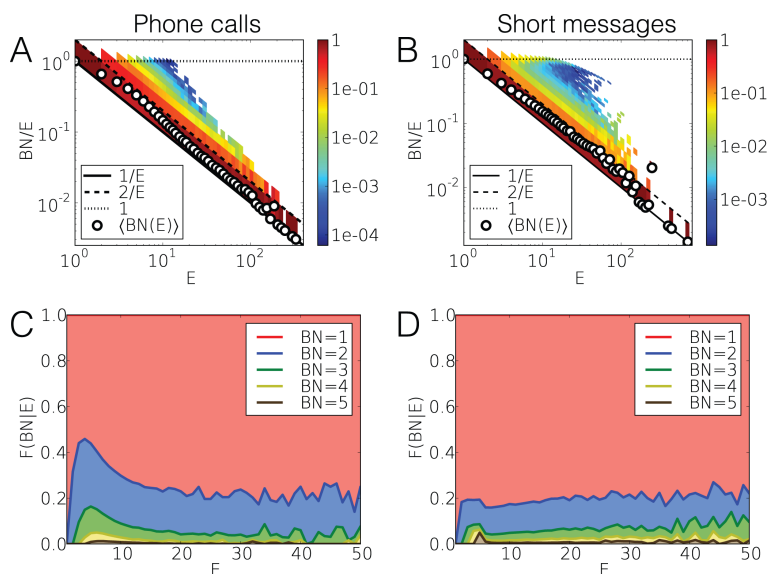


Válaszok Dr. Palla Gergely bírálata:

1. A 2.4.1 fejezet a humán egocentrikus kommunikációs hálózatok „bursty” jellegével foglalkozik, ahol azt olvashatjuk, hogy a kommunikációs események sorozatában fellelhető „bursts” jelleg egyik intuitívnek tűnő magyarázata lehetne az, hogy egy egyén által kezdeményezett, de több különböző másik személlyel folytatott interakciók között erős időbeli korreláció léphet fel. Azonban az empirikus adatok elemzése erre rácsúfol és azt mutatja, hogy a „bursty” jelleg elsősorban az egyes éleken zajló kommunikáció sajátosságainak következménye. Ezzel kapcsolatban szeretném kérdezni azt, hogy mennyire jó modellje az emberi kommunikációnak az, ha egy „ego” éleit egymástól függetlennek tekintjük? Van-e bármi olyan effektus amihez mégis szükség van az egy csúcsból induló de különböző éleken folytatott kommunikációk közti korrelációra?

Ahogy a Bíráló összefoglalta, a 2.4.2-es fejezetben arra a meglepő megfigyelésre jutottunk, hogy egy egyén klaszteresedett villanásos („bursty”) kommunikációs eseménysorozatában, nagyrészt csak egy másik hálózati szomszéd vesz részt. Ez a megfigyelés azt sugallja, hogy az általános jelenség kialakulása mögött nagyobb szerepe van a két egyén közötti kommunikációnak, mint csoportos kommunikációs jelenségeknek. Ez a jelenség egyébként a 3.4.1-es fejezetben is visszaköszön (valamint a [223]-as publikációban), ahol a mobil-kommunikációs hálózatokban leggyakrabban megfigyelt temporális motívum pont az a bursty eset, amikor egy élen több egymást követő kommunikációs esemény történik két egyén között.

Azonban, a Bíráló kérdésére válaszolva, számos más esetben előfordul, hogy bursty eseménysorozatok ez ego és több szomszéd között alakulnak ki. Az ilyen elosztott bursty klaszterek előfordulási számára a 2.4.2-es fejezetben tárgyalt [203]-as cikkem 3.a-d ábrája



1. ábra: Bursty eseménysorozatokban résztvevők számának statisztikai leírása

(a, b) A bursty vonatokban résztvevő szomszédok BN számának a vonat E méretéhez viszonyított arányainak eloszlása és átlagértékei az E függvényében (a) telefonhívások és (b) sms-ek esetén. A folytonos és szaggatott vonalak a $BN=1$ és 2 esetet jelölik. (c, d) Az E méretű bursty vonatokban résztvevő szomszédok számának (BN) kumulatív tört része, (c) telefonhívások és (d) sms-ek esetén.

ad egy jól értelmezhető választ. Ez terjedelmi okokból nem került be a dolgozatba, de a bírálói kérdésre adott válaszhoz itt csatolom.

