KÖZGAZDASÁGI SZEMLE, LX. ÉVF., 2013. DECEMBER (1333–1353. o.)

CSATÓ LÁSZLÓ

Rangsorolás páros összehasonlításokkal

Kiegészítések a felvételizői preferencia-sorrendek módszertanához

A Közgazdasági Szemle márciusi számában *Telcs és szerzőtársai* [2013] a felvételizők preferenciái alapján új megközelítést javasolt a felsőoktatási intézmények rangsorolására. Az alábbi írás új szempontokat biztosít ezen alapötlet gyakorlati megvalósításához. Megmutatja, hogy az alkalmazott modell ekvivalens az alternatívák egy aggregált páros összehasonlítási mátrix révén végzett rangsorolásával, ami rávilágít a szerzők kiinduló hipotéziseinek vitatható pontjaira. A szerző röviden áttekinti a hasonló feladatok megoldására javasolt módszereket, különös tekintettel azok axiomatikus megalapozására, majd megvizsgálja a *Telcs és szerzőtársai* [2013] által alkalmazott eljárásokat. Végül említést tesz egy hasonló megközelítéssel élő cikkről, és megfogalmaz néhány, a vizsgálat továbbfejlesztésére vonatkozó javaslatot.*
Journal of Economic Literature (JEL) kód: C44, O71.

A Közgazdasági Szemle 2013. márciusi számában Telcs András, Kosztyán Zsolt és Török Ádám egy tanulmányt közölt Hallgatói preferencia-sorrendek készítése az egyetemi jelentkezések alapján címmel (*Telcs és szerzőtársai* [2013]). A cikk kiindulópontjának – a felsőoktatási intézmények felvételi jelentkezésekből meghatározott páros összehasonlítások alapján történő rangsorolásának – megértése után első gondolatom az volt, milyen szerencse egy ilyen innovatív megközelítéssel találkozni. Közgazdászkörökben tartott előadásaim során már többször szembesültem azzal a váddal, hogy az általam választott kutatási téma ugyan matematikailag szép, de nem igazán kapcsolódik a tudományterület fő kérdéseihez. Ennek cáfolatára folyamatosan keresem az alkalmazási lehetőségeket, és *Telcs és szerzőtársai* [2013] ötlete, amivel korábban nem találkoztam, valóban kiválónak ígérkezett.

* A kutatás a TÁMOP 4.2.4.A/1-11-1-2012-0001 sz. Nemzeti Kiválóság Program – Hazai hallgatói, illetve kutatói személyi támogatást biztosító rendszer kidolgozása és működtetése országos program című kiemelt projekt által nyújtott személyi támogatással valósult meg. A projekt az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósul meg. A kutatás szakmailag kapcsolódik az OTKA K-77420 sz. projekthez. Köszönet illeti az anonim bírálót tanácsaiért és a szerkesztőt rugalmasságáért. Természetesen a leírtakkal kapcsolatos minden felelősség az enyém.

Csató László, Budapesti Corvinus Egyetem Operációkutatás és Aktuáriustudományok Tanszék.

A cikk végigolvasása után azonban némi hiányérzetem támadt. A szerzők egyik megjegyzése szerint ez „egy hosszabbra tervezett cikksorozat első eleme" (292. o.), így éppen emiatt tartottam volna fontosnak a módszertani háttér alapos, hivatkozásokkal bőségesen alátámasztott tárgyalását, külön kitérve más, szintén jelentkezői preferenciákon alapuló vizsgálatok főbb hipotéziseinek áttekintésére.

Ebben a vitairatban – némileg esszéisztikus jelleggel – a *Telcs és szerzőtársai* [2013] cikkel kapcsolatos gondolataimat, módszertani kiegészítéseimet kívánom megosztani az olvasókkal. Célom egyrészt az, hogy a páros összehasonlításokon alapuló rangsorolás módszertani áttekintésével segítsem a hasonló problémákkal szembesülő, a matematikai hátteret esetleg kevésbé ismerő társadalomtudományi kollégák munkáját, másrészt felhívjam a hozzám hasonló doktoranduszok, fiatal kutatók figyelmét a szakirodalmi megalapozottság, a matematikai szabatosság és az elméleti háttér ismeretének fontosságára. Emiatt talán többször úgy tűnhet, hogy egyes megjegyzéseim túlságosan kritikusak, kötekedők, kevéssé kapcsolódnak az elemzett tanulmány alapvető mondanivalójához. Semmiképpen sem állt szándékomban, hogy a tisztelt szerzők szakmai érdemeit vagy az általuk végzett kutatás erényeit kisebbítsem; azonban úgy vélem, célszerűbb a hibalehetőségekre rámutató megjegyzésekkel segítséget nyújtani, mint a dicsérő jelzőket sorolni.

Először az egyéni preferenciák aggregálásának, illetve az ezek alapján történő rangsorolás kérdését vizsgálom, amit a *Telcs és szerzőtársai* [2013] által javasolt módszerek hosszabb, illetve a tanulmány többi megállapításának rövid kritikája követ. Ezután bemutatom az általam feltárt, a cikkben kevésbé hangsúlyos nemzetközi előzményeket, majd *Telcs és szerzőtársai* [2013] legfontosabb következtetéseit tárgyalom. Végül összefoglalom a tanulmánnyal kapcsolatos meglátásaimat, javaslatot teszek annak továbbfejlesztésére, és megfogalmazok néhány szubjektív gondolatot a vizsgált tanulmány hivatkozásairól, nemzetközi beágyazottságáról.

Módszertani háttér – rangsorolás páros összehasonlításokkal

A szerzők a 2001–2011 közötti, a http://felvi.hu honlap által nyilvántartott felsőoktatási jelentkezéseket tartalmazó adatbázist használták. Emiatt szigorú értelemben véve nem beszélhetünk *hallgatói* jelentkezésekről, a felvételi folyamat résztvevőit sokkal találóbb *felvételizőnek* hívni, hiszen nem mindegyikükből lesz egyetemi hallgató. A felhasznált adatbázis kétségtelenül impozáns: minden év több mint 400 ezer rekordot és 100 ezer jelentkezést tartalmaz.[1]

[1] Ezzel kapcsolatban legfeljebb az a kérdés merülhet fel, hogy a hibás kitöltés miatt az adatbázisból kimaradó, de egyébként értelmezhető preferenciák módosítanának-e a képen; a papíralapú dokumentumok feltételezhetően nagyobb eséllyel tévesek, mint az elektronikusan beadottak, a jelentkezés módja pedig minden bizonnyal nem független a felvételizők társadalmi hátterétől.

A páros összehasonlítási mátrix megadása

Telcs és szerzőtársai [2013] a preferenciarendezések felállítása során minden egyes jelentkezési lap alapján egy irányított gráfot definiál a következő módon:[2]

1. a gráf csúcsai a jelentkezési lapon szereplő objektumok;
2. az i-edik objektum pontosan akkor preferált a j-edikkel szemben, ha a gráfban az i-edik csúcsból egy irányított él fut a j-edikkel csúcsba;
3. s jelentkezési lapon előrébb szereplő objektum minden hátrébb levővel szemben preferált;
4. a nem megjelölt objektumok egyenrangúak;
5. a megjelölt objektumok preferáltak az összes kihagyotthoz képest.

Telcs és szerzőtársai [2013] többször hangsúlyozza, hogy a gráfreprezentáció redundanciamentesen és információveszteség nélkül tárolja a jelentkezések alapján kialakuló preferenciákra vonatkozó adatokat, ami kétségtelenül igaz, de nem tekinthető önálló eredménynek. Az első három elv elfogadása szinte egyértelműnek tűnik, annál inkább kérdéses lehet a negyedik és az ötödik. Ehhez azonban először némi kitérőt szükséges tennünk a páros összehasonlításokon alapuló rangsorolás irodalmába.

Alternatívák közötti választáskor gyakran merülnek fel hasonló, az egyéni döntéshozók preferenciáira vonatkozó kérdések. Itt többnyire négy lehetőség áll a p-edik szavazó rendelkezésére az i-edik és j-edik alternatíva közötti viszony megadásakor, amelyeket számokkal „kódolhatunk” az $\mathbf{M}^{(p)}$ egyéni páros összehasonlítási mátrixban (lásd például *Chebotarev–Shamis* [1999]):[3]

- az i-edik jobb a j-ediknél, $m_{ij}^{(p)} = 1$ és $m_{ji}^{(p)} = 0$;
- az i-edik ekvivalens a j-edikkel, $m_{ij}^{(p)} = 0{,}5$ és $m_{ji}^{(p)} = 0{,}5$;
- a döntéshozó nem nyilvánít véleményt – itt a transzformáció nem egyértelmű, legszerencsésebbnek az $m_{ij}^{(p)} = m_{ji}^{(p)} = 0$ választást tartom;
- a elfődiagonálisban szereplő $m_{ii}^{(p)}$ elemeknek nincs jelentőségük, én kényelmi okokból az $m_{ii}^{(p)} = 0$ megoldással élek.

Vegyük észre, hogy a p-edik felvételizőhöz a fenti módon rendelt egyéni páros összehasonlítási mátrix éppen megegyezik a *Telcs és szerzőtársai* [2013] által definiált irányított gráf szomszédsági mátrixával. Most térjünk vissza az általuk választott megoldásra. A nem megjelölt objektumok egyenrangúként besorolása vitatható megoldás, sokkal helyesebb lenne azt mondani, hogy ekkor a két objektum összehasonlításának eredménye ismeretlen, azaz $m_{ij}^{(p)} = m_{ji}^{(p)} = 0$, nem pedig $m_{ij}^{(p)} = m_{ji}^{(p)} = 0{,}5$.

