

Bevezetés

Munkánk során a komplex hálózatok topológiája és a hiperbolikus tér geometriája közötti kapcsolat [1] különböző vonatkozásait vizsgáljuk: egyrészt hiperbolikus beágyazást (vagyis a hálózat pontjainak a hiperbolikus térben történő, a hálózati topológiát jól tükröző elrendezését) megvalósító algoritmusokat, másrészt pedig hiperbolikus geometriára építő hálózatgeneráló modelleket.

Az aktuális félévben elvégzett kutatások ismertetése

Az első év végére megszületett egy kézirat a hálózatokat a hiperbolikus síkban dimenzió redukció segítségével elhelyező ncMCE beágyazó módszerből [2] kinyert szögkoordináták optimalizálására kidolgozott, a hiperbolikus síkon hálózatokat generáló E-PSO modellre [3] vonatkozó likelihood [3] maximalizációját végző algoritmusunkról, mely nem csak a likelihood, hanem a greedy routing score [2] szempontjából is képes javítani az ncMCE módszerből kapott beágyazások minőségén. Jelen félév során a kéziratot a kapott bírálatok figyelembevételével átdolgoztuk, a cikk a félév végére elfogadásra került a Scientific Reports folyóirat által.

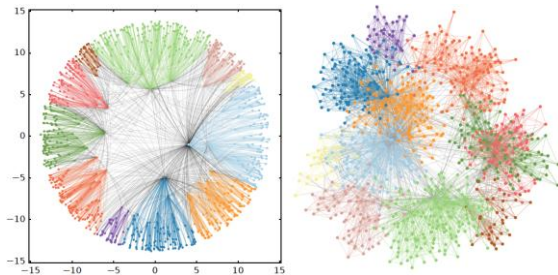
Az első évben elkezdtem foglalkozni a különböző dimenziójú hiperbolikus terekben történő hálózatgenerálással. Jelen félév során rájöttem, hogy hogyan lehet a hiperbolikus távolságot magas dimenziós terek esetén is a hiperbolikus tér natív reprezentációjában [1], a Poincaré ball modellre való áttérés nélkül kiszámítani. A natív reprezentációban az \underline{u} és a \underline{v} helyvektorral jellemzett pontok által bezárt $\Delta\Psi$ szög az (1) képlet alapján, a helyvektorok $\underline{u} \cdot \underline{v}$ euklideszi skalárszorzatából és r_u , r_v euklideszi normáiból kiszámítható, így a két pont közti $x(\underline{u}, \underline{v})$ hiperbolikus távolságot tetszőleges dimenziójú, $K = -\zeta^2$ görbületű hiperbolikus térben a hiperbolikus koszinusztételből származó (2) kifejezés adja meg. A félév végén megkezdjük az E-PSO hiperbolikus síkbeli hálózatgeneráló modell [3] tetszőleges dimenziós térre történő kiterjesztéséről szóló kézirat elkészítését, mellyel kapcsolatos főbb eredmények az előző féléves beszámolóban már említésre kerültek.

$$\cos(\Delta\Psi) = \frac{\underline{u} \cdot \underline{v}}{r_u r_v} \quad (1)$$

$$x(\underline{u}, \underline{v}) = \frac{1}{\zeta} \cdot \operatorname{arccosh} \left(\cosh(\zeta r_u) \cdot \cosh(\zeta r_v) - \sinh(\zeta r_u) \cdot \sinh(\zeta r_v) \cdot \frac{\underline{u} \cdot \underline{v}}{r_u r_v} \right) \quad (2)$$

