TANULMÁNYOK

A hálózati szerkezet és a hálózati csúcsok jellemzőit integráló tudáshálózati mutató: az ENQ-index

ENQ Index: an Integrated Measure of Network Structure and Node Characteristics in Knowledge Networks

Sebestyén Tamás
MTA-PTE Innováció és Gazdasági Növekedés Kutatócsoport
Pécsi Tudományegyetem, Közgazdaságtudományi Kar

E-mail: sebestyent@ktk.pte.hu

Varga Attila
MTA-PTE Innováció és Gazdasági Növekedés Kutatócsoport
Pécsi Tudományegyetem, Közgazdaságtudományi Kar
E-mail: vargaa@ktk.pte.hu

A tanulmány az ún. ego-hálózat minőség (Ego Network Quality: ENQ) indexet mutatja be. Ez az index kompozit mércében számszerűsíti a tudáshálózatokból hozzáférhető tudás értékét, a mind a közvetlen, mind a közvetett hálózati partnerek figyelembevételével. Az index a hálózati pozíció három aspektusát integrálja. Egyrészt figyelembe veszi a partnerek tudásszintjét, amely újdonság a standard hálózati mutatókkal szemben. Másrészt a tudásszinteket a partnerek közötti kapcsolatrendszer szerkezetét leíró mutatószámmal súlyozza, így a hálózati szerkezet figyelembevételét is lehetővé teszi. Harmadrészt nemcsak a közvetlen, hanem a közvetett szomszédok tudásszintjét és kapcsolódási struktúráját is számba veszi. A szerzők rámutatnak arra, hogy az ENQ-index ismert hálózati mutatókra vezethető vissza. Szimulációs és vizsgálataink azt támasztják alá, hogy az ENQ-index robosztus módon képes megjeleníteni egy csúcs pozícióját a hálózatban, a távolságra és a szerkezetre alkalmazott súlyozás megválasztása pedig másodlagos jelentőségű.

Kulcsszavak:
Ego Network Quality,
hálózatelemzés,
tudáshálózatok,
regionális innováció

In this paper we introduce the Ego Network Quality (ENQ) index, which quantifies the value of knowledge accessible from knowledge networks through, taking into consideration both direct and indirect partners in the network. The index integrates three aspects of network position. First, it takes into account the knowledge level of partners, which is a novelty in comparison with standard network indicators. Second, it weighs knowledge levels with an indicator reflecting the structure of links between partners, focusing on network structure. Third, not only the knowledge level and network structure of direct partners but also those of the indirect ones are accounted for. It is shown that the ENQ index can be linked to already known network indicators in special cases. Simulation exercises are used to verify that the ENQ index can reflect the position of a node i within the network in a robust way, and the selection of distcane and structure weighting is only secondary importance.

Keywords:
Ego Network Quality, network analysis, knowledge networks, regional innovation

Beküldve: 2015. július 6.
Elfogadva: 2015. december 11.

Bevezető

A hálózatelmélet eszközeit egyre szélesebb körben alkalmazzák a különböző dimenziókban jelentkező tudásáramlások vizsgálatára két, egymástól viszonylag elkülönülten fejlődő területen. A tudomány és a technológia folyamatait elemző, a térbeli jellemzők hatásaival tipikusan nem foglalkozó szakirodalomban a fő kutatási irány a különböző hálózati jellemzők (például a méret, a centralitás, a sűrűség, a kapcsolatok erőssége) vagy a tudás sokféleségének, tipikusan valamilyen, a hálózat elemeinek teljesítményét megragadó mutatóra (publikációk, hivatkozások, profit stb.) történő hatásvizsgálata. Ezt általában vállalatok, kutatóintézetek, tudományos publikációk szerzői közötti hálózatokra végzik el, melynek során a hálózat tagjainak teljesítménye és az egyes szereplők ego-hálójának szerkezete között keresnek összefüggést. A hálózati szereplők ego-hálójának tipikusan négy elemét emeli ki a szakirodalom: (1) a közvetlen partnerek jellemzői és a közvetlen partnerekkel való kapcsolat intenzitása (a partnerek száma, a kapcsolatok erőssége, a partnerek tudása); (2) a partnerek közötti interakciók intenzitása; (3) a hálózati kapcsolatokon keresztül elérhető tudás sokfélesége; (4) a szereplők pozíciója a teljes hálózaton belül (például centrális, periferikus pozíció).

A szakirodalmi eredmények pozitív kapcsolatot mutatnak ki a kutatás eredményessége és a partnerek száma (Powell et al. 1999, Hopp et al. 2010, Van Der Deijl et al. 2011), a kapcsolatok erőssége (Van Der Deijl et al. 2011), valamint a partnerek tudása (Maggioni et al. 2007, Hoekman et al. 2009, Ponds et al. 2009, 2010, Varga et al. 2014) között. Ugyanakkor a partnerek közötti kapcsolatok erősségének szerepéről a szakirodalomban nem találtunk egyértelmű eredményeket. Salmenkaita (2004), valamint Cross és Cummings (2004) tanulmányai pozitív kapcsolatot mutattak ki a szabadalmi tevékenység és a feltalálók partnerei közötti kapcsolatok erőssége között. Van Der Deijl et al. (2011) fordított U-alakú kapcsolatot mutatott ki, míg Rumsey–Wairepo (2006) és Cainelli et al. (2010) negatív kapcsolatot talált a partnerek közötti kapcsolat erőssége és az akadémiai publikációs tevékenység között. A hivatkozott tanulmányok azt is megerősítették, hogy mind a hálózati partnerek tudásának sokfélesége (Powell et al. 1999, Cainelli et al. 2010, van Der Deijl et al. 2011), mind pedig a szereplők centralitása a teljes hálózatban (Powell et al. 1999, Cainelli et al. 2010, van Der Deijl et al. 2011, Hopp et al. 2011) pozitívan befolyásolja a tudástermelést.

A tudásáramlásokban a tér szerepét középpontba állító regionális közgazdaságtani és a gazdaságföldrajzi szakirodalom más gyökerekkel és motivációval közelít. E vizsgálatokban a térbeliség jelentősége a tudáshoz való hozzáférés szempontjából válik központivá. Közel kell-e egymáshoz települniük a tudás termelésében részt vevő szereplőknek (vállalatok, egyetemek, vagy különböző üzleti szolgáltatók) a sikeres innováció érdekében, vagy az agglomerációs előnyök távoli forrásokból származó tudáselemekhez való hozzájutással is kiválthatóvá válnak? Ugyanazon iparágba tartozó vállalatok vagy egymástól különböző szektorokhoz tartozó vállalatok térbeli tömörülése előnyösebb-e az innováció szempontjából? Ez a terület tehát kevésbé a tudáshálózatok felépítését, mint inkább a partnerek tudásához való hozzáférésben a tér szerepét kutatja. A partnerek tudásszintje tehát az egyik legfontosabb meghatározója a hálózatokban történő tudásáramlásnak. Ezt a kutatási irányt nemcsak a kérdésfeltevésben, de a modellezésben alkalmazott technikák tekintetében is jelentősen befolyásolták azok a térökonometriai módszerek, amelyeket a térbeli tudásáramlás első vizsgálatai alkalmaztak (például Anselin, Varga és Ács, 1997). Gyakran alkalmazott módszer a térbeli hálózatok kutatásában, hogy a térökonometria térbeli súlymátrixait a régiók közötti együttműködési kapcsolatokat leíró mátrixra cserélik ki (Maggioni és Uberti, 2011, Varga, Pontikais és Chorafakis, 2014, Ponds, van Oort és Frenken, 2010).

Tanulmányunkban az ENQ- (Ego Network Quality – ego-hálózat minőség) indexet mutatjuk be, amely a fent érintett két területen (a hálózati szerkezetre fókuszáló tér nélküli elemzések és a partnerek tudásához való hozzáférést középpontba állító térbeli elemzések területein) alkalmazott megközelítéseket integrálja. A megközelítés alapja az a megállapítás, hogy egy innovációs szereplő (egyén, vállalat, kutatóhely, régió stb.) pozíciójának értéke, vagyis a szereplő hálózatából hozzáférhető tudás minősége függ mind a hálózat, a kapcsolatrendszerek szerkezetétől, mind a partnerek jellemzőitől. Az index célja, hogy segítsen számszerűsíteni a hálózat egy adott pontján elhelyezkedő szereplő számára a hálózatból hozzáférhető tudás minőségét. Habár az indexet térbeli hálózatok vizsgálatára fejlesztettük ki, a módszert általánosíthatónak tartjuk tetszőleges dimenzióban értelmezett tudáshálózatokra is.

Az e tanulmányban bemutatott ENQ-indexet korábban több elemzésben alkalmaztuk. Sebestyén és Varga (2013a) a hálózati minőség hatását vizsgálják európai régiók közötti szabadalmi és kutatási kooperációkon keresztül, megmutatva, hogy a hálózati minőség fontos meghatározója a K+F-ráfordítások termelékenységének. Varga és Sebestyén (2015, 2016) szintén az ENQ-index használatával kimutatták, hogy markáns különbség van a közép-kelet-európai régiók, valamint a nyugat-európai régiók között a tekintetben, ahogyan a helyi és a hálózatokon keresztül elérhető tudást használják: míg az utóbbiak esetében a két tudásforrás inkább kiegészítő, addig az előzőek esetében helyettesítő viszonyban áll egymással. Hau, Sebestyén és Varga (2016) egy integrált gazdaságpolitikai hatáselemző modellben alkalmazzák az indexet, melynek segítségével különböző, a hálózatformálódást befolyásoló beavatkozások hatása tesztelhető.

A tanulmány első fejezete az ENQ-indexet mutatja be, a második fejezet rávilágít az index több, gyakran használt hálózati mutatószámmal való kapcsolódására. A harmadik fejezet az index robusztusságát vizsgálja, különböző súlyozási megoldások mellett. A tanulmányt összegzés zárja.

Ego Network Quality index

Az innovációelmélet az innovációban részt vevő szereplők közötti interakciók fontosságát hangsúlyozza. Ezek az interakciók rendszerbe szerveződnek és a rendszer jellemzői nagymértékben meghatározzák az új tudás termelésének hatékonyságát (Lundvall 1992, Nelson 1993). Kiterjedt, felméréseken alapuló szakirodalom dolgozza fel, hogy az innováció valóban kollektív folyamatként értelmezhető, ahol a partnerek szakértelme, tudása és a közöttük zajló együttműködések intenzitása nagymértékben befolyásolja az új, gazdaságilag hasznos tudás létrejöttét (például Diez 2002, Fischer és Varga, 2002). Az innováció szereplőit csúcsokként, a közöttük kialakuló kapcsolatokat élekként értelmezve az együttműködő szereplők viszonyrendszere gráffá, hálózattá képezhető le. Erre építve a hálózatelemzés eszköztára lényegesen kibővíti a tudáskapcsolatok vizsgálatának spektrumát, túllépve a hagyományos felméréseken alapuló módszerek lehetőségein.

