

re irányuló folyamatos vásárlásaink elemzésével – összekapcsolva a SuperShop kártya-kérelmünkhöz megadott adatainkkal – konkrét ajánlatot kapjunk a tavaszi virág- vagy zöldségpálánta értékesítési akciójukról, véletlenül sem tévesztve el, hogy virágoskertünk vagy konyhakertünk van, esetleg mind a kettő.

A collaborative filtering módszer alkalmazásának egyszerűsített, sematikus példáját mutatja be az alábbi táblázat. Tegyük fel, hogy egy videotékában ismerik és feljegyzik a kölcsönző vásárlók adatait és véleményét. Ezt tartalmazza az alábbi sematikus táblázat:

film / vevő	1. vevő	2. vevő	3. vevő
A film	4	3	4
B film	5	3	5
C film	2	4	2
D film	5	2	? (nyilván 5)

Tetszési index, értékelés személyenként  
(1= nagyon rossz, 5 = kiváló)

Minthogy a 3. vevő már három film kapcsán hasonló véleményt fogalmazott meg, mint

#### IRODALOM

- Kandikó József (2005): Kapcsolat a vevővel. Harvard Businessmanager. 3, 55–63.  
Kotler, Philip (2000): *Kotler a marketingről. Jönni, látni, győzni – a piacon*. Park, Budapest  
Kotler, Philip – Jain, D. C. – Maesincee, S. (2003): *Marketing lépések. Nyereséges növekedés és megújulás a 21. században*. Park Könyvkiadó, Budapest

az 1. számú vásárló, érthető, hogy ha a 3. vevő ajánlatot kér, hogy mit kölcsönözzön ki legközelebb, az egyébként hasonló ízléssel rendelkező 1. vevő által kitűnőnek ítélt D filmet kockázat nélkül ajánlhatjuk neki.

Képzeld el, hogy ez a mechanizmus egy nagy adatbázis alapján működik, és ezeket az információkat a számítógép automatikusan szolgáltatja, mi több – a vevőkről szóló információk alapján az előre elkészített ajánlószövegekből a megfelelőt kiválasztva – az ajánló e-mailt (vagy sms-üzenetet) el is küldi a vevő címére. A piacbefolyásolás, rábeszélés e formáját – ha erre a reklámlevél szövegében kellően ügyelünk – a vásárló a vevőkiszolgálás magas színvonalú megnyilvánulásaként értelmezi és értékeli. Ez a XXI. századi marketing-sajátosságok egyik legkézenfekvőbb stratégiája és megnyilvánulása.

Kulcsszavak: *törzsvásárló rendszerek, vonalkódos árubeazonosítás, vásárlók beazonosítása, microtargeting, holisztikus marketingkonceptió, collaborative filtering*

- Majó Zoltán – Révész Balázs (2007): *A collaborative filtering szerepe az online marketing munkában: egy időszoros kutatás magyar vonatkozásai*. MMSz Marketing Oktatók Klubja 2007. évi konferenciája, Gyöngyös  
Riedl, John – Konstan, J. – Majó Z. – Révész B. (2004.): *Szajtpropaganda*. KJK-Kerszöv, Budapest

## KAPCSOLATOK ÉS TÁVOLSÁGOK: A HAZAI VEZETÉKES HÍVÁS-SZOKÁSOK ELEMZÉSE\*

Kurucz Miklós

doktori ösztöndíjas

Siklósi Dávid

doktori ösztöndíjas

Csalogány Károly

doktori ösztöndíjas

Lukács László

doktori ösztöndíjas

Benczúr András

PhD, laborvezető

Lukács András

PhD, csoportvezető

MTA SZTAKI Informatika Kutató Laboratórium Adatbányászat és Webes Keresés Kutatócsoport  
{mkurucz, lacko, sdavid, benczur, cskaresz, lukacs}@ilab.sztaki.hu

Az elmúlt tíz évben ismereteink jelentősen bővültek a szociális hálózatok tulajdonságainak, a kapcsolatok kialakulásának és változásainak és a hálózati résztvevők osztályozásának területein, amelyek kiemelt fontosságúak a telefontulajdonosok szegmentációja, a marketing célcsoportok kiválasztása, az információterjedés vagy egyéb viselkedés előrejelzése céljából is. A szakirodalomban több összefoglaló munkát találhatunk, amelyek közül kiemeljük Barabási Albert-László (2002) magyarul is megjelent művét.

Tanulmányunkban, amely a (Kurucz et al., 2008) cikk válogatott kivonata, a szociális hálózatok modellezésének és vizsgálatának legújabb eredményeit alkalmazva bemutatjuk a hazai vezetékessé telefonhívások által megrajzolt kapcsolati hálózat elhelyezkedését

\* A TEXTREND NKFP-07-A2 és az OTKA NK 72845 támogatásával.

az ország területén belül. Egyedülálló adathalmazunk több millió anonimizált telefon-tulajdonos hosszú időszakon keresztül történő összesített viselkedését tartalmazza. Eredményeink alapján intuíciót nyerhetünk a szociális kapcsolatok térbeliségéről. Három fő területet vizsgálunk: a kapcsolatok távolságának eloszlását és ezen keresztül az ún. „kis világ” létrejöttét; az egymással szorosabb kapcsolatban álló csoportok települések szerinti hierarchikus tagozódását; végül pedig a hálózati kapcsolatok alkalmazhatóságát a résztvevők tulajdonságainak modellezésében és előrejelzésében.

