

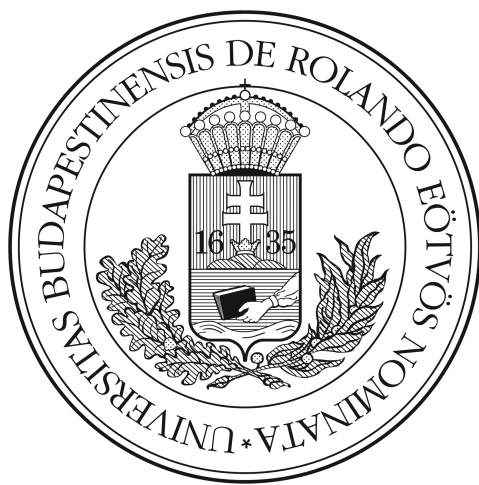
BSc SZAKDOLGOZAT

Járványterjedés vizsgálata gráfokon

Kovács-Deák Dániel

ELTE TTK Fizika BSc

Informatikus fizikus specializáció



Témavezető:

Stéger József

ELTE Komplex Rendszerek Fizikája Tanszék

„Negyvenkettő - mondta Bölcs Elme végtelen méltósággal és hidegvérrel.”

— Douglas Adams

Tartalomjegyzék

Tartalomjegyzék	3
1. Bevezetés	4
2. A dolgozat tudományterülete	5
2.1. „Kicsi a világ!”	6
2.2. Klaszterezettség	7
2.3. Skálafüggetlenség	7
2.4. Napjaink kutatási területei	10
3. A dolgozatban tárgyalt modell	10
3.1. A gráfot felépítő program	11
3.2. A SIS-modell	12
4. Futási idő	13
4.1. A gráfot felépítő program	13
4.2. A SIS-modell	14
5. A szimulációs eredmények értelmezése	15
5.1. Az állandósult állapot kezdetének detektálása	18
6. A paraméter tér vizsgálata	21
6.1. A hálózat éleinek hatása a kihalás/felfutás arányra	25
6.2. A határvonal görbájének vizsgálata	27
7. Összegzés	30
8. Továbbfejlesztési lehetőségek	33
Irodalomjegyzék	35

1. Bevezetés

A hálózatok átszövik a világunkat. Mind gazdasági, politikai és – nem utolsó sorban – tudományos szempontból is egyre jobban vizsgált a hálózatkutatás területe. A dolgozatom célja bevezetni az olvasót a hálózatkutatás egy speciálisabb területére a skálafüggetlen hálózatok világába. A skálafüggetlen hálózatok felépítése után járványok terjedésének egy modelljét az ún. SIS modellt mutatom be. A terjedési folyamatok vizsgálata számítógépes szimuláción keresztül zajlik, python nyelven írtam egy programot, szimuláció alatt ennek a programnak a futtatását kell érteni. A szimulátor két fő elemből áll: az egyik a hálózatot felépítő kód, a másik maga a SIS modell. A felépített hálózaton a SIS modell különböző paramétereinek a terjedési folyamatra gyakorolt hatásait mutatom be.

A hálózatok alapvetően valamilyen egységek – az emberek, a sejtek, a számítógépek – és közöttük fellelhető kapcsolatokról állnak. Hétköznapi példát nézve hálózat az osztályterem, ahol az egyedek a diákok, és kapcsolatok a barátságok. Hálózatot alkotnak az Interneten a honlapok, melyeket a linkek kötnek össze, ezekre rákattintva egy másik honlapra juthatunk. A hálózatokat a matematika gráfokkal modellezi. A hálózat a vizsgált dolog, a modell mögötte pedig a gráf. A kettő megfeleltethető egymásnak. A továbbiakban hálózatot és a gráfot szinonimaként fogom kezelni. Az egyedek a gráf pontjai / csúcsai, a csúcsok közötti kapcsolatok pedig az élek.

A gráfok lehetnek irányítottak vagy irányítatlanok. Ha irányítottak, akkor az éleknek az egyik irányban más tulajdonságai vannak, mint a másik irányban, például az egyik irányba terjedhet információ, a másikba nem. A dolgozatban irányítatlan gráfokat vizsgálunk, a vizsgált járványterjedési példában nincs irányhoz kötött tulajdonsága az éleknek.

Egy adott csúcsból kiinduló élék számát fokszámnak nevezzük. Ha a csúcsokat fokszámuk szerint csoportosítjuk és megszámloljuk, hogy egy-egy csoportba hány csúcs kerül, akkor megkapjuk a gráf fokszámeloszlását. A dolgozatban vizsgált hálózatok ún. skálafüggetlen hálózatok. Ez azt jelenti, hogy a fokszámeloszlásuk hatványfüggvény szerinti eloszlás. Ez az eloszlás a hálózatra nézve azt jelenti, hogy a legtöbb csúcs alacsony fokszámú, azonban léteznek kiemelkedően magas fokszámú csúcsok is. Az internetes példában a Google egy ilyen kiemelkedően magas fokszámú csúcs, míg az én saját honlapom a rámutató hivatkozások számát tekintve átlagos.

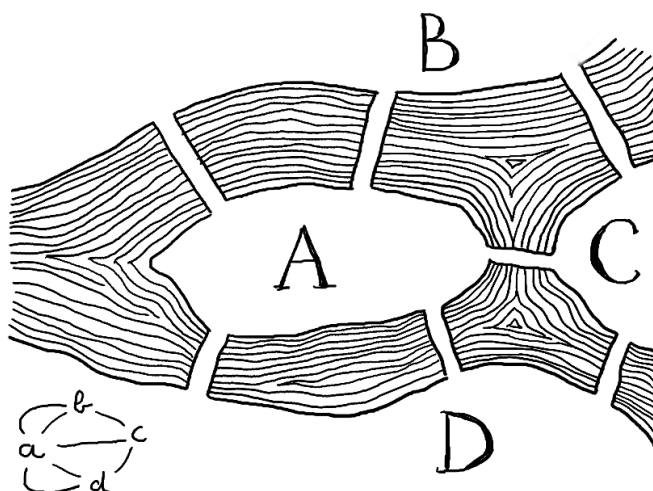
A világban sok olyan dologgal találkozunk, ami egy hálózaton terjed: betegségek, számítógépes vírusok, gondolatok, nézetek, stb. Mivel ezek a terjedési folyamatok mindig egy hálózaton belül terjednek (emberek, számítógépek, stb. hálózatán), ezért célszerű a gráfokra alapozva modellezni őket. A SIS-modell ilyen terjedési folyamatokat modellez gráfokon. A modell a fenti gráfokat azzal egészíti ki, hogy a gráf minden csúcsához rendel egy S vagy I értéket, ami azt jelzi, hogy az adott csúcs meg van-e fertőzve. Az éppen fertőzött csúcsok az éleiken keresztül tudnak fertőzni, a csúcsok meg tudnak gyógyulni és újra meg is fertőzhetőek. Ebben a modellben nem alakul ki a betegséggel szemben védelem. Az immunitást bonyolultabb, több állapotot

megkülönböztető modellek tárgyalják, mint például a SIR-modell.

2. A dolgozat tudományterülete

Ebben a fejezetben bemutatom a dolgozatban vizsgált skálafüggetlen hálózathoz vezető út néhány tudománytörténeti mérföldkövét. Olyan tudósokról lesz szó, akiknek a munkái alapvető fontosságúak a bonyolult, nagy méretű komplex hálózatok vizsgálatához. Láthatjuk majd, hogy a hálózatok kutatásában fontos szerep jutott magyar származású tudósoknak is.

Az első elhíresült probléma, amely bevezette a hálózatok gráfelméleti alapjait a Königsbergi hidak problémája volt. Az 1. ábrán illusztrált feladvány arra keresi a választ, hogy megszakítás nélkül bejárható-e az összes híd, mindegyiken pontosan egyszer áthaladva. A problémát Euler oldotta meg, megoldását 1735-ben a Szentpétervári Akadémián mutatta be [1]. Euler ebben a megoldásában vezette be a pontok és élek fogalmát.



1. ábra. A Königsbergi hidak problémája. Euler a bal alsó sarokban látható gráffal modellezte a problémát. Belátta, hogy zárt vonallal, amelynek a kezdeti és végpontja ugyanaz, csak akkor rajzolható meg a gráf, ha mindegyik pont fokszáma (az adott pontba befutó élek száma) páros. Nem zárt vonallal (ha nem követeljük meg, hogy a kezdő pont egyben a végpont legyen) pedig csak akkor, ha pontosan 2 db páratlan fokszámú csúcs van. Mivel a hidak esetében mind a négy csúcs páratlan fokszámú (5, 3, 3, 3) ezért sem nyílt sem zárt vonallal nem rajzolható meg a gráf, következésképpen nem járható be az összes híd mindegyiken pontosan egyszer áthaladva.

A XX. században több magyar kötődésű eredmény született a gráfelméletben. König Dénes a szerzője az első gráfelméleti könyvnek [2]. Erdős Pál és Rényi Alfréd vezette be 1959-ben a véletlen gráf fogalmát. A véletlen gráf N csúcsból áll és bármelyik két csúcs között ugyanazon

p valószínűséggel húzunk be élt. Az élek számának várható értéke $pN(N-1)/2$, az elhelyezkedésük pedig véletlenszerű.

A nagy, összetett hálózatok vizsgálata megmutatott három fontos tulajdonságot, melyeket komplex hálózatokkal folytatott vizsgálódások során – a konkrét problémától függően – figyelembe kell venni. Ezek a kisvilág-tulajdonság, a klaszterezettség és a skálafüggetlenség.

2.1. „Kicsi a világ!”

A komplex hálózatoknak – a sok élből és csúcsból álló nagy, összetett hálózatoknak – egy fontos és érdekes tulajdonsága a kisvilág-tulajdonság. Sokan megtapasztaljuk ezt az emberi társadalom ismeretségi hálózatában, amikor egy újonnan megismert emberrel közös ismerőstök fedezünk fel. Egy komplex hálózatban jellemzően két véletlenszerűen megválasztott pont között meglepően rövid út vezet.¹ Kisvilág-tulajdonságúnak nevezzük egy hálózatot, ha a csúcsok átlagos távolsága nagyon kicsi a rendszermérethez képest. Magyar kötődés, hogy Karinthy Frigyes Láncszemek című novellájában 1929-ben vetette fel annak a gondolatát, hogy a Föld bármely két lakosa között közvetlen ismeretségi kapcsolatokon keresztül el lehet jutni pusztán 5 emberen keresztül. Az elmélettel, amely később a hatlépésnyi távolság néven vált ismertté, Stanley Milgram a szociálpszichológia híres kutatója foglalkozott [3]. Kísérletében Milgram azt próbálta megmérni, hogy az Egyesült Államokban mekkora két véletlenszerűen választott személy átlagos távolsága. A kísérletben kansasi és nebraskai lakosokat kért meg, hogy egy levelező lapot juttassanak el egy bostoni célszemélynek személyes ismeretségek láncán keresztül. Munkája során arra kérte a résztvevőket, hogy amennyiben személyesen nem ismerik a célszemélyt, úgy küldjék el a küldeményt az instrukciókkal együtt egy olyan ismerősüknek, aki valószínűbb, hogy ismeri a célszemélyt, végül az összes elindított küldemény 21%-a érkezett meg szerencsésen [4]. Ugyanakkor a beérkező levelező lapok átlagosan 5-6 lépés alatt értek célba.

Nemcsak az emberi kapcsolatok hálózata kisvilág-tulajdonságú, hanem számos más komplex hálózat is, köztük az Erdős-Rényi-modell is.

A terjedési folyamatok vizsgálatánál azért fontos a kisvilág-tulajdonság, mert ez a tulajdonság biztosítja, hogy a járvány (gondolat, vírus, stb.) könnyen el tudjon jutni a hálózat bármely részébe. Mivel a valóságban a vártnál könnyebb kapcsolatot találni két távolinak ítélt pont között egy hálózatban, ezért jogos elvárás a hálózat mögötti modelltől – a gráftól – hogy szintén bírjon ilyen tulajdonsággal.

Az Erdős-Rényi-modell hosszú időn át a legelfogadottabb modellje volt a komplex hálóz-

¹ A hálózatban az út hossza alatt azt értjük, hogy legkevesebb hány csúcsot kell érinteniünk ahhoz, hogy elérjünk a kiinduló pontból a célba.

toknak. A komplex hálózatok vizsgálatai a kisvilág-tulajdonságon túl két további fontos jellemzőt fedeztek fel: a klaszterezettség és a skálafüggetlenség. A számítógépek elterjedésével egyre könnyebbé vált az adatok feldolgozása, egyre több hálózatban fedezik fel ezeket a tulajdonságokat.

2.2. Klaszterezettség

Sok vizsgált hálózat klaszterezett. Ez azt jelenti, hogy egy csúcs szomszédjai közül sokan egymásnak is szomszédai. Laza megfogalmazásban ez a hálózatnak egy fajta „klikkesedése”. Tekintheünk a társadalomra is, mint egy nagy hálózat, ahol baráti társaságok kis csoportjai vannak, akik mind ismerik egymást. Ebben a hálózatban a véletlenszerű gráfnál nagyobb várható értékkel találunk kapcsolatokat egy-egy klikken belül. Duncan Watts és Steven Strogatz 1998-ban a *Nature*-ben megjelent cikke ad egy olyan modellt, ami ezt a jelenséget is képes reprodukálni, miközben a kisvilág-tulajdonságot is megtartja [5].