Az ENQ-index (Sebestyén és Varga 2013a, b) koncepciója mögött három intuíció húzódik meg, amelyeket közvetlen módon az innováció elmélete motivál. Az első szerint egy szereplő hálózatában fellelhető tudás szintje, nagysága pozitív kapcsolatban áll a szereplő új tudás létrehozásában mutatott termelékenységével, hatékonyságával. A második intuíció, hogy egy szereplő partnerei közötti kapcsolatok intenzitása további pozitív hozadékkal bírhat a hálózatból kinyerhető tudás szempontjából. A harmadik intuíció, hogy a partnerek nemcsak a hozzáférhető tudás nagysága szempontjából fontosak, hanem a tudás sokféleségéhez, diverzitásához is hozzájárulnak azáltal, hogy további, közvetlenül nem hozzáférhető tudásokat csatornáznak be a vizsgált szereplőhöz.

Ennek megfelelően az ENQ-indexet alapvetően két koncepció köré építjük. Az egyik koncepció a tudáspotenciál, amely a partnerek tudásszintjét fejezi ki, a másik koncepció a lokális struktúra, amely a partnerek kapcsolódási szerkezetét számszerűsíti. Így tehát a tudáspotenciál az egyéni jellemzőknek, míg a lokális struktúra a kapcsolódási szerkezetnek felel meg a hálózatban. Ezt a két koncepciót a közvetlen és közvetett partnerek egymást követő „szomszédságaira" alkalmazzuk. Az index koncepciója a következő:

1. Válasszunk egy tetszőleges csúcsot a hálózatban.
2. A csúcs körül képezzünk koncentrikus „szomszédságokat": egy szomszédságba a vizsgált csúcstól adott távolságra lévő csúcsok tartoznak.
3. Minden egyes szomszédságra (az ahhoz tartozó csúcsokra) végezzük el a következőket:
   a. Határozzuk meg és adjuk össze a csúcsok tudásszintjét (tudáspotenciál).
   b. Határozzuk meg a csúcsok kapcsolódási struktúráját és képezzünk egy alkalmas súlyszámot (lokális struktúra).
   c. Számítsuk ki a szomszédság minőségét úgy, hogy a csúcsok tudását (tudáspotenciál) súlyozzuk a kapcsolati szerkezettel (lokális struktúra).
4. Az egyes szomszédságok minőségét súlyozzuk a szomszédságok vizsgált csúcstól vett távolságával és összegezzük az így súlyozott minőségeket.

Az ENQ-index bemutatása során a következő jelöléseket alkalmazzuk. Legyen adott egy N elemű hálózat, amelyben a kapcsolódási struktúrát az $\mathbf{A} = [a_{ij}]$ NxN-es szomszédsági mátrix írja le. A mátrix a_{ij} általános eleme az i és j csúcsok közötti kapcsolatot számszerűsíti. Általánosan a mátrix elemei a kapcsolat intenzitását leíró súlyokként értelmezhetőek, amely súlyokat a 0 és 1 közötti intervallumra normálunk. Egy speciális eset ekkor a bináris kapcsolatmátrix, ahol a mátrix elemei vagy egységnyiek, vagy nullák, attól függően, hogy a két csúcs között van-e kapcsolat vagy sem. Bár a jelen tanulmányban irányítatlan hálózatokat (szimmetrikus kapcsolati mátrixokat) vizsgáljuk, az ENQ-index számítható és értelmezhető irányított hálózatokon is. A szomszédsági mátrix meghatározza a csúcsok közötti legrövidebb utakat tartalmazó $\mathbf{R} = [r_{ij}]$ mátrixot is. Annak érdekében, hogy a hálózat elemeinek tudásszintjét is figyelembe tudjuk venni, bevezetjük továbbá a $\mathbf{k} = [k_i]$ vektort, amely az egyes elemek tudásszintjét tartalmazza.

Az ENQ-index koncepciója alapján a következő formában írhatjuk fel az ENQ értékét a vizsgált i csomópontra:

$$ENQ^i = \sum_{d=1}^{N-1} W_d \cdot LS_d^i \cdot KP_d^i \quad (1)$$

Az (1) egyenletben i jelöli a vizsgált csúcs indexét, d a hálózatban mért távolságot jelenti, N a hálózat mérete, W_d egy súlyfaktor, amelyet az i csúcstól pontosan d távolságra lévő szomszédságra alkalmazunk,[1] míg LS_d^i és KP_d^i rendre az i csúcstól d távolságra lévő szomszédságra kiszámolt tudáspotenciál és lokális struktúra értéke. A továbbiakban a tudáspotenciál és a lokális struktúra meghatározásának módját mutatjuk be.

Tudáspotenciál

A tudáspotenciál, ahogy korábban már bevezettük, a hálózatban elérhető partnerek tudásszintjét tükrözi. Az ENQ-index koncepciójának megfelelően valamennyi szomszédságra értelmezhetjük a tudáspotenciált, amely az adott szomszédsághoz tartozó csúcsok tudásszintjeinek összege:

$$KP_d^i = \sum_{j:r_{ij}=d} k_j \quad (2)$$

Lokális struktúra

A lokális struktúra funkciója, hogy a (2) egyenletben meghatározott tudásszinteket az adott szomszédsághoz tartozó szereplők kapcsolódási szerkezetével súlyozzuk. Ebben az esetben rögtön felmerül a „jó" struktúra kérdése. Az index kiszámításához tudnunk kell, hogy melyek azok a kapcsolódási szerkezetek, amelyek kedvezőek a tudásáramlás szempontjából, vagyis melyek kapnak nagyobb súlyt az indexben.

Egy olyan megközelítést mutatunk be, ami a kapcsolatok intenzitására fókuszál: feltételezzük, hogy minél intenzívebb együttműködési struktúrát találunk a partnerek között, az annál kedvezőbb a vizsgált csúcs szempontjából. Úgy tekintjük, hogy a partnerek közötti együttműködés olyan potenciális új tudás létrejöttét valószínűsíti az adott szomszédságban, amelynek pozitív a hozadéka a vizsgált csúcs számára is. A későbbiekben kitérünk más megközelítési lehetőségekre is.

Specifikusan azt feltételezzük, hogy minden egyes kapcsolat, amely egy adott szomszédságon belül vagy a szomszédságok között létrejön, pozitív hatással van a hálózati tudás minőségére, így az ENQ-indexre. Ezt a következő formulával fejezzük ki:

$$LS_d^i = \frac{1}{N_d^i}\left(\sum_{j:r_{ij}=d-1}\sum_{l:r_{il}=d} a_{jl} + \frac{\sum_{j:r_{ij}=d}\sum_{l:r_{il}=d} a_{jl}}{2}\right) \quad (3)$$

ahol N_d^i az i csúcstól d távolságra lévő csúcsok száma (a szomszédság mérete). A zárójelben lévő kifejezés két tagból áll. Az első tag a kapcsolatok számát (súlyát) adja meg a vizsgált i csúcstól $d-1$ és d távolságra lévő szomszédságok között. A

[1] A távolságsúlyokat úgy definiáljuk, hogy értékük pontosan egységnyi, ha $d = 1$, és szigorú monoton csökken d növekedésével.

kifejezésnek ez a része tükrözi a két szomszédság összekapcsolódásának intenzitását. A második tag a d távolságra lévő csúcsok közötti kapcsolatok számát (súlyát) adja meg.[2] Másként fogalmazva az első tag a szomszédságok közötti, a második tag a szomszédságokon belüli kapcsolatok intenzitását adja meg. Ennek eredményeként a lokális struktúra fenti mutatója azt méri, hogy a d távolságra lévő partnerek milyen erősen kapcsolódnak egymáshoz és az egy lépéssel közelebb lévő partnerekhez.

A (3) formula pontosabb megértéséhez emlékezzünk vissza, hogy a lokális struktúra mutató súlyfaktorként használatos, tehát egy speciális referenciaesetben értéke célszerűen 1-et vesz fel. Egyrészt tegyük fel, hogy az i csúcs N_1^d számú másik csúcshoz kapcsolódik közvetlenül és hogy ezeknek az összekötő kapcsolatoknak a súlya 1. Amennyiben a partnerek nincsenek kapcsolatban egymással, a zárójelben lévő második tag nulla, az első tag pedig pontosan N_1^i (mivel minden kapcsolat súlya 1). Ebben az esetben $LS_1^i = 1$. Ha az összekötő kapcsolatok kisebb intenzitásúak lennének ($a_{ij} < 1$ legalább valamelyik közvetlen szomszéd esetében), akkor a lokális struktúra értéke csökkenne, ami azt fejezi ki, hogy a partnerek tudása nem teljesen hozzáférhető. Másrészt tegyük fel, hogy a partnerek további kapcsolatokat hoznak létre egymás között. Ekkor a zárójeles kifejezés második tagja növekszik, a lokális struktúra értéke nő, ami azt tükrözi, hogy a partnerek erősebb együttműködése kedvező a tudáshoz való hozzáférés, a tudás termelése szempontjából.

A közvetett szomszédságok esetén (vagyis, ha $d > 1$) ez a normalizálás némileg más értelmet kap. A vizsgált i csúcstól d távolságra lévő csúcsok legalább annyi kapcsolaton keresztül kapcsolódnak a megelőző, $d - 1$ távolságra lévő csúcsokhoz, amekkora a d távolságra lévő szomszédság mérete, vagyis legalább N_d^i számú ilyen kapcsolatnak kell léteznie. Következésképpen, ha a kapcsolatok erőssége egységesen 1, akkor a zárójelben lévő első tag értéke a normalizálást elvégezve éppen 1. Vagyis továbbra is igaz, hogy ha a szomszédság tagjait nem köti össze kapcsolat, akkor az LS_d^i mutató értéke akkor egységnyi, ha (1) a szomszédságot az eggyel közelebbi szomszédsághoz a minimális számú kapcsolat köti hozzá és (2) e kapcsolatok intenzitása egységnyi. A kapcsolatintenzitás csökkenése a lokális struktúra értékét csökkenti, további kapcsolatok kialakulása a szomszédságon belül pedig növeli. Az egyetlen különbség, hogy a közvetett szomszédságok esetén a szomszédságokat összekötő kapcsolatok száma nagyobb is lehet, mint N_d^i.

