

BIG DATA A TERÜLETI KUTATÁSOKBAN:

avagy néhány példa a nagy adatbázisokra épülő térbeli elemzési lehetőségek köréből

Jakobi Ákos – ELTE TTK FFI Regionális Tudományi Tanszék

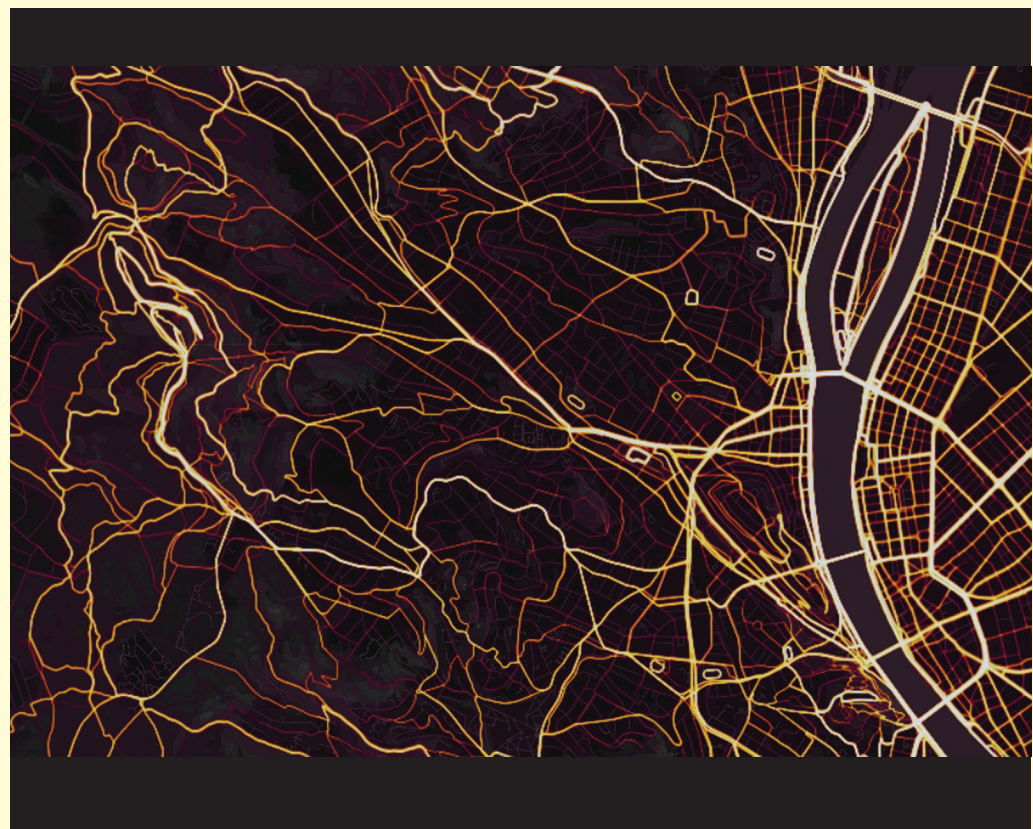


Geolocalizált Twitter üzenetek (kék) és Flickr képfeltöltések (piros) sűrűsége, valamint ezek együttes előfordulása (fehér) Londonban (forrás: Eric Fischer 2011)

VÁSÁRLÁSI ADATOK TÉRBELI MINTÁZATAI

A példa egy magyarországi bank ügyfeleinek POS- (Point-of-Sale) terminálokon feljegyzett vásárlásainak aktivitási adatai alapján készült. A terminálok az adatrögzítő eszközök, míg az ügyfelek azok, akiknek a viselkedéséről információ szerezhető. A bankok üzleti érdekeik miatt és saját tevékenységükből adódóan folyamatosan rögzítik ügyfeleik tranzakcióit. Leszámítva az online, a fióki vagy az ATM tranzakciókat a banki ügyfelek legtöbbször a POS-terminálokon keresztül (a helyszíni bankkártyás fizetés során) hagynak digitális nyomokat maguk után. Ezek a digitális nyomok nemcsak azt jegyzik fel, hogy az egyes bankkártya-tulajdonosok mikor és