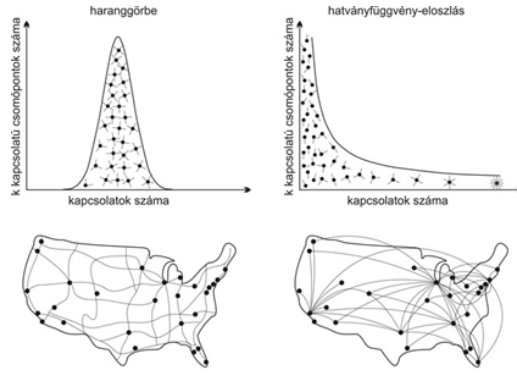
Habár a klaszterezettséget ebben a dolgozatban vizsgált modell nem biztosítja, ez a tulajdonság általában a komplex hálózatoknak egy fontos jellemzője.

2.3. Skálafüggetlenség

A fokszámeloszlás megadja, hogy adott fokszámú csúcsból hány darab van a hálózatban. Az Erdős-Rényi-modell fokszámeloszlása Poisson-eloszlást követ, amelynek alakja a haranggörbéhez hasonlít. Egy ilyen hálózatban a legtöbb csúcsnak átlagos a fokszáma, kiugróan sok él bekötéssel jellemezhető csúcsok nincsenek. Ezzel szemben a nagy komplex hálózatok, mint például a tudományos cikkek társszerzőségi hálói (különösen a nagyobb fokszámú tartományon) hatványszerűen lecsengő eloszlást mutatnak [6, 7]. Az ilyen hálózatokat skálafüggetlen hálózatoknak nevezzük. A kétféle fokszámeloszlást hasonlítja össze a 2. ábra. Barabási Albert-László és Albert Réka publikált először olyan algoritmust, amely skálafüggetlen véletlen gráfot generál [8]. Az általuk javasolt módszerben a gráf egy növekedési folyamat eredményeként születik.

2.3.1. A hatványfüggvény szerinti fokszámeloszlás levezetése

A Barabási-Albert-modellben lépésenként adunk csúcsokat a hálózathoz. Minden lépésben 1 db csúcsot kötünk be m db éllel, ezeket az éleket nevezem bekötő éleknek. Az m db él ún. preferenciális kapcsolódással kötődik be. A gráf minden csúcsához tartozik egy valószínűség, ami az adott csúcs fokszámával arányos. A bekötő él egy csúcshoz az adott csúcshoz tartozó valószínűséggel kötődik be. A következő levezetésben az $m = 1$ esetről lesz szó, amikor minden új csúcs 1 éllel kötődik be a hálózatba. Jelölje N a csúcsok számát a gráfban, L az élek számát,



2. ábra. A városok az azokat összekötő közutak hálózatában a fokszámeloszlás haranggörbét követ, a legtöbb csúcs átlagos fokszáma, és nincsenek nagyon kiugró fokszámaú pontok. Ezzel szemben a légitözkedés hálózatában a fokszámeloszlás már hatványfüggvény szerinti. Kevés kiugróan nagy fokszámaú csúcs (pl. nemzetközi repölőterek) és sok kis fokszámaú csúcs (helyi repterek) alkotják a hálózatot. [9]

N_k pedig a k fokszámaú csúcsok számát. Minden új csúcs egy éllel kötődik be a már bent lévő csúcsok valamelyikéhez, a bekötődés valószínűsége általános esetben: $\frac{k^\alpha}{\sum_{l=1}^N l^\alpha N_l}$, ahol k az adott csúcs fokszáma, a nevezőben pedig az összfokszám szerepel. A dolgozatban vizsgált modellben és a levezetés során is $\alpha = 1$. Legyen $A = \sum_{l=1}^N l N_l$, az összfokszám. A folyamat Master-egyenletei:

$$N_k(N+1) = N_k(N) - \frac{k}{A} N_k(N) + \frac{k-1}{A} N_{k-1}(N) \quad (1)$$

$$N_1(N+1) = N_1(N) - \frac{1}{A} N_1(N) + 1 \quad (2)$$

A 2. egyenlet értelmezése: az 1 fokszámaú csúcsok száma az $N+1$. időlépésben egyenlő az 1 fokszámaú csúcsok száma a megelőző időlépésben plusz 1, hiszen az új, frissen bekötött csúcs fokszáma mindig 1, viszont ha ez pont egy 1 fokszámaú csúcshoz kötődik be, akkor csökkenteni kell 1-gyel. Mivel $1/A$ eséllyel kötődik be az új él az N_1 db 1 fokszámaú csúcs valamelyikéhez, a kettő szorzata lesz a várható értéke azon csúcsoknak, amelyek eddig 1 fokszámaúak voltak, de mert az új csúcs hozzájuk kötődött be, ezért 2 fokszámaúvá váltak. És mivel minden lépésben egy új csúcs 1 éllel bekötődik, ezért az 1 fokszámaúak számát ez minden lépésben 1-el biztosan növeli. Az 1. egyenlet hasonlóan értelmezhető.

Ha a kezdeti gráfhoz képest sok élt hozzáadtunk már a gráfunkhoz, akkor - mivel minden lépésben 1 élt adunk hozzá a hálózathoz, ezért minden lépésben az élek össz száma 1-gyel nő, jó közelítéssel $L \approx N$. Mivel $A = \sum_{l=1}^N l N_l = 2L \approx 2N$ A fenti egyenleteket átírva differencia egyenletekké:

$$\Delta N_k = -\frac{k}{A} N_k + \frac{k-1}{A} N_{k-1}, \quad (3)$$

$$\Delta N_1 = -\frac{1}{A}N_1 + 1. \quad (4)$$

Behelyettesítve az $A \approx 2N$ közelítést, és bevezetve a $P_k = \frac{N_k}{N}$ jelölést (P a fokszámoszlás):

$$\frac{dN_k}{dN} = -\frac{k}{2}P_k + \frac{k-1}{2}P_{k-1}, \quad (5)$$

$$\frac{dN_1}{dN} = -\frac{1}{2}P_1 + 1. \quad (6)$$

Az egyenletek bal oldalát írjuk át a következő képpen: $\frac{dN_k}{dN} = \frac{d(NP_k)}{dN} = P_k + N\frac{dP_k}{dN}$

$$P_k + N\frac{dP_k}{dN} = -\frac{k}{2}P_k + \frac{k-1}{2}P_{k-1}, \quad (7)$$

$$P_1 + N\frac{dP_1}{dN} = -\frac{1}{2}P_1 + 1. \quad (8)$$

Átrendezve:

$$N\frac{dP_k}{dN} = -\frac{k+2}{2}P_k + \frac{k-1}{2}P_{k-1}, \quad (9)$$

$$N\frac{dP_1}{dN} = -\frac{3}{2}P_1 + 1. \quad (10)$$

A 10. egyenlet stacionárius esetben (amikor elértük az állandósult állapotot, és már nem változik P_1 értéke): $P_1^{stac} = \frac{2}{3}$, míg P_k :

$$P_k^{stac} = \frac{k-1}{k+2}P_{k-1}^{stac} \rightarrow P_k^{stac} = \frac{k-1}{k+2} \frac{k-2}{k+1} \cdot \dots \cdot P_1^{stac}$$

Ahol P_1^{stac} együtthatója átírható:

$$\frac{k-1}{k+2} \frac{k-2}{k+1} \dots = \frac{(k-1)(k-2) \cdot \dots \cdot 1}{(k+2)(k+1) \cdot \dots \cdot 4} = \frac{(k-1)!}{\frac{(k+2)!}{3!}} = \frac{3!}{(k+2)(k+1)k}$$

Így tehát stacionárius állapotban:

$$P_k^{stac} = \frac{4}{(k+2)(k+1)k} \sim \frac{4}{k^3}$$

A kifejezés logaritmusát véve $\ln(P_k^{stac}) \sim -3\ln(k)$, amit dupla logaritmikusan ábrázolva -3 meredekségű egyenest kapunk.

2.4. Napjaink kutatási területei

A klaszterezettséghez hasonlóan a skálafüggetlenség is egy tapasztalati tény. A vizsgált komplex hálózatok jellemzően skálafüggetlenek, ezért a modellezésnél ezt a tulajdonságot figyelembe kell venni. A terjedési folyamatok szempontjából a skálafüggetlenség a kisvilág-tulajdonsággal összefüggő, azzal rokon szereppel bír: biztosítja a rövid utat két tetszőleges csúc között. Egy terjedési folyamatban a nagy foksámú csúcsok hatása jelentősebb. Ha egy olyan ember lesz influenzás, aki sok időt tölt sok ember között, akkor ő várhatóan jóval több embert meg tud fertőzni, mint aki egymagában otthon tölti az idejét. Egy politikus, aki a sajtón keresztül nagyon sok emberhez el tudja juttatni az üzenetét (a sajtónak hála nagyon megnövekszik az üzenet célba juttatását biztosító élek száma) egy adott nézetről, véleményről több embert tud meggyőzni, mint az a hétköznapi átlagember, aki csak a jóval kevesebb számú személyes ismerősével tud kapcsolatba lépni.

Napjaink kutatásainak alapvető fontosságú alapköve a Barabási-Albert-modell és Watts-Strogatz-modell, a skálafüggetlenség és a klaszterezettség. A hálózattudomány legfontosabb fejezeteit közérthetően és részletesen taglalja Barabási Albert-László tankönyve [10].

A hálózatkutatás területe napjainkban igen népszerű, éves szinten több ezer publikáció születik a komplex hálózatokkal kapcsolatos témákban. A kutatások irányai is különbözőek, például hálózati csoportkeresés (hálózatklaszterezés), időfejlődő hálózatok, vírusterjedés, emberi kapcsolatok modellezése, csoportdinamika vagy biológiai hálózatok kutatása [11].

3. A dolgozatban tárgyalt modell

A gyakorlatban vizsgált komplex hálózatok jellemzően igen sok pontot és élt tartalmaznak. Gyakran egy-egy konkrét hálózat (mint például az emberi kapcsolatok) adatgyűjtése nehézkes. Éppen ezért a rendszert nem mikro szinten, hanem makro szinten, annak statisztikáit vizsgálva írják le. A dolgozatban tárgyalt általános tulajdonságok a konkrét speciális hálózatok vizsgálata esetén is érvényesek.

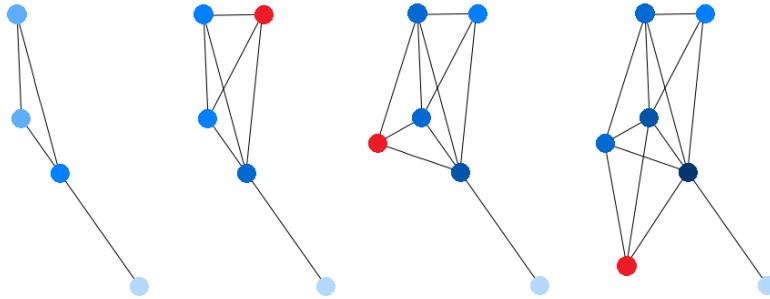
A munkám során egy programot írtam Python nyelven. Ezzel a programmal számítógépes szimulációt végeztem, mely a Barabási-Albert-modellben megismert algoritmussal felépít egy skálafüggetlen hálózatot, majd azon vírusok terjedését szimulálja a SIS-moddal. A szimulált folyamat megfigyelhető lehet akár betegségek, akár nézetek, vélemények, vagy éppen számítógépes vírusok terjedési folyamatának. Amennyiben ismerjük egy konkrét vírus terjedési és gyógyulási valószínűségét, úgy a felépített modellel megvizsgálhatjuk a terjedés a hálózat egészét jellemző tulajdonságait.

3.1. A gráfot felépítő program

A Barabási-Albert-modellben a gráf egy növekedési folyamat eredményeként jön létre. Kiindulunk egy tetszőleges gráfból, és minden lépésben egy csúcsot és vele új élteket adunk hozzá. Minden új csúcs azonos m darab éllel kapcsolódik be a már meglévő gráfba. Azonban az új élek bekötődése nem teljesen véletlenszerű: egy adott már meglévő csúcshoz annak fokszámával arányos valószínűséggel fog az új csúcs bekötődni. Minél nagyobb tehát a gráfban egy csúcs fokszáma, annál nagyobb eséllyel fog egy következő új csúcs is hozzákötődni. Ez a preferenciális kapcsolódás biztosítja a hálózat skálafüggetlenségét, a hatványfüggvény szerinti elszólást a fokszámeloszlást $\log\text{-}\log^2$ ábrán ábrázolva ellenőrizhetjük. Hatványfüggvény szerinti eloszlás esetén az eloszlás pontjai $\log\text{-}\log$ ábrán egyenesre kell illeszkedniük. Az algoritmus egy egyszerű kezdeti gráfot igényel, amire lépésenként ránőveszti a nagy skálafüggetlen gráfot.

