

# Időben inhomogén fertőzés hálózatokon

TDK-DOLGOZAT

Horváth Dávid

Konzulens: Kertész János

Elméleti Fizika Tanszék

# Tartalomjegyzék

<b>1. Bevezetés</b>	<b>2</b>
1.1. Motiváció . . . . .	2
1.2. A modell ismertetése . . . . .	3
1.3. Az alkalmazott valószínűségi eloszlások . . . . .	4
1.3.1. Eseményközidő-eloszlások . . . . .	4
1.3.2. Várakozásiidő-eloszlások . . . . .	7
<b>2. A fontosabb eredmények</b>	<b>8</b>
2.1. Eredmények stacionárius dinamika esetén . . . . .	8
2.1.1. Terjedés Barabási-Albert-hálózatokon . . . . .	8
2.1.2. Terjedés Barabási-Albert-fákon . . . . .	13
2.1.3. Terjedés Cayley-fán . . . . .	17
2.2. Nemstacionárius dinamika esete . . . . .	20
<b>3. Összefoglalás</b>	<b>23</b>
<b>4. Függelék</b>	<b>24</b>

# 1. Bevezetés

## 1.1. Motiváció

A komplex rendszerek és hálózatok tudományának különösen hangsúlyos területe a legkülönbözőbb terjedési, perkolációs jelenségek vizsgálata. Ezen jelenségek fontossága mindenki számára nyilvánvaló, ha a különböző vírusok, járványok terjedésének megértésére és a terjedés előrejelezhetőségének lehetőségére gondolunk. A járványterjedés matematikai modellezésének kiterjedt irodalma van [1], de a hálózatok kutatás újabb fejleményeire támaszkodva sok szép eredmény született a közelmúltban - említést érdemel az Alessandro Vespignani vezette csoport munkája, melyben a kutatók megdöbbentően pontos előrejelzést tudtak adni a H1N1 vírus 2009-es terjedéséről. (A csoport eredményéről lásd az alábbi internetes összefoglalót:[2])

Természetesen nemcsak betegségek terjednek hálózatokon, hiszen számos olyan emberalkotta vagy természetes hálózat van, melyben valaminek a szétáramlása a rendszerhez kapcsolódó folyamatok közül a legfontosabbak közé tartozik. Ennek a TDK-dolgozatnak a fő motivációját elsősorban emberi, kommunikációs hálózatok adják, ahol valamilyen információ, legyen az hír, innováció vagy egyszerűen pletyka elterjedésére vagyunk kíváncsiak. A hírek, innovációk terjedése önmagában is fontos, de talán még inkább az az emberi kommunikációhoz köthető számítógépes, vagy a mobiltelefonokon terjedni képes vírusok terjedése.

Az elmúlt néhány évben a kutatók nagy erőfeszítéseket tettek a hálózatokon zajló folyamatok időfüggésének tanulmányozására. Megállapították, hogy az esetek többségében - és ez különösen igaz a kommunikációs hálózatokra, a dinamika rendkívül inhomogén, villanásos (angolul „bursty”), vagyis aktív periódusok néha nagyon sűrűn követik egymást, máskor pedig igen jelentős időtartamok telnek el közöttük. Az ilyen dinamikának egyik legfontosabb jellemzője az eseményközidő-eloszlás, mely valóságos rendszerekre (legalábbis széles tartományban) nem exponenciális eloszlással, hanem hatványfüggvénnyel írható le [3]. Ennek a villanásos dinamikának a terjedésre gyakorolt hatását a világ számos vezető laboratóriumában vizsgálják.

Szokás egy új képből gondolkodni az ilyen rendszerekről. A kommunikációs hálózatokban, de akár az emberi találkozások epidemiológiai szempontból meghatározó hálózatában valójában kötések csak akkor vannak jelen, amikor a kapcsolat él, vagyis hálózat élei időfüggőek. Könnyű tanulmányozni ezeket „temporális hálózatokat” [4] ha Poisson-folyamatokkal lehet az aktivitást jellemezni. A bennünket érdeklő kérdést tehát meg lehet úgy is fogalmazni, hogy milyen a terjedés a nem-poissoni temporális hálózatokon.

Az eddigi eredmények ellentmondásosak. Megfigyelték, hogy az események közötti idő széles, hatványfüggvény-szerű lecsengése lelassítja a terjedést [5], de vannak ezzel ellentétes eredmények is, elsősorban Esteban Moro és munkatársai részéről [6]. A képet tovább bonyolítja, hogy egy átlagtér-jellegű számítás a terjedés kezdeti szakaszára a hatványfüggvény-eloszlású események közötti időre mindig gyorsítást ad [7].

