A BSc-képzés szakdolgozati témái

Valószínűségelméleti és Statisztika Tanszék

2020/2021

1. *Téma:* **Véletlen permutációk modellezése**
Témavezető: Arató Miklós
Rövid leírás: Szociológiai, pszichológiai kutatásokban igen gyakran fordulnak elő rangsorolási feladatok. A véletlen permutációkra az utóbbi években különböző modelleket javasoltak. A dolgozat bemutatná a véletlen permutációk különböző ábrázolási lehetőségeit és modelleit, továbbá néhány kapcsolódó feladatot szimulációkon keresztül is vizsgálna.
Ajánlott irodalom:
[1] M. A. Fligner and J. S. Verducci(eds.): *Probability Models and Statistical Analyses for Ranking Data* Springer-Verlag, New York (1993)
[2] Gladkich, Alexey and Peled, Ron: On the cycle structure of Mallows permutations,*Ann. Probab.*, Volume 46, no. 2 (2018) 1114-1169.
Szak: matematikus, alkalmazott matematikus

2. *Téma:* **Szűrési módszerek véletlen hálózatokban** (a téma már foglalt)
Témavezető: Backhausz Ágnes
Rövid leírás: Nagyméretű valós hálózatok statisztikai elemzésében felmerülnek az alábbi kérdések: (i) hogyan állapíthatjuk meg, hogy a gráf eltér egy homogén véletlen gráftól, például egy Erdős–Rényi-gráftól, és szignifikáns inhomogén szerkezettel rendelkezik? (ii) melyek azok az élek, amik a gráf szerkezetének szempontjából különösen fontosak, és amelyek együttesen a hálózat gerincét alkotják? (iii) milyen algoritmussal találhatjuk meg ezen élek halmazát? A szűrési módszerek témaköre ezekre a kérdésekre ad válaszokat, módszereket, többek között a Pólya-féle urnamodellen alapuló statisztikai eljárások segítségével.
A szakdolgozat célja a Pólya-féle urnamodell elméletének és az erre alapuló szűrési eljárások megismerése, feldolgozása a szakirodalom alapján, lehetséges biológiai (például az agy hálózatával kapcsolatos) alkalmazások keresése, ezek alapján valós hálózatok statisztikai elemzése.
Ajánlott irodalom:
[1] H. Mahmoud, Pólya urn models, Chapman&Hall, 2008.
[2] R. Marcaccioli, G. Livan, A Pólya urn approach to information filtering in complex networks, Nature Communications **10** (2019), 745.
[3] R. van der Hofstad, Random Graphs and Complex Networks I. Cambride University Press, 2017.

Szak: elemző

3. *Téma:* **Perkoláció rácsokon és véletlen gráfokon** (a téma már foglalt)
Témavezető: Backhausz Ágnes
Rövid leírás: A perkoláció elméletének [1, 2] alapkérdéseit fizikai példák motiválják. Kérdezhetjük például, hogy egy szilárd anyagban a szomszédos részecskék közötti kötéseket mennyi valószínűséggel lehet felbontani (a kötések felbontását egymástól függetlenül kisorsolva), hogy az anyag még szilárd maradjon, azaz található legyen benne egy nagyméretű összefüggő komponens. Hasonló kérdés akkor is érdekes, ha eredetileg is egy véletlen gráfból indulunk ki: például mondhatjuk, hogy egy nagyméretű valós hálózatot, az internetet modellezünk egy preferencialapú véletlen gráffal, majd azt kérdezzük, hogy mennyire robusztus a kapott gráf, vagyis az így kialakított éleket milyen valószínűséggel lehet felbontani (egymástól függetlenül), hogy maradjon nagy összefüggő komponens [3].
A szakdolgozat célja a perkolációelmélet alapjainak és véletlen gráfokra vonatkozó eredményeinek megismerése, feldolgozása a szakirodalom alapján, a témához kapcsolódó egy-egy állítás önálló bizonyításával, valamint saját számítógépes szimulációk készítése, melyek a fenti témakörök elméleti szempontból nyitott kérdéseit vizsgálják.
Ajánlott irodalom:
[1] H. Duminil-Copin, Introduction to Bernoulli percolation, 2018.
[2] J. E. Steif, A mini course on percolation theory, 2009.
`http://www.math.chalmers.se/ steif/perc.pdf`
[3] E. Jacob and P. Mörters, Robustness of scale-free spatial networks, Ann. Probab. **45** (2017), no. 3, 1680–1722.
Szak: alkalmazott matematikus