A hálózat növesztő program két paramétert kér, a kiinduló gráfhoz házzáadandó csúcsok számát, és a bekötőélek számát (m). Fontos, hogy a kiinduló gráfban legyen annyi csúcs, ahány bekötőéllel (m) szeretnénk az új csúcsokat bekötni. A terjedési folyamatok szimulációihoz a 3. ábrán látható 4 csúcsú kiindulási gráfhoz $N = 10000$ db csúcs lett hozzáadva. A dolgozatban tárgyalt legtöbb szimuláció egy $m = 4$ bekötő élszámmal növesztett gráfon futott.

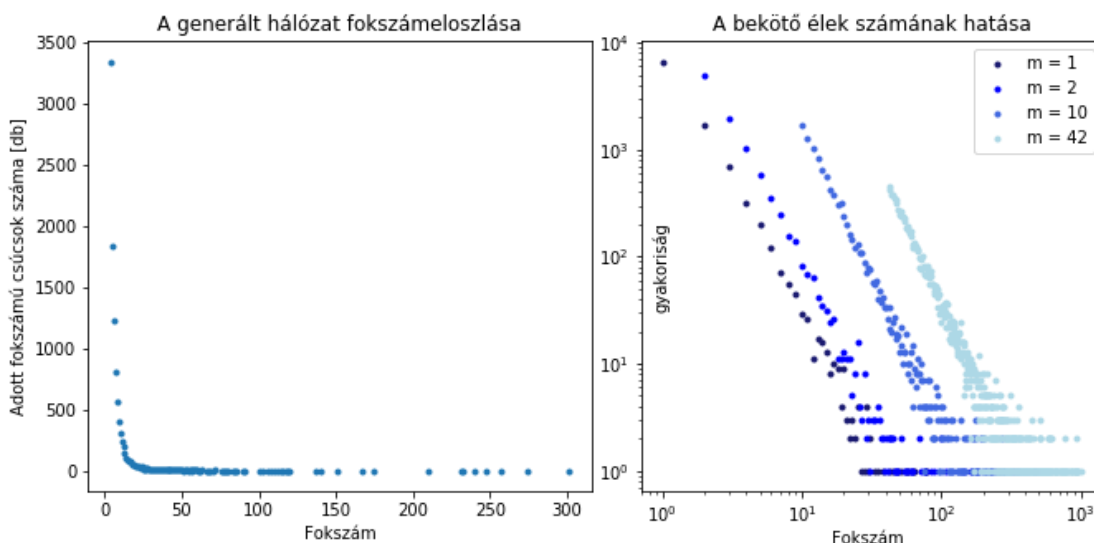
A preferenciális kapcsolódás miatt egy csúcs, aminek a többihez képest már az elején nagyobb a fokszáma várhatóan kiemelkedik a többi közül. Magasabb m érték választással ez kissé egyenletesebbé tehető, várhatóan több nagy fokszámú csúcs is kialakul.



3. ábra. A növesztési folyamat első három lépése, $m = 3$ él esetén. Pirossal a frissen bekötött új csúcs, a többi a fokszámmal arányosan egyre sötétebb kékre színezve.

Ha megvizsgáljuk a generált gráf fokszámeloszlását, akkor megkapjuk a várt hatványfüggvény szerinti eloszlást. A 4. ábrán láthatjuk a fokszámeloszlás alakját. Azt, hogy valóban hatványfüggvény szerinti ez az eloszlás, azt a $\log\text{-}\log$ ábrán megjelenő egyenes trend bizonyítja.

² Olyan ábrán, amelyen mindkét változónak (x és y) a logaritmusa szerepel



4. ábra. Az ábrán $N = 10000$ csúcsúra növesztett gráfok fokszámeloszlásai láthatóak. A bal oldalin látszik a hatványfüggvény görbéje. Ahogy az a 2.3.1-es rész levezetéséből kiderült, logaritmikus tengelyekkel ábrázolva a fokszámeloszlás negatív meredekségű egyenes lesz. A jobb oldali ábrán különböző bekötő élszámmal (m) növesztett gráfok fokszámeloszlása látszik. A bekötő élek számával egyre nagyobb negatív meredekségű egyenesekre illeszkednek a pontok.

3.2. A SIS-modell

A vírusok, betegségek, nézetek terjedésének modellezéséhez jó alapot jeltenek a gráfok. Ugyanis egy terjedés vizsgálatánál ugyanaz a két alapegység van, mint a gráfoknál: a csúcsok (emberek, egyedek, számítógépek, stb), amik meg tudnak fertőződni, és az őket összekötő élek, amik a fertőzés lehetséges csatornái. A modellben a csúcsok két állapotban lehetnek: S („susceptible”) fertőzhető és I („infected”) fertőzött, melyek egymásba alakulhatnak. Minden fertőzött meg tud gyógyulni és újra meg tud fertőződni. A modellben a fertőzés folyamata véletlenszerű. Az algoritmusban minden fertőzött egyed először is P_{gy} gyógyulási valószínűséggel meggyógyul, ha nem gyógyul meg, akkor minden egyes élén P_f fertőzési valószínűséggel megfertőzi az él túlsó végén lévő csúcsot, amennyiben az még nem fertőzött.

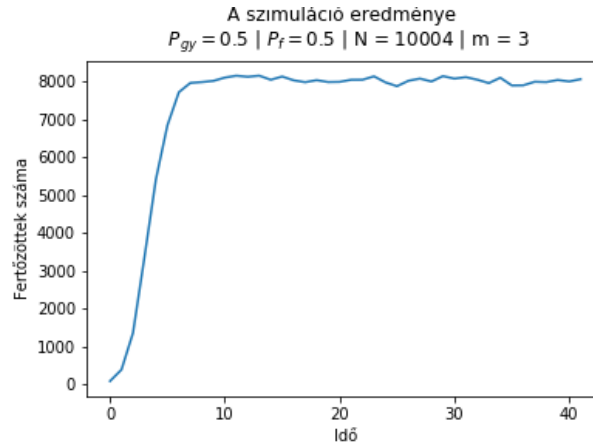
A program három paraméterrel dolgozik: a gyógyulási és fertőzési valószínűséggel, és a szimuláció időlépésekben mért hosszával.

A járvány csak akkor tud terjedni, ha van a hálózatban legalább egy darab fertőzött egyed. Minden szimuláció előtt adott számú, a gráfban véletlen pozícióban elhelyezkedő csúcsot fertőzött állapotúra kell állítani. Ezt egy külön függvény kezeli a programban, ami paraméterül kapja a kezdeti fertőzött pontok számát. A legtöbb esetben 100 kezdeti fertőzöttről indított³

³ A 100 kezdeti fertőzött a gráf csúcsainak 1%-a.

szimulációkat vizsgálunk, ahol a rendszert leíró különböző paraméterek hatásait hasonlítjuk össze.

Az 5. ábrán egy tipikus időfejlődés látható. A fertőzöttek száma gyorsan emelkedik, a 10. időlépésre elér egy állandósult állapotot. Itt a csúcsok 80% fertőződik meg.



5. ábra. Egy próbafuttatás eredménye, a hálózat $N = 10004$ csúcsot tartalmaz, a bekötő élek száma $m = 3$, a fertőzési és gyógyulási valószínűségek: $P_f = 0.5$ és $P_{gy} = 0.5$

4. Futási idő

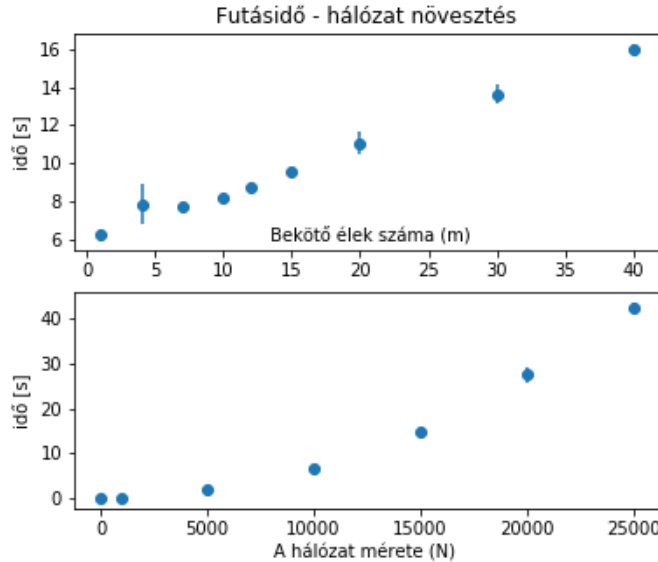
Általában nagy rendszereket szeretnénk megvizsgálni, ezért – mint a legtöbb program esetében – itt sem feledkezhetünk meg a futási idő kérdéséről. A későbbiekben látni fogjuk, hogy ez az átfogóbb vizsgálatoknál már igen jelentős kérdés lesz.

A hálózat építő és a SIS-modell algoritmusa valószínűségi sorsolásokon alapul. Nagyon sok véletlen szám sorsolása pedig megjelenik a futás idejében is. Minél több véletlen számot kell sorsolni, annál lassabb lesz a program.

4.1. A gráfot felépítő program

A hálózat növesztésének ideje két dologtól függ, a rendszer méretétől (N) és a bekötő élek számától (m). A futási idő egy adott paraméter mellett természetesen nem egységes, ezért több futtatást kell egymás után megmérni, hogy pontosabb értéket kaphassunk. Amikor a bekötő élek számának hatását vizsgáljuk fontos odafigyelni, hogy a végső mérete a hálózatnak egységes legyen, és hogy legyen annyi csúcs kiindulási hálózatban, amennyi az adott esetben a bekötő élek száma. Mivel a hálózatot a mindig ugyanabból a 4 csúcsú gráfból indult ki (látsd a 3. ábrán), ezért az $m > 4$ esetekben még az időmérés megkezdése előtt a ciklus m méretűre növesztette

a gráfot ($m = 4$ -es növesztéssel). A 6. ábrán különböző m értékekhez egymás után 5 futtatás idejét megmérve, az 5 futtatás átlagát és szórását láthatjuk.



6. ábra. A gráfépítés futásidők vizsgálata. A felső ábrán lineárisan nő a bekötő élek számával a futási idő. Az alsó ábrán a hálózat méretével a futási idő parabola szerűen növekszik.

A hálózat méretének hatását egyszerűbb megmérni. A standard kiinduló gráfot $m = 4$ bekötőélel $N = 10000$ csúcsúra hízalva, szintén 5 párhuzamos futtatás átlag és szórásértékeivel adódik az eredmény.

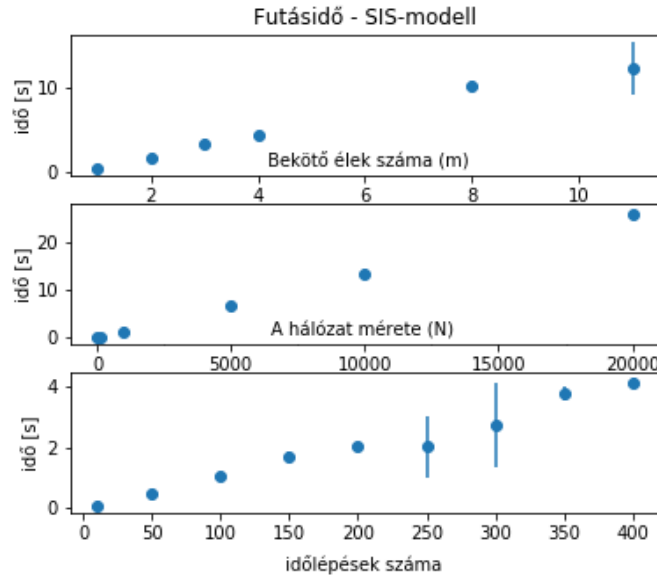
4.2. A SIS-modell

A SIS-modell esetében egyel több paraméter jön szóba, mégpedig az időlépések száma. A fertőzési és gyógyulási valószínűségek hatása is számít ez azonban csak konkrét folyamatok esetén vizsgálható.

Az ábrák lineáris trendet mutatnak, kivéve a hálózat növekedése N függvényében ábrát, ami a lineárisnál nagyobb mértékben növekszik. Mivel a terjedési szimuláció sokkal több alkalommal fut, mint a hálózat építés, ezért a vizsgálni kívánt hálózat méretét alapvetően ez alapján célszerű eldönteni. A későbbiekben lesz szó olyan szimulációkról, amelyek a teljes $P_f - P_{gy}$ valószínűség téren végig iterál⁴. Ezeknek a nagyobb volumenű szimulációknak már több óra hosszú a futási idejük. Az $N = 10000$ csúcsú gráf bizonyult még kezelhető méretűnek, ezért a kutatásban erre a méretre alapoztam. A futási idő ceteris paribus⁵ alacsony bekötő élszám mellett lesz rövidebb.

⁴ Adott lépésközzel lépkedve az összes lehetséges $P_f - P_{gy}$ valószínűség pároshoz futtat szimulációt.

⁵ Minden más tényező változatlansága mellett.



7. ábra. A járványterjedés szimulátor futásidejének vizsgálata. A SIS-modell futási ideje minden paraméterben lineárisan növekszik.

