

3. félévi beszámoló

Asztalos Bogdán (*abogdan@caesar.elte.hu*)

Statisztikus Fizika, Biológiai Fizika és Kvantumrendszerek Fizikája
PhD program

Témavezető: Pollner Péter

A dolgozat címe: Erősen kölcsönható, történetfüggő, nem-egyensúlyi komplex rendszerek vizsgálata

Bevezetés

Az előző félévekben bemutatott témák mellett ebben a félévben részletesebben foglalkoztam a hierarchikus generatív folyamatok korrelációinak leírásával. Ezen téma olyan komplex rendszerekre fókuszál, melyeknél a releváns változók értékei egy hierarchikus gráffal reprezentálható folyamat eredményei. Ide sorolható többek között a legtöbb időfejlődéssel kapcsolatos fizikai vagy biológiai folyamat, vagy az olyan rendszerek, ahol az elemek egyértelmű alá-fölé rendelt kapcsolatban vannak valamilyen struktúrális viszony miatt. A lehetséges alkalmazások halmaza tehát igen széles, evolúciós folyamatoktól kezdve génszekvenciák fejlődésén át akár a spinüvegekig vagy folyadékturbulenciákig. A kutatás során a fő célunk egy konkrét, viszonylag egyszerű hierarchikus minimálmodell viselkedésének feltérképezése a rendszert jellemző alapvető paraméterek különböző esetei mellett, és ezen keretrendszeren belül igyekszünk feltérképezni korábbi kutatási eredmények alapján egy rendszer korrelációja és entrópiája közötti kapcsolatot [1, 2].

Az aktuális félévben elvégzett kutatás

A félév során főként két témával foglalkoztam. Egyrészt folytattam az előző félévben is végzett munkát a CounteR nemzetközi konzorciumnak a tanszéken működő csoportjában, másrészt elmélyedtem a hierarchikus generatív folyamatok korrelációit leíró elméleti kutatásban.

Hálózatkészítés internetes közösségi adatokból

A CounteR együttműködés keretében közösségimédia-felületről letöltött adatokból készített hálózatokon végzett elemzések segítségével igyekszünk kiszűrni radikalizálódásra utaló tartalmakat. Ebben a félévben én főként azon dolgoztam, hogy más csoportok elemzőmoduljai felhasználásával olyan eljárást találjak, amivel egy letöltött adathalmazból értelmes hálózatot kapunk, és amin elemzés végezhető. A probléma nem magától értetődő, hiszen egy adathalmaz esetén többféle objektumra tekinthetünk úgy, mint egy hálózat csúcsai, és a közöttük lévő kapcsolatokat is sokféle módon definiálhatjuk, így rengeteg hálózatkészítési eljárás képzelhető el. A fő feladat tehát az, hogy megtaláljuk azt a fajta viszonylag kevés, de széles körben alkalmazható módszert, ami nem csak értelmes, de hasznos eredményre vezet.

Az adatok jellegéből és az együttműködés többi moduljának funkciójából fakadóan leginkább az olyan típusú adatfeldolgozásokat alkalmaztunk, ahol megkülönböztethetünk különböző típusú csúcsokat (pl. felhasználók, posztok, helyszínek, témák, nevek, hashtagek stb. . .), és közöttük valamilyen jól definiálható kapcsolatot tekintve egy bipartit hálózatot készíthetünk. A

in egyesével végiglépkedve felismeri és kigyűjti a benne előforduló név-entitásokat. Bár egyik feladat pontossága sem egyértelműen definiálható, és magam is találtam problémás elemeket a program tesztelése során, de a jelenleg aktuális kutatási eredményeknek megfelelő hatékonysággal dolgozik [3]. Az utóbbi funkciót használtam az 1. ábrán látható ábra elkészítéséhez is.

Hierarchikus korrelációk vizsgálata

Már a mesterképzés alatt is foglalkoztam a hierarchikus kapcsolati struktúrával rendelkező fejlődési folyamatok során létrejövő korrelációk osztályozásával. A szakdolgozatomban azt tárgyaltam, hogy egy rendszer Shannon-entrópiájának sérülésével jellemezhető egy rendszerben lévő elemek közötti korreláció. Ez nagyon hasonló kvalitatív viselkedést mutat a szomszédok közötti korrelációs függvényvel kapható osztályozással, ugyanakkor az osztályozás konkrét kvantitatív jellemzői jelentősen eltértek. Ebben a félévben ezt a problémát kezdtem újra vizsgálni, és a korábban csupán szemléleti alapon bevezetett közelítő megoldásokkal kiszámolt mennyiségeket igyekeztem egzakt számolásokkal megalapozottá és konzekvensen használhatóvá tenni.

