za slednje je ključno, da omrežja niso predstavljena v obliki matrike, pač pa v obliki, kjer so (običajno v obliki seznamov) zabeležene le prisotne povezave. V ta namen smo razvili podporno knjižnico Nets (<https://github.com/bavla/Nets>). Ker posplošeno BM temelji na lokalni optimizaciji, lahko postopek

pohitrino tako, da zelo omejimo možne lokalne transformacije, na primer tako, da jih omejimo samo na predstavitev enote iz njene skupine v eno izmed skupin njenih sosedov. Program s to možnostjo in uporabo seznamske predstavitve je na voljo na GitHub-u (<https://github.com/bavla/Linked>).

V R paketu *blockmodeling* [6] je implementirana tudi pohitritev, ki pri lokalni optimizaciji poskrbi, da se pri zaporednih korakih na novo izračunajo le deli kriterije funkcije, ki se spremenijo. Pri lokalni optimizaciji imamo več možnosti, kdaj se premaknemo v novo rešitev. Ena je, da se premaknemo v prvo boljšo rešitev, ki jo najdemo, druga pa, da preizkusimo vse možnosti in se premaknemo v najboljšo. Obe možnosti sta implementirani v [6], v [7] pa smo pokazali, da je pri povezanih (in na splošno večjih) omrežjih bistveno bolj primerna prva.

Rezultati:

- [1] DOREIAN, P., BATAGELJ, V., FERLIGOJ, A. (ur.). *Advances in network clustering and blockmodeling*. Hoboken (NJ); Chichester: J. Wiley & Sons, 2020. [COBISS.SI-ID [1538492867](#)]
- [2] BATAGELJ, V., FERLIGOJ, A., DOREIAN, P.. Bibliometric analyses of the network clustering literature. V: [1]. [COBISS.SI-ID [36547421](#)]
- [3] BATAGELJ, V.. Clustering approaches to networks. V: [1]. [COBISS.SI-ID [18884953](#)]
- [4] NORDLUND, C., ŽIBERNA, A.. Blockmodeling of valued networks. V: [1]. [COBISS.SI-ID [36547933](#)]
- [5] ŽIBERNA, A.. Blockmodeling linked networks. V: [1]. [COBISS.SI-ID [36548957](#)]
- [6] MATJAŠIČ, M., CUGMAS, M., ŽIBERNA, A.. Blockmodeling : an R package for generalized blockmodeling. *Metodološki zvezki*. 2020, vol. 17, no. 2, str. 49-66. [COBISS.SI-ID [53164547](#)].
- [7] ŽIBERNA, A.. K-means-based algorithm for blockmodeling linked networks. *Social Networks*. [Printed.]. May 2020, vol. 61, str. 153-169. [COBISS.SI-ID [36567901](#)]

II. Prilagoditev hitrejših pristopov za bločno modeliranje povezanih omrežij (DP6-8 + DP9)

Znotraj tega cilja smo načrtovali prilagoditev hitrejših metod za BM, in sicer stohastičnega BM in BM na podlagi metode voditeljev (k-means), za povezana omrežja. Te metode zaradi hitrejšega delovanja omogočajo BM bistveno večjih omrežij.

Kljub temu, da smo načrtovali le razvoj ene izmed teh dveh metod (zaradi polovičnega financiranja), smo se lotili dela na obeh metodah (deloma namesto dela na prvemu cilju). Osnovne ideje tako za k-means kot tudi stohastično BM smo navedli že v predstavitvi na EUSN 2017 [8].

Začeli smo z razvojem metod za k-means BM povezanih omrežij, zato imamo pri tej metodi tudi več oz. bolj zaključene (že objavljene) rezultate. Najprej smo k-means metodo za BM z dvovrstnih razširili še na enovrstna omrežja, nato pa še na povezana omrežja. Obe različici sta predstavljeni v članku, objavljenem v ugledni reviji *Social networks* [9]. Z metodo za povezana omrežja smo analizirali tudi večnivojsko omrežje v dveh časovnih točkah. Razvite metode so tudi implementirane v R paketu *kmBlock*, kjer je za hitrejšo delovanje glavnina implementirana v programskem jeziku C++ [10]. Za povezana omrežja smo prilagodili tudi metodo za stohastično BM in jo predstavili v članku z naslovom "Stochastic blockmodeling of linked networks", ki je trenutno v recenziji v reviji *Social Networks*. Metode so implementirane v R paketu *StochBlockTest*, (jedro v programskem jeziku C++) (https://r-forge.r-project.org/R/?group_id=203).

Obe razviti metodi (na podlagi k-means in stohastičnega BM) omogočata tudi vnaprejšnjo specifikacijo dela rešitve v smislu omejitev na povprečjih po blokih v primeru k-means metode ali na verjetnostih v primeru stohastičnega BM.

Izvedli smo primerjavo metod za povezana omrežja na podlagi k-means metode in posplošenega BM. Rezultati, objavljeni v [9], so pokazali, da je za večja omrežja (ne glede na strategijo optimizacije pri posplošenem BM) primernejši k-means pristop. Metode smo primerjali tudi na enovrstnih omrežjih in rezultate predstavili na več konferencah, na primer na konferenci EUSN 2019, ter v članku, ki je bil v poslan v *Journal of Classification*.