[2] Egyelőre tekintsünk el attól, hogy az objektumok most intézményeket, karokat, esetleg szakokat jelentenek. A továbbiakban a *szakot* úgy tekintem, mint a felvételiző által választható, jól definiált alapobjektumot, azaz a szigorú értelemben vett szak nevén kívül különbséget téve a képzés formája (alap, mester vagy osztatlan), módja (nappali vagy levelező) és finanszírozása (állami vagy költségtérítéses) szerint is.

[3] Ezt az angol terminológia a *paired comparison matrix* névvel illeti, ami nem ugyanaz, mint a később szintén előforduló *pairwise comparison matrix*. Én a páros összehasonlítási mátrix kifejezést fogom használni.

Ettől függetlenül utóbbi is alkalmazható, csak világosan meg kell indokolni, hogy miért. Vagy még jobb lenne mindkét megoldás mellett elvégezni a számításokat, majd összehasonlítani azok végeredményét.

Hasonló megállapítás érvényes az 5. pont, a megjelölt objektumok összes kihagyotthoz képesti előnyben részesítése tekintetében. Több dolog is visszatarthat egy felvételizőt attól, hogy az általa legjobbnak gondolt szakot megjelölje. Egyrészt – bár a tényleges szabályok évről évre változtak –, egy további szakra történő jelentkezés gyakran pénzbeli (és pénzben nehezen kifejezhető adminisztrációs) költséggel járt, ami egy szerényebb képességű hallgatót visszatarthat attól, hogy beírja az álmaiban szereplő, de számára a magas ponthatárok miatt lényegében elérhetetlen szakot. Extrém esetben elképzelhető, hogy egy intézmény a jelentkezők sikeres szelekciója révén az összes magas pontszámmal rendelkező felvételizőt megszerzi, mivel azonban ez a többi diák előtt is ismert, ők meg sem jelölik a kérdéses elitszakot. Ekkor a *Telcs és szerzőtársai* [2013] által javasolt módszer az adott szakot indító „második legjobb" intézménynek kedvez, hiszen a felvételizők jelentős része ezt fogja első helyre rangsorolni, miután előre tudják, hogy a „legjobbra" nem tudnak bekerülni. Másrészt, ha például egy vidéki felvételiző nem képes vállalni a budapesti élet jelentette anyagi és az időigényes utazással járó egyéb terheket, annak ellenére sem jelöl meg egy intézményt, hogy az adott szak tekintetében azt tartja a legjobbnak. Emiatt azon intézmények szakjai fognak sokszor az első helyen szerepelni, amelyeknek „vonzáskörzetük" nagy, az ország jelentős nagyságú területeiről vonzzák a jelentkezőket. Harmadrészt, egyáltalán nem tartom elképzelhetetlennek, hogy néhány felvételizőben él az az – egyébként teljességgel alaptalan – tévhit, hogy a felsőoktatási intézmények „látják" a jelentkezési lapjukat, és később hátrányt szenvedhetnek, ha olyan helyre kerülnek, amit nem az elsők között jelöltek meg. Végül, az utóbbi években elterjedt az a gyakorlat, hogy korlátozzák a jelentkezési lapon maximálisan szereplő objektumok számát; idén például legfeljebb öt helyet lehetett beírni, bár a finanszírozási forma megkülönböztetése nélkül.[4] Könnyen belátható, hogy ha egy szakot ennél több intézmény indít, akkor a gyengébb diákoknak nem célszerű valós sorrendjüket megadni, hiszen ezzel bekerülési esélyeik csökkennek.

Természetesen a preferenciák aggregálása az ilyen furcsa esetek kiszűrésére is alkalmas lehet, de mindenképpen hasznos megvizsgálni, hogy mi történik, ha nem élünk ezzel a feltételezéssel sem, a megjelölt és kihagyott szakok közötti összehasonlítást ugyancsak ismeretlennek tekintjük.[5] Tehát *Telcs és szerzőtársai* [2013] javaslata (a továbbiakban a T felső indexű megoldás) mellett három másik alternatív megközelítéssel élhetünk:

– a nem megjelölt objektumok összehasonlításainak eredménye hiányzik, de rosszabbak minden megjelöltnél (H);

– a nem megjelölt objektumok egymás közötti összehasonlításainak eredménye döntetlen, a beírtakkal való kapcsolatuk viszont ismeretlen (D);

[4] Lásd a Felsőoktatási felvételi tájékoztató jelentkezési sorrendre vonatkozó részét: http://www.felvi.hu/felveteli/jelentkezes/felveteli_tajekoztato/FFT_2014K/2_menetrend_es_szabalyok/22_jelentkezes_modja/225_jelentkezesi_sorrend.

[5] A két feltevés közül nyilván a 4., a nem megjelölt objektumok egyenrangúsága inkább vitatható, mint a megjelölt objektumok összes kihagyotthoz képesti preferáltsága.

– a kihagyott szakok egymás közötti, illetve a megjelöltekkel szembeni relációi is ismeretlenek (*HD*).

A rangsorolási módszerek többsége nem az egyéni, hanem az aggregált $\mathbf{M} = \sum_{p=1}^{s} \mathbf{M}^{(p)}$ páros összehasonlítási mátrixon alapul, ahol $s \in \mathbb{N}$ a döntéshozók (itt felvételizők) jól definiált, véges száma; ennek megfelelően itt csak ezeket tárgyalom.[6] Ezzel alapvetően nem térünk el *Telcs és szerzőtársai* [2013] eljárásától, hiszen ők is az aggregált szomszédsági mátrixot használják.

Rangsorolási módszerek

Vagyis – immár feltételezve, hogy sikerült megnyugtató megoldást találni a felvételiző által kihagyott szakok egymás közötti, illetve a megjelöltekhez való viszonyára – eljutottunk egy aggregált szomszédsági vagy páros összehasonlítási mátrixhoz. A kérdés már „csak" az, hogy milyen módszerek révén lehet ebből a szakok egy, valamilyen szempontból optimális rangsorát megkapni.

A páros összehasonlításokon alapuló rangsorolás meglehetősen komoly múltra tekint vissza, miután számos, egymástól független területen bukkannak fel lényegében hasonló problémák: a két legrégebbi talán a pszichológia (*Thurstone* [1927]) és a sport (*Zermelo* [1929]), de nem feledkezhetünk meg a statisztikáról sem (*Éltető–Köves* [1964], *Szulc* [1964]). A szakirodalomban két eltérő rangsorolási megközelítéssel találkozhatunk. Az egyik egy adott választási eljárás sorozatos alkalmazásán alapul (minden lépésben megkapjuk a legjobb alternatívák halmazát, majd azokat elhagyva, a redukált problémára megismételjük a kiválasztást), míg a másik közvetlenül az objektumok egy sorrendjét adja (*Bouyssou* [2004]). *Bouyssou* [2004] megmutatja, hogy az első módszer, a sorozatos kiválasztás nagyon gyakran monotonitási problémákhoz vezet, ezért ajánlott a második megközelítés alkalmazása.

Erre szintén két lehetőség áll rendelkezésre: az alternatívák rangsorának közvetlen felállítása, például az egyéni irányított gráfok *lineáris rendezéssel* közelítése révén (*Kemeny* [1959], *Slater* [1961]), illetve az aggregált páros összehasonlítási mátrix alapján minden objektum egy számszerű értékelésének meghatározása, majd az ezek alapján történő rangsorolás. Legyen $\mathcal{M} \in \mathbb{N}^{n \times n}$ a fent definiált aggregált páros összehasonlítási mátrixok tere. Ekkor az utóbbi egy $f: \mathcal{M} \rightarrow \mathbb{R}^n$ függvény megadását jelenti, amit a következőkben *pontozási eljárásnak* fogok nevezni.

Miután a *Telcs és szerzőtársai* [2013] cikkben mindkét megközelítéssel találkozhatunk, nem haszontalan végigvenni *Bouyssou* [2004] érveit a lineáris rendezés ellenében a pontozási eljárások alkalmazása mellett:

– a feladat gyakran nehéz kombinatorikus problémákhoz vezet, amelyekre csak lassú megoldó algoritmusok léteznek (*Hudry* [2009]);

– egyszerre több optimális rangsort adhat, az ezek közötti választás nem egyértelmű;

– a normatív tulajdonságok vizsgálata meglehetősen bonyolult.

[6] *Chebotarev–Shamis* [1999] meggyőzően érvel amellett, miért nem célszerű az *egyéni* páros összehasonlítási mátrixokból kiindulni.

Ezen érvek alapján *Bouyssou* [2004] felveti az igényt a pontozási eljárások alapos elemzése iránt. Jóllehet egy ilyen átfogó vizsgálat a mai napig várat magára (ami a kiterjedt irodalom miatt nem meglepő), legígéretesebbnek egyfajta axiomatikus megközelítés alkalmazása tűnik: néhány megkövetelt tulajdonság definiálása után jól láthatóvá válnak azok ellentmondásai, kölcsönös kapcsolatai, illetve az egyes módszerek eltérései. Ez nagyon hasonlít a kooperatív játékelmélet megoldási koncepciói esetén követett kutatási irányhoz; például a tavaly közgazdasági Nobel-díjat nyert Lloyd S. Shapley nevéhez fűződő Shapley-értékre több, egymástól részben független axiomatizáció született, nem is említve a különböző játékosztályokon megfigyelhető eltéréseket (*van den Brink–Pintér* [2012]).