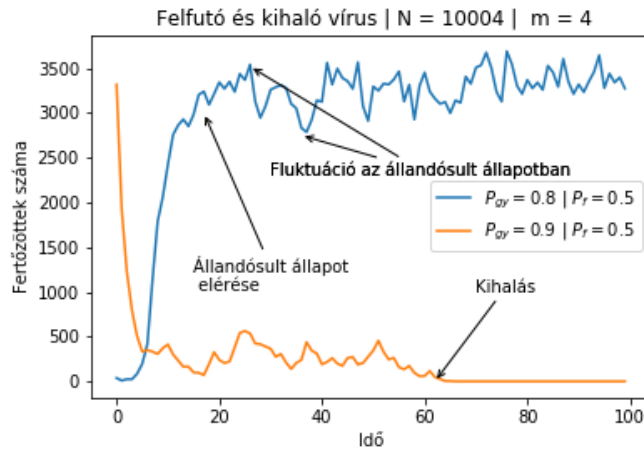
5. A szimulációs eredmények értelmezése

A szimuláció algoritmusa különböző paramétereket használ: az új csúcsokat bekötő élek számát (m), a hálózat végső méretét (N), a gyógyulási és fertőzési valószínűségeket (P_f és P_{gy}). Egy adott paraméter négyeshez a szimuláció egy fertőzöttek száma – idő adatpár sorozatot ($I(t)$) ad eredményül. Alakja alapján $I(t)$ különböző lehet. Egy általános folyamat, hogy a szimulált vírus elterjed. Valami alacsony magból⁶ kiindulva egyre több csúcsot fertőz meg, végül beáll egy állandósult állapotba („steady state”). Mint az eredmények értékelésénél látni fogjuk, ez a leggyakoribb szcenárió, amikor a folyamat „felfut”. Azonban olyan eset is előfordulhat – ha elég nagy a gyógyulási valószínűség –, hogy a vírus „kihal”, a fertőzöttek száma nullává válik. Mivel a modellben csak a 0. pillanatban fertőződnek spontán egyedek, ezért, ha a szimuláció során bármikor eléri a 0-t a fertőzöttek száma, akkor onnantól a szimuláció végéig 0 is marad.

⁶ Kezdeti fertőzöttek egy csoportjából.

Az $I(t)$ vizsgálata az alábbi kérdésekre ad választ:

1. Milyen arányban fut fel vagy hal ki a vírus?
2. Hányadik lépésnél éri el az állandósult állapotot, avagy hányadik lépésnél hal ki?
3. Mennyire ingadozik, fluktuál a betegek száma az állandósult állapotban?
4. Az állandósult állapotban a fertőzöttek számának mi a várható értéke?

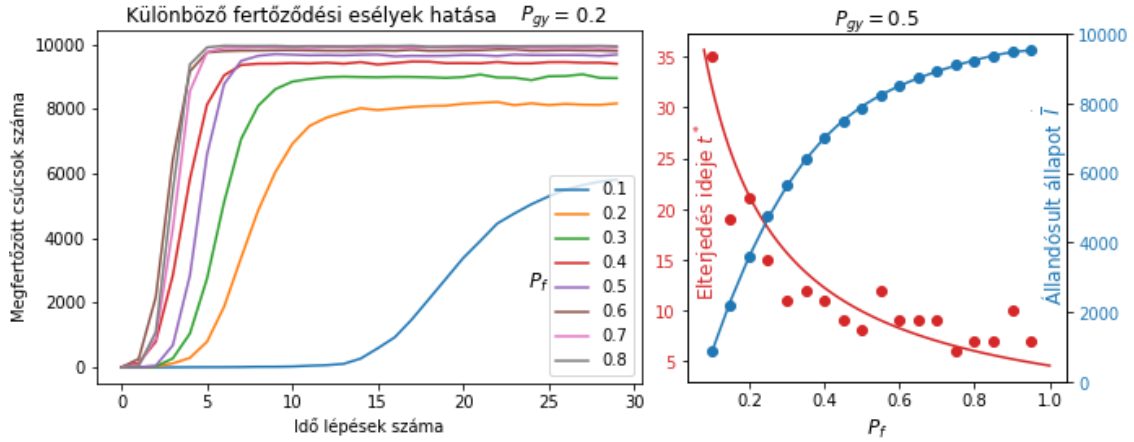


8. ábra. A fertőzött csúcsok számának időbeli fejlődése. A folyamatok végállapotukat tekintve vagy felfutnak és a paraméterektől függő várható érték körül mozognak, vagy lecsengenek.

A különböző paraméterekkel futtatott szimulációk eredményeit megvizsgálva megsejthető néhány törvényszerűség a terjedés lefolyásával kapcsolatban. A gyógyulás és a fertőzés valószínűsége érthető módon egymással ellentétesen hat. Nagy fertőzési valószínűség esetén inkább felfut, nagy gyógyulási valószínűség esetén inkább kihal a vírus. Azonban a két hatás nem ugyanolyan mértékben változik a valószínűségek változtatásának hatására. A fertőzött csúcsok egy körben több élen is fertőzhetnek, viszont gyógyulás esetén, csak az az adott csúcs gyógyul meg, a többi fertőzött csúcs – a gyógyulást tekintve – tőle független. Emiatt sokszor akkor is felfut a vírus, ha a gyógyulás valószínűsége magasabb a fertőzés valószínűségénél. Amiatt, hogy minden él mentén fertőzhetnek a csúcsok nem körönként csak egy élükön, ezért az összes valószínűség-pár többségében a folyamat fel fog futni és csak kisebb részében fog kihalni.

A fent említett paraméterek közül egyet változtatgatva, a többi fixen tartva képet kaphatunk az adott paraméter hatásáról. A 9. - 10. ábrákon a hálózat minden szimuláció esetén ugyanaz a standard $N = 10000$ méretű, $m = 4$ bekötő élszámú gráf látható. A valószínűség

paraméterek egyikét állandó értéken tartva a másik valószínűség különböző értékei különböző lefolyást eredményeznek. Ezek a különböző lefolyások különböző színekkel jelennek meg az ábrákon.



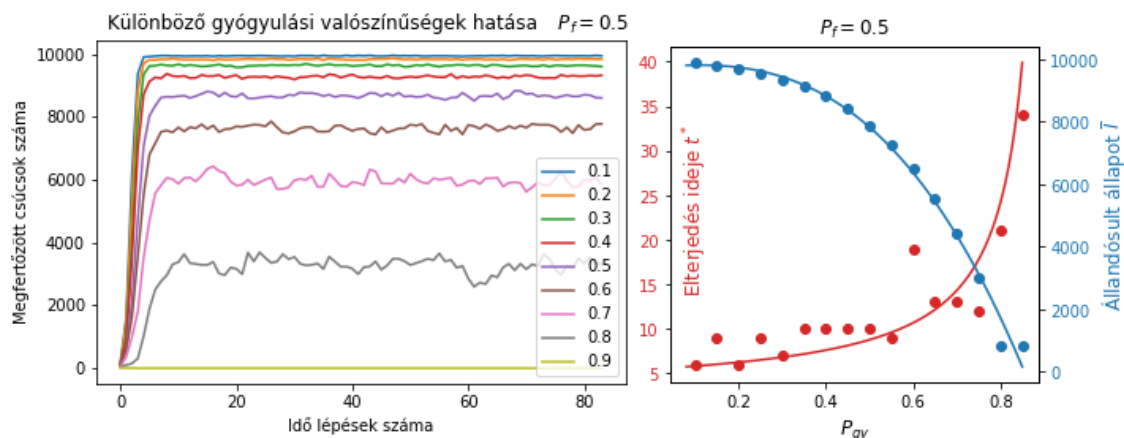
9. ábra. A fertőzési valószínűség hatása a fertőzöttek számára. Az egyre nagyobb fertőzési valószínűségek egyre magasabb állandósult állapotba viszik a folyamatot. Ráadásul minél nagyobb ez a fertőzési valószínűség, annál rövidebb idő alatt terjed el a fertőzés. A jobb oldalon – egy másik gráfon megmérve – láthatóak a konkrét trendek.

A gyógyulási valószínűség és a hálózat állandósága mellett a különböző fertőzési valószínűségek hatása a folyamatra. Minél nagyobb a fertőzési valószínűség, annál magasabbra kerül az állandósult állapot, és annál kevesebb idő kell ahhoz, hogy a folyamat elérje azt. Hasonlóan, ha adott hálózat esetén nem a P_{gy} hanem a P_f állandó, akkor a gyógyulási valószínűség a lefolyásra gyakorolt hatásáról kaphatunk képet.

Adott rögzített paraméterek esetén sem tökéletesen ugyanazt az eredményt kapjuk. Ennek a valószínűségi megközelítés az oka. Azonos paraméterekkel több párhuzamos futtatást nézve látható, hogy a felfutáshoz szükséges idő változik, az állandósult állapotok elég közel vannak, de különösen a fluktuáció hatására mégsem esnek egybe. A folyamatok jellemzése ezzel szemben egy olyan megközelítést igényelne, ahol egységes állítások tehetők az adott paraméterekhez tartozó folyamatra. Ha több folyamatot futtatunk párhuzamosan⁷, akkor ezek eredményeinek átlagai már egy egységesebb, univerzális eredményhez közelítenek.

Fontos továbbá, hogy az átlagok vizsgálatánál elkülönítsük az összes esetet (a kihalt eseteket is beleértve) és a csak felfutó esetekből álló átlagot. Az átlagba a kihaltakat is beleszámító

⁷ Párhuzamos futtatás alatt a dolgozatban azt értjük, hogy egy konkrét gráfon, állandó gyógyulási és fertőzési valószínűséggel, illetve szimulációs hosszal egymás után több szimulációt is elvégezzünk. A párhuzamos futtatásokkal vizsgálható meg, hogy adott valószínűségek esetén milyen arányban fut fel vagy hal ki a folyamat. Több párhuzamosan futtatott terjedési folyamat alapján az állandósult állapotban becsülhető fertőzöttek száma várható érték hibája jelentősen csökkenthető.



10. ábra. A gyógyulási valószínűség hatása a fertőzöttek számára. A többi paramétert rögzítve, növelve a gyógyulás valószínűségét az állandósult állapot egyre alacsonyabb szintre kerül, a kilengések pedig egyre nőnek. A gyógyulási valószínűséget növelve az állandósult állapot eléréséhez egyre több idő szükséges. A jobb oldalon – egy másik gráfon megmérve – láthatóak a konkrét trendek.

esetben az átlag tartalmaz egy a kihalások arányára vonatkozó információt is, azonban ez abból a szempontból nem szerencsés, hogy a felfutók állandósult állapot értékét torzítja és nehéz visszanyerni belőle a kihalások arányát. A felfutás / kihalás arányt éppen ezért érdemes külön kezelni.

Nagy fluktuációk esetén arra is ügyelni kell, hogy a szimuláció elegendően sok időt töltsön az állandósult állapotban. Ha túl rövid ideig szimulálunk, akkor az állandósult állapot értéket a véletlenszerű ingadozás számottevően torzíthatja.

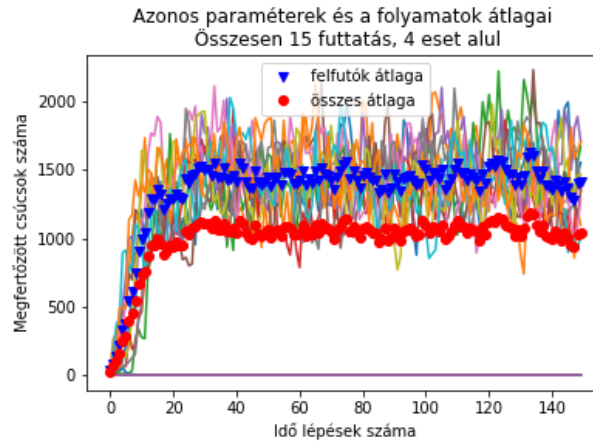
5.1. Az állandósult állapot kezdetének detektálása

Egy adott paraméternégyessel⁸ futtatott szimuláció egyik jellemzője az állandósult állapot értéke (\bar{I}), ami viszont fluktuáció miatt ingadozik (σ_I). Hogy egy konkrét értéket kaphassunk átlagolnunk kell a fertőzöttségi értékeket attól a ponttól, ahol már beállt az állandósult állapot. Grafikusan ábrázolva a folyamatot szemmel általában könnyű eldönteni, hogy hol van ez a pont, azonban célszerű ezt automatizálni.

Nemcsak az \bar{I} meghatározásához fontos ez a pont, hanem a szimuláció hosszának megválasztásához is. Mint azt a futási idő rész is bemutatja, a fölöslegesen hosszú szimuláció költséges.

Keressük tehát minden szimulációs folyamat azon pontját, ahonnan úgy ítéljük, hogy elérte

⁸ A négy paraméter: a gráf mérete (N), a bekötő él szám (m), a fertőzési és gyógyulási valószínűség (P_{gy} és P_f).