4. *Téma:* **Véletlen gráfok spektruma** (a téma már foglalt)
Témavezető: Backhausz Ágnes
Rövid leírás: A véletlen gráfokat (elsősorban a nagyméretű hálózatokkal való szoros kapcsolata miatt) az utóbbi évtizedekben sokféle szempontból vizsgálták. Egy lehetőség a véletlenül sorsolt gráf adjacenciamátrixának sajátértékeinek és sajátvektorainak viselkedését, eloszlását vizsgálata, illetve ennek következményeinek megértése a gráf szerkezetére nézve. Itt a véletlen gráf sokféle lehet, többek között tekinthetünk fokszám szerinti preferencián alapuló véletlen gráfmodelleket (amikor nagyobb fokszámú csúcsok nagyobb valószínűséggel kapnak új éleket, mint például a Barabási–Albert-modellben), vagy véletlen reguláris gráfokat, amikor minden csúcs fokszáma

azonos, a gráfot pedig ezzel a feltétellel véletlenszerűen, egyenletesen választjuk, illetve különböző sűrűségű Erdős–Rényi-gráfokat, amikor a csúcspárokat egymástól függetlenül adott valószínűséggel kötjük össze. Ha a gráfot kisorsoltuk, kérdés például, hogy a gráf sajátértékei hogyan oszlanak el, vannak-e például olyan értékek, amik pozitív valószínűséggel előfordulnak sajátértékként, vagy hogy a sajátvektorok koordinátái hogyan viselkednek.
A feladat a téma szakirodalmának feldolgozása, illetve olyan esetekben, amikor nincs elméleti eredmény, a fenti kérdések vizsgálata számítógépes szimulációk segítségével.
Ajánlott irodalom:
[1] C. Bordenave, Spectrum of random graphs. Jegyzet. https://www.math.univ-toulouse.fr/∼bordenave/coursSRG.pdf
[2] A. E. Litvak, A. Lytova, K. Tikhomirov, N. Tomczak-Jaegermann and P. Youssef, Structure of eigenvectors fo random regular digraphs, Transactions of the AMS, 371 (2019), no. 11., 8097–8172.
[3] D. Montealegre and V. Vu, Spectrum of complex networks. Kézirat. arXiv: 1809.05469
Szak: alkalmazott matematikus

5. *Téma:* **Algebrai statisztika a kontingenciatáblák elemzésében**
Témavezető: Csiszár Villő
Rövid leírás: Kérdőívet tölttetünk ki emberekkel: dohányoznak-e, magas-e a vérnyomásuk, hányan laknak egy háztartásban, néznek-e TVsorozatokat? A válaszokból kontingenciatáblát készítünk, majd megvizsgáljuk, hogy az egyes kérdésekre adott válaszok függetlenek-e egymástól. Találunk-e legalább feltételes függetlenséget? Kis számú megfigyelés esetén az algebrai statisztikát hívhatjuk segítségül a klasszikus khi-négyzet próba helyett.
Ajánlott irodalom: Móri Tamás - Székely Gábor (szerk.): Többváltozós statisztikai analízis XI. - XII. fejezete, illetve az algebrai statisztikához angol nyelvű cikkek (rövid összefoglalás a témavezet? disszertációjának 7-13. oldalán).
Szak: matematikus, alkalmazott matematikus.

6. *Téma:* **Az EM algoritmus néhány alkalmazása**
Témavezető: Csiszár Villő
Rövid leírás: Az EM (Expectation-Maximization) algoritmust a maximum likelihood becslés meghatározására használhatjuk, amennyiben rejtett változók is vannak. Számos szép alkalmazása van a statisztikában, pl. keverékeloszlások felbontása, Bayes-hálók tanulása, stb. A dolgozat ezekből az alkalmazásokból válogatna a hallgató érdeklődésének megfelelően.

