

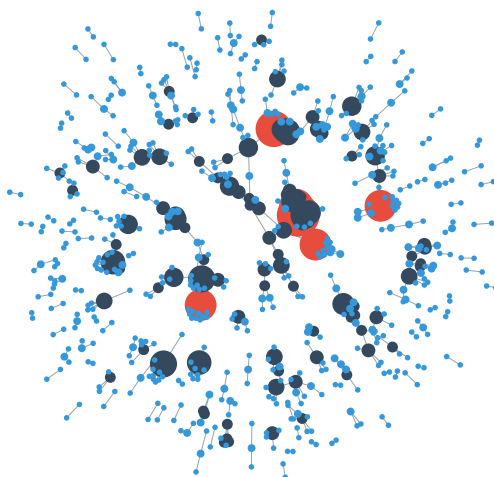


M Ű E G Y E T E M 1 7 8 2

REVEALING INFORMATION NETWORKS

PH.D. ÉRTEKEZÉS

TÉZISFÜZET



Szerző	Pálovics Róbert
	Számítástechnikai és Automatizálási Kutatóintézet
	Magyar Tudományos Akadémia
Témavezető	Benczúr András
	Számítástechnikai és Automatizálási Kutatóintézet
	Magyar Tudományos Akadémia

BUDAPESTI MŰSZAKI ÉS GAZDASÁGTUDOMÁNYI EGYETEM

2017

A Web 2.0 megjelenésével olyan közösségi hálózatok terjedtek el globálisan, mint például a Facebook vagy a Twitter. Ezen közösségi oldalak lehetővé teszik a felhasználók közötti azonnali tartalommegosztást, ezzel pedig katalizálják a globális kommunikációt. Napjaink technológiai fejlődése lehetővé tette, hogy a közösségi hálózatok mögött nagy adatbázisok gyűljenek fel, melyek a felhasználók múltbéli online tevékenységeit tartalmazzák. Mindezek kivételes lehetőséget adnak az ember közösségen belüli viselkedésének a tanulmányozására. Az ún. „*computational social science*” [7] megannyi fizikus, matematikus, szociológus és biológus kutató érdeklődését keltette fel.

A *komplex hálózatok* napjaink egy szintén népszerű kutatási ága, melynek eszköztára jól alkalmazható az emberi közösségbeli viselkedésének tanulmányozására. Megannyi hálózatkutatás célja az emberi közösségeket minél pontosabban leíró gráfmodellek fejlesztése. Ezen modellekben személyek a gráf csúcsai, míg a közöttük lévő kapcsolatok – például barátság, vagy információközlés – a gráf éleit adják. A hálózatok tudományának néhány népszerű feladata a/az (1) online közösségi hálózatok elemzése, (2) nagy gráfok klaszterezése, csoportok, közösségek detektálása sűrű részgráfok keresésével, (3) jövőben létrejövő kapcsolatok, élek predikciója.

Míg a web egy kivételes „online laboratórium” a kutatók számára, addig az óriási mennyiségű naponta generált online tartalom kihívás elé állítja a személyt mint felhasználót. „We are drowning in information but starved for knowledge” [9]. A web felhasználóinak egyik legfontosabb problémája a releváns cikkek, hírek és termékek szűrése. A webes keresés mellett erre a számítástudományi feladatra az *ajánlórendszerek* adnak megoldást. Az ajánlórendszerek a „The Netflix Prize” [2] adatbányászati versennyel lettek népszerűek ipari alkalmazásokban. Egy tipikus ajánlási feladat során az ajánlórendszer feladata egy toplista készítése mely a felhasználó számára releváns entitásokat (termékeket) tartalmazza, ezek lehetnek például könyvek, hírek, filmek, vagy akár emberi kapcsolatok. A toplista generálásához az ajánlóalgoritmusok számos adattípust használhatnak, például a felhasználó attribútumait, múltbéli online aktivitását, vagy aktuális kontextusát.

