

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR

BSC SZAKDOLGOZAT

Járványterjedés vizsgálata hipergráfokon

Huszárik Nikolett

Matematika BSc
Alkalmazott matematikus szakirány

Témavezető: **Bodó Ágnes**, Doktorandusz
Alkalmazott Analízis és Számításmatematikai Tanszék



Budapest, 2017.

Köszönetnyilvánítás

Ezúton szeretnék köszönetet mondani témavezetőmnek, Bodó Ágnesnek, aki felkeltette érdeklődésemet a téma iránt. Köszönöm, hogy igényeim szerint mindig időt szakított rám. Tanácsai és útmutatása nélkül dolgozatom jelenlegi formája nem jöhetett volna létre.

Hálás köszönettel tartozom családomnak a tanulmányaim alatt nyújtott kitartó támogatásukért. Valamint köszönettel tartozom tanárainknak, hogy megszerettették velem a matematika világát.

Budapest, 2017. május 31.

Huszárik Nikolett

Tartalomjegyzék

Bevezetés	3
1. Általános összefoglaló	4
1.1. Történeti áttekintés és példák	4
1.2. A járványterjedés dinamikája	5
1.3. A járványterjedés struktúrája	6
2. Járványterjedés hipergráfokon	8
2.1. A modell	8
2.2. Alapegyenlet és a várható értékre vonatkozó egyenletek	10
2.2.1. Alapegyenlet	10
2.2.2. Várható értékre vonatkozó differenciálegyenletek	14
3. Szimulációk	17
3.1. A szimulációk során alkalmazott gráfmodellek	17
3.1.1. Két állapotú hipergráf	17
3.1.2. Barabási–Albert klikk modell	18
3.1.3. Konfigurációs modell	18
3.2. A szimulációs algoritmus	20
3.3. Szimulációs eredmények	20
3.3.1. Az f függvény hatása	20
3.3.2. A struktúra hatása	22
3.4. Várható értékre vonatkozó egyenletek lezárása	23
4. Valós adatok	27
4.1. A küszöbérték meghatározása	28
4.2. A struktúra paramétereinek hatása	29
Összefoglalás	31
Irodalomjegyzék	32

Bevezetés

Jelen dolgozat középpontjában hálózatokon zajló folyamatok állnak, amelyek matematikai leírására például Barrat, Barthélemy és Vespagnani [1] tanulmánya egy kiváló összefoglaló. Célunk a terjedési folyamatok, azon belül is a járványterjedés vizsgálata. Ezt a folyamatot hagyományos értelemben vett gráfok helyett hipergráfokon fogjuk tekinteni, amelynek alapjául a [2] cikk szolgál.

A dolgozat felépítése a következő. Az első fejezetben rövid betekintést nyújtunk a témakörbe és megemlítjük a terjedési folyamatokkal kapcsolatos alapvető modelleket. A második fejezetben egy konkrét járványterjedési modell hipergráfokon való leírásával foglalkozunk, levezetjük a modellt leíró alapegyenletet és a várható értékre vonatkozó egyenleteket. A harmadik fejezetben az általunk létrehozott különböző struktúrák és paraméterek mellett vizsgáljuk a folyamatok viselkedését, szimulációs eredményekkel és valós adatokkal zárjuk a dolgozatot.

1. fejezet

Általános összefoglaló

1.1. Történeti áttekintés és példák

Az elmúlt néhány évtizedben a számítógépek teljesítményének növekedése és a hatalmas adatállományok elérhetősége új szintre emelték a hálózati rendszerek tanulmányozását. Az internet és azon belül a világháló feltérképezése lehetőséget adott nagy és komplex hálózatok elemzésére. Napjainkban ez egy igen fontos kutatási terület, ugyanis a hálózatok a mindennapi életünkben, a tudományban és a gazdaságban is egyaránt jelen vannak. Többet között az emberi agy idegsejtek komplex hálózata, amelynek vizsgálata az agyi funkciók megértésének kulcsa, vagy például egy adott ökoszisztémán belül élő fajok tápláléklánc hálózata, amely leírja az egyes fajok közti táplálkozási kapcsolatokat [1]. Gondolhatunk továbbá az áramellátást biztosító villamosenergia-hálózatra, amely generátorok és villamos távvezetékek rendszere, vagy a kereskedelmi- és közlekedési hálózatra, amelyek hosszú idők óta elengedhetetlen részei a társadalomnak [15]. Ugyanakkor maga a társadalom is leírható hálózatként az ismeretségek által. Ezen komplex rendszerek megértéséhez elkerülhetetlen a mögöttük húzódó hálózatok szerkezetének megismerése és az ezeken végbemenő folyamatok vizsgálata. Hálózati folyamatnak tekinthető például egy számítógépes vírus térhódítása a világhálón vagy az aktivitás terjedése neurális hálózatokon [8].

A szakdolgozatban a járványterjedés matematikai leírásával foglalkozom. Betegségterjedés esetén a hálózati folyamat egy populáción belüli fertőzés terjedésének időbeli lefolyását jelenti. A történelem során számos járvány sújtotta az emberiséget. Az egyik legpusztítóbb, fekete halálként is emlegetett betegség a pestis, amely megharmadolta Európa lakosságát a 14. században [16]. Egy másik, gyakorlatilag halálos ítéletet jelentő fertőző betegség a spanyolnátha, amely 1918 és 1919 között több áldozatot követelt, mint az egész első világháború [17]. Az orvostudomány fejlődésével egyenes arányban fejlődik a járványok ellen vívott harc is. Példaként említhető a feketehimlő, amely a védőoltások felfedezésének köszönhetően végleg megsemmi-

sült. Ugyanakkor vannak olyan pusztító járványok, mint a malária, az AIDS, a kolera vagy a tífusz, amelyek napjainkban is nagymértékben szedik áldozataikat. Az egyes betegségek feltérképezésében nagy szerepet vállalnak az epidemiológiai modellek. A klasszikus betegségterjedési modellt Kermack és McKendrick alkotta meg 1927-ben [3], amelyet a fertőzések matematikai modellezésének kiindulópontjaként tekintünk. Ezen modellek megértése, elemzése többek között az alábbi kérdésekre adhat választ: Egy új járvány megjelenésekor a populációnak várhatóan mennyi része lesz fertőzött? A járvány eléri-e a reprodukciós számot, azaz hosszútávon elterjed a vírus vagy kihal a populációból? Ha elterjed a vírus, akkor egy bizonyos idő elteltével várhatóan hányan lesznek fertőzöttek? Illetve a populáció hányad részét kell beoltatni, hogy megállítsuk a terjedést?

1.2. A járványterjedés dinamikája

A járványterjedés matematikai leírásához legyen adott egy N csúcsú irányítatlan, egyszerű gráf, ahol a gráf csúcsai a társadalom egyedeinek felelnek meg és két csúcs között akkor vezet él, ha a csúcsoknak megfelelő egyedek a fertőzést átadhatják egymásnak. A járvány típusától függően ez több mindent jelenthet, például nem mindegy, hogy szexuális úton vagy cseppfertőzéssel terjedő betegségről van szó. A csúcsok lehetséges állapotait az $\{a_1, a_2, \dots, a_m\}$ állapotter tartalmazza, ahol egy adott csúcs állapotának megváltozása Poisson-folyamattal írható le. Tehát ennek értelmében kis Δt idő alatt egy a_i állapotban lévő csúcs $1 - \exp(-\lambda_{ij}\Delta t)$ valószínűséggel változhat a_j állapotban lévő csúccsá, ahol λ_{ij} egy pozitív szám, amit az átmenet rátájának nevezünk. A modellezés során tehát meg kell adni az állapotteret és az átmenetek rátáit, amiket együttesen dinamikának szoktunk nevezni. Az egyik legegyszerűbb, úgynevezett SIS típusú járványterjedés olyan fertőzések leírására alkalmas, ahol a fertőzésen átesett egyedek nem nyernek immunitást, azaz újra megfertőződhetnek. Ilyenek általában a nemi úton terjedő betegségek, mint például a gonorrhoea [4]. Ebben az esetben az állapotter két elemet tartalmaz: S , mint egészséges, vagy más néven fertőzhető (az angol 'susceptible' szóból származik), illetve I , mint fertőző (az angol 'infected' szóból származik). Továbbá az állapotok között kétféle átmenet lehetséges. Az egyik az $S \rightarrow I$ átmenet, azaz a fertőzés folyamata, ahol egy egészséges egyed csak a fertőző szomszédai által válhat fertőzővé. Ekkor az átmenet rátája $k\tau$, ahol k az adott csúcs fertőző szomszédainak száma, τ pedig egy pozitív szám, amit fertőzési rátának nevezünk. A másik pedig az $I \rightarrow S$ átmenet, azaz a gyógyulás folyamata, ahol a fertőzőek a környezetüktől függetlenül meggyógyulhatnak, így az átmenet rátája megegyezik egy pozitív γ számmal, amelyet gyógyulási rátának nevezünk. Amennyiben a fertőzésen átesett egyedek immunitást nyernek,

SIR típusú járványterjedésről beszélünk. Ilyenkor a fertőzésen átesett egyedek egy új, immunis (nem fertőző és nem fertőzhető) állapotba kerülnek, amelyet R jelöl (az angol 'recovered' szóból), ahol a gyógyulás az $I \rightarrow R$ átmenetet jelenti. Ilyen például az ebola [5]. Ha az immunitás csak ideiglenes, akkor SIRS típusú járványterjedést alkalmazunk, ahol az állapottér ugyanúgy három elemű, de az $R \rightarrow S$ átmenet is lehetséges, azaz egy idő után az immunis egyedek újra fertőzhetővé válnak. Továbbá előfordulhat, hogy a fertőzésnek van lappangási ideje, azaz egy egészséges egyed a betegséget elkapva nem lesz azonnal fertőzött, ilyenkor érdemes bevezetni egy új állapotot, amelyet lappangó osztálynak nevezünk és E -vel jelölünk (az angol 'exposed' szóból). Ennek megfelelően az $S \rightarrow I$ átmenetet felváltják az $S \rightarrow E$ és az $E \rightarrow I$ átmenetek. Ezt a dinamikát az immunitástól függően SEIS-nek vagy SEIR-nek nevezzük. Például SEIR dinamikával modellezhető a malária [6].

Természetesen számos további dinamika létezik, amelyeket nem csak járványterjedés esetén alkalmaznak. Hasonló dinamikákkal írhatók le például a híresztelések terjedése társadalmi hálózatokon [7], vagy az aktivitás terjedése biológiai neurális hálózatokon [8].