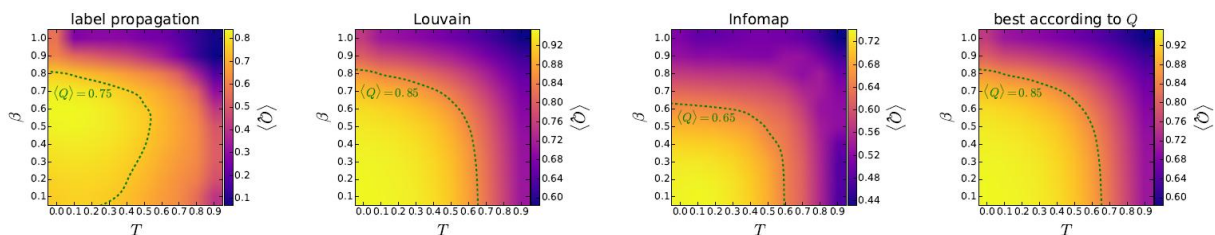
A félév során részletesen tanulmányoztam a kis-világ tulajdonsággal, skálafüggetlen fokszámeloszlással és magas átlagos klaszterezettséggel rendelkező hálózatokat generálni képes E-PSO [3,4] és $\mathbb{S}^1/\mathbb{H}^2$ modellek [5,6] egy kevésbé ismert tulajdonságát, miszerint az ezekkel generált hálózatok a pontok szög szerinti eloszlásának egyenletessége és bármiféle direkt csoportgeneráló lépés hiánya ellenére a modellparaméterek széles tartományában moduláris struktúrájúak. Célunk az volt, hogy megmutassuk, hogy a különböző csoportkereső algoritmusok (Louvain [7], Infomap [8], asynchronous label propagation [9]) valóban releváns strukturális egységeket képesek detektálni a generált hiperbolikus hálózatokban, valamint hogy feltérképezzük a modellparaméterek hatását a generált hálózatok csoportstruktúrájára.

Azt találtuk, hogy amennyiben a hálózat pontjai közötti kötések viszonylag jól lokalizáltak (tehát a kötési valószínűség a pontok közötti hiperbolikus távolságnak erősen csökkenő függvénye, ami miatt a hálózat átlagos klaszterezettségi együtthatója nagy) és a hiperbolikus síkot reprezentáló korongon legbelül elhelyezkedő, azaz a legnagyobb várható fokszámmal rendelkező pontok radiális koordinátája nem túl kicsi (ekkor a várt fokszámeloszlás gyorsan lecsengő hatványfüggvény), akkor minden külsőbb pont viszonylag egyértelműen csak bizonyosakhoz képes hozzákötni ezen egymástól viszonylag távoli, jól elkülönülő legbelső pontok közül, aminek eredményeképp alapvetően a hiperbolikus korong különböző szögtartományain belül alakulnak ki kötések, így az egyes szögtartományokban elhelyezkedő pontok más szögtartományok pontjaival kevésbé kötött, különálló modulokat képeznek annak ellenére is, hogy a hálózat pontjainak szög szerinti eloszlása teljesen egyenletes. Az 1. ábrán látható példa szerint a hálózat pontjainak a szögtartományok szerinti összetartozása csoportkereső algoritmusok által detektálható, és a talált csoportstruktúra még a hálózat hiperbolikus eredetét teljesen figyelmen kívül hagyó, euklideszi síkbeli standard ábrázolás mellett is jól kirajzolódik. Mivel a PSO és az $\mathbb{S}^1/\mathbb{H}^2$ modell viselkedése a csoportstruktúra szempontjából nagyon hasonló, ezért a beszámoló terjedelmi korlátaira való tekintettel az alábbiakban csak a PSO modellre vonatkozó legalapvetőbb eredményeket mutatom be.