Kutatásunkban a globális információterjedést vizsgáljuk közösségi hálózat alapú adatok feldolgozásával. Célunk a következő kérdéseket megválaszolni:

- Hogyan keletkeznek az információs hálózatok? Milyen effektusok határozzák meg az információterjedést?
- Mi alapján detektálható az ún. közösségi hatás? Hogyan alkalmazható a közösségi hatás ajánlórendszerekben?
- Hogyan tudjuk a globális szintű információterjedést kihasználni ajánlórendszerekben?

Eredményeink a komplex hálózatok és az ajánlórendszerek kutatási témakörökhöz kapcsolódnak. Kísérleteinkben olyan közösségi média adatbázisokat használunk, melyek a felhasználók online tevékenységeinek idősorát tartalmazzák. Ezen felül a vizsgált adatokban rendelkezésre áll két kulcsfontosságú információ, a felhasználók közösségi hálózata, valamint a tevékenységeikhez köthető geolokációs információ.

Tézispontok

Az alábbiakban összefoglaljuk a disszertáció legfontosabb eredményeit. Minden egyes tézispont után feltüntetjük a kapcsolódó publikációkat, melyeket végül az utolsó fejezetben egy összesített listában is megadunk.

1. Információs hálózatok kialakulása

A komplex hálózatok területén számos kutatási eredmény foglalkozik egyetlen alapvető statisztikával, a gráf fokszámeloszlásával. A legtöbb valós adatok alapján konstruált gráf fokszámeloszlása skálafüggetlen [1], tehát hatványfüggvény eloszlást követ. Jelölje $p(d(i) = k)$ annak a valószínűségét, hogy az i csúcs fokszáma k . Ekkor a hatványfüggvény eloszlás

$$p(d(i) = k) = Ck^{-\alpha}, \quad \alpha > 2.$$

A legtöbb népszerű gráfmodell hatványfüggvény fokszámeloszlású hálózatokat generál.

Egy másik szintén alapvető, és széles körben vizsgált statisztika a gráf átlagfokszáma. Az olyan növekvő rendszerekben, mint pl. egy épülő közösségi hálózat, az átlagfokszám természetes módon növekszik a gráf növekedésével. A hálózatok tudományának egyik aktuális eredménye, hogy az átlagfokszám a gráf méretével hatványfüggvény szerint növekszik,

$$\bar{d}(n) = cn^b,$$

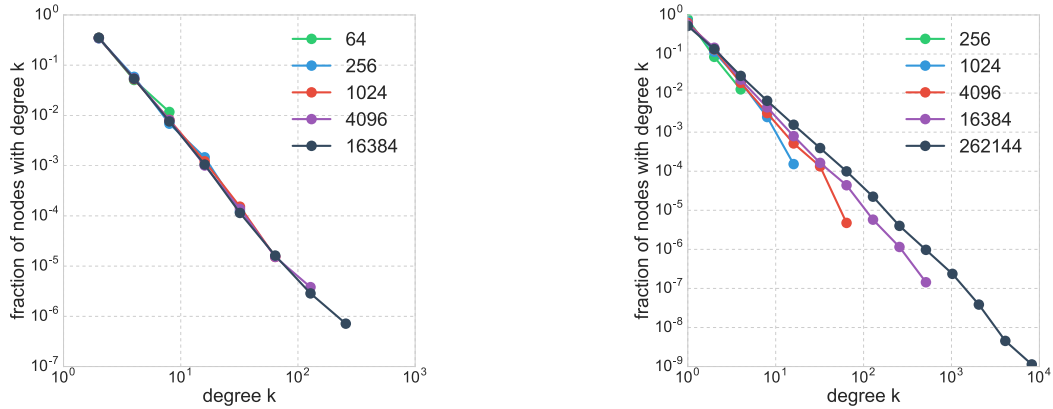
ahol n a gráf csúcsainak száma. Az effektus két népszerű elnevezése az „accelerated growth” [3] illetve a „densification law” [8].

