

# KOMPLEX HÁLÓZATOK A TERMÉSZETBEN ÉS A TÁRSADALOMBAN

Kertész János

az MTA levelező tagja, BME Fizikai Intézet – kertesz@phy.bme.hu

Vicsek Tamás

az MTA rendes tagja, ELTE Fizikai Intézet – vicsek@angel.elte.hu

## 1. Bevezetés

Ma már közhely, hogy körbe vagyunk véve hálózatokkal. Egyfelől tudatosan megéljük, hogy az internetet mint hálózatot használjuk, ugyanakkor talán sokan nem is tudják, hogy azokat a bonyolult biológiai, közgazdasági és egyéb rendszereket, amelyekbe ágyazva élünk, leginkább úgy lehet megérteni, ha hálózatoknak tekintjük őket.

Sokféleképpen lehet egy elemekből (pontokból, csúcsokból) és az azokat összekötő kapcsolatokból (élekből) álló hálózatot (gráfot) jellemezni. Ha az elemek száma nagy, akkor kézenfekvő, hogy a sok kölcsönható részecskéből álló rendszerek tulajdonságainak megértése céljából kifejlesztett statisztikus fizikai módszereket használjuk a hálózatok megismerésére.

Magyarországon a hálózatok modem elméletéhez kapcsolódó témákban széleskörű és sikeres kutatás folyik. Ebben a cikkben – egy bevezetést követően – két kiragadott területre: a súlyozott élekkel és a csoportosulásokkal rendelkező hálózatokra vonatkozó eredményeket tekintjük át.

## 2. A természetben, a társadalomban előforduló hálózatok és univerzális tulajdonságaik

A „komplex rendszer” fogalmát a fizikában vezették be olyan rendszerekre, ahol az alko-

tóelemek nagy száma és a közöttük lévő kölcsönhatás révén a rendszer viselkedése az egyes egységektől lényegesen eltérő sajátosságokat mutat. Egyszerűen fogalmazva: az egész több mint részeinek összege. A fizikából számos példát lehet idézni komplex rendszerekre a mágnesektől az üvegeken keresztül a szemcsés anyagokig. A fogalom nyilvánvalóan túlmutat a fizikán: a sejt, az agy, a gazdaság, a társadalom csak néhány további példa.

A komplex rendszerek esetében a kölcsönhatások részleteiről gyakran beható ismeretekkel rendelkezünk, de ez nem vezet az egész igazi megértéséhez. Mi történik, ha egy nagy ugrással teljesen figyelmen kívül hagyjuk a kölcsönhatások természetét, és csak az általuk generált topológiával, a komplex rendszer vázával foglalkozunk?

Nevezzük a komplex rendszer alkotó-elemeit csúcsoknak. Ha két ilyen alkotóelem kölcsönhat, akkor azt mondjuk, hogy a csúcsok között van él; az egy csúcsához tartozó élék száma a fokszám. (Az élék lehetnek irányítottak is.) Az így nyert objektumot a komplex rendszerhez tartozó hálózatnak tekintjük. A matematikában a gráfelmélet foglalkozik ilyen objektumokkal, és óriási ismeretanyag halmozódott fel ezekről az elmúlt 250 év során. A gráfelméletet sikerrel alkalmazták például a villamosmérnöki, a kémiai tudományokban vagy a szociológiában. Az

<i>Jelenség</i>	<i>Csúcsok</i>	<i>Élek</i>
1. sejt metabolizmus	molekulák	kémiai reakciók
2. tudományos együttműködés	tudósok	közös cikkek
3. WWW	honlapok	URL mutatók
4. légi közlekedés	repülőterek	repülőjárat
5. gazdaság	vállalatok	kereskedés
6. nyelv	szavak	hasonló jelentés
7. társadalom	emberek	ismeretség

1. táblázat • Példák hálózatokra

1990-es évek végén bekövetkezett látványos fejlődés elsősorban annak köszönhető, hogy új eszközöket és módszereket alkalmaztak a hálózatok tanulmányozására, erősen támaszkodva a statisztikus fizikára (Albert – Barabási, 2002; Dorogovtsev – Mendes, 2002; Newman, 2003).