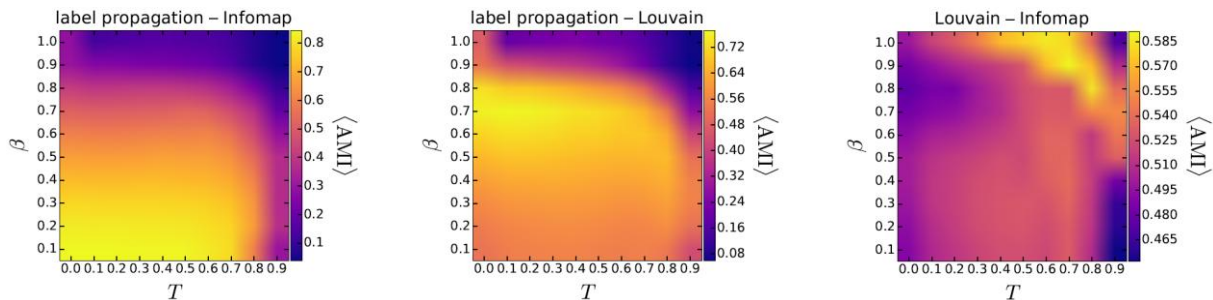


1. ábra: Egy, a PSO modellel generált hálózat a hiperbolikus sík natív reprezentációjában (bal) és az euklideszi síkon (jobb) megjelenítve. A hálózat pontjainak színezése a Louvain algoritmus által detektált csoportosításnak megfelelő, a talált csoportokon belül húzódó élek színe az adott csoport pontjainak színével megegyező, a különböző csoportok tagjai között húzódó élek pedig szürkék.

A 2. és a 3. ábrán látható, hogy a hiperbolikus hálózatokban talált csoportosításokat egy viszonylag kiterjedt paramétertartományon magas modularitás [10] jellemzi, valamint hogy teljesen különböző elveken alapuló csoportkereső algoritmusok egymáshoz igen hasonló csoportokba sorolják a hálózati pontokat (az egyes csoportosítások közötti adjusted mutual information [11] viszonylag nagy). Mindezek arra engednek következtetni, hogy a detektált csoportok valóban egyfajta strukturális egységeit képezik a hiperbolikus hálózatoknak.



2. ábra: A PSO modellel generált $N = 10000$ pontból álló, $\langle k \rangle = 10$ várható átlagfokszámú hálózatokban különböző algoritmusokkal talált csoportstruktúrák Q modularitásának a T hőmérséklettől (a hőmérséklet csökkentésével az átlagos klaszterezettségi együttható nő) és a $P(k) \sim k^{-\gamma}$ alakú fokszámeloszlás hatványkitevőjét a $\gamma = 1 + 1/\beta$ képlet szerint meghatározó β popularity fading paramétertől való függése. Minden paraméterbeállításnál 100 hálózatra vett átlag látható. Ugyan bizonyos paraméterbeállítások mellett magas modularitás érték érhető el pl. a valójában nem moduláris hálózatokat generáló Barabási–Albert-moddellel is, de az itt vizsgáltakkal azonos méretű és átlagfokszámú Barabási–Albert-féle random hálózatokban detektálható csoportosítások modularitása csupán 0.28 körül van, aminél a legtöbb $\beta - T$ paraméterbeállítás mellett lényegesen nagyobb értékeket kaptunk.



3. ábra: A PSO modellel generált $N = 10000$ pontból álló, $\langle k \rangle = 10$ várható átlagfokszámú hálózatokban a Louvain, az Infomap és az asynchronous label propagation algoritmusok által detektált csoportstruktúrák közti hasonlóságot kifejező adjusted mutual information-nek az átlagos klaszterezettségi együttthatót szabályozó T hőmérséklettől (a hőmérséklet csökkentésével a klaszterezettség nő) és a $P(k) \sim k^{-\gamma}$ hatványfüggvény alakú fokszámeloszlás hatványkitevőjét a $\gamma = 1 + 1/\beta$ képlet szerint meghatározó β popularity fading paramétertől való függése. Minden paraméterbeállítás mellett a 100 hálózatra vett átlag látható. Két véletlenszerű csoportosításra AMI = 0, teljes egyezés esetén pedig AMI = 1.

Publikációk

A félév során az ncMCE módszerrel előállított hiperbolikus beágyazásoknak az E-PSO modellre vonatkozó likelihood maximalizációján alapuló optimalizálásáról szóló kézirat *Optimisation of the coalescent hyperbolic embedding of complex networks* címmel a *Scientific Reports* folyóirat által **elfogadásra került**.

Emellett a hiperbolikus geometrián alapuló, egyenletes szög szerinti eloszlást alkalmazó E-PSO és $\mathbb{S}^1/\mathbb{H}^2$ modellekkel generált hálózatok csoportszerkezetét bemutató munkánkról elkészítettünk egy kéziratot, ami *The inherent community structure of hyperbolic networks* címmel már elérhető az arXiv weboldalon.