milyen összegben vásároltak, de a POS-terminálok lokalizációja révén azt is, hogy mindez hol történt. Az adatokat aggregáltan térképezve látható, hogy a bank ügyfélköre általában hol költi el a pénzét. Ha ez az ügyfélkör, avagy a vizsgált minta kellően nagy, akkor a bolti kereskedelmi aktivitás társadalmi szintjeibe is bepillantást nyerhetünk. Az ábra a banki ügyfelek bolti bankkártya-használati aktivitásának térbeli eloszlásviszonyait mutatja Budapest belső területén. Jól láthatók a főváros frekvenciált kereskedelmi zónái, amelyek főként a Nagykörút, illetve a Kiskörút vonalaihoz igazodnak. Sűrűsödés jellemző a pesti belváros környékén, valamint egyes sugárirányú utak (Rákóczi út) vonalában. Budán a vonalas szerkezetek kevésbé jellemzőek, inkább lokális kereskedel-



A szabadidős sporttevékenységek (futás, kerékpározás) városi térpályái a főváros budai oldalán és a belvárosban (részlet) (forrás: www.strava.com, 2019)(világosabb színek = gyakrabban használt útszakaszok)

ONLINE KÖZÖSSÉGI HÁLÓK FÖLDRAJZI SAJÁTÓSÁGAI

Az online közösségi hálók (online social networks: OSN) lényegében virtuális kreálmányok, melyek a való élet baráti kapcsolatainak virtuális térbeli kifejezésére szolgálnak. Ezek a közösségi hálók egyidejűleg tükrözik vissza a kibertér és az offline földrajz társadalmi-területi jelenségeit. A közösségi oldalak arra is használhatók, hogy új személyeket ismerjünk meg, azonban döntő többségben azokkal az emberekkel való kapcsolatainkat dokumentáljuk rajtuk, akikkel az eddigi offline világ hálózataiban kapcsolatba kerültünk. Ezek a kapcsolati hálók viszont, úgy tűnik, térbeli megkötöttségeket is mutatnak. Az iWiW („International Who is Who”) közösségi oldal 2002 és 2014 között működött Magyarországon. A mellékelt ábra ezen egykori közösségi hálózat szerkezetét mutatja: a hálózatban a csúcsok a felhasználókat, míg az élek a köztük lévő baráti kapcsolatokat reprezentálják. A teljes hálózat 4 millió felhasználója között 785 millió él volt definiálható. A felhasználókat települési szinten aggregálva végül a településközi kapcsolati hálózat is felrajzolhatóvá vált, aminek eredményeként 1,3 millió élbe lehetett tömöríteni

a kapcsolati információkat. Az eredmények vizuális megjelenítésekor ugyanakkor nem tanácsos mind az 1,3 millió település-település kapcsolatot egyidejűleg felrajzolni, a mellékelt ábrán ezért csak a legfontosabb összeköttetések szerepelnek: az ábra azokat a hálózati éleket mutatja, melyek legalább 15000 személyes kapcsolatot képviselnek. A térkép, nem meglepő módon a legnépesebb településeink között mutatja a legsűrűbb összeköttetéseket (a legtöbb kapcsolat, szám szerint 1,8 millió, a budapesti és a debreceni felhasználók között volt megfigyelhető). A településközi abszolút kapcsolatszám egyértelműen a településhierarchia mentén rendeződik Magyarországon, így a térképen hub-ként jelennek meg a regionális központok, s természetesen a legjelentősebb hubnak Budapest adódott. Fontos tudnunk, hogy az ábra nem ad tájékoztatást a kapcsolatok szorosságáról, azaz nem tükrözi, hogy az egyes települések kapcsolati szerkezetében melyek a fontos partnerek, az erős szálakkal összekötött esetek, és melyek a gyenge kapcsolatok. Erre más, súlyozott modellek adhatnak választ.