1.3. A járványterjedés struktúrája

A járványterjedést leíró dinamika mellett a másik fontos tényező a folyamatot leíró gráf. A 3.3.2. szakaszban látni fogjuk, hogy a különböző gráftípusok hogyan befolyásolhatják a járványterjedést. Ezen gráfok lehetnek időben állandók, de változhatnak is az úgynevezett adaptív gráfok esetében. Továbbá társadalmi hálózatok esetén szinte lehetetlen a pontos kapcsolati struktúra ismerete, részben az élek és csúcsok száma, másrészt a tényleges kapcsolatokról való ismeret szegényessége miatt. Ebből adódóan a hálózat leírására úgynevezett véletlen gráfokat használunk, ugyanis úgy gondoljuk, hogy ezek modellezik leginkább a valóságot. Számos ilyen gráfmodell létezik, ezek között is vannak, amelyek jobban alkalmasak a valós hálózatok modellezésére, és vannak olyanok is, amelyek ehhez túl szabályosak, ezért kevésbé életszerűek. Most ezekre nézünk egy-egy példát. A járványterjedés vizsgálata során alkalmazott gráfmodelleket a 3.1. szakaszban mutatjuk be.

Erdős–Rényi véletlen gráf

Ez talán az egyik legegyszerűbb gráfmodell, amely több változatban is jelen van. Az eredeti, Erdős Pál és Rényi Alfréd által 1959-ben leírt modellben [9] adott N csúcs és M él, és az összes ilyen gráf közül egyenlő valószínűséggel választjuk az egyiket. Tehát például $N = 3$ és $M = 2$ esetén három különböző gráf létezik és ezek mindegyikét egyharmad valószínűséggel választjuk. A másik változatban, amelyet

ugyancsak Erdős-Rényi véletlen gráfnak hívunk, az élek száma nem előre meghatározott. Ez a modell Edgar Gilbert [10] nevéhez fűződik, amelyben adott N csúcs esetén a gráfot úgy kapjuk, hogy minden élet egymástól függetlenül p valószínűséggel húzunk be. Ha jobban belegondolunk, ez a gyakorlatban ritkán teljesül. Például nem igaz, hogy bármelyik ember egyforma, p valószínűséggel adhat át bármely másik embernek egy betegséget, illetve, hogy valaki egyforma eséllyel lesz fertőzött, akkor is ha otthon, egy kis közösségben vagy a munkahelyén fertőződik meg. Így ez a gráfmodell kevésbé alkalmas valós hálózatok modellezésére.

Barabási–Albert véletlen gráf

Ez a növekedő gráfmodell Barabási Albert-László és tanítványa, Albert Réka nevéhez fűződik, amelyet 1999-ben dolgoztak ki [11]. Tulajdonképpen ez egy véletlenszerű algoritmus olyan típusú gráfok megkonstruálására, amelyek gyakran jelennek meg a valós társadalmi hálózatok esetén. Az algoritmus lényege a következő: egy kis gráfból kiindulva csúcsok egyenkénti hozzávételével hozzuk létre az N csúcsú gráfot, ahol egy új csúcs mindig adott fokszámmal rendelkezik, és egy már meglévő csúcs annál nagyobb valószínűséggel kap új élet, minél nagyobb a fokszáma. A valóságban ez egy gyakori jelenség, például egy munkahelyi közösségbe gyakran kerülnek be új emberek, ezáltal növekszik az ismeretségek száma, de nem mindenkinél egyforma mértékben, ugyanis egy ismertebb személy, mint például egy vezető nagyobb valószínűséggel szerez új ismeretséget.

2. fejezet

Járványterjedés hipergráfokon

2.1. A modell

Járványterjedés esetén a betegségre fogékony, vagy más néven egészséges egyénekre gyakorolt nyomás a fertőző szomszédok számával növekszik. Ez tulajdonképpen minden terjedési folyamat alapja, hiszen minél összefüggőbb egy hálózat, annál effektívebb a terjedés. Például sűrűbb úthálózat esetén általában könnyebben tudunk eljutni egy adott pontból egy másikba, míg ritkább hálózatnál nehezebben. Valós hálózatokban gyakran megjelenik a közösségi struktúra, ahol jelentős különbség van az egyes csoportokon belüli, illetve a csoportok közötti kapcsolatok számában. Ennek megfelelően a terjedési folyamat bizonyos csoportokon belül sokkal erősebb, míg más csoportok között kevésbé meghatározó. Ezt vizsgálják az úgynevezett háztartás típusú modelleknél. Következésképpen járványterjedés esetén jelentős szerepet játszanak a közösségek, mint például háztartások, iskolák, munkahelyek, stb. Ugyanakkor megfigyelhető, hogy míg a közösségek kapcsolati struktúrája kulcsfontosságú a terjedési folyamatok szempontjából, addig az egyes csoportok belső szerkezete szinte irreleváns [12]. Ezen megfigyelések adták motivációul a járványterjedés hipergráfokon való vizsgálatát.

Egy társadalmi hálózat modellezése rendszerint gráfok segítségével történik. Ebben a dolgozatban a hagyományos értelemben vett gráfok helyett ezt hipergráfokkal szeretnénk megvalósítani. A hipergráf egy (V, E) pár, ahol $V = \{v_1, v_2, \dots, v_N\}$ a csúcsok nem-üres halmaza, $E = \{h_1, h_2, \dots, h_M\}$ pedig az élek halmaza, ahol minden h_i hiperél V -nek egy részhalmaza. A csúcsokat másképpen hipercsúcsoknak nevezhetjük. Hagyományos értelemben vett gráffal ellentétben itt egy él kettőnél több csúcsot is összeköthet. A populáció egyedei a hipergráf csúcsainak, míg a csúcsokat összekötő hiperélek egy-egy háztartásnak vagy munkahelynek felelnek meg. Ebben a dolgozatban az 1.2. szakaszban definiált SIS (susceptible-infected-susceptible) típusú járványterjedést vizsgáljuk, amely például nemi úton terjedő betegségek leírására

alkalmazható. Egy másik módosítás a modellünkben a következő megállapításon alapszik. Valós modellek esetében a fertőzési ráta nem feltétlenül arányos a fertőző szomszédok számával, azaz az átmenet rátája fertőzés esetén nem szükségszerűen egyenlő a $k\tau$ értékkel. Példának okáért, ha egy 9 fős munkahelyi közösségben a fertőzők száma 4-ről 8-ra növekszik, nem feltétlenül jelenti azt, hogy a fertőzési nyomás is dupla annyi lesz ezen a helyen. Így ebben a modellben bevezetünk egy f függvényt, amelynek segítségével a nemlineáris függés leírható. Ekkor a fertőzési rátája $f(k)\tau$ értékre módosul. Tegyük fel, hogy egy egyed két hiperélhez tartozik, azaz például van családja és munkája, ekkor a fertőzési rátája $(f(k_1) + f(k_2))\tau$, ahol k_1 az otthoni, k_2 pedig a munkahelyi betegek száma. Hagyományos értelemben vett gráf esetén, azaz ha a hiperéleket egy-egy klikknek tekintenénk, akkor ez az érték $(f(k_1 + k_2))\tau$ lenne, és ennél fogva nem tudnánk megmondani, hogy a $k_1 + k_2$ beteg hogy aránylik szét az egyes csoportokban. Ebből is látható, hogy hipergráfok használatával pontosabb leírást adhatunk az egyes közösségekről.

Mindent összevetve egy S típusú csúcs kis Δt idő alatt $1 - \exp(-r\Delta t)$ valószínűséggel változhat I típusúvá, ahol az átmenet rátája

$$(2.1) \quad r = \tau \sum_h f(k_h),$$

ahol azon $h \in E$ hiperélekre összegzünk, amelyek tartalmazzák az adott S típusú csúcsot és k_h jelöli az adott csúcs fertőző szomszédainak számát a h hiperélben, f pedig egy adott függvény. Gráfoknál ez az f függvény sok esetben az identitást jelöli, így a fertőzési ráta megegyezik a korábbival, azaz $r = \tau k$, ahol k az összes fertőzött szomszéd száma. Egy másik lehetséges, neurális hálózatok [13] esetén gyakran alkalmazott függvény a tangens hiperbolikus, vagy egy ahhoz hasonló függvény, ami a 0 környezetében közel van az identitás függvényéhez, azonban nagy értékek esetén már konstans függvényként viselkedik. Utóbbi megfelel az elvárásainknak, miszerint sok beteg szomszéd már nem növeli az átmenet valószínűségét. Tehát az például, hogy 10 kollégánk közül 8 vagy 9 beteg, nagyjából egyforma kockázattal van ránk a betegség terjedése szempontjából. Alapos vizsgálatokkal a [2] cikkben a szerzők arra jutottak, hogy a folyamatok megértéséhez elegendő a legegyszerűbb, egy szakaszon lineáris, majd konstans függvény használata. Ennek értelmében a (2.1)-beli f függvény legyen a következő:

$$(2.2) \quad f(x) = \begin{cases} x, & \text{ha } 0 \leq x \leq c, \\ c, & \text{ha } x > c. \end{cases}$$

Ilyenkor a $c \in \mathbb{R}^+$ paramétert küszöbértéknek nevezzük. Ez azt jelenti, hogy ha egy hiperélben a fertőző csúcsok száma kisebb, mint ez a c érték, akkor a hagyományos

értelemben vett eset következik be, azaz a fertőzési ráta arányos a fertőzött szomszédok számával. Ugyanakkor, ha a fertőző csúcsok száma nagyobb, mint a küszöbérték, akkor a fertőzők száma már nem növeli a fertőzés rátáját az adott hiperélben.

2.2. Alapegyenlet és a várható értékre vonatkozó egyenletek

Célunk az állapotvalószínűségekre vonatkozó lineáris differenciálegyenlet-rendszer levezetése, amit alapegyenletnek nevezünk, illetve ennek segítségével az úgynevezett mean-field típusú, azaz átlagolt egyenletek levezetése, ami hagyományos gráfokra a [14] cikkben található meg.

2.2.1. Alapegyenlet

Tekintsünk egy N csúcsú tetszőleges hipergráfot, amelyen SIS típusú járványterjedést vizsgálunk. Az állapottér ennek megfelelően a 2^N elemet tartalmazó $\{S, I\}^N$ halmaz, ugyanis bármely csúcs kétféle állapotban lehet. Ezt a halmazt célszerű a fertőző csúcsok száma szerint $N + 1$ részhalmazzra bontani. Legyen S^0 az az állapot, amikor nincs fertőző csúcs a hipergráfban, ezt az $S^0 = (S, S, \dots, S)$ vektor jelöli, ahol az egyes koordináták az egyes csúcsok állapotait mutatják. Jelölje S^k azon állapotok halmazát, amelyekben k darab fertőző csúcs található. Végül S^N legyen az az állapot, amikor minden csúcs fertőző, azaz $S^N = (I, I, \dots, I)$.

