

SIS TÍPUSÚ JÁRVÁNYTERJEDÉS VIZSGÁLATA HIPERGRÁFOKON

BODÓ ÁGNES

Hipergráfokon történő járványterjedés matematikai vizsgálata kerül bemutatásra. A cél, hogy a modellezés során egyrészt figyelembe lehessen venni, hogy általában a populáció apróbb közösségekből áll, másrészt, hogy a terjedést meghatározó fertőzési ráta nem feltétlenül lineárisan függ a beteg szomszédok számától. Levezetésre kerül a járványterjedésre vonatkozó alapegyenlet tetszőleges hipergráf esetén. Ennek segítségével bevezethető az átlagtér közelítés, mint az alapegyenlet egy lehetséges egyszerűsítése, amely a sztochasztikus szimuláció eredményeivel összevethető. A szimuláció során mind a hipergráf szerkezete, mind a matematikai leírásban alkalmazott paraméterek hatásai vizsgálhatóak.

1. Bevezetés

Történelmünk során számos súlyos járvány sújtotta a világ népességét, többek között a fekete halálnak emlegetett pestis, amely már az 5. században jelen volt, és azóta folyamatosan szedi áldozatait a világ különböző területein, vagy például az 1918-ban felbukkanó spanyolnátha, amely a Föld teljes lakosságának 2-4%-át betegítette meg, és amely csak az 1918-as évben több áldozatot követelt, mint az egész első világháború. Az orvostudomány előrehaladásával folyamatosan fejlődik a járványokkal vívott küzdelem, amelyben nagy szerepet kapnak a járványterjedési modellek.

A járványterjedés matematikai modellezésének kezdetei a 20. század elejére nyúlnak vissza, az első ilyen jellegű mű Kermack és McKendrick 1927-ben megjelent munkája. Az ebben a dolgozatban megjelenő modell még igen kezdetleges volt, azonban az évek során a tudományág folyamatosan fejlődött és az alkalmazott matematika egyik fontos kutatási területévé vált. Ezen modellek segítségével olyan kérdések válaszolhatóak meg, mint például el fog-e terjedni a járvány egy adott populáción belül, és ha igen, akkor várhatóan a populáció hányadát fogja érinteni a fertőzés, de meghatározhatóak olyan mennyiségek is, mint például a reprodukciós szám vagy a kritikus oltási küszöb [3].

2. A modell

A járványterjedés leírására számos dinamika létezik, az egyik legegyszerűbb az úgynevezett SIS típusú járványterjedés, amely olyan fertőzések leírására alkalmas, ahol a betegségen átesett egyedek nem nyernek immunitást, hanem újra fertőzhetővé válnak. Ilyenek általában a nemi úton terjedő betegségek, mint például a kankó. Ezen betegségeknel az egyedek kétféle állapotban lehetnek: egészséges és fertőző, amelyeket a továbbiakban jelöljön S és I (az angol *susceptible* és *infected* szavakból). Egy csúcs állapota kétféleképpen változhat: egy egészséges egyed a fertőző szomszédai hatására fertőzötté válhat, ez a fertőzés folyamata ($S \rightarrow I$), illetve egy fertőző egyed meggyógyulhat, ez a gyógyulás folyamata ($I \rightarrow S$). Mind a fertőzés, mind a gyógyulás folyamatát Poisson-folyamattal írjuk le, amelynek értelmében annak a valószínűsége, hogy kis Δt idő alatt egy egyed S állapotból I állapotba kerül $1 - \exp(-\tau k \Delta t)$, illetve annak, hogy I állapotból S állapotba $1 - \exp(-\gamma \Delta t)$, ahol k jelöli az adott egyed fertőzött szomszédai számát, és τ , γ pozitív számok. A továbbiakban a τk és γ számokat a fertőzési és gyógyulási folyamatok rátáinak nevezzük.

A betegségterjedés matematikai leírásához célszerű megállapítani a kapcsolati struktúrát, amelyet hagyományosan gráfokkal szoktak jellemezni. Legyen tehát adott egy N csúcsú irányítatlan, egyszerű gráf, ahol a gráf csúcsai a populáció egyedeinek felelnek meg. Továbbá két csúcs között akkor van él, ha a fertőzés terjedhet köztük, ami a különböző betegségek esetében más és más lehet, attól függően, hogy például cseppfertőzéssel vagy nemi úton terjedő betegségről van szó.