Kutatásunkban az információs hálózatok növekedését olyan rendszerekben vizsgáljuk, ahol a hálózat minden csúcsa és éle egyszer adódik hozzá a hálózathoz, és nincsenek eltűnő csúcsok vagy élek a folyamatban. Méréseinkben először a fent említett két statisztikát, a fokszámeloszlást és az átlagfokszámot vizsgáljuk növekedő hálózatokban. Méréseinket három Twitter-hez kapcsolódó, valamint három ettől teljesen különböző, a publikus „Koblenz Network Collection” [5] adatbázisban megtalálható gráf adaton ismételtük meg.

Eredményeink:

1. tézispont: *Egy olyan új gráfmodellt adunk melynek célja az információs hálózatok növekedésének szimulációja. Modellünk képes két, általunk valós adatokon kimért alapvető statisztika reprodukálására. Kísérleteink szerint a valós információs gráfok átlagfokszáma $a + cn^b$ függvény szerint nő, eközben pedig a gráf skálafüggetlen fokszámeloszlásának exponense lassan csökken.*

- Méréseink szerint az átlagfokszám $a + cn^b$ függvény szerint nő. A hatványfüggvény mellett megjelenő konstans pontosabb illeszkedést ad a valós adatokra, valamint fontos lesz az általunk ajánlott gráfmodell esetén.
- A gráf fokszámeloszlása hatványfüggvény eloszlást követ, azonban a gráf növekedésével az eloszlás exponense folyamatosan csökken. A legtöbb ismert növekedő gráfokat leíró gráfmodellben az eloszlás exponense fix a folyamat során. Például a talán legismertebb, Barabási-Albert modellben is fix az eloszlás exponense, ahogy azt a 1. ábra bal oldali grafikonján is szemléltetjük. Ezzel szemben a valós adatokon végzett méréseinkben a fokszámeloszlásra jellemző, log-log skálán adódó egyenes meredeksége lassan csökken a méret növekedésével (1. ábra jobb oldali grafikon).
- A fenti két jelenséget körültekintő méréseinkre alapozva egy természetesen adódó elméleti összefüggéssel kapcsoljuk össze.
- Egy teljesen új gráfmodellt adunk, mely valós adatokon végzett méréseinkkel összhangban olyan hálózatokat eredményez, melyekben az átlagfokszám $a + cn^b$ szerinti növekedésével a gráf hatványfüggvény fokszámeloszlásának exponense csökken. A modell alapja a hálózatok



1. ábra. Növekedő információs hálózat fokszámeloszlása a hálózat különböző méreteinél (csúcsok száma). **Bal:** A Barabási-Albert modell esetén a hatványfüggvény eloszlás exponense fix. Log-log skálán a hatványfüggvényre jellemző egyenesek meredeksége megegyezik **Jobb:** A valós Twitter adaton a fokszámeloszlás szintén hatványfüggvény, azonban az exponens csökken a mérettel. A hatványeloszlásra jellemző egyenes meredeksége változik a gráf méretével.

tudományában közismert preferenciális kapcsolódás illetve az exponenciális növekedés. A modell minden egyes lépésében kétféle éltípust generálunk. *(i)* Egyrészt ún. „*random*” éleket, melyek olyan csúcsokat (személyeket) kötnek össze, akik az éllel (személyek közötti interakcióval) együtt jelennek meg a gráfban, valamint *(ii)* *homofília* éleket, melyek már megjelent csúcsokat kötnek össze. Részletesebben, minden egyes lépésben

- $r \cdot n(t)$ random élet generálunk, ahol r konstans a modell egy paramétere, $n(t)$ pedig az aktuális csúcsok száma.
- Minden egyes már megjelent i csúcs új homofília éleket hoz létre. Az i csúcs által létrehozott ilyen élek várható száma $s \cdot d_h(i)$, ahol s konstans a modell egy paramétere, $d_h(i)$ pedig a csúcs által eddig létrehozott homofília élek száma. Az létrehozott új kapcsolatok végpontjai preferenciális kapcsolódással [1] kerülnek kiválasztásra.