Összefoglalva, a lokális struktúra egy olyan súlyfaktor, amely a kapcsolatok struktúráját ragadja meg. Az itt bemutatott speciális esetben a súlyozás a kapcsolatok intenzitásával (a kapcsolatok számával és azok súlyával) történik. A lokális struktúra mutató azt méri, hogy milyen erősen kapcsolódik egy csúcs a közvetlen partnereihez, mennyire

[2] A kettővel történő osztás szimmetrikus kapcsolatok esetén szükséges, hogy elkerüljük a kapcsolatok dupla beszámítását. A kifejezés első tagja csak a $d - 1$ távolságra lévő csúcsoktól a d távolságra lévő csúcsokhoz „mutató" kapcsolatokat összegzi, visszafelé nem, így itt a kettővel történő osztás nem szükséges.

erősen kapcsolódnak ezek a partnerek további csúcsokhoz és egymáshoz stb. A súlyozás ugyanakkor egy referenciastruktúra köré szerveződik. A lokális struktúra értéke akkor egységnyi, ha a hálózat faszerkezetű a vizsgált csúcs körül, vagyis minden csúcsot egyetlen kapcsolat köt az egy lépéssel megelőző (a vizsgált csomóponthoz közelebb lévő) csúcsok valamelyikéhez és nincsenek átkötő kapcsolatok, továbbá, ha a meglévő kapcsolatok intenzitása egységnyi.

Ego Network Quality

A (2) és (3) formulákkal adott tudáspotenciál és lokális struktúra értékeket az (1) formula alapján összegezzük az ENQ-indexben. A fenti logikát követve tehát az index a közvetlen és a közvetett partnerektől elérhető tudás nagyságát számszerűsíti, és két tényezővel súlyozza ezt a tudást. Egyrészt, a partnerek kapcsolódási szerkezetével, másrészt pedig az egyes tudáselemek távolságával. Minél erősebben kapcsoltak a partnerek és minél közelebb van a tudásuk, annál nagyobb súllyal szerepelnek a számításban. Összességében tehát az ENQ-index megfogalmazható úgy, mint a hálózat szereplőinek tudásszintje a távolsággal és a hálózati szerkezettel súlyozva. A struktúra (3) egyenletben bemutatott operacionalizálása mellett ezt úgy is értelmezhetjük, hogy az ENQ annál nagyobb értéket vesz fel, minél inkább közel helyezkedik a vizsgált csúcs a hálózat sűrű szövésű (intenzív kapcsolódásokkal jellemezhető) és nagy tudást magában foglaló részeihez.

Az ENQ-index értelmezésének speciális esetei a hálózati mutatók rendszerében

A továbbiakban megvizsgáljuk az ENQ-index néhány speciális esetét, amelyek segítenek az index értelmezésében és a különféle hálózati mutatók rendszerében történő elhelyezésében. Az eddigiek alapján megállapíthatjuk, hogy az ENQ-indexben egy csúcs pozícióját három dimenzió mentén értékeljük: a partnerek tudása, a távolsága és a közöttük lévő kapcsolatok szerkezete alapján. Mindhárom dimenzió felfogható egyfajta súlyozásként, attól függően, hogy figyelembe vesszük-e a következőket:

(1) a hálózat csúcsai különböző tudással rendelkeznek, vagyis súlyozzuk-e a csomópontokat a tudásuk alapján;

(2) a csúcsokat milyen kapcsolatrendszer köti össze, vagyis súlyozunk-e a struktúra alapján;

(3) a hálózat csúcsai milyen távol esnek a vizsgált csúcstól, vagyis súlyozunk-e a távolság alapján.

A különböző súlyozások egyes kombinációit foglalja össze az 1. táblázat.

1. táblázat

Az ENQ-index speciális esetei és kapcsolatuk a különféle hálózati mutatókhoz

Special cases of ENQ index and their relationship with different network indicators

Eset	Tudás-súly	Struktúra-súly	Távolság-súly	ENQ-index	Értelmezés
1. Nulleset	×	×	×	$N-1$	A hálózati partnerek száma (hálózat mérete)
2. Tudásszint	✓	×	×	$\sum_{j:j\neq i} k_j$	A hálózatban (partnereknél) található összes tudás
3. Kapcsoltság	×	✓	×	$\sum_j \sum_k a_{jk}/2$	A hálózatban lévő kapcsolatok száma (súlya)
4. Közelség	×	×	✓	$\sum_{j:j\neq i} 1/r_{ij}$	A partnerekhez mért közelségek (inverz távolságok) átlaga
5. Centralitás	×	✓	✓	~ Sajátérték-centralitás	A hálózatban mért centralitás
6. Autokorreláció	✓	×	✓	~ „Hálózati" késleltetés	A partnerek tudásszintjeinek „hálózati" késleltetése

A vizsgált esetek közül kimaradt a tudással és a struktúrával együttesen történő súlyozás. Ennek oka, hogy a bemutatott esetekkel szemben más mutatóhoz ez közvetlenül nem köthető, ugyanakkor a bemutatott esetek intuícióját követve könnyen értelmezhető ez is. A lehetséges kombinációk közül az utolsó értelemszerűen a mindhárom súlyt használó verzió, ezt az előző pont mutatta be.

Nulleset

Ebben az esetben az ENQ-index egy nagyon szélsőséges verzióját kapjuk: sem a tudásszintekben lévő különbségeket, sem a kapcsolatok struktúráját, sem a távolságot nem vesszük figyelembe. Formálisan, minden csúcs tudásszintjét azonosan egységnyinek tekintjük, vagyis $k_i = 1$ minden i-re. Ebből az is következik, hogy a tudáspotenciál értéke valamennyi szomszédságban megegyezik az adott szomszédság méretével: $KP_d^i = \sum_{j:r_{ij}=d} k_j = N_d^i$. A távolságsúlyokat szintén egységnyinek tekintjük minden esetben: $W_d = 1$, minden d-re, továbbá a struktúrával való súlyozást is „kikapcsoljuk", vagyis $LS_d^i = 1$ minden i-re és d-re. Ekkor az ENQ-index értéke a következő:

$$ENQ^i = \sum_{d=1}^{N-1} W_d \cdot LS_d^i \cdot KP_d^i = \sum_{d=1}^{N-1} N_d^i = N-1 \quad (4)$$

Vagyis ebben a speciális esetben az ENQ-index összegzi a hálózatban lévő csúcsok számát, leszámítva a vizsgált csúcsot: a kapott eredmény az összes közvetlen és közvetett partner száma.

Tudásszint

Tegyük fel, hogy figyelembe vesszük a csúcsok tudásszintbeli különbségeit, azonban nem súlyozunk a kapcsolati struktúrával és a távolsággal, ekkor továbbra is fennáll, hogy $W_d = 1$, minden d-re, és $LS_d^i = 1$ minden i-re és d-re. Ugyanakkor most $k_i \in R^+$ minden i-re. Ekkor az ENQ-indexet a következő formában írhatjuk fel:

$$ENQ^i = \sum_{d=1}^{N-1} W_d \cdot LS_d^i \cdot KP_d^i = \sum_{d=1}^{N-1} \sum_{j:r_{ij}=d} k_j = \sum_{j:j\neq i} k_j \qquad (5)$$

Ezek szerint csak a tudássúlyokat „bekapcsolva” az ENQ-index az adott csúcs számára hálózatban elérhető összes tudást számszerűsíti.

Kapcsoltság vagy sűrűség

Ebben az esetben a tudásszintekkel és a távolsággal nem, csak a kapcsolati szerkezettel súlyozunk, vagyis igaz, hogy $k_i = 1$ minden i-re és ebből következően $KP_d^i = \sum_{j:r_{ij}=d} k_j = N_d^i$, továbbá $W_d = 1$, minden d-re. Ugyanakkor a lokális struktúra értékét a (3) egyenletben felírtak szerint határozzuk meg. Az ENQ-index értéke ekkor:

$$ENQ^i = \sum_{d=1}^{N-1} W_d \cdot LS_d^i \cdot KP_d^i = \sum_{d=1}^{N-1} \frac{1}{N_d^i} \cdot \left(\sum_{j:r_{ij}=d-1} \sum_{l:r_{il}=d} a_{jl} + \frac{\sum_{j:r_{ij}=d} \sum_{l:r_{il}=d} a_{jl}}{2} \right) \cdot N_d^i = \sum_{d=1}^{N-1} \sum_{j:r_{ij}=d-1} \sum_{l:r_{il}=d} a_{jl} + \frac{\sum_{d=1}^{N-1} \sum_{j:r_{ij}=d} \sum_{l:r_{il}=d} a_{jl}}{2} = \frac{\sum_j \sum_l a_{jl}}{2} \qquad (6)$$

Az utolsó egyenlőségnél kihasználjuk azt a tényt, hogy a zárójelben lévő első tag a két szomszédságot összekötő kapcsolatokat számszerűsíti, a második tag pedig a szomszédságon belüli tagokat összegzi. Ha ezeket valamennyi szomszédságra összeadjuk, a hálózatban lévő kapcsolatok számát kapjuk. Amennyiben ezt az értéket a lehetséges kapcsolatok számával normáljuk, úgy a hálózatelemzésből ismert sűrűségi mutatót kapjuk eredményül.[3]

Közelség

Tegyük most fel, hogy a tudás- és strukturális súlyokat kikapcsoljuk, tehát $k_i = 1$ minden i-re, és ebből következően $KP_d^i = \sum_{j:r_{ij}=d} k_j = N_d^i$, valamint $LS_d^i = 1$ minden i-re és d-re. A távolsághoz rendelt súlyokról tegyük fel, hogy $W_d = 1/d$. Ekkor az ENQ-index a következőképpen írható fel:

$$ENQ^i = \sum_{d=1}^{N-1} W_d \cdot LS_d^i \cdot KP_d^i = \sum_{d=1}^{N-1} \frac{1}{d} \cdot N_d^i = \sum_{j:j\neq i} \frac{1}{r_{ij}} \qquad (7)$$

A fenti összefüggés alapján az ENQ-index ebben az esetben a vizsgált csúcs többi csúcshoz mért közelségének (amelyet a távolság reciprokaként értelmezünk) összege.

[3] Megjegyzendő, hogy súlyozott kapcsolatok esetén ez a speciális eset a kapcsolati súlyok összegét adja.

Ez pontosan megfelel a hálózatelemzésből ismert közelségcentralitás mutatónak, vagyis ebben a speciális esetben az ENQ-index a vizsgált csúcs többi csúcshoz vett átlagos közelségét adja meg.