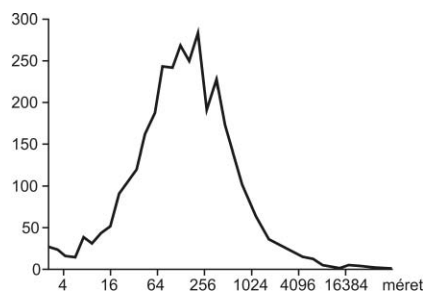
Legfontosabb megfigyeléseink között említhető a közösségek hierarchiájának erőteljes, felülről lefelé történő szerveződése, amelyben a nagyvárosok egyetlen, nehezen továbbosztható csoportként jelennek meg. A hívások hálózatának fő jellemzői, hogy a kapcsolatok

száma hatványeloszlást követ, aminek következtében a közösségek a sok kapcsolattal rendelkező központok köré csoportosulnak. A hálózat résztvevőinek jelentős része ugyanakkor a központokból induló hosszú nyúlványokban helyezkedik el. A nagyléptékű struktúra ugyanakkor csak a sűrű magok és a nyúlványok mint lokális jelenségek tisztítása és előfeldolgozása után ismerhetők meg, ez önálló kutatási eredményünk.

#### Az adathalmaz

Az általunk vizsgált, több mint kétmillió telefonszámra alapuló előálló hálózat követi a szociális hálózatok általános viselkedését. A kapcsolatok száma hatványeloszlást követ (Barabási et al., 2000), azaz a  $k$  darab kapcsos-

	nagy	kicsi
tagok (ezer)	2,072	74
tagok óriás komponensén kívül	130	15
időtartam (hónap)	8	12
kapcsolatok, irányított (ezer)	48,400	3,965
kapcsolatok, kölcsönös (ezer)	10,800	706



1. ábra • A két vizsgált híváshálózat fő tulajdonságai és a telefonszámra alapuló előálló hálózat eloszlása településméret szerint, a közel 600 ezer felhasználót számláló Budapest elhagyásával

lattal rendelkezők száma  $k^Y$ -val arányos, ami azt is jelenti, hogy meglepően sok nagy kapcsolati számmal rendelkező telefonszámra alapuló hálózatot találunk. A hálózat átmérője kicsi (Watts – Strogatz, 1998; Albert et al., 1999), azaz tetszőleges két személy között általában néhány lépésen keresztül kapcsolatot tudunk teremteni. Végezetül a résztvevők kisszámú kivételtől eltekintve egyetlen hatalmas összefüggő hálózat részét képezik. A bemutatásra kerülő mérések egy kisebb és egy nagyobb hálózaton történnek, amelynek adatait és a települések méret szerinti eloszlását az 1. ábra tartalmazza. Az eloszlás a kiugróan magas, majdnem 600 ezer telefonvonalal rendelkező Budapest elhagyása után közel lognormális.

#### Navigáció a kis világban

Az ún. „kis világ” jelenséget elsőként Stanley Milgram (1967) mutatta be híres kísérletében, amelyben az Egyesült Államok polgárai közötti kapcsolati távolságok átlagos értékeként nem egész hat lépés adódott. A kicsi átmérő tulajdonságot azóta több hálózat, így például a World Wide Web esetében is megfigyelték (Albert, 1999), és az általunk vizsgált hálózatnak is sajátja. Néhány híresebb modell (Watts – Strogatz, 1998) között számunkra kiemelkedő fontosságú Jon Kleinbergé (2000), amely a kis átmérő mellett Milgram kísérletének egy további kulcsmozzanatát is modellezi. A tetszőleges két személy közötti hat lépéses távolság ugyanis nemcsak elméletben létezik, hanem a kísérlet résztvevői képesek is azt pusztán lokális kapcsolati információjuk segítségével megtalálni. Valójában Milgram kísérletében néhány nem lokális információ is adott volt, például a célszemély lakhelye vagy foglalkozása. Kleinberg modelljében a navigáció a lakhelyen, azaz egészen pontosan a célszemély földrajzi koordinátáin alapul, s

az útvonal minden lépésben a földrajzilag legközelebbi vivő kapcsolatokon keresztül halad.

Kísérletünkben megvizsgáltuk Kleinberg (2000) útvonalkereső módszerének lépésszámát és az általa adott modell paramétereinek érvényességét a hazai telefonhívások által rajzolt kapcsolati hálózaton. Összességében megállapítottuk, hogy ha Budapest torzító hatását kiszűrjük, akkor a modell kellően jól illeszkedik a hazai kapcsolati hálóra, ám Budapesttel együtt ugyanez már kevésbé mondható el. Valójában Budapesten belül már az adatok hiányossága miatt sem tudtuk az útvonalkereső algoritmust alkalmazni, mivel utca- és házszám szintű földrajzi pozíció nem állt rendelkezésünkre. Ezért általánosságban is megelégedtünk azzal, hogy a céltelepülést elérjük, azaz például két budapesti (vagy azonos településen lakó) személy között azonnal, nulla lépésben véget ér az útvonalkeresés.