További érdekes probléma, hogy a modellszámolásoknál általában stacionárius állapotból indulnak ki. Ez Poisson-folyamatnál mindegy, de bármilyen más esetben nem, és főleg

lassan lecsengő eloszlásoknál várható erős effektus. Ugyanakkor, ha az eredményeket össze akarjuk vetni a megfigyelésekkel, tudomásul kell vennünk, hogy - akár a napi biológiai ritmus miatt [8] - nem várható stacionárius állapot.

A komplex hálózatok elméletének egyik fontos felismerése, hogy létezik egyfajta kvalitatív univerzalitás. A rendkívül sokféle rendszert leíró hálózatoknak közös vonásaik vannak. Tudjuk, hogy a fokszámeloszlás az ilyen rendszerekben jó közelítéssel skálafüggetlen, hogy e rendszerek közös sajátossága a kisvilág-tulajdonság, a magas klaszterezés és általában valamilyen moduláris tagozódás is jelen van. Ismeretes, hogy a skálafüggetlen hálózatok különleges perkolációs tulajdonságokkal bírnak [9], várható tehát, hogy a hálózat fokszám-eloszlása jelentősen befolyásolhatja a terjedést, de az is kézenfekvő, hogy a modularitás megléte is komoly hatással van a fertőzés lefolyására. Stratégiánk az, hogy kiválasztunk néhány fontos modellt és azokon vizsgáljuk a terjedési jelenségeket, bízva abban, hogy az említett univerzalitás miatt az így nyert eredményeink szélesebb körben is alkalmazhatók lesznek.

Ebben a dolgozatban olyan kérdésekre keressük a választ, hogy mi a nem-poissoni jellegnek a terjedésre gyakorolt hatásánál tapasztalt ellentmondások oka? Hogyan hat a topológia a terjedésre? Milyen következményei vannak a nemstacionaritásnak? A dolgozat célja tehát a valóságos rendszerekben tapasztalható topológiai elemek és az ilyen rendszerekben lejátszódó időben inhomogén (stacionárius és nemstacionárius) folyamatok terjedési jelenségekre gyakorolt hatásának a bemutatása, alapvetően számítógépes szimulációk eredményeire építve.

A dolgozat felépítése a következő: először bemutatjuk a modellt, majd áttérünk az alkalmazott valószínűségi sűrűségek ismertetésére. Ezután eredményeink leírása következik, amelyeket végül összefoglalunk és diszkutálunk. A gondolatmenet folytonosságának érdekében néhány számítást a függelékben közlünk.

## 1.2. A modell ismertetése

Ennek a TDK-dolgozatnak a célja tehát egy olyan modell eredményeinek bemutatása, mely hálózatokon lejátszódó, időben alapvetően meglehetősen inhomogén folyamatok okozta terjedési jelenségek leírására alkalmas. A modell lényegében egy terjedési jelenségekre általánosan használt SI modell. Az SI modellben a szokásos terminológia szerint fertőzéstől, annak terjedéséről, beteg és egészséges csúcsokról beszélünk, azonban nyilvánvaló, hogy a modell érvényességi tartománya nem korlátozódik kizárólag járványok esetére. Az SI modell lényege, hogy valamilyen statikus hálózaton a hálózat csúcsainak két állapota lehetséges: egészséges (S mint susceptible) és fertőzött (I mint infected). Egy fertőzött csúcs az egészséges szomszédait megfertőzheti, s ha egy csúcs megfertőződött, akkor maga is képes a továbbiakban fertőzni, azonban meggyógyulni egyik csúcs se tud, vagyis ami egyszer megfertőződött, az fertőzött is marad. A legegyszerűbb esetben az, hogy egy fertőzött csúcs milyen eséllyel betegít meg egy szomszédot, valamilyen időegységre eső transzmissziós valószínűséggel jellemezhető. Ilyen esetekben a hálózat szomszédsági mátrixának ismeretében különböző, többnyire numerikus módszerekkel lehetséges jóslatokat mondani a fertőzés lezajlására.

A bemutatandó modell a fenti SI modell kiterjesztése. Képzeljünk el egy kommunikációs hálózatot, melyben a „fertőzés” (információátadás) a szomszédos csúcsok közötti kommunikáció, kapcsolatteremtés eredményeképp jön létre. Legyen tehát egy olyan hálózatunk, melyben a szomszédos csúcsok olyan személyeknek feleltethetők meg, akik pl. rendszeresen beszélnek egymással telefonon. Egy ilyen rendszerben szeretnénk megvizsgálni, hogy valamilyen információ vagy pletyka, hogyan jut el a hálózat többi eleméhez. Úgy tekintjük, hogy a hír csak telefonbeszélgetéseken keresztül juthat el valakihez, és aki birtokában van az információnak, az tovább is adja. A terjedés lényege tehát, hogy ha valakihez eljut a hír, akkor a szomszédaival legközelebb esedékes telefonhívások alkalmával a hír eljut a szomszédokhoz is.