Az 1. táblázat különböző kölcsönhatásokra épülő hálózatokat tartalmaz. Van közöttük irányított (például a 3.) és irányítatlan (például a 2.), ember által alkotott és tervezett (például 4.), természeti fejlődés eredménye (1.), vagy emberi közreműködéssel, de spontán fejlődéssel keletkezett (például 3., 5.).

A hálózatok empirikus tanulmányozásának a lehetőségét óriási mértékben növeli, hogy a komputerizációnak köszönhetően az elérhető adatok mennyisége hihetetlen módon megnőtt. Néhány példa: A hollywoodi filmszínészek adatbankja 450 ezernél több személyt tartalmaz, valamennyi filmjükkel együtt. A *Humán Genom Projekt* egy harmincezer génről és hárommilliónál több bázispár szekvenciájáról hozott létre adatbankot. A WWW mérete meghaladja a hatmilliárd honlapot. A tőzsdén valamennyi tranzakciót feljegyzik egy állandóan növekvő adathalmazt alkotva, amelyből gazdasági hálózatok térképezhetők fel. Az ilyen hatalmas adatmennyiségek tanulmányozása fontos felfedezésekhez vezetett. A fő kérdés, hogy van-e remény arra, hogy ennyire különböző eredetű és funkciójú hálózatok közös vonásokat mutassanak.

*Kicsi a világ* – mondjuk, ha váratlanul kiderül, hogy könnyű kapcsolatot találni egy ismeretlen személyhez ismeretségi láncon keresztül. Már Karinthy Frigyes rámutatott, a 60-as évektől pedig szociológusok igazolták, hogy a társadalom szerkezete ilyen: Az alkotóelemek nagy száma ellenére az átlagos lépésszám („távolság”) a hálózaton meglepően kicsiny. A „kis világ” tulajdonság számos hálózat sajátossága, a kutatók társszerzői hálózatától a WWW-n keresztül a sejtekig.<sup>1</sup>

Komplex hálózatok egy másik érdekes, szintén a szociológiából ismert tulajdonsága az, amit röviden a „barátok barátai könnyen barátkoznak” kifejezéssel lehet érzékeltetni. Más szóval, a hálózatokban a háromszögek gyakorisága nagy (például egy véletlen hálózathoz képest). További, igen széles körben jelentkező tulajdonsága a komplex hálózatoknak az ún. skálamentesség. Ez azt jelenti, hogy a csúcsok fokszámának nincs egy jellemző értéke, aminél sokkal kisebb illetve sokkal nagyobb fokszámú csúcsot csak elenyésző hányadban lehet találni. Matematikailag ez abban nyilvánul meg, hogy a fokszámeloszlás hatványfüggvény jellegű. A skálamentesség a komplex rendszerek fizikai elméletében központi szerepet játszik. A komplex rendszerekre jellemző, hogy hierarchikus szerveződésűek, azaz

<sup>1</sup> Matematikailag egy összefüggő hálózat „kis világ”, ha a két pont közötti legrövidebb távolságok átlaga nem növekszik logaritmikusan gyorsabban a rendszer méretével.

egymásba ágyazott egységeket tartalmaznak. Ez a tulajdonság az emberi társadalomra és az egyes emberekre is igaz. Munkahelyünkön az egyes szinteknek megfelelően, például csoportok, osztályok, főosztályok, leányvállalatok vannak. Életünket szerveink (májunk, agyunk, lépünk stb.) összehangolt működése tartja fent. Ezek a szervek sejtekből állnak, amelyekben sejtsejtszervecskék találhatók. A legalsó, már konkrét biológiai jelentéssel/funkciókkal rendelkező építőelem a több mint százezerféle fehérje.

A nagy, véletlen hálózatok leírásában a 90-es évek végéig az Erdős–Rényi-modell volt az uralkodó. A fenti sajátosságok közül azonban ez a teljesen véletlen kapcsolatokat feltételező modell sem a nagy csomósodási tulajdonságot, sem a skálamentességet nem adja vissza. 1998-ban Duncan Watts és Steven Strogatz egyszerű modellt vezettek be: egy nagy csomósodási tulajdonságú (sok háromszöget tartalmazó), szabályos hálózatból (rácsból) indultak ki, és ebbe viszonylag kis számú, véletlen élt helyeztek el. Az így létrejött hálózat nagy csomósodású kis világ. Barabási Albert-Lászlónak kulcsszerepe volt a skálamentesség általános jellegének felismerésében. Az általa és Albert Réka által kidolgozott, kivétező csatlakozáson alapuló növekedési hálózatmodell a skálamentes kis világ modellek paradigmájává vált.