A mérési eredményeket a 2. és a 3. ábrán mutatjuk be. Az első a települések elhelyezkedését ábrázolja (Budapest tehát egyetlen pont), illetve Budapest hatását kiszűrő kiemeli a kiválasztott nyugati országrészt. A 2. ábrán az adott távolságra levő kapcsolatok számát láthatjuk. Ez Kleinberg modellje szerint a távolság  $-d$  hatványával arányosan csökken, ahol  $d$  a földrajzi tér dimenziószáma – normál esetben tehát kettő. Miközben a teljes ország területén az illeszkedés meglehetősen zajos, és a  $-1$  hatvány grafikonja közelebb áll az egyeneshez, mint a  $-2$  hatványé, a nyugati országrészben már a modell által jósolt  $-2$  hatvány illeszkedést látjuk. A teljes ország talán egy torz, Budapesttől mért egydimenziós távolságú képet mutat, azonban Budapestet kihagyva már egészséges kétdimenziós kapcsolati mozgást látunk. A 3. ábrán azt is megfigyelhetjük, hogy egymillió véletlen telefonszámra alapuló pár között a nagy

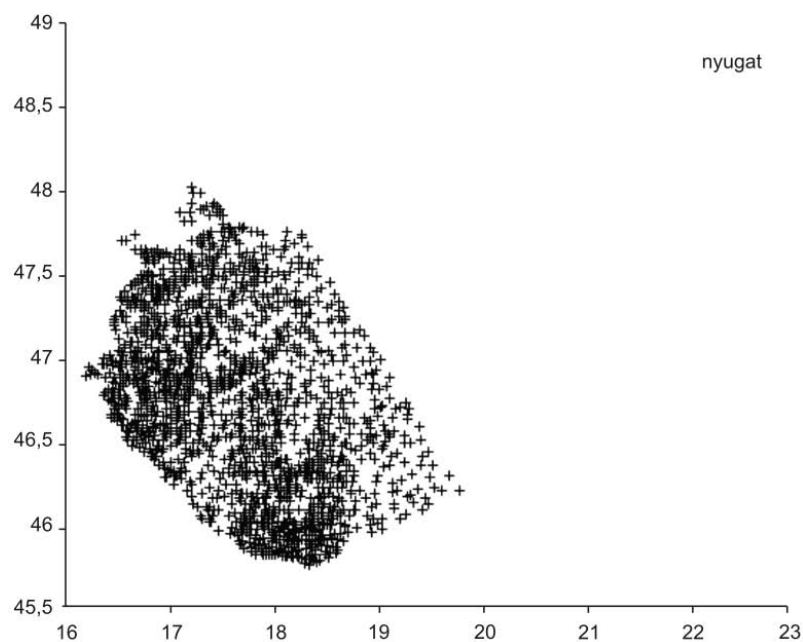
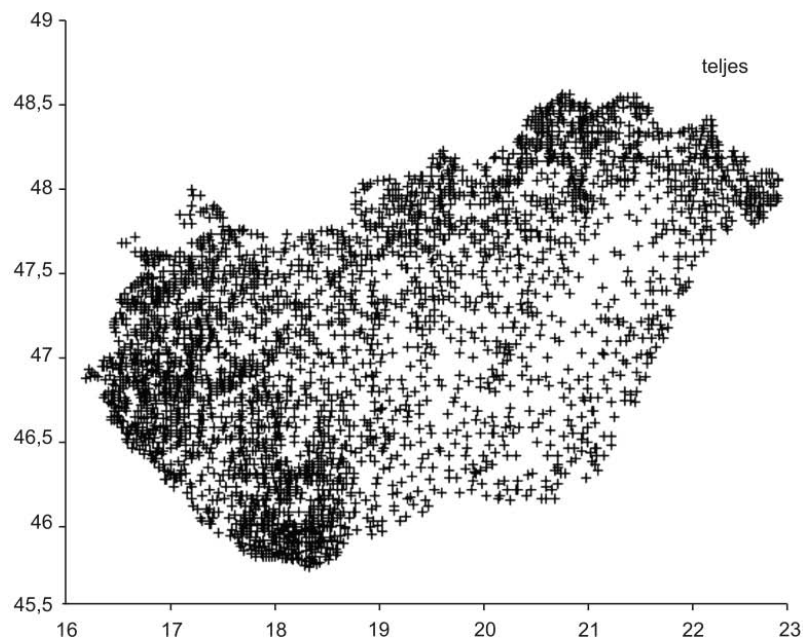
kapcsolati távolságok exponenciális sebességgel tűnnek el, és a tíz lépésen belül nem elérhető párok száma százas nagyságrendű marad. Ne felejtjük azonban el, hogy az útvonal keresése a települést elérve megáll, ami jogos lehet egy falu esetében, de értelmetlen egy nagyvárost, vagy Budapestet tekintve.