11. ábra. A tipikus folyamat meghatározása több futtatás alapján. Magasabb gyógyulási valószínűség mellett az állandósult állapotban nagy kilengések tapasztalhatóak, 15 párhuzamos futtatás látszik az ábrán különböző színekkel. Ezek az egymásba gabalyodó vékony vonalak. Habár a lefolyások az ingadozás miatt nem fedik egymást, de elég sok futtatás átlaga már egészen sima eredményhez közelít. Az átlagokat jelölik a vastag kék és piros pontok. A kék pontok a felfutó folyamatok átlagai, míg a pirosakba a kihaló esetek is bele vannak számolva.

az állandósult állapotot (t^*).

Mivel magát az állandósult állapot értékének megadásához keressük az állandósult állapot elérésének helyét, ezért a kereséshez nem használhatjuk fel magát az állandósult állapot értékét. A folyamat általános lefolyásából kell kiindulnunk. A szimulált terjedés mindig egy relatíve kicsi (0-1%-nyi) fertőzött magból indul ki. Ezután – kivéve az kifejezetten magas relatív gyógyulási valószínűségű eseteket – elkezd növekedni a fertőzöttek száma és eléri a keresett (de egyelőre nem ismert) szintet ami körül fluktuál. Mivel kicsi a kiindulási mag, ezért az állandósult állapot magasabban van, tehát ha a teljes folyamatot kiátlagoljuk, akkor egy jó alsó becslését⁹ kapjuk az állandósult állapotnak, az a lépés pedig amikor a folyamat eléri ezt a fertőzöttségi szintet (a teljes folyamat átlagát) egy jó becslése a keresett pontnak.

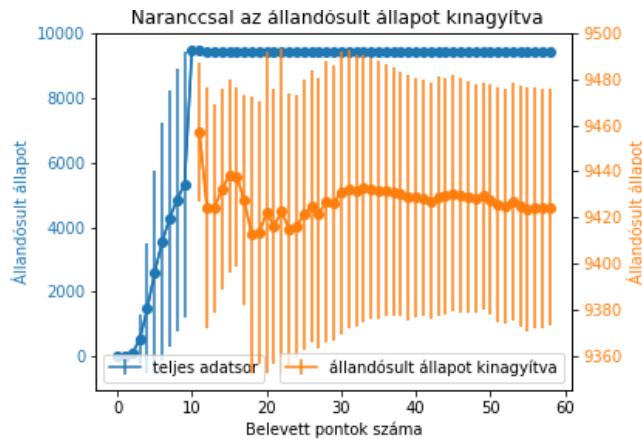
Ha egy elég hosszú folyamat¹⁰ esetén, végigiterálunk folyamaton, és minden lépésben az adott indextől a végéig kiszámoljuk az átlagot, akkor egy másik alkalmas ponthoz is eljuthatunk. Jelöljük \bar{I}_i -vel azt az átlagot, amibe az i . ponttól kezdve az összes pont értékét belevesszük.

$$\bar{I}_i = \frac{\sum_{t>i} I_t}{\sum_{t>i} 1}$$

⁹ A dolgozatban vizsgált esetekben a fertőzöttek kezdeti száma mindig (leszámítva a 8. ábrán a sárga, kihalást szemléltető esetet) maximum 100 volt. A vizsgált esetekben nem volt olyan állandósult állapot, ami ne lenne ennél magasabb. Amikor az állandósult állapot 100 alatti, olyankor a folyamat rendre eléri a 0-t, kihal. Látsd a 15. ábrán.

¹⁰ Legalább 100-150 lépés hosszú

Amíg a folyamat elején egyre növekszik a fertőzöttek száma, addig az i -vel lépkedve ez az \bar{I}_i átlag is egyre csak nagyobb lesz, hiszen egyre inkább elhagyjuk a folyamat elejéről az alacsony fertőzöttségű pontokat. Ezzel szemben, amikor már az állandósult állapotban vagyunk, akkor a fluktuáció miatt \bar{I}_i néha növekszik máskor viszont csökkenni fog, ahogyan az a 12. ábrán is látszik. Így tehát ha ennek az \bar{I}_i -kből álló sorozatnak megkeressük az első inflexiós pontját¹¹ akkor egy olyan pontot kapunk ami még pontosabban megtalálja az állandósult állapot beálltát¹².



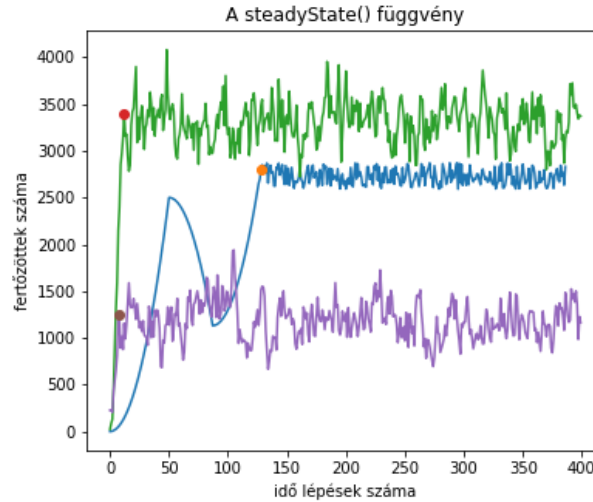
12. ábra. Az ábrán egy terjedési folyamatra kiszámolt átlag sorozat látható. Az i . pontban a folyamat i -től a végéig a pontok átlaga és szórása látható. Naranccsal az állandósult állapot elérését követő pontok láthatóak kinagyítva (felettük ugyanazok késsel az eredeti skálán). Az ábrán szépen látszik, hogy a kezdetben monoton növekvő átlagokból álló sorozat az állandósult állapotot elérve fluktuálni kezd.

A módszer a felfutás közbeni esetleges visszahanyatlásokat is szépen kezeli, viszont egy nagyobb visszaesés esetén csak akkor ad jó eredményt, ha az állandósult állapot elegendően hosszú ahhoz, hogy dominálja az átlagot. Ahol először éri el a folyamat az így megtalált állandósult állapot szintjét (\bar{I}), az a keresett t^* pont. Ezt a t^*, \bar{I} párt jelöli egy-egy ponttal a különböző folyamatok esetén a 13. ábra.

Ha megvan az állandósult állapot kezdete, akkor ettől a ponttól számítva az átlag ad egy fertőzöttségi értéket. A fluktuáció miatt minél több pontot átlagolunk – minél hosszabb a szimuláció – annál pontosabban kapjuk meg a keresett értéket. A futási időt vizsgáló résznél láthattuk, hogy az időlépések számával a szimuláció lineárisan egyre hosszabb. Azonban jobban megéri egy nagyon hosszú szimuláció helyett több rövidebbet futtatni, és ezek állandósult állapotait átlagolni. Bizonyos gyógyulási és fertőzési valószínűségek esetén a folyamat kihal, speciálisabb esetben

¹¹ Ahol növekvőből csökkenőbe vált.

¹² Megjegyzés: számos szimuláció esetén megvizsgálva a módszer kielégítőnek bizonyult. Azonban a módszer korlátait, sajátosságait érdemes alaposabban megvizsgálni.



13. ábra. Az állandósult állapot kezdetének detektálása. A zöld és lila idősor szimuláció eredménye, a kék terjedési folyamat kézzel szerkesztett patológikus eset. Ebben az esetben a kék folyamat elég időt van az állandósult állapotban, hogy egy ilyen mély beszakadást is túldomináljon. A lila folyamat kezdetén is van egy kis visszaesés, ami nem téveszti meg a programot.

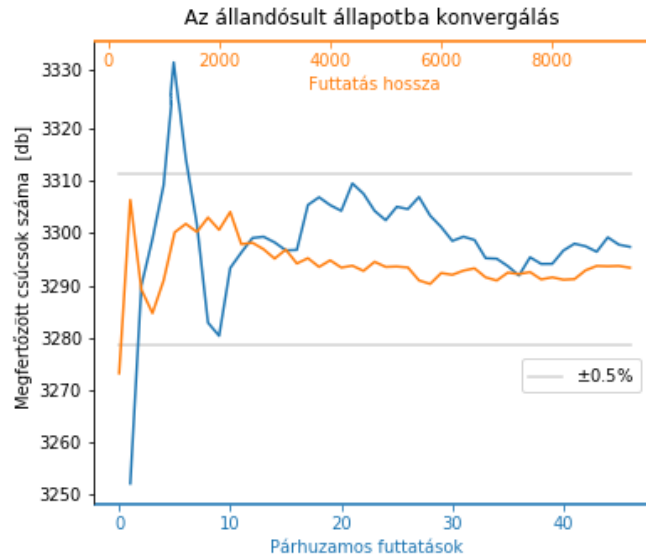
egy olyan alacsony állandósult állapot körül olyan nagy amplitudóval¹³ fluktuál, ami elérheti a 0-t, ezzel kihalhat a vírus. Ha egy hosszú szimulációt végeznénk, akkor az egy ilyen esetben könnyen kihalhat, ami az átlagot eltorzítja. Ezzel szemben több rövidebb párhuzamos futtatás átlagát nézve, külön kezelve a kihalt eseteket pontosabb képet kaphatunk. A párhuzamos futtatások azért is célravezetőek, mert csak így mérhető a felfutás / kihalás arány adott valószínűségek esetén.

Az ábrával kapcsolatban fontos megjegyezni, hogy a 0.5%-nyi tartomány függ az állandósult állapot értékétől is de a fluktuáció nem. Ezért egy alacsonyabb állandósult állapot esetén nem várható ekkora pontosság. A későbbiekben viszont látni fogjuk, hogy a valószínűség párok legtöbbje esetén az állandósult állapot ennél a kb. 3300-as értéknél magasabb.

6. A paraméter tér vizsgálata

Az eddigiekben láthattuk, hogy hogyan működik a program, milyen általános jellemzői vannak a terjedési folyamatoknak. Megfigyeltünk trendeket, sikerült automatizálni az állandósult állapot megtalálását, és láthattuk, hogy a paraméterekben legalább lineárisan nő a futási idő. Ezekkel az ismeretek szükségesek egy átfogóbb szimuláció sorozat futtatása előtt. Ebben a feje-

¹³ Kilengésekkel



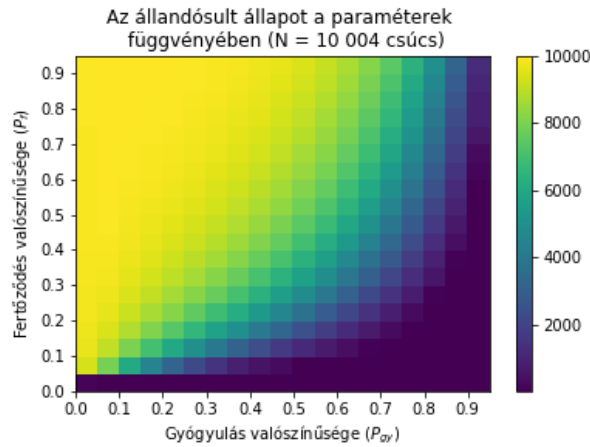
14. ábra. Az időátlagok és a szimulációs sokaságátlagok összevetése. A szimuláció paramétere: $N = 10000$, $m = 4$, $P_f = 0.5$, $P_{gy} = 0.8$. Narancssal az egy hosszú futtatás átlagai látszanak. A felső vízszintes tengelyen látható, hogy az adott pontban a szimuláció első hány pontját számolta bele az átlagba, minél nagyobb ez a szám, annál jobban közelíti meg a helyes értéket. Kékkel azt láthatjuk, hogy az adott pontba hány párhuzamos futtatás eredményét átlagolta ki (alsó vízszintes tengely) és ezeknek mennyi az eredményük. Szürkével a két utolsó pont (a két legpontosabb érték) átlagától $\pm 0.5\%$ -nyi sáv látható. Ha el tudunk fogadni ennél valamivel nagyobb, 1% -nyi hibát az eredményben, akkor mindkét módszer elég hamar eléri ezt.

zetben olyan szimulációkról lesz szó, melyek során a hálózat mérete adott (a 4 csúcsú kiindulási gráf + 10 000 további csúcs) és „letapogatva” a teljes paraméter teret vizsgáljuk a folyamat jellemzőit (állandósult állapot értéke és elérésének ideje, fluktuáció és a kihalás aránya). Ez azt jelenti, hogy egy bizonyos lépésközzel (ami kezdetben 0.05 később pedig 0.025 lesz) 0-ról léptetgetjük a gyógyulási és fertőződési valószínűségeket, és minden lehetséges valószínűség pár esetén futtatunk több párhuzamos szimulációt, ezek eredményeit összesítve fogunk általános érvényű tulajdonságaira bukkanni a modellnek.

Korábban megvizsgáltuk, hogy fix hálózat és gyógyulási (másik esetben meg fertőzési) valószínűség mellett, ha a másik valószínűséget változtatjuk, annak mi lesz a hatása. Ekkor egy ábrán kirajzolva a különböző szimulációs eredményeket láthattuk az adott paraméter változtatásának hatását. Amikor viszont a teljes paraméter teret vizsgáljuk az ábrák két tengelyén a két valószínűségi paraméter értékeit tüntetjük fel, ezzel egy négyzetrácsot képezve, és a négyzetrács színezése fogja a szimulációs eredmény egy konkrét jellemzőjét leírni (pl. az állandósult állapot értékét).

