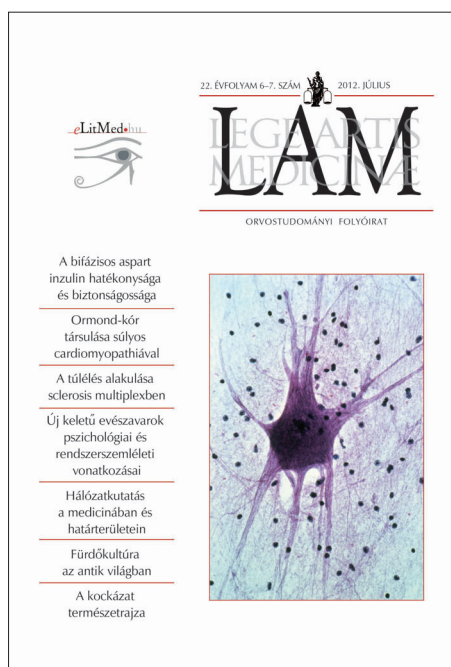




Title:

Network research in medicine and at its interdisciplinary boundaries

Authors:

Zoltan BRYs, Bela BUDA, Andras PLUHAR

Corresponding author:

Andras PLUHAR, University of Szeged, Institute of Informatics
H-6720 Szeged, Arpad ter 2. Hungary - EU
pluhar@inf.u-szeged.hu

Summary: Network analysis is an increasingly used research method in medicine. Network analysis is a research and interpretation model at the same time. Our article outlines the various applications of network analysis in medicine and at its interdisciplinary boundaries. General introduction to the notions of small world graphs, cluster analysis, protein-protein interaction network and clinical opinion leaders are presented along with a overview of the field.

Keywords: network research, network analysis, medicine, small world, graphs, cluster analysis, system biology, network medicine

Language: Hungarian

Journal: Lege Artis Medicinae (peer-reviewed, indexed in Scopus, and selectively indexed in Pubmed, NLM ID:9425319)

Citation: Lege Artis Med. 2012 Jul;22(6-7):4455-449

"Not for reproduction, distribution or commercial use".

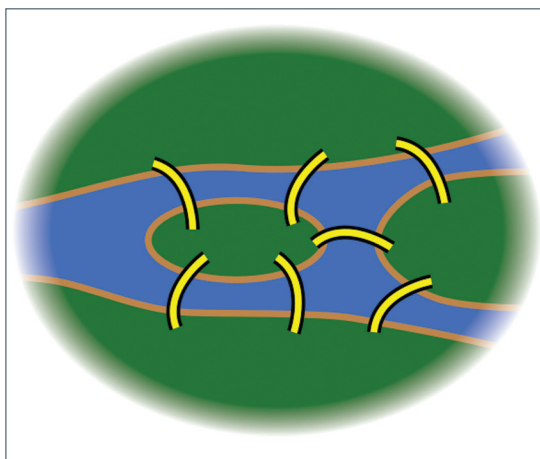
Hálózatkutatás a medicinában és határterületein

Brys Zoltán, Buda Béla, Pluhár András



Az elmúlt években robbanásszerűen megnőtt a hálózatelemzési módszereket használó tudományos közlemények száma a medicinában, és várhatóan a közeljövőben tovább nő. A módszer tekintetében megosztott az értelmiségi közvélemény – részben jogosan –, hiszen a közlemények színvonala rendkívül szórta. Időnként túlzott elvárásokat, máskor pedig jelentős alulértékelést hallani a tudományos közéletben. A terület tudományos fejlődése és tudományos marketingje is intenzív. Fogalmi és értelmezési zavarok is tetten érhetők az új diszciplínában. Néhány magas impaktfaktorú orvostudományi folyóiratban megjelent, hálózatelemzést alkalmazó közlemény (1–3) jelentős szakmai vitát váltott ki, ezzel is rávilágítva a hálózatkutatási módszertan fiatalságára és értelmezési nehézségeire.