A legfontosabb különbségeket az irodalomban eddig ismertetett eredményekhez képest az alábbiakban foglaljuk össze.

- A skálafüggetlen fokszámeloszlás exponensének változását ugyan már észlelték, azonban legjobb tudomásunk szerint eddig nem modellezték.
- Modellünk az információs hálózatok születésétől fogva képes azok növekedését modellezni.
- Az általunk javasolt modell képes reprodukálni két legfontosabb mérésünket valós hálózatokban, a csökkenő hatványfüggvény eloszlás exponensét, valamint a $a + cn^b$ szerint növekedő átlagfokszámot.

Összefoglalásként eredményeink szerint állítjuk, hogy az információs hálózatok kezdetben véletlenszerűen, a már említett random élek megjelenésével keletkeznek, mint egy Erdős-Rényi gráf. Ezek után egy folyamatosan erősödő szabály lassan elnyomja az előző folyamatot, és ezzel előidéz a gráfban a skálafüggetlen fokszámeloszlást. A szabály erősödésével az eloszlás exponense csökken. Ez egyben azt is jelenti, hogy log-log skálán a skálafüggetlen fokszámeloszlásra jellemző

egyenes meredeksége lassan csökken. A skálafüggetlen hálózatot előidéző szabály modellünkben a preferenciális kapcsolódás, azonban elképzelhető, hogy más gráfépítő szabályok, például a háromszög lezárás is előidézhethet hasonló viselkedést.

Legfontosabb eredményünkben ezt véletlen gráfból egy magasabb szinten rendeződő hálózatba való átalakulást modellezzük. Modellünk alapja a hálózat kutatásban közismert preferenciális kapcsolódás és exponenciális növekedés. Eredményeinket korábban az alábbi publikációban ismertettük,

- I. Róbert Pálovics and András A Benczúr. Raising graphs from randomness to reveal information networks. In *Proceedings of WSDM 2017*, 2017.

2. Az online top- k ajánlás probléma

A továbbiakban ismertetet eredményeink az ajánlórendszerek kutatási területéhez kapcsolódnak. Az ajánlórendszerek információszűrő algoritmusok, melyek célja a felhasználók preferenciáinak predikciója. Képesek a felhasználó számára releváns entitások (termékek) szűrésére abban az esetben, ha ezek közül túl sok áll rendelkezésre. Ajánlórendszerekre néhány példa (1) zenei ajánlás olyan ún. „streaming” zenei szolgáltatásokban, mint a Spotify, (2) film ajánlás olyan online filmkatalógusokban, mint a Netflix, (3) klasszikus termékajánlás olyan online webshopokban, mint az Amazon.

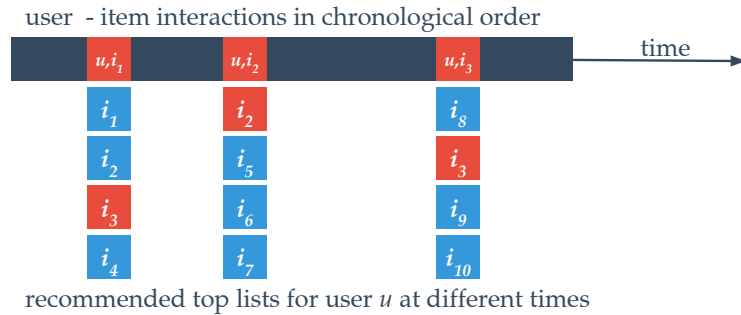
A számítástudományban az ajánlórendszerek kutatása a 2009-ben megrendezett „Netflix Prize” adatbányászati versennyel vált kiemelten népszerűvé. A verseny feladata ún. *kötegelt értékelés predikció* volt, melyben egy véletlen módon kettéosztott adaton folyt a verseny. A versenyzők az adat egyik felén tanították modelljeiket, melyekkel cél volt minél pontosabban megjósolni a másik részben számukra ismeretlen felhasználó-termék értékeléseket.