Ahhoz, hogy egy ilyen modellt minél könnyebben tudjunk vizsgálni és szimulálni, különböző feltevésekkel kell élnünk. Feltesszük például, hogy a hívások időtartama elhanyagolható, s az egymást követő telefonhívások közti időkülönbség, az eseményközi idő egy valószínűségi változó, melyet egy megfelelő sűrűségfüggvénnyel jellemezhetünk. Feltételezzük, hogy tetszőleges két szomszédos csúcs esetén ez a függvény megegyezik, és az egyszerűség kedvéért lényegében mindenféle korrelációtól eltekintünk: azaz az egymást követő hívások közötti véletlen idők függetlenek egy adott pár esetén, továbbá feltételezzük, hogy a különböző éleken (tehát olyan személypárok között, ahol legalább az egyik személy különbözik) történő események sincsenek hatással egymásra. A modellt tehát a következő módon lehet összefoglalni. Adott egy statikus hálózat, melynek élein véletlenszerűen kis villanások történnek, minden élen egymástól és a múltbeli eseményektől függetlenül, ugyanolyan valószínűségi eloszlás szerint, s ha egy fertőzött és nem fertőzött csúcs között van villanás, akkor a fertőzetlen csúcs megbetegszik.

Kevés szó esett még arról, hogy milyen hálózatmodellekkel kívánunk dolgozni. A könnyű előállíthatóság miatt a modularitással és klaszterezéssel nem, de kisvilág-tulajdonsággal és skálafüggetlen fokszámeloszlással rendelkező Barabási-Albert-hálózatok szolgálnak az itt bemutatandó eredmények alapjául. Természetesen az ilyen hálózatok nem teljesen hű modelljei a valóságban létező hálózatoknak, de kétségtelenül sokkal közelebb állnak hozzájuk, mint pl. ahogy erre később kitérünk, a [7]-ben alkalmazott teljes gráf.

## 1.3. Az alkalmazott valószínűségi eloszlások

### 1.3.1. Eseményköziidő-eloszlások

A valóságos rendszerekben a kommunikációs folyamatok eseményközi ideje alapvetően hatványfüggvényes, ám hogy ennek a ténynek a jelentősége és szerepe világos legyen, célszerű exponenciális és alsó levágással rendelkező exponenciális függvényekkel is szimulációkat végezni. Ahhoz, hogy a különböző típusú eloszlásokhoz köthető fertőzési dinamikát összehasonlíthassuk, előírjuk, hogy ugyanaz legyen az eseményközi idők  $\tau$ -val jelölt várható értéke. Fontos megjegyezni, hogy a fertőzésre vonatkozó eredmények nyilván függetlenek  $\tau$  választásától, feltéve, hogy az időt következetesen ilyen, azaz  $\tau$  egységekben mérjük, számolások során ezért adott esetben  $\tau$  értékét 1-nek is rögzíthetjük. Az alábbiakban összefoglaljuk a dolgozatban szereplő eloszlásfüggvényekről tudni érdemes ismereteket. Ahogy azt az

imént mondtuk, a hatványfüggvény- más néven Pareto-eloszlások különösen fontosak, hiszen a valóságos dinamika ilyen függvényekkel jellemezhető. Ha az eseményközi idő egy  $\xi_{pow}$  Pareto-eloszlású valószínűségi változó, akkor az ő sűrűségfüggvénye

$$p_{pow}(t) = \begin{cases} t_{min}^\alpha \alpha \frac{1}{t^{\alpha+1}}, & \text{ha } t \geq t_{min} \\ 0, & \text{egyébként} \end{cases} \quad (1.1)$$

alakú.

Ebben a képletben  $t_{min}$  az alsó levágást jelöli, s jelölje továbbá  $\tau$  a  $\xi$  eseményközi idő várható értékét az előbb leírtaknak megfelelően. Ekkor  $\tau$  kifejezhető  $t_{min}$  és  $\alpha$  segítségével az alábbiak szerint:

$$\tau = t_{min} \frac{\alpha}{\alpha - 1} \quad (1.2)$$

Ha osztjuk (1.2)-t  $t_{min}$  -nel és bevezetjük a  $v$  dimenziótlan paramétert mint a  $\tau$  és  $t_{min}$  hányadost, akkor a Pareto-eloszlás  $\alpha$  paraméterét kifejezhetjük (1.3) szerint.

$$\alpha = \frac{v}{v - 1} \quad (1.3)$$

Vagyis  $\tau$ ,  $t_{min}$  és  $\alpha$  paraméter közül tetszőleges kettőből a harmadik nagyon könnyen adódik.