A tudományos módszer alapvetően redukcionista, a különböző entitások között kölcsönhatásokat állapít meg, és a vizsgálatok során igyekszik az egyéb zavaró hatások kiküszöbölésére. Ennek a hozzáállásnak megvannak a korlátai, ezt kívánja részben orvosolni a hálózatelemzés. Természetesen az egyszerűsítés itt sem kerülhető el, a megfigyelhető, rendkívül bonyolult és szerteágazó jelenségeket csomópontokra és élekre egyszerűsíti azért, hogy új, eddig nem ismert rejtett összefüggéseket tudjon megragadni, felfedezni, esetleg eseményeket előre jelezni. Az egyszerűsítés miatt sokszor kérdéses, hogy a megragadott törvényszerűség valós vagy az egyszerűsítés módszertanából eredeztethető. A rendkívül különböző színvonalú közlemények ismeretelméleti értelmezése tehát sokszor nem triviális feladat – még szakembereknek sem. (Az adatbányászatban „adatszivárgásnak” nevezett jelenség sok tekintetben hasonlóságot mutat azzal, amikor a vélt topológiai szabályszerűség módszertani okokból következik, nem pedig a jelenség valódi természetéből.)

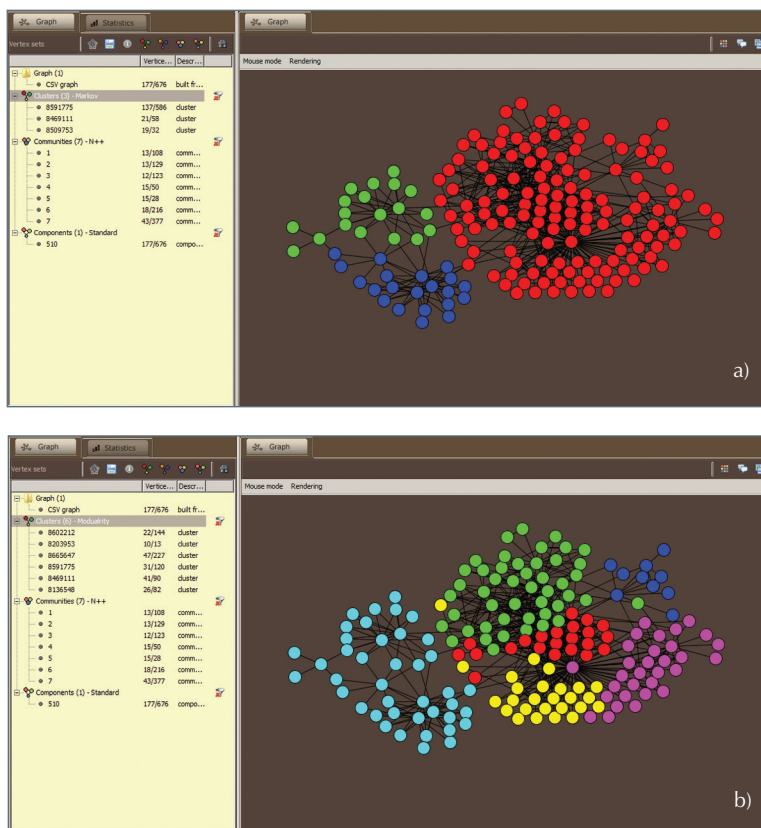


1. ábra. A königsbergi hidak (Forrás: Wikipedia)

A hálózatelemzés kutatási és értelmezési modell is egyben (4), alapja a gráfelmélet, amely a kombinatorika önálló területe. Az első gráfelméleti közleményt 1741-ben Euler publikálta (5) (a königsbergi hidak problémája). Königsberg városában hét híd ívelt át a várost átszelő folyón úgy, hogy ezek a folyó két szigetét is érintették. A legenda szerint a helyiek azzal a kérdéssel fordultak Eulerhez, vajon végig lehet-e menni az összes hídon úgy, hogy mindegyiken csak egyszer haladjanak át, és egyúttal visszaérjenek a kiindulópontba (1. ábra)?