A továbbiakban az adatok elemzésének néhány új módszerét ismertetjük.

### 3. Súlyozott hálózatok és elemzésük

A komplex rendszerek hálózati alapú tanulmányozása holisztikus megközelítés, szemben a kölcsönhatások tulajdonságaira összpontosító redukcionista leírással. A valódi megértéshez a két szélsőséges álláspont egyesítésére van szükség. Az első lépés a hálózati irányból, ha a kölcsönhatásokat nem tekintjük egyformának, és erősségüket valamilyen skálán jellemezzük, vagyis súlyokat rendelünk a kölcsönhatásokat jelképező élekhez.

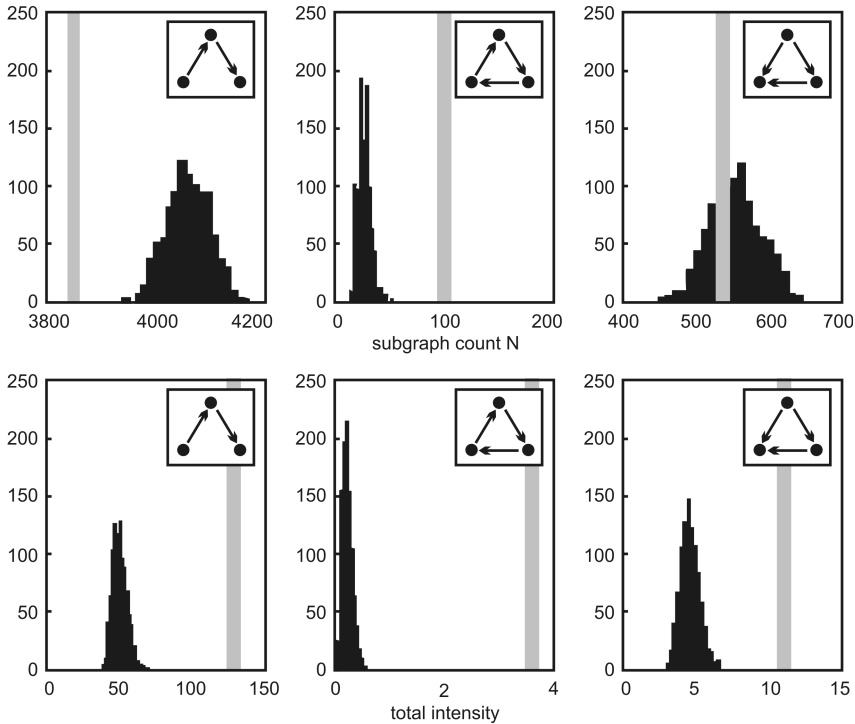
Természetes súlyként adódik az időegységre vetített forgalom például a közlekedési hálózatoknál, az interneten vagy egy telefonhálózatban. Kémiai reakciókban a szereplő anyagáramok mértéke használható fel súlyok formálására (lásd például Almaas et al., 2004). A hálózatokat gyakran valamilyen komplex rendszerben lejátszódó folyamatok korrelációs függvényeinek segítségével nyerjük – ilyenkor nyilván a korreláció mértéke azonosítható az élhez tartozó súllyal.

A súlyozatlan hálózatok jellemzéséhez célszerű volt az egyes motívumok statisztikáját tanulmányozni. Motívum ebben a vonatkozásban a topológiaiilag azonos részgráfok összessége.<sup>2</sup> Például a korábban már említett háromszögek ilyen motívumnak tekinthetők. Ha egy motívum szignifikánsan gyakori egy hálózatban (például a véletlen hálózathoz képest), akkor felmerül, hogy funkcionálisan fontos lehet (Milo et al., 2002).