Ezekből az eredményekből az látszik, hogy habár minden vonat méret esetén leggyakrabban a bursty sorozat az ego egy szomszédjával kommunikálva alakul ki, ezért a vonatok egy nem elhanyagolható hányada két vagy akár több szomszéd bevonásával jelenik meg, függetlenül a kommunikációs csatornától. Számos oka lehet a megfigyelt vonatokban résztvevő események ilyen fajta eloszlásának, azonban ezek kiderítéséhez a kommunikációk tartalmát is ismernünk kellene, ami adatvédelmi szempontok miatt nem érhető el. Így csak intuitív választ tudok adni erre a kérdésre. Amíg a legtöbbet megfigyelt, párok közötti bursty vonatok valószínűleg egy adott téma megbeszélése vagy feladat megoldása indukál, a több ember között kialakuló bursty vonatok csoportos szervezés vagy csoportos koordináció következménye lehet. Ugyanakkor a több embert bevonó bursty vonatok viszonylagos hiánya a vizsgált kommunikációs csatornák természetéből is adódhatnak. Míg a telefonhívások megengednek két irányú kommunikációt, sms kommunikáció szigorúan egyirányú, így két ember közötti beszélgetés természetesen egy hosszabb bursty vonathoz vezet. Ugyanakkor, habár technikailag lehetséges, ezeket a csatornákat általában nem szokták csoportos kommunikációra használni. Így egy egyén befektetett kommunikációs erőfeszítése a csoport méretével lineárisan nőhet, ritkábbá téve a több embert bevonó bursty vonatok megjelenését.

2. Formálisan úgy tűnik, mintha gond lenne a (2.15) egyenlettel: Az egyenlet felett lévő szöveg nem mondja, hogy b_e függene E -től és a b_e -re megadott formula sem mutat explicit E függést. Azonban ha b_e nem függene E -től, akkor a (2.15) egyenletben kiemelhetnénk a szumma elé ami biztosan nem a kívánt eredményre vezetne. Esetleg itt a b_e (véltetően E -től függő) definícióját szerencsés lenne ismertetni.

A Bíráló által említett egyenlet a következőképpen jelenik meg a disszertációban:

$$b_{All}(E) = \langle b_e \rangle = \frac{\sum_e n_e(E) b_e}{\sum_e n_e(E)} \quad (1)$$

ahol $n_e(E)$ egy e élen kialakult E hosszúságú bursty vonatok számát jelöli, és ezek a nevezőben szereplő szummája a hálózatban megfigyelt összes E hosszúságú vonat teljes számát jelöli. Ugyanakkor, $b_e = \max(N_A, N_B) / (N_A + N_B)$ egy adott $e = \{A, B\}$ él teljes kommunikációs egyensúlyát méri az egész megfigyelt periódusra. Ebben a definícióban N_A az A->B irányába kezdeményezett kommunikációs események számát, míg N_B a B->A irányába kezdeményezett események számát jelenti az e élen. Így $1/2 \leq b_e \leq 1$ értékeket vehet fel, ami határesetben a kommunikáció teljesen kiegyensúlyozottságát (amikor $N_A =$

N_B vagyis $b_e = 1/2$), vagy a teljes kiegyensúlyozatlanságát (amikor $N_A = 0$ vagy $N_B = 0$ vagyis $b_e = 1$) jelenti.

Ezen definíciók alapján b_e nem függ E -től, hanem egy minden e élre külön kiszámolt mennyiség. Így nem emelhető ki az éleken végigfutó szumma elé, hanem mint a súlyozott-összegzés súlyaként értelmezhető minden egyes élre. Az így számolt $b_{All}(E)$ mennyiség azt adja meg, hogy egy adott kommunikációs (esetünkben mobilhívás vagy sms) hálózatban kialakuló E hosszúságú bursty vonat átlagosan mennyire van egyensúlyban.