Euler a hidakat éleknek és a földrészeket csomópontoknak tekintette, és az így definiált gráf fokszámaira vezette vissza a problémát. (Egy v pont $d(v)$ fokszáma a v pontba összefutó élek száma, a fokszámsorozat pedig a fokszámok növekvő sorba rendezése $d(1) \leq \dots \leq d(n)$.) Euler elemi érveléssel bebizonyította, hogy a feladat nem megoldható, illetve általánosítva a problémát, szükséges és elegendő feltételt adott a később Euler-körnek nevezett objektum létezésére.

A terület tudományos fejlődése és marketingje is intenzív.



2. ábra. Fiktív közösségi hálózatban (social network) nem átfedő közösségek (klaszterek) detektálása és vizualizációja kétféle módszerrel. Elkészítve a Sixstep program használatával. Teljes adathalmaz: 177 csomópont (személy) és 676 kapcsolat (él)

Újszerű és vitatott gondolat az elhízás „járvány-szerű” terjedése a szociális hálózatokban.

Jelenleg leginkább leíró, modellező szándékkal és eredménnyel alkalmazható a hálózatelemzés, de már mutatkoznak a prediktív jellegű megközelítések jelei is. A klasszikus gráfelmélet kialakítását számos magyar kutató segítette, a teljesség igénye nélkül említhető König Dénes, Egerváry Jenő, Erdős Pál, Turán Pál, Gallai Tibor, Rényi Alfréd, Lovász László és Szemerédi Endre.

Kisvilág gráfok

A hálózati elemzések gráfjai többnyire egy adathalmazból, valamely szabály szerint származtattak, adott pontok közt többfajta él is elképzelhető. (Karinthy Frigyes a kézfogásokat, Milgram a levélátadást, Zachary az önbevalláson alapuló barátságot tekintette annak. Lehet ez a kapcsolat www-link, telefonhívás, pénzügyi tranzakció, betegségátadás, közös DNS-szekvencia stb.) Az igazán meglepő tény, hogy a teljesen különböző jelenségekhez rendelt gráfok-

nak vannak közös vonásaik, s ezek alapján *kisvilág gráfoknak* (Small World Graphs) hívják őket, bár pontos definícióban még nem állapodott meg a kombinatorika. A kisvilág gráfoknak – az alkalmazások szempontjából is fontos – tapasztalati tulajdonságai az alábbiak:

1. A gráfnak viszonylag kevés éle van az összes lehetségeshez képest.
2. Jelentős sűrűsödések találhatók bennük.
3. Kicsik a bennük levő átlagos távolságok.
4. A fokszámok eloszlása a hatványtörvénnyel írható le.

Az orvosi gyakorlatban a standard normális eloszlás a leggyakoribb (például vérnyomás, hormonszintek, jellemző laboratóriumi értékek stb.), itt a nullától x -szel való eltérés esélye gyorsan csökken $\exp(-x^2/2)$. A hatványtörvényt követő eloszlásban bármely k érték elég nagy eséllyel előfordul, ck^a , ahol c pozitív, a negatív értékű állandó.

A kisvilág gráfok felfedezésének egyik hatása a klasszikus járványterjedés újragondolása. Az 1. és 2. pont szerint hiába tűnik korlátozottnak egy vírus elterjedése, a 3. miatt valójában gyorsan elérhet bárkit. A 4. pont arra utal, hogy vannak az átlagostól nagyon eltérő, óriási fokszámú (sok kapcsolattal rendelkező) egyedek, és jobbra ők felelősek a járvány terjedéséért. Tehát elsősorban ezeknek az egyedeknek az immunizációja fékezheti meg az adott járványt.