Centralitás

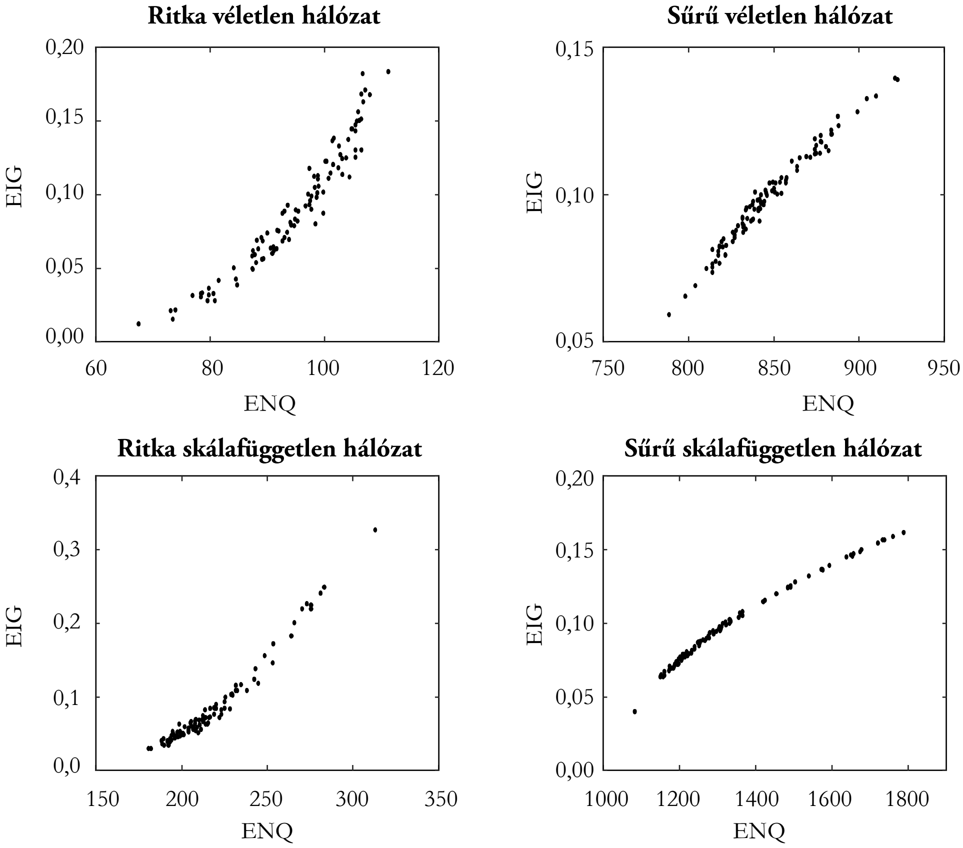
Ebben a speciális esetben használjuk a távolsággal és a struktúrával való súlyozást is, azonban a tudásszinteket egységnyinek tekintjük. A kapcsoltság vagy sűrűség intuíciójának alkalmazásával ebben az esetben egy adott szomszédságra a szomszédságban található kapcsolatok távolsággal súlyozott számát kapjuk, és ezeket az értékeket öszszegezzük. Kicsit más értelmezésben ez úgy is megfogalmazható, hogy az ENQ-index ebben az esetben különböző távolságban lévő partnerek fokszámait összegzi, a távolsággal súlyozva. Ebben a speciális esetben az ENQ-index értéke szoros együttmozgást mutat a hálózatelméletből szintén jól ismert sajátérték-centralitással (lásd: Bonacich, 1972, 2007).

A sajátérték-centralitás (EIG) a következő rekurzív definícióra épül. Legyen x_i az i csúcs centralitását mérő mutató. Tegyük fel, hogy ezt az értéket a közvetlen szomszédok hasonló mutatója határozza meg (vagyis egy csúcs annyira centrális, amennyire a közvetlen szomszédai is azok): $x_i = 1/\lambda \cdot \sum_j a_{ij} \cdot x_j$. Valamennyi csúcsra felírva ez a definíció a következő mátrixegyenletbe írható: $\mathbf{x} = 1/\lambda \cdot \mathbf{A} \cdot \mathbf{x}$, amely egy standard sajátérték-feladat. A legnagyobb sajátértékhez tartozó sajátvektor adja a megfelelő centralitásindexeket. Könnyen látható, hogy ez a megoldás exponenciálisan diszkontálja a távolságot (amennyiben $\lambda > 1$), továbbá ha feltesszük, hogy a közvetlen szomszédok centralitásindexe azonos, akkor a kapott x_i érték arányos az i csúcs fokszámával. Ha feloldjuk a közvetlen szomszédok azonos centralitására vonatkozó feltételt, de fenntartjuk ezt a másodkörös partnerekre vonatkozóan, akkor megmutatható, hogy a centralitás értéke a közvetlen partnerek fokszámainak összege és így tovább. Ez a gondolatmenet természetesen nem bizonyítja formálisan az ENQ-index e speciális esete és a sajátérték-centralitás közötti azonosságot, de rámutat a kettő logikájának hasonlóságára.

1. ábra

Az ENQ-index (vízszintes tengely) és a sajátérték-centralitás (függőleges tengely) korrelációja különböző hálózati struktúrák és sűrűségek mellett

Correlation between ENQ index (horizontal axis) and the own value centrality (vertical axis) by different network structures and densities



Az 1. ábra azt a korábban kifejtett gondolatmenetet szemlélteti, amelyben különböző hálózati szerkezetek: az Erdős-Rényi féle véletlen hálózat (Erdős–Rényi 1959) és a preferenciális kapcsolódáson alapuló skálafüggetlen hálózat (Barabási–Albert 1999), illetve különböző hálózati sűrűségek (ritka és sűrű) mellett szimulált hálózatokra vizsgáltuk meg az ENQ-index és a sajátérték-centralitás korrelációját. Az ered-

mények azt mutatják, hogy ez a korreláció annál erősebb, minél sűrűbb a hálózat és minél inkább jellemző rá a skálafüggetlen struktúra.[4]

Autokorreláció

Bár meglehetősen távoli az analógia, de érdemes kitérni arra az esetre is, amikor a távolság és a tudásszint szerinti súlyozást használjuk, de a strukturális súlyozást kikapcsoljuk, tehát $LS_d^i = 1$ minden i-re és d-re. Tegyük fel, hogy a távolsági súly exponenciális: $W_d = \rho^d$, ahol $0 < \rho < 1$ adott skalár. Ekkor az ENQ-index értékét a következőképpen írhatjuk fel:

$$ENQ^i = \sum_{d=1}^{N-1} W_d \cdot LS_d^i \cdot KP_d^i = \sum_{d=1}^{N-1} \rho^d \cdot \sum_{j:r_{ij}=d} k_j \qquad (8)$$

Definiáljuk az $\mathbf{A}_d$ mátrixot, amely az egyes csúcsoktól d távolságra lévő további csúcsok indikátormátrixa, vagyis $\mathbf{A}_d$ (i,j) cellájában akkor szerepel 1, ha a j csúcs az i csúcstól éppen d távolságra van és 0 egyébként. Legyen továbbá $\mathbf{q}$ az ENQ-indexek vektora valamennyi csúcsra. Ekkor a (8) egyenlet a következő mátrix formában adható meg:

$$\mathbf{q} = \sum_{d=1}^{N-1} \rho^d \cdot \mathbf{A}_d \cdot \mathbf{k} \qquad (9)$$

A fenti összefüggés érdekes hasonlóságot mutat a térbeli autokorrelációt kezelő, térökonometriából ismert SAR-specifikációval. Ez utóbbi alakja (a hibatagokat elhagyva) a következő:

$$\mathbf{y} = \rho \cdot \mathbf{W} \cdot \mathbf{y} + \beta \cdot \mathbf{x} \qquad (10)$$

ahol az $\mathbf{y}$ vektor értékeit magyarázzuk az $\mathbf{x}$ vektor értékeivel, valamint az $\mathbf{y}$ értékek térbeli késleltetésével. A térbeli késleltetést írja le a $\mathbf{W}$ súlymátrix.

Átalakítva és kihasználva a mátrix inverzre vonatkozó felírási szabályt:

$$\mathbf{y} = (\mathbf{I} - \rho \cdot \mathbf{W})^{-1} \cdot \beta \cdot \mathbf{x} = \sum_{d=0}^{\infty} \rho^d \cdot \mathbf{W}^d \cdot \beta \cdot \mathbf{x} \qquad (11)$$

A fenti összefüggés jobb oldala hasonló az ENQ-indexre a (9) egyenletben kapott összefüggéshez. Ez alapján megállapíthatjuk, hogy a strukturális súlyozás hiányában az ENQ-index hasonló jelenséget ragad meg, mint a térbeli autokorreláció, a térbeli szomszédsági viszonyok helyett a hálózati kapcsolatokra építve. Ugyanakkor nagyon fontos felhívni a figyelmet arra, hogy csak hasonlóságról beszélhetünk, ugyanis a (9) és (11) összefüggések két nagyon fontos ponton különböznek. Egyrészt, amíg a (11) egyenletben a $\mathbf{W}$ súlymátrix hatványai szerepelnek, a (9) összefüggés speciálisan képzett $\mathbf{A}_d$ mátrixokkal számol. Az ugyanakkor igaz, hogy mindkettő valamilyen módon a távolság szerepét tartalmazza, de az $\mathbf{A}_d$ mátrix egy adott eleme azt fejezi ki, hogy a két csúcs d távolságra van-e egymástól, a $\mathbf{W}$ súlymátrix megfelelő hatványai pedig azt

[4] Mindegyik ábra 100 elemű hálózatokra készült. A ritka hálózat 5%-os sűrűséget jelent, a sűrű hálózat 30%-os sűrűséget. Az 5%-os érték mögött az a logika húzódik, hogy ez az a határérték, amely mellett egy 100 elemű Erdős-Rényi féle véletlen hálózat már nagy valószínűséggel összefüggővé válik. A 30%-os érték megfelel az európai régiók közötti szabadalmi feltalálói együttműködési hálózat sűrűségének, lásd például Sebestyén és Varga (2013a) tanulmányát.

mutatják meg, hogy két csúcs között mennyi d hosszúságú utat találunk. Az előbbi értelmezés a legrövidebb utakra fókuszál, az utóbbi pedig az összes útra két csúcs között. Másrészt, a (11) egyenlet szummájában megtalálható a $d = 0$ eset is, az ENQ-index viszont ezzel nem számol, ami azt jelenti, hogy az ENQ-index csak a közvetlen és a közvetett hálózati partnerekre fókuszál, a vizsgált csúcs jellemzői kimaradnak, ezzel szemben a térökonometriai megközelítés (már a mátrixhatványokból is következő) „körkörös" visszacsatolásokra épülő logikája a vizsgált csúccsal is számol. Összességében tehát valamennyi hasonlóság feltárható az ENQ-index e speciális esete és a térökonometriából ismert térbeli (vagy hálózati) autokorreláció adatgeneráló folyamata között, ez a hasonlóság azonban nem teljes, számos különbség is jelen van a két megközelítés között.