Rezultati:

[8] ŽIBERNA, A.. Multilevel blockmodeling for "larger" networks. V: *Book of abstracts*. Mainz: Johannes Gutenberg-University, 2017. Str. 191. [COBISS.SI-ID [35683165](#)]

[9] ŽIBERNA, A.. K-means-based algorithm for blockmodeling linked networks. *Social Networks*. [2020, vol. 61, str. 153-169. [COBISS.SI-ID [36567901](#)],

[10] ŽIBERNA, A.. *kmBlock : K-means like blockmodeling of one-mode and linked networks*. [S. l.: s. n.], 2020. https://r-forge.r-project.org/R/?group_id=203. [COBISS.SI-ID [61863683](#)]

III. Uporaba razvitih metod za bločno modeliranje empiričnih omrežij (DP10)

Namen tega cilja je poglobiti vpogled v zgradbo znanstvenega sodelovanja slovenskih raziskovalcev z upoštevanjem soodvisnosti več ravni znanstvenega sodelovanja (individualne in organizacijske); z uporabo razvitih metod za BM povezanih omrežij.

S k-means metodo BM za povezana omrežja smo analizirali podatke za vedo Družboslovje za obdobje med 2006 in 2015. Rezultati za raven raziskovalcev so v skladu s predhodnimi ugotovitvami. Na ravni organizacij pa je zaznati visoko stopnjo sodelovanja pri projektih – v splošnem so organizacije močno povezane tako znotraj skupin, kakor tudi med skupinami organizacij. Na podlagi dela omrežja, ki povezuje raven raziskovalcev in raven organizacij, je mogoče sklepati, da se formalna znanstvena sodelovanja pri projektih med organizacijami vedno ne odražajo v soobjavah znanstvenih publikacij med raziskovalci. Rezultati raziskave so bili objavljeni v reviji *Scientometrics* [11], v bolj poljudni obliki pa tudi v obliki objave na blogu [12].

Predhodno smo poskusno uporabili metodo za posplošeno BM na časovnem omrežju znanstvenega sodelovanja v disciplini Procesno strojništvo (disciplina je bila izbrana priložnostno), kjer smo pokazali, da se lahko metode za BM povezanih omrežij uporabi za analizo strukture znanstvenega sodelovanja v času. Analiza je predstavljena v poglavju o bločnem modeliranju povezanih omrežij [5] v knjigi [1].

Rezultati:

[11] CUGMAS, M., MALI, F., ŽIBERNA, A.. Scientific collaboration of researchers and organizations : a two-level blockmodeling approach. *Scientometrics*. Dec. 2020, vol. 125, no. 3, str. 2471-2489. [COBISS.SI-ID [28497411](#)]

[12] CUGMAS, M., MALI, F., ŽIBERNA, A.. The scientific collaboration of researchers and organizations in Slovenia. YoungStats: [blog]. 1st Oct. 2020, <https://youngstats.github.io/post/2020/10/01/cugmas/>. [COBISS.SI-ID 34176003]

[5] ŽIBERNA, A.. Blockmodeling linked networks. V: [1]. [COBISS.SI-ID [36548957](#)]

IV. Spoznavanje lokalnih mehanizmov, ki vodijo v določeno globalno (bločno strukturo) oziroma spremembo globalne strukture in ki omogočajo generiranje omrežij z določeno bločno strukturo (DP2)

V [13] smo (z uporabo MCMC algoritma in predlaganega algoritma prestavljanja povezav) potrdili, da je večino osnovnih bločnih modelov mogoče generirati zgolj z upoštevanjem različnih vrst triad.

Nadalje smo [14,15] preučevali lokalne omrežne mehanizme pri vzpostavljanju prijateljstev in/ali interakcij med predšolskimi otroci. V okviru tega dela smo predlagali tudi središčno-kohezivni bločni model, ki se lahko pojavi na podlagi takih mehanizmov v omrežjih interakcij. Z uporabo simulacij smo potrdili, da predlagana vrsta bločnega modela lahko nastane kot rezultat delovanja takih mehanizmov.

Z analizo rastočih časovnih omrežij smo v empiričnih omrežjih prenosa znanja v slovenskem podjetju potrdili prisotnost hierarhične strukture omrežja, ki nastane iz dveh relativno nepovezanih središčno-perifernih zgradb [16]. Tovrstna zgradba omrežja bi lahko nastala pod vplivom mehanizmov zaznane vrednosti in mehanizmov zaznanih stroškov.

Rezultati:

[13] CUGMAS, M., FERLIGOJ, A., ŽIBERNA, A.. Generating global network structures by triad types. *PloS one*. May 2018, vol. 13, iss. 5. [COBISS.SI-ID [35624541](#)]

[14] CUGMAS, M., ŽIBERNA, A., FERLIGOJ, A.. Mechanisms generating asymmetric core-cohesive blockmodels. *Metodološki zvezki*. 2019, vol. 16, no. 1, str. 17-41.

[15] CUGMAS, M., DELAY, D., ŽIBERNA, A., FERLIGOJ, A.. Symmetric core-cohesive blockmodel in preschool children's interaction networks. *PloS one*. 2020, vol. 15, no. 1, 19 str. [COBISS.SI-ID [36567645](#)]

[16] CUGMAS, M., FERLIGOJ, A., ŠKERLAVAJ, M., ŽIBERNA, A.. Global structures and local network mechanisms of knowledge-flow networks. *PloS one*. 2021, vol. 16, iss. 2, 23 str. [COBISS.SI-ID [52090371](#)]