3. A 2.10-es ábra f) paneljén az $E \rightarrow 0$ határesetnél a görbék látszólag 2/3-hoz tartanak, illetve a g) panelnél és a 2.12 ábra c) paneljénél meg 3/4-hez ugyanebben a határesetben. Van-e ezeknek a viselkedéseknek egy egyszerű szemléletes magyarázata? Illetve van-e olyan triviális kommunikációs mintázat, amikor a teljes E tartományon 2/3 illetve 3/4-en marad a vonatkozó érték?

Nagyon érdekesek a Bíráló megjegyzései, amikre a következők a válaszaim:

- A 2.10(f) ábra a fenti (1)-es egyenletben definiált $b_{All}(E)$ egyensúlyi mértéket mutatja telefonhívás és sms szekvenciákra számolva. Ahogy az előző válaszban tárgyaltam, $b_{All}(E)$ két további mértéktől függ, amik közül b_e egy él teljes kommunikációs egyensúlyát jelöli és $n_e(E)$ pedig az adott élen kialakult E hosszúságú vonatok számát. Ezek alapján $b_{All}(E)$ azt méri, hogy adott hosszúságú vonatok átlagosan milyen egyensúllyal rendelkező éleken alakulnak ki. A tanulmány szempontjából fontos megfigyelés az volt, hogy $E \gg 1$ esetben a telefonhívás szekvenciákban megfigyelt hosszabb vonatok egyre kiegyensúlyozatlanabb éleken alakulnak ki, míg az ellenkezője igaz sms vonatokra. A Bíráló által említett $E \rightarrow 1$ esetben $b_{All}(E)$ egyre kiegyensúlyozottabb értéket mutat. Ez azért van, mert minél rövidebbek a vonatok, annál több van belőlük, ami a 2.10(d) ábrán mutatott eloszlásból is következik. Így egyre reprezentatívabb mintát közvetítenek az élek teljes halmazáról, amibe egyre több jobban kiegyensúlyozottabb él is beletartozik ahogy $E \rightarrow 1$. Azonban a $b_{All}(E)$ mérték még az $E = 1$ határesetben sem lesz egyenlő az élek átlagos b_e egyensúlyi mértékével, mert függ attól, hogy hány $E = 1$ hosszúságú vonat alakul ki a különböző egyensúlyú éleken. Ezen logika alapján nem szükségszerű, hogy $b_{All}(E)$ a 2/3 közeli értékeket vegyen fel az $E \rightarrow 1$ esetben. Azonban, természetesen lehetséges, hogy $b_{All}(E)$ ezt, vagy bármely más $0.5 \leq b_{All}(E) \leq 1$ értéket vegyen fel tetszőleges E -re, ha az élek egyensúlyi mértéke és az adott hosszúságú vonatok a különböző éleken való eloszlása megfelelően viszonyban van. Azonban, ahogy ez a disszertációból kiderül, ez nem jellemző az megfigyelt empirikus rendszerekre.

- A 2.10(g) ábrán mutatott eredmények az egyes E hosszúságú vonatokat jellemző átlagos $p_m(E) = \max(n_A, n_B)/n_A + n_B$ egyensúlyt mutatják meg mobilhívás és sms, valamint a hozzájuk tartozó korrelálatlan modellek esetén. Itt n_A (resp. n_B) egy E hosszúságú vonatban szereplő A (ill. B) által indított események számát jelöli. Az egyik fontos konklúzió ebben az esetben az, hogy $E \gg 1$ határesetben minél hosszabb egy vonat, annál kiegyensúlyozatlanabb, ha mobilhívásokból áll (és az ellenkezője igaz sms-ekre). $E = 1$ esetben $p_m(E)$ és annak minden ilyen méretű vonatra számolt átlaga triviálisan $p(E) = \langle p_m(E) \rangle_m = 1$, ahogy ez az ábrán látszik is. Az $E = 2$ esetben azonban p_m két lehetséges értéket vehet fel. Ha az m indexsel jelölt vonatban szereplő két interakciót ugyan az a fél indította (vagy A vagy B), akkor $p_m = 1$, valamint, ha az egyik interakciót A és a másik eseményt B indította, akkor $p_m = 1/2$. Tehát az ebben az esetben kialakult $p(E)$ átlagos egyensúly 0.75 értéke azt jelentené, hogy a kettő hosszúságú vonatok fele teljesen kiegyensúlyozott és a másik fele pedig teljesen kiegyensúlyozatlan. Azonban közelebbről megtekintve az ábrát, egyik függvény sem veszi fel pontosan a $p(E = 2) = 0.75$ értéket. Hívások esetén már $E = 2$ esetén is $p(E) > 0.75$ értéket vesz fel, tehát egy gyenge átlagos kiegyensúlyozatlanságot jelöl. Ezzel szemben sms-ek esetén $p(E) < 0.75$, vagyis már a rövid vonatok is átlagosan jobban kiegyensúlyozottabbak. Mivel mind $p_m(E)$ mind $p(E)$ bármely értéket felvehet $1/2$ és 1 között, természetesen lehetséges, hogy tetszőleges E esetén az átlaguk 0.75 -ös értéket mutasson, például, ha egyik felük teljesen kiegyensúlyozott és a másik felük teljesen kiegyensúlyozatlan. Azonban az eredmények alapján kijelenthetjük, hogy ez egyik vonathossz esetén sem igaz.