Ajánlott irodalom: Dempster, Laird, Rubin: Maximum likelihood from incomplete data via the EM algorithm,
`http://project.mit.bme.hu/mi almanach/books/aima/ch20s03`
Szak: matematikus, alkalmazott matematikus, elemző.

7. *Téma:* **Információelmélet és portfóliók**
Témavezető: Csiszár Villő
Rövid leírás: Tőzsdei folyamatokat vizsgálunk, optimális portfóliókat illetve univerzális portfóliókat keresünk, információelméleti módszereket felhasználva. A téma a lóverseny-fogadások optimális stratégiáival is összefügg. A témát Cover és Thomas alapművének két fejezete alapján lehet feldolgozni, melyek számos tételbizonyítást, ugyanakkor példákat is tartalmaznak, illetve a fejezetek végén feladatok is szerepelnek. Természetesen a hallgatónak a felhasznált információelméleti fogalmakkal is meg kell ismerkednie.
Ajánlott irodalom: Cover, Thomas: Elements of Information Theory (2nd edition), 6. és 16. fejezet
Szak: matematikus, alkalmazott matematikus

8. *Téma:* **Markov-lánc CHT, algebrai megközelítések**
Témavezető: Csiszár Villő
Rövid leírás: A diszkrét idejű ergodikus Markov-láncok funkcionáljaira érvényes a centrális határeloszlás-tétel. A dolgozat egyik része a tétel bizonyításának precíz megértése lenne (szerepel a témavezet? egyetemi jegyzetében, de idő hiányában nem szokott sorra kerülni a mesterszakos előadáson). A dolgozat másik része a CHT-ben szereplő szórást vizsgálná: hogyan lehet algebrai eszközökkel kiszámítani az átmenetmátrixból; min múlik, hogy ez a szórás a független változók esetéhez képest kisebb vagy nagyobb? A téma feldolgozása során hasznos lehet szimulációs vizsgálatok lefolytatása is (pl. R-ben).
Ajánlott irodalom: Csiszár Villő: Diszkrét és folytonos idej? Markov-láncok (jegyzet)
Szak: matematikus, alkalmazott matematikus

9. *Téma:* **Véletlen bolyongások hatékonysága**
Témavezető: Gerencsér Balázs
Rövid leírás: A szimmetrikus véletlen bolyongás természetes választás lehet, ha egy komplikált térből véletlen mintát akarunk venni (pl. egy gráf jólszinezéseiből egyenletesen).
Ez kiemelten érdekes lehet, ha a téren egy (szép) függvény integrálját akarjuk közelíteni véletlen pontokon vett értékek átlagával, így tudunk jól véletlen

pontokat generálni. Ez az ún. Markov chain Monte Carlo algoritmuscsalád alapja.
Előfordul azonban, hogy sokat gyorsíthatunk az átmenetvalószínűségek módosításán, a futásidőt leíró *keverési idő* a töredékére csökken.
Irodalmi áttekintés mellett kedvenc gráfjainkra kereshetünk ilyen lehetőséget.
Ajánlott irodalom:
[1] P Diaconis, S Holmes, R M Neal, Analysis of a nonreversible Markov chain sampler, Ann. Appl. Probab., 10 (2000), pp. 726-752.
[2] P Diaconis, The Markov chain Monte Carlo revolution, Bull. Amer. Math. Soc., 46 (2009), pp. 179-205.
[3] S Boyd, P Diaconis, J Sun, L Xiao, Fastest mixing Markov chain on a path, The American Mathematical Monthly, 113(1), (2006), pp. 70-74,
[4] D A Levin, Y Peres, Markov chains and mixing times, vol. 107, American Mathematical Society, 2017
Szak: matematikus, alkalmazott matematikus

10. *Téma:* **A Pólya-fele urnamodell** (a téma már foglalt)
Témavezető: Móri Tamás
Rövid leírás: A Pólya-fele urnamodell legegyszerűbb változatában egy urnából, amelyben néhány fehér és fekete golyó van, golyókat húzunk ki véletlenszerűen, és a kihúzott golyót visszatesszük további c ugyanolyan színű golyóval együtt. Kérdés az urna összetételének leírása az egymás utáni húzások során, illetve aszimptotika adása. Ennek az egyszerű modellnek nagyon sok általánosítása van, amelyeket különböző területeken alkalmaznak. A dolgozat ezeket mutatná be.
Ajánlott irodalom:
[1] N. L. Johnson, S. Kotz: *Urn Model and Their Application*, Wiley, 1977
Szak: matematikus, alkalmazott matematikus