Az eddig bemutatott szélsőséges esetek rávilágítanak arra, hogy az ENQ-index a csúcsok hálózatban elfoglalt pozícióját fejezi ki, amely a hálózat vizsgált csúcs körüli struktúrájából fakad, figyelembe véve a közvetlen és a közvetett szomszédságokat is. Ebben az értelemben az ENQ tekinthető egyfajta centralitásmutatónak (lásd a 4. és az 5. speciális eseteket). Ugyanakkor az index alkalmas arra is, hogy a csúcsok egyedi jellemzőit, esetünkben tudásszintjeit is figyelembe vegye (2. speciális eset). Ez a dimenzió nem része a tradicionális hálózatelemzési mutatóknak, amelyek jellemzően csak a kapcsolódási struktúrára koncentrálnak. Összefoglalva tehát az ENQ-index egy olyan mérce, amely egyszerre számszerűsíti a hálózati szerkezetből és az egyedi jellemzőkből fakadó hatásokat. Tehát értelmezhető egyrészt úgy, mint a hálózati pozíció mércéje, amely figyelembe veszi a partnerek egyéni jellemzőit is, másrészt úgy is, mint a hálózatból hozzáférhető tudás mércéje, amely figyelembe veszi a hálózat egészének struktúráját is.

A különböző súlyozások szisztematikus összehasonlító elemzése

Az előző fejezetben bemutattuk az ENQ-index néhány speciális esetét, kiemelve, hogy a különböző dimenziók mentén történő súlyozás hogyan jelenik meg az indexben. Ugyanakkor azt is fontos kiemelni, hogy az egyes dimenziók mentén alkalmazott súlyozás megválasztásában jelentős szabadsági fok található. A tudásszint közelítésére alkalmazott mutató megválasztása külön tanulmányt igényelne, és mivel kevéssé kötődik a hálózati jellemzőkhöz, itt ennek megválasztásával nem foglalkozunk.[5] Ugyanakkor fontos kérdés, hogy a másik két súlyt, a távolságot és a hálózati szerkezetet

[5] A szakirodalom számos megoldást kínál a tudásszint becslésére. A leggyakrabban alkalmazott megoldások közé tartozik a szabadalmi bejelentések, vagy az elfogadott szabadalmak száma, az adott időszakban bevezetett új termékek és termelési eljárások száma, a kutatás-fejlesztési (K+F) ráfordítások értéke, a K+F-szektorban foglalkoztatottak száma, vagy az innovációs felmérések alapján becsült innovációk mennyisége (bővebben például Acs, Anselin és Varga (2002) tanulmánya foglalkozik a kérdéssel).

hogyan határozzuk meg. A következőkben e két dimenzió mentén mutatunk be alternatív lehetőségeket, vagyis azt vizsgáljuk, hogy a hálózati szerkezet mérésére alkalmazott feltevések/módszerek hogyan érintik az ENQ-index értékét.

Az elemzési keret: a módosított preferenciális kapcsolódási modell

Az ENQ-index különböző súlyozási szisztémák mellett történő vizsgálatát egy általános hálózati modell keretein belül végezzük el, amely a jól ismert preferenciális kapcsolódás modelljének (Barabási–Albert 1999) egy kiegészített változata. E hálózatelméleti keret alkalmazásának az az oka, hogy a segítségével az ENQ-indexet különböző hálózati szerkezetek mellett tudjuk vizsgálni.

Az alkalmazott hálózati modell segítségével egy folytonos átmeneten keresztül generálhatunk hálózatokat az Erdős-Rényi algoritmusnak megfelelő véletlen topológiától a centralizált struktúráig, amelyben egy erősen kapcsolt központi mag mellett a periferikus szereplők csak a központi szereplőkhöz kapcsolódnak és a periferikus elemek között csak sporadikus kapcsolatok léteznek. E két szélső eset között a modell a preferenciális kapcsolódás alapján kapott skálafüggetlen topológiát adja vissza. A modell részletes leírása megtalálható Sebestyén (2011) munkájában. A modell segítségével olyan hálózati szerkezeteket tudunk vizsgálni, amelyek empirikus szempontból relevánsak. A skála köztes tartományaira jellemző skálafüggetlen topológia számos valós hálózat jellemzője (lásd például Barabási 2003 vagy Csermely 2006), a skála egyik végletét jelentő centralizált szerkezet is több megfigyelt hálózat esetén releváns. A másik végletet jelentő véletlen struktúra kevéssé jellemző a valós hálózatokra, azonban természetes és széles körben használt viszonyítási pontként használatos a hálózati szerkezetek vizsgálata során.

A távolságsúlyok megválasztása

Habár az egy logikus állításnak tűnik, hogy a hálózat távolabbi részeinek tulajdonságai kisebb relevanciával bírnak egy adott csúcs szempontjából, mint a közelebbiek, nyitott kérdés, hogy pontosan mennyivel is kisebb ez a relevancia. Az ENQ-index esetében ez azt jelenti, hogy az (1) egyenletben szereplő W_d távolságsúlyt miként definiáljuk. Általánosságban a távolsági diszkont meghatározása tetszőleges, számos függvényformával feltölthető, amelyek megfelelnek bizonyos alapkövetelményeknek. Ebben az alfejezetben azt elemezzük, hogy a távolsági diszkont megválasztása miként befolyásolja az ENQ-index viselkedését különböző topológiák mellett. Három alapvető függvényformát vizsgálunk: a lineáris, a hiperbolikus és az exponenciális súlyozást.

Lineáris távolsági diszkont esetén feltesszük, hogy egy lépés távolság minden esetben azonos információs vagy tudásbeli veszteséget okoz, függetlenül attól, hogy ez a lépés a kérdéses csúcshoz közel vagy távol következik be. A lineáris súlyozást a következőképpen formalizálhatjuk:

$$W_d = \frac{N-d}{N-1} \qquad (12)$$

Ez a formula rendelkezik a következő, a távolsági diszkonttól elvárt tulajdonságokkal: $d = 1$ esetén az értéke 1, az értéke pedig nullára csökken, ahogy a távolság átlépné a hálózat határait, nevezetesen $d = N$-et. Másképpen, a lehetséges legtávolabbi csúcs, amely legfeljebb $d = N - 1$ lépés távolságra lehet, még kis pozitív súllyal szerepel. Ez a lineáris forma különböző diszkontokat definiál különböző hálózati méretek esetén, azonban kiküszöböli a negatív súlyokat.

A hiperbolikus súlyozást a következő módon definiáljuk:

$$W_d = \frac{1}{d} \qquad (13)$$

Könnyen látható, hogy $d = 1$ esetén ez a forma is 1-et ad, nemnegatív és független a hálózat méretétől. Ezen felül, ahogy korábban megmutattuk, ezt a fajta súlyozást használva az ENQ-index speciális esetben a közelségcentralitást adja. Ebből következően a hiperbolikus súlyozás akkor a megfelelő választás, ha az index ilyen szélsőséges esetekben mutatott tulajdonságait helyezzük előtérbe.

A harmadik lehetőség az exponenciális súlyozás, amelynek egy alkalmas képlete a következő:

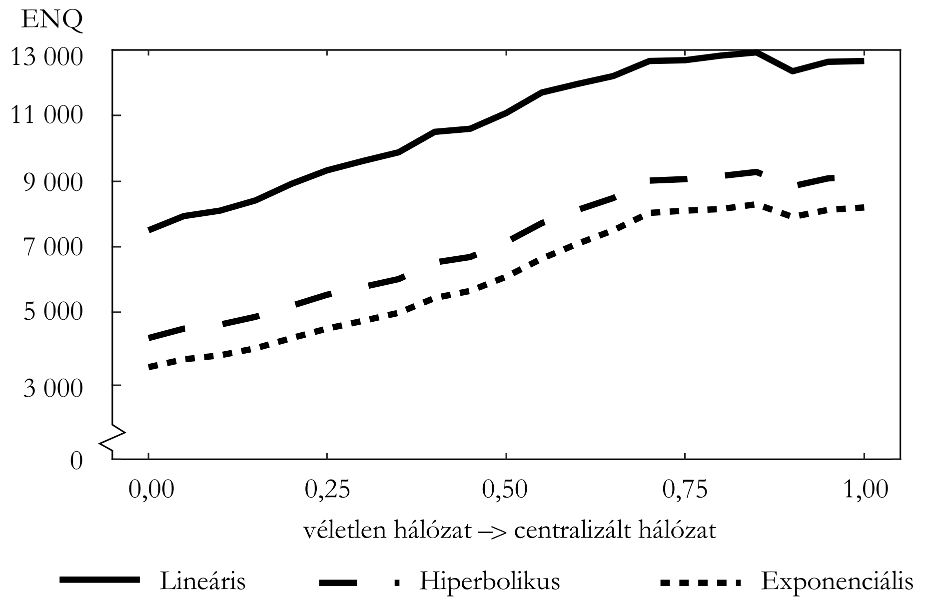
$$W_d = e^{1-d} \qquad (14)$$

Az exponenciális súlyozás hasonló tulajdonságokkal rendelkezik, mint a hiperbolikus, azonban a csökkenés sebessége nagyobb. Ez az eset akkor releváns, ha feltesszük, hogy az információ vagy tudásveszteség a hálózatban mért lépésenként relatív (százalékos) értelemben azonos.

2. ábra

Átlagos ENQ-értékek különböző távolságsúlyok és hálózati topológiák mellett

Average ENQ values at different distance weights and network topologies



A 2. ábrán látható az ENQ alakulása különböző hálózati topológiák és különböző távolságsúlyok esetén. Az ábra elkészítéséhez a következő szimulációs kísérletet végeztük el. A módosított preferenciális kapcsolódási modellnek megfelelően különböző topológiájú hálózatokat generáltunk, összesen 21 lépésben a véletlen hálózattól a centrális struktúráig – ezt mutatja a vízszintes tengely: a 0 értéknél az Erdős-Rényi féle véletlen hálózati topológiát kapjuk, míg az 1 értéknél egy szélsőségesen centralizált, centrum–periféria topológiát. Valamennyi topológia esetén 100 független futtatást végeztünk, hogy a hálózatgeneráló algoritmusban található sztochasztikus hatásokat kiszűrjük, az ábra valamennyi topológiára a 100 futtatás átlagát mutatja – a hálózat mérete minden esetben 100. Valamennyi topológia esetén kiszámítottuk az ENQ-értékeket (mindhárom különböző távolságsúly alapján), majd ezeket átlagoltuk – ezeket az átlagos, a hálózat egészére jellemző ENQ-értékeket mutatja a függőleges tengely.