Legyen ezek után  $\xi_{exp}$  egy másik valószínűségi változó, melynek sűrűségfüggvénye exponenciális, vagy eltolt exponenciális. Így ez a valószínűségi változó csak  $t_{min} \geq 0$ -nél nagyobb értékeket vehet fel. Érdeemes most is kifejezni a  $\lambda$  paramétert  $\tau$  és  $t_{min}$  függvényében.

$$c\lambda \int_{t_{min}}^{\infty} e^{-\lambda t} dt = ce^{-\lambda t_{min}} = 1 \quad (1.4)$$

és

$$c\lambda \int_{t_{min}}^{\infty} te^{-\lambda t} dt = ce^{-\lambda t_{min}} \left( t_{min} + \frac{1}{\lambda} \right) = \tau \quad (1.5)$$

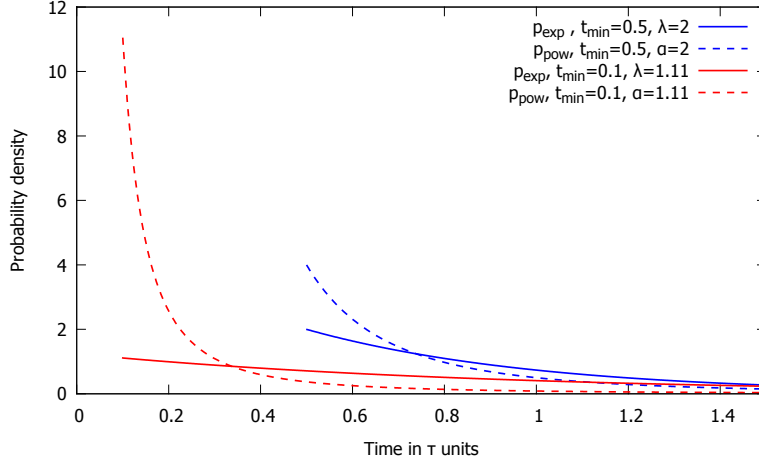
A fenti sorokban, (1.4) biztosítja a megfelelő normát és (1.5) az előírt várható értéket. Ezeket az egyenleteket megoldva a következő adódik:

$$p_{exp}(t) = \begin{cases} \lambda e^{-\lambda(t-t_{min})}, & \text{ha } t \geq t_{min} \\ 0, & \text{egyébként} \end{cases} \quad (1.6)$$

ahol

$$\lambda = \frac{v}{v - 1} \frac{1}{\tau} \quad (1.7)$$

A következő ábrán Pareto- és alsó levágással rendelkező exponenciális sűrűségfüggvények láthatók azonos várható értékkel és különböző alsó levágásokkal.



1. ábra. Pareto és exponenciális valószínűségi sűrűségfüggvények különböző alsó levágásokkal.

Az 1. ábráról látható, hogy hatványfüggvény-eloszlásoknál, amellet, hogy az alsó levágással az  $\alpha$  exponens csökken (1.3)-nak megfelelően, a  $t_{min}$  alsó levágás körüli függvényértékek megnőnek. Csökkentve az alsó levágást és összehasonlítva az azonos levágáshoz tartozó Pareto- és exponenciális függvényeket, a hatványfüggvények jóval konvexebb alakkal bírnak, és  $t_{min}$  körül egy kis intervallumban nagyobb értékeket vesznek fel, mint a módosított exponenciális társaik, majd hirtelen az exponenciális függvény alá kerülnek, s természetesen a végtelenben lassan csengenek le, míg az exponenciálisok exponenciálisan gyorsan. A kis alsó levágás környékén tapasztalható viselkedést támasztja alá a következő becslés is. Írjuk fel (1.1)-et a  $v = \tau/t_{min}$  paraméterrel. Ekkor (1.8)-at kapjuk.

$$p_{pow}(t) = \begin{cases} \left(\frac{\tau}{v}\right)^{\frac{v}{v-1}} \frac{v}{v-1} \frac{1}{t^{\frac{v}{v-1}+1}}, & \text{ha } t \geq \frac{\tau}{v} \\ 0, & \text{egyébként} \end{cases} = \begin{cases} \frac{1}{\tau} \left(\frac{1}{v}\right)^{\frac{v}{v-1}} \frac{v}{v-1} \frac{1}{\frac{t}{\tau}^{\frac{v}{v-1}+1}}, & \text{ha } t \geq \frac{\tau}{v} \\ 0, & \text{egyébként} \end{cases} \quad (1.8)$$