Megfigyelhető, hogy a betegségterjedés egyes csoportokon belül jóval erősebb, míg a különféle csoportok között kevésbé meghatározó, ilyen csoportra természetes példa egy család vagy egy munkahelyi közösség. Ezen folyamatokat tanulmányozzák az úgynevezett háztartás típusú hálózatok segítségével [2]. Ezen kutatások alapján a járványterjedés során a populációt nem gráffal, hanem úgynevezett hipergráffal, vagy más néven általánosított gráffal reprezentáltuk. A hipergráf egy $(\mathcal{V}, \mathcal{E})$ pár, ahol $\mathcal{V} = \{v_1, v_2, \dots, v_N\}$ a hipergráf csúcsait jelöli, $\mathcal{E} = \{e_1, e_2, \dots, e_M\}$ pedig a hiperéleket, ahol $e_i \subset \mathcal{V}$ teljesül minden $i = 1, 2, \dots, M$ esetén. Hagyományos értelemben vett gráfoktól eltérően hipergráfoknál egy él két-től több csúcst is összeköthet.

A továbbiakban célunk az SIS típusú járványterjedés hipergráfokon való vizsgálata. A hipergráf csúcsai tehát kétféle állapotban lehetnek: S , illetve I . A gyógyulás rátája a gráfos esethez hasonlóan γ maradt, azonban a fertőzés rátáját az alábbiak szerint módosítottuk:

$$\lambda_{SI} = \tau \sum_h f(k_h), \quad (1)$$

ahol az összegzés azon $h \in \mathcal{E}$ hiperélekre történik, amelyek tartalmazzák az adott csúcst, k_h jelöli a fertőzött csúcsok számát a h hiperélben, és f adott függvény. A

hagyományos gráfos megközelítésben, ha f az identitásfüggvény, akkor visszkapjuk a korábban említett $k\tau$ rátát, ahol $k = \sum_h k_h$ a fertőzött szomszédok száma. A modellezés során fontos kérdés volt, hogy az (1) rátában milyen f függvénnyel dolgozzunk. A hipergráfokon való vizsgálat mellett egy másik jelentős változtatás volt ezen f függvény bevezetése, amivel azt a jelenséget szerettük volna figyelembe venni, hogy sok esetben a fertőzés rátája nem lineárisan függ a fertőzött szomszédok számától. Példának okáért biológiai neurális hálózatok modellezésénél egy inaktív neuron aktívvá válásának rátája nem lineárisan függ az aktív szomszédok számától, hanem a tangens hiperbolikus függvényhez hasonló telítődés szerint. Tehát nagyon sok szomszéd nem növeli az átmenet valószínűségét, ami betegségterjedés során is életszerű. Az [1] cikkben a szerzők arra jutottak, hogy a folyamat megértéséhez elegendő a tangens hiperbolikus függvénynél numerikusan könnyebben kezelhető függvény alkalmazása, nevezetesen

$$f(x) = \begin{cases} x, & \text{ha } 0 \leq x \leq c, \\ c, & \text{ha } x > c, \end{cases} \quad (2)$$

amelyet csupán a c küszöbértéknek nevezett paraméter jellemez.

3. Eredmények

3.1. Alapegyenlet és várható értékre vonatkozó egyenletek

Célunk volt gráfokhoz hasonlóan az állapotvalószínűségekre vonatkozó lineáris differenciálegyenlet-rendszer levezetése [5], amelyet alapegyenletnek neveznek. Egy N csúcsú hipergráf során az állapottér az $\{S, I\}^N$ halmaz, ugyanis bármely csúcson ezen kétféle állapotban lehet. Célszerű ezt a halmazt $N + 1$ részhalmazra osztani a fertőző csúcsok száma szerint. A továbbiakban legyen $S^0 = (SSS \dots S)$ az az állapot, ahol a hipergráfban nincs fertőző csúcs. Hasonlóan jelölje S^k azon állapotokat, amelyeknél k fertőző csúcs található a hipergráfban. Végezetül legyen $S^N = (III \dots I)$ az az állapot, ahol minden csúcs fertőző. Az S^k halmaz elemeit jelölje $S_1^k, S_2^k, \dots, S_{c_k}^k$, ahol $c_k = \binom{N}{k}$. Ezen jelölések segítségével felírható a fertőzés és a gyógyulás folyamata, lásd [1].

