

Hogyan rangsoroljunk, ha muszáj?

*Gondolatok Amy N. Langville és Carl D. Meyer könyve nyomán**

Schubert András

MTA Könyvtár és Információs Központ, Tudománypolitikai és Tudományelemzési Osztály, Budapest

Napjainkban életünket minden eddigénél jobban átszövik a rangsorok. Ez a tudomány világában sincs másként. A rangsorok készítőin igen nagy a felelősség, mert eredményeiken sokak karrierjét, egzisztenciáját befolyásoló döntések múlhatnak. Az ismertett könyv – ugyan csak a rangsorok egy speciális területén – példát mutat arra, hogy hogyan lehet a módszerek egy rendszeres gyűjteményét összeállítani, hogyan kell elméleti próbákat és gyakorlati példákat gondosan kiválasztani, és hogyan lehet különféle módszereket egy összefoglaló, többdimenziós értékelési és rangsorolási rendszerre kombinálni. Orv. Hetil., 2015, 156(32), 1298–1300.

Kulcsszavak: tudománymetria, rangsorok

How to rank if you must?

Thoughts about the book of Amy N. Langville and Carl D. Meyer

More than ever, our life is permeated by rankings. This is true also in the world of scientific research. The responsibility of the producers of such rankings is significant, since their results may influence decisions determining human fates and careers. The reviewed book - in a rather special area of ranking only - exemplifies how a set of methodologies can be systematically compiled, and how various methods can be combined into comprehensive, multidimensional rating and ranking systems.

Keywords: scientometrics, ranking

Schubert, A. [How to rank if you must? Thoughts about the book of Amy N. Langville and Carl D. Meyer]. Orv. Hetil., 2015, 156(32), 1298–1300.

(Beérkezett: 2015. május 20.; elfogadva: 2015. május 21.)

Az írás címe *Dubins* és *Savage* klasszikus könyvére utal [2]. Az a könyv azt mutatta meg, hogy mi az optimális stratégia a rulettjátékban, ha feltétlenül játszani akarunk, miközben tudjuk, hogy a tényleges optimális stratégia az, ha egyáltalában nem játszunk. A várható veszteség ekkor a legkisebb (nulla), minden más esetben – tehát ha játszunk – nullánál nagyobb. Ha azonban a cél nem a veszteség minimalizálása, hanem egy bizonyos nyereség elérési esélyének maximalizálása, akkor a helyzet bonyolultabb; erről szólt *Dubins* és *Savage* könyve.

A rangsorolások esetében a „játékos” (a rangsorolás készítője) lehetséges vesztesége általában nem anyagi, inkább erkölcsi. Napjainkban életünket minden eddigénél

jobban átszövik a rangsorok. A Google találati listái, az Árukereső ajánlatai, a könyv- és zenekiadás népszerűségi listái mindannyiunk mindennapi döntéseit alapvetően befolyásolják. Nincs ez másként a tudomány világában sem. Kutatók, kutatóegységek, egyetemek publikációs, idézettségi rangsorain súlyos támogatási pénzek múlhatnak; a rangsorolások készítői „csak” szakmai hitelüket kockáztatják. Számukra a veszteség minimalizálásának az optimális stratégiája itt is az, ha egyáltalán nem készítenek rangsorokat. Sok irányadó szakértő igyekszik is a lineáris rangsorok iránti igényt féken tartani, és az összetettebb, többdimenziós elemzések irányába terelni az érdeklődést. Egy bizonyos ponton, a döntéshozatal előtt

*Ez az írás a *Scientometrics* című folyóiratban megjelent könyvismertetés [1] egyik részének kibővített változata.

azonban már óhatatlan az „egydimenziós” vagy éppen séggel „bináris” rendezés igénye. Ilyenkor lenne szükség egy *Dubins-* és *Savage*-féle iránymutató segédletre: „Hogyan rangsoroljunk, ha muszáj?”

Valami ilyet reméltem *Amy N. Langville* és *Carl D. Meyer* könyvétől [3].