Elsőként azt látjuk a 2. ábrán, hogy a különböző súlyozási sémák módszeresen különböző ENQ-értékeket adnak. Ez egy logikus eredmény, ha figyelembe vesszük, hogy a különböző súlyozások tulajdonképpen különböző módon értékelik a hálózat távolabbi elemeit, $d = 2$ távolságtól kezdődően a lineáris megoldás adja a legnagyobb súlyt a távolabbi szomszédságoknak, az exponenciális megoldás a legkisebbet, míg a hiperbolikus súlyozás középen helyezkedik el. Az ábra ezt a tendenciát tükrözi, amikor a lineáris súlyozásnál magasabb, a hiperbolikusnál közepes és az exponenciálisnál alacsony ENQ-értékeket látunk: mindössze arról van szó, hogy tetszőleges tudásszintek és struktúrák mellett a lineáris súlyozás magasabb, az exponenciális súlyozás alacsonyabb, a hiperbolikus súlyozás pedig köztes értékeket ad az egyes szomszédságokra számolt részindexeknek.

Emellett azt is látjuk, hogy az ENQ-értékek – függetlenül a súlyozástól – növekednek a magasabb centralizáltság felé haladva. Ezt a jelenséget könnyen megmagyarázhatjuk a módosított preferenciális kapcsolódási modell segítségével. Ahogy a centralizált struktúrák felé haladunk, két jelenséget figyeltünk meg: egyrészt az átlagos távolságok csökkentek a hálózatban, másrészt a klaszterezettség növekedett. Mindkét tendencia az ENQ növekedését eredményezte. Ha rövidebbek a távolságok, közelebb kerülnek egymáshoz a csúcsok, így kevesebb tudás és kapcsolat esik a hálózat távolabbi, erősebben diszkontált részeibe, függetlenül a választott távolságsúlytól. Másrészt az erősebb klaszterezettség azt jelenti, hogy a szomszédságok erősebben kapcsoltak lesznek, vagyis a lokális struktúra mutatója a távolságtól függetlenül növekszik, emelve az ENQ-értékét. Ez a két jelenség erősíti egymást, így a centralizáltabb struktúrák esetében magasabb az ENQ-érték.

Továbbá az is megállapítható, hogy a különböző súlyozások mellett számolt ENQ-értékek közötti különbség nem változik a struktúrával, vagyis a különbség valóban csak a különböző súlyozások matematikai különbségéből fakad. A távolságsúly megválasztása tehát az ENQ-indexet pusztán „eltolja", különböző hálózati topológiák esetén az egyes súlyozási sémák közötti különbség megmarad. Más megközelítésben arra a következtetésre jutunk, hogy az ENQ-index képes tükrözni a hálózati topológiák különbségeit, ugyanakkor a távolságsúlyfüggvény megválasztása – különféle topológiákra alkalmazva – nem torzítja az ENQ-értéket (eltekintve a technikai jellegű eltolódástól). Ez azt is jelenti, hogy a távolságsúlyok megválasztása technikai jellegű, alapvetően az határozhatja meg, hogy az információ vagy tudás távolságon keresztül történő „amortizációját" milyen erősnek tekintjük.

Strukturális lyukak – mi a jó struktúra?

Az eddigiekben az ENQ-index strukturális súlyát a (3) egyenletnek megfelelően használtuk, vagyis azt feltételeztük, hogy a partnerek közötti kapcsolatok pozitívan járulnak hozzá a hálózatból nyerhető tudáshoz. Ez a feltételezés logikusnak tűnik, amenynyiben arra gondolunk, hogy ha egy egyén/szervezet/régió partnerei is közösen dolgoznak valamilyen innováción, akkor az e kapcsolatban létrejövő új tudás, információ hasznos lehet a vizsgált egyén/szervezet számára is. Ezen felül a lokális struktúrára formulája egybevág a társadalmi tőke irodalmából ismert kohéziós megközelítéssel, amelyet többek között Coleman (1986) definiál. Ez a megközelítés azt emeli ki, hogy a zárt társadalmi struktúrák a bizalom felépítésén keresztül elősegítik az egyéni cselekvést. Később azonban Burt (1992) hívta fel a figyelmet arra, hogy a kohézióval szemben az ún. strukturális lyukaknak van jelentős szerepük a társadalmi tőke szempontjából, vagyis olyan speciális helyzeteknek, amikor valaki a hálózat olyan más tagjait köti össze, akik másképpen nem kapcsolódnak. E nézet szerint az információkereskedés (*information brokerage*), vagy másként egyfajta gatekeeper szerep jelenti azt az előnyt, amely e speciális pozícióból származik. Ennek a hálózati szerkezetnek pedig az innovációban kiemelkedő jelentősége lehet. Mint számos példa mutatja, igen sok újdonság keletkezik azáltal, hogy egyes innovációs szereplők (feltalálók, kutatók, innovatív cégek) olyan, egymástól külön fejlődő tudáselemeket kötnek össze, melyek összekapcsolása addig elképzelhetetlen volt, és az innováció pontosan ennek az összekapcsolásnak az eredménye. Ez utóbbi logika tehát az általunk korábban definiált lokális struktúra mutatóval szemben azt állítja, hogy nem az a szomszédság jelent nagyobb értéket, amely szorosan összekapcsolt, hanem amelyik sok, egyébként nem kapcsolódó csoportra bomlik.

Bár a strukturális lyukak koncepciója intuitíve logikus, a mérése egy nagyon fontos nyitott kérdést tartalmaz. Mivel a koncepció egy adott szomszédságban található, össze nem kapcsolt klikkek számát próbálja megragadni, azonnal felvetődik a kérdés, hogy az összekapcsoltság milyen küszöbértékét határozzuk meg akkor, amikor a csomópontok egy csoportját egy klikknek tekintjük (és fordítva, amikor a csoportokat különbözőkként definiáljuk). A pozícióelemzéstől a blokkmodellezésig számos módszer létezik a társadalmi kapcsolatháló elemzésében, amelyek erre a problémára adnak valamilyen megoldást (lásd például Wasserman–Faust, 1994), azonban valamennyiben közös, hogy ezt a küszöböt valamilyen külső, módszertől független döntés határozza meg.

Ugyanakkor az is látható, hogy a kohézió és a strukturális lyukak nem független topológiai jelenségek. Ha egy csúcs valamely szomszédsága sűrűn kapcsolt (kohézív), akkor kicsi a valószínűsége annak, hogy össze nem kapcsolt csoportokat találunk. Fordítva, ahhoz hogy számos össze nem kapcsolt csoportot találjunk, a sűrűségnek (kohéziónak) alacsonynak kell lennie. Ebből következően, nem tudunk egymástól független mércéket kialakítani a két jelenségre.

Mindezeket egybevetve az ENQ-index egy rugalmas lehetőséget biztosít arra, hogy különböző strukturális jellemzőket figyelembe vegyünk. A (3) képletben szereplő lokális struktúra pusztán annyit ír elő, hogy definiáljunk egy alkalmas súlyszámot, amely az egyes szomszédságok szerkezetét értékeli: a kedvező szerkezeteket magasabb, a kevésbé kedvezőbbeket alacsony súllyal. Ebben a keretben csupán egyetlen (a kohéziós logikába illeszkedő) megoldás a (3) egyenletben definiált mérce, amely a sűrűbb kapcsolatokat pozitívan súlyozza. Ugyanígy alkalmazható olyan mérce is, amely a strukturális lyukakat súlyozza pozitívan.

Annak érdekében, hogy a strukturális lyukak szerepét mérhetővé tegyük, egy egyszerű eljárást javaslunk. A fent említett küszöbválasztási problémát egy szélsőséges esettel oldjuk fel, nevezetesen a lokális struktúra mutatóját úgy definiáljuk, hogy az megegyezik az adott szomszédságban, mint részhálózatban található komponensek számával. Bár az összekapcsoltsági küszöb itt is exogén (zérus), egyfajta referenciaesetnek tekinthető. Jelölje az i csúcstól d távolságra található csúcsok által definiált részhálózat kapcsolati mátrixát $\mathbf{A}_d^i$. Definiáljuk e kapcsolati mátrix Laplace-mátrixát:

$$\mathbf{L}_d^i = \mathbf{D}_d^i - \mathbf{A}_d^i \qquad (15)$$

ahol $\mathbf{D}_d^i$ egy olyan diagonális mátrix, amely a főátlón az $\mathbf{A}_d^i$ kapcsolati mátrix által leírt hálózat csúcsainak fokszámát tartalmazza. Határozzuk meg a Laplace-mátrix sajátértékeit:

$$\mathbf{v}_d^i = [v_i] = eig\left(\mathbf{L}_d^i\right) \qquad (16)$$

Ismert, hogy a vizsgált hálózatban található komponensek számát a Laplace-mátrix zérus sajátértékeinek száma adja meg (lásd például Godsil–Royle 2001). Ezek alapján a lokális struktúrára a következő alternatív mutatót vezethetjük be:

$$LS_d^i = \|\{v_i\}|v_i = 0\| \qquad (17)$$

ahol $\|X\|$ az X halmaz számosságát jelöli. Ez a formula azonban a lokális struktúra mutatóját a (3) egyenlethez képest egy másik végletbe mozdítja. Amíg a (3) definíció szerint az a legtöbbet érő szomszédság, amelyik a lehető legtöbb kapcsolattal rendelkezik, a (13) definíció szerint éppen ellenkezőleg, az a legjobb szomszédság, amelyben egyetlen kapcsolat sincsen. A két megközelítést kombinálva azonban az ENQ-index képes arra is, hogy egy lényegesen finomabb képet közvetítsen az egyes szomszédságok struktúrájáról. Jelölje a (3) egyenlet szerinti strukturális súlyt LC_d^i (Local Connectivity – helyi kapcsoltság), míg a (13) egyenlet szerint strukturális súlyt CC_d^i

(Connected Components – kapcsolt komponensek). Egy lehetséges harmadik formája a struktúra súlyozásának e kettő kombinációja:

$$LS_d^i = LC_d^i \cdot CC_d^i \qquad (18)$$

Mivel a hálózati szerkezettel foglalkozó szakirodalom nem egységes abban a tekintetben, hogy a kohézív vagy a strukturális lyukakat tartalmazó szerkezet kedvezőbb-e a hálózati csúcsok teljesítménye szempontjából, logikusnak tűnik azt feltételezni, hogy sem a nagyon összekapcsolt, sem a nagyon ritka szerkezetek nem a legkedvezőbbek, hanem a köztes topológiák jelentik a legnagyobb hozzáadott értéket. Másként fogalmazva, egy tudáshálózat szempontjából a partnerek közötti szoros kapcsoltság kedvező, mert az ott áramló információ, ötlet kedvező hatással lehet a vizsgált szereplő számára is. Ugyanakkor a különböző csoportokhoz való hozzáférés is kedvező, hiszen ezáltal különböző típusú tudáselemekhez lehet hozzájutni a hálózati kapcsolatokon keresztül, ami szintén pozitív hatással lehet az új tudás létrehozására. A kapcsolatok sűrűségének növekedésével az a veszély áll fenn, hogy a partnerek ugyanazt fogják tudni (tanulnak egymástól), így a diverzitás elvész. Ugyanakkor az egyik típusú struktúra – a korábban említettek szerint – a másik rovására erősíthető, így logikusan egyfajta köztes megoldás tűnik optimálisnak, amely tartalmazza a kohéziót, de nem extrém módon, hanem úgy, hogy a hozzáférhető tudás sokfélesége is megmaradhasson. A (14) egyenletben adott forma éppen ezeket a köztes eseteket súlyozza felül a szélsőséges esetekkel szemben.