Számos tanulmányban találkozunk egy-egy kiválasztott pontozási eljárás karakterizációjával; itt nem szeretném ezeket felsorolni, *Telcs és szerzőtársai* [2013] módszereivel kapcsolatban amúgy is több elő fog kerülni. Ugyanakkor jóval kevesebb az egyszerre több eljárást axiomatikus megközelítéssel vizsgáló munka. A rangsorolási eljárások karakterizációjáról kiváló áttekintést nyújt *Chebotarev–Shamis* [1998] (összesen 40-nél több módszert vizsgál!), amely máig minden páros összehasonlításon alapuló rangsorral foglalkozó kutatás nélkülözhetetlen kiindulópontja.

Chebotarev–Shamis [1999] egy, a cikkben bevezetett tulajdonság, az önkonzisztens monotonitás szempontjából vizsgálja a pontozási eljárások két osztályát, megmutatva, hogy minden győzelem–vereség kombinálásán alapuló eljárás megsérti azt.[7]

Végül *González-Díaz és szerzőtársai* [2013] megjelenés alatt álló tanulmánya néhány elterjedt pontozási eljárás által meghatározott rangsorolási módszer összehasonlítására alkalmazza az axiomatikus megközelítést. E kutatási irány haszna elsősorban az lehet, hogy megbízható alapot szolgáltat a pontozási eljárások közötti választáshoz, így elkerülhetővé válik a *Telcs és szerzőtársai* [2013] tanulmányra is érvényes probléma – az alkalmazandó módszerekről történő *ad hoc* döntés – felmerülése.

Szemléltető példák

A páros összehasonlítási mátrix definíciójának és hozzákapcsolódóan a rangsorolási módszer kiválasztásának jelentőségét érdemes egyszerű példákon keresztül szemléltetni. Legyen tehát két felvételizőnk: 1. és 2., valamint négy szak: *a, b, c, d,* továbbá az 1. felvételiző jelentkezési lapja $\mathbf{t}^1 = [a, b]$, a 2. felvételizőé pedig $\mathbf{t}^2 = [b, a]$. Ekkor a négyféle módszerrel kapott aggregált páros összehasonlítási mátrixok:

$$\mathbf{M}^T = \begin{bmatrix} 0 & 1 & 2 & 2 \\ 1 & 0 & 2 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \mathbf{M}^H = \begin{bmatrix} 0 & 1 & 2 & 2 \\ 1 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \mathbf{M}^D = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \mathbf{M}^{HD} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

[7] *Telcs és szerzőtársai* [2013] hivatkozik erre a cikkre, ezért számomra érthetetlen, hogy a szintén a győzelem–vereség kombinálásán alapuló PageRank-módszer esetén miért nem tér ki *Chebotarev–Shamis* [1999] legfontosabb eredményére, fő üzenetére.

A négy megközelítés jelentős mértékben eltérő következtetésekre vezethet, ami egy újabb egyszerű példával illusztrálható. Legyen négy felvételiző, illetve három szak (*a*, *b* és *c*) az alábbi jelentkezési lapokkal: $\mathbf{t}^1 = [a, b, c]$, $\mathbf{t}^2 = [b]$, $\mathbf{t}^3 = \mathbf{t}^4 = [c]$:

$$\mathbf{M}^T = \begin{bmatrix} 0 & 2 & 1{,}5 \\ 2 & 0 & 2 \\ 2{,}5 & 2 & 0 \end{bmatrix}, \mathbf{M}^H = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix}, \mathbf{M}^D = \begin{bmatrix} 0 & 2 & 1{,}5 \\ 1 & 0 & 1 \\ 0{,}5 & 0 & 0 \end{bmatrix}, \mathbf{M}^{HD} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

A rangsor meghatározására használjuk a klasszikus legkisebb négyzetek módszerét (*Horst* [1932], *Mosteller* [1951], *Morrissey* [1955], *Gulliksen* [1956], *Kaiser–Serlin* [1978]).[8] Miután *Telcs és szerzőtársai* [2013] az M aggregált páros összehasonlítási mátrixban (az itteni jelölésekkel $\mathbf{M}^T$-ben) nem enged meg hiányzó elemeket, esetükben körmérkőzés-problémáról *(round robin)* beszélhetünk, azaz $\mathbf{M}_{ij} + \mathbf{M}_{ji} = s$ minden i, j pár esetén. *González-Díaz és szerzőtársai* [2013] megmutatták, hogy ekkor a legkisebb négyzetek eljárással kapott rangsor azonos a (későbbiekben még előkerülő) sorösszegmódszer által szolgáltatott megoldással.

Az így kapott értékelővektorok (a súlyok összegét 0-ra normálva):

$$\mathbf{x}^T = [-0{,}0208; 0; 0{,}0208];$$

$$\mathbf{x}^H = [-0{,}0427; 0{,}0085; 0{,}0342];$$

$$\mathbf{x}^D = [0{,}1465; 0{,}1919; -0{,}3384];$$

$$\mathbf{x}^{HD} = [0{,}6667; 0; -0{,}6667],$$

az ezeknek megfelelő sorrendek pedig $\mathbf{r}^T = \mathbf{r}^H = [c, b, a]$, $\mathbf{r}^D = [b, a, c]$ és $\mathbf{r}^{HD} = [a, b, c]$. Tehát a legkevesebbszer megjelölt *a* intézmény egyre jobban jár, ahogy eltekintünk a nem megjelölt objektumok egymás közötti, illetve a beírtakkal szembeni összehasonlításaitól. A négy megoldás három különböző sorrendet eredményez, ezek közül kettő teljesen ellentétes egymással. Ennek analógiájára számos más olyan eset található, amikor ugyanazon bemenő adatok alapján különböző felsőoktatási rangsorokat kapunk.[9] Az ezek közötti választás kérdésére itt nem szeretnék kitérni; úgy vélem, ez nem kizárólag a módszertani hátteret biztosító kutatók feladata.

[8] Más kontextusban, de ugyanez a módszer szerepel *Éltető–Köves* [1964]-ben, illetve *Bozóki és szerzőtársai* [2010]-ben. A konkrét számítás során *Chebotarev–Shamis* [1999] megoldását, a mátrix ferdén szimmetrikussá alakítását alkalmaztam..

[9] Olyan példát sem nehéz találni, ahol az $\mathbf{r}^T$ és $\mathbf{r}^H$ sorrendek különböznek. Ha a beadott jelentkezési lapok

$\mathbf{t}^1 = \mathbf{t}^2 = [a, b, c]$, $\mathbf{t}^3 = [b, c]$ és $\mathbf{t}^4 = [c]$,

akkor

$\mathbf{x}^T = [0{,}0208; 0{,}0208; -0{,}0417]$ és $\mathbf{x}^H = [0{,}0292; 0{,}0125; -0{,}0417]$,

vagyis az első esetben az *a* és *b* intézmény viszonya döntetlen, míg a másodikban az előbbi jobb.

A cikkben használt módszerek kritikája

A következőkben a *Telcs és szerzőtársai* [2013] által tárgyalt módszerek szakirodalmi hátterét igyekszem alaposabban bemutatni.[10] A tanulmány nem foglalkozik a rangsoroló eljárások alkalmazásának három kulcskérdésével: létezés, egyértelműség és számítási igény.

Hibafüggvény, inhomogenitási index

A szerzők által bevezetett hibafüggvény elég természetes módon adódik, ennek minimalizálásáról a későbbiekben még lesz szó. Mindenesetre alkalmazását nem ártott volna néhány hivatkozással alátámasztani, indokolni. Az inhomogenitási indexet talán célszerűbb lenne *heterogenitási* mutatónak hívni. Számomra az is furcsa, miért $s[m(m-1)/2]$ szerepel ennek maximumaként. Ugyanis, amennyiben a hibafüggvény értéke egy adott rangsorra ennek felét meghaladja, egyszerűen kapható olyan sorrend, amely alacsonyabb hibafüggvényértékkel jellemezhető. Bár ez csak egy konstanssal való transzformációt jelent, mégis szerencsésebb lenne $s[m(m-1)/4]$-gyel osztani.

> Ha egy rangsorra a hibafüggvény értéke az elméleti maximum felét meghaladja, a rangsort eredményező módszer helyett ennek „inverzét" lenne célszerű használni. Emiatt, ha az egyik eljárással kapott hibafüggvény értéke például 0,9, egyszerűen meg kell fordítani a sorrendet, és máris a sokkal kedvezőbb 0,1-es értéket kapjuk. A probléma némileg hasonló a tőzsdei árfolyamok változásának előjelét (növekedését vagy csökkenését) előrejelző modellekéhez: ezek is csak akkor használhatatlanok, ha a későbbi valós adatokon éppen 50 százalék körül találati aránnyal dolgoznak. Ha ugyanis ennél lényegesen alacsonyabb érteket adnak, a javaslat nem elvetendő, egyszerűen az általa jósolt iránnyal ellentétesen kell cselekedni. Ezért lehet megtévesztő az elméleti maximum használata *Telcs és szerzőtársai* [2013]-nál. Az inhomogenitási indexnek a 305. oldal 4. táblázatában kapott 47,08 százalékos értéke első ránézésre ugyan nem tűnik rossznak, valójában viszont azt jelenti, hogy a végső rangsor lényegében semmilyen információt sem ad az egyéniekről. Ez két okból fordulhat elő: vagy az egyéniek egymással sem konzisztensek (a témában laikusként ez tűnik a kevésbé valószínű esetnek, azt gondolnám, a középiskolás diákok nagyjából egyetértenek abban, hogy a Szegedi Tudományegyetem vagy a Budapesti Corvinus Egyetem gazdasági informatika képzése színvonalasabb, mint a Zsigmond Király Főiskoláé vagy a Nyugat-magyarországi Egyetemé), ami más eszközökkel (akár a javasolt stresszteszttel?) is kimutatható lenne, vagy a cikkben használt módszerek, például a hibás hipotézisek miatt, alkalmatlanok a konszenzusos vélemény visszaadására. E tekintetben különösen fájóak a 305. oldalon lévő 6. táblázatban látható 50 százaléknál nagyobb értékek, miszerint a rangösszegmódszer fordítottja jobb eredményt adna, vagyis a helyes végső sorrendhez azt kellene díjazni, ha egy diák hátul szerepeltette az intézményt.

