

Járványterjedési folyamatok paraméterbecslése többrétegű véletlen gráfokon

Krajnyák Bálint

Témavezető: Backhausz Ágnes

Bevezetés

A projekt keretében járványterjedési szimulációkat hajtottam végre, és arra voltam kíváncsi, hogy bizonyos paraméterek megváltoztatása milyen hatással van a járványterjedésre vonatkozóan. Szakirodalom tekintetében elsősorban a [2]-es sorszámú forrásra támaszkodtam. Ezen cikkben négyrétegű hálózaton lettek szimulációk elvégezve MATLAB szoftver segítségével, amely programkódok velem is meg lettek osztva, és ezen kódok jelentették az általam elvégzett szimulációk kiindulópontját.

A modell

A járványterjedést kétrétegű véletlen gráfok segítségével modelleztük. Tekintünk egy populációt, amelyben minden egyes személyt megfeleltetünk a véletlen gráf egy csúcsának. A csúcsok között súlyozott élek mennek, mely élek megmutatják, hogy a populáció mely tagjai között van kapcsolat, és ezen kapcsolatnak mekkora az erőssége. Minél nagyobb ez a szám, annál erősebb a kapcsolat a 2 személy között.

A hálózat első rétegébe a háztartások tartoznak. Ezen réteg pontosan ugyanúgy lett megalkotva, mint ahogy a [2]-es cikkben a háztartások rétege elkészült. A szimulációk 1600 háztartásból álló hálózatra lettek lefuttatva, amely nagyjából 4000 fős populációt jelentett. Ez a létszám szimulációnként kicsit változott, mivel a hálózatban eltérő létszámú háztartások voltak, az egyes háztartások létszamai pedig valós magyarországi adatok alapján lettek kisorsolva. Mivel egy háztartás tagjai között elég erős kapcsolat van, ezért egy háztartáson belül bármely két személy között 1 súlyú él megy.

A második réteget pedig úgy alkottuk meg, hogy a populáció tagjait véletlenszerűen kisebb csoportokba soroltuk. Ezen csoportokat tekinthetjük például munkahelyeknek, iskoláknak. A sorsolást úgy végeztük el, hogy minden csoportba tartozzon, de az sem kizárható, hogy egy háztartásból több személy is ugyanazon csoportba került kisorsolásra. Ezen csoportokra is teljesül, mint ahogyan a háztartásokra is, hogy a csoport bármely 2 tagja között vezet él. Azonban ezen élek súlya szimulációnként eltérő.

Az SIR modell

A járványterjedés szimulációjára az SIR modellt használtuk. Ez azt jelenti, hogy a populáció egyedeit 3 csoportba soroljuk:

- **S** (Susceptible): fogékonyak
- **I** (Infected): fertőzők
- **R** (Recovered): gyógyultak

Minden olyan személyt, aki nem betegedett még meg, a fogékony kategóriába sorolunk. Miután megbetegszik, akkor átkerül a fertőző kategóriába, és miután meggyógyul, akkor pedig átkerül a gyógyultak kategóriába. Ezen modell szerint csak az előbb leírt sorrendben lehet egyik kategóriából a másikba átkerülni, tehát például, aki már meggyógyult nem tud újra megfertőződni. A személyek egyes állapotokban töltött idejét a fertőzési és gyógyulási ráták segítségével adhatjuk meg. Ha van egy él, amely egy fogékony és fertőző egyén között megy, akkor ezen élhez tartozik egy τ paraméterű, exponenciális eloszlású valószínűségi változó, amely megadja, hogy mikor adódik át a fertőzés ezen a csúcson. Ennek várható értéke $\frac{1}{\tau}$. Továbbá minden fertőző csúcshoz tartozik egy γ paraméterű, exponenciális eloszlású valószínűségi változó, amely a gyógyulás idejét adja meg, és melynek várható értéke $\frac{1}{\gamma}$.