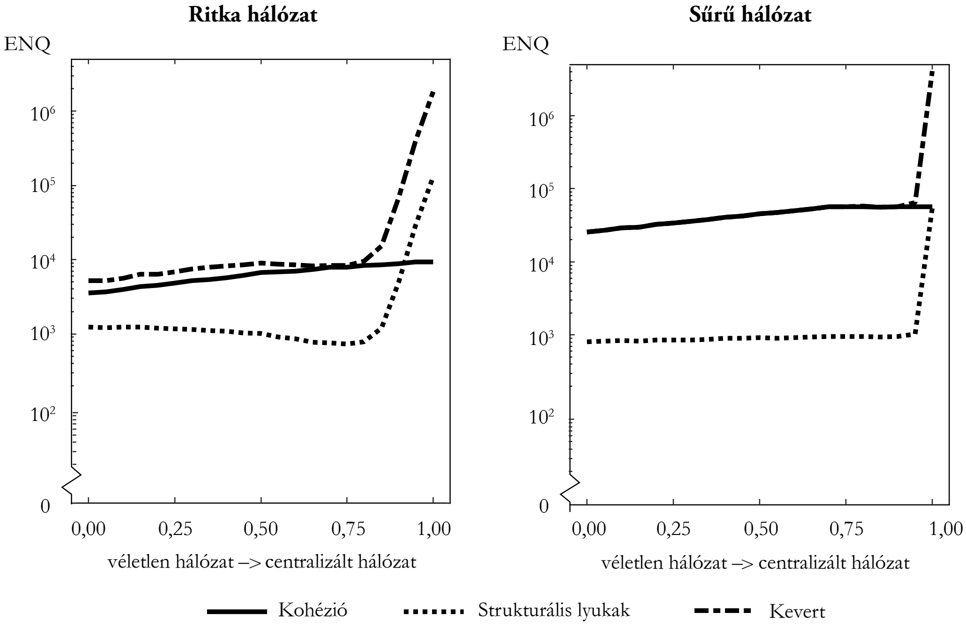
Annak érdekében, hogy a hálózati topológia súlyozásának megválasztását vizsgáljuk, a távolságsúlyok esetében alkalmazott elemzéshez hasonló kísérletet mutatunk be. Ebben a módosított preferenciális kapcsolódási modellt használjuk ismét, és most azt vizsgáljuk, hogy a strukturális súly megválasztása milyen hatással van az ENQ-index értékére különböző topológiák mellett. Három súlyozást hasonlítunk össze: a kohéziós gondolaton alapuló súlyt a (3) egyenlet szerint, a strukturális lyukak alapján képzett súlyt a (13) egyenlet szerint és a kevert megoldást a (14) egyenlet szerint. A szimulációs vizsgálat eredményeit mutatja a 3. ábra, ritka és sűrű hálózatokra.[6]

[6] A szimulációkat ebben az esetben 500 elemű hálózatokon végeztük, mivel nagyobb hálózatokon a megfigyelhető tendenciák markánsabban láthatóak. További szimulációk megmutatták, hogy az eredmények robosztusak a hálózat méretére, kisebb hálózatokon minőségileg azonos eredményeket kaptunk. A korábbiaknak megfelelően a ritka hálózat 5, a sűrű 30%-os sűrűséget jelent. Valamennyi topológiára 100 független szimulációt végeztünk és ezek eredményeit átlagoltuk. A függőleges tengely logaritmikus skálát használ a tendenciák jobb láthatósága érdekében. Az ENQ-indexet hiperbolikus súlyozás mellett számoltuk ki.

3. ábra

Átlagos ENQ-értékek különböző strukturális súlyok és hálózati topológiák mellett

Average ENQ values at different structural weights and network topologies



Az eredmények azt mutatják, hogy a ritka és a sűrű hálózaton kapott eredmények minőségileg hasonlóak, de a megfigyelt tendenciák élesebben kijönnek a sűrű hálózaton. A fekete folytonos vonalak mutatják a kohéziós megközelítés mellett kapott ENQ-értékeket, ezek megegyeznek a 2. ábrán a hiperbolikus súlyozás mellett kapott értékekkel.

Elsőként azt látjuk, hogy amikor csak a komponensek számával súlyozunk (szaggatott vonal), az ENQ-értékek tipikusan alacsonyabbak a topológiák jelentős részére. Ezt az a tény magyarázza, hogy a lokális struktúra értékek tipikusan alacsonyabbak ebben az esetben, mivel ez a komponensek számát mutatja, míg a kohéziós megközelítésben a kapcsolatok számával súlyozunk, utóbbiból lényegesen több van. A szélsőségesen centralizált topológiák esetén (például csillagszerkezet) ez a tendencia megfordul, a komponensekre kapott értékek lesznek magasabbak. Ez a tendencia is könnyen megérthető, ha arra gondolunk, hogy a szélsőségesen centralizált hálózatban a szomszédságok tipikusan sok komponensből állnak (a csillaghálózat középpontja egymással nem kapcsolt elemekkel áll összeköttetésben), így a CC_d^i érték ezekben a hálózatokban tipikusan magas. Ugyanakkor kis elmozdulás e topológiától a véletlen hálózatok irányába számottevő csökkenést mutat a komponensek számában, vagyis a szomszédságok hamar elkezdenek összefüggővé válni, jelentős mértékben csökkentve a különálló komponensek számát.

A második megfigyelés az, hogy a topológiák nagyobb hányadán a kohéziós és a kevert megközelítés azonos ENQ-értékekhez vezet. Ez abból következik, hogy e topológiák mellett a szomszédságokban található különálló komponensek száma jellemzően kicsi, egyhez közeli, így a kevert esetben a CC_d^i értékekkel történő addicionális súlyozás nem vezet szignifikáns eltérésekhez a sima, kohéziós elvű súlyozáshoz képest. Ez a jelenség azonban a jelentősen centralizált topológiák esetén már nem jelentkezik, mivel esetünkben a komponensek száma az egyes szomszédságokban tipikusan és jelentősen eltér egytől. Ezen felül a sűrű hálózat esetében ez az egyezőség a kohéziós és a kevert megközelítések között nagyon erős, tulajdonképpen hiba nélküli, aminek az a magyarázata, hogy a nagyobb sűrűség, vagyis a nagyobb valószínűséggel megjelenő szomszédságon belüli kapcsolatok kizárják annak a lehetőségét, hogy különálló komponenseket figyeljünk meg, míg néhány ritka hálózatban ezek nagyobb valószínűséggel találhatók meg.

Összefoglalva megállapíthatjuk, hogy egyrészt a kohéziós és a kevert súlyozás között nincs számottevő különbség a módosított preferenciális kapcsolódás modelljében adódó legtöbb hálózati topológia esetén, kivéve a szélsőségesen centrális, csillagtípusú topológiákat. Másrészt viszont az a kérdés, hogy a tisztán kohéziós, a tisztán strukturális lyukakra építő vagy a kevert megoldás jelent-e „jobb" hálózati szerkezetet továbbra is megválaszolásra vár, ami empirikus alátámasztást igényel. Az ENQ-index azonban rugalmasan tudja kezelni bármelyik esetet, aminek következtében arra is alkalmas, hogy a releváns, hatékony struktúrákat a segítségével empirikusan teszteljük.

Összefoglalás

Tanulmányunkban az Ego Network Quality indexet mutattuk be, amely azt a célt szolgálja, hogy egy kompozit mércében számszerűsítse a tudáshálózatokból hozzáférhető tudás értékét, figyelembe véve mind a közvetlen, mind a közvetett partnereket a hálózatban. Az index a hálózati pozíció három aspektusát integrálja. Egyrészt figyelembe veszi a közvetlen és közvetett partnerek tudásszintjét, amely újdonság a standard hálózati mutatókkal szemben. Másrészt a tudásszinteket a partnerek közötti kapcsolatrendszer szerkezetét leíró mutatószámmal súlyozza, így a hálózati szerkezet figyelembevételét is lehetővé teszi. Harmadrészt nemcsak a közvetlen, hanem a közvetett szomszédok tudásszintje és kapcsolódási struktúráját is számba veszi.

A hálózatok tudástermelésre gyakorolt hatását vizsgáló kutatások vagy a hálózati szerkezetre (tér nélküli vizsgálatok), vagy a partnerek tudásszintjére, egyéb jellemzőire (térbeli vizsgálatok) fókuszáltak, azonban nem integrálták a két megközelítést. Ebben a tanulmányban azt mutattuk meg, hogy az ENQ-index alkalmas mindkét tényező megjelenítésére.

Az ENQ-index a vizsgált csúcs körül a távolság alapján egymást követő szomszédságokra osztja a hálózatot, majd az egyes szomszédságokból elérhető tudásszintet súlyozza a szomszédságban megfigyelt kapcsolatok szerkezetével és a szomszédság távolságával. Tanulmányunkban a különböző súlyozási sémák szerepét vizsgáltuk meg. A távolság esetén három különböző diszkontfüggvényt vizsgáltunk (lineáris, hiperbolikus és exponenciális), a struktúra esetén pedig, a társadalmi tőkéből merített intuícióra építve szintén három különböző variációt elemeztünk: a kohézióra (intenzív kapcsolódásra), a strukturális lyukakra (összekötő szerepre) fókuszáló eseteket, valamint ezek kombinációját.

Megmutattuk, hogy speciális esetekben az ENQ-index jól értelmezhető, adott esetben ismert hálózati mutatókra vezethető vissza, majd szimulációs vizsgálatokkal támasztottuk alá, hogy az ENQ-index valóban alkalmas a hálózati szerkezet bizonyos tulajdonságainak kimutatására anélkül, hogy a különböző súlyozási sémák alkalmazása jelentősebb torzítást okozna, kifejezetten a hálózati topológiák egy empirikus szempontból releváns skáláján. Azt is megmutattuk, hogy az ENQ-index rugalmasan kezeli a tudástermelés szempontjából kedvező hálózati szerkezet különböző definícióit.