Az alapegyenlet felírásához jelölje $\mathcal{X}_j^k(t)$ annak a valószínűségét, hogy a rendszer a t időpillanatban S_j^k állapotban van. Továbbá legyen

$$\mathcal{X}^k(t) = (\mathcal{X}_1(t), \mathcal{X}_2(t), \dots, \mathcal{X}_{c_k}(t)), \quad k = 0, 1, \dots, N$$

a k beteget tartalmazó állapotok valószínűségeiből álló vektor. A fenti átmenetek az $\mathcal{X}_j^k(t)$ változókra az alábbi lineáris, állandó együtthatós differenciálegyenlet-rendszert határozzák meg:

$$\dot{\mathcal{X}}^k = A^k \mathcal{X}^{k-1} + B^k \mathcal{X}^k + C^k \mathcal{X}^{k+1}, \quad k = 0, 1, \dots, N,$$

ahol A^0 és C^N nulla mátrixok. Az A^k mátrixok a fertőzés, a C^k mátrixok a gyógyulás folyamatát írják le, amelyek részletes leírása megtalálható az [1] cikkben. Az egyenlet az alábbi alakban is írható:

$$\dot{\mathcal{X}} = P\mathcal{X}, \tag{3}$$

ahol P blokk-tridiagonális mátrix, amelynek főátlójában a B^k mátrixok szerepelnek, a főátló alatt az A^k , a főátló felett pedig a C^k mátrixok.

A (3) alapegyenlet nagy N esetén kezelhetetlenné válik, ezért célszerű az alapegyenletet leegyszerűsíteni, amelynek egyik módja az úgynevezett átlagtér közelítés. Ilyenkor nem az egyes állapotvalószínűségekre írunk fel differenciálegyenleteket, hanem bizonyos várható értékekre, jelen esetben a fertőző és egészséges típusú csúcsok számának várható értékeire, amelyeket $[I]$ és $[S]$ jelöl, ahol

$$[I](t) = \sum_{k=0}^N k \sum_{j=1}^{c_k} \mathcal{X}_j^k(t), \quad [S](t) = \sum_{k=0}^N (N - k) \sum_{j=1}^{c_k} \mathcal{X}_j^k(t).$$

Az egyenletek felírásához vezessük be a következő mennyiségeket: jelölje $N_h(\mathcal{S}_j^k)$ a h hiperélben lévő fertőzött csúcsok számát az \mathcal{S}_j^k állapotban, továbbá legyen

$$N_{SI}^f(\mathcal{S}_j^k) = \sum_{l: \mathcal{S}_j^k(l)=S} \sum_{h: l \in h} f(N_h(\mathcal{S}_j^k)),$$

azaz az \mathcal{S}_j^k állapotban lévő egészséges csúcsokra vonatkozó $\sum_{h: l \in h} f(N_h(\mathcal{S}_j^k))$ értékek összege. Ekkor az $[I](t)$ és $[S](t)$ függvényekre az alábbi differenciálegyenletek írhatóak fel.

3.1. TÉTEL. Az $[I](t)$ és $[S](t)$ várható értékek az alábbi differenciálegyenleteket elégítik ki tetszőleges hipergráf esetén:

$$\dot{[S]} = \gamma[I] - \tau[SI], \tag{4}$$

$$\dot{[I]} = \tau[SI] - \gamma[I], \tag{5}$$

ahol

$$[SI] = \sum_{k=0}^N \sum_{j=1}^{c_k} N_{SI}^f(\mathcal{S}_j^k) \mathcal{X}_j^k(t).$$