A legsűrűbb településközi kapcsolatokat az iWiW közösségi hálóban (2013; csak a 15.000 db feletti kapcsolatszámú összeköttetéseket jelölve)

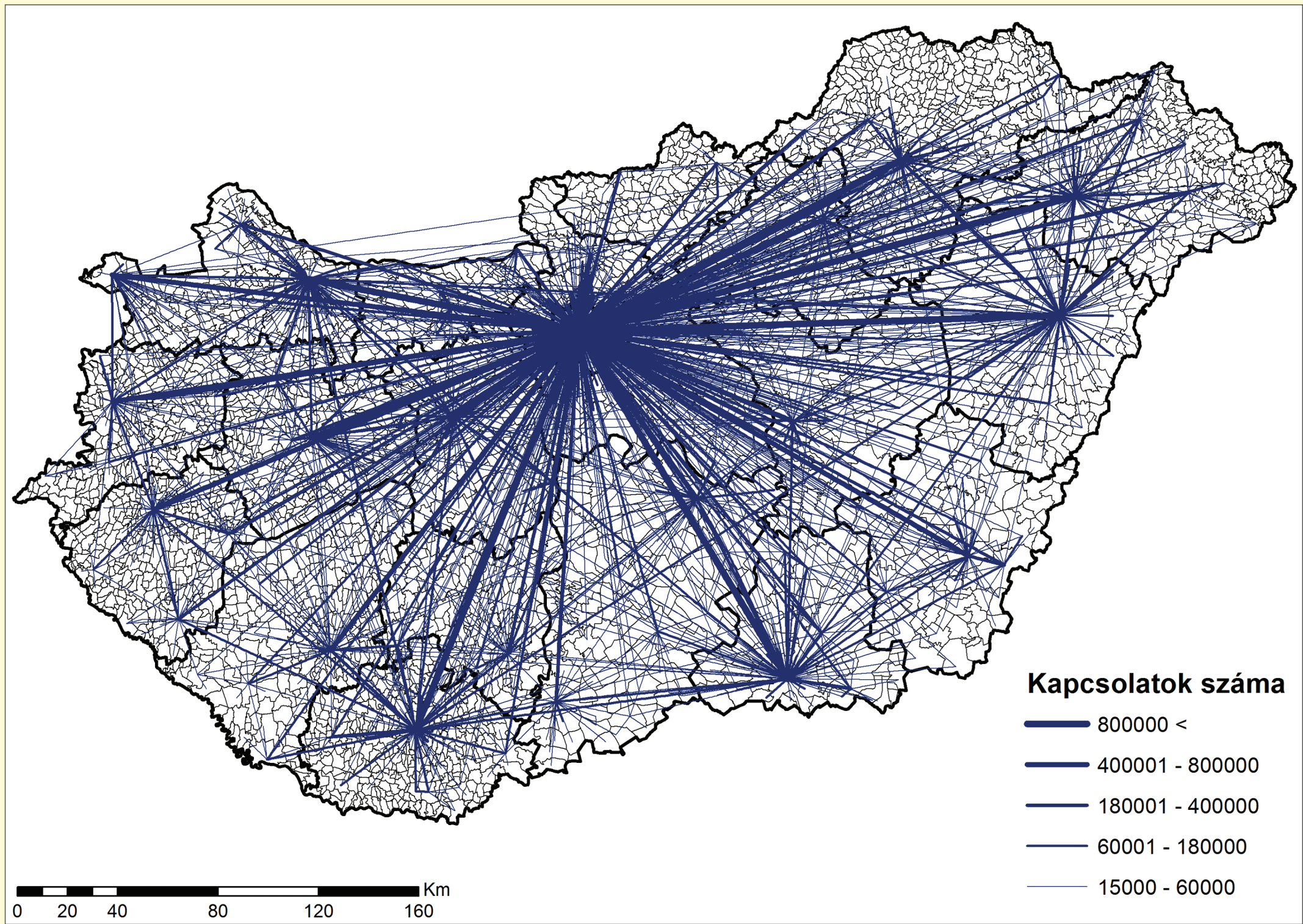
ADATROBBANÁS A TÉRELEMZÉSEKBEN

Szakmai körökben a Big Data (szabad fordításban „nagy adathalmaz”) néven ismert fogalom arra a hatalmas adatmennyiségre utal, amely információs világunkban gyorsan és folyamatosan keletkezik, s melynek feldolgozása a hagyományos kapacitásokkal és módszerekkel szinte megoldhatatlan kihívást jelent. A Big Data ennek ellenére mégis nagy lehetőségeket kínál. A sokáig csak virtuális melléktermékként számon tartott napi információhalmaz akkor válik értékessé, amikor a különböző adatokat összekötjük, összefüggéseket, felismerhető mintázatokat találunk közöttük, s ebből értékelhető következtetéseket vonunk le. A különféle helyeken (kormányzati szerveknél, internetes és telekommunikációs cégeknél, vállalatoknál, stb.) összegyűlő hatalmas adatmennyiség a társadalom és a gazdaság folyamatainak vizsgálata számára valóságos aranybánya. Az adatok az emberi viselkedés egyedi és csoportos (társadalmi) szintjeibe is betekintést nyújtanak. Az egyre szélesebb körben terjedő, térbeli információkat

mi góccok láthatók például a Széna tér, a dél-budai Móricz Zsigmond körtér és Újbuda-Központ környezetében. E gócterületek utólag is jól beazonosítható a bevásárlóközpontokat (Mammut, Allée, Westend, Aréna Pláza, Corvin Pláza stb.), amelyek a kiskereskedelmi fogyasztás urbánus centrumai. A modell megerősíti a városi kereskedelmi térszerkezettel kapcsolatos eddigi ismereteinket, de ki is egészíti őket azzal, hogy bizonyítékot szolgáltat a valós kereskedelmi térhasználat alakulásáról. Ilyen léptékű lokalizált kérdőíves vagy statisztikai felmérés más módon amúgy nem lenne elérhető.