[10] *Telcs és szerzőtársai* [2013] említést tesz arról, hogy tanulmányuk ezen fejezete részben azonos a *Telcs és szerzőtársai* [2012] cikk módszertani fejezetével. Utóbbi interneten hozzáférhető változata sem foglalkozik az általam tárgyalt kérdések többségével (http://www.cs.bme.hu/~telcs/PUBS/uur.pdf).

A inhomogenitási index (heterogenitási mutató) értékének javasolt maximum azonban csak akkor érhető el, ha minden felvételiző teljes preferencia-sorrendet adott meg (ami gyakorlatilag kizárt), vagy ha – *Telcs és szerzőtársai* [2013]-hoz hasonlóan – a hiányzó összehasonlításokat döntetlenekkel (a nem megjelölt objektumok esetén), illetve szigorú preferenciákkal (a megjelölt és a kihagyott objektumok vonatkozásában) helyettesítjük. Utóbbi esetben viszont a minimum sem nulla, hiszen könnyen belátható, hogy két, a jelentkezési lapon nem szereplő szak közül az egyik biztosan előrébb kerül az aggregált preferencia-sorrendben, ezért a hibafüggvény értéke ezen páros összehasonlítás következtében garantáltan 0,5-tel növekszik.

Ennek szemléltetésére vegyük az egyik korábbi példát, azaz legyenek a beadott jelentkezési lapok $\mathbf{t}^1 = [a, b, c]$, $\mathbf{t}^2 = [b]$, $\mathbf{t}^3 = \mathbf{t}^4 = [c]$. Az inhomogenitási index maximuma $s[m(m-1)/2)] = 4 \times 3 = 12$, értékei a hat lehetséges sorrendben:

a, b, c	*a, c, b*	*b, a, c*	*b, c, a*	*c, a, b*	*c, b, a*
6,5	6,5	6,5	5,5	5,5	5,5

Az elvi maximum ugyan valóban 6,5, de egy ekkora értéket eredményező rangsor (mondjuk az $[a, b, c]$) teljességgel elfogadhatatlan lenne, hiszen a sorrend megfordításával az inhomogenitási index értéke 5,5-re csökkenne. Ezért tartanám indokoltnak a maximumot $s[m(m-1)/4)] = 6$-nak, a *Telcs és szerzőtársai* [2013] által adott érték felének választani. Hasonlóan, a 0 csupán egy elméleti alsó korlát, hiszen egyszerűen belátható, ha egy felvételiző az összesen három intézmény közül csupán egyet jelölt meg, az inhomogenitási index lehetséges legkisebb értéke is 0,5. Ez rögtön látszik az előző példa módosításával: $\mathbf{t}^1 = \mathbf{t}^2 = [c, b, a]$, $\mathbf{t}^3 = [c, b]$ és $\mathbf{t}^4 = [c]$ esetén a $[c, b, a]$ sorrend hiába írja le tökéletesen a felvételizők preferenciáit, az index értéke nem nulla, hanem 0,5.

Emiatt a mutatószám nem informatív, hiszen alsó korlátja nehezen meghatározható, lényegében ismeretlen. Ez is amellett szól, hogy a nem megjelölt objektumok egyenrangúsága igencsak problémás feltevés, célszerűbb lenne ezen összehasonlítások kimenetelét hiányzónak tekinteni, amikor a hibafüggvény legkisebb lehetséges értéke nulla, ekkor azonban a maximum kiszámítása válik problémássá – de ez kevesebb nehézséggel jár, mint a minimum nem meghatározott volta. Az viszont kétségtelen, hogy két, ugyanabból az aggregált páros összehasonlítási mátrixból adódó rangsorhoz tartozó hibafüggvényértékek használhatók a közülük történő választásra, de inkább csak sorrendi, ordinális információt hordoznak.

Az oszlopösszeg-, a sorösszeg- és a rangszámösszeg-módszer gráfreprezentációja

A sorösszegmódszer annyira nyilvánvaló megoldása a problémának, hogy eredete szinte felderíthetetlen, az irodalomban legtöbbször talán *Copeland* [1951]-re hivatkoznak. A *Telcs és szerzőtársai* [2013] által definiált „körmérkőzéses” *(round robin)* esetben azonban ez az eljárás egyáltalán nem nevezhető „heurisztikusnak”, legalább három független axiomatizáció létezik rá: *Young* [1974], az általa adott bizonyítást egyszerűsítő *Hansson–Sahlquist* [1976], valamint *Nitzan–Rubinstein* [1981] és *Bouyssou* [1992] (utóbbi esetén valamivel általánosabb) változatai. Amennyiben, mint a vizsgálatom tárgyát képező tanulmányban, más módszerek is felhasználásra kerülnek, tanácsos érthetően kifejteni, ezek közül melyik tulajdonság elhagyása indokolt, és pontosan miért.

Illusztrációként érdemes kettőt közülük jobban megvizsgálni. Az axiómák (mindegyiket egy $f: \mathcal{M} \to \mathbb{R}^n$ pontozási eljárásra megfogalmazva):

1. SEMLEGESSÉG *(neutrality):* $f(\mathbf{M}) = f(\sigma\mathbf{M})$, ahol $\sigma\mathbf{M}$ az objektumok tetszőleges permutációjával kapott aggregált páros összehasonlítási mátrix;

2. KONZISZTENCIA *(consistency):* legyen i és j két különböző objektum, valamint $M_1, M_2 \in \mathcal{M}$. Ha $f_i(\mathbf{M}_1) \geq f_j(\mathbf{M}_1)$ és $f_i(\mathbf{M}_2) \geq f_j(\mathbf{M}_2)$ tetszőleges $j = 1, 2, ..., s$ mellett, akkor $f_i(\mathbf{M}_1 + \mathbf{M}_2) \geq f_j(\mathbf{M}_1 + \mathbf{M}_2)$. Ha a két egyenlőtlenség közül legalább az egyik szigorú, akkor $f_i(\mathbf{M}_1 + \mathbf{M}_2) > f_j(\mathbf{M}_1 + \mathbf{M}_2)$ is fennáll;

3. HIHETŐSÉG *(faithfulness):* ha egyetlen felvételiző van, akkor a jelentkezési lapjának első helyén szereplő objektum lesz a legjobb;

4. TÖRLÉS *(cancellation):* ha $\mathbf{M}_{ij} = \mathbf{M}_{ji}$ minden $j = 1, 2, ..., s$ mellett, akkor $f_i(\mathbf{M}) = f_j(\mathbf{M})$;

5. MONOTONITÁS *(monotonicity):* legyen i és j két különböző objektum, valamint $M, M' \in \mathcal{M}$. Ha $f_i(\mathbf{M}) \geq f_j(\mathbf{M})$, és az $\mathbf{M}'$ aggregált páros összehasonlítási mátrix azonos $\mathbf{M}$-mel, eltekintve attól, hogy létezik egy olyan k objektum, amire $M'_{ik} > M_{ik}$, akkor $f_i(\mathbf{M}) > f_j(\mathbf{M})$.

A semlegesség azt jelenti, hogy a rangsor független a felvételizők „nevétől", míg a konzisztencia szerint, ha kettébontjuk a felvételizők halmazát, és mindkét csoportban ugyanaz az intézmény lesz a legjobb, akkor ez a teljes halmazra is igaz. A hihetőség talán nem igényel bővebb magyarázatot, míg a törlés egy természetes indifferenciakritériumnak felel meg. Egy konzisztens és hihető pontozási eljárás Pareto-tulajdonságú, ha minden jelentkezési lapon ugyanaz az objektum szerepel első helyen, akkor a végső rangsorban is ez lesz a győztes. A monotonitás azt követeli meg, hogy egy objektum ne kaphasson egy másikkal szemben alacsonyabb relatív értékelést akkor, ha javul egy harmadikhoz viszonyított teljesítménye.

Egy „körmérkőzéses" aggregált páros összehasonlítási mátrixban a sorösszegmódszer az egyetlen rangsorolási eljárás, ami:

– semleges, konzisztens, hihető és teljesíti a törlési tulajdonságot (*Young* [1974]);
– semleges, monoton, konzisztens és teljesíti a törlési tulajdonságot (*Nitzan–Rubinstein* [1981]).