A félév végén nekiláttunk az E-PSO hiperbolikus síkbeli hálózatgeneráló modell tetszőleges dimenziós térre történő kiterjesztéséről szóló kézirat elkészítésének is.

Tanulmányi tevékenység az aktuális félévben

A félév során a PhD programomban meghirdetett alábbi két tárgyat végeztem el:

- Klaszterezés hálózatokkal (FIZ/3/064E)
- Preklinikai modellek a daganatkutatásban (FIZ/3/082)

Emellett a Biológiai Doktori Iskola által meghirdetett két tárgyat is elvégeztem:

- Válogatott fejezetek a modern ökológiából EA (BIO/09/02E)
- Ökológiai és evolúciós téridő modellek EA (BIO/10/14E)

Konferenciák az aktuális félévben

Szeptemberben részt vettem az online megrendezett *NetSci 2020* konferencián, ahol poszter prezentáció formájában bemutattam az ncMCE módszerből kapott hiperbolikus beágyazásoknak az E-PSO modellre vonatkozó likelihood maximalizációjára irányuló szövekkorrekciójával elért eredményeimet.

Decemberben részt vettem egy online konferencián (*Complex Networks 2020 – The 9th International Conference on Complex Networks and their Applications*), ahol szóbeli

prezentáció formájában bemutattam a hiperbolikus síkon generált hálózatoknak a pontok egyenletes szög szerinti eloszlása mellett kialakuló csoportstruktúráját elemző munkámat.

Oktatási tevékenység az aktuális félévben

A Biológiai Fizika Tanszék által meghirdetett *Modern fizika laboratóriumi gyakorlatok* című tárgy *Spektrofotometria* mérését vezettem 16 alkalommal, illetve részt vettem a tanszék *Valószínűségszámítás és statisztika a fizikában* című tárgyhöz tartozó házi feladatok javításában is.

Hivatkozások

- [1] Dmitri Krioukov, Fragkiskos Papadopoulos, Maksim Kitsak, Amin Vahdat, Marián Boguñá, *Hyperbolic geometry of complex networks*, Physical Review E, 2010, 82, 036106
- [2] Alessandro Muscoloni, Josephine Maria Thomas, Sara Ciucci, Ginestra Bianconi, Carlo Vittorio Cannistraci, *Machine learning meets complex networks via coalescent embedding in the hyperbolic space*, Nature Communications, 2017, 8, 1615
- [3] Fragkiskos Papadopoulos, Constantinos Psomas, Dmitri Krioukov, *Network Mapping by Replaying Hyperbolic Growth*, IEEE/ACM Transactions on Networking, 2015, 23, 198-211
- [4] Fragkiskos Papadopoulos, Maksim Kitsak, M. Ángeles Serrano, Marián Boguñá, Dmitri Krioukov, *Popularity versus similarity in growing networks*, Nature, 2012, 489, 537-540
- [5] M. Ángeles Serrano, Dmitri Krioukov, Marián Boguñá, *Self-Similarity of Complex Networks and Hidden Metric Spaces*, Phys. Rev. Lett., 2008, 100, 078701
- [6] G García-Pérez, A Allard, M Á Serrano, M Boguñá, *Mercator: uncovering faithful hyperbolic embeddings of complex networks*, New J. Phys., 2019, 21(12):123033
- [7] VD Blondel, JL Guillaume, R Lambiotte, E Lefebvre, *Fast unfolding of communities in large networks*, Journal of Statistical Mechanics: Theory and Experiment, 2008, 2008(10):P10008
- [8] M Rosvall, CT Bergstrom, *Multilevel Compression of Random Walks on Networks Reveals Hierarchical Organization in Large Integrated Systems*, PLOS ONE, 2011, 6(4):1-10
- [9] UN Raghavan, R Albert, S Kumara, *Near linear time algorithm to detect community structures in large-scale networks*, Physical Review E, 2007, 76:036106
- [10] MEJ Newman, *Analysis of weighted networks*, Physical Review E, 2004, 70:056131
- [11] NX Vinh, J Epps, J Bailey, *Information Theoretic Measures for Clusterings Comparison: Variants, Properties, Normalization and Correction for Chance*, Journal of Machine Learning Research, 2010, 11(95):2837-2854