Ugyan vitathatatlan, hogy a verseny kulcsfontosságú volt az ajánlórendszerek elterjedése szempontjából, azonban az akkor definiált feladat mára már kevésbé aktuális. Napjaink online szolgáltatásaiban az ajánlórendszereknek a felhasználó számára általában egy toplistát kell prezentálniuk. Továbbá a felhasználók preferenciái, valamint az elérhető termékek is gyorsan változhatnak időben, ilyen feladat például a hír ajánlás. Végül érdemes megjegyezni, hogy az algoritmusok olyan kontextus adatokat is kihasználhatnak, melyek időben változnak, ilyen például a felhasználó geolokációja.

A fentiek miatt egy valós szolgáltatásban az ún. *online tanulás alapú top- k ajánlás* relevánsabb feladat, mint a klasszikus kötegelt értékelés ajánlás. Ez a feladat az ajánlórendszerek kutatásában eddig kevés figyelmet kapott. Eredményeinkben célunk minél pontosabb top- k ajánlás adása időben gyorsan változó, ún. instacioner rendszerekben. Célunk az általunk javasolt ajánlórendszer modellek azonnal frissítése amint új információ áll rendelkezésünkre a felhasználóról. Eredményeink:

2. tézispont: *Definiáljuk az online top- k ajánlás feladatot, mely a szakirodalomban eddig ismertetett kísérleti elrendezéseknél jobban modellezi a napjaink online szolgáltatásaiban adódó ajánlási problémákat. Definiáljuk az online sztochasztikus gradiens alapú mátrixfaktorizációs algoritmust, amit későbbi méréseinkben baseline módszerként alkalmazunk.*

- Formálisan definiáljuk az instacioner rendszerekben való top- k ajánlást, melyre a továbbiakban „online ranking prediction” problémaként hivatkozunk. Ez esetben feltételezzük, hogy az ajánlórendszer algoritmus számára elérhető a felhasználói online aktivitások egy



2. ábra. Az online top- k ajánlás probléma.

idősora, melyet a rendszer kronologikus sorrendben dolgoz fel, ezzel folyamatosan frissítve modelljét. Minden egyes új felhasználó-entitás interakció esetén a 2. ábrának megfelelően

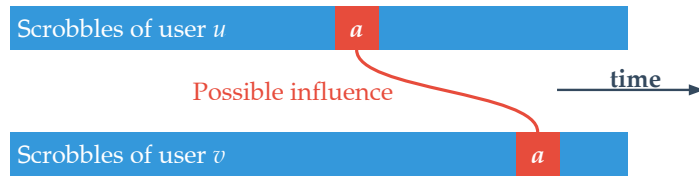
- lekérdezzük a felhasználó számára egy top- k ajánlást az ajánlórendszerből,
 - a generált toplistát az egyetlen aktuális felhasználói interakció alapján értékeljük ki,
 - végül tanítjuk az ajánló modellt felfedve a megtörtént interakciót.
- A szakirodalomban ismert, iparban is népszerű mátrixfaktorizációs módszernek [4] egy robusztus online verzióját adjuk meg. Ez a modell azért is releváns, mert kísérleteink nagy részében a legerősebb baseline módszerként teljesített. A sztochasztikus gradiens (SGD) alapú módszerek a konvergencia érdekében hagyományosan többször iterálnak egy adaton. Ezzel szemben egy valós idejű ajánlás esetén a rendszerbe kerülő új információt (amikor egy felhasználó valamilyen új terméket adoptál) azonnal fel kell dolgozni, és nincs idő az eddig modell teljes újratanítására. Az online SGD alapú mátrixfaktorizáció esetén ezért minden egyes felhasználói interakciót egyszer dolgozunk fel, és időrendben iterálunk végig az adaton.
 - Későbbi méréseinkben szükség lesz a különböző algoritmusok predikcióinak kombinálására. Ehhez egy online SGD alapú modell kombináló keretrendszerre adunk mely képes differenciálható és nem differenciálható modellek hatékony, perszonalizált online kombinálására.