A karakterizációk mindegyike esetén belátható a négy megadott tulajdonság függetlensége. A két axiomatizáció egyike sem könnyen támadható. A semlegesség, a hihetőség és a monotonitás megkövetelése ellen nehéz érveket találni. A törlési tulajdonság előírása szintén egyfajta közömbösséget jelent: ha az egyének preferenciái alapján nem tudunk különbséget tenni az objektumok között, akkor a végső rangsor is tükrözze ezt. Az egyetlen bizonytalan pont a konzisztencia, de emellett is könnyen érvelhetünk azzal, hogy a felvételizők bizonyos ismérvek szerinti csoportokra bontása a vizsgálat szempontjából indokolt lehet (például lakóhely, életkor, érettségi eredmény stb. szerint), és nem szeretnénk, ha az így adódó rangsorok ellentmondanának egymásnak. Ha pedig ezeket elfogadjuk, akkor már nem marad más megoldás, mint az oszlopösszeg-eljárás alkalmazása. Ennek fényében egyenesen megdöbbentő, hogy *Telcs és szerzőtársai* [2013] tanulmánya a többi javasolt módszer esetén nem tér ki arra a kérdésre, hogy a fentiek közül melyik tulajdonság megsértése fogadható el.

Az eljárásnak minden aggregált páros összehasonlítási mátrix esetén létezik egyértelmű megoldása, ráadásul számítási igénye minimális. *Telcs és szerzőtársai* [2013] azt is megemlíti, hogy hiányos preferencia-sorrendek esetén az ezzel kapott sorrend nem feltétlenül eredményez minimális hibafüggvényt. Nem triviális azonban, hogy ellenkező esetben miért adna azt (ha tényleg igaz, ezt érdemes lenne egy állításban, bizonyítással együtt megfogalmazni, esetleg megadni a pontos hivatkozást), illetve nem lett volna haszontalan egy ellenpélda bemutatása sem.

A *Telcs és szerzőtársai* [2013] cikkben tárgyalt rangszámösszeg-módszerrel, mivel az oszlopösszeg-eljárással azonos eredményre vezet, nem foglalkozom.

A páros összehasonlítás módszere hiányos rangsorokra

Az ezt bemutató alfejezetben hivatkozott *Fedrizzi–Giove* [2007], illetve *Bozóki és szerzőtársai* [2010] azzal foglalkozik, hogyan számítható ki a súlyvektor egy *hiányos* páros összehasonlítási mátrixból, ezzel szemben a tárgyalt tanulmányban, ahogy az az (5) képletből (298. o.) látható, erről szó sincs. Érzésem szerint elég megtévesztő a szóhasználat, hiszen a felvételizők preferenciái esetén ugyan valóban megjelennek a hiányzó összehasonlítások, ezt azonban *Telcs és szerzőtársai* [2013] egy megfelelően definiált páros összehasonlítási mátrixszal lényegében eltüntetik. Ezek alapján *Fedrizzi–Giove* [2007] tanulmánya inkább csak ötletet adhatott a bemutatott módszerhez.

Utóbbi mindenképp egy új, innovatív megoldásnak is tekinthető, ekkor azonban bővebb tárgyalást igényelne. A páros összehasonlítási mátrix elemeinek kiszámítására alkalmazott hányadosképző transzformáció komoly torzítást tartalmazhat, hiszen nem tesz különbséget például az $m_{ij} = 2$, $m_{ji} = 1$, illetve az $m_{ij} = 200$, $m_{ji} = 100$ kimenetelek között, holott 300 felvételiző értékelése nyilván sokkal megbízhatóbbnak tekinthető, mint három véleménye. Itt sem kapunk információt a megoldás létezésének, egyértelműségének feltételeiről, annak számításigényéről. Ha a szerzők ténylegesen alkalmazni szeretnék a javasolt eljárást, érdemes lenne részletesebben foglalkozni vele, megvizsgálni legfontosabb tulajdonságait és összehasonlítani más, jól ismert rangsorolási eljárásokkal.

PageRank-módszer

A PageRank-eljárás egyre népszerűbb a páros összehasonlításon alapuló rangsorolás irodalmában, éppen ezért kicsit szegényesnek érzem az erre való hivatkozások számát és minőségét; a szerzők által megadott *Page–Brin*-féle leírást (*Teles* *és szerzőtársai* [2013] 300. o. és 317. o.) nem tudtam elérni. Ennél komolyabb probléma, hogy a képletek, a matematikai háttér bemutatása hiányában az sem derül ki, pontosan miként történik az értékelővektor kiszámítása.

A módszer első megjelenéseként a *Brin–Page* [1998] és a *Page és szerzőtársai* [1999] tanulmányokat szokás említeni.[11] Az általam javasolt axiomatikus tárgyalás nem hiányzik a PageRank esetén sem (*Altman–Tennenholtz* [2008]), sőt bizonyos karakterizációs eredmények is ismertek (*Altman–Tennenholtz* [2005], *Slutzki–Volij* [2006]).[12] Vagyis megint tanácsos lenne kitérni ezek relevanciájára, esetleges megkérdőjelezhetőségére az egyetemi jelentkezések körében.

Az eljárás az aggregált páros összehasonlítási mátrix egy transzformáltjának sajátvektorán alapul. Legyen $\mathbf{C} = \mathrm{diag}(\mathbf{e}^T\mathbf{M})$, ahol $\mathbf{e} \in \mathbb{R}^n$ az n-dimenziós csupa 1-esből álló vektor, tehát a $\mathbf{C}$ mátrix diagonálisában az egyes objektumok – az irányított gráfon vett – elődeinek (*predecessors;* azok az objektumok, amelyek hozzá képest preferáltak) száma szerepel. Ekkor a PageRank-rangsor alapjául szolgáló értékelővektor az $\mathbf{r} = (1 - \mathbf{d})\mathbf{M}\mathbf{C}^{-1}\mathbf{r} + d\mathbf{e}$ módon számítható, a $d \in [0, 1]$ paraméter inverz módon az elődök szerepének fontosságát méri. A feladatnak $d > 0$ esetén mindig létezik egyértelmű megoldása, $d = 0$ mellett csak akkor, ha az aggregált preferenciákat leíró irányított gráf erősen összefüggő, azaz bármely csúcsból minden másikba vezet irányított út (*Altman–Tennenholtz* [2008]). Itt nem kívánunk kitérni a módszer pontos magyarázatára, ez számos alkalmazással foglalkozó cikkekben – például *Radicchi* [2011] vagy *Dingle és szerzőtársai* [2013] – megtalálható. Ezek több hasznos tanáccsal is szolgálnak a *Telcs és szerzőtársai* [2013] által vizsgált egyetemi jelentkezésekkel analóg problémák vonatkozásában.

Összefoglalva, a PageRank-módszer ugyan valóban jól alkalmazható az egyetemi szakok rangsorolására, a vizsgált cikkből azonban nem kapunk információt a számítás részleteiről és a d paraméter értékéről sem. Ha a szerzők a $d = 0$ megoldással éltek, akkor pedig ki kellett volna térni az erős összefüggőség kérdésére, bár a jelentkezések nagy számának köszönhetően ez nagy valószínűséggel teljesül.

Genetikus algoritmusok alkalmazása

A hibafüggvényt minimalizáló rangsor *(minimum violations ranking)* kérdéskörét a döntetleneket kizáró esetben *Ali és szerzőtársai* [1986] tárgyalja, míg *Coleman* [2005] vegyes bináris egészértékű programozási feladata a döntetleneket és a hiányzó összehasonlításokat is megengedi. Az egyetemi jelentkezések nagyméretű problémájában azonban valószínűleg elengedhetetlen a genetikus algoritmus használata. *Telcs és szerzőtársai* [2013] legkisebb hibaértéket generáló leírása ugyanakkor erősen vázlatos, nem derül ki belőle a leállási kritérium, és az algoritmus (várható) pontosságáról sem kapunk információt, akár néhány

[11] Előbbire a GoogleScholar kereső szerint több mint 10 ezer, utóbbira pedig több mint 5 ezer hivatkozás található.

[12] Az axiomatizációk a $d = 0$ esetre vonatkoznak, lásd a következő bekezdést.

kisméretű szimuláción keresztül. A módszert kevésbé ismerő hazai kutatók számára ajánlható *Benedek* [2003] disszertációja, amely széles körű áttekintést nyújt a genetikus algoritmusról.

A feladatnak a probléma természetéből adódóan biztosan létezik megoldása, de a számítások meglehetősen erőforrás-igényesek. Ettől eltekintve sem mellőzhető *Bouyssou* [2004] figyelmeztetése a többszörös optimum lehetőségéről; egy ekkora méretű probléma esetén akár százas, ezres nagyságrendű optimumok is elképzelhetők, és *Pasteur* [2010] példája arra utal, hogy gyakran még a legjobb alternatívák halmaza sem egyértelmű. Ekkor az sem világos, hogy a genetikus algoritmus éppen melyik megoldást próbálja megtalálni, így az eredmény erősen függhet a kezdeti feltételezésektől.

A hibafüggvény minimalizálásának egyik legnagyobb problémája tehát a *Telcs és szerzőtársai* [2013] által a sorösszegmódszernél említett jelenséghez kapcsolódik, miszerint a kapott sorrend nem feltétlenül egyértelmű. Kétségtelen, hogy az összeadás miatt az eljárás sok olyan objektumot is ekvivalensnek tekinthet, amelyek a felvételizők preferenciái alapján jól megkülönböztethetők. Ugyanakkor a másik véglet, a hibafüggvény szempontjából optimális egyértelmű rangsor is nehezen elfogadható eredményekre vezet, hiszen két, az irányított gráfban teljesen azonos csúcs (például egy adott objektum és ennek „duplikáltja") között is sorrendet állít fel, ami biztosan nemkívánatos. E tekintetben, intuitív alapon, a PageRank-módszer kedvezőbbnek tekinthető, de hasonlóképpen javasolható a *González-Díaz és szerzőtársai* [2013] cikkben szereplő módszerek többsége, melyek – a *Telcs és szerzőtársai* [2013] által nem használt hiányzó összehasonlítások megengedésével – nagy biztonsággal képesek egyértelmű sorrend felállítására.