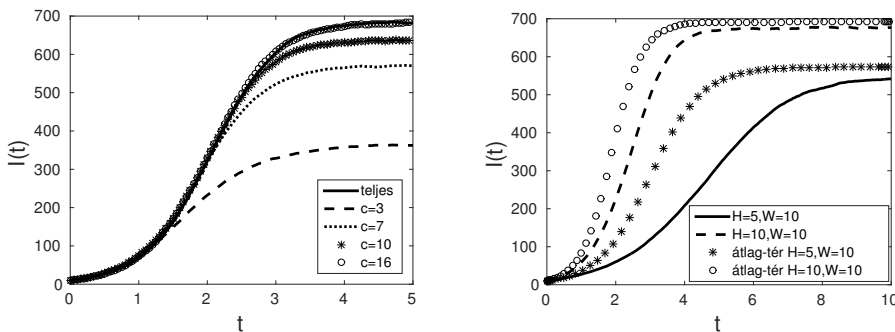
A bizonyítás megtalálható az [1] cikkben.

3.2. Szimulációk

A (4), (5) típusú egyenletek korlátozottak abban az értelemben, hogy nem tudják figyelembe venni a fertőzés útját, azaz hogy a betegség pontosan melyik

egyedről melyik egyedre jut át. Ennek leírásához célszerű megállapítani a kapcsolati struktúrát. A hálózat szerkezetének ismeretében Monte-Carlo-szimulációval lehet vizsgálni a betegség terjedésének folyamatát, például lásd [4].

A gyakorlatban a pontos kapcsolati struktúra feltérképezése általában nehezen kivitelezhető, ezért gyakran alkalmaznak véletlen gráfokat a folyamatok leírására. Ebből fakadóan a szimulációkat különféle véletlen hipergráfokon végeztük. Az első hipergráftípus a valós életnek egy nagyon leegyszerűsített modellje, ahol feltételezzük, hogy minden egyednek van otthona és munkahelye, tehát minden csúcs pontosan két hiperélben szerepel. A második hipergráftípus során a konfigurációs modell segítségével generáltunk véletlen hipergráfokat. A pontos részletek megtalálhatóak az [1] cikkben.



(a) A teljes gráfos és a hipergráfos megközelítés az (1) rátában szereplő függvény c küszöbértékének különböző megválasztásai esetén 16 nagyságú, 500 hiperéllal rendelkező véletlen hipergráfon $\tau = 0,02$, $\gamma = 1$ paraméterek esetén.

(b) A hipergráfos megközelítés $c = 7$ küszöbérték választással, és az átlagtér közelítésnek megfelelő (6) egyenlet megoldása különböző otthon (H) és munkahely mérettel (W) rendelkező hipergráfokon $\tau = 0,18$, $\gamma = 1$ paraméterek esetén.

1. ábra. A fertőzött csúcsok számának időbeli változása $N = 1000$ esetén.

A szimulációk során a folyamatot meghatározó mennyiséget, azaz a betegszám időbeli változását követhetjük nyomon. Egyrészt megvizsgáltuk a (2)-ben definiált f függvény hatását, azaz azt, hogy a küszöbérték változása milyen hatással van a betegszámra. Kétféle megközelítést használtunk: teljes gráfok és hipergráfok. A teljes gráfos megközelítésben minden hiperélt teljes gráffal helyettesítettünk. Azt az eredményt kaptuk, hogy hipergráftípustól függetlenül a küszöbérték elég nagy értékére a teljes gráfos és a hipergráfos szimuláció közelítőleg megegyezik, lásd az 1a. ábrát. Másrészt megvizsgáltuk a struktúra hatását, azaz azt, hogy a hipergráfok homogenitását változtatva hogyan alakul a betegszám időfüggése. Homogenitás alatt azt értjük, hogy a hipergráf uniform és reguláris. Hipergráftípustól függetlenül arra az eredményre jutottunk, hogy minél homogénebb a folyamatot

leíró hipergráf, annál lassabb kezdetben a járvány terjedése (azaz annál kevesebb kezdetben a betegszám), azonban annál hatékonyabban terjed el idővel a járvány (azaz a betegszám kellő idő elteltével nagyobb lesz). A homogenitás hatása az [1] cikkben van illusztrálva. Végül pedig összehasonlítottuk a szimulációt a várható értékre vonatkozó egyenlettel, lásd az 1b. ábrát, ahol az első hipergráftípusnál az (5) egyenletre az alábbi közelítést írhatjuk fel:

$$\dot{I} = \tau(N - I) \left[f\left(\frac{H - 1}{N}I\right) + f\left(\frac{W - 1}{N}I\right) \right] - \gamma I, \quad (6)$$

ahol H az otthonok, W a munkahelyek méretét jelöli.