Igen helyesen a szerzők nem azt a címet adták könyvüknek, hogy „Minden, amit tudni akartál az értékelésről és rangsorolásról...”. Az csak vágyálom lehet, hogy egy olyan könyvet talál az olvasó, amely átfogó megoldást ad mindarra az értetlenségre és érthetlenségre, ami az értékelést és a rangsorolást körülveszi akár a tudománymetriában, akár másutt. A „Ki az első?” nem orákulum, hanem egy nagyon jó könyv, amely több kérdést vet fel, mint amennyit megválaszol. E kérdések közül pedig néhány sokkal érdekesebb, mint amikre választ vártunk.

A szerzők az algebrai rangsorolás-elmélet jeles szakértői. Előző könyvük, a „Google’s PageRank and Beyond” (Princeton University Press, 2006) több mint 200 idézetet kapott a Web of Science-ben. Az Amazon Best Seller rangsorban a 254 349-ik helyezést foglalja el. Ez első látszatra csalódást keltő, de ha tekintetbe vesszük, hogy *Derek de Solla Price* klasszikusa, a „Little Science, Big Science...and Beyond” a 4 025 308-ik helyen található, a hat számjegyű helyezés már figyelemre méltó. A „Ki az első?” néhány hónappal a kiadása után már a 282 234-ik helyezésnél járt, tehát bizonyára hamarosan megelőzi a szerzők előző sikerét. Akar valaki fogadni?

Ami azt illeti, az Amazon értékelései és rangsorai, valamint a fogadási stratégiák a könyv visszatérő témái, ami azt jelzi, hogy a szerzők ügyesen hígítják az algoritmusok elméletét a népszerű kultúrából és mindenekelőtt a sportból vett példákkal. Én magam régi lelkes híve vagyok a tudomány és a sport közötti analógiák feltárásának, de sajnos igen tájékozatlan vagyok az amerikai futball világában, amely a könyv fő edzőpályája. A könyvet végigkísérő példaadatok a keleti parti konferencia csapatai közötti mérkőzések egy halmazára vonatkoznak az NCAA 2005-ös évadából. Mire a könyv végére értem, már kezdtem hozzászokni...

A könyv az első részében egy sor rangsorolási algoritmust mutat be. *Massey*, *Colley* és *Keener* módszereit valós példákon használja rangsorolási és előrejelzési feladatokra. A szerzők saját kutatási területe, a Markov-láncok és a támadási-védekezési értékelési módszer megfelelő terjedelemben és érthető szenvedéllyel kerül terítékre. Azok számára, akiknek a sakk közelebb áll a szívéhez, mint az amerikai futball, nagyon tanulságos látni, hogy az Élő-féle pontozási rendszer hogyan illeszkedik a könyvben vázolt rendszertanba. *Élő Árpád* honfitársaiként különösen dicséretesnek találhatjuk, hogy a szerzők szinte minden alkalommal hibátlanul írják le a nevét. Természetesen ez a módszer sem menekül meg egy amerikai futballpéldán történő alkalmazástól. Figyelemre méltóan sikeresnek bizonyul előrejelzési eszközként, ami nemcsak nemzeti büszkeségünket ébresztheti, hanem érdek-

lődésünket is e zseniális módszernek a tudományos értékelés és rangsorolás területén való lehetséges alkalmazásai iránt.

A holtversenyek eldöntésére és a súlyozási módszerekre vonatkozó speciális fejezetek után két terjedelmes fejezet foglalkozik a rangsorok egyesítésének a kérdésével. Ez a rész is kizárólag az algebrai módszerekkel foglalkozik, ezen belül megkülönböztetve a heurisztikus és az optimumkereső módszereket. A könyv technikai részét a rangsorok összehasonlításával (a konszenzus mérésével) foglalkozó fejezet zárja le. Az elérhető adatforrásoknak egy rövid, de izgalmas listája, valamint a nem algebrai (főként statisztikai) értékelési és rangsorolási módszerek áttekintése egészíti ki a könyvet.