11. *Téma:* **Statisztikai módszerek, alkalmazások**
Témavezető: Pröhle Tamás
Rövid leírás: A jelentkező hallgatók érdeklődésének megfelelő, szabadon választott statisztikai téma
Ajánlott irodalom: Tipikusan angol nyelvű cikkek
Szak: alkalmazott matematikus, elemző

12. *Téma:* **Az elektronikus aláírások kriptográfiai algoritmusai** (a téma már foglalt)
Témavezető: Szabó István

Rövid leírás: Az informatikai biztonság fontos résztelülete a letárolt dokumentumok, illetve a kommunikáció (küldő fél és küldött dokumentumok) hitelességének biztosítása. Ennek egyik fontos eszköze az elektronikus aláírás (melynek keretrendszerét az AZ EURÓPAI PARLAMENT ÉS A TANÁCS 910/2014/EU RENDELETE határozza meg. A hitelesség matematikai, kriptográfiai megalapozását szabványosított algoritmusok biztosítják, melyek között fontos szerepet játszanak a kulcs nélküli (ún. hash) algoritmusok, és a diszkrét logaritmus problémán és a prím-faktorizáció nehézségén alapuló, valamint az elliptikus görbe kriptográfián alapuló nyilvános kulcsú algoritmusok. A szakdolgozatban elvárás az alkalmazott, szabványosított algoritmusok és biztonsági szempontjaik rövid ismertetése.
Ajánlott irodalom:
[1] Rivest, Shamir és Adleman: A method for obtaining digital signatures and public-key críptosystems, Comm. of. ACM, 1978
[2] Federal Information Processing Standard (FIPS) 186-3, The Digital Signature Standard, dated June, 2009.
[3] NIST Special Publication 800-57 Part 1 Revision 4 Recommendation for Key Management Part 1: General, 2016
[4] A. J. Menezes, P. C. van Oorshchor, S. A. Vanstone: Handbook of Applied Cryptography, http://www.cacr.math.uwaterloo.ca/hac/
[5] További saját forráskutatás
Szak: mind

13. *Téma:* **Az RSA kriptográfiai algoritmus paraméter-követelményeinek változása** (a téma már foglalt)
Témavezető: Szabó István
Rövid leírás: Miután Rivest, Shamir és Adleman publikálta az RSA algoritmust, nagyon gyorsan általánosan alkalmazásra került, és ma is jelentős szerepe van az internetes kommunikáció biztonsági megoldásaiban. A publikálás után jelentős kutatások indultak, publikációk sokasága jelent meg az alkalmazások támadási lehetőségeiről, ezek alapján a biztonsági szempontok folyamatos érvényesítése érdekében folyamatosan változtak az RSA paramétereivel szembeni követelmények, melyek érintették mind a paraméterek nagyságrendi elvárásait, mind a paraméterek véletlen választásának és tesztelésének eljárásait. A szakdolgozatban kerüljenek áttekintésre a kezdeti követelményektől a legújabb mértékadó szabványkövetelmények és néhány változás indokolása
Ajánlott irodalom:
[1] Rivest, Shamir és Adleman: A method for obtaining digital signatures and public-key cryptosystems, Comm.of ACM, 1978;

https://people.csail.mit.edu/rivest/Rsapaper.pdf
[2] NIST (National Institute of Standards and Technology) követelmények (pl. 800-56B, 800-57, 800-131A,...;
https://csrc.nist.gov/publications/sp800
[3] Matthieu Caneill, Jean-Loup Gilis: Attacks again the WiFi protocols WEP and WPA, 2010,
https://matthieu.io/dl/wifi-attacks-wep-wpa.pdf
[4] CCITT (Comité Consultatif Internationale de Télégraphique et Téléphonique) X:509, 11/1988 (C6 és C7 fejezet)
[5] SOG_IS (Senior Officials Group Information Systems Security) SOG-IS Crypto Working Group SOG-IS Crypto Evaluation Scheme Agreed Cryptographic Mechanisms (2018);
https://www.sogis.org/documents/cc/crypto/SOGIS-Agreed-Cryptographic-Mechanisms-1.1.pdf
[6] További saját forráskutatás
Szak: matematikus, alkalmazott matematikus