A következő eredmények olyan hálózaton születtek, ahol a bekötő élek száma $m = 3$, az

adott valószínűség párral futtatott „párhuzamos futtatások” száma 10 volt.



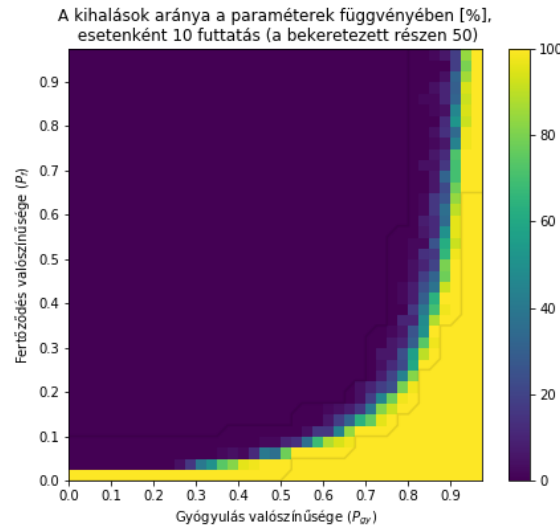
15. ábra. A fertőzöttek számának várható értéke különböző szimulációs paraméterválasztás mellett. Az ábrán a valószínűségek lépésköze 0.05, tehát a 0-1 tartomány 20-20 egyenlő részre van bontva.

Az állandósult állapot kereséséről szóló részben láthattunk egy összehasonlítást a több párhuzamos futtatás és a nagyon hosszú futtatás eredményeiről, hogy milyen hosszú futtatás átlaga milyen állandósult állapot értéket ad és hány párhuzamos, rövid futtatás eredményének átlag ettől miben tér el. A 14. ábrán be volt jelölve egy $\pm 0.5\%$ -nyi sáv, amiről azt állítottam, hogy habár a százalékos tartomány függ az állandósult állapot értékétől, viszont abban az esetben az 3000 körüli volt, és a legtöbb valószínűség esetén ennél magasabb az állandósult állapot. A 15. ábra megerősíti, hogy valóban így van. Az ábrán csupán a sötétkek terület van 3000 alatt, ám mint azt látni fogjuk azon a területen jellemzően a folyamat ki is hal.

A kihalások arányához mindenképpen több párhuzamos futtatás szükséges, minél több szimulációt futtatunk egymás után adott paraméterek mellett, annál pontosabb eredményt kapunk. Az egymás után futtatott „párhuzamos futtatásokból” valahány esetben a folyamat felfut, más-kor eléri a 0-t, kihal. Ha n_f -fel jelöljük ezek közül a felfutó esetek számát, míg n_{gy} -vel a kihalókat, akkor ezzel a jelölésrendszerrel a vizsgált „felfutás / kihalás” arány alatt a $\frac{n_f}{n_f+n_{gy}}$ hányadost értjük. A 0.05-ös lépésközzel lefuttatva a szimulációt, a kihalások aránya olyan képet mutat, hogy a fenti ábrán a sötétkek részen minden esetben kihal, a többi területen pedig minden esetben felfut. A sötétkek rész határvonalán pedig egy két négyzet átmenet a kettő között, amikor csak az esetek egy bizonyos hányadában hal ki.

Viszont amikor apróbb lépésközzel vizsgáljuk meg a jelenséget, akkor egy nagyobb felbontású képet kapunk, amiben már szépen látszik a vékony átmenet a „mindig kihal” és „mindig felfut” területek között. Egy ilyen nagyobb felbontású kép a 16. ábra.

Az állandósult állapot megtalálásáról szóló részben szó volt arról, hogy bizonytalan, hogy



16. ábra. A lehetséges végállapotok viselkedése a fertőzési és a gyógyulási valószínűség paramétereik függvényében. A sárga területen a folyamat mindig kihal, a sötétlila területen pedig mindig felfut. A kettő között egy vékony sáv az átmenet, ahol az esetek egy bizonyos hányadában fut fel vagy hal ki. A 10 párhuzamos futtatás eredménye alapján körbe lehetett keríteni az átmeneti sávot (összegyűjteni azon valószínűség párokat, amelyek esetén átmeneti eredményt kaptunk). Ezekre a valószínűség párokra már több párhuzamos futtatást vizsgálva pontosabb eredményt kapunk. Az ábrán egy vékony szürke keret jelzi azokat a pontokat, ahol 10 helyett 50 párhuzamos futtatás eredménye látható.

hány lépés után érjük el az állandósult állapotot, továbbá ahhoz, hogy az állandósult állapot kereső program jó eredményt adjon, ahhoz az állandósult állapot beállta után még „elegendő időt” kell ott töltenie a folyamatnak, hogy az állandósult állapotban töltött rész elég domináns legyen az átlagok vizsgálatához. Ebből a megfontolásból az első szimuláció sorozat minden valószínűség pár esetén 200 idő lépésen keresztül zajlott.

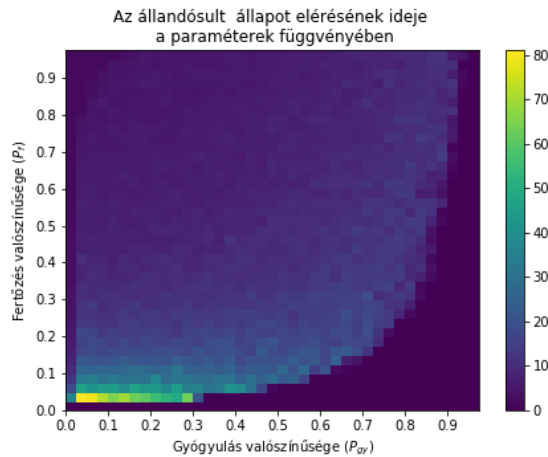
Mindkét valószínűség 20 különböző értéket vehet fel 0.05-ös lépésköznél, ami összesen 400 esetet jelenet. Mindegyik valószínűség párral 10 különböző futtatás ezt 4000 db szimulációra növeli. Ilyen nagy számú szimuláció már 8-10 órás futási időt produkál. A fenti ábrán ráadásul a lépésközt az eredeti fele, csupán 0.025, ami a 4000-ret 16 000-é növeli.

Két módszer adódik a teljes szimuláció futás idejének rövidítésére:

1. Minden szimuláció időlépéseinek számát 200-ról visszavenni valami alacsonyabb értékre, ehhez fel kell térképezni, hogy minden valószínűség pár esetén hány időlépés alatt éri el a folyamat az állandósult állapotot.
2. A szimulációk darabszámának csökkentése. A kihálás arányának vizsgálatakor kirajzoló-

dott az a tartomány, ahol átmenetet találunk. A szimulációt csak ezekre a valószínűség párokra futtatva a szimulációk számát lehet csökkenteni. Nevezzük ezt bekeretezésnek!

A két módszert egyszerre alkalmazva már számottevően gyorsabban, 1-2 óra alatt lefuttatható egy szimuláció sorozat, még úgy is, ha a párhuzamos futtatások számát 10-ről 25-re növeljük. A szimuláció hosszának csökkentéséhez meg kell mérni a különböző valószínűségek esetén az állandósult állapot elérésének idejét. A 17. ábrán egy ilyen mérésnek az eredménye látható, ez alapján a futtatási hossza a legtöbb esetben több mint a felére csökkenthető.



17. ábra. Az állandósult állapot beálltának vizsgálata. Az ábrán 0.025-ös lépésközzel az állandósult állapot elérésének ideje (t^*) látható. Az ábra alapján látszik, hogy a legtöbb esetben 30 lépés alatt a folyamat eléri az állandósult állapotot, vagy kihal. Csupán az egészen alacsony fertőzési valószínűségek esetén kell 40-80 lépést várni erre.

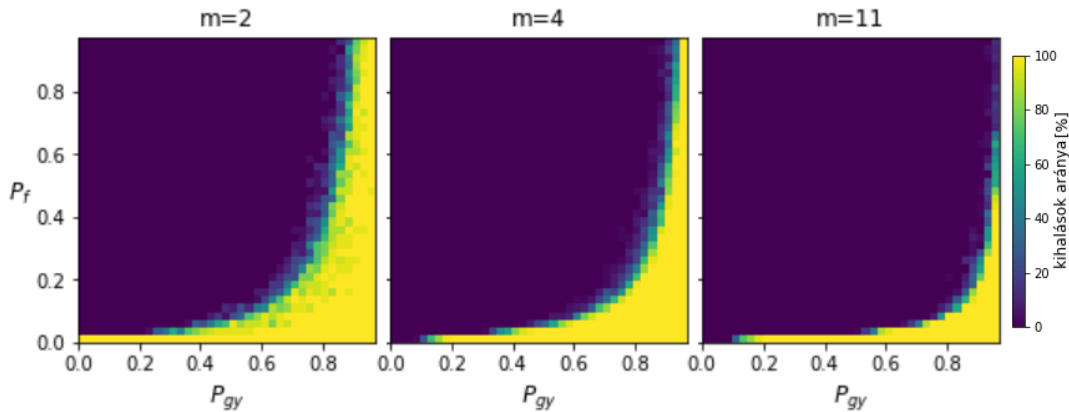
6.1. A hálózat éleinek hatása a kihalás/felfutás arányra

A teljes valószínűség teret letapogató szimulációk a modell különböző érdekes tulajdonságaira világítottak rá. Ezek a vizsgálatok mindig ugyanazon az $N = 10004$ csúcsú, $m = 3$ bekötő élszámú gráfon folytak. Felvetődik a kérdés, hogy miképp változik az eredmény, ha ugyanannyi csúcsú, de más szerkezetű gráfokat vizsgálunk. A más szerkezet alatt két dolgot kell érteni: egyrészt más bekötő élszámmal felépített gráfot, másrészt egy adott konkrét gráfból éleket elvéve vagy éppen újabbak behúzásával kapott gráfot.

Összehasonlítva az $m = 2$, $m = 4$ és $m = 11$ bekötő éllel felépített gráfot (18. ábra), az tapasztaljuk, hogy az $m = 11$ esetben a határvonal jobbra lejjebb, az $m = 2$ esetben pedig balra feljebb tolódik el. A bekötő él számának növelésével a hálózatok egyre sűrűbbé válnak abban az értelemben, hogy egyre több él fog ugyanannyi csúcshoz tartozni. A sűrűbb hálózatokban ez

alapján a folyamat több esetben fog elterjedni, míg a ritkább hálózatok esetében gyakrabban hal ki a terjedési folyamat.

Láthattuk, hogy a határvonalhoz egy átmeneti sáv is tartozik, ahol bizonyos esetekben kihal, máskor pedig felfut a terjedési folyamat. Ez az átmeneti sáv a sűrűbb hálózat esetén elvékonyodik, míg a ritkább hálózatban kiszélesedik.



18. ábra. A gráfépítés során választott bekötő élszám hatásának vizsgálata. A középen látható kiinduló $m = 4$ bekötő éllel felépített hálózat átmenetéhez képest az $m = 2$ eset balra-felfelé tolódik és az átmenet kiszélesedik. Ezzel szemben a sűrűbb $m = 11$ -es hálózat a határvonalat jobbra-lefelé tolja, az átmenetet pedig szűkebbé teszi.

A vizsgálódásból felsejlik egy szabályszerűség a határvonal görbáját illetően. Ezek a vonalak szabályosnak tűnnek, felmerül az igény, hogy valahogyan jellemezzük azokat. Egy másik izgalmas kérdés pedig, hogy a hálózatoknak milyen tulajdonságai hogyan befolyásolják a határvonal eltolódását. Ennek vizsgálatára tekintsük az $m = 4$ bekötő éllel felépített gráf néhány speciálisan módosított verzióját. A módosított gráfokba újabb élek kerülnek be vagy éppen ellenkezőleg meglévő éleket húzunk ki. A kihúzandó élek és a csúcs párok kiválasztása (melyek közé az új élek kerültek) preferenciálisan véletlenszerűen zajlik, hasonlóan a hálózatonövesztés preferenciális kapcsolódásához.

A kiválasztás kétféle lehet: a nagy foksámokat lineárisan preferáló¹⁴ vagy anti-preferenciális. Az anti-preferenciális kiválasztás lényege, hogy éppen az alacsony foksámú csúcsoknak ad nagyobb esélyt a kiválasztódásra. A korábban tárgyalt preferenciális kiválasztáshoz hasonlóan itt is minden csúcshoz tartozik egy valószínűség. Ez a valószínűség: $w_k = \frac{k^{-2}}{A}$, ahol¹⁵ $A = \sum_l l^{-2} N_l$. Tehát itt is az összfoксámmal osztunk, viszont nem önmagában a foksámot, hanem annak -2. hatványát.

¹⁴ Pontosan úgy, ahogy a hálózat növesztésénél folyt: minden csúcshoz tartozik egy valószínűség, ami az adott csúcs foksáma elosztva az összfoксámmal.

¹⁵ A képletben szereplő változók: k az adott csúcs foksáma, míg N_l az l foksámú csúcsok száma.