Az S^k halmaz elemeit jelölje $S_1^k, S_2^k, \dots, S_{c_k}^k$, ahol $c_k = \binom{N}{k}$. Ekkor az SIS típusú járványterjedésnek megfelelően az S_j^k állapotban az l -edik csúcs típusa $S_j^k(l) = I$ vagy $S_j^k(l) = S$, továbbá a rendszer állapota kétféleképpen változhat. A kérdés, hogy ezek hogyan írhatóak le. Fertőzés esetén egy S típusú csúcs I típusú lesz. Az átmenetek során a fertőző csúcsok száma legfeljebb eggyel változhat, így ez egy olyan $S_j^k \rightarrow S_i^{k+1}$ átmenet, ahol i és j olyanok, hogy létezik l csúcs, hogy $S_j^k(l) = S$, $S_i^{k+1}(l) = I$ és $S_j^k(m) = S_i^{k+1}(m)$ minden $m \neq l$ esetén. Továbbá a fertőzés csak fertőző szomszéd esetén lehetséges, azaz létezik $r \neq l$, amelyre $S_j^k(r) = I$ és l, r egy hiperélben vannak. Gyógyulás esetén egy I típusú csúcs S típusú lesz. Ez egy olyan $S_j^k \rightarrow S_i^{k-1}$ átmenet, ahol i és j olyanok, hogy létezik l csúcs, hogy $S_j^k(l) = S$, $S_i^{k-1}(l) = I$ és $S_j^k(m) = S_i^{k-1}(m)$ minden $m \neq l$ esetén. Több feltétel nem szükséges, mert ezt a folyamatot a szomszédok állapota már nem befolyásolja.

Jelölje $X_j^k(t)$ annak a valószínűségét, hogy a t időpillanatban a rendszer az S_j^k állapotban van. Ekkor legyen $k = 0, 1, \dots, N$ esetén $X^k(t) = (X_1^k(t), X_2^k(t), \dots, X_{c_k}^k(t))^T$ egy olyan c_k nagyságú vektor, amely a k darab fertőzött csúcsot tartalmazó állapotok valószínűségeit tartalmazza. A fenti átmenetek az $X_j^k(t)$ változókra egy li-

neáris, állandó együtthatós differenciálegyenlet-rendszert határoznak meg. Ezt nevezzük alapegyenletnek, ahol a differenciálegyenlet jobb oldala az alábbi speciális blokk-tridiagonális alakban írható fel, ugyanis az átmenetek során a fertőzött csúcsok száma legfeljebb egyel változhat:

$$(2.3) \quad \dot{X}^k = A^k X^{k-1} + B^k X^k + C^k X^{k+1}, \quad k = 0, 1, \dots, N,$$

ahol az A^0 és C^N nulla mátrixok. Az egyenlet mátrix alakban is felírható:

$$\dot{X} = PX,$$

ahol

$$P = \begin{pmatrix} B^0 & C_0 & 0 & 0 & 0 & 0 \\ A^1 & B^1 & C^1 & 0 & 0 & 0 \\ 0 & A^2 & B^2 & C^2 & 0 & 0 \\ 0 & 0 & A^3 & B^3 & C^3 & 0 \\ \vdots & \vdots & \dots & \dots & \dots & \vdots \\ 0 & 0 & \dots & \dots & A^N & B^N \end{pmatrix}.$$

Az A^k mátrixok a fertőzés, a C^k mátrixok pedig a gyógyulás folyamatát írják le. Az alapegyenlet hagyományos gráfok esetén ugyanígy írható fel, azonban a benne szereplő mátrixok leírása módosul. Ezen mátrixok elemeit a következő módon határozhatjuk meg.

Az A^k egy olyan $c_k \times c_{k-1}$ -es mátrix, amelynek i -edik sorának j -edik eleme megadja az S_j^{k-1} állapotból az S_i^k állapotba történő átmenet rátáját, ezt jelölje $A_{i,j}^k$. Ez az elem pontosan akkor nem nulla, ha létezik l csúcs, amelyre $S_j^{k-1}(l) = S$, $S_i^k = I$ és $S_j^{k-1}(m) = S_i^k(m)$ minden $m \neq l$ csúcs esetén, továbbá létezik egy $r \neq l$ csúcs, amelyre $S_j^{k-1}(r) = I$ és l és r egy hiperélben vannak. Ekkor

$$(2.4) \quad A_{i,j}^k = \tau \sum_{h:l \in h} f(N_h(S_j^{k-1})),$$

ahol $N_h(S_j^{k-1})$ megadja, hogy a h hiperélnek az S_j^{k-1} állapotban hány fertőzött csúcsa van, f pedig a korábban definiált (2.2) függvény. Ha ezeket az értékeket összegezzük a sorokra, akkor minden $j \in \{1, 2, \dots, c_{k-1}\}$ esetén teljesül, hogy

$$(2.5) \quad \sum_{i=1}^{c_k} A_{i,j}^k = \tau N_{SI}^f(S_j^{k-1}),$$

ahol

$$(2.6) \quad N_{SI}^f(S_j^k) = \sum_{l:S_j^k(l)=S} \sum_{h:l \in h} f(N_h(S_j^k)).$$

Ennek értelmében $N_{SI}^f(S_j^k)$ megadja az S_j^k állapotban egészséges l csúcsokra vonatkozó $\sum_{h:l \in h} f(N_h(S_j^k))$ értékek összegét.

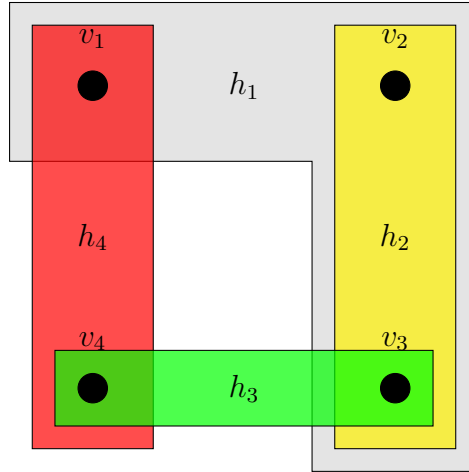
A C^k egy olyan $c_k \times c_{k+1}$ -es mátrix, amelynek i -edik sorának j -edik eleme megadja az S_j^{k+1} állapotból az S_i^k állapotba történő átmenet rátáját, ezt jelölje $C_{i,j}^k$. Ez az elem pontosan akkor nem nulla, ha létezik l csúcs, amelyre $S_j^{k+1}(l) = I$, $S_i^k = S$ és $S_j^{k+1}(m) = S_i^k(m)$ minden $m \neq l$ csúcs esetén. Ekkor $C_{i,j}^k = \gamma$. Az S_j^{k+1} állapotban $k+1$ fertőző csúcs van, ennél fogva a C^k mátrix j -edik oszlopában összesen $k+1$ nem nulla elem van, amelyek mindegyike γ -val egyenlő, így minden $j \in \{1, 2, \dots, c_{k+1}\}$ esetén fenáll

$$(2.7) \quad \sum_{i=1}^{c_k} C_{i,j}^k = \gamma(k+1).$$

A B^k egy olyan $c_k \times c_k$ -s négyzetes mátrix, amelynek i -edik sorának j -edik eleme megadja az S_j^k állapotból az S_i^k állapotba történő átmenet rátáját, ezt jelölje $B_{i,j}^k$. Ez az elem pontosan akkor nem nulla, ha $i = j$. Ugyanis ha $i \neq j$, akkor egy időben egyszerre legalább két csúcs állapota változna és ezt kizártuk. A P mátrix minden oszlopösszege 0, ennek megfelelően a diagonális mátrix elemei a következők:

$$(2.8) \quad B_{i,i}^k = -\sum_{j=1}^{c_{k+1}} A_{j,i}^{k+1} - \sum_{j=1}^{c_{k-1}} C_{j,i}^{k-1}.$$

2.1. Példa. Írjuk fel az alapegyenletet az alábbi $N = 4$ csúcsú hipergráfra.



A hipergráf illeszkedési mátrixa:

$$\mathcal{J} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Négy csúcsú hipergráf esetén az állapottér 2^4 elemet tartalmaz, amelyeket a fertőző csúcsok száma szerint $4 + 1$ osztályba sorolhatunk:

$$\begin{aligned} S^0 &= \{SSSSS\}, \\ S^1 &= \{S_1^1, S_2^1, S_3^1, S_4^1\} = \{ISSS, SISS, SSIS, SSSI\}, \\ S^2 &= \{S_1^2, \dots, S_6^2\} = \{IISS, ISIS, ISSI, SIIS, SISI, SSII\}, \\ S^3 &= \{S_1^3, S_2^3, S_3^3, S_4^3\} = \{IIIS, IISI, ISII, SIII\}, \\ S^4 &= \{IIII\}. \end{aligned}$$

Ekkor például $X_1^2(t)$ annak a valószínűsége, hogy a t időpillanatban a rendszer az $IISS$ állapotban van, azaz v_1 és v_2 fertőző csúcsok, a többi pedig egészséges. $X^2(t) = (X_1^2(t), \dots, X_6^2(t))$ pedig a két fertőző csúcsot tartalmazó állapotok valószínűségeit tartalmazza. $X = (X^0, X^1, X^2, X^3, X^4)$ és a P mátrix ebben az esetben az alábbi alakú:

$$P = \begin{pmatrix} B^0 & C_0 & 0 & 0 & 0 \\ A^1 & B^1 & C^1 & 0 & 0 \\ 0 & A^2 & B^2 & C^2 & 0 \\ 0 & 0 & A^3 & B^3 & C^3 \\ 0 & 0 & 0 & A^4 & B^4 \end{pmatrix},$$

ahol a részmátrixok a következők:

$$A^1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, A^2 = \tau \begin{pmatrix} f(1) & f(1) & 0 & 0 \\ f(1) & 0 & f(1) & 0 \\ f(1) & 0 & 0 & f(1) \\ 0 & f(1) & f(1) & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & f(1) & f(1) \end{pmatrix},$$

Itt például az A^2 ötödik sorának minden eleme nulla, amelyek az alábbi átmenetek rátáit adják: $ISSS \rightarrow SISI$, $SISS \rightarrow SISI$, $SSIS \rightarrow SISI$, $SSSI \rightarrow SISI$. Az első és a harmadik átmenet azért nem valósulhat meg, mert az állapotok között több, mint egy helyen van különbség, de az átmenetek során a fertőzött csúcsok száma legfeljebb eggyel változat. A második átmenet esetén a v_4 csúcs fertőződik meg, de az $SISS$ állapotban csak v_2 fertőző, amely nem szomszédja v_4 -nek, így ez az átmenet sem lehetséges. Az utolsó az előbbiekből adódik v_2 és v_4 felcserélésével.