Véve (1.8)-et, látható, hogy  $\int_{\frac{1}{v}}^1 p_{pow}(t) dt = 1 - \left(\frac{1}{v}\right)^{\frac{v}{v-1}} \approx 1$ , ha  $v$  nagy. ( $\tau$ -t vehetjük egynek.)  $p_{pow}(1) = \left(\frac{1}{v}\right)^{\frac{v}{v-1}} \frac{v}{v-1} \approx 0$  és  $p_{pow}\left(\frac{1}{v}\right) = \left(\frac{1}{v}\right)^{\frac{v}{v-1}} \frac{v}{v-1} \frac{1}{\left(\frac{1}{v}\right)^{\frac{v}{v-1}+1}} = \frac{v}{v-1} v \approx v$  ugyancsak nagy  $v$ -re. Mivel a terület közel van 1-hez, világos, hogy  $p_{pow}$  valóban nagyon konvex, ha  $v$  nagy, míg  $p_{exp}$  ekkor lényegében olyan, mint egy levágás nélküli tiszta exponenciális, ami nem változik sokat a  $[0,1]$  intervallumon, így tehát a  $p_{pow}$  függvény  $p_{exp}$  alatt tartózkodik  $[0,1]$  nagy részén.

A kis értékekre megnövekvő valószínűsége-sűrűség hozzájárulhat a terjedés felgyorsulásához. A kép természetesen bonyolultabb; mindenképp szükségünk van a várakozási idők eloszlására.

### 1.3.2. Várakozásiidő-eloszlások

A fertőzési modellünkben, amikor egy csúcs megfertőződik, akkor a szomszédai közötti, időben következő telefonhívásokkal adhatja tovább a fertőzést. Azonban, hogy ezek az események mikor játszódnak le, az a modellben független a fertőződés tényétől. A kérdés tehát az, hogy ha két, A és B szomszéd között van egy eseménysorozat, melyet jellemez egy eseményköziidő-eloszlás, akkor az eseménysorozat egy véletlen időpontjából, amikor is az egyik csúcs megfertőződik valamely másik C szomszéd által, mennyit kell még várni a következő eseményig A és B között. Ha az események közötti idők tiszta exponenciális eloszlásúak, akkor a helyzet egyszerű, hiszen a várakozási idő paradoxonának értelmében ilyenkor ugyanannyit kell várnunk a következő eseményig, mint ha épp a fertőzéskor történt volna egy esemény a két csúcs között, vagyis a várakozási idő eloszlása egy ugyanolyan időállandójú exponenciális, mint az eseményközi idő eloszlása. Azonban ha az eseményközi idő eloszlása nem tiszta exponenciális, akkor a várakozási időre egy másfajta jellegű eloszlás adódik, ezt nevezzük várakozásiidő-eloszlásnak, mely sűrűségfüggvényét az eseményköziidő-eloszlás sűrűségfüggvényéből a következő formulával határozhatjuk meg.

$$p'(t) = \frac{1}{\tau} \int_t^\infty p(t') dt' \quad (1.9)$$

Ezt követően kiszámolhatjuk a kérdéses sűrűségfüggvényeket a számunkra érdekes esetekben. Ha  $p(t)$  hatványfüggvény ( $p(t) = t_{min}^\alpha \alpha \frac{1}{t^{\alpha+1}}$ ), akkor

$$p'_{pow}(t) = \begin{cases} \frac{1}{\tau}, & \text{ha } 0 \leq t \leq t_{min} \\ \frac{1}{\tau} t_{min}^\alpha \frac{1}{t^\alpha}, & \text{ha } t > t_{min} \end{cases} \quad (1.10)$$

(1.10) azért érdekes, mert ha  $\alpha$  nagyobb 1-nél, az eseményközi időnek létezik várható értéke, a várakozási időnek azonban csak akkor, ha  $\alpha$  2-nél is nagyobb, hiszen (1.10)-ben nagy időkre  $1/t'$  az  $\alpha$ -dik hatványon, míg (1.1)-ben  $1/t$  az  $\alpha + 1$ -dik hatványon szerepel. Ennek ellenére ezt a függvényt kell használnunk a szimulációkban, ha az események közti idő Pareto-eloszlású.