Köszönetnyilvánítás

A szerző köszönetét fejezi ki az Országos Tudományos Kutatási Alapprogramok, OTKA (pályázat száma: 115926) támogatásáért.

Hivatkozások

- [1] BODÓ, A., KATONA, Y. G., AND SIMON, L. P.: *SIS Epidemic Propagation on Hypergraphs*, Math. Biol., Vol. **78**, pp. 713-735 (2016).
- [2] HOUSE, T. AND KEELING, M. J.: *Deterministic epidemic models with household structure*, Math. Biosci., Vol. **213**, pp. 29-39 (2008). DOI: [10.1016/j.mbs.2008.01.011](https://doi.org/10.1016/j.mbs.2008.01.011)
- [3] PASTOR-SATORRAS, R., CASTELLANO, C., VAN MIEGHEM, P., AND VESPIGNANI, A.: *Epidemic processes in complex networks*, Rev. Mod. Phys., Vol. **87**, pp. 925-979 (2015). DOI: [10.1103/RevModPhys.87.925](https://doi.org/10.1103/RevModPhys.87.925)
- [4] SZABÓ-SOLTICZKY, A. AND SIMON, L. P.: *The Effect of Graph Structure on Epidemic Spread in a Class of Modified Cycle Graphs*, Math. Model. Nat. Phenom., Vol. **9** No. **2** pp. 89-107 (2014). DOI: [10.1051/mmnp/20149206](https://doi.org/10.1051/mmnp/20149206)
- [5] TAYLOR, M., SIMON, L. P., GREEN, M. D., HOUSE, T., AND KISS, Z. I.: *From Markovian to pairwise epidemic models and the performance of moment closure approximations*, Math. Biol., Vol. **64** No. **6** pp. 1021-1042 (2012). DOI: [10.1007/s00285-011-0443-3](https://doi.org/10.1007/s00285-011-0443-3)



Berzlánovichné Bodó Ágnes 1991-ben született Szombathelyen. 2009-ben nyert felvételt az Eötvös Loránd Tudományegyetem matematika szakára. Az alapszakos diplomáját 2012-ben szerezte meg, ezt követően 2015-ig az Alkalmazott Matematikus MSc mesterképzésének hallgatója volt. 2014-ben elnyerte a Természettudományi Kar „Kar Kiváló Hallgatója” díját. A diploma megszerzése után 2018-ig az Eötvös Loránd Tudományegyetem Matematika Doktori Iskola, Alkalmazott Matematika Program ösztöndíjas doktoranduszaként tanult, jelenleg doktorjelölt. 2018 szeptemberétől 2019 júniusáig tudományos segédmunkatársként dolgozott az Eötvös Loránd Tudományegyetem Alkalmazott Analízis és Számításmatematikai Tanszékén. Kutatási területe a közönséges differenciálegyenletek és alkalmazásai témakörén belül különféle járványterjedési modellek analitikus és numerikus vizsgálata. Doktori képzésének megkezdése óta három angol nyelvű folyóiratcikke jelent meg és egy magyar nyelvű népszerűsítő cikke.

BODÓ ÁGNES

Eötvös Loránd Tudományegyetem
Alkalmazott Analízis és Számításmatematikai Tanszék
Numerikus Analízis és Nagy Hálózatok Kutatócsoport
1117 Budapest, Pázmány Péter sétány 1/C
bodoagi@cs.elte.hu

SIS EPIDEMIC PROPAGATION ON HYPERGRAPHS

ÁGNES BODÓ

Mathematical modeling of epidemic propagation on hypergraphs is considered in this paper. The goal is to model the community structure with greater accuracy and to describe the dependence of the infection pressure on the number of infected neighbours with a nonlinear function. The master equation describing the process is derived for an arbitrary hypergraph. The mean-field equations are introduced as an approximation to the master equation and are compared against the stochastic simulations. Simulation results can be used to analyze the effects of the hypergraph structure and the model parameters.

Keywords: SIS epidemic; mean-field model; exact master equation, hypergraph

Mathematics Subject Classification (2000): 05C65, 60J28, 90B15, 92D30