$$A^3 = \tau \begin{pmatrix} f(2) + f(1) & f(2) + f(1) & 0 & f(2) & 0 & 0 \\ f(1) & 0 & f(1) & 0 & 2f(1) & 0 \\ 0 & 2f(1) & 2f(1) & 0 & 0 & 2f(1) \\ 0 & 0 & 0 & f(1) & 2f(1) & f(1) \end{pmatrix},$$

$$A^4 = \tau \left(2f(1), 2f(1) + f(2), f(1) + f(2), f(1) + f(2) \right).$$

Nézzük meg az $A_{1,2}^4 = \tau(2f(1) + f(2))$ elemet, amely az $S_2^3 = IISI \rightarrow IIII = S^4$ átmenet rátája. Ekkor a v_3 csúcshoz fertőződik meg, amelynek az $IISI$ állapotban több fertőző szomszédja is van, erre számítjuk ki az $N_{SI}^f(IISI)$ értéket. A h_1 élben mindkét szomszédja fertőző, ebből adódik az $f(2)$, továbbá a h_2 és h_3 élekben egy-egy fertőző szomszédja van, innen kapjuk a $f(1) + f(1)$ értéket.

$$B^0 = (0), C^0 = (\gamma, \gamma, \gamma, \gamma, \gamma), C^1 = \begin{pmatrix} \gamma & \gamma & \gamma & 0 & 0 & 0 \\ \gamma & 0 & 0 & \gamma & \gamma & 0 \\ 0 & \gamma & 0 & \gamma & 0 & \gamma \\ 0 & 0 & \gamma & 0 & \gamma & \gamma \end{pmatrix}, C^2 = \begin{pmatrix} \gamma & \gamma & 0 & 0 \\ \gamma & 0 & \gamma & 0 \\ 0 & \gamma & \gamma & 0 \\ \gamma & 0 & 0 & \gamma \\ 0 & \gamma & 0 & \gamma \\ 0 & 0 & \gamma & \gamma \end{pmatrix}, C^3 = \begin{pmatrix} \gamma \\ \gamma \\ \gamma \\ \gamma \end{pmatrix}.$$

Itt például $C_{2,1}^2 = \gamma$, mert az $S_1^3 = IIIS \rightarrow ISIS = S_2^2$ átmenet lehetséges, viszont az $S_1^3 = IIIS \rightarrow ISSI = S_3^2$ nem az, így a $C_{3,1}^2 = 0$.

Ekkor az alapegyenlet $k = 2$ esetén: $\dot{X}^2 = A^2 X^1 + B^2 X^2 + C^2 X^3$, ahol például

$$\begin{aligned} \dot{X}_1^2 = \dot{X}_{IISS} &= \tau f(1) X_{IISS} + \tau f(1) X_{SISS} - \\ &- 2\gamma X_{IISS} - \tau(f(2) + f(1)) X_{IISS} - \tau f(1) X_{IISS} + \\ &+ \gamma X_{IIIS} + \gamma X_{IISI} \end{aligned}$$

Látható, hogy már kis N esetén is sok számolással jár a P mátrix kiszámítása. Ilyenkor a rendszer 2^N egyenletből áll, amely az összes lehetséges állapot valószínűségének időbeli változását adja meg. Valós modellek esetén, nagy N -re a megoldása előállítása numerikus módszerekkel sem lehetséges. Így ez az eljárás főként elméleti szempontból hasznos. Az alapegyenlet egyszerűsítésének egy lehetséges módja az úgynevezett mean-field típusú egyenletek.

2.2.2. Várható értékekre vonatkozó differenciálegyenletek

A mean-field elmélet alapja, hogy az egyes állapotok valószínűségei helyett bizonyos várható értékeket vizsgálunk. Ezek közül a két legfontosabb: a fertőzők, illetve az egészségesek számának várható értéke, ezeket jelölje $[I]$ és $[S]$, ahol

$$(2.9) \quad [I](t) = \sum_{k=0}^N k \sum_{j=1}^{c_k} X_j^k(t), \quad [S](t) = \sum_{k=0}^N (N - k) \sum_{j=1}^{c_k} X_j^k(t).$$

Ezen mennyiségekre szeretnénk differenciálegyenleteket felírni, amelyekhez szükségünk van az SI típusú él számának várható értékére:

$$(2.10) \quad [SI](t) = \sum_{k=0}^N \sum_{j=1}^{c_k} N_{SI}^f(S_j^k) X_j^k(t),$$

ahol N_{SI}^f a korábban definiált (2.6) mennyiség. Ekkor az $[I](t)$, illetve $[S](t)$ függvényekre az alábbi differenciálegyenleteket kapjuk.

2.2. Tétel. *Az $[I](t)$ és $[S](t)$ várható értékek hipergráfok esetén kielégítik az alábbi differenciálegyenleteket:*

$$(2.11) \quad [\dot{S}] = \gamma[I] - \tau[SI],$$

$$(2.12) \quad [\dot{I}] = \tau[SI] - \gamma[I].$$

Bizonyítás. Legyen $e_k = (1, 1, \dots, 1)$ egy csupa 1-esekből álló, c_k dimenziójú sorvektor. Ezt felhasználva

$$\sum_{j=1}^{c_k} X_j^k = e_k X^k$$

adódik. Ennélfogva (2.9) az alábbi alakban írható:

$$(2.13) \quad [I](t) = \sum_{k=0}^N k e_k X^k, \quad [S](t) = \sum_{k=0}^N (N - k) e_k X^k.$$

Továbbá a (2.8) egyenlet a következő alakot ölti:

$$B_{i,i}^k = -(e_{k+1} A^{k+1})_i - (e_{k-1} C^{k-1})_i,$$

ahol B^k diagonális mátrix, ezért

$$B_{i,i}^k = (e_k B^k)_i.$$

Így minden $i = 1, \dots, c_k$ esetén

$$(e_k B^k)_i = -(e_{k+1} A^{k+1})_i - (e_{k-1} C^{k-1})_i.$$

Ebből adódóan minden $k = 0, 1, \dots, N$ esetén teljesül, hogy

$$(2.14) \quad e_{k+1} A^{k+1} + e_k B^k + e_{k-1} C^{k-1} = 0,$$

ahol A^{N+1} és C^{-1} nulla mátrixok.

Innentől a (2.12) egyenletet szeretnénk belátni, ugyanis az $[S]$ mennyiségre vonatkozó (2.11) egyenlet hasonlóan vezethető le.

Az $[I](t)$ függvényt deriválva, és a (2.3) egyenletet felhasználva:

$$\begin{aligned} [\dot{I}] &= \sum_{k=0}^N k e_k \dot{X}^k = \sum_{k=0}^N k e_k \left(A^k X^{k-1} + B^k X^k + C^k X^{k+1} \right) = \\ &= \sum_{k=1}^N k e_k A^k X^{k-1} + \sum_{k=0}^N k e_k B^k X^k + \sum_{k=0}^{N-1} k e_k C^k X^{k+1} = \\ &= \sum_{k=0}^{N-1} (k+1) e_{k+1} A^{k+1} X^k + \sum_{k=0}^N k e_k B^k X^k + \sum_{k=1}^N (k-1) e_{k-1} C^{k-1} X^k = \\ &= \sum_{k=0}^N \left((k+1) e_{k+1} A^{k+1} + k e_k B^k + (k-1) e_{k-1} C^{k-1} \right) X^k. \end{aligned}$$

Ekkor (2.14) alapján az alábbi differenciálegyenletet kapjuk:

$$[\dot{I}] = \sum_{k=0}^N \left(e_{k+1}A^{k+1} - e_{k-1}C^{k-1} \right) X^k.$$

Ezt követően (2.12) az alábbi állításból következik. □

2.3. Állítás. *Az A^k és C^k mátrixok esetén az alábbiak teljesülnek:*

1. $e_{k-1}C^{k-1} = \gamma k e_k$,
2. $\sum_{k=0}^N e_{k-1}C^{k-1}X^k = \gamma[I]$,
3. $\sum_{k=0}^N e_{k+1}A^{k+1}X^k = \tau[SI]$.

Bizonyítás. A (2.7) egyenlet szerint minden $j \in \{1, 2, \dots, c_k\}$ esetén

$$(e_{k-1}C^{k-1})_j = \sum_{i=1}^{c_{k-1}} C_{i,j}^{k-1} = \gamma k.$$

Ebből az állítás 1. része már következik, azaz $e_{k-1}C^{k-1} = \gamma k e_k$.

Az állítás 2. pontja a (2.13) egyenlet felhasználásával, az 1. részből következik.

A (2.5) egyenlet szerint minden $j \in \{1, 2, \dots, c_k\}$ esetén

$$(e_{k+1}A^{k+1})_j = \sum_{i=1}^{c_{k+1}} A_{i,j}^{k+1} = \tau N_{SI}^f(S_j^k),$$

amelyből az állítás 3. pontja következik, ugyanis:

$$\sum_{k=0}^N e_{k+1}A^{k+1}X^k = \sum_{k=0}^N \sum_{j=1}^{c_k} (e_{k+1}A^{k+1})_j X_j^k = \tau \sum_{k=0}^N \sum_{j=1}^{c_k} N_{SI}^f(S_j^k) X_j^k = \tau[SI].$$

□

A 2.2. Tétel egy-egy differenciálegyenletet biztosít számunkra, amelyek megoldásával megkapjuk a fertőző, illetve az egészséges csúcok számának várható értékét. Viszont ez az egyenlet nem zárt, ugyanis mindkét differenciálegyenlet tartalmazza az $[SI]$ mennyiséget, amit nem ismerünk. Az egyenletrendszer lezárásnak egyik lehetséges módja, hogy az $[SI]$ értékre is differenciálegyenletet írunk fel, ezzel a dolgozatban nem foglalkozunk. Egy másik lehetőség, hogy közelítjük az SI élek számának várható értékét $[I]$ és $[S]$ függvényében, amiről a 3.4. alfejezetben lesz szó.

3. fejezet

Szimulációk

A mean-field típusú egyenlet korlátozott abban az értelemben, hogy nem tudja figyelembe venni a fertőzés útját, vagyis hogy a betegség melyik egyedről melyik egyedre jut át. Ennek leírásához meg kell állapítani a kapcsolati struktúrát, amely megadja, hogy melyik egyed melyikkel van kapcsolatban a fertőzhetőség szempontjából. Ha ennek leírása lehetséges, akkor a struktúrát gráffal adjuk meg. A hálózat pontos ismeretében Monte-Carlo szimulációval lehet vizsgálni a betegség terjedésének folyamatát. Az algoritmus pontos leírásával a 3.2. alfejezetben foglalkozunk. A numerikus vizsgálatokhoz MATLAB programot használtunk.