Ha  $p(t)$  exponenciális alsó levágással, akkor

$$p'_{exp}(t) = \begin{cases} \frac{1}{\tau}, & \text{ha } 0 \leq t \leq t_{min} \\ \frac{\tau - t_{min}}{\tau} \lambda e^{-\lambda(t - t_{min})}, & \text{ha } t > t_{min} \end{cases} \quad (1.11)$$

Itt  $\lambda$  megegyezik az eseményközi idő időállandójával, azaz (1.7)-nek megfelelően  $\lambda = v/(v - 1) \cdot 1/\tau$ .

Végezetül egy fontos megjegyzéssel zárjuk ezt a fejezetet. Az előbbieken meghatároztuk a várakozási idők eloszlását, melyet arra szeretnénk használni, hogy ha egy adott  $t$  időben megfertőződik egy A csúcs, valamely C csúcs által, akkor meg tudjuk mondani, mennyit kell várni az A-val szomszédos, de fertőzetlen B csúcs A általi fertőzési kísérletéig. A várakozási idő fenti eloszlása azonban csak stacionárius esetben alkalmazható, a várakozási idő általánosan függ attól, hogy a  $t$  időpont előtt mennyi ideje zajlik a kommunikációs



folyamat A és B között [10]. A (1.9) definiálta függvény azonban sokszor jó közelítés nemstacionárius esetben is, ha a folyamat elegendően hosszú idejére tart. A (1.9) függvény használata teljesen elterjedt az ilyen modellekben, és a következő fejezetben is ezen eloszlással adódó eredményeket mutatunk be, azonban ilyenkor lényegében impliciten azt feltételezzük, hogy a terjedést lehetővé tevő dinamikai folyamatok már végtelen ideje zajlanak, másképpen fogalmazva a rendszer stacionárius állapotban van. A későbbiekben vizsgálni fogjuk a nemstacionárius jellegből fakadó jelenségeket is.