A fenti okok miatt én mindenképp egy vagy több *pontozási eljárás* alkalmazását javasolnám a felsőoktatási rangsorok felállítására, az inhomogenitási index inkább azok egymással, esetleg a genetikus algoritmussal kapott, optimumhoz közeli megoldással való összevetésére lehet alkalmas.

A csomópontok aggregálása

Ehhez a ponthoz nincs nagyon mit hozzátenni, az objektumok aggregálása az intuíció által javasolt módon működik. Esetleg annyit érdemes megjegyezni, hogy a hurokélek elhagyása nem feltétlenül szükséges: léteznek olyan rangsorolási módszerek, amelyek ezek kezelésére is alkalmasak, például a *Herings és szerzőtársai* [2005] cikkben említett, súlyozott irányított gráfon alapuló modell.

Néhány további megjegyzés

Célom elsősorban a rangsorolás problémájának vizsgálata volt, ezért *Telcs és szerzőtársai* [2013] tanulmányának fennmaradó részével kapcsolatban csak néhány gondolatot osztanék meg az olvasóval.

A szerzők viszonylag hosszan, csaknem egy oldal terjedelemben foglalkoznak azzal a kérdéssel, miért jobb a preferencia-sorrendek alapján történő rangsorképzés, mint a jelentkezők számának figyelembevétele. Kétségtelen, hogy lehetnek olyan nemzetközi felsőoktatási teljesítménymutatók, amelyek ilyen jellemzőket is figyelembe vesznek, de ezekre kár szót vesztegetni. Így az alapos kifejtés célja egyáltalán nem volt világos számomra, egyszerűen érthetetlen, miként merülhet fel bárkiben, hogy a felvételizők puszta tömege megfelelő mutató lenne. Egy közgazdasági hasonlattal élve, ez körülbelül olyan, mintha egy ország gazdasági teljesítőképességét nem a GDP-vel, hanem a lakosság számával szeretnénk mérni – ennek abszurditásáról, azt hiszem, senkit nem kell meggyőzni. Az illusztrációként bemutatott, két hallgatót és három intézményt tartalmazó példa két bekezdésben, a 301. és a 302. oldalon is feltűnik.

A 7. táblázatból nem derül ki, pontosan melyik mesterszak jelentkezésein alapul, míg a 309. oldalon a főszövegben téves a gazdálkodási és menedzsment-alapszak intézményi rangsorát tartalmazó táblázatra való hivatkozás.

Itt olvasható egy érdekesnek tűnő felvetés is: „A páros összehasonlítás eredményeként pedig az intézmények közötti távolságokat is láthatjuk." (309. o.) Egy rangsorolással foglalkozó cikkben nem szerencsés a távolság szó köznyelvi jelentésének használata, hiszen matematikai értelemben távolságfüggvényről csak akkor beszélhetünk, ha teljesül a háromszög-egyenlőtlenség, erre azonban nem találunk indoklást. Célszerűbb lett volna a távolság szót idézőjelbe tenni, ahogy az egy bekezdéssel feljebb látható.

Tanulságos kitérni az inhomogenitási index értékeire is. A fentiek értelmében érvényes alsó korlát nem adható, ugyanakkor a 0,5 biztosan alkalmas felső határ; egyszerűen belátható, ha az I inhomogenitási index értéke ennél nagyobb, akkor a rangsor teljes megfordításával a mutató értéke éppen $1 - I$ lenne, azaz javulna. Ehhez képest megdöbbentően nagynak tűnnek a számított értékek, a 4. táblázatban 47,08 százalék (itt nincs megadva külön-külön az egyes módszerekre, ezért csak feltételezhető, hogy az elvileg optimális rangsort megtaláló genetikus algoritmushoz tartozik), az 5. táblázatban legalább 46,78 százalék, a 6. táblázatban minimum 51,05 százalék (!) (ráadásul nem a genetikus algoritmusé a legkisebb), a 7. táblázatban 42,19 százalék. Ennek két oka lehet: vagy a felvételizők sorrendjei lényegében minimális konzisztenciával rendelkeznek (e tekintetben az sem sokat segítene, ha a 6. táblázatban megfordítanánk a rangsort), ez azonban kevéssé tűnik valószínűnek; vagy a *Telcs és szerzőtársai* [2013] definíciója, a megjelölt szakok összes kihagyottal szembeni preferáltsága és két utóbbi ekvivalenssé nyilvánítása teljességgel alkalmatlan megoldás.

Érdemes ezeket összevetni a robusztusságvizsgálat eredményeivel. Az inhomogenitási index gyakorlatban megfigyelt értékeihez hasonlók a $\lambda = 0{,}01$ csillapítási tényező mellett fordulnak elő, amikor az „elméletitől" gyökeresen eltérő sorrendek is megjelenhetnek. Ennek fényében vitatható az az állítás, miszerint a „stresszteszt alacsony értéke azt mutatja, hogy a páros összehasonlítás eredményeképpen kapott Z értékekből nagy pontossággal visszaállítható az eredeti preferenciamátrix" (305. o.). Bár a kifejtés elégtelensége miatt a stresszteszttel nem foglalkoztam, a fentiek tükrében az aggregált sorrendből valószínűleg nem kaphatók meg az eredeti preferenciák.

Rövid nemzetközi kitekintés

Telcs és szerzőtársai [2013] tanulmányát megelőzően más cikkek ugyancsak javasoltak hasonló, a felvételizők preferenciáin alapuló megközelítést az oktatási intézmények rangsorolására, bár ezt a szerzők nem említik. A legelső ilyen tanulmány valószínűleg *Avery és szerzőtársai* [2004] műhelytanulmánya. Ez azóta egy rangos folyóiratban, a The Quarterly Journal of Economicsben is megjelent (*Avery és szerzőtársai* [2013]). Itt a korábbi változatot szeretném vázlatosan bemutatni. A szerzők 3240 amerikai diák körében végeztek kérdőíves felmérést, így számos egyéb jellemzőjük ismert, azonban az adatbázis *Telcs és szerzőtársai* [2013]-val szemben (részben szándékosan) nem reprezentatív. A megközelítés hasonlít a játékosok egy bajnokság alapján történő rangsorolására: a jelentkezőt elfogadó intézmények közötti döntés egy-egy mérkőzésnek felel meg, melynek győztese a végül kiválasztott egyetem. *Avery és szerzőtársai* [2004] a diákok anyagi körülményeiről, családi hátteréről, lakóhelyéről stb. rendelkezésre álló kiegészítő információk révén erre a bináris eredményváltozóra egy multinomiális logit modellt épít, majd azt bayesi keretbe helyezi, ami *a posterior* eloszlásokon végzett szimulációk révén az intézmények sorrendjének megbízhatóságáról is információval szolgálhat.

A *Telcs és szerzőtársai* [2013] által elemzett országos adatbázis esetén ugyan ez az út nem járható, *Avery és szerzőtársai* [2004] számos felvetése azonban erre a tanulmányra is érvényes.[13] Elsőként, az így kialakított rangsor az intézmények oldaláról rendkívül nehezen manipulálható, míg a korábbi mutatók befolyásolása egyáltalán nem bizonyult lehetetlennek. Ez a fogolydilemmához hasonló rossz egyensúly kialakulásához vezetett, ahol – miután egyénileg minden intézmény érdekében állt a manipuláció, ezért azt gyakorolták is – végül senki sem járt jobban, mint kezdetben, miközben minden érdekelt egyetem erőforrásokat áldozott a csalás érdekében. A manipulálhatóság hiánya vagy legalábbis erősen költséges volta a Felvi-adatbázisra ugyanúgy érvényes: az intézményeknek azt kellene elérniük, hogy az őket hátrébb soroló diákok ezt a tényt titkolják el a felvételi folyamat során, ami azonban nekik nem áll érdekükben.

Avery és szerzőtársai [2004] cikkében ugyancsak előkerül a hiányzó páros összehasonlítások problémája. Ez azonban megfelelő adatsűrűség esetén nem okoz problémát a rangsor felállításában, hiszen, bár az *a* egyetem nem feltétlenül szerepelt együtt *c*-vel egyetlen jelentkezési lapon sem, várhatóan található olyan *b* intézmény, amely már mindkettővel össze lett hasonlítva, így az *a* és *c* kapcsolata közvetett módon megállapíthatóvá válik (az is előfordulhat, hogy az *a* és *c* intézmények összehasonlítása csak egy hosszabb láncolaton keresztül lehetséges). A feladat hasonló a svájci rendszerű sakkversenyekhez, és valóban ismertek ilyen irányú alkalmazási kísérletek (*Csató* [2013]).

[13] Emellett számos, az amerikai példában felmerülő nehézség a hazai körülmények között csak kisebb gondokat okoz. Ezek közül legfontosabb a tandíj hiánya, de a felsőoktatás folyamatos átalakulása következtében valószínűleg az alumniszervezetek befolyása is jóval csekélyebb, tehát a szülők intézményválasztása nem igazán befolyásolja a diákok preferenciáit.