3.1. A szimulációk során alkalmazott gráfmodellek

Ahogy korábban említésre került, társadalmi hálózatok modellezése véletlen gráfok segítségével történik. Ezért véletlen hipergráfok segítségével szeretnénk a járványterjedést modellezni. Ebben a fejezetben három módszert mutatunk be, amelyek segítségével véletlen hipergráfokat generálhatunk. Az első modell gráfok generálásának egy általánosított algoritmus, ahol adott a társadalmi szerkezet. A másik két módszer egy-egy ismert modellen alapszik.

3.1.1. Két állapotú hipergráf

Az első általunk generált hipergráfban minden egyed pontosan két hiperélhez tartozik. Az egyik megfelel az egyén háztartásának, a másik pedig a munkahelyének. Tegyük fel, hogy minden háztartás különálló, illetve minden egyednek pontosan egy munkahelye van. Ekkor az élek egy bizonyos része megfelel egy-egy háztartásnak, amelyek V -nek, vagyis a csúcsok halmazának egy partícióját adja, ezt jelöljük \mathcal{H} -val. Az egyszerűség kedvéért tegyük fel, hogy minden háztartás pontosan H tagból áll, így (V, \mathcal{H}) egy H -uniform hipergráf, amelyben bármely két hiperél egymástól

független és minden hiperél pontosan H csúcsot tartalmaz. Hasonlóan a munkahelyeket jelentő élek partíciót alkotnak, ezt jelöljük \mathcal{M} -mel. Feltesszük, hogy minden munkahely pontosan M tagból áll. Ekkor a csúcsok száma egész számú többszöröse H -nak és M -nek, amelyeket az alábbiak szerint véletlenszerűen generálunk. Az első háztartást jelöljük \mathcal{H}_1 -el, amelynek tagjait véletlenszerűen választjuk ki a csúcsok V halmazából. Ezt követően a második háztartás tagjait ugyancsak véletlenszerűen választjuk ki, de már csak a $V \setminus \mathcal{H}_1$ halmazból. Ezt az eljárást addig folytatjuk, amíg nincs meg az összes \mathcal{H}_i háztartás, ahol $i = \{1, 2, \dots, \frac{|V|}{H}\}$. A véletlenszerű kiválasztás az alábbiak szerint könnyen megvalósítható. Először permutáljuk a csúcsok halmazát, majd az első H elemet tegyük bele \mathcal{H}_1 -be, a második H elemet \mathcal{H}_2 -be, és így tovább. A \mathcal{M}_j munkahelyek elkészítése hasonló módon történik, ahol $j = \{1, 2, \dots, \frac{|V|}{M}\}$. Befejezésékképpen a két uniform hipergráf egyesítéséből megkapjuk a $(V, \mathcal{H} \cup \mathcal{M})$ két állapotú hipergráfot.

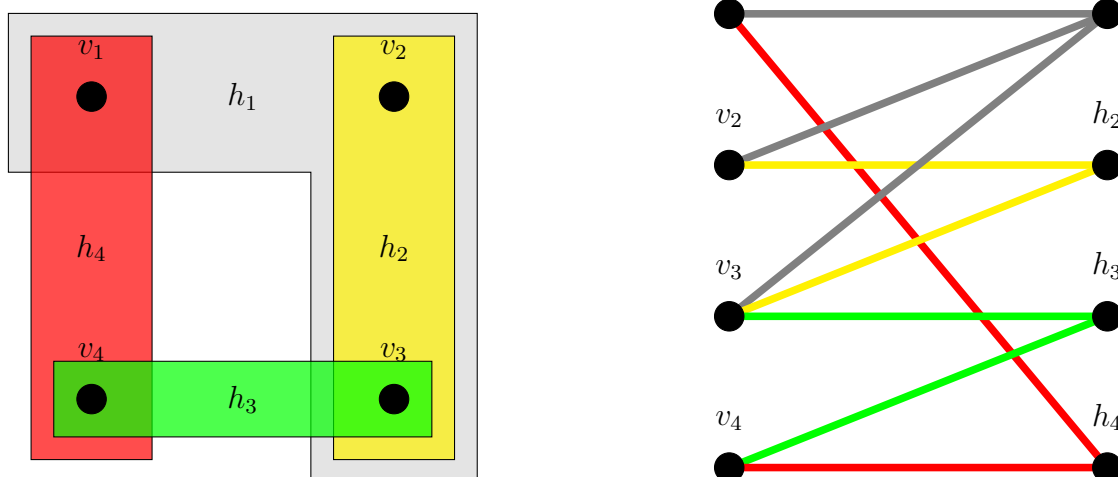
3.1.2. Barabási–Albert klikk modell

Második modellünk a Barabási–Albert-féle preferált kapcsolódási (angolul: preferential attachment) modellen alapszik. A módszer a következő: Barabási–Albert gráfokat generálunk az 1.3. alfejezetben leírt algoritmus alapján, majd ezeket a gráfokat hipergráffá alakítjuk át az alábbiak szerint. Hagyományos értelemben vett gráf esetén minden él legfeljebb két csúcsot köt össze, amely nemigazán alkalmas több tagból álló közösségi struktúrák ábrázolására. Ugyanakkor feltételezhető, hogy az egyes közösségeken belül bármely két tag kapcsolatban áll egymással, így gráfok esetén a teljes részgráfok, más néven klikkek egy-egy közösséget határoznak meg. Az MTA-ELTE *Statisztikus és Biológiai Kutatócsoport* által létrehozott CFinder 2.0.6. szoftver [18] segítségével könnyedén megtalálhatóak ezek a klikkek, amelyeket a hipergráf hiperéleinek tekintünk. A preferált kapcsolódási modellnek köszönhetően ez sok kis méretű és néhány nagyméretű hiperélet eredményez a hipergráfban.

3.1.3. Konfigurációs modell

Ebben a modellben adott számú csúcs esetén egy olyan véletlen hipergráfot szeretnénk generálni, amely uniform, azaz minden hiperél ugyanannyi csúcsot tartalmaz, illetve reguláris, azaz a csúcsok fokszáma azonos. Rögzítsük ezeket az értékeket, vagyis a csúcsok száma legyen N , a hiperélek mérete e és a csúcsok fokszáma d . Ekkor a hiperélek száma $M = Nd/e$. Egy általánosabb megközelítésben minden hiperél mérete és minden csúcs fokszáma egyénileg van meghatározva. Például az élek felének a mérete e_1 , míg a másik felének e_2 , és ugyanígy csúcsok esetében az értékek d_1 és d_2 .

Ennél a módszernél az alapötlet a következő: minden hipergráf azonosítható egy páros gráffal. A páros gráf első színosztálya (V) legyen az adott hipergráf csúcsainak halmaza, a második színosztály (U) pedig azon pontokból álljon, amelyek egy-egy hiperélnak felelnek meg. Ekkor egy V -beli csúcs pontosan akkor szomszédos egy U -beli csúccsal, ha az U -beli csúcsnak megfelelő hiperél tartalmazza a V -beli csúcsot a hipergráfban. Ezt a reprezentálást a 3.1. ábra szemlélteti.



3.1. ábra. Hipergráf azonosítása páros gráffal

A mi esetünkben minden csúcs d hiperélben van benne, azaz a V színosztályban minden csúcs fokszáma d , továbbá minden él e csúcsot tartalmaz, azaz az U színosztályban minden csúcs fokszáma e . Az általános esetben a V -beli csúcsok előre meghatározott fokszámmal rendelkeznek, továbbá az U -beli csúcsok fokszáma megegyezik a nekik megfelelő hiperélek előre meghatározott méretével.

Ennek következtében elegendő generálni egy véletlenszerű páros gráfot, amihez a konfigurációs modellt használjuk. Először is húzzunk be minden egyes csúcshoz fél-éleket az előre meghatározott fokszámok szerint, majd véletlenszerűen kössük össze az egyes éleket a két színosztály között. Ily módon előfordulhatnak párhuzamos élek is, amelyeket kitörölhetünk a gráfból ugyanis ezen élek várható értéke viszonylag kicsi, különösen akkor, ha az előírt fokszámok N -hez viszonyítva kicsik. Ennek eredményeképpen a hipergráf néhány éle e -nél kevesebb csúcsot fog tartalmazni és néhány csúcsa d -nél kevesebb élben lesz benne, de ennek elhanyagolható hatása van a szimulációs eredményeknél. Általános esetben hasonlóan járunk el, a párhuzamos élek törlése után a hipergráf néhány éle e_1 -nél, néhány éle pedig e_2 -nél kevesebb csúcsot fog tartalmazni, illetve néhány csúcsa d_1 -nél, néhány csúcsa pedig d_2 -nél kevesebb élben lesz benne.

3.2. A szimulációs algoritmus

Jelölje a rendszer t időpontbeli állapotát $x(t) \in \{0, 1\}^N$. Ennek k -adik koordinátája $x_k(t)$, amelynek értéke 0, ha a k -adik csúcs egészséges és 1, ha a k -adik csúcs fertőző. A hipergráf illeszkedési mátrixát jelölje \mathcal{J} , amelynek sorai egy-egy csúcsnak, oszlopai pedig egy-egy élnek felelnek meg, továbbá $\mathcal{J}_{ij} = 1$, ha a j -edik hiperél tartalmazza az i -edik csúcsot, ellenkező esetben $\mathcal{J}_{ij} = 0$. Ekkor az $x(t)\mathcal{J}$ szorzat megadja a fertőző csúcsok számát az egyes hiperélekben, amelynek j -edik koordinátája, $(x(t)\mathcal{J})_j$ megadja a j -edik hiperélben található fertőző csúcsok számát. Mint már említettük, a fertőzést és a gyógyulást is független Poisson folyamatnak tekintjük. Ennek értelmében egy h hiperélben lévő egészséges egyed, k_h fertőző szomszédal, egy rövid Δt idő alatt $1 - \exp(-\tau \sum_h f(k_h)\Delta t)$ valószínűséggel válik fertőzővé. Hasonlóan egy fertőző egyed $1 - \exp(-\gamma\Delta t)$ valószínűséggel válik egészségesé. Így, ha feltételezzük, hogy az i -edik csúcs egészséges a t időpontban, azaz $x_i(t) = 0$, akkor a csúcs megfertőződésének rátája

$$\tau \sum_{j=1}^M \mathcal{J}_{ij} f((x(t)\mathcal{J})_j).$$

Széles körben elterjedt sztochasztikus szimulációt alkalmazunk. Adott t időpontban generálunk egy $r \in [0, 1]^N$ véletlen számokból álló vektort, és ezután futtatjuk az algoritmust minden $i = \{1, \dots, N\}$ csúcsra. Ha az i -edik csúcs a t időpontban egészséges, azaz $x_i(t) = 0$, akkor ez a csúcs a $t + \Delta t$ időpontban fertőzővé válik, ha

$$r_i < 1 - \exp\left(-\tau \sum_{j=1}^M \mathcal{J}_{ij} f((x(t)\mathcal{J})_j)\Delta t\right).$$

Illetve, ha az i -edik csúcs a t időpontban fertőző, azaz $x_i(t) = 1$, akkor ez a csúcs a $t + \Delta t$ időpontban egészségesé válik, ha

$$r_i < 1 - \exp(-\gamma\Delta t).$$

Ha az eljárást kellően kicsi Δt számmal futtatjuk, akkor a szimulációk sokszori megismétlése után átlagot vonva az eredmény jól közelíti a Poisson-folyamatot.