#### Klaszter-analízis

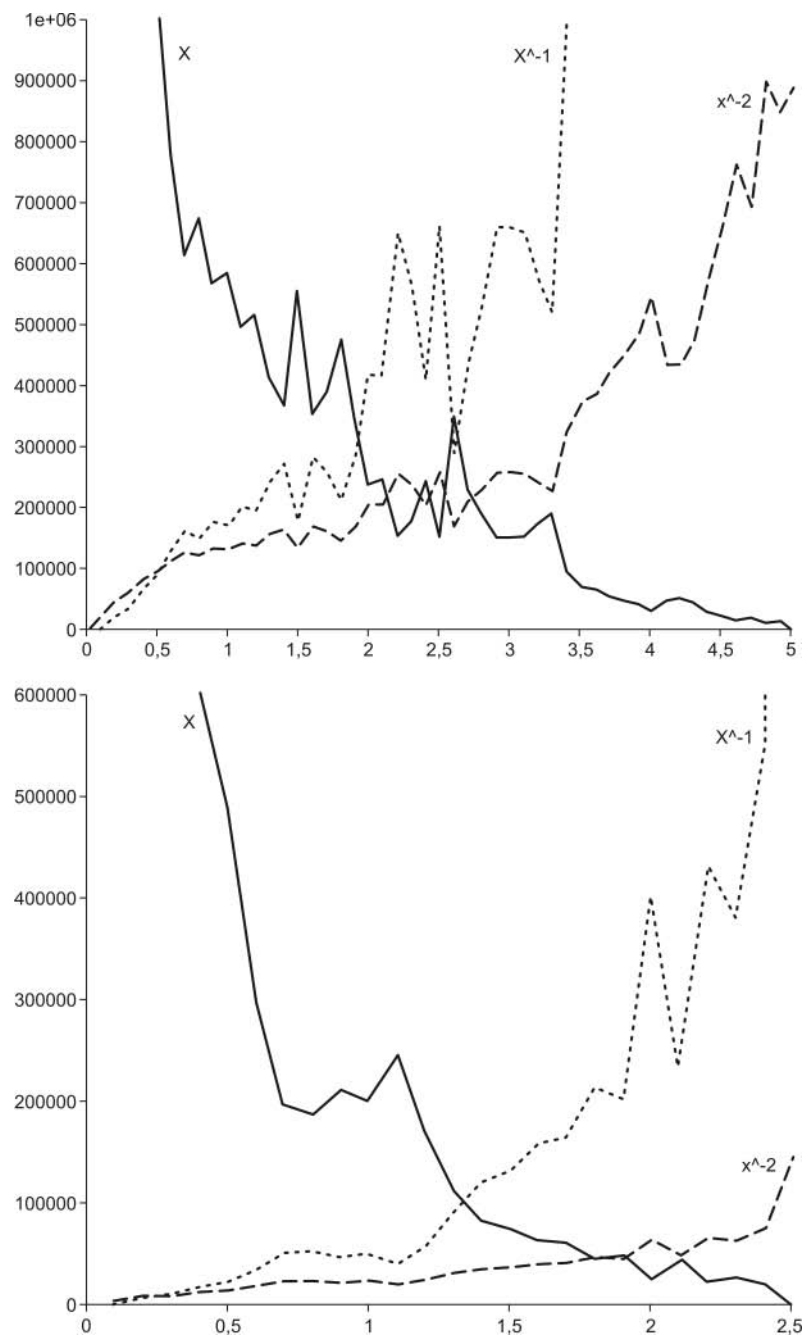
A klaszterezés vagy csoportosítás, (felügyelet nélküli) osztályozás célja az ügyfelek értelmezhető, fontos információkat bemutató, a teljes állomány áttekintését megkönnyítő felosztása. Telefonszámra alapú esetében így az ügyfeleket szegmentálhatjuk, kívánatos vagy elkerülendő csoportokat határozhatunk meg, amelyekben például nagy a hajlandóság valamely új szolgáltatás kipróbálására – vagy éppen lemondására, fizetési késedelemre.

Telefonhívás-hálózatok klaszterezése esetén azzal a nehézséggel szembesülünk, hogy a kis világ tulajdonság (Milgram, 1967; Watts–Strogatz, 1998) következtében az egyes közösségek környezete erősen keveredő, átfedő. A kapcsolatok számának hatványeloszlása (Barabási et al., 2000) következtében pedig sok olyan csomópontot találunk, amelyek nagyon sok szomszédal rendelkeznek, és így sok közösséget kapcsolnak össze. A résztvevők jelentős része ugyanakkor távolabb van a központi magtól és hosszas nyúlványokban kötődik azokhoz. A sűrű közösségi magok és a hosszú nyúlványok jelenlétét az 5. ábrán két valós példán mutatjuk be.

Ebben a fejezetben két önkényesen választott klaszterező algoritmust mutatunk be, a klikk perkolációt (Derényi et al., 2005) és a saját heurisztikáinkat (Kurucz et al. 2008) alkalmazó hierarchikus spektrál módszereit. Megvizsgáljuk a fellelt közösségek az algoritmus paramétereitől való függőségét és a kap-

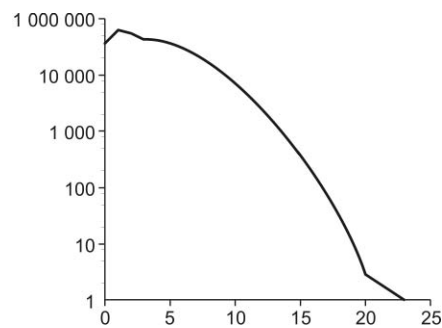


2. ábra • Fent: A hazai települések elhelyezkedése földrajzi szélesség és hosszúság szerint  
Lent: a kiválasztott nyugati országrész



3. ábra • Fent: kapcsolatok száma a két résztvevő (szélesség és hosszúság szerinti) földrajzi távolságának függvényében. Lent: ugyanez a nyugati országrészen (2. ábra). A lineáris illeszkedés érzékeltetésére mindkét grafikonon feltüntettük a darabszámok  $x^1$  és  $x^2$  transzformációját is.