A rangsort kétségtelenül befolyásolhatja az összehasonlítási struktúra, ezért *Avery és szerzőtársai* [2004]-hoz hasonlóan érdemes lenne az adatleírás során megemlíteni az egyes intézménypárok előfordulásának (azaz az $m_{ij} + m_{ji}$ összegek) átlagát és mediánját, illetve a speciális eseteket, szűk keresztmetszeteket. Noha a felvételizőkről rendelkezésre álló adatok ismeretének hiányában a rangsor megbízhatóságának *Avery és szerzőtársai* [2004] nyomán végrehajtott elemzése nem lehetséges, a páros összehasonlítások véletlen változókként kezelésével, majd azokra valamilyen eloszlás illesztésével talán vizsgálható lenne az érzékenység.[14] Könnyen előfordulhat például, hogy az amerikai rangsorral analóg módon az első helyek viszonylag stabilak, és a középmezőny válik kevésbé robusztussá; ezt szintén befolyásolhatja az összehasonlítások struktúrája.

Végezetül érdemes áttekinteni *Avery és szerzőtársai* [2004] meggyőző érveit a felvételizői preferenciákon alapuló intézményi rangsorokról, hiszen első ránézésre valóban nehezen érthető, miért lenne alkalmas ez a módszer az egyetemek értékelésére.

– Minden jelentkező igyekszik alkalmazkodni a többi, hozzá hasonló diák viselkedéséhez. Egyfelől a felvett hallgatók képességei befolyásolják a tanítás minőségét, másrészt társaiktól is tanulhatnak, végül a tanulmányok során kialakított kapcsolati háló értékesebb lehet, ha az ismerősök magasabb közigazgatási és vállalati pozíciókat töltenek be.

– A felvételizők jelentős része nem érzéketlen az intézmények minőségi, versenyképességi mutatói iránt. Információkat gyűjtenek már oda járó ismerőseiktől, szüleiktől, az egyetemek tanáraitól és saját tapasztalataik alapján. Miután a pályaválasztás az életpálya egyik döntő pontja, nyilván komoly erőfeszítést hajlandók tenni az optimális döntés meghozatala érdekében. E tekintetben több párhuzam figyelhető meg más iparágakkal, az éttermek, hotelek, könyvek, utazási célpontok vagy internetes kereskedők fogyasztói értékeléseivel; a választás során itt is fontos szempont lehet a többi ember véleménye.

– Ha a diploma a munkavállaló vállalatok által nem megfigyelhető termelékenységére vonatkozó jelzés (*Spence* [1973]), akkor a várható bér az adott intézménybe korábban járó hallgatók képességeinek függvénye lehet. Emiatt a jól tanuló felvételizőknek érdekük a hasonló eredményekkel rendelkező diákok társaságát keresni.

Természetesen felmerülhet, hogy a középiskolások egy része külföldi felsőoktatási intézményt választ, ezért nem is ad be jelentkezési lapot. Ez azonban alig kezelhető, a vizsgált időszakban valószínűleg nem volt tömeges jelenség, és az sem egyértelmű, hogy befolyásolná a hazai intézmények rangsorát: ha a legjobb középiskolások mennek el az országból, az tekinthető úgy, mintha az itthon maradók mindegyike kedvezőbb helyzetbe kerülne, akik ezt felismerve immár eredetileg számukra elérhetetlen egyetemeket jelölnek meg.

A felvételizők kinyilvánított preferenciáin alapuló felsőoktatási rangsorok lehetősége nem ismeretlen az ezzel a területtel foglalkozó áttekintésekben (*Myers–Robe*

[14] Ilyen elemzésekről sajnos nincs tudomásom.

[2009]), és a megközelítést orvosrezidensek kórházválasztására is alkalmazták (*Machado és szerzőtársai* [2012]). E kutatási irányok mindenképpen említést érdemelnének egy felvételizői preferencia-sorrendet javasló tanulmányban.

A cikk következtetései és hivatkozásai

Bizonyos mértékben kénytelen vagyok vitába szállni *Telcs és szerzőtársai* [2013] összegzésével, az abban kiemelten szereplő megállapításokkal. „Az alkalmazott gráfreprezentáció segítségével információvesztés nélkül tároltuk a hallgatói jelentkezéseket." (314. o.) E mondat kétségtelenül igaz állítás, de a megfeleltetés triviális. A javasolt módszerek robusztussága iránt komoly kétségeket támaszt a hibafüggvény magas értéke. A további kutatási tervekkel, irányokkal viszont teljes mértékben egyetértek. Tehát *Telcs és szerzőtársai* [2013] valóban mutatott *egy* lehetséges eljárást az egyéni preferencia-sorrendek összegzésére, ami azonban több ponton vitatható, ráadásul a bemutatott számítási eredmények sem meggyőzők. Sok igazság van a bekezdésben szereplő állításban: „Tanulmányunkban preferencia-sorrend felállítására *teszünk kísérletet*, ami új irány lehet, szakítva az elődök által kicsiszolt módszertan hagyományaihoz képest." (292. o.; kiemelés tőlem – *Cs. L.*) A cikk ilyen szempontból valóban egy nagyon innovatív, ám módszertanilag támadható kísérlet.

Abban viszont egyértelműen igazuk van a szerzőknek, hogy az alkalmazási lehetőségek sokkal szélesebbek, mint az itt bemutatott felsőoktatási rangsor felállításának problémája. Éppen ezért lett volna célszerű említést tenni a más tudományterületeken – a játékelmélet, a sport, a szavazáselmélet, a pszichológia stb. terén – elért eredményekről is. Ekkor talán könnyebben feltűnt volna a páros összehasonlítás során fellépő döntetlen (ekvivalenciareláció) és hiányzó információ (ismeretlen reláció) közötti döntő különbség, amit a tanulmány lényegében említés nélkül hagy.

Egy cikk értékéről, szakirodalmi beágyazottságáról gyakran már első ránézésre pontos képet ad a hivatkozási lista, ezért befejezésül álljon itt egy rövid számvetés *Telcs és szerzőtársai* [2013] referenciáiról. A tanulmány összesen 30 ilyet sorol fel, ezek közül tíz önhivatkozásnak tekinthető. Három a nem teljesen kitöltött páros összehasonlítások hiányzó elemeinek becslésével foglalkozik, ami a szerzők által használt definíció miatt a tanulmányban nem jelenik meg, két másik viszont relevánsnak tekinthető a páros összehasonlítási mátrixok vonatkozásában. További nyolc hivatkozott mű a hazai felsőoktatási rangsorokkal kapcsolatos, egy szóbeli közlés, míg a kanadai tanulmány egy, a felvételizők jelentkezési döntésére alkalmazott logit modellt mutat be, ami a tárgyalt rangsorolás szempontjából irreleváns (emellett szövegközi utalás sem szerepel rá). A fennmaradó öt cikk közül egy a hibafüggvény minimalizálására használt genetikus algoritmus technikai megvalósítását tárgyalja, egy, a PageRank-módszert tárgyaló hivatkozás pedig számomra elérhetetlennek bizonyult, de amúgy sem a szakirodalomban „szokásos" cikk. Végül, három tanulmány ugyan kötődik a páros összehasonlításokon alapuló rangsorolás problémájához (*Chebotarev–Shamis* [1999], *Kóczy–Nichifor* [2013], *Slikker és szerzőtársai* [2012]), viszont mindegyikben kiemelten szerepel a hiányzó összehasonlítások lehetősége, amire *Telcs és szerzőtársai* [2013] egyáltalán nem tér ki, a nem megjelölt szakok viszonyát eldöntetlennek tekinti, ami közel sem azonos az előbbivel.

A tanulmány továbbfejlesztésére vonatkozó javaslatok

A fent megfogalmazott kritikák alapján úgy tűnhet, a vizsgált tanulmányban alkalmazott megközelítés erősen kifogásolható, ez azonban egyáltalán nem igaz. A bevezetésben említett célokkal csak egyetérteni tudok: a felvételizők preferenciái alapján kapott rangsor sok szempontból alkalmas az önkényesség vádjának elutasítására, hiszen nem merülnek fel a mutatók szubjektív kiválasztásából és súlyozásából vagy az intézmények eltérő méretéből fakadó torzítások; az aggregált páros összehasonlítási mátrix definiálása mindössze kisebb nehézségekkel jár (lásd a Módszertani háttér: rangsorolás páros összehasonlításokkal című fejezetet), míg a pontozási eljárás kiválasztása axiomatikus alapon jól védhetővé tehető. Emiatt az eddig ismertek közül a leginkább objektív mutatónak tűnik, széles körben történő publikálása és felhasználása mindenképpen kívánatos lenne. Vitairatom kizárólag módszertani megközelítésből vizsgálta a tanulmányt, ekkor pedig nem kérdés, hogy mit sikerült megmérnünk a felvételizői preferencia-sorrend felállításával. Ennek eldöntését véleményem szerint célszerű az oktatáskutatókra, illetve nem utolsósorban az érintett szülőkre és diákokra bízni.

Napjainkban élénk vita folyik a magyar felsőoktatás versenyképességéről, az egyes intézmények kutatóegyetemi vagy egyéb szempontból kiemelt besorolásáról, így akár azt is elképzelhetőnek tartom, hogy a politikai döntéshozók – a felülről erőltetett elképzelések helyett – a „szavazók”, vagyis a felvételizők preferenciáit vegyék figyelembe. Nagy valószínűséggel állíthatjuk ugyanis, hogy százezer olyan döntéshozó (ami több év jelentkezéseinek felhasználása esetén a milliós nagyságrendet közelítheti), akinek életpályáját, foglalkoztatási esélyeit nagymértékben befolyásolja a megfelelő felsőoktatási intézmény kiválasztása, az egyetemek, szakok sokkal megbízhatóbb értékelésére képes, mint a témát kutató bármelyik szakmai műhely vagy az államapparátus megannyi hivatalnoka. Ezért, amennyiben nem merülnek fel adatvédelmi aggályok, valamilyen formában tanácsos lenne az egész adatbázis közzététele. A rangsorolás alapjául szolgáló (legáltalánosabb, szakonkénti) aggregált preferenciákat tartalmazó mátrix ugyancsak hasznos kutatási segédanyagként szolgálhat, emellett lehetővé tenné a szerzők számításainak reprodukálását.