- 4. A 2.12 ábra c) paneljén úgy tűnik, mintha a kék négyzetből két görbe is lenne, egy a sárga négyzet görbét szorosan követve, illetve egy másik görbe ami valamivel lejjebb halad. Ez két külön görbe, vagy pedig egy darab erősen fluktuáló görbe, ami hol felugrik a sárga négyzet görbéjéhez közel, hol vissza a lenti burkolórá? Akármelyik eset is áll fent, mi ennek a magyarázata?**

A Bíráló által megfigyelt jelenséget egy erősen fluktuáló görbe alakítja ki, ami az E hosszúságú kommunikációs vonatok átlagos $p(E)$ egyensúlyát jeleníti meg egy modellezett rendszerben. A valós megfigyelések alapján, egy sms szekvenciákat modellező folyamat definíciója során azt a feltételezést tesszük, hogy a nagyobb E értékekre egyre erősödő egyensúlyt az sms irányított kommunikációs tulajdonsága alakítja ki. Ezt úgy modellezzük, hogy minél hosszabb egy vonat, annál nagyobb valószínűséggel folytatjuk, mitöbb, a következő esemény annál nagyobb valószínűséggel jelenik meg ellentétes paritással, mint a megelőző (ahogy ez a disszertáció 2.20-as egyenletében formálisan is

definiálva van). Ez a fordított paritási preferencia okozza azt, hogy a páros hosszúságú vonatok átlagosan jobban kiegyensúlyozottabbak, míg a páratlan hosszúságúak kevésbé. Fontos megjegyezni, hogy ez a jelenség nem csak modellezett sms szekvenciák esetén jelenik meg, hanem kisebb mértékben a 2.10(g)-es ábrán is megfigyelhető valós sms szekvenciák esetén, így erősítve meg a modellfeltevést.

5. A 3.3 alfejezet időfüggő hálózatokon zajló SI terjedési folyamatok érdekes tulajdonságait mutatja be. A statikus hálókon zajló terjedési folyamatokhoz képest itt egy fontos különbség, hogy a terjedésnek tiszteletben kell tartania a kapcsolatok megjelenésének időbeliségét is. Ezzel kapcsolatban kérdezném meg, hogy a vizsgált hálózatoknak van-e „időtükrozási szimmetriája” a következő értelemben: Tegyük fel, hogy először megfordítjuk az adatokban az időbeliséget, ami olyan, mintha ha a hálózat időfejlődését visszafelé játszanánk el. Ezután ezen a megfordított hálón végezzük el ugyanazokat az SI terjedési modellre épülő vizsgálatokat mint a dolgozatban. Vajon az így tapasztalt viselkedés ugyanaz lesz (statisztikai értelemben) mint eredetileg, vagy jól elkülöníthetően más?