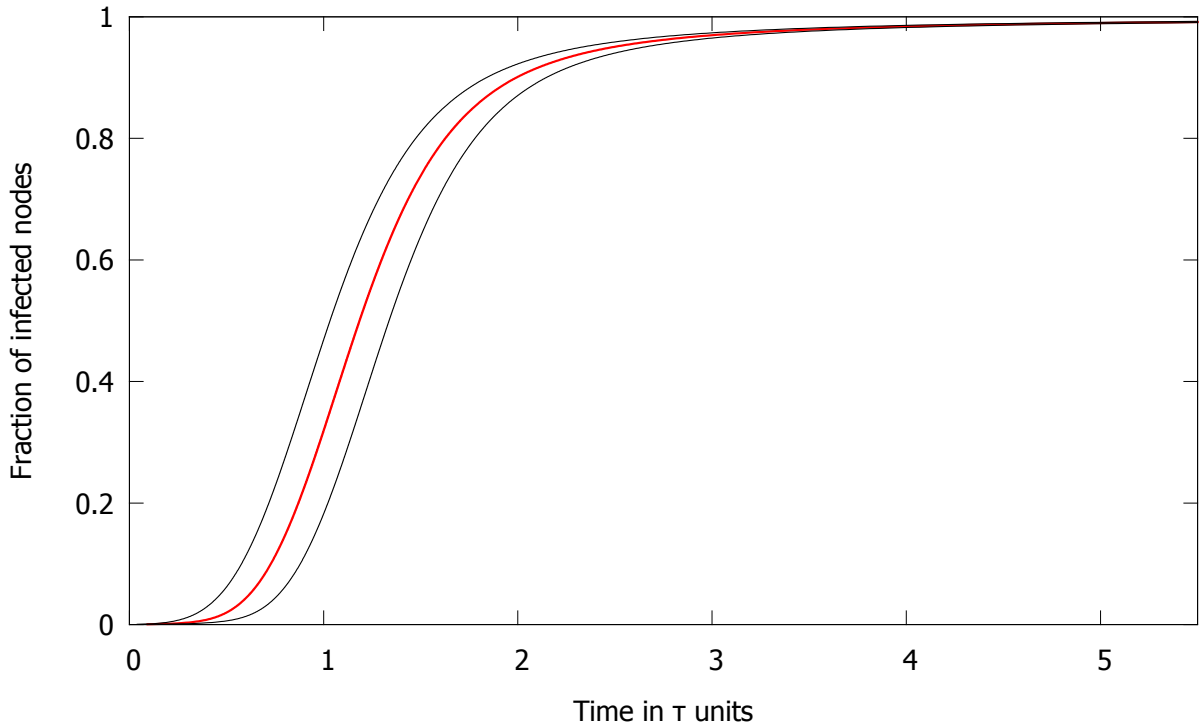
## 2. A fontosabb eredmények

### 2.1. Eredmények stacionárius dinamika esetén

#### 2.1.1. Terjedés Barabási-Albert-hálózatokon

Az előző fejezet zárásaként utaltunk arra, hogy mit értünk stacionárius dinamikán. Ilyenkor használhatjuk az (1.9) által megadott sűrűségfüggvényt a szimulációkban, ebben az alfejezetben ilyen esetekben tapasztalható fertőződés bemutatása a cél. A fertőzés jellemzésére a fertőzött csúcs/összes csúcs mennyiség függését vizsgáljuk az eltelt időtől. Egy ilyen függvény természetesen szimulációs futások átlagolásából adódik. A szimulációk során egy adott fertőződési arányhoz tartozó átlagos idők és azok szórása határozható meg, de az ábrázolás a szokásoknak megfelelően történik, vagyis az alábbi ábrákon az  $x$  tengelyen a nulladik csúcs megfertőzésétől eltelt idő (ezt a csúcstól külsőleg fertőzi meg valami, vagyis egyetlen csúcs fertőzésével indul a folyamat), az  $y$  tengelyen pedig a fertőzött hányad szerepel.

Arról is esett szó a *Bevezető*ben, hogy megfertőzendő hálózatnak skálafüggetlen fokszámeloszlással bíró Barabási-Albert-hálózatokat (BA) használunk, melyek számítógépes generálása könnyű feladat. Az ilyen hálózatok létrehozása során minden lépésben egy új csúcs adódik a rendszerhez, melynek  $m$  új éle kapcsolódik a már meglévő csúcsokhoz a preferált kapcsolódás (preferential attachment) szerint. Az  $m$  paraméter a BA hálózat egy jellemzője. Az  $m = 1$  esetben BA fákat kapunk, míg  $m \geq 2$  esetén egymástól lényegileg nem különböző, csak sűrűbb vagy kevésbé sűrű, másképpen kisebb vagy nagyobb átlagos fokszámmal rendelkező hálózatokat. A BA fákkal, vagyis az  $m = 1$  esettel egy külön alfejezetben még foglalkozunk, a most következő szimulációs eredmények  $m = 2$ , azaz  $\langle k \rangle = 4$  átlagos fokszámú BA hálózatokon születtek egy önkényes  $\alpha$  paraméterű Pareto-eloszlású eseményközi idővel jellemezhető dinamikánál. (Természetesen stacionaritást feltételezve, a szimulációkban az eseményközi eloszlásból származtatható várakozási időket használtuk.)



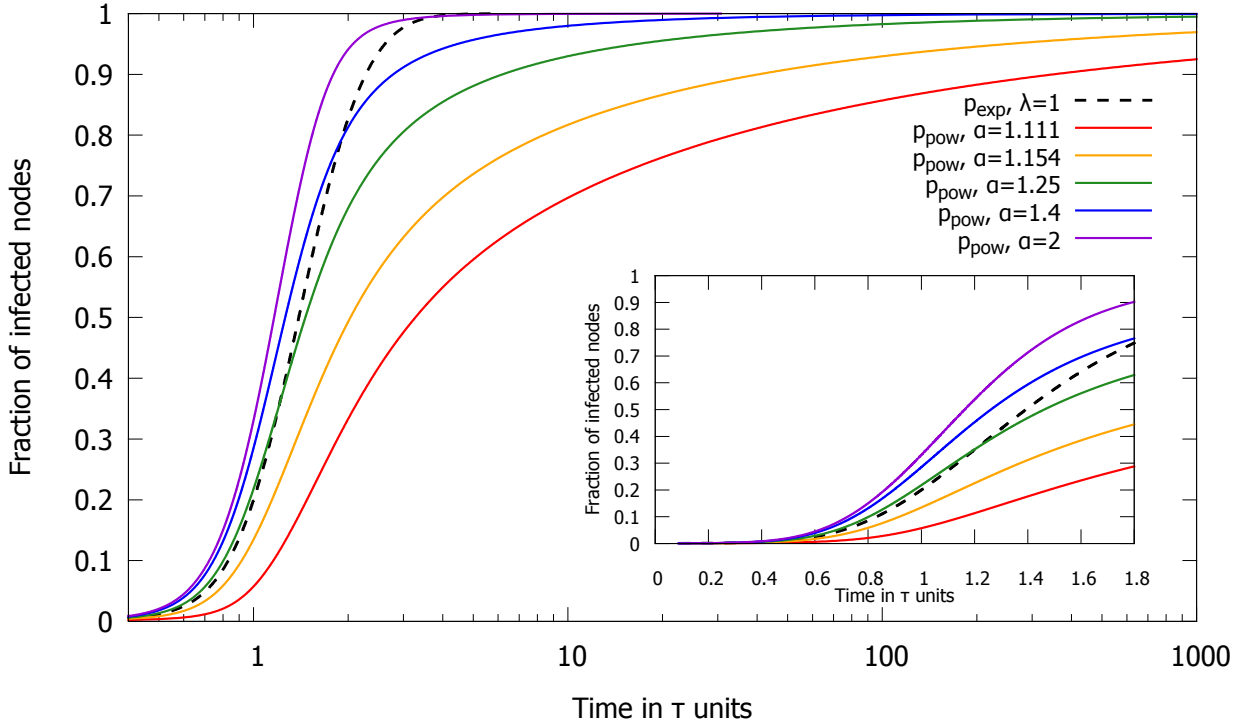
2. ábra. Egy tipikus fertőzési görbe (piros vonal) és a fluktuációk (fekete vonalak)  $\alpha = 1.66$ , azaz  $t_{min} = 0.4$  paraméterű Pareto-eloszlással jellemezhető eseményközi idejű, stacionárius dinamika esetén  $10^4$  csúcsból álló BA hálózaton.  $\langle k \rangle = 4$ , és az elsőnek megfertőzött csúcs  $k_0$  fokszáma 25 és 125 közé esett.  $10^3$  futás átlaga.