POS-terminálokon (bankkártya leolvásokon) jegyzett vásárlások sűrűsége térképe egy magyarországi bank ügyfeleinek adatai alapján (Budapest belső területein, részlet)(vörös = magasabb értékek; zöld = alacsonyabb értékek)

rögzített tér-idő adatok különösen akkor válnak hasznossá, amikor nagyon sok felhasználó kiemelkedően nagy számú hasonló információját sikerül egyszerre, áttekinthető módon értékelni. A Big Data rendszerek éppen ezt kínálják. A sokfelhasználós közegekben ugyanis nemcsak az egyedi térpályák, hanem a jellemző útvonalak, térbeli csoportosulások meghatározására is lehetőség nyílik. A mellékelt ábra az egyre divatosabbá váló sport és szabadidős applikációk egyike alapján készült. Az adatbázis alapját képező okoseszközökre (leginkább okostelefonokra) telepíthető alkalmazás az alapkészülékbe épített helymeghatározó rendszer segítségével megállapított helyadatokat gyűjti össze a felhasználó számára a szabadidős sporttevékenység (futás, kerékpározás) közben. A sportteljesítmény monitorozása mellett az eszköz rögzíti a sportoló által bejárt útvonalat is. Ha nagyszámú felhasználó rögzíti az adatait, akkor a bejárt útvonalak egyidejű térképi ábrázolásával lehetővé válik a sport és szabadidős tevékenységek városi térpályáinak és frekvenciált helyeinek meghatározása is. Az ábra (háttértérkép hiányában is) felismerhetően kirajzolja a fővárosi utak hálózatát, illetve azokat az utakat is, amelyek csak a szabadidős sporttevékenység (futás, kerékpározás) szempontjából fontosak. Ilyen útszakaszokat főként a budai hegyvidéki részekben találhatunk. A belvárosi és a Dunához közeli területek frekvenciáltabb útszakaszokkal jellemezhetők, főleg ott, ahol kerékpársávok és sportolási területek is találhatók. A térképen egy-egy sportpálya futóköre is beazonosítható.



Kapcsolatok száma

- 800000 <
- 400001 - 800000
- 180001 - 400000
- 60001 - 180000
- 15000 - 60000