A Bíráló javaslatát követve tegyük fel, hogy a 3.3 alfejezetben tárgyalt időfüggő hálózat kölcsönhatásait, melyeket először irányítatlannak feltételezünk, fordítva játszunk le az időben. Ha az így kapott hálózaton determinisztikus SI terjedési folyamatokkal (ahol a β fertőzési valószínűség 1) elvégezzük ugyanazokat a kísérleteket, természetesen az egyes kísérletek kimenetele nem lesz ugyanaz, még akkor sem, ha ugyanabból a csúcspontból indítjuk a folyamatot. Ez azért van, mert a fordított hálózaton más sorrendben jelennek meg az események, így más időfüggő utak alakulnak ki, melyek más csúcspontokhoz, más sorrendben és más sebességgel juttatják el a terjedő fertőzést. Ez akkor még inkább igaz, ha az időbeli kölcsönhatásokat irányítottan feltételezzük, és az idővel együtt nem fordítanánk meg az események orientációját. Ha megfordítjuk, akkor az irányítatlan esettel ekvivalens problémát kapunk.

Azonban, a fordított időben hálózatban is jelen lesznek azok strukturális és időbeli korrelációk (például csoport struktúra, időbeli klaszteresedés, időfüggő utak, vagy akár a napi periodicitás), melyek befolyásolják a szimulált terjedési folyamat sebességét. Tehát, ha a kísérleteket ugyanazon a véletlen referencia modelleken elvégezzük el, a kvalitatív konklúzió ugyanaz lesz annak ellenére, hogy kvantitatív módon nem feltétlen ugyan azt az eredményt fogjuk kapni. Másképp fogalmazva, mivel az egyes szimulált SI folyamatok más időfüggő utak mentén és más sebességgel futnak le az időben fordított hálózaton, az empirikus és véletlen referencia modelleken eredetileg mért terjedési sebességeket nem fogjuk visszakapni. Azonban a strukturális és időbeli korrelációknak ugyanolyan hatása kell, hogy legyen a terjedési folyamatokra, tehát az őket elimináló véletlen referenciakísérletekből

származó átlaggörbék ugyan olyan sorrendben kell, hogy megjelenjenek, mint az eredeti kísérletben.

6. A 3. fejezet végén az egocentrikus hálók időfejlődésével kapcsolatos eredményekről, modellekről olvashatunk. Ezzel kapcsolatban szeretném kérdezni azt, hogy az egocentrikus háló időbeli bővülését összekapcsolták-e már korábban a nyelvészetben ismert Heap-törvénnyel (vagy Heaps Herdan-törvénnyel), mely azt mondja ki, hogy az egymástól különböző szavak száma egy dokumentumban a dokumentum méretétől hatványszerűen függ? Első ránézésre egy olyan háló bővülést leíró modell melyben az ego véletlenszerűen vagy már meglévő kapcsolatokat erősít meg, vagy teljesen új kapcsolatokat hoz létre nagyon hasonlít azon egyszerű szöveggeneráló modellekre, melyben véletlenszerűen vagy egy már használt szóval folytatjuk a dokumentumot, vagy egy teljesen új szót írunk le.

A Bíráló által javasolt kapcsolat az egocentrikus hálózatok időbeli bővülése és a Heaps Herdan-törvény között egy nagyszerű ötlet. Ez a kapcsolat a disszertációban összefoglalt tudományos közleményeknek nem tárgya, azonban a közelmúltban két korábbi társszerző által történtek ilyen irányú kutatások:

- Ubaldi, E., Burioni, R., Loreto, V., & Tria, F. (2021). Emergence and evolution of social networks through exploration of the Adjacent Possible space. *Communications Physics*, 4(1), 28.

Ebben a cikkben a szerzők felismerik az egocentrikus hálózat fejlődő mérete és a Heaps-törvény közötti kapcsolatot és kiszámolják az egocentrikus hálózatokban mért Heaps-exponensek eloszlását különböző valós hálózatokon. Ez alapján egy modellt vezetnek be ami a Pólya-féle urnamodell egy több ágensre épülő változata, amelyet sikeresen alkalmaznak az egocentrikus hálózatokból kifejlődő hálózat legfontosabb statisztikai tulajdonságainak (Zipf-, Heaps- és Taylor-törvények) és dinamikus korrelációinak reprodukálására.