4. ábra • Egymillió véletlen hívó és hívott esetén az útvonalkeresés lépésszámának eloszlása

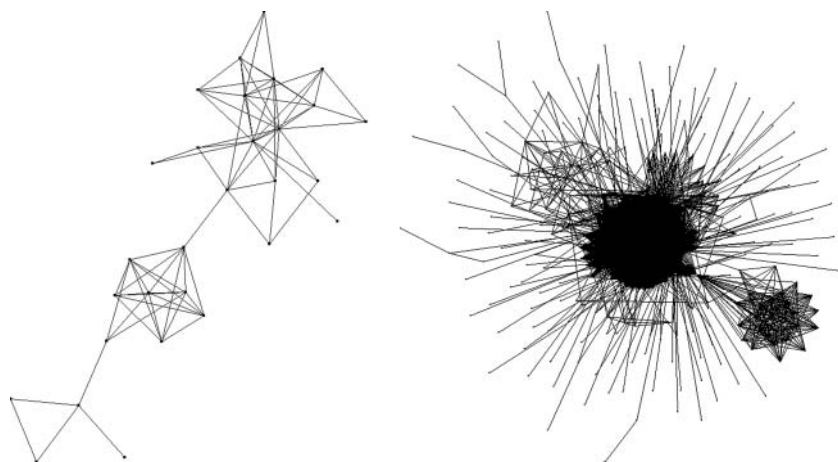
csolatok súlyozásainak lehetőségeit. Megmutatjuk, hogy a sűrű közösségek és hosszú nyúlványok léte mennyire megnehezíti a klaszterezés feladatát, és a spektrál esetben olyan heurisztikákat mutatunk, amelyek megszüntetik ezek torzító hatását.

*Klikk perkoláció: egy friss klaszterező algoritmus*

A klikk perkoláció (Derényi et al., 2005) alapötlete az olyan  $k$  résztvevős csoportok, ún.  $k$ -klikkek vizsgálata, amelyekben minden részt vevő pár között előfordul hívás. Ezek a

magok az elképzelhető legsűrűbben összekapcsolt kis közösségek. A módszer az ilyen  $k$ -klikkekből növeszt lépésről lépésre növekvő közösségeket úgy, hogy minden lépésben valamelyik  $k$ -klikk egyik résztvevőjét megpróbálja egy, a már feltárt közösségen kívüli, de a maradék  $k-1$  résztvevővel újra  $k$ -klikket alkotó elemmel lecserélni. A már tovább nem bővíthető, esetlegesen egymással átfedő közösségek képezik a végeredményt.

A következőkben a fent vázolt módszer a 6. ábra nagy hálózatán, illetve általában olyan nagyméretű hálózatokon való viselkedését mutatjuk be, amelyekben sok sűrű kis közösséget hosszú nyúlványok kötnek össze. A 6. ábra felső táblázatában láthatjuk, hogy a klikkek száma rendkívül magas, és az ezeket kezelő algoritmus megvalósítása így nem egyszerű. A legnagyobb közösség mérete 3-klikkek esetében szintén a klikkek nagy száma miatt óriási. Ugyanakkor a közösségek száma áttekinthetetlenül nagy – ez még 5-klikkek esetén is igaz. Ahogy  $k$  növekszik, ugyanakkor csökken a felhasználható klikkek száma, és egyre kevesebb résztvevőt választunk bele



5. ábra • Egy 82 (bal) és egy 317 (jobb) résztvevős szociális hálózat két-két sűrű maggal és számos rövid nyúlvánnyal