Mivel az online top- k ajánlás problémát az új kontextusfüggő modelljeinket tartalmazó cikkjeinkben vezettük be, ezért ezen tézisponthoz vonatkozó publikációkat a következő fejezetekben adjuk meg.

3. Közösségi hatás detektálása

Kutatásunkban egyik célja az ún. közösségi hatás detektálása olyan rendszerekben, ahol a közösségi hatásról nincs direkt információnk. Méréseinket a Last.fm online szolgáltatás adatain végeztük, mely egy zenehallgatás köré szerveződő közösségi hálózat. A továbbiakban azon eseményre, amikor egy Last.fm felhasználó adott időben meghallgat egy adott előadót, a Last.fm terminológiáját használva „scrobble”-ként hivatkozunk. Méréseinkben a A Last.fm „scrobble” bejegyzések idősorát és a felhasználók statikus közösségi gráfját vizsgáljuk.

Ahogy az a 3. ábrán látható, olyan eseményeket detektálunk, amikor két felhasználó egymás után ugyanazt a zenei előadót hallgatja meg valamilyen késleltetéssel (közös meghallgatás). Egy



3. ábra. Közösségi hatáshoz kapcsolódó esemény, mely esetén u és v felhasználók hallgatták meg egymás után a előadót.

közös meghallgatás oka lehet a két felhasználó közötti közösségi hatás. Eredményeinkben két felhasználó közötti közösségi hatás valószínűségére adunk becslést abban az esetben, amikor detektálunk egy közös meghallgatást. Fontos megjegyezni, hogy a hatás valószínűségének becslése egy ismert és nehéz feladat, mivel az adatban detektált közös meghallgatásokat indukálhatják más effektusok, pl. homofília is. A hatás valószínűségére adott összefüggésünket később felhasználjuk egy hatás alapú ajánlórendszer fejlesztésére. Mivel a felhasználók közötti hatás egy időfüggő effektus, az általunk adott ajánlórendszert az online top- k ajánlás probléma keretein belül vizsgáljuk.

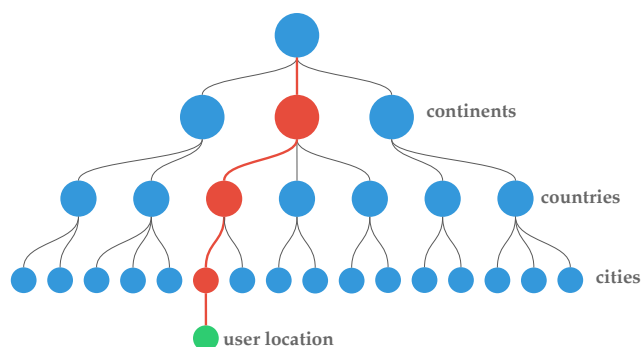
Eredményeink:

3. tézispon: *Az ún. közösségi hatást detektáljuk olyan rendszerekben, ahol az effektusról nincs közvetlen információnk. A hatás valószínűségének becslésére adott módszerünk alapján egy ajánlóalgoritmust fejlesztünk az online top- k feladatra.*