A használt jelölések:

- τ : a csoporttagok közötti fertőzési ráta (élek súlya a gráfban)
- γ : gyógyulási ráta
- R_{inf} : a fertőzésen átesettek száma
- I_{max} : a fertőzöttek maximális száma, vagyis egyazon időpontban a betegek maximális száma

A Gillespie-algoritmus

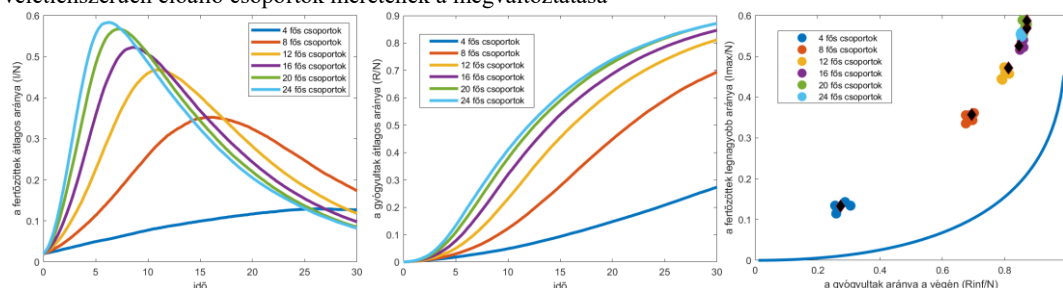
A járványterjedés szimulációjára a Gillespie-algoritmust futtattuk, amely az SIR modellt valósítja meg véletlen gráfon a következőképpen:

Minden egyes lépésben meghatározza, hogy a következő esemény mikor fog bekövetkezni. Ezen időt egy olyan exponenciális eloszlású valószínűségi változó adja meg, amelynek paramétere a fertőzési ráták, illetve gyógyulási ráták összege. Ezután pedig kisorsolja, hogy ezen esemény fertőzés vagy gyógyulás lesz, a fertőzési és gyógyulási ráták összege alapján. Gyógyulás esetén egyenletes eloszlás alapján választ egy csúcst a fertőzöttek közül, amelyik meggyógyul, fertőzés esetén, pedig a fogékony csúcsok közül választ egyet, annál nagyobb valószínűséggel, minél nagyobb az adott csúcshoz tartozó fertőző csúcsokból menő élsúlyok összege.

A szimulációk

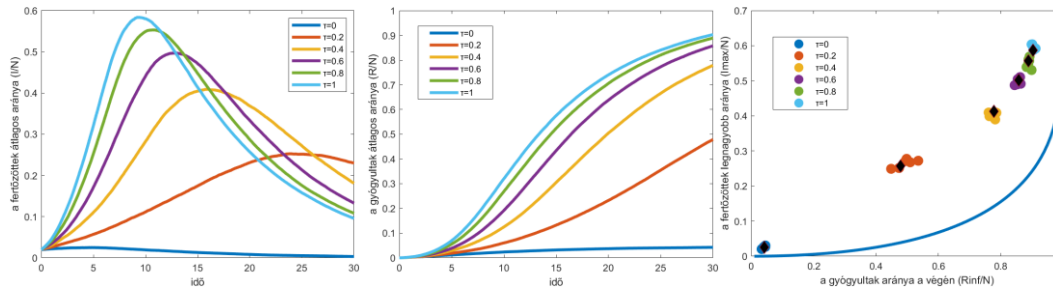
A járványterjedés vizsgálata céljából szimulációkat futtattunk le a MATLAB szoftver segítségével az elkészített véletlen gráfon. A szimulációk kiindulásakor az egyének 2%-át tekintettük fertőzöttnek, a többi ember a fogékony kategóriába tartozott az SIR modellben. Minden esetben 4 gráfot sorsoltunk ki, és minden egyes gráfon 6-6 szimulációt futtattunk le, vagyis az ábrákon ezen szimulációk eredményeinek átlaga látható. Minden szempont vizsgált szempont esetén 3-3 ábrát készítettünk. Ezek közül az első ábra mutatja a fertőzöttek arányát az idő függvényében, a második ábra a gyógyultak arányát az idő függvényében. A harmadik ábra pedig azt mutatja, hogy miképp alakul a fertőzöttek maximális száma, illetve a fertőzésen átesettek száma a szimulációk során. A kis körök mutatják a szimulációk során kapott eredményeket, ezek átlagát (különböző paraméterek esetén nézve) pedig a fekete rombuszok mutatják. Az ábrán látható görbe pedig azt mutatja, hogy milyen összefüggést ad a differenciálegyenletek segítségével felírt modell ugyanezen feladatra, amely azt veszi figyelembe, hogy az egyes kategóriákba hány ember tartozik az idő függvényében, de nem veszi figyelembe az egyének kapcsolatainak szerkezetét, mint a véletlen gráf modell. Azt vehetjük észre, hogy R_{inf} és I_{max} értékek mind magasabbak a véletlen gráf modell esetén.