Hogyan lehet ezt a megközelítést a súlyozott hálózatokra kiterjeszteni (Onnella et al., 2005)? Be kell vezetni egy mérőszámot, ami a részgráf súlyát jellemzi. Erre (pozitív súlyok esetében) a részgráfban szereplő élek súlyainak mértani közepe kínálkozik, amit a részgráf intenzitásának nevezünk. Jobb a mértani közeget választani a számtani helyett, hiszen az előbbi automatikusan szolgálja a 0 intenzitást, ha valamelyik él hiányzik a részgráfból.

Az 1. ábra bemutatja, hogy a súlyok figyelembe vétele hogyan változtathatja meg akár az egyes motívumokra vonatkozó következtetéseket is. A vizsgált irányított hálózat az *Escheria coli* baktérium metabolizmusa alapján készült, ahol a kémiai reakciók során fellépő eredő fluxus segítségével lehet a súlyokat meghatározni. Az ábra három egyszerű motívum statisztikáját mutatja be: a függőleges vonalak jelentik a mért értéket. Az összehasonlításként feltüntetett el-

<sup>2</sup> Részgráf a hálózat olyan része, amely maga is egy hálózat (gráf).



1. ábra • Súlyozatlan (felső sor) és súlyozott motívumok statisztikája az E. coli baktérium metabolikus hálózatában.

oszlások a viszonyítási rendszernek felelnek meg, olyan véletlen hálózatnak, amelynek a fokszámeloszlása megfelel a megfigyeltnek, és amelyen az empirikus súlyoknak véletlen permutációi valósultak meg. A tanulság, hogy bizonyos ritka motívumok igen nagy intenzitással lehetnek képviselve, ami a funkcionalitásra vonatkozó következtetéseket befolyásolja.

A súlyozott hálózatok elemzésének további, hatékony és egyszerű módszere a „küszöbölés”. Egyik megvalósítási lehetősége, hogy a hálózatban figyelmen kívül hagyjuk az egy bizonyos küszöbérték alatti súlyral rendelkező éleket. A küszöbérték mozgatásával egy képsorozat vagy mozgókép keletkezik, amelyik az elszigetelt csúcsoktól elindulva, az élek fokozatos, erősség szerinti bekapcsolásával végül a teljes hálózatig jut el. Ezzel a módszerrel

elemelve a New York-i tőzsde legnagyobb vállalatai árfolyamváltozásainak korrelációs függvénye alapján létrehozott hálózatról (Onnela et al., 2004) kiderül, hogy bizonyos gazdasági ágazatokon belül (például ilyen az energiaszektor) nagyon erős a korreláció, míg más szektorok (például a pénzügy) elemeire ágazatokon átívelő szoros kapcsolatok jellemzőek. Az ilyen és hasonló módszerekkel szerzett információk egyrészt hozzájárulnak a piac alaposabb megértéséhez, de felhasználhatók például a portfóliók optimalizálásánál is.

A szociális hálózatok vizsgálatánál régóta hangsúlyozzák a súlyok fontosságát, de ezek meghatározása nem könnyű feladat. A modern telekommunikáció korában új, a korábbinál jóval objektívebb módszerek alkalmazására nyílik lehetőség. Így például

a telefontársaságok minden beszélgetésről feljegyzik, hogy ki hívott kít, mikor, és mennyi ideig tartott a beszélgetés. Kézenfekvő a kapcsolatok intenzitását az egymással folytatott beszélgetésekre fordított idővel mérni. Egy ilyen hatalmas, több millió csomópontot tartalmazó rendszerre alkalmazva a küszöbölés módszerét, érdekes kép tárul elénk (Onnela et al., preprint): ha a legerősebb kapcsolatok behelyezésével kezdjük a hálózat felépítését, majd fokozatosan csökkentjük a küszöböt, azt tapasztaljuk, hogy először intenzív kapcsolatban lévő egyedekből álló, de egymástól elszigetelt közösségek azonosíthatók. Csak a viszonylag gyengébb kapcsolatokat jelképező élek megjelenésével alakul ki az egész rendszerre (társadalomra) kiterjedő hálózat. Ily módon igazolni lehet Mark Granovetter 70-es években felállított hipotézisét a gyenge kötések fontosságáról (Granovetter, 1973): bizonyos értelemben ezek tartják össze, stabilizálják a társadalmat.