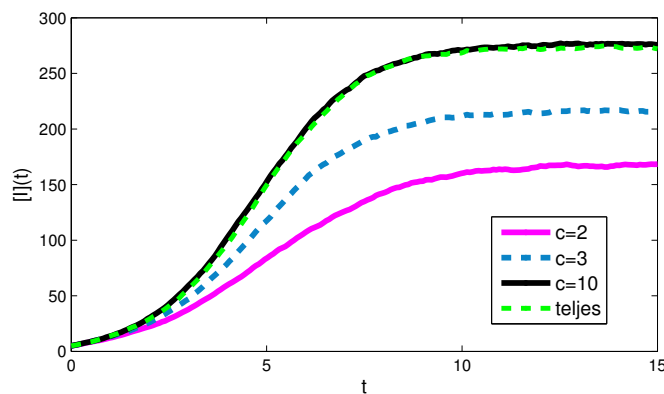
3.3. Szimulációs eredmények

3.3.1. Az f függvény hatása

Ebben a szakaszban szimulációk segítségével arra a kérdésre próbálunk választ adni, hogy az f függvény egyes változatai milyen hatással vannak a terjedési folyamatra. Az általunk használt (2.2) függvény egyetlen paramétertől függ, amelyet

küszöbértéknek neveztünk. Ha egy hiperélben a fertőző egyedek száma kisebb, mint ez a küszöbérték, akkor a fertőzés arányos a fertőző szomszédok számával. Ha pedig a fertőző egyedek száma nagyobb, mint ez a küszöbérték, akkor a fertőző szomszédok száma már nem növeli a fertőzés valószínűségét az adott hiperélben. A szimulációk során kétféle megközelítést használunk. Egyrészt a szimulációkat a korábban bemutatott hipergráf modelleken futtatjuk, másrészt hagyományos értelemben vett gráfokat is alkalmazunk a szimulációs eredmények összehasonlításának céljából. Ezeket a gráfokat a hipergráfokból készítjük el úgy, hogy a hipergráf hiperéleit klikkekre, másnéven teljes részgráfokra cseréljük. Előfordulhat, hogy két csúcs ugyanazon háztartáshoz és munkahelyhez tartozik, ennek következtében a hagyományos értelemben vett gráf esetén a két csúcs között párhuzamos élek lesznek. Azért, hogy az eredeti hipergráfhoz hasonló hálózatot készítsünk, súlyozott gráfokat alkalmazunk, ahol a 2 súllyal rendelkező élek párhuzamos éleket helyettesítenek. Ennek megfelelően a továbbiakban minden szimulációs ábra tartalmazni fogja a súlyozott gráfként való megközelítést is.

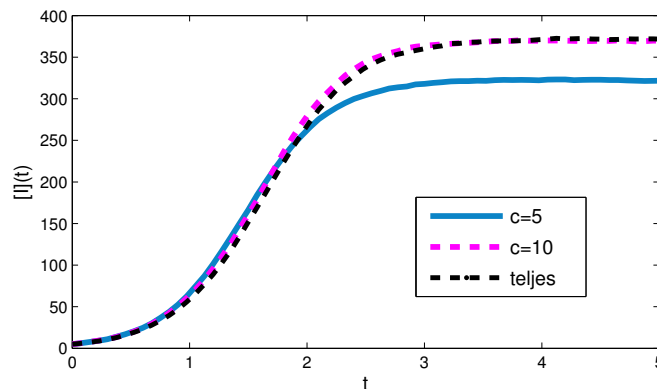
Tekintsük először a két állapotú hipergráf modellt, ahol minden csúcs pontosan két hiperélhez tartozik. A szimulációt $\gamma = 1$ gyógyulási és $\tau = 0.18$ fertőzési rátával különböző c küszöbértékekre futtattuk, $H = 5$ fős háztartások és $M = 10$ fős munkahelyek esetén. A fertőző csúcsok várható számának időbeli változását a 3.2. ábra mutatja három különböző c küszöbérték esetén. Ahogy az várható volt, nagyobb c érték erősebb járványt eredményez. Továbbá $c = 10$ esetén az eljárás a hipergráfon alapvetően megegyezik a hipergráfnak megfelelő súlyozott gráfon. Ennek oka világos, hiszen ilyenkor az f függvénynek nincs hatása. Ugyanis az egyes hiperélek legfeljebb 10 csúcsot tartalmaznak, így $f(k) = k$ minden olyan k értékre, amely a hipergráfban előfordulhat. Emlékeztetőül, k a fertőző szomszédok számát jelöli.



3.2. ábra. Fertőző csúcsok számának időbeli változása $N = 500$ csúcsú hipergráf, $H = 5$ fős háztartások és $M = 10$ fős munkahelyek esetén, $\gamma = 1$ gyógyulási és $\tau = 0.18$ fertőzési rátával. Az ismétlések száma 200, $\Delta t = 0.075$.

Tekintve, hogy a Barabási–Albert klikk modellel a későbbiekben nem foglalkozunk, így csak megemlítjük, hogy ennél a modellnél az előbbihez teljesen hasonló viselkedés tapasztalható. Az egyedüli különbség abban mutatkozik meg, hogy ennek a modellnek jelentősebb a hatása a járvány korai szakaszában, ami azzal magyarázható, hogy ilyenkor a nagyobb klikkek valószínűleg már fertőzőek, és az ezeknek megfelelő hiperélek mérete nagyobb, mint a c küszöbérték.

Végül tekintsük azt a hipergráf modellt, amelyet a konfigurációs modell segítségével készítettünk el. Olyan reguláris hipergráfot hoztunk létre, amelynek minden hiperéle $e = 10$ csúcsot tartalmaz és minden egyes csúcs fokszáma $d = 8$, azaz minden csúcs pontosan 8 hiperélhez tartozik. A szimulációt különböző c küszöbértékekre futtattuk $\gamma = 1$ gyógyulási és $\tau = 0.05$ fertőzési rátával. A fertőző csúcsok várható számának időbeli változását a 3.3. ábra mutatja két különböző c küszöbérték esetén. Továbbra is fennáll, hogy nagyobb c érték erősebb járványt eredményez. A kisebb, $c = 5$ érték itt is az átlagos hiperélméret alatt van, amelynek hatása a korábbiaknak megfelel. Ugyanakkor a nagyobb, $c = 10$ érték megegyezik a hiperélek méretével, amelynek köszönhetően a struktúrának nincs hatása a súlyozott gráffal szemben, mivelhogy $f(k) = k$ minden olyan k értékre, amely a hipergráfban előfordulhat.

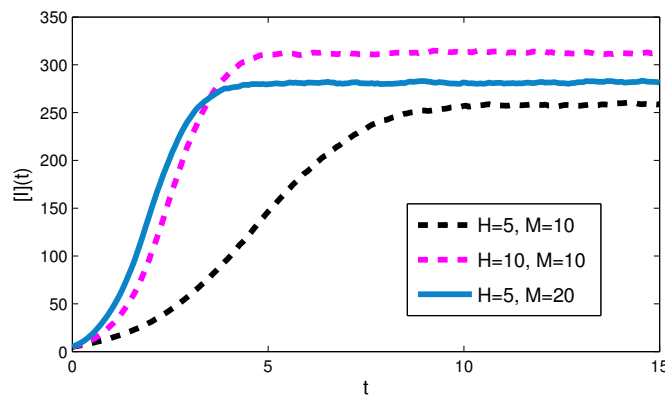


3.3. ábra. Fertőző csúcsok számának időbeli változása a konfigurációs modell alapján generált $N = 500$ csúcsú reguláris hipergráf esetén, amelyben minden csúcs $d = 8$ hiperélhez tartozik, melyek mérete $e = 10$. Továbbá a gyógyulási ráta $\gamma = 1$, míg a fertőzési ráta $\tau = 0.05$. Az ismétlések száma 200, $\Delta t = 0.025$.

3.3.2. A struktúra hatása

Már láttuk, hogy az f paraméterének megválasztása hogyan befolyásolja a terjedési folyamatot. Most rögzítsük le ezt a c paramétert és vizsgáljuk meg, hogy a kétállapotú hipergráf modell milyen hatással van a folyamatra. Ebben a modellben minden csúcs pontosan két hiperélhez, pontosabban egy háztartáshoz és egy mun-

kahelyhez tartozik, ahol a háztartások mérete H , a munkahelyek mérete pedig M . A szimulációt különböző H és M értékekre futtattuk $\gamma = 1$ gyógyulási és $\tau = 0.18$ fertőzési rátával. A fertőző csúcsok várható számának időbeli változását a 3.4. ábra mutatja három különböző H , M értékpár esetén. Megfigyelhető, hogy ha növeljük a hiperélek méretét, például $M=10$ értéket lerögzítjük és H értékét 5-ről 10-re növeljük, akkor a fertőzés erősebbé válik. Ugyanakkor az is látható, hogy ha a hiperélek mérete heterogénebb, azaz $H = M = 10$ helyett $H = 5$ és $M = 20$ esetet nézzük, akkor kezdetben a terjedés folyamat gyorsabb, de az állandósult állapotban alacsonyabb értéket eredményez, mint a homogénebb esetben.



3.4. ábra. Fertőző csúcsok számának időbeli változása $N = 500$ csúcsú hipergráf, H fős háztartások és M fős munkahelyek esetén $\gamma = 1$ gyógyulási és $\tau = 0.18$ fertőzési rátával. A (2.2) függvénybeli küszöbérték $c = 5$. Az ismétlések száma 200, $\Delta t = 0.075$.

A különböző fokszámeloszlások, valamint a különböző méretű élek hasonló hatással vannak a terjedési folyamatra: a kezdeti fázisban a heterogénebb hipergráf gyorsabb terjedést, ugyanakkor a végső stádiumban kevesebb betegszámot eredményez.