- a véletlenszerűen előálló csoportok méretének a megváltoztatása



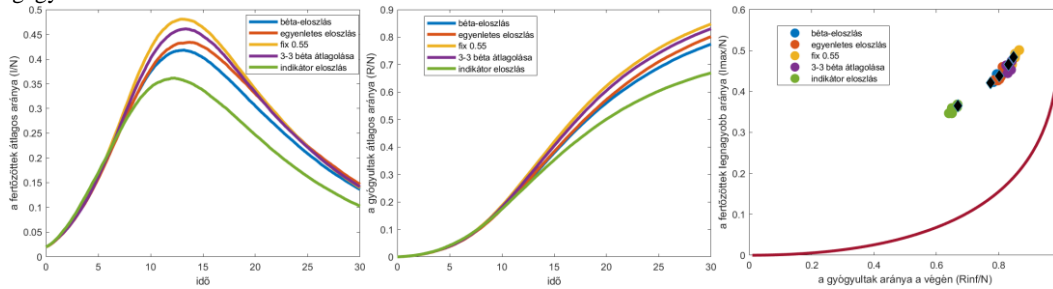
1-3. ábra: τ értékek: 0,55 és 0,45 paraméterű béta-eloszlásúak

- az egyes csoporttagok közötti fertőzési ráta megváltoztatása



4-6. ábra: csoportméret fix 10 fő, a csoportokon belüli τ értékek változnak

- a csoporttagok közötti fertőzési ráta értékét különböző eloszlásokból vett minta alapján határozzuk meg, melyeknek a várható értéke megegyezik



7-9. ábra: csoportméret fix 10 fő, várható érték fix 0.55, eloszlások különbözőek

A szimulációk eredménye

- **a csoportméret hatása:** A csoportméretek megváltoztatásának hatását több szimuláció során is vizsgáltuk, melyeknél a csoporttagok közötti τ értékeket különböző eloszlások alapján választottuk ki, azonban várható értékük egységesen 0.55 volt. Az első 3 ábrán azon szimuláció eredményei láthatóak, amelyeknél ezen értékeket béta-eloszlás alapján választottuk. A többi eloszlás esetén nagyon hasonló járványgörbéket kaptunk, azonos csoportméretek esetén R_{inf} és I_{max} esetén láthattunk kisebb eltéréseket a különböző eloszlások között, amelyek az eltérő szórásokból adódhatnak, de ezen eltéréseket, majd egy későbbi szimuláción vizsgáljuk még. Viszont ennél nagyobb eltéréseket tapasztalhattunk a különböző csoportméretek esetén. Az ábrákból egyértelműen kivethető, hogy R_{inf} és I_{max} is nagyobb értéket vesz fel, ha nagyobb a csoportméret. Emellett az is látható, hogy a járvány csúcsát később érjük el kisebb csoportok esetén, továbbá I_{max} értéke is kisebb lesz. Viszont érdemes alaposabban megnézni a középső ábrát is, amelyen a fertőzöttek átlagos számát láthatjuk az idő elteltével. Azt vehetjük észre, hogy a görbék metszik egymást, a csoportméretek változtatva másképp fut le a járvány. Kisebb csoportok esetén később érjük el a járvány csúcsát, a járvány lassabb tempóban fut le.
- **a fertőzési ráta hatása:** Ezen szimuláció során azt vizsgáltuk, hogy milyen hatással van a járvány lefutására a τ értékek megváltoztatása, rögzített csoportméret mellett. A szimuláció eredményei a következő 3 ábrán láthatóak. Észrevehető, hogy a járványgörbék hasonló alakot vesznek el, mint amikor τ volt fix, és a csoportméret változott. Ezen esetben azt tapasztalhatjuk, hogy nagyobb τ értékek esetén a járvány hamarabb éri el a csúcsát, és összességében is gyorsabban fut le, illetve I_{max} és R_{inf} értékek is nagyobbak.
- **a szórás hatása:** A következő szimuláció során arra voltunk kíváncsiak, hogy milyen hatással van a járvány lefutására, ha a csoporttagok közötti τ értékeket azonos várható értékű, de eltérő szórású eloszlások alapján választjuk ki. Ezen szimulációk eredményei az utolsó 3 ábrán láthatóak. Az ábrákon észrevehető, hogy a járványgörbék alakja nagyon hasonlít egymáshoz, viszont különböző eloszlások esetén különbséget láthatunk a járvány csúcsában, illetve a gyógyultak számában a járvány végén. Azt láthatjuk, hogy mind az R_{inf} , mind az I_{max} akkor volt a legnagyobb, amikor a τ értékeket fixen 0.55-re állítottuk be. A legkisebb értékeket pedig, amikor csoportonként különböző indikátor eloszlás szerint választottuk ki.