A vizsgált minimálmodellt a félév során sikerült kétféle kiterjesztéssel is kezelhetővé tenni, két külön irányból megközelítve a problémát. Az egyik irány az volt, hogy egy probabilsztikus modellel megadtam a rendszerben lévő változók relatív gyakoriságának várható értékét, amiből könnyen áttérhetünk a korrelációk számolására. Ezzel bár analitikus eredményeket kaptam, a korábbi kétféle osztályozási módszerhez képest egy harmadik módszernek bizonyult, és eddig nem bizonyítható, hogy bármelyiknek általánosítása vagy rokon-modellje lenne. A másik irány esetén a rendszer véges számú változóit végtelennek tekintettem, és a relatív előfordulásukat folytonos eloszlásfüggvénnyel írtam fel. Ennek egy nagy előnye volt, hogy sikerült a modell viselkedését visszavezetni a klasszikus diffúziós egyenletre, így szokásos differenciálegyenlet megoldási módszerekkel kezelhető lett. Ugyanakkor a folytonos modell viselkedése nagyban függött a kezdőfeltételektől, amely arra utal, hogy az időfejlődés azon szakasza, amíg a változók nem tekinthetők folytonosnak kritikus későbbiek szempontjából.

Így tehát azt kaptuk, hogy bár a modellhez léteznek olyan kezelési módszerek, amellyel konzekvensen kezelhetők, a kézzelfogható eredményekhez meg kell oldani az átmenetet a korábban is ismert probléma és a matematikailag kényelmes kezelési mód között. A továbbiakban ennek megoldására fogom koncentrálni a munkát.

Publikációk

A félév során két kéziratot adtunk be tudományos folyóirathoz megjelentetésre:

- Dr. Palla Gergellyel és Czégel Dániellel közösen adtuk be a „Subdiffusive Semantic Evolution in Indo-European Languages” című cikket a *Scientific Reports* című újsághoz. A kézirat emellett elérhető az arXiv szerveren is [4].
- Emellett Dr. Bányász Péter, a Nemzeti Közszolgálati Egyetem adjunktusának társszerzőségével küldtük be a „Monitoring the semantic change of Covid-19-related expressions using dynamic word embeddings” című cikket a *The Journal of Popular Culture* című laphoz.

Jelenleg mindkét cikk szakmai felülvizsgálat alatt áll.

Tanulmányi tevékenység

A félév során az alábbi tárgyat végeztem el a doktori programból:

- Adatmodellek és adatbázisok a tudományban (FIZ/3/086)

Emellett pedig két olyan tárgyat végeztem, amelyek nem részei a doktori programon képzési tervének:

- Markov láncok (MAT/200E): Ez a tárgy az ELTE Matematika Doktori Iskolájának keretében került meghirdetésre.
- Renormálás (rpprenormf20em): Ezt a tárgyat az ELTE Elméleti Fizika Tanszéke hirdette meg, és a Fizikus MSc. 2020-tól érvényes tanterveinek a része.

Mindkét tárgy teljesítéséről szóló igazolást továbbítottam Dr. Horváth Gábor programvezető részére, kérve, hogy a tárgyakat elismerjék a doktori képzésem keretében.

Oktatási tevékenység

A félév során kettő, a Fizika BSc szakos hallgatóknak meghirdetett tárgy lebonyolításában is részt vettem:

- A Valószínűségszámítás és statisztika a tudományban (valszamf19va) című tárgyhoz gyakorlatvezetéssel (4 alkalommal tartottam 2 órás gyakorlatot) járultam hozzá,
- a Modern fizika laboratórium (fizlab3f19la) című tárgyon pedig gyakorlatvezetést (6 alkalommal vezettem a 3+1 órás gyakorlatot) és a *kvantumradír*, valamint a *granuláris anyagok* mérésekhez készült jegyzőkönyvek (összesen 34 darab) értékelését végeztem.

Szakmai közéleti tevékenység

A doktori tanulmányaim mellett igyekszem aktívan részt venni a középiskolások természettudományos tehetséggondozásában is. Ennek keretében ebben a félévben a Középiskolai Matematikai és Fizikai Lapok (Kömal) által kiírt levelező pontverseny feladatjavítói között voltam, illetve részt vettem a Dürer verseny fizika kategóriáinak (F és F+ kategória) országos forduló feladatsorának összeállításában.

Hivatkozások

- [1] D. Czégel, S. G. Balogh, P. Pollner, G. Palla. *Phase space volume scaling of generalized entropies and anomalous diffusion scaling governed by corresponding non-linear Fokker-Planck equations*. Scientific reports 8.1 (2018): 1-8.
- [2] S. G. Balog, G. Palla, P. Pollner, D. Czégel. *Generalized entropies, density of states, and non-extensivity*. Scientific reports 10.1 (2020): 1-12.
- [3] A. Ushio, J. Camacho-Collados. *T-NER: An All-Round Python Library for Transformer-based Named Entity Recognition*. arXiv preprint. arXiv:2209.12616
- [4] B. Asztalos, G. Palla, D. Czégel. *Subdiffusive semantic evolution in Indo-European languages*. arXiv preprint. arXiv:2209.04701