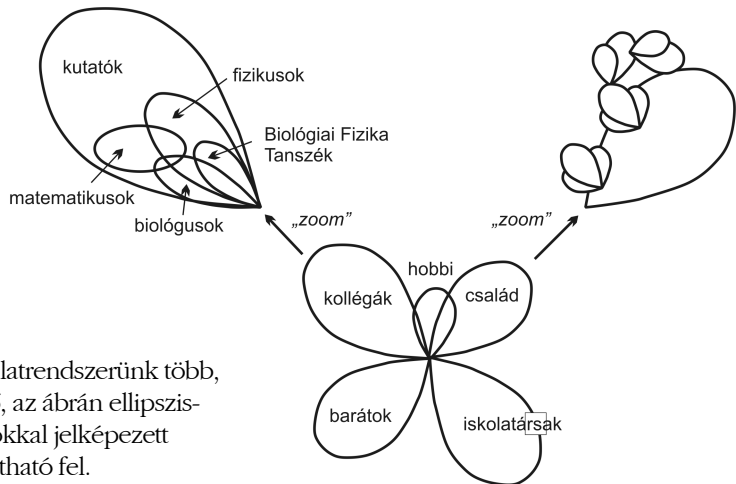
4. *Csoportosulások/közösségek a hálózatokban*

A valós hálózatok rendkívül összetettek, a sok ezer csúcs és a sok tízezer él között szinte reménytelen feladatnak tűnik az eligazodás. Vajon egy komplex hálózat (például

génhálózat) is olyan, mint egy „szervezet”? Vannak benne jól meghatározható, különálló egységek hasonló típusú egységekből, vagy a szerveződés kibogozhatatlan?

Vizsgáljuk meg, hogy egy nagy, biológiai vagy szociológiai eredetű hálózat belső szerkezete hogyan jellemezhető a bennük levő sűrűbben összekapcsolt részek (közösségek) feltárása segítségével.

Atémakör lényegét legjobban egy konkrét példán lehet bemutatni. A szemléltetéshez alkalmasabb a szociológiai megközelítés (Faust, 2005; Newman, 2004). A gráf pontjai itt az egyének, és akkor van köztük él, ha két egyén valamilyen kapcsolatban van egymással. Tekintsük tehát egy embertársunkat, egy kutatót, aki például biológiai fizikai kutatásokkal foglalkozik. Ennek a személynek a kapcsolatrendszere jól meghatározott csoportokból áll, amelyeken belül a tagok között viszonylag több a kapcsolat, mint a tagoknak a külvilág felé. Ilyen csoportok vagy másképpen közösségek a család, a barátok, a kollégák, az iskolatársak stb. Ezek a közösségek egyszerre *különállóak, és át is fednek, valamint több szíjjal kötődhetnek* a hálózat fennmaradó részéhez. Átfedés az, hogy volt iskolatárs lehet kolléga, sőt barát is, de nyilván vannak további, más közös-



2. *ábra* • Kapcsolatrendszerünk több, egymással átfedő, az ábrán ellipsziszszerű tartományokkal jelképezett részhálózatra osztható fel.

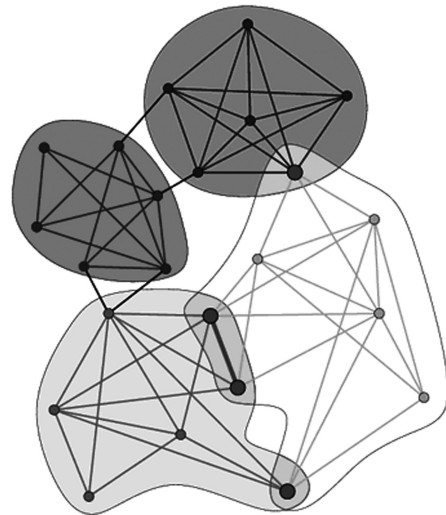
ségekkel is kapcsolatai. A szociális kapcsolati háló tehát rendkívül összetett, egymással átfedő, izolált vagy éppen egymásba ágyazott csoportosulások együttes halmaza, és az egész társadalom működésének megértéséhez ezt a komplex hálózatot is át kell tudnunk tekinteni. Ezt próbálja szemléltetni a 2. ábra.