14. *Téma:* **Poszt-kvantum kriptográfia** (a téma már foglalt)
Témavezető: Szabó István
Rövid leírás: Jelentős matematikai kutatások alapozzák meg a klasszikus, az internetes kommunikációban is alkalmazott kriptográfiai algoritmusok biztonságát, melyek azon alapulnak, hogy nem ismertek hatékony, gyors törési algoritmusok, melyek reális időben támadhatóvá teszik ezen rendszereket.
Ezen algoritmusok addig védenek, ameddig nem lesznek elérhetőek nagy teljesítményű (ú.n. qbiteken operáló) kvantum-számítógépek, melyek fejlesztése, rész-sikerekkel hatalmas erőforrásokkal folyik. Ha a támadók is rendelkezhetnek kvantum-számítógépekkel, akkor ezekkel az eddigi nyilvános kulcsú algoritmusok (pl. a faktorizáció nehézségén alapuló RSA algoritmus, vagy az elliptikus görbéken alapuló algoritmusok) biztonsága megkérdőjeleződik, új elvű algoritmusokra lesz szükség.
A szakdolgozatban a hallgató tekintse át, hogy a kvantum-számítógépek korában mely algoritmusok maradnak biztonságosak, melyek nem, hogyan lehet gyorsan faktorizálni kvantum-számítógépen, milyen biztonságos algoritmusokat javasolnak, milyen matematikai alapokra építenek ezen algoritmusok. A feladat része a szakirodalom-kutatás is.
Ajánlott irodalom:
[1] Michael A. Nielsen and Isaac L. Chuang: Quantum Computation and Quantum Information,
http://mmrc.amss.cas.cn/tlb/201702/W020170224608149940643.pdf

[2] Samuel J. Lomonaco, JR.: A lecture on Shor's quantum factoring algorithm Version 1.1 ;
`https://arxiv.org/pdf/quant-ph/0010034.pdf`
[3] NISTIR 8105: Report on Post-Quantum Cryptography
`https://nvlpubs.nist.gov/nistpubs/ir/2016/nist.ir.8105.pdf`
Szak: matematikus, alkalmazott matematikus

15. *Téma:* **A budapesti ingatlanpiac változásai** (a téma már foglalt)
Témavezető: Zempléni András
Rövid leírás: Igen aktuális téma: a közelmúlt árrobbanásának kisebb részben közgadasági, de főleg matematikai elemzése a feladat. Historikus külföldi adatokat bevonva, analógiákat keresve a közeljövőben várható folyamatok előrejelzése zárná a dolgozatot.
A feladatot az R programnyelv segítségével elvégzett elemzéssel célszerű megoldani.
Ajánlott irodalom:
[1] Ádám Banai - Nikolett Vágó - Sándor Winkler: The MNB's house price index methodology. MNB Occasional Papers 127, 2017
Szak: elemző

16. *Téma:* **Európai városok levegőminőségi adatainak elemzése – különös tekintettel a COVID járvány hatásaira** (a téma már foglalt)
Témavezető: Zempléni András
Rövid leírás: Igen aktuális téma: a járvány "mellékhatásaként" a városokban javult a levegő minősége. Kérdés, hogy a karantén után mire lehet számítani. A dolgozatban a [2] cikkben összefoglalt legfontosabb statisztikai módszereket feldolgozva, a friss adatok alapján készülnének modellek.
A feladatot az R programnyelv segítségével elvégzett elemzéssel célszerű megoldani.
Ajánlott irodalom:
[1] `https://discomap.eea.europa.eu/map/fme/AirQualityExport.htm`
[2] Hamid Taheri Shahraiyni and Sahar Sodoudi: Statistical Modeling Approaches for PM10 Prediction in Urban Areas; A Review of 21st-Century Studies (Atmosphere, 2016)
Szak: elemző, alkalmazott matematikus