Klaszteranalízis

Amíg a *komponens*, *korrelációk*, a *csomósodás* és a *köztiség* (az értéke egy v pontra arányos azzal, hányszor fordul elő v egy legrövidebb út részeként, azaz mennyire „megkerülhetetlen” a pont) számításának módjai és algoritmusai nagyrészt tisztázódtak, addig a klaszteranalízis, azaz a hálózatban található csomópontok homogén csoportokba történő sorolása továbbra is népszerű kérdés a hálózatkutatásban (2. ábra). A különböző távolságok (*Euklideszi távolság*, *abszolút eltérés*, *Mahalanobis-távolság*, *Hamming-féle távolság*, *Pearson-távolság* stb.) vagy hasonlóságmértékek alapján végzett csoportosítás lehet hierarchikus vagy nem hierarchikus.

Átfedő *klaszterek* (vagy *közösségek*) – amelyekben egy csomópont több csoportba is tartozhat – képzésében pedig továbbra is sok nyitott kérdés ismert (3. ábra). A bioinformatikában gének és fehérjék vonatkozásában is számos klaszterezési eljárást használnak. Irányított és súlyozott gráfoknál viszont számos további tényező nehezíti a klaszterek meghatározását, itt továbbra is sok a megválaszolatlan kérdés.

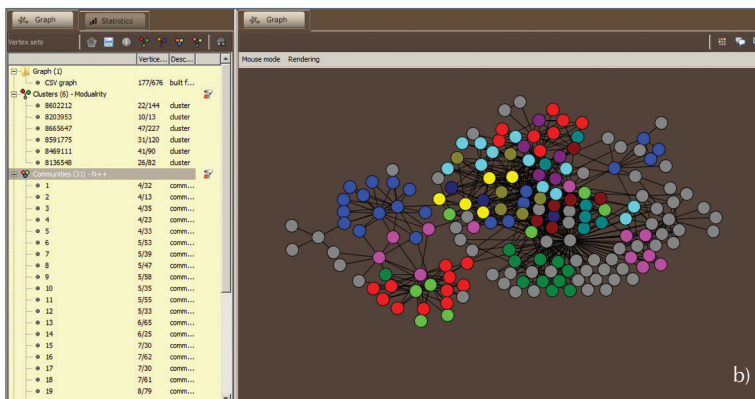
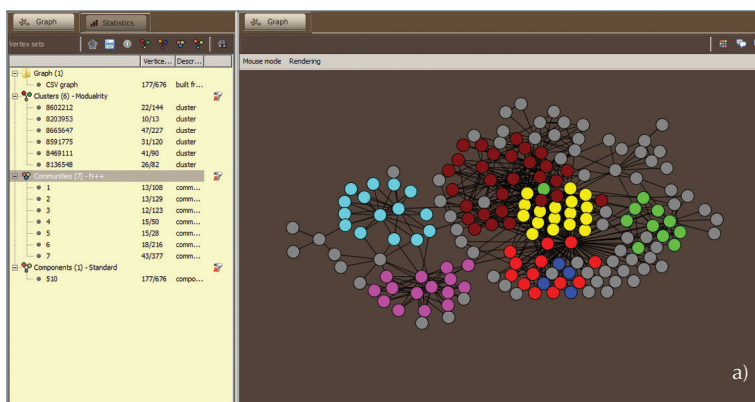
Hálózatelemzés a tudományban

Szinte minden tudományterületen található hálózatelemzéssel vizsgált vagy azzal értelmezni próbált jelenség. Az ökológiában a tápláléklánc topológiai jellegzetességei (6), a biokémiában a fehérje-fehérje hálózatok (7) a nyelvészetben a szóasszociációs hálózatok (8), a szociológiában és szociálpszichológiába pedig az emberi kapcsolatok hálózata, a tudományometriában a citációs hálózatok (9), mérnöki tudományokban az elektromos hálózatok (10), az úthálózatok stb. elemzésére használják a hálózatelemzés adta lehetőségeket.