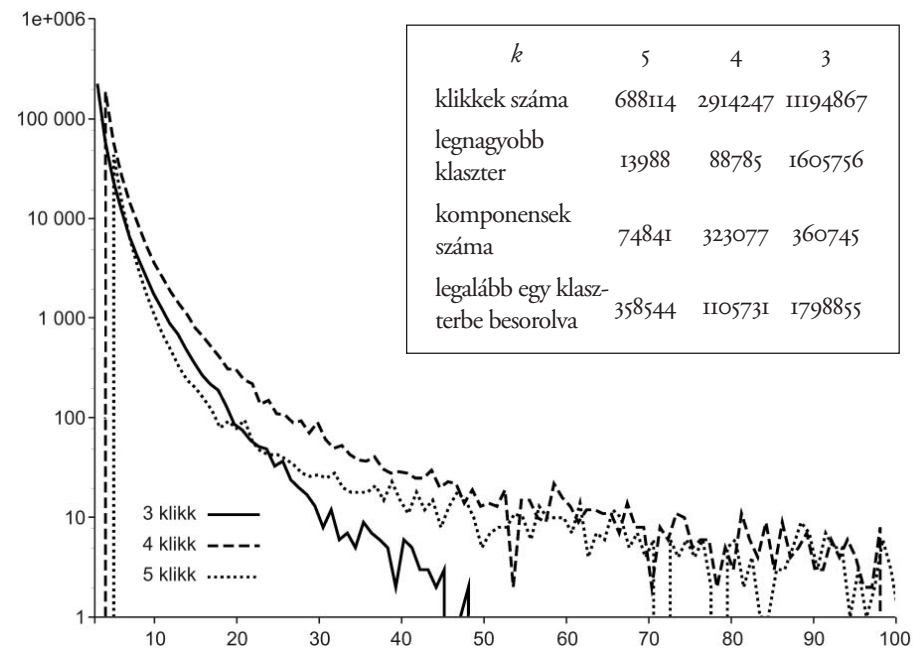
egyáltalán valamelyik közösségbe. A nyúlványokba tartozó, sűrű magoktól távolabb eső elemekkel nem tudunk igazán mit kezdeni. Az ábra alsó részén láthatjuk, hogy a közösségek túlnyomó többsége rendkívül kicsi, azaz nagyon sok a nagyon sűrű kis csoport.

*Spektrál klaszterezés*

A spektrál klaszterezés olyan heurisztikák gyűjtőneve, melyek a hálózat elemeit az első főirányokra vetítik, és a kapott alacsony dimenziós térben sűrűség alapján keresnek csoportokat. A legjobban működő módszerek, amint azt a szakirodalom részletesebb elemzésével és kísérletekkel megmutattuk (Kurucz et al., 2008), egy viszonylag magas, pár tízdimenziós térbe vetíti a hálózat részt-

vevőit, és utána tíz körüli darab csoportot próbál képezni. Siker esetén a kapott csoportokon a módszer megismételhető, így egy klaszterhierarchia áll elő.

A spektrál klaszterezés is érzékeny a sűrű magok és hosszú nyúlványok problémájára. Igen gyakran kimenetként a tagok sűrű magokba, illetve nyúlványokba tartozás szerinti kettéosztását kapjuk. A sűrű magok legalább összefüggő csoportot alkotnak, a másik oldal azonban egymástól izolált apró nyúlványokra bomlik. Ráadásul ezek a darabkák nagyon erősen kötődnek a másik oldal bizonyos részeihez, egy utófeldolgozási lépésben tehát ezeket egyenként áthelyezhetjük a túloldalra, amíg végül visszakapjuk az eredeti teljes adathalmazt mint egyetlen klasztert.



6. ábra • Keretben: A nagy híváshálózatban a  $k = 3, 4$  és  $5$  értékekre vett klikk perkolációs algoritmus néhány tulajdonsága. Grafikon: A klaszterek méretének eloszlása  $k = 3, 4$  és  $5$  értékekre. A vízszintes tengelyen a klaszter elemszáma, a függőlegesen az ennyi elemből álló klaszterek száma látható. Az extra nagy klasztereket levágtuk az ábráról.

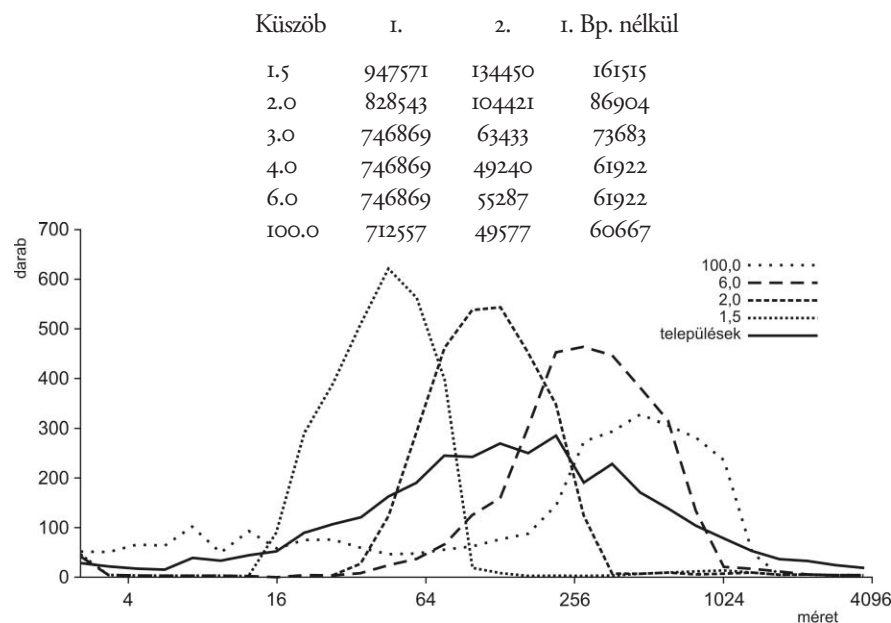
Az általunk javasolt módszer lényege, hogy a vetítési dimenziók számát mindaddig növeljük, amíg a nem összefüggő részek szétosztása után kapott legkisebb és legnagyobb összefüggő klaszter méreteinek aránya egy küszöb alá nem kerül. Nagyon nehezen bontható hálózatok esetében a sűrű részek és a nyúlványok összevonását is alkalmazhatjuk, például Budapest sűrű szociális kapcsolattrendszereit ilyen módon lehet szétbogarozni. Megjegyezzük, hogy a dimenziók számának növelése nagyon megnöveli a számításgigánt, tehát gondos egyensúlyra kell törekednünk a paraméterek megválasztásakor.