Az $m = 4$ bekötő éllel felépített hálózatban összesen 40 004 él van, hiszen a kiindulási gráf 4 élű, amihez összesen 10 000 db új csúcsot fejenként 4 éllel kötünk be. Az $m=2$ -es gráfban 20 004, míg az $m=11$ -es gráfban 109 976 (ez nem a hagyományos kiindulási gráfra épült rá, hiszen abban nincs 11 csúcs, és annyi csúcsra minden növesztésnél szükség van, ahány éllel be akarjuk kötni az új csúcsokat.) A 18. ábrán az $m=2$ és $m=11$ esetben a mindig felfut és mindig kihál tartományok közötti határvonal már számottevően arrébb mozdult. Hogy jól látható legyen az eredeti $m = 4$ -es gráf módosított verzióinak az eltérése, ezért vizsgáljunk 20 000 éllel megnövelt / csökkentett hálózatokat.

A szimuláció tervezésénél itt már építhetünk az egyedi szimulációs hosszakra¹⁶ és a szimulációk számának a bekeretezéssel¹⁷ csökkentett számára.

Az eredmények igazolják azon várakozásainkat, hogy a ritkított hálózatban a határvonal balra-felfelé tolódik el, a sűrített pedig épp ellenkezőleg jobbra-lefelé. Az eltolódás mutatja, amit a különböző bekötő éllel felépített hálózatoknál láttunk. A több él több felfutó terjedési folyamatot jelent. A 19. ábrán megfigyelhetjük, hogy más eredményt ad a magas fokszámot preferáló és az anti-preferenciális kiválasztás eredménye. A magas fokszámot preferáló kiválasztás mindkét esetben (elveszük / hozzáadunk élt) a határvonal nagyobb eltolódását eredményezi. A különböző bekötő éllel növesztett hálózatok eredményeinek összehasonlítása rámutatott, hogy a határon az átmenet az $m = 2$ bekötő élű, ritkább hálózatban szélesebb. Ezt a megfigyelést támasztják alá a ritkított $m = 4$ -es hálózatok eredményei is. A ritkított esetekben a határvonal jobban eltolódik.

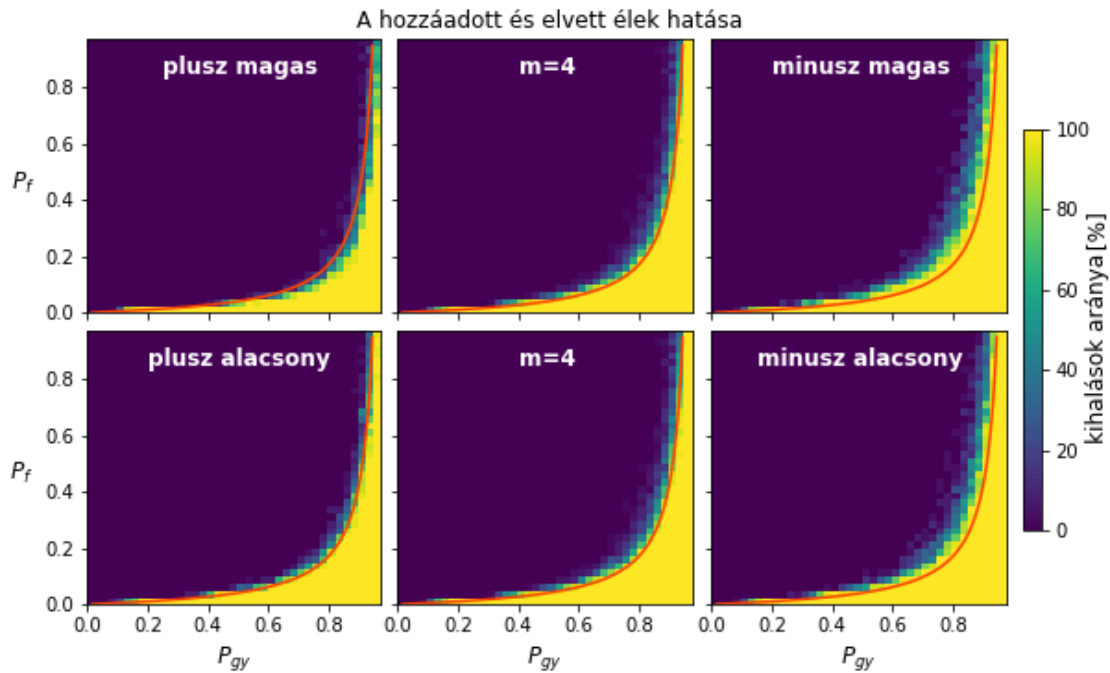
6.2. A határvonal görbéjének vizsgálata

A különböző gráfokra megvizsgált kihálási arányok ábráin a határgörbék alakja szabályos-ságot sugall. A határ átmenetes, egymás mellett futó hasonló görbék-ből áll, ezek a görbék az azonos kihálási arányú pontokat kötik össze.

Ahhoz, hogy a különféle gráfokat össze tudjuk hasonlítani, mindig az ugyanahhoz a kihálási arányhoz tartozó görbét kell vizsgálni. A továbbiakban a legnagyobb kihálási arányhoz tartozó görbét vizsgáljuk. Ha az eredeti görbét úgy ábrázoljuk, hogy a függőleges tengelyt logaritmizáljuk, akkor az azonos görbéhez tartozó pontok egyenesre fognak illeszkedni (20. ábra). Ez alapján az eredeti görbe egy hiperbola volt. A hiperbola alak azt jelzi számunkra, hogy a két valószínűség aránya ha egy adott állandó, akkor a kihálások aránya is egy állandó érték lesz. A

¹⁶ Minden valószínűség pár esetén a korábban megmért (17. ábra) állandósult állapot eléréséhez szükséges idő lépések száma +42 lépésen keresztül fut a szimuláció.

¹⁷ Az első kihálás / felfutás ábrán kirajzolódó átmenetes határ környezetében lévő pontokat mérjük csak meg. Ehhez persze a konkrét átmeneten túl annak környékét sem árt vizsgálni, hogy ki ne hagyjuk átmenetes pontokat. A szimuláció végén az ábra alapján úgy ellenőrizhető, hogy elégséges pontot kereteztünk-e be, hogy az ábrán az átmenet nem törik meg.



19. ábra. A skálafüggetlen hálózat éleinek perturbációjának hatása a fertőzési folyamatra. Mindkét sorban középen az eredeti $m = 4$ -es gráf eredménye látható. Ehhez képest bal oldalon a 20000 új él behúzásának hatása, míg a jobb oldalon a 20000 él kitörlésének hatása látható. Minden ábrán ugyanaz a piros vonal látható, amely az eredeti $m = 4$ -es eset sárga színű kontúrja. Ehhez viszonyítva láthatók az eltolódások. A felső sorban a magas fokszámú csúcsokat, míg az alsó sorban az alacsony fokszámú csúcsokat preferálva lettek kiválasztva a csúcs párok az új él behúzásához, és az élek a kitörléshez. A magas éleket preferáló eset jelentősebb eltolódást okozott.

görbékre illesztett függvény:

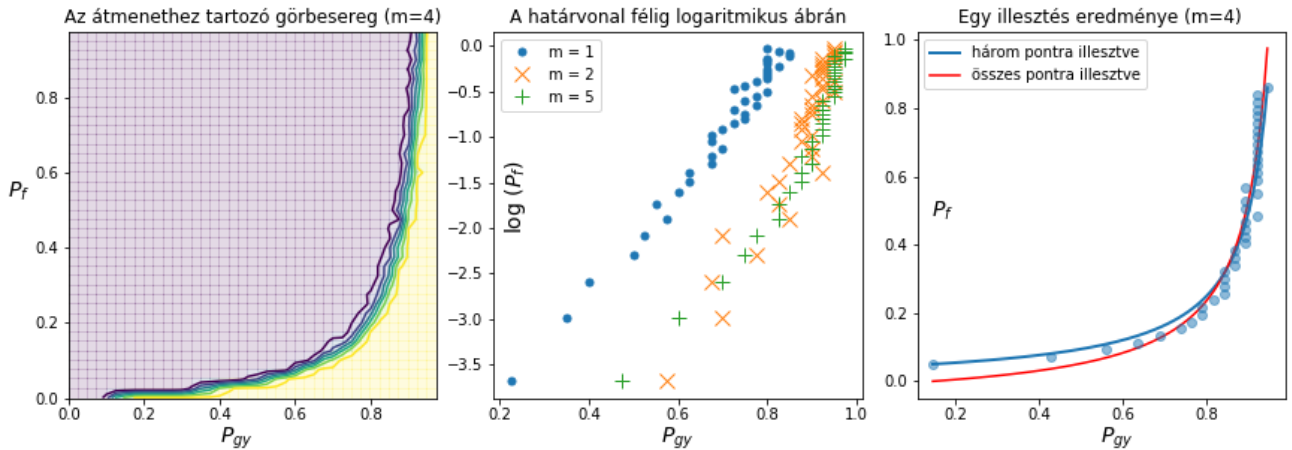
$$y = \frac{Ax + B}{x - C} \quad (11)$$

Az illesztéshez szükség van még egy kiindulási A,B,C paraméter hármasra. Ha az azonos kihalási arányhoz tartozó pontok közül kiválasztunk hármat (egyét az elejéről (x_0, y_0) , egyet a közepéről (x_1, y_1) és egyet a végéről (x_2, y_2) és azokra illesztünk, akkor a három paraméter egy jó kezdeti becsléséhez jutunk. A három pontra a fenti függvényt illesztve egy lineáris egyenletrendszerhez jutunk.

$$\begin{pmatrix} x_0 & 1 & y_0 \\ x_1 & 1 & y_1 \\ x_2 & 1 & y_2 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \end{pmatrix} = \begin{pmatrix} x_0 y_0 \\ x_1 y_1 \\ x_2 y_2 \end{pmatrix}$$

Ennek megoldása lesz az A,B,C paraméter hármas kezdeti becslése. A Python `scipy.optimize`

csomagjának fitfunc függvényével a kezdeti paraméter hármas ismeretében a (11) függvényalakot az összes pontra illeszthetjük. A 6.1. részben láttuk, hogy amennyiben egy másik bekötő



20. ábra. A kihalási arány ábrák görbéjének alakja. A bal oldali ábrán az $m = 4$ -es gráf átmenete látható, az azonos kihalási arányú pontokat összekötve megkapjuk az átmenetet alkotó görbesereget. Ezek közül a jobb oldali legszélső görbét vizsgáljuk az összes gráf esetén. Három különböző gráf ezen görbéit félig logaritmikus ábrán ábrázolva olyan pontsorokat kapunk, melyek egy egyenesre illeszkednek. Ez azt jelzi, hogy az eredeti görbe alakja hiperbola alakú. A jobb oldali ábrán egy illesztés eredménye látható, kékkel a paraméter hármas kezdeti becsléséhez számolt görbe, pirossal pedig az összes pontot figyelembe vevő alak.

élszámmal növesztett gráfra cseréljük az eredeti $m = 4$ -est, akkor a kihalások arányát mutató ábrán a határvonal eltolódik. Az eltolódást a függvényillesztés segítségével már számszerűsíthetjük. Sok különböző bekötő élszám mellett lefuttatva a kihalások arányát mérő szimulációt, a különböző m -ekhez adatpontokat kapunk melyekre illeszthetünk. Az illesztések paramétereit eltárolva képet kapunk a gráfok bekötő élszámának (m) az illesztési paraméterekre gyakorolt hatásáról. A hatás vizsgálatához a 11-szerinti függvényalakot alakítsuk át:

$$y = \frac{A \cdot x + B}{x - C} = \frac{A(x - C) + B + AC}{x - C} = A + \frac{B + AC}{x - c} \quad (12)$$

Vezessünk be új jelölést:

$$a = B + AC$$

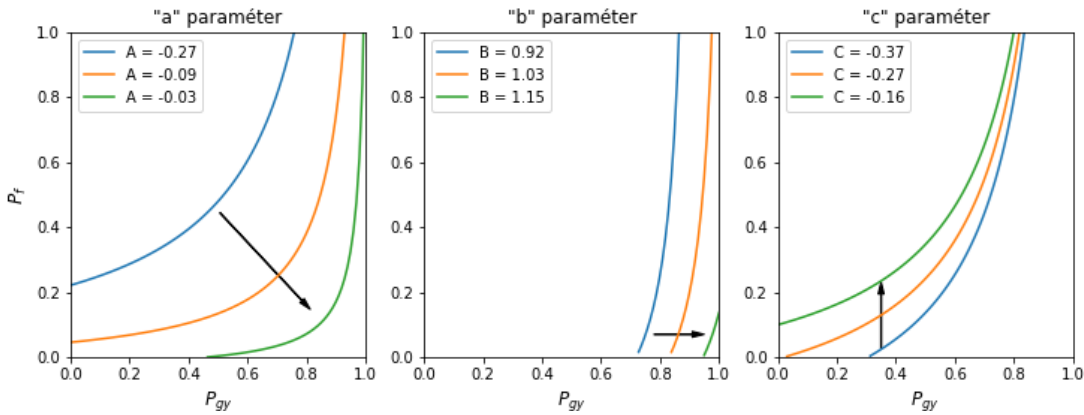
$$b = C$$

$$c = A$$

Ezeket a 12-be beírva, a következő hiperbola egyenlethez jutunk:

$$y = \frac{a}{x - b} + c \quad (13)$$

Egy ilyen függvényalak esetén már könnyebb felismerni a paraméterek hatását. Az a paraméter a hiperbolát görbébbé fogja tenni, a b paraméter vízszintesen, míg a c paraméter függőlegesen tolja el. A 21. ábrán a görbe hármasok úgy vannak kirajzolva, hogy mindig az adott ábrának



21. ábra. Az illesztett függvény paramétereinek hatása a függvény alakjára. Az a meghajlítja a görbét, a b vízszintesen, míg a c paraméter függőlegesen tolja el.