A társadalomtudományokban kezdettől fogva érintett kérdés a hálózatos hatások vizsgálata, elsősorban a vélemények és beállítódások, és főleg a viselkedés társadalmi méretű változásának elemzésekor. A civilizációval együtt járó gyors társadalmi átalakulás szembeütnővé vált és magyarázatot igényelt. Gabriel Tarde (11) az utánzást feltételezte az érték- és mintaváltozás mechanizmusának. Tarde szerint a gazdag és hatalmas réteg alakítja ki az életvezetési és fogyasztási értékeket, és ezt veszik át, a rétegződésen lefelé haladva az emberek. A sajtó és a horizontális mobilitás (utazás, migráció) a közvetítő. A rádió és a televízió társadalmi hatásának vizsgálata vetette fel az opinion leader (12) fontosságát, az a műsor hat, amely az egyént körülvevő kapcsolatszerkezet adott témában öntudatlanul elfogadott tekintélyszemélyének elnyeri a jóváhagyását. Általában minimális jelzésekkel, észrevétlenül, tudattalanul zajlik. Úgy tűnik, hogy a divatban és a tudományos mintadiffúzióban, hírdiffúzióban is hasonló folyamat megy végbe. A második világháború után kutatták a telefon szerepét, ez is mutatta az egyén körüli hálózatot. A szociometria megjelenésével (13) és később a gyenge kapcsolatok szerepéről megalkotott elképzelés (14) mind implikálják a hálózatos hatásokat, ugyanakkor ezeket a felülről lefelé vonatkoztatás eszközével próbálta megragadni. A hálózatelemzés megnyithatja a lentől felfelé történő vonatkoztatás lehetőségét, amely más eszközöket és szemléletmódot is kíván.

Hálózatelemzés a medicinában

Az orvostudományon belül számos orvosi szakterület vonta be a kutatási eszköztárába a hálózatelemzést. Főbb területek: a neurológiában az agyműködés vizsgálatára (15), genetikában a betegségek asszociációs analízisére és az RNS- és mRNS-szintű szabályozás hálózatos működése, valamint a génhálózatok elemzésére, az epidemi-



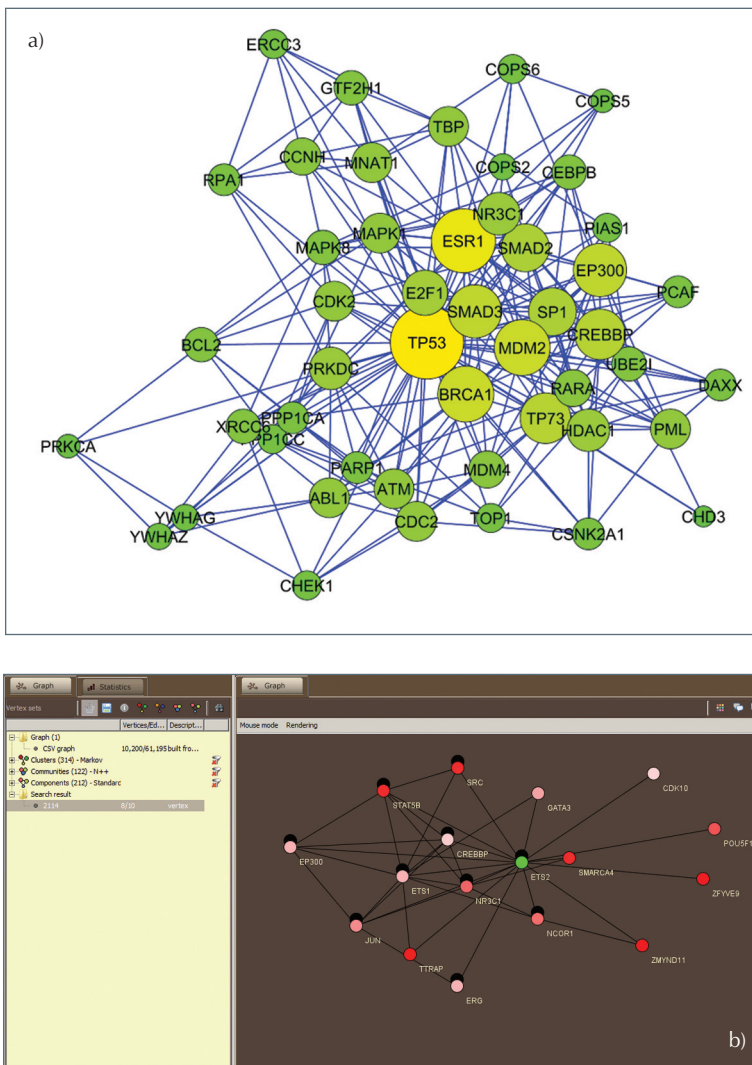
3. ábra. Fiktív közösségi hálózatban (social network) átfedő közösségek detektálása vizualizációja kétféle parametrizálással. Elkészítve a Sixstep program használatával. Teljes adathalmaz: 177 csomópont (személy) és 676 kapcsolat (él).