3.4. Várható értékre vonatkozó egyenletek lezárása

Célunk a (2.10) kifejezésben megadott $[SI]$ függvény közelítése, amely a (2.6) mennyiségen alapszik. Tekintsük az első hipergráf modellünket, amelyben minden csúcs pontosan két hiperélhez tartozik. A háztartások mérete legyen H , a munkahelyek mérete pedig M . Ekkor egy csúcs fertőző szomszédainak átlagos száma a háztartásban $\frac{H-1}{N}I$, a munkahelyen pedig $\frac{M-1}{N}I$. Jelöljük k -val a hipergráfban található összes fertőző csúcs számát. Így a (2.6) egyenlet második összegzését közelíteni

tudjuk az alábbival:

$$f\left(\frac{H-1}{N}k\right) + f\left(\frac{M-1}{N}k\right),$$

amely egy egészséges típusú csúcsra vonatkozik. Így

$$N_{SI}^f(S_j^k) \approx (N-k) \left[f\left(\frac{H-1}{N}k\right) + f\left(\frac{M-1}{N}k\right) \right],$$

hiszen az S_j^k állapotban $N-k$ darab egészséges csúcs van. Ha ezt behelyettesítjük a (2.10) egyenletbe, akkor az alábbi közelítést kapjuk:

$$\begin{aligned} [SI](t) &\approx \sum_{k=0}^N (N-k) \left[f\left(\frac{H-1}{N}k\right) + f\left(\frac{M-1}{N}k\right) \right] \sum_{j=1}^{c_k} X_j^k(t) \\ &= E \left[(N - I(t)) \left[f\left(\frac{H-1}{N}I(t)\right) + f\left(\frac{M-1}{N}I(t)\right) \right] \right] \\ &\approx (N - [I](t)) \left[f\left(\frac{H-1}{N}[I](t)\right) + f\left(\frac{M-1}{N}[I](t)\right) \right]. \end{aligned}$$

Ennélfogva (2.12) közelíthető az alábbiak szerint:

$$(3.1) \quad [\dot{I}] \approx \tau(N - [I]) \left[f\left(\frac{H-1}{N}[I]\right) + f\left(\frac{M-1}{N}[I]\right) \right] - \gamma[I].$$

Végezetül az $[I]$ várható értéket közelítjük az \tilde{I} mennyiséggel, amelyre a következő differenciálegyenlet írható fel:

$$(3.2) \quad \dot{\tilde{I}} = \tau(N - \tilde{I}) \left[f\left(\frac{H-1}{N}\tilde{I}\right) + f\left(\frac{M-1}{N}\tilde{I}\right) \right] - \gamma\tilde{I}.$$

Ezzel a (2.12) egyenletet két állapotú hipergráfok esetén lezártuk. A (3.2) egyenlet megoldását és a szimulációs eredményeket két különböző hipergráf esetén hasonlítottuk össze, amely a 3.5. ábrán látható. Megfigyelhető, hogy a közelítő egyenlet homogénebb hipergráf esetén jó egyezést mutat a szimulációval.

Most tekintsük a reguláris véletlen hipergráf modellünket, amelyben minden csúcs d hiperélhez tartozik és minden hiperél e csúcsot tartalmaz. Ekkor egy hiperélben lévő egészséges csúcs fertőző szomszédainak átlagos száma $\frac{1}{N}(e-1)$. Jelöljük k -val a hipergráfban található összes fertőző csúcs számát. Így a (2.6) egyenlet második összegzése $df\left(\frac{e-1}{N}k\right)$ értékkel közelíthető, amely egy egészséges típusú csúcsra vonatkozik. Így

$$N_{SI}^f(S_j^k) \approx (N-k)df\left(\frac{e-1}{N}k\right),$$

hiszen az S_j^k állapotban $N-k$ darab egészséges csúcs van, és minden csúcs d hiperélhez tartozik. Ha ezt behelyettesítjük a (2.10) egyenletbe, akkor az alábbi közelítést

kapjuk:

$$\begin{aligned} [SI](t) &\approx \sum_{k=0}^N (N-k) f\left(\frac{e-1}{N}k\right) d \sum_{j=1}^{c_k} X_j^k(t) \\ &= E\left[(N-I(t))f\left(\frac{e-1}{N}I(t)\right)d\right] \approx (N-[I](t))f\left(\frac{e-1}{N}[I](t)\right)d. \end{aligned}$$

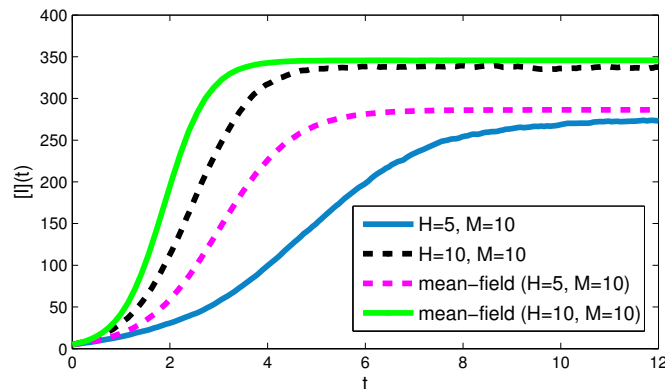
Ennélfogva, reguláris véletlen hipergráfok esetén a (2.12) egyenletet közelíteni tudjuk az alábbiak szerint:

$$(3.3) \quad [\dot{I}] \approx \tau(N-[I])df\left(\frac{e-1}{N}[I]\right) - \gamma[I].$$

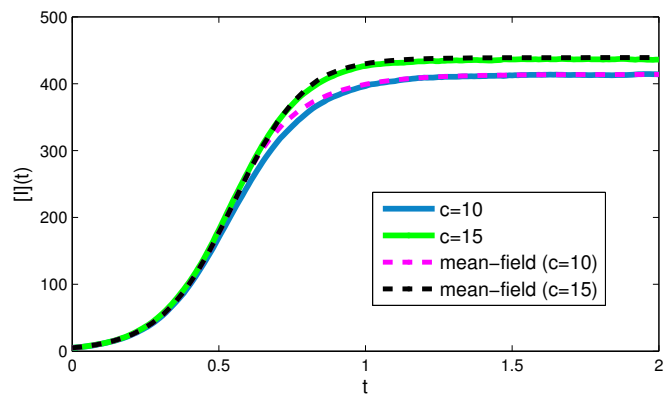
Végezetül jelölje \tilde{I} az $[I]$ közelítését, amelyre a következő differenciálegyenletet írhatjuk fel:

$$(3.4) \quad \dot{\tilde{I}} = \tau(N-\tilde{I})df\left(\frac{e-1}{N}\tilde{I}\right) - \gamma\tilde{I}.$$

Ezzel a (2.12) egyenletet reguláris véletlen hipergráfok esetén lezártuk. A (3.4) egyenlet megoldását és a szimulációs eredményeket két különböző c küszöbérték esetén hasonlítottuk össze, amely a 3.6. ábrán látható. Megfigyelhető, hogy reguláris véletlen hipergráfok esetén a közelítő egyenlet a korai szakaszban pontos, majd gyorsabban emelkedik, végül az állandósult állapotban ismét közel kerül a szimulációs eredményekhez.



3.5. ábra. Fertőző csúcsok számának időbeli változása $N = 500$ csúcsú hipergráf, H fős háztartások és M fős munkahelyek esetén $\gamma = 1$ gyógyulási és $\tau = 0.18$ fertőzési rátával. Az ábrán a (3.2) egyenlet megoldása is látható. A (2.2) függvénybeli küszöbérték $c = 7$. Az ismétlések száma 200, $\Delta t = 0.06$.



3.6. ábra. Fertőző csúcsok számának időbeli változása konfigurációs modell alapján generált $N = 500$ csúcsú és $M = 400$ élű reguláris hipergráf esetén, ahol minden csúcs pontosan $d = 16$ hiperélhez tartozik és minden hiperél mérete $e = 20$. A szimulációs görbék $c = 10$ és $c = 15$ küszöbértékekhez tartoznak, amelyek a (3.4) egyenlet megoldását is mutatják. A gyógyulási ráta $\gamma = 1$, míg a fertőzési ráta $\tau = 0.03$. Az ismétlések száma 200, $\Delta t = 0.01$.

4. fejezet

Valós adatok

A 3.3. alfejezetben az általunk létrehozott különböző struktúrák és paraméterek mellett vizsgáltuk a terjedési folyamat viselkedését. Ebben a fejezetben két állapotú hipergráfokkal foglalkozunk és célunk olyan paraméterek meghatározása, amelyek mellett a modellünk minél pontosabban illeszkedik a valós adatokhoz. Ehhez a Pittsburghi Egyetem oldalán [19] található, regisztrációt követően bárki számára elérhető járványterjedési adatokat használtuk, amely 58 féle betegség, köztük az SIS típusú gonorrhoea betegszámának alakulását tartalmazza az Amerikai Egyesült Államokban, hetekre lebontva 1888 és 2014 között. Az adatok egy részét a 4.1. ábra

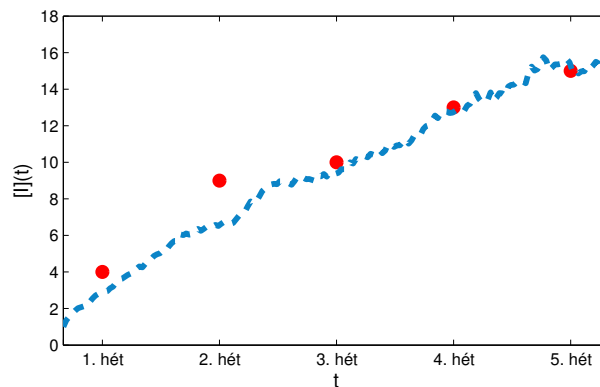
number	location	country	from date	to date	reporting period	epi week	disease
682	NEW YORK CITY	US	1979.03.04	1979.03.10	7	197910	GONORRHEA
390	OHIO	US	1979.03.04	1979.03.10	7	197910	GONORRHEA
2022	TEXAS	US	1979.03.04	1979.03.10	7	197910	GONORRHEA
51	IDAHO	US	1979.03.04	1979.03.10	7	197910	GONORRHEA
173	WASHINGTON	US	1979.03.04	1979.03.10	7	197910	GONORRHEA
3599	CALIFORNIA	US	1979.03.04	1979.03.10	7	197910	GONORRHEA
33	HAWAII	US	1979.03.04	1979.03.10	7	197910	GONORRHEA

4.1. ábra. 1979. március 4-10. közötti gonorrhoea betegszámok államokra lebontva

szemlélteti. Fontos megjegyezni, hogy a gonorrhoea egy nemi úton terjedő bakteriális fertőzés, amely nők esetében (ritkán férfiaknál is) tünetmentes lehet, ennél fogva az adatok pontossága megkérdőjelezhető, ráadásul nincs adatunk azokról a fertőző betegekről, akik a tünetek ellenére sem fordultak orvoshoz. Maga a betegség hozzávetőlegesen a férfiak 0.6, a nők 0.8 százalékát érinti [20], így a [21] weboldalon megtalálható népességszám adatokat ennek figyelembevételével kellett csökkentenünk az N meghatározásához, amely a korábbiaknak megfelelően egy adott társadalomban a fertőzhető: S és a fertőzők: I együttes számát jelenti. A fertőzés terjedését tekintve a 3.1.1. alfejezetben definiált két állapotú hipergráfban az egyed háztartásának megfelelő hiperélre a továbbiakban úgy gondolunk, mint az egyed baráti körére.