A könyv stílusa bámulatos. Az egzakt matematika simán keveredik a világos magyarázatokkal és az illusztratív példákkal. Különleges adalékként szórakoztató kitérők szolgáltatnak kiegészítő információkat, néha csak lazán, de mindig odaillően kapcsolódva a fejezet fő témájához.

Minden elragadtatás és gyönyörködés mellett nem mondanám, hogy a könyv a tudománymetriával foglalkozók számára közvetlenül hasznos lenne. A bemutatott módszerek bevallottan mind a páronkénti összehasonlításra (két ellenfél közötti mérkőzések eredményein) alapulnak. Számos példa esetében (például az Amazon példáján vagy más felhasználói értékelésen alapuló rangsorolás esetén) nem ez a helyzet, de megkísérlelhető a probléma párok közötti „virtuális mérkőzések” való transzformálása. Ez az ügyes trükk néha tökéletesen működik, de például a tudománymetria legtöbb gyakorlati értékelési és rangsorolási feladata esetében meglehetősen természetellenes és indokolatlan, esetenként egyenesen kivitelezhetetlen lenne.

Egy másik szempont, amit a szerzők nem hangsúlyoznak kellőképpen a sportból vett példáikon sem, az, hogy minden értékelés vagy rangsorolás csak valamilyen jól meghatározott cél ismeretében értelmezhető. Egészen mások a követelmények egy értékelési/rangsorolási módszerrel szemben, ha a kérdés tényleg az, hogy „Ki az első?”, mint ha például egy verseny kiemelési rendszerét dolgozzuk ki. Tapasztalatom szerint a valóságos tudománymetriai értékelések/rangsorolások legfőbb fogyatékossága a jól meghatározott cél hiánya. Mutatószámokat javasolnak anélkül, hogy pontosan definiálnák, hogy mit is kellene mutatni, ranglistákat közölnek anélkül, hogy bármit is tudnánk arról, hogy mi a rangsorolás célja. Ezt a gyakorlatot nem lehet eléggé elítélni, és ez nem orvosolható a legpontosabb és kidolgozottabb értékelési és rangsorolási algoritmussal sem.

A tudománymetria közössége számára e könyv haszna közvetett: egy hozzáállást tanulhatunk, hogyan kell a lehetséges módszerek egy rendszeres gyűjteményét összeállítani, hogyan kell elméleti próbákat és gyakorlati példákat gondosan kiválasztani, és hogyan lehet különféle módszereket egy összefoglaló, többdimenziós értékelési

és rangsorolási rendszerre kombinálni. Ebben az értelemben tekinthetjük ezt a kötetet a tudománymetria iránt érdeklődő olvasók számára nagyon ajánlatos könyvnek.

Anyagi támogatás: A szerző köszöni az FP7-SSH-2013-2 #613202 (IMPACT-EV) projekt támogatását.

A cikk végleges változatát a szerző elolvasta és jóváhagyta.

Érdekltségek: A szerzőnek nincsenek érdekltségei.

Irodalom

- [1] *Schubert, A.*: How to rank if you must: two useful guiding books. *Scientometrics*, 2014, 98, 763–765.
- [2] *Dubins, L. E., Savage, L. J.*: How to gamble if you must: inequalities for stochastic processes. McGraw-Hill series in probability and statistics. McGraw-Hill, New York, 1965.
- [3] *Langville, A. N., Meyer, C. D.*: Who's #1? The Science of rating and ranking. Princeton University Press, Princeton and Oxford, 2012.

(Schubert András
e-mail: schuba@iif.hu)

25 éves orvostalálkozó

A SZOTE Általános Orvosi Karán 1990-ben végzett orvosok évfolyamtalálkozót szerveznek
Szegeden 2015. október 10-én.

Szervező iroda: Top Congress Kft. Szeged

Telefon: 30-968-2725

E-mail: info@top-congress.hu