- Di Bona, G., Bellina, A., De Marzo, G., Petralia, A., Iacopini, I., & Latora, V. (2023). The dynamics of higher-order novelties. arXiv preprint arXiv:2307.06147.

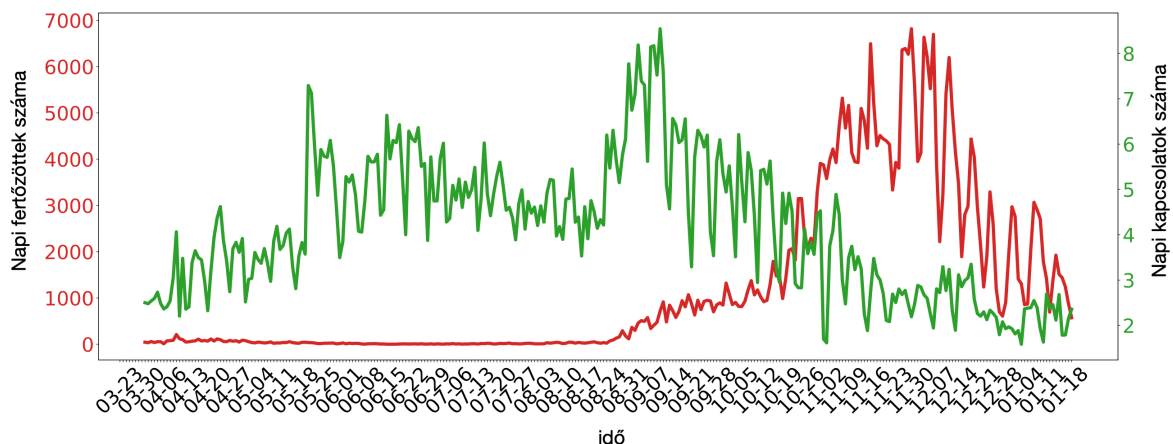
Ebben a kéziratban a szerzők több attribútumú elemek rendezett sorozatát vizsgálják, ami egy egocentrikus hálózatban több tulajdonsággal leírt új szomszédok megjelenési dinamikáját is leírja. Azt ismerik fel, hogy a magasabb rendű újdonságokat két vagy több tulajdonság kombináció első megjelenéseként jellemezve, az elemek felfedezési dinamikáját egy magasabb rendű Heaps-exponenssel lehet leírni. Azonban valós adatok széleskörű

elemzéséből kiderül, hogy a magasabb rendű Heaps-exponensek eltérhetnek a rendszerekben mért hagyományos, alacsonyabb rendű Heaps-exponens értékektől.

Mindezek alapján megállapítható, hogy a Heaps-Herdan-törvény jelentőségét már felismerték az egocentrikus hálózat fejlődésével kapcsolatban, azonban ennek teljes megértése további kutatást igényel.

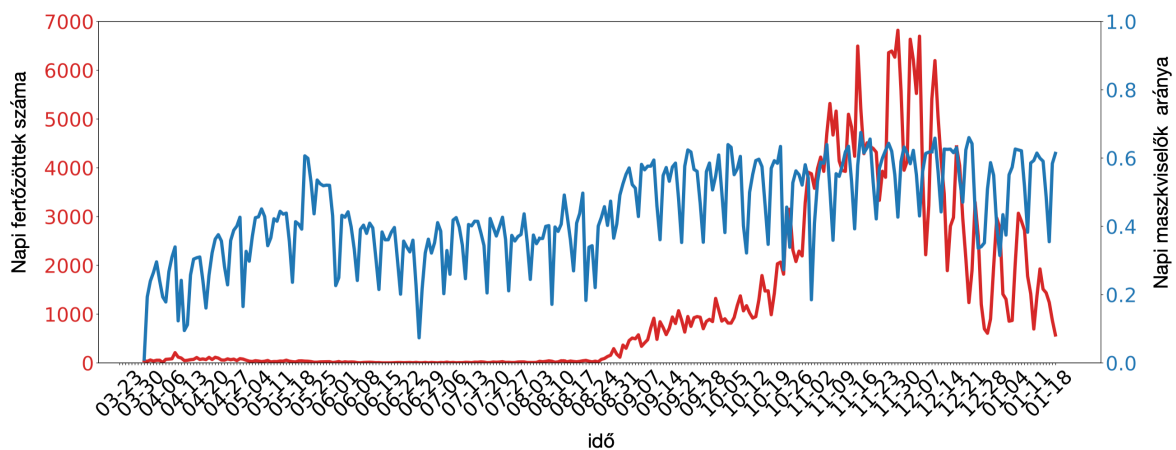
7. A 4. fejezet érdekes eredményeket mutat be a magyarországi Covid-19 terjedés első két hullámának összehasonlításáról. Ennek kapcsán felmerül az az aspektus is, hogy esetleg a második hullám alatt általánosan lazább lehetett a fegyelem a lakosság részéről, hiszen az első hullámot már átvészelte az ország. Voltak-e erre utaló jelek bármelyik hozzáférhető adatforrásban (p.I MASZK), illetve hogyan befolyásolja az eredmények értelmezését ha egy ilyen lehetőséget is figyelembe akarunk venni?