A 7. ábrán láthatók a spektrál particionálással kapott eredményeink. Megfigyelhető, hogy a küszöb szigorításával egyre nagyobbá válik a legnagyobb, tovább nem bontható

közösség mérete. Ez a nagy közösség azonban a kapcsolatokkal túl sűrűn ellátott főváros következménye, és Budapestet elhagyva már jól kezelhető méreteket látunk. Az ábra alsó részén a kisebb közösségek darabszámának eloszlását is láthatjuk, amelyre rávetítettük az 1. ábra településméret-eloszlását is. Itt a 2 és 6 közötti küszöbértékek választása a településméretekhez nagyon hasonló közösség-méreteket eredményez.

#### Lemorzsolódás és hálózaton alapuló kategorizáció

Utolsó alkalmazásunkban megmutatjuk, hogyan aknázható ki a hálózati kapcsolatokban rejlő információ a churn, azaz az előfizetés lemondásának előrejelzésében. Kísérletünket a kis adathalmazon folytatjuk olyan



7. ábra • Fent: A küszöb változtatásával kapott legnagyobb és második legnagyobb klaszter, illetve ugyanez a budapesti ügyfelek elhagyása után. Lent: a vízszintes tengelyen a közösség vagy település mérete, a függőlegesen az ekkora közösségek vagy települések száma, különféle küszöb választásával

módon, hogy nyolc hónap tanító időszak után négy hónappal előre, a 12. hónapra adjuk az előrejelzést. Hipotézisünk szerint, amelyet méréseink igazolnak, az egymással kapcsolatban álló előfizetők befolyásolják egymás véleményét, és csoportokban együtt jutnak elhatározásra, hogy előfizetésüket lemondják.

A következőkben bemutatott módszer az ún. félig felügyelt osztályozás (Zhu, 2005) csoportjába tartozik, amelyben az ismeretlen, felcímkézetlen egyedek tulajdonságait is hasznosítjuk a modellépítés során. Ebben a csoportban található az általunk alkalmazott *hálózati elemek osztályozása*, amely iteratív tanulási folyamat. Kezdőlépésként a meglévő forgalmi, előfizetési és szociodemográfiai adatok alapján meghatározzuk minden résztvevő churn értékét. Ezután megnézzük minden előfizető esetén a vele kapcsolatban álló többi előfizető előrejelzett értékét és ezekből újabb jellemzőket gyártunk, amelyekkel kibővíve, akár több iterációban folytatjuk a modellépítést. Többféle jellemzőt állítunk elő, amely a szomszédok churn előrejelzésének  $c$  átlagán, illetve kapcsolati erősségekkel súlyozott  $c$  átlagán alapulnak. Ezek közül a  $c$  olyan változata bizonyul legjobbnak, amely a kevés kapcsolattal rendelkező csúcson „regularizációt” alkalmaz, azaz a semleges átlag felé tolja el a jellemző értékét.

A tulajdonságok hálózati kapcsolatokon keresztül történő továbbítása a churn előrejelzés mellett számos további alkalmazással bír,

amelyek között megtalálható az internet dokumentumainak osztályozása és a bizalom vagy bizalmatlanság megerősítése. Talán a legfontosabb, a módszert egyik elsőként alkalmazó terület a web spamszűrés, azaz a kizárólag a keresőrendszerek találati rangsorának manipulálására létrehozott, értelmetlen vagy kifejezetten megtévesztő tartalmak fellelése (Benczúr, 2008). A web spam esetében a továbbítás például azon alapul, hogy becsületes tartalom elvétele hivatkozik spamre, míg a spam oldalakra azok egymást erősítő szándékai következtében mindig nagyon sok spam hivatkozik.

A klasszifikációt kísérletünkben a Weka ún. *cost sensitive* C4.5 módszerével végeztük a különböző típusú hívások (helyi, távolsági, csúcsidő, alternatív szolgáltatón keresztül) összerétekeivel mint jellemzőkkel. Az alap klasszifikátorhoz képest egy lépésben 25 % javulást eredményezett a hálózati információk kihasználása, a második iteráció azonban már rontott a minőségen. Érdekességként megjegyezzük, hogy a  $c$  érték számításakor nem feltétlenül a hívások darabszámát vagy idejét érdemes használni, hanem két ügyfél kapcsolati erősségét inkább a közös hívottak számával, vagy két hívott csoport által meghatározott vektorok által bezárt szöggel érdemes súlyozni.

Kulcsszavak: *szociális hálózatok, klaszteranalízis*

#### IRODALOM

- Albert Réka – Jeon, H. – Barabási A-L. (1999): Diameter of the World Wide Web. *Nature*. 401, 130–131.  
 Barabási Albert-László (2002): *Linked*. Perseus Publishing  
 Barabási Albert-László – Albert, R. – Jeong, H. (2000): Scalefree Characteristics of Random Networks: The

Topology of the Word-Wide Web. *Physica A*. 281, 69–77.