Azt az igényt, hogy keressük meg egy gráf sűrűbben összekötött részeit, sokkal egyszerűbb megfogalmazni, mint ténylegesen teljesíteni. Már maga a *közösség fogalma sem egyértelműen definiált*, hiszen a „sűrűbb összekötöttség” kritériuma többféleképpen is teljesülhet.

Ha egy adott hálózaton belül van  $k$  darab olyan pontunk, amelyek egymással páronként mind össze vannak kötve, akkor ez a pontthalmaz egy  $k$ -klikket képez. Két  $k$ -klikk pedig akkor szomszédos, ha van egy közös  $k-1$  pontot tartalmazó althalmazuk (amely szintén klikk). Definíció: *két elem akkor tartozik ugyanahhoz a közösséghez, ha össze lehet kötni őket szomszédos  $k$ -klikkeken keresztül*. A  $k$ -klikk közösség tehát az összes olyan egyed halmaza, akik ily módon egymással összeköthetőek. Egy ilyen csoportosulás lehet kicsi, de mérete, ha a gráf elég sűrű, összemérhetővé válhat a teljes hálózatával (ilyenkor mondjuk, hogy az alhálózat „perkolál” [Derényi et al., 2005]).

A fenti típusú közösségek megkeresése egy trükkös algoritmus alapján történik. Van tehát egy módszerünk, most kipróbálhatjuk, milyen eredményeket ad valós adatokra (Palla et al., 2005). A fizikusok működtetnek egy elektronikus cikkarchívumot, amelyen több tízezer cikk adatai megtalálhatóak. Ez az adathalmaz egy gráfnak tekinthető: *a szerzők a pontok, és két szerző akkor van összekapcsolva, ha van közös cikkük*. Ha erre a gráfra lefuttatjuk a közösségedetektáló algoritmusunkat, azt tapasztaljuk, hogy eredményeképpen feltérképeződnek a szerzők tágabb értelemben vett (több országon átívelő) kutatókollektívái.

A legfontosabb alkalmazás azonban valószínűleg a biológiai hálózatok terén várható. Amint arról a fentiekben részletesebben is írtunk, legfontosabb építőköveink, a sejteink működéséről egyelőre a legtöbbet éppen a bennük található fehérjék kölcsönhatási hálózatainak megismerése útján tudhatunk meg. A fehérjék közösségeinek feltérképezésétől sokat várunk: az egyes csoportosulások feltárása segíthet a még nem ismert funkciójú fehérjék szerepének tisztázásában. Többek között megvizsgálható, hogy egy adott  $F$  fehérje milyen közösségek tagja. Ha  $F$  benne van egy olyan csoportosulásban, amelyben a többi fehérje mondjuk épp a sejtosztódás egy bizonyos folyamatában játszik fontos szerepet, akkor sikerült kimutatnunk, hogy  $F$  valószínűleg szintén fontos ennek a sejtosztódási folyamatnak a szempontjából. Amennyiben ezt nem tudtuk erről a fehérjéről valamilyen más forrásból



3. ábra • Az egyes alhálózatokon belül egy teljes algráf, itt például egy  $k=4$  pontból álló klikk, szomszédos  $k-1$  klikkeken (háromszögeken) keresztül végiggördíthető. A fenti négy alhálózat két helyen egy-egy pontban, egy helyen pedig egy élben fed át.

(és ez ma még gyakori), akkor máris kiderül, hogy a módszert használni tudtuk annak a megjelölésére, hogy mi lehet a fehérje szerepe a sejtben. Egy ilyen predikció fontossága gyógyszertervezési, illetve gyógyászati szempontból nyilvánvaló.