eloszlás	fix 0.55	3 béta eloszlás átlagolása	egyenletes eloszlás	béta eloszlás	indikátor eloszlás
szórásnégyzet	0	0.0412	0.0675	0.1238	0,2475

A táblázatban látható eloszlások alapján választottuk ki a csoporttagok közötti τ értékeket. Azt tapasztaltuk, hogy nagyobb szórásnégyzet esetén kisebbek voltak az R_{inf} és I_{max} értékek.

- **a ferdeség hatása:** Az egyik szimuláció keretében azt vizsgáltuk, hogy a járványterjedésre milyen hatással van az eloszlás ferdesége. Ezen szimuláció során minden egyes csoportra a τ értéket 0.55 várható értékű béta-eloszlás alapján választottuk ki, azonban az α és β értékeket különbözőképpen választottuk ki, vagyis különböző ferdeségű, de azonos várható értékű eloszlásokból indultunk ki. A szimulációk eredményében lényeges különbséget nem tapasztaltunk az eltérő ferdeségű eloszlások esetén. Ezen ábrák nem kerültek bele a beszámolóba.

Következtetések: A szimulációk eredményei alapján arra következtethetünk, hogy korlátozó intézkedések bevezetésével lehet lassítani a járványt. Azon szimulációk esetén, amikor a csoporttagok közötti τ értékeket változtattuk, akkor azt tapasztaltuk, hogy kisebb τ értékek esetén a járvány lefutása lassabb volt, a csúcsán pedig kevesebben voltak betegek. Ez azt jelenti, hogy a maszkviselés egy hatásos védekezés lehet a járvánnyal szemben. Az eltérő τ értékek jelenthetik azt, hogy a csoport tagjai mennyire profi maszkot viseltek. Továbbá a csoportméretek csökkentésénél is hasonló tapasztalatokat fogalmazhatunk meg, vagyis hatásos védekezés lehet, ha a nagyobb csoportokat több kisebb csoportra osztják fel, például az iskolákban jobban odafigyelnek rá, hogy elkülönüljenek egymástól az osztályok.

A jövőbeni tervek: Ebben a félévben járványterjedési szimulációkat futtatunk, hogy a modell paraméterérzékenységét vizsgáljuk. Következő félévre pedig azt tervezzük, hogy paraméterbecsléssel fogunk foglalkozni, vagyis a járványgörbe alapján szeretnénk majd megbecsülni a feladat különböző paramétereit, például a fertőzési rátát, a gyógyulási rátát, vagy a csoportméretet.

- [1] Backhaus Ágnes, Bognár Edit: Járványterjedés paramétereinek becslése többretegű véletlengráf-modellekben, Kézirat
- [2] Ágnes Backhaus, István Z. Kiss, Péter L. Simon: The impact of spatial and social structure on an SIR epidemic on a weighted multilayer network
- [3] Wasirur R. KhudaBukhsh, Boseung Choi, Eben Kenah, and Grzegorz A. Rempala: Survival dynamical systems: individual-level survival analysis from population-level epidemic models
- [4] Kiss, I.Z., Miller, J.C., Simon, P.L., Mathematics of Epidemics on Networks; From Exact to Approximate Models, Springer, Interdisciplinary Applied Mathematics 46, (2017).