- Egy új módszert adunk a közösségi hatás detektálására a Last.fm-hez hasonló adatokon. Mivel az adat felépítése elég általános, ezért úgy gondoljuk, hogy eredményeink más rendszerekben is alkalmazhatóak. Méréseink mellett a közösségi hatás valószínűségének közelítésére elméleti összefüggést is adunk.
- Egy olyan ajánlórendszer keretrendszert fejlesztünk, mely képes a felhasználók közötti múltbéli hatások detekciója alapján jövőbeli hatásokat pontosan predikálni. Ezt kihasználva a rendszer közösségi hatás alapú top- k zenei ajánlás generálására képes. Az új ajánlórendszert az online top- k ajánlás problémán keresztül vizsgáljuk.
- Egy kifinomultabb verzióját adjuk a hatás alapú ajánlórendszerünknek kontextus-alapú faktorizációs módszerek alkalmazásával.
- Sikeresen kombináljuk az új ajánlóalgoritmust több baseline algoritmussal, melyek közül a legrelevánsabb a már említett online SGD alapú mátrixfaktorizációs módszer.

A fenti eredményeket az alábbi cikkekben publikáltuk.

- II. Róbert Pálovics and András A Benczúr. Temporal influence over the Last.fm social network. In *Proceedings of IEEE/ACM ASONAM 2013*, pages 486–493. ACM, 2013,
- III. Róbert Pálovics and András A Benczúr. Temporal influence over the Last.fm social network. *Social Network Analysis and Mining*, 5(1):1–12, 2015,
- IV. Róbert Pálovics, András A Benczúr, Levente Kocsis, Tamás Kiss, and Erzsébet Frigó. Exploiting temporal influence in online recommendation. In *Proceedings of RecSys 2014*, pages 273–280. ACM, 2014.



4. ábra. A geolokáció alapú hierarchikus ajánlórendszer modell.

4. Hierarchikus geolokációt használó ajánlórendszer modellek

Végső eredményeinkben olyan közösségi hálózatot vizsgálunk, ahol a felhasználók által generált eseményekhez geolokációs információ is tartozik. A Twitter [6] mikroblog szolgáltatás adatain kísérletezünk, ahol a felhasználók különböző híreket oszthatnak meg egymással. Minden egyes felhasználó által generált bejegyzéshez elérhető a felhasználó aktuális GPS pozíciója. Az adatot felhasználva egy olyan GPS pozíció alapú ajánlórendszert fejlesztünk, mely a hírek globális elterjedése alapján képes időfüggő top- k hírajánlásra. Mivel a felhasználók számára releváns és elérhető hírek folyamatosan változnak időben lokációtól függően, a feladatot az online top- k ajánlásnak megfelelően definiáljuk. Eredményeink:

4. tézispont: *Olyan hierarchikus, lokáció alapú ajánló algoritmusokat definiálunk, melyek képesek a hírek globális terjedésének a detektálására. Új algoritmusainkat az online top- k ajánlás probléma kereteiben vizsgáljuk.*

- Olyan frissesség és népszerűség alapú időfüggő ajánlórendszer algoritmusokat definiálunk, akkor is alkalmazhatóak, amikor a túlságosan ritka adat miatt a mátrixfaktorizációs módszerek nem használhatóak.
- A fenti algoritmusokat egy geolokáció alapú hierarchikus modellben alkalmazzuk a 4. ábrának megfelelően. A modellezés alapötlete, hogy kontinens/ország/város/... szintű lokális frissesség és népszerűség modelleket definiálunk és tanítunk. Ezek után a felhasználó aktuális lokációját használva a megfelelő szinteken lévő modellek kimenetét kombináljuk a végső top- k ajánlás generálásához.
- Modelljeink képesek globális szinten detektálni az egyes hírek és trendek közösségi hálózatban való diffúzióját.
- Az egyes szinteken tanult lokális modelleket online SGD használatával kombináljuk.
- Minden egyes hierarchikus modellhez baseline-ok egy sokaságát adjuk, melyekkel célunk annak eldöntése, hogy az adott modell esetén a hierarchikus modellezés mennyit ad hozzá az ajánlórendszer végső teljesítményéhez.