### 5. Magyar hálózatkutatók

A hálózatok kutatása igazi multidiszciplináris feladat. A magyar kutatók Erdős Pál és Rényi Alfréd klasszikus munkái óta kulcsszerepet játszanak ezen a területen. Hadd említsük itt meg néhány komplex hálózatokkal foglalkozó kolléga nevét a különböző szakterületekről, a teljesség igénye nélkül, és elnézést kérve a kihagyottaktól: Bollobás Béla (Cambridge, UK és Memphis, USA) a véletlen gráfok matematikai elméletének kiemelkedő művelője. Barabási Albert-László (Notre Dame, USA) egyike volt a komplex hálózatok elméletében bekövetkezett robbanásszerű fejlődés elindítóinak. Tusnády Gábor (Rényi Intézet) a hálózatok matematikájával, Benczúr András (SZTAKI)

WWW modellezéssel, Csemely Péter (SOTE) (Csemely, 2004) biokémiai hálózatokkal, Vedres Balázs (CEU) szociális hálózatokkal foglalkozik igen sikeresen.

Befejezésül megemlíti, hogy magyar nyelven három ismeretterjesztő könyv is megjelent a komplex hálózatok elméletéről. Mark Buchanan *Nexus* című műve mellett két neves magyar szakember is írt könyvet a szélesebb közönség számára. Csemely Péter *A rejtett hálózatok ereje* című könyve néhány hónappal ezelőtt jött ki. Különösen Barabási Albert-László *Behálóztva* című művét ajánljuk az olvasók figyelmébe, mert nemcsak olvasmányos módon ismerteti meg a tudományág új eredményeivel, de ízelítőt ad a nemzetközi szinten kiemelkedően elismert kutatómunka izgalmából is. Nem véletlen, hogy a könyv eredeti, angol változata felkerült a *New York Times* bestsellerlistájára.

Kulcsszavak: *komplex rendszer, gráf, statisztikus fizika, súlyozott hálózatok, skálamentesség, perkoláció*

### IRODALOM

- Albert Réka – Barabási, Albert-László (2002): Statistical Mechanics of Complex Networks. *Reviews of Modern Physics*. **74**, 47.
- Almaas, Eivind – Kovács B. – Vicsek T. – Oltvai Z. – Barabási A.-L. (2004): Global Organization of Metabolic Fluxes in the Bacterium *Escherichia Coli*. *Nature*. **427**, 839–843.
- Csemely Péter (2004): A gyenge kölcsönhatások ereje a stresszfehérjéjéktől a szociális hálózatokig. *Magyar Tudomány*. **12**, 1318–1324.
- Derényi Imre – Palla G. – Vicsek T. (2005): Clique Percolation in Random Networks. *Physical Review Letters*. **94**, 160202.
- Dorogovtsev, Sergei N. – Mendes, José F. F. (2002): Evolution of Networks. *Advances in Physics*. **51**, 1079.
- Granovetter, Mark (1973): The Strength of Weak Ties. *American Journal of Sociology*. **78**, 1360.
- Faust, Katherine (2005): Using Correspondence Analysis for Joint Displays of Affiliation Networks. In: Carrington, Peter J. – Scott, J. – Wasserman, S. (eds.):

- Models and Methods in Social Network Analysis* Ch. 7. Cambridge University Press, New York
- Newman, Mark E. J. (2003): The Structure and Function of Networks. *SIAM Review* **45**, 167–256.
- Newman, Mark E. J. (2004): Detecting Community Structure in Networks. *European Physical Journal B*. **38**, 321–330.
- Milo, Ron et al. (2002): Network Motifs: Simple Building Blocks of Complex Networks. *Science*. **298**, 824.
- Onnela, Jukka-Pekka – Kaski, K. – Kertész J. (2004): Clustering and Information in Correlation Based Financial Networks. *European Physical Journal B*. **38**, 353.
- Onnela, Jukka-Pekka – Saramäki, J. – Kertész J. – Kaski, K. (2005): Intensity and Coherence of Motifs in Weighted Complex Networks. *Physical Review E*. **71**, 065103(R).
- Onnela, Jukka-Pekka – Hyvönen, J. – Saramäki, J. – Kaski, K. – Kertész J. – Szabó G. Barabási A.-L.: Weak Ties in the Structure and Function of Social Networks (preprint)
- Palla Gergely – Derényi L. – Farkas, L. – Vicsek, T. (2005): Uncovering the Overlapping Community Structure of Complex Networks in Nature and Society. *Nature*. **435**, 814.