Ahogy a Bíráló megjegyezte, a pandémia ellen bevezetett intézkedések és a vészhelyzet percepciója is nagyban változott a magyarországi első két COVID-19 hullám alatt. Míg az első hullám alatt általános lezárásokkal és kijárási korlátozásokkal próbálták az emberek közötti fertőzés terjedésének rátáját csökkenteni, ezek az intézkedések együtt jártak az emberek viszonylagos normakövetésével. A második hullám alatt kevésbé szigorúbb és jobban differenciált intézkedések léptek életbe, amik főleg személyes védekezést előíró intézkedéseket tartalmaztak (maszkviselés, távolságtartás), és így igyekeztek a pandémiát kordában tartani. Habár a második hullámot megelőzően az általános veszélyérzet alacsony lehetett, mivel a második hullám társadalomra gyakorolt hatása (magas megbetegedési, kórházba kerülési, és halálozási arányok) sokkal nagyobb volt, mint az első, a hullám elindulása után a mérhető viselkedési mintázatok ismét az intézkedések viszonylagos betartását tükrözték.



2. ábra: Fertőzöttek és átlagos közelségi kapcsolatok számának napi alakulása Magyarországon a COVID-19 első két hulláma alatt. (forrás: MASZK)

Erre egyik jó példa az átlagos kapcsolatszámok alakulása (2. ábra), amiből egyértelműen látszik a kapcsolatok emelkedése az első hullámot követő enyhítések során, és azok csökkenése amint a második hullám sohasem látott megbetegedési számokat okozott. Egy másik jól követhető viselkedési forma a maszkot viselők arányának napi alakulása (3. ábra). Míg az első hullám után az így védekezők aránya lecsökkent, igazából ez az arány még 2020 nyarán sem csökkent nullára, és a második hullám alatt bevezetett kötelező maszkviselési rendeletek hatására emelkedett ismét magasabbra, habár a méréseink alapján sosem érte el a 100%-ot.



3. ábra: Fertőzöttek számának és maszkviselők arányának napi alakulása Magyarországon a COVID-19 első két hulláma alatt. (forrás: MASZK)

A Bíráló második kérdésére válaszolva, természetesen a viselkedési normák változásának fontos hatása van a megfigyelt fertőzési folyamat alakulására és ezek valamilyen módon való beépítése prediktív modellekbe nagyban segíti azok pontosságának növelését. Ebben az irányban a közelmúltban mi is további kutatásokat végeztünk, ahol egyén szintű adaptív kapcsolattartási mechanizmusokat figyelembe véve újabb érdekes jelenségeket figyeltünk meg. Azonban a disszertációba összefoglalt kutatások során nem erre a jelenségre koncentráltunk. Ott egy olyan egyszerű modellt próbáltuk felállítani, amely képes a hálózat magjából vagy perifériájáról elindított terjedési folyamatok által okozott átkapcsolási (“switchover”) jelenségek reprodukálására, és célunk annak alapvető matematikai megértése volt.

Budapest, 2023. december 4.

Dr. Karsai Márton