ológiában a járványterjedések elemzésére, különösen a szexuális úton terjedő betegségek elemzésére (16).

Számos orvosi szakterületet érintő kutatás a metabolikus hálózatok elemzése (17), valamint a sejttanyagszere hálózatos elemzése. Onkológiában a rák természetének megértésében (különös tekintettel azokra a hálózatokra, amelyben a TP53 gén érintett) és a metasztázisok valószínűségének kutatására (18) használják. Újszerű és vitatott gondolat az elhízás „járványszerű” terjedése a szociális hálózatokban (2). A gyógyszerkutatásokban pedig új molekulacsoportok lokalizálására használható a hálózatelemzés. Kulcsfogalom a Network Medicine (19), amely a sejtben belüli és kívüli hálózatok (az erős funkcionális és folyamatbeli összekötöttségre épített) rendszer-szemléletű megközelítése és kutatási módszertana (4. ábra).

A hálózatelemzés ikonikus figurája, Barabási Albert-László véleménye szerint a hálózatelemzés jelentős lendületet adhat a medicinának. Hatékonyabb vakcinációs stratégiák mellett különösen a személyre (genetikai) és a személy adott betegségére (differenciáldiagnosztikai szem-

Az informal educators/mentors kórházon belüli lokalizálása fontos tényező lehet a klinikum minőségbiztosítási folyamataiban.



4. ábra. a) Fent a TP53 gén közvetlen (egylépéses) környezetében található, kizárólag magas fokszámú (>10) csomópontok hálózatos vizualizációja. Elkészítve a Cytoscape (21) program használatával. Szemléletes adat, hogy az adott adathalmazban a TP53-tól két lépésben már 3497, legalább egy éllel rendelkező csomópont és 38 019 él található! b) ETS2 gén közvetlen (egylépéses) környezetében található 16 csomópont hálózatos vizualizációja. Elkészítve a Sixtep program használatával.

Mindkét esetben a használt egyesített adathalmaz: A merged human interactome (Unilever, 2007). A használt adathalmaz tartalmazza a IntAct, DIP, BIND és HPRD, Rual (PMID: 16189514) és Stelzl (PMID: 16169070), valamint Marcotte (PMID: 15892868) adatait is.

pontból) szabott medikamentumok (ideértve a DNS- és fehérjehipecet is) irányába (20).

Ugyanakkor a biokémiai hálózatok ismert kutatója, Csermely Péter több előadásában arra is felhívta a figyelmet, hogy a legtöbb élettani folyamat sematikus rendszerszemléleti modellje folyamatvezéreltnek (kondicionálnak) tűnik. Azaz némileg leegyszerűsítve: „C” aktor akkor és csak akkor aktiválódik „D” módon, ha „A” aktor kapcsolatba kerül „B” aktoral. A jelenlegi hálózatelemzés még csak ritkán használ kondicionált kapcsolatokat, azaz olyanokat, amelyek

csak akkor jönnek létre, ha valamilyen más feltétel is teljesül. A kondicionált kapcsolatok részletes elemzése még fontos kihívás tehát a kutatók számára.