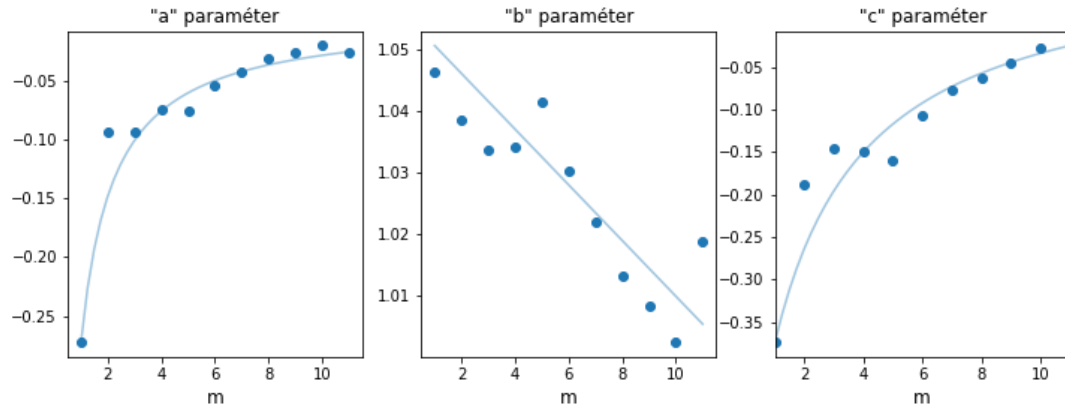
megfelelő paramétert változtatjuk, a másik kettő állandósága mellett.

A különböző bekötő éllel felépített hálózatokhoz tartozó görbék illesztése után láthatóvá válik, hogy milyen trend rajzolódik ki a különböző paraméterekben.

A 22. ábrán a pontsorokra illesztett görbék a szem vezetését szolgálják, nem hivatottak a valós függvényalak feltárására. Látható, hogy a hiperbolák az m növekedése mellett egyenletesen tolódnak jobbra. Ezt a b paraméter lineáris csökkenése jelzi. Az a paraméter növekedése a hiperbolák egyre nagyobb görbülését jelzik, miközben a görbe felfelé tolódik (c paraméter növekszik).

7. Összegzés

A dolgozat rövid betekintést nyújt a hálózat kutatás világába. Definiáltuk a skálafüggetlen gráfokat, melyek fokszám eloszlása hatványfüggvényt követ. Ez az Erőds-Rényi-féle véletlen gráfokkal szemben tartalmaz kiugróan nagy fokszámú csúcsokat. Egy egyszerűbb esetet vizsgálva levezettük, hogy a fokszámok eloszlása valóban hatványfüggvényt követ. Láthattuk, hogy az



22. ábra. A bekötő élek számának ($m = 1$ -től $m = 11$ -ig) változtatásának hatására a görbék eltolódnak. Az eltolódás során a (12) függvényalak paraméterei eltolódnak. Az a paraméter növekvő, a b csökkenő, a c paraméter növekvő trendet mutat. A trendtől való eltérés a növesztési és terjedési folyamat valószínűségeken alapuló bizonytalanságából eredhet.

ilyen fokszámeloszlás log-log ábrán ábrázolva egy egyenes lesz, amit le is teszteltünk a saját gráfjainkon.

Megvizsgáltuk a SIS-modellt, amely a gráfokon modellezi különféle terjedési folyamatok (vírusok, nézetek, trendek, vélemények, betegségek, stb.) lefolyását. A SIS-modellnek azt az esetét vizsgáltuk meg, amelynek algoritmusában először az adott fertőzött csúcsra sorsolunk, hogy meggyógyul-e, majd – amennyiben nem gyógyult meg – minden egyes élére külön sorsolunk, hogy fertőz-e az adott élen. Az algoritmus minden időlépésében az összes fertőzött csúcsra elvégezzük ezeket a sorsolásokat.

A gráfokat felépítő Barabási-Albert-modell és a SIS-modell számítógépes programjait megvizsgáltuk futásidő szempontjából. A paraméterek függvényében a futási idők mindegyike legalább lineárisan skálázódik. A futási idők ismerete a nagyobb volumenű szimulációk tervezésében segít.

A SIS-modellben a betöltött gráfon kívül a modellnek három paramétere változtatható: a szimuláció hossza, a fertőzési és a gyógyulási valószínűség. Egy terjedési folyamatot négy dolog jellemez:

1. Milyen arányban fut fel vagy hal ki a vírus?
2. Hányadik lépésnél éri el az állandósult állapotot, avagy hányadik lépésnél hal ki?
3. Mennyire fluktuál az állandósult állapotban?
4. Milyen magasan van az állandósult állapot?

A dolgozat középpontjában a felfutás / kihalás aránya állt. A kihalás aránya mellett megvizsgáltuk azt is, hogy egy adott gráf esetén mekkora lesz az állandósult állapot értéke, és annak elérése ideje a valószínűség párok függvényében. Ezek vizsgálatának automatizálásához megírtam egy függvényt, ami az állandósult állapotot hivatott megtalálni. A függvény lényege, hogy a szimulációs pontok közül megkeresi az elsőt, amitől számítva a szimulációs pontoknak az átlaga csökkenőbe vált. A függvény mögött az a gondolat áll, hogy a vizsgált esetekben a terjedési folyamat mindig a kezdeti fertőzöttek egy szűk csoportjából indul ki. Mivel ez a csoport szűk ezért, ha felfut a folyamat, akkor a fertőzöttek száma egy darabig valóban növekvő trendet mutat. Az állandósult állapotot elérve a folyamat fluktuálni kezd, innentől tehát az átlag is néha növekszik máskor pedig csökken.

A függvény segítségével az $m=3$ -as gráf esetén a valószínűségeket 0.025-ös lépésközzel léptetve a teljes valószínűségi téren végigiterálva képet kaptunk a kihalások arányának és az állandósult állapot eléréséhez szükséges idő alakulásáról. A szimulációk hosszát kétféleképpen csökkentettük:

1. A megkapott állandósult állapot eléréséhez szükséges idő adatokra alapozva minden pont-hoz egyedi szimulációs hosszt rendelve.
2. A szimulációs pontok leszűkítésével, ezt neveztük „bekeretezésnek”.

A kihalások arányát megvizsgáltuk más gráfok esetén is. Azt láttuk, hogy a „sűrűbb” gráfok (nagyobb bekötő élszámmal (m) növesztett vagy behúzott extra éllel sűrített gráfok) esetén az addig tapasztalt határvonal jobbrább-lefelé tolódik, és az átmenet keskenyedik. A ritkább gráfok esetén pont ezeknek az ellenkezőjét kaptuk. Az élkiszéledésével / behúzásával módosított gráfok vizsgálatából az derült ki, hogy amennyiben a magasabb foksámú csúcsokat preferáljuk a kiválasztásnál¹⁸ akkor az eltolódás mértéke jelentősebb. Végül a határvonal alakját és eltolódását vettük vizsgálat alá. Az átmeneti részen összekötve az azonos kihalási arányhoz tartozó csúcsokat az adott átmenet kontúrvonalait kapjuk. Ezek közül mindig a legmagasabb kihalási arányhoz tartozót vizsgáltuk¹⁹. Ezzel egy kihalási arány képet egy pontsorra szűkítettünk. A pontsor félig logaritmikus ábrán egyenesre illeszkedett, ezért hiperbola függvényt illesztettünk az eredeti pontsorra.

Különböző bekötő élű hálózatokra illesztve ugyanazt a függvényalakot megfigyeltük, hogy a görbék hogyan tolódnak el m változtatásának hatására. A kapott összefüggések mindegyike trendet jelzett.

¹⁸ Annak kiválasztásánál, hogy hova húzzunk be új élt / honnan vegyünk el

¹⁹ Bármelyik vonalat választhattuk volna, a lényeg az, hogy a különböző gráfok esetén mindig ugyanahhoz a kihalási arányhoz tartozót válasszuk.

8. Továbbfejlesztési lehetőségek

A dolgozat elején felmerült az Erdős-Rényi-féle véletlen gráf modell és a gyakorlatban tapasztalt komplex hálózatok között két jelentős különbség, a klaszterezettség és a skála-függetlenség. A dolgozatban a gráfok skálafüggetlenek voltak, de a klaszterezettségük nem számított. Érdekes lehet megvizsgálni a modellnek egy olyan bővített verzióját, amiben a klaszterezettség is biztosított.

A SIS-modellben egy adott csúcs minden élen tud fertőzni egy időlépésben. Ez tovább árnyalható, ha - akár véletlenszerűen - az élek ki-be kapcsolhatóak lennének. Például, ha én influenzás vagyok, hiába van számos ismerősöm, nem tudok mindenkit fertőzni személyesen, csak a környezetemben lévőket. Tehát a valós kapcsolataimnak csupán egy töredéke az, amit egy adott időben tudok fertőzni.

A dolgozatban a kihalás arányát vizsgáltuk meg alaposabban. A folyamatok négy jellemzőjéből a másik hármat²⁰ is meg lehetne vizsgálni abból a szempontból, hogy a hálózat m paraméterének, illetve az új élek behúzásának / meglévők kihúzásának milyen hatása van.

A folyamat jellemzői közül a további három vizsgálatát a dolgozatban tárgyalt módszerek közül az állandósult állapotot megtaláló függvénnyel lehetne megvizsgálni. A kihalások arányával szemben ezeknek a vizsgálatához ezt a függvényt, a korlátait és a gyengéseit alaposabban meg kellene vizsgálni. Hogy mennyi adatpontot igényel az állandósult állapot helyes meghatározásához? Hány pont kell ahhoz, hogy „elég” dominálja az átlagot?

A kihalások arányát bemutató ábrákon a határvonal valójában egy átmenet. A dolgozatban vizsgált szimulációk azt mutatják, hogy a „ritkább” gráfok esetén ez az átmenet szélesebb. Egy külön vizsgálat lehetne annak a felderítése, hogy hogyan jellemezhető és hogyan változik ez az átmenet különböző gráfok esetén.

²⁰ Az állandósult állapot értékét, az állandósult állapot eléréséhez szükséges időt, és az állandósult állapotban a szórást

„ - Negyvenkettő! - ordította Loonkawl. - Ez minden, amit
hét és félmillió év alatt kiagyaltál?

- Viszont gondosan ellenőriztem - mondta a komputer - és
egész egyértelműen ez a Válasz.”

— Douglas Adams

Irodalomjegyzék

- [1] L. Euler. Solutio problematis ad geometriam situs pertinentis. *Commentarii academiae scientiarum Petropolitanae*, 8:128–140, 1741.
- [2] D. König. *Theorie der endlichen und unendlichen Graphen*. AMS Chelsea Publishing, Providence, Rhode Island, 1936.
- [3] S. Milgram. The small world problem. *Psychol. Today*, pages 61–67, 1967.
- [4] G. Prazsák. Network society 2.0, virtuelle gemeinschaft - mobiltársadalomkutatás. *Világosság*, (9):78, 2007.
- [5] D. J. Watts-S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393:440–442, 1998.
- [6] S. Redner. How popular is your paper? an empirical study of the citation distribution. *Eur. Phys. J. B*, 4:131–134, 1998.
- [7] D. J. De Sola Price. Networks of scientific papers. *Science*, 14:510–515, 1965.
- [8] A.-L. Barabási-R. Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.
- [9] A.-L. Barabási. A hálózatok tudománya: a társadalomtól a webig. *Magyar Tudomány*, (11):1298, 2006.
- [10] A.-L. Barabási. *Network Science*. Cambridge University Press, 2015.
- [11] G. Palla. *Komplex hálózatok szerkezetének és dinamikájának feltárása és modellezése statisztikai fizikai módszerekkel*. PhD thesis, 2016.

Nyilatkozat

Név: Kovács-Deák Dániel

ELTE Természettudományi Kar, szak: Fizika BSc.

NEPTUN azonosító: T1M6IS

Szakdolgozat címe: Járványterjedés vizsgálata gráfokon

A szakdolgozat szerzőjeként fegyelmi felelősségem tudatában kijelentem, hogy a dolgozatom önálló munkám eredménye, saját szellemi termékem, abban a hivatkozások és idézések standard szabályait következetesen alkalmaztam, mások által írt részeket a megfelelő idézés nélkül nem használtam fel.

Budapest, 2019. június 22.

a hallgató aláírása