Eredményeinket az alábbi publikációban közöltük:

- V. Róbert Pálovics, Péter Szalai, Júlia Pap, Erzsébet Frigó, Levente Kocsis, and András A Benczúr. Location-aware online learning for top- k recommendation. *Pervasive and Mobile Computing*, 2016.

Kapcsolódó publikációk

Az alábbi lista összefoglalja a disszertációhoz kapcsolódó publikációkat.

- I. Róbert Pálovics and András A Benczúr. Raising graphs from randomness to reveal information networks. In *Proceedings of WSDM 2017*, 2017
- II. Róbert Pálovics and András A Benczúr. Temporal influence over the Last.fm social network. In *Proceedings of IEEE/ACM ASONAM 2013*, pages 486–493. ACM, 2013
- III. Róbert Pálovics and András A Benczúr. Temporal influence over the Last.fm social network. *Social Network Analysis and Mining*, 5(1):1–12, 2015
- IV. Róbert Pálovics, András A Benczúr, Levente Kocsis, Tamás Kiss, and Erzsébet Frigó. Exploiting temporal influence in online recommendation. In *Proceedings of RecSys 2014*, pages 273–280. ACM, 2014
- V. Róbert Pálovics, Péter Szalai, Júlia Pap, Erzsébet Frigó, Levente Kocsis, and András A Benczúr. Location-aware online learning for top-k recommendation. *Pervasive and Mobile Computing*, 2016

Irodalomjegyzék

- [1] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.
- [2] James Bennett and Stan Lanning. The Netflix prize. In *KDD Cup and Workshop in conjunction with KDD 2007*, 2007.
- [3] Sergey N Dorogovtsev and JFF Mendes. Accelerated growth of networks. *arXiv preprint cond-mat/0204102*, 2002.
- [4] Yehuda Koren, Robert Bell, Chris Volinsky, et al. Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37, 2009.
- [5] Jérôme Kunegis. Konect: the Koblenz network collection. In *Proceedings of the 22nd international conference on World Wide Web companion*, pages 1343–1350. International World Wide Web Conferences Steering Committee, 2013.
- [6] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. What is Twitter, a social network or a news media? In *Proceedings of the 19th international conference on World wide web*, pages 591–600. ACM, 2010.
- [7] David Lazer, Alex Sandy Pentland, Lada Adamic, Sinan Aral, Albert Laszlo Barabasi, Devon Brewer, Nicholas Christakis, Noshir Contractor, James Fowler, Myron Gutmann, et al. Life in the network: the coming age of computational social science. *Science (New York, NY)*, 323(5915):721, 2009.
- [8] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. Graph evolution: Densification and shrinking diameters. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 1(1):2, 2007.

- [9] John Naisbitt and J Cracknell. Megatrends: Ten new directions transforming our lives. Technical report, Warner Books New York, 1984.
- [10] Róbert Pálovics and András A Benczúr. Temporal influence over the Last.fm social network. In *Proceedings of IEEE/ACM ASONAM 2013*, pages 486–493. ACM, 2013.
- [11] Róbert Pálovics and András A Benczúr. Temporal influence over the Last.fm social network. *Social Network Analysis and Mining*, 5(1):1–12, 2015.
- [12] Róbert Pálovics and András A Benczúr. Raising graphs from randomness to reveal information networks. In *Proceedings of WSDM 2017*, 2017.
- [13] Róbert Pálovics, András A Benczúr, Levente Kocsis, Tamás Kiss, and Erzsébet Frigó. Exploiting temporal influence in online recommendation. In *Proceedings of RecSys 2014*, pages 273–280. ACM, 2014.
- [14] Róbert Pálovics, Péter Szalai, Júlia Pap, Erzsébet Frigó, Levente Kocsis, and András A Benczúr. Location-aware online learning for top-k recommendation. *Pervasive and Mobile Computing*, 2016.