A hálózat kutatás egészére elmondható, hogy új eljárások, sőt új eljárás családok megjelenése várható, amelyek továbblépnek a jelenlegi korlátokon. Mindezekből érezhető, hogy a hálózat kutatás jellege és számításgénye miatt a kutatáshoz használt algoritmus és számítógépes szoftver kulcsszerepet tölt be. A humán biológiában és orvostudományban talán a Cytoscape (21) nevű szoftver a legismertebb, amely ingyenesen elérhető a kutatók számára. Számos farmakológiai, biológiai kutatás már ezzel a programmal történt. Az elméleti kutatásokban viszont inkább a CFinder, a Pajek és a Sonia az ismert eszköz. A hazai fejlesztésű Sixtep szoftvert – amely egyedi módon járványterjedést is képes szimulálni – pedig (többek közt) a Szegedi Tudományegyetem Orvosi Genetikai Intézetében használják.

Metabolikus kaszkádok

Számos – hálózat kutatással támogatott – felfedezés látott napvilágot a sejtbológia területéről. A modern rendszerbológia (system biology) megjelenésével nyilvánvalóvá vált, hogy a fehérjék közötti összetett kölcsönhatások, valamint a sejtek közötti kommunikációs és szabályozási rendszer is jól megragadható hálózatos értelmezésben. A sejtbológiában például a metabolikus kaszkádok felfedezése világított rá, hogy számos molekula (például a secretin) nem csak egyfajta módon (útvonalon) képződhet. A kutatók a PPI (protein-protein interaction) hálózat topológiai sajátosságainak elemzésével beavatkozási pontokat helyettesítő útvonalakat, megvédendő pontokat, indirekt kapcsolatokat, stressztűrő képességet, robusztusságot és megbízhatóságot vizsgálnak. A klaszterezési eljárások pedig például a komorbiditási összefüggések feltárásában használt módszer, hiszen a közvetlenül érintett géneken felül két vegyület közötti indirekt kapcsolatok feltárása már hasonlóan fontos feladat, mint a direkt kapcsolatoké.

Clinical Opinion Leader – határterületi fogalom

A kommunikációtudományon kívül a kommitológia, a lobbikutatás és a marketing is felfigyelt a hálózat kutatásra. A közösségek meghatározásán túl az egyik divatos irány a véleményvezérek és az „up-rising-star”-ok detektálása. Közel egy tu-

cat módszert dolgoztak ki a véleményvezérek detektálására (22). Az eredmények értelmezése nehéz, a véleményvezérek erősen kontextusfüggők, és a szakemberek szerint jelentősen más dinamika érezhető a hálózat növekvő, stagnáló és csökkenő időszakában. Az ismert nehézségek ellenére néhány éve a medicinában is megjelent a COL, a Clinical Opinion Leader (23), amely kulcsfogalomként vált a klinikusi közegben. A COL más néven az „informal educators/mentors”-nak (23) számító „gatekeeper” orvos, aki jelentős számú kollégára van hatással. A kutatások célja, hogy az új szakmai protokollok végrehajtását gyorsítani lehessen, és csökkenteni a korszerű klinikai gyakorlatba történő elterjedés időtartamát. Habár számos spekulatív elemet tartalmaz az elképzelés és sok kulturális elem is szerepet játszhat, néhány kísérlet eredménye szerint az *informal educators/mentors* kórházon belüli lokalizálása fontos tényező lehet a klinikum minőségbiztosítási folyamataiban (23).

A hálózatelemzés-hype

A hálózatelemzést alkalmazó közlemények színvonaluk rendkívül szórta. Időnként úgy tűnik,

hogy más módszertannal jobban megragadható jelenségeket is hálózatként értelmeznek kutatók, vagy egy vélt hálózat topológiai jellemzőinek kimutatását publikálják. Megfelelő lektor hiányában ezeket talán kellő gondosság nélkül is közlik tudományos folyóiratok. A tudományos közéletben tehát egyfajta „hálózatkutatás-hype” tapasztalható. Több kutató a hálózatelemzést valamiféle „theory of everything”-ként emlegeti. Meglátásunk szerint a „hype” káros is lehet, mivel nehezíti a módszer fejlődését és a valós lehetőségeinek felfedezését.