Fel kell hívnunk azonban a figyelmet mind az ENQ-index, mind bemutatott elemzésének korlátaira is. Egyrészt, a szimulációs elemzések kereteként használt módosított preferenciális kapcsolódási modell csupán a hálózati topológiák egy korlátozott skáláját fedi le. Bár amellett érveltünk, hogy ez a skála releváns a valós hálózatok esetében, természetesen a modell által alkalmazott struktúráktól eltérő, komplexebb szerkezetek is elképzelhetőek. Az ilyen szerkezetekre értelemszerűen nem feltétlenül érvényesek a tanulmány megállapításai. Másrészt, az elemzés az ENQ-index átlagos értékeire fókuszált, különböző súlyozási sémák és topológiák mellett az index egyedi, csúcsszintű jellemzőit nem vizsgáltuk, ugyanakkor fontos megjegyezni, hogy kifejezetten a centralizált hálózati topológiákon, ahol a csúcsok markánsan központi és periferikus csoportokba sorolódnak, az ENQ-indexek különbségei jelentősek lehetnek az egyes csúcsok szintjén.

A bemutatott index, bár elsősorban tudáshálózatok vizsgálatára lett kifejlesztve és alkalmazva, rugalmassága okán más típusú hálózatok vizsgálatára is használható. A hálózati elemek egyedi jellemzőinek és az őket összekötő kapcsolatrendszer integrált vizsgálata és számszerűsítése több területen is hasznos lehet. Bár ebben a tanulmányban az egyedi jellemzőnek a tudásszintet tekintettük, tulajdonképpen bármelyik jellemzőt alkalmazhattuk volna. Ezen felül, a strukturális súlyozás rugalmassága arra is lehetőséget teremt, hogy az adott vizsgálat szempontjából releváns szerkezeti felépítést értékeljünk az indexszel.

Köszönetnyilvánítás

A jelen tudományos közleményt a szerzők a Pécsi Tudományegyetem alapításának 650. évfordulója emlékének szentelik. A kutatást az OTKA (OTKA K101160) támogatta. Ezúton is köszönjük Nicolas Carayol, Claude Raynaut és Mario Maggioni értékes megjegyzéseit a tanulmány korábbi változataihoz.

IRODALOM

ACS, Z. – ANSELIN, L. – VARGA, A. (2002): Patents and innovation counts as measures of regional production of new knowledge *Research Policy* 31 (7): 1069–1085.

ANSELIN, L. – VARGA, A. – ACS, Z. (1997): Local geographic spillovers between university research and high technology innovations *Journal of Urban Economics* 42 (3): 422–448.

BARABÁSI, A. L. – ALBERT, R. (1999): Emergence of scaling in random networks *Science* 286: 509–512.

BARABÁSI, A. L. (2003): *Linked: How Everything is Connected to Everything Else what It Means for Business, Science and Everyday Life* Penguing Group, USA.

BONACICH, P. (1972): Factoring and weighting approaches to clique identification *Journal of Mathematical Sociology* 2 (1): 113–120.

BONACICH, P. (2007): Some unique properties of eigenvector centrality *Social Networks* 29 (4): 555–564.

BURT, R. S. (1992): *Structural Holes* Harvard University Press. Cambridge, MA.

BURT, R. S. – HOGARTH, R. M. – MICHAUD, C. (2000): The Social Capital of French and American Managers *Organization Science* 11 (2):123–147.

BURTON, P. – YU, W. – PRYBUTOK, V. (2010): Social network position and its relationship to performance of IT professionals *Informing Science: the International Journal of an Emerging Transdiscipline* 13: 121–137.

CAINELLI, G. – MAGGIONI, M. – UBERTI, E. – DE FELICE, A. (2010): The strength of strong ties: co-authorship and productivity among Italian economists *'Marco Fanno' working papers* 125, Department of Economics, University of Padova.

COLEMAN, J. S. (1986): Social Theory, Social Research, and a Theory of Action *American Journal of Sociology* 91 (6): 1309–1335.

CROSS, R. – CUMMINGS, J. N. (2004): Tie and Network Correlates of Individual Performance in Knowledge-Intensive Work *The Academy of Management Journal* 47 (6): 928–937.

CSERMELY, P. (2006): *Weak Links: Stabilizers of Complex Systems from Proteins to Social Networks.* Springer-Verlag, Berlin.

DIEZ, R. (2002): Metropolitan Innovation Systems - A comparison between Barcelona, Stockholm, and Vienna *International Regional Science Review* 25 (1): 63–85.

DONCKELS, R. – LAMBRECHT, J. (1997): The Network Position of Small Businesses: An Explanatory Model *Journal of Small Business Management* 35 (2): 13–26.

ERDŐS, P. – RÉNYI, A. (1959): On Random Graphs I. *Publ. Math* 6: 290–297.

FISCHER, M. – VARGA, A. (2002): Technological innovation and interfirm cooperation. An exploratory analysis using survey data from manufacturing firms in the metropolitan region of Vienna *International Journal of Technology Management* 24 (7-8): 724–742.

GODSIL, C. – ROYLE, G. F. (2001): *Algebraic Graph Theory* Springer-Verlag, New York.

HAU-HORVÁTH, O. – SEBESTYÉN, T. – VARGA, A. (2016): Tudáshálózatok szerepe a regionális fejlődésben – egy integrált modell alkalmazásának tapasztalatai a magyar régiók esetében *Statisztikai Szemle* (megjelenés alatt).

HOEKMAN J. – FRENKEN, K. – VAN OORT, F. (2009): Collaboration networks as carriers of knowledge spillovers: evidence from EU27 regions *Annals of Regional Science* 43 (3): 721–738.

HOPP, W. J. – IRAVANI, S. – LIU, F. – STRINGER, M.J. (2010): *The Impact of Discussion, Awareness, and Collaboration Network Position on Research Performance of Engineering School Faculty* Ross School of Business Paper No. 1164.

JONES, C. (1995): R&D-based models of economic growth, *Journal of Political Economy* 103 (4): 759–784.

KRETSCHMER, H. (2004): Author productivity and geodesic distance in bibliographic co-authorship networks , and visibility on the Web. *Scientometrics* 60 (3): 409–420.

LUNDVALL, B. A. (1992): *National Systems of Innovation* Pinter Publishers, London.

MAGGIONI M. – NOSVELLI M. – UBERTI, E. (2007): Space versus networks in the geography of innovation: a European analysis *Papers in Regional Science* 86 (3): 471–493.

MAGGIONI, M. – UBERTI, T. (2011): Networks and geography in the economics of knowledge flows *Quality and Quantity* 45 (5): 1031–1051.

NELSON, R. R. (ed.) (1993): *National Innovation Systems: A Comparative Analysis* Oxford University Press, Oxford.

OECD (2009): *REGPAT Database* October 2009.

PONDS, R. – VAN OORT, F. – FRENKEN, K. (2009): Internationalization and regional embedding of scientific research in the Netherlands In: Varga, A. (ed): *Universities, knowledge transfer and regional development: geography, entrepreneurship and policy* pp 109–137., Edward Elgar Publishers, Northampton.

PONDS, R. – VAN OORT, F. – FRENKEN, K. (2010) Innovation, spillovers and university--industry collaboration: an extended knowledge production function approach *Journal of Economic Geography* 10 (2): 231–255.

POWELL, W. W. – KOPUT, K. W. – SMITH-DOERR, L. – OWEN-SMITH, J. (1999): Network Position and Firm Performance: Organizational Returns to Collaboration in the Biotechnology Industry *Research in the Sociology of Organizations* 16 (1): 129–159.

ROMER, P. M. (1990): Endogenous technological change *Journal of Political Economy* 5 (98): S71–S102.

RUMSEY-WAIREPO, A. (2006): *The association between co-authorship network structures and successful academic publishing among higher education scholars* Brigham Young University, Provo UT.

SALMENKAITA, J. P. (2004): Intangible capital in industrial research: Effects of network position on individual inventive productivity In Bettis, R. (Ed.): *Strategy in transition* pp. 220–248., Blackwell Publishing, Malden, MA.

SEBESTYÉN, T. (2011): *Knowledge Networks and Economic Performance. Approaches for Modeling and Empirical Analysis* VDM Verlag, Saarbrücken.

SEBESTYÉN, T. – VARGA, A. (2013a): Research productivity and the quality of interregional knowledge networks *Annals of Regional Science* 51 (1): 155–189.

SEBESTYÉN, T. – VARGA, A. (2013b): A novel comprehensive index of network position and node characteristics in knowledge networks: Ego Network Quality In: Scherngell,

T. (Ed.) *The geography of networks and R&D collaborations* pp. 71–97.,Springer, New York.

TSAI, W. (2001): Knowledge Transfer in Intraorganizational Networks: Effects of Network Position and Absorptive Capacity on Business Unit Innovation and Performance *The Academy of Management Journal* 44 (5): 996–1004.

VAN DER DEIJL H. – KELCHTERMANS S. – VEUGELERS, R. (2011): Researcher networks and productivity Paper presented at the DIME-DRUID ACADEMY winter conference, Aalborg, Denmark.

VARGA, A. (2000): Local academic knowledge transfers and the concentration of economic activity *Journal of Regional Science* 40 (2): 289–309.

VARGA, A. (2006): The spatial dimension of innovation and growth: empirical research methodology and policy analysis *European Planning Studies* 14 (9): 1171–1186.

VARGA, A. – PONTIKAKIS, D. – CHORAFAKIS, G. (2014): Metropolitan Edison and cosmopolitan Pasteur? Agglomeration and interregional research network effects on European R&D productivity *Journal of Economic Geography* 14: 229–263.

VARGA, A. – SEBESTYÉN, T. (2015): Innováció Kelet-közép Európában: Az EU Keretprog-ramokban való részvétel szerepe az innovációs teljesítményben *Közgazdasági Szemle* 62 (9): 881–908.

VARGA, A. – SEBESTYÉN, T. (2016): Does EU Framework Program participation affect regional innovation? The differentiating role of economic development *International Regional Science Review* (megjelenés alatt).

WASSERMAN, S. – FAUST, K. (1994): *Social Network Analysis – Methods and Application.* Cambridge University Press, Cambridge, UK.

WATTS, D.J. – STROGATZ, S. H. (1998): Collective dynamics of 'small-world' networks *Nature* 393 (6684): 409–410.

ZAHEER, A. – BELL, G. G. (2005): Benefiting from network position: firm capabilities, structural holes and performance *Strategic Management Journal* 26 (9): 809–825.