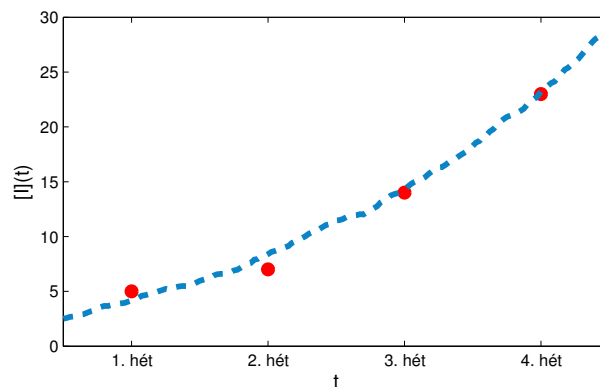
4.1. A küszöbérték meghatározása

Egy adott állam és adott időszak kiválasztását követően rögzítettük a H és M értékeket, majd olyan fertőzési rátát és olyan (2.2) függvénybeli küszöbértéket kerestünk, amellyel a modellünk jól közelíti a tényleges folyamatot. Az alábbi ábrákon a piros pontok mindig a valós adatokat, a szaggatott kék vonal a 3.2. alfejezetben tárgyalt szimulációs eredményt jelöli. A gyógyulási rátát mindenütt $\gamma = 1$ értékkel rögzítettük. A 4.2. ábra a fertőzők számának időbeli változását mutatja, $H = 5$ fős háztartások és $M = 10$ fős munkahelyek esetén, $\tau = 0.131$ fertőzési rátával és



4.2. ábra. Wyoming, 1972. október 22 - november 25.

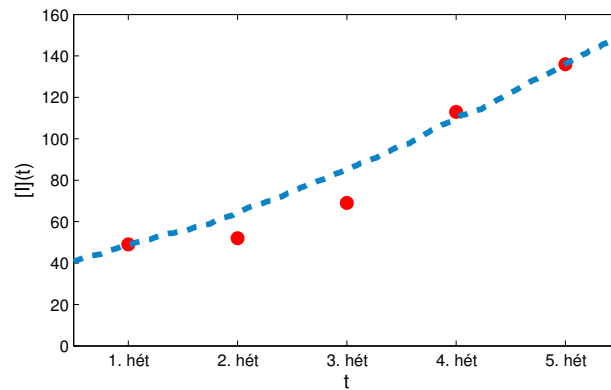
$c = 1$ értékkel. A szimulációk futtatása során arra jutottunk, hogy a szóban forgó c küszöbérték gonorrhoea esetén alacsony, ugyanis $c = 2$ és $c = 3$ értékek mellett a modellünk még jó közelítéseket adott a járvány kezdeti növekedésére, viszont $c = 4$ és afeletti értékek esetén már jóval több fertőző beteget mutatott a valós adatokhoz képest. A 4.3. ábra egy ilyen kezdeti növekedést szemléltet $H = 4$ fős háztartások és $M = 15$ fős munkahelyek, valamint $\tau = 0.13$ fertőzési ráta és $c = 3$ érték esetén.



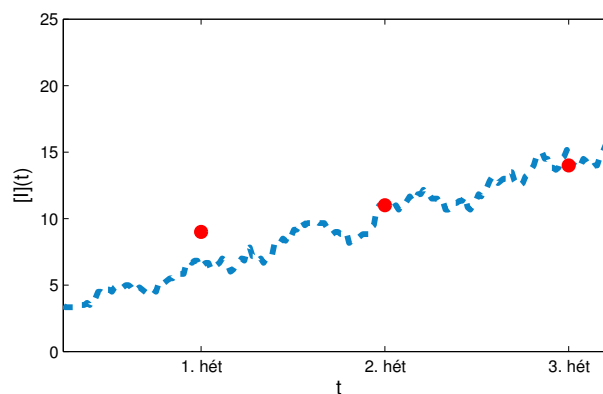
4.3. ábra. Vermont, 1972. május 7 - június 3.

4.2. A struktúra paramétereinek hatása

Ebben az esetben két adott állam és egy közös kezdetű időszak kiválasztását követően rögzítettük a fertőzési rátát, illetve a (2.2) függvénybeli küszöbértéket, majd olyan H , M értékpárokat kerestünk, amelyekkel a modellünk jól közelíti a tényleges folyamatokat. A 4.4. és a 4.5. ábrákon látható modelljeinkben Nevadánál $H = 5$ és $M = 10$, míg Vermontnál $H = M = 8$ fős háztartások és munkahelyek szerepelnek. A fertőzési és gyógyulási ráta, valamint a küszöbérték mindkét esetben azonos: $\tau = 0.117$, $\gamma = 1$ és $c = 2$. A betegszámkülönbségekből jól látható, hogy a terjedési folyamat Nevada esetében gyorsabb, ugyanis az első héten 40, a harmadik héten már 55 volt a differencia Vermont betegszámához képest. Tekintve, hogy a fertőzési ráta és a küszöbérték mindkét esetben azonos, ez a heterogénebb hiperélekek köszönhető.



4.4. ábra. Nevada, 1977. február 6 - március 12.



4.5. ábra. Vermont, 1977. február 6 - február 26.

Látható, hogy a modellünk viszonylag jól közelíti a tényleges terjedési folyamatot, valamint alátámasztja a 3.3. alfejezetben látottakat. Ugyanakkor fontos megjegyezni, hogy a valós adatok [19] pontatlannak bizonyultak, így nem kaphatunk teljes képet a járvány terjedési folyamatáról. Erre megoldás lehet pontosabb adatsor használata. További lehetőség a modell általánosítása SIR típusú járványterjedésre, ugyanis az ilyen típusú járványok esetén komplexebb és megbízhatóbb adatok találhatóak.

Összefoglalás

A szakdolgozatban a járványterjedés matematikai modellezésével foglalkoztunk. A hagyományos értelemben vett gráfokon való vizsgálat helyett a folyamatot hipergráfokon tekintettük. Ezt a háztartás típusú hálózatok adták motivációul. A hipergráfok segítségével számításba vehettük, hogy egyrészt a populáció kisebb közösségekből áll össze, másrészt azt, hogy ezeken a közösségeken belül a fertőzés nem feltétlen arányos a fertőzött szomszédok számával. Ezt követően levezettük az alapegyenletet, illetve ismertettük a várható értékekre vonatkozó differenciálegyenleteket. A hálózatokra használt szimulációs algoritmust kiterjesztettük hipergráfokra, majd a szimulációt három különböző típusú hipergráfon vizsgáltuk: egyrészt a két állapotú hipergráfon, ahol minden egyén véletlenszerűen lett szétosztva egy háztartásba és egy munkahelyre, másrészt a konfigurációs véletlen hipergráfon, illetve egy olyan hipergráfon, ahol az élek egy Barabási–Albert véletlen gráf klikkjeiből lettek létrehozva. A különböző struktúrák és a modell paramétereinek hatásait a szimulációs eredmények alapján vizsgáltuk. Ezek után a várható értékekre vonatkozó egyenletek lezárásával foglalkoztunk, amelyeket összehasonlítottuk a szimulációs eredményekkel. Végezetül pedig valós adatok alapján, két állapotú hipergráfok esetén vizsgáltuk a modell paramétereit.

Irodalomjegyzék

- [1] A. Barrat, M. Barthelemy, A. Vespignani: *Dynamical processes on complex networks*, Cambridge University Press Cambridge, 2008.
- [2] Á. Bodó, Gy. Y. Katona, P. L. Simon: SIS Epidemic Propagation on Hypergraphs, *Bulletin of Mathematical Biology* **78** (2016), 713–735.
- [3] W. O. Kermack, A. G. McKendrick: A Contribution to the Mathematical Theory of Epidemics, *Proceedings of the Royal Society of London A* **115** (1927), 700–721.
- [4] V. Capasso: *Mathematical Structures of Epidemic Systems*, Springer, 1993.
- [5] A. Rachah, Delfim F. M. Torres: Predicting and controlling the Ebola infection, *Math. Meth. Appl. Sci.* (2016), DOI: 10.1002/mma.3841.
- [6] N. H. Shah, J. Gupta: SEIR Model and Simulation for Vector Borne Diseases, *Applied Mathematics* (2013) **4**, 13–17.
- [7] V. Isham, S. Harden, M. Nekovee: Stochastic epidemics and rumours on finite random networks, *Physica A* **389** (2010), 561–576.
- [8] A. V. Goltsev, F. V. de Abreu, S. N. Dorogovtsev, J. F. F. Mendes: Stochastic cellular automata model of neural networks, *Phys. Rev. E* **81** (2010), 061921.
- [9] P. Erdős, A. Rényi: On Random Graphs I., *Publ. Math. Debrecen* **6** (1959), 290–297.
- [10] E. N. Gilbert: Random graphs, *Ann. Math. Stat.* **30** (1959), 1141–1144.
- [11] A-L. Barabási, R. Albert: Emergence of scaling in random networks, *Science* **288** (5439) (1999), 509–512.
- [12] C. Stegehuis, R. van der Hofstad, J. S. H. van Leeuwen: Epidemic spreading on complex networks with community structures, *Scientific Reports* **6** (2016), 29748.

- [13] M. Benayoun, J. D. Cowan, W. van Drongelen, E. Wallace: Avalanches in a stochastic model of spiking neurons, *PLoS Computational Biology* **6(7)** (2010), e1000846
- [14] M. Taylor, P. L. Simon, D. M. Green, T. House, I. Z. Kiss: From Markovian to pairwise epidemic models and the performance of moment closure approximations, *Journal of Mathematical Biology* **64** (2012), 1021–1042.
- [15] <http://barabasi.com/networksciencebook> (2017. május)
- [16] [https://en.wikipedia.org/wiki/Plague_\(disease\)](https://en.wikipedia.org/wiki/Plague_(disease)) (2017. május)
- [17] <https://virus.stanford.edu/uda> (2017. május)
- [18] <http://www.cfinder.org> (2017. május)
- [19] <http://www.tycho.pitt.edu> (2017. május)
- [20] <https://en.wikipedia.org/wiki/Gonorrhoea> (2017. május)
- [21] <https://www.census.gov> (2017. május)