A rendszerelméletből régóta következik, hogy a jelenségek szüntelen kölcsönhatása hálózatos módon is értelmezhető. Megmutatkoztak az ehhez eszközt adó hálózatelemzés előnyei: számos multidiszciplináris kutatást indikált, és bizonyította, hogy a mennyiségi és minőségi szabályszerűségek mellett a gráfelmélet és a kombinatorika segítségével megragadhatóak új jelenségek és szabályszerűségek a medicinában. A molekuláris biológiában ez különösen tetten érhető. Az elkövetkező években – a megfelelő kutatási eljárásrendek kialakulása után – várhatóan számos újabb, hálózatelemzéssel kimutatott eredményről fogunk értesülni a medicinából és a határterületeiről is.

A rendszerelméletből régóta következik, hogy a jelenségek szüntelen kölcsönhatása hálózatos módon is értelmezhető.

Irodalom

- Christakis NA, Fowler JH. The Collective Dynamics of Smoking in a Large Social Network. *N Engl J Med* 2008;358:2249-2258.
- Christakis NA, Fowler JH. The Spread of Obesity in a Large Social Network over 32 Years. *N Engl J Med* 2007;357:370-379.
- Fowler JH, Christakis NA. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study. *BMJ* 2008;337:a2338.
- Buda B. Connected Networks. *LAM* 2011;21(5):402-403.
- Euler L. Solutio problematis ad geometriam situs pertinentis Commentarii academiae scientiarum. *Petropolitanae* 8(1741)128-140.
- Dunne JA, Williams RJ, Martinez ND. Food-web structure and network theory: The role of connectance and size. *PNAS* 2002;99(20):12917-12922.
- Lim J, Hao T, Shaw C, Patel AJ, Szabó G, Rual JF et al. A protein-protein interaction network for human inherited ataxias and disorders of Purkinje cell degeneration. *Cell* 2006;125:801-814.
- Nelson DL, McEvoy CL, Schreiber TA. The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods* 2004;36(3):402-407.
- Newman ME. Coauthorship networks and patterns of scientific collaboration. *PNAS* 2004;101(Suppl 1):5200-5205.
- Strogatz SH. Exploring complex networks. *Nature* 2001;410:268-276.
- Tarde G. Les lois de l'imitation. *Félix Alcan*. 1890.
- Katz E, Lazarsfeld PL. Personal Influence: The Part Played by the People in the Flow of Mass Communication. *New York; Free Press*. 1955.
- Moreno JL. Sociometry, Experimental Method and the Science of Society. An Approach to a New Political Orientation. *New York; Beacon House, Beacon*. 1951.
- Granovetter M. The Strength of Weak Ties: A Network Theory Revisited. *Sociological Theory* 1983;1:201-233.
- Vértes PE, Alexander-Bloch AF, Gogtay N, Giedd JN, Rapoport JL, Bullmore ET. Simple models of human brain functional networks. *PNAS* 2012;109(15):5868-5873.
- Liljeros F, Edling CR, Nunes LA, Stanley HE, Aberg Y. The Web of Human Sexual Contacts. *Nature* 2001;411:907-908.
- Motter AE, Gulbahce N, Almaas E, Barabási AL. Predicting synthetic rescue in metabolic networks. *Molecular Systems Biology* 2008;168(4):1-10.
- Chen LL, Blumm N, Christakis NA, Barabási AL, Deisboeck TS. Cancer metastasis networks and the prediction of progression patterns. *British Journal of Cancer* 2009;101:749-758.
- Barabási AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nature Reviews Genetics* 2011;12:56-68.
- Barabási AL. Behálózva – a hálózatok új tudománya. *Magyar Könyvklub; Budapest*. 2003. pp. 213-214.
- Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 2011;27(3):431-2.
- Valente TW, Pumpuang P. Identifying Opinion Leaders To Promote Behavior Change. *Health Educ Behav* 2007;34(6):881-896.
- Borbas C, Morris N, McLaughlin B, Asinger R, Gobel F. Role of Clinical Opinion Leaders in Guideline Implementation and Quality Improvement. *Chest* 2000;118 (Suppl 2):24S-32S.