Telcs és szerzőtársai [2013] vizsgálata tehát egyedülállónak tekinthető abban, hogy egy széles körű, országos adatbázis felhasználásával készült, a páros összehasonlításokon alapuló rangsorolás alkalmazásának részletei viszont (még) nem minden esetben tisztázottak. Mivel a szerzők megjegyzése szerint ez „egy hosszabbra tervezett cikksorozat első eleme” (292. o.), megfontolásra ajánlanám a kiindulási hipotézisek alapos felülvizsgálatát, kiegészítését, nehogy a későbbi eredmények emiatt megkérdőjelezhetővé váljanak. Az elvégzendő feladat kétségtelenül óriási, hiszen egyszerre követeli meg a matematikai háttér és a felsőoktatási rendszer alapos ismeretét; nem véletlen, hogy a témában született amerikai tanulmány is négy neves szakértő közreműködésével készült (*Avery és szerzőtársai* [2013]).

Hivatkozások

Ali, I.–Cook, W. D.–Kress, M. [1986]: On the minimum violations ranking of a tournament. Management Science, Vol. 32. No. 6. 660–672. o.

Altman, A.–Tennenholtz, M. [2005]: Ranking systems: the PageRank axioms. Megjelent: Proceedings of the 6th ACM conference on Electronic commerce. 1–8. o.

Altman, A.–Tennenholtz, M. [2008]: Axiomatic foundations for ranking systems. Journal of Artificial Intelligence Research, Vol. 31. No. 1. 473–495. o.

Avery, C.–Glickman, M.–Hoxby, C.–Metrick, A. [2004]: A revealed preference ranking of U.S. colleges and universities. National Bureau of Economic Research Working Paper, No. 10803.

Avery, C.–Glickman, M.–Hoxby, C.–Metrick, A. [2013]: A revealed preference ranking of U.S. colleges and universities. The Quarterly Journal of Economics, Vol. 128. No. 1. 425–467. o.

Benedek Gábor [2003]: Evolúciós gazdaságok szimulációja. PhD-értekezés, Budapesti Corvinus Egyetem, Budapest.

Bouyssou, D. [1992]: Ranking methods based on valued preference relations: a characterization of the net flow method. European Journal of Operational Research, Vol. 60. No. 1. 61–67. o.

Bouyssou, D. [2004]: Monotonicity of “ranking by choosing”: a progress report. Social Choice and Welfare, Vol. 23. No. 2. 249–273. o.

Bozóki Sándor–Fülöp János–Rónyai Lajos [2010]. On optimal completion of incomplete pairwise comparison matrices. Mathematical and Computer Modelling, Vol. 52. No. 1–2. 318–333. o.

Brin, S.–Page, L. [1998]: The anatomy of a large-scale hypertextual web search engine. Computer networks and ISDN systems, Vol. 30. No. 1. 107–117. o.

Chebotarev, P. Yu.–Shamis, E. [1998]: Characterizations of scoring methods for preference aggregation. Annals of Operations Research, Vol. 80. 299–332. o.

Chebotarev, P. Yu.–Shamis, E. [1999]: Preference fusion when the number of alternatives exceeds two: indirect scoring procedures. Journal of the Franklin Institute, Vol. 336. No. 2. 205–226. o.

Coleman, B. J. [2005]: Minimizing game score violations in college football rankings. Interfaces, Vol. 35. No. 6. 483–496. o.

Copeland, A. H. [1951]. A reasonable social welfare function. Seminar on applications of mathematics to social sciences. University of Michigan.

Csató László [2013]: Ranking by pairwise comparisons for Swiss-system tournaments. Central European Journal of Operations Research. Vol. 21. No. 4. 783–803. o.

Dingle, N.–Knottenbelt, W.–Spanias, D. [2013]: On the (page) ranking of professional tennis players. Megjelent: *Tribastone, M.–Gilmore, S.* (szerk.): Computer Performance Engineering. Lecture Notes in Computer Science. Springer, Berlin–Heidelberg, 237–247. o.

Éltető Ödön–Köves Pál [1964]: Egy nemzetközi összehasonlításoknál fellépő indexszámítási problémáról. Statisztikai Szemle, 42. évf. 5. sz. 507–518. o.

Fedrizzi, M.–Giove, S. [2007]: Incomplete pairwise comparison and consistency optimization. European Journal of Operational Research, Vol. 183. No. 1. 303–313. o.

González-Díaz, J.–Hendrickx, R.–Lohmann, E. [2013]: Paired comparisons analysis: an axiomatic approach to ranking methods. Social Choice and Welfare. Megjelenés alatt. DOI 10.1007/s00355-013-0726-2.

Gulliksen, H. [1956]: A least squares solution for paired comparisons with incomplete data. Psychometrika, Vol. 21. No. 2. 125-134. o.

Hansson, B.–Sahlquist, H. [1976]: A proof technique for social choice with variable electorate. Journal of Economic Theory, Vol. 13. No. 2. 193–200. o.

Herings, P. J.-J.–van der Laan, G.–Talman, D. [2005]: The positional power of nodes in digraphs. Social Choice and Welfare, Vol. 24. No. 3. 439–454. o.

Horst, P. [1932]: A method for determining the absolute effective value of a series of stimulus situations. Journal of Educational Psychology, Vol. 23. No. 6. 418–440. o.

Hudry, O. [2009]: A survey on the complexity of tournament solutions. Mathematical Social Sciences, Vol. 57. No. 3. 292–303. o.

Kaiser, H. F.–Serlin, R. C. [1978]: Contributions to the method of paired comparisons. Applied Psychological Measurement, Vol. 2. No. 3. 423–432. o.

Kemeny, J. G. [1959]: Mathematics without numbers. Daedalus, Vol. 88. No. 4. 577–591. o.

Kóczy Ádám László–Nichifor, A. [2013]: The intellectual influence of economic journals: quality versus quantity. Economic Theory, Vol. 52. No. 3. 863–884. o.

Machado, M. P.–Mora, R.–Romero-Medina, A. [2012]: Can we infer hospital quality from medical graduates' residency choices? Journal of the European Economic Association, Vol. 10. No. 6. 1400–1424. o.

Morrissey, J. H. [1955]: New method for the assignment of psychometric scale values from incomplete paired comparisons. Journal of the Optical Society of America, Vol. 45. No. 5. 373-378. o.

Mosteller, F. [1951]: Remarks on the method of paired comparisons: I. The least squares solution assuming equal standard deviations and equal correlations. Psychometrika, Vol. 16. No. 1. 3–9. o.

Myers, L.–Robe, J. [2009]: College rankings: History, criticism and reform. Center for College Affordability and Productivity, Washington.

Nitzan, S.–Rubinstein, A. [1981]: A further characterization of Borda ranking method. Public Choice, Vol. 36. No. 1. 153–158. o.

Page, L.–Brin, S.–Motwani, R.–Winograd, T. [1999]: The PageRank citation ranking: Bringing order to the web. Technical report. Stanford InfoLab.

Pasteur, R. D. [2010]: When perfect isn't good enough: Retrodictive rankings in college football. Megjelent: *Gallian, J. A.* (szerk.): Mathematics & Sports. Dolciani Mathematical Expositions 43. Mathematical Association of America, 131–146. o.

Radicchi, F. [2011]: Who is the best player ever? A complex network analysis of the history of professional tennis. PloS one, Vol. 6. No. 2. e17249.

Slater, P. [1961]: Inconsistencies in a schedule of paired comparisons. Biometrika, Vol. 48. No. 3–4. 303–312. o.

Slikker, M.–Borm, P.–van den Brink, R. [2012]: Internal slackening scoring methods. Theory and Decision, Vol. 72. No. 4. 445–462. o.

Slutzki, G.–Volij, O. [2006]: Scoring of web pages and tournaments – axiomatizations. Social Choice and Welfare, Vol. 26. No. 1. 75–92. o.

Spence, M. [1973]: Job market signaling. The Quarterly Journal of Economics, Vol. 87. No. 3. 355–374. o.

Szulc, B. [1964]: Indeksy dla porownan wieloregionalnych. Przeglad Statysticzny, Vol. 3. 239–254. o.

Telcs András–Kosztyán Zsolt Tibor–Török Ádám [2012]: Unbiased one dimensional University Ranking – application based preference ordering. Közlésre benyújtva.

Telcs András–Kosztyán Zsolt Tibor–Török Ádám [2013]: Hallgatói preferencia-sorrendek készítése az egyetemi jelentkezések alapján. Közgazdasági Szemle, 60. évf. 3. sz. 290–317. o.

Thurstone, L. L. [1927]: A law of comparative judgment. Psychological Review, Vol. 34. No. 4. 273–286. o.

van den Brink, J. R.–Pintér Miklós [2012]: On axiomatizations of the Shapley value for assignment games. Tinbergen Institute Discussion Paper, TI 2012-092/II.

Young, H. P. [1974]: An axiomatization of Borda's rule. Journal of Economic Theory, Vol. 9. No. 1. 43–52. o.

Zermelo, E. [1929]: Die Berechnung der Turnier-Ergebnisse als ein Maximumproblem der Wahrscheinlichkeitsrechnung. Mathematische Zeitschrift, Vol. 29. No. 1. 436–460. o.