- Benczúr András A. – Siklósi D. – Szabó J. – Bíró I. – Fekete Z. – Kurucz M. – Pereszlényi A. – Rác S. – Szabó A. (2008): Web Spam: A Survey with Vision for the Archivist. In: *Proceedings of the International Web Archiving Workshop*. <http://iwaw.net/08/IWAW2008-Benczur.pdf>

Derényi Imre – Palla G. – Vicsek T. (2005): Clique Percolation in Random Networks. *Physical Review Letters*. 94, 49–60. [http://arxiv.org/PS\\_cache/cond-mat/pdf/0504/0504551v1.pdf](http://arxiv.org/PS_cache/cond-mat/pdf/0504/0504551v1.pdf)

Kleinberg, Jon (2000): Navigation in a Small World. *Nature*. 406, 845.

Kurucz M. – Siklósi D. – Lukács L. – Benczúr A. A. – Csalogány K. – Lukács A. (2008): Telephone Call Network Data Mining: A Survey with Experiments. In: *Handbook of Large-Scale Random Networks*. Springer Verlag in conjunction with the Bolyai Mathematical Society of Budapest

Milgram, Stanley (1967): The Small World Problem. *Psychology Today*. 2, 1, 60–67.

Watts, Duncan J. – Strogatz Steven H. (1998): Collective Dynamics of Small-World Networks. *Nature*. 393, 6684, 440–442.

Zhu, Xiaojin (2005): *Semi-supervised Learning Literature Survey*. Technical Report. 1530, Computer Sciences, University of Wisconsin-Madison [http://pages.cs.wisc.edu/~jerryzhu/pub/ssl\\_survey.pdf](http://pages.cs.wisc.edu/~jerryzhu/pub/ssl_survey.pdf)

WebSpamChallenge: [\texttt{http://webspam.lip6.fr/}](http://webspam.lip6.fr/)

Weka nyílt forráskódú gépi tanulási eszköz: [\texttt{http://www.cs.waikato.ac.nz/ml/weka/}](http://www.cs.waikato.ac.nz/ml/weka/)



## Tanulmány

# SZÁZ ÉVE SZÜLETETT GOMBÁS PÁL

Szépfalusy Péter

az MTA rendes tagja, professor emeritus  
ELTE TTK Fizikai Intézet

A 20. század első évtizedének végén, száz éve született Gombás Pál, Selegsántón. Sopronban érettségizett, majd 1932-ben matematika–fizika szakos tanári oklevelet szerzett a budapesti tudományegyetemen. 1934-ben lett bölcsészdoktor. Pályáját az *Ortvay Rudolf* által vezetett elméleti fizikai tanszéken kezdte gyakornokként. Életútjának további állomásai: 1939-től a szegedi egyetem, 1940-től a kolozsvári egyetem rendkívüli tanára, 1941-től ugyanott nyilvános rendes tanár, majd 1944-től a budapesti műegyetemen a fizika tanszéket vezette 1971-ig, haláláig.

A centenáriumi emlékezés Gombás Pál tudományos munkájára koncentrálni kívánja érzékeltetni, hogy ő mindmáig az egyik legeredményesebb, Magyarországon tevékenykedő elméleti fizikus kutatónak számít.

Egyetemi hallgatóként élte át a kvantummechanika kiteljesedését. Meghatározó volt számára, hogy Ortvay a modern fizika szellemét hozta az egyetemre (az már a véletlen játéka volt, hogy Ortvay tizenkilenc évi kolozsvári, majd szegedi egyetemi tevékenység után abban az évben lett a budapesti egyetem elméleti fizikai intézetének igazgatója, amikor Gombás egyetemi tanulmányait elkezdte).

Gombás érdeklődése az új elmélet alkalmazásai felé fordult. Felismerte, hogy ebben a folyamatban a többrészes rendszerek különleges helyet foglalnak el. A mérésekkel összehasonlítható számszerű eredmények elérése a fejlődés számára alapvetően fontos feladat volt, és ebbe a nemzetközi erőfeszítésbe Gombás nagy intenzitással kapcsolódott be.

Első cikke 1933-ban jelent meg az atomok diamágneses szuszceptibilitásáról, az akkor egyik vezető folyóiratnak számító *Zeitschrift für Physik*-ben. A következő három évben ugyanebben a folyóiratban további tizenegy(!) cikket publikált különböző témakörökben (kettőt *Neugebauer Tiborral* közösen). Kutatómunkáját évekre meghatározó legfontosabb terület a fémek statisztikus elmélete volt.

Fém-modelljét a *Nature* című folyóiratban is bemutatta 1936-ban. A modell választ tudott adni arra a nagy kihívásra, hogy miért tekinthető viszonylag kicsinek a fémelektronok és az ionok közötti effektív kölcsönhatás. Nevezetesen, Gombás modelljének sarkalatos pontja annak a pozitív energiának a bevezetése volt, amely a fémelektronoknak az iontörzsekbe való behatolásából származik, és eredete a Pauli-féle betöltési tilalomra ve-