A 2. ábrán tehát egy tipikus fertőződés alakulását mutatjuk be. Jól látható, hogy  $\tau$  egységekben gondolkodva a fertőzés alapvetően gyors, hiszen másfél időegység alatt a hálózat majdnem 80 %-a megfertőződik. Látható továbbá, hogy az első néhány fertőzés is igen gyorsan megtörténik, laposan indul a görbe. A megfelelő kezdő csúcsot minden futásnál véletlenszerűen választottuk ki. Az ábráról az is látszik azonban, hogy a gyors kezdeti fertőzést lassuló telítődés követi. A legutolsó csúcs átlagos fertőződése pl. csak kb. a  $400 \tau$  idő után történik meg. A telítődési jelleg oka a kis fokszámú csúcsokban keresendő, amelyek közül nagy eséllyel találunk olyat, amelyhez csak lassan jut el a fertőzés, mivel ezeknek kevés szomszédjuk van. Ez a megfigyelés összhangban van tapasztalati tényekkel is, gondoljunk csak arra, hogy mindig vannak olyan személyek, akikhez valamilyen információ csak nagyon későn jut el. A szórás alakulásáról is érdemes néhány szót szólni, hiszen az a terjedés legnagyobb része során kb. állandónak tekinthető, természetesen a telítési szakaszban gyorsan és meredeken kezd növekedni.

Láttuk tehát, hogy egy hatványfüggvényes dinamikához köthető terjedésnek milyen tulajdonságai vannak, azonban érdemes különböző  $\alpha$  paraméterű skálafüggetlen, és exponenciális eloszlás esetén is összehasonlítani az eredményeket. Ahogy azt a *Bevezető*ben említettük, [7]-ben egy a miénkhez hasonló dinamikájú, de alapvetően teljes gráfra épülő modell vizsgálatára került sor. Ebben a modellben azt követően, hogy egy csúcs megfer-

tőződik, a csúcs minden fertőzési kísérletben véletlenszerűen választ egy másik csúcsot a teljes rendszerből. Az első ilyen választásig itt is a várakozási időnek kell eltelnie, ám azt követően a következő próbálkozásig mindig csak az eseményközi időt kell kivárni. Ebben a modellben a hatványfüggvényes eseményközi idők a terjedés szempontjából gyorsítanak, ez adódik a végtelen rendszerek esetében lehetséges analitikus számolásokból. Amint azt a szimulációs eredmények megerősítik, a korai terjedésre ez a megfigyelés véges rendszerek esetén is fennáll [7]-ben közölt modell esetén.

Fontos kérdés tehát, hogy modellünk, amely nemcsak véges, hanem figyelembe veszi a skálafüggetlen gráftopológiát, mit eredményez. Jóllehet a mi modellünk a fertőzés to-vaterjedését és a csúcsok közötti kommunikációt tekintve nem teljesen azonos a [7]-ben bemutatott modellel, a folyamatok időbeliségét tekintve alapvetően ugyanazon jelenségkör leírására szolgálnak, ezért érdekes tanulsága a mostani összefoglalónak, hogy BA hálózatokon a hatványfüggvényes dinamika lassulást okozhat. Ezt az állítást támasztja alá a következő ábra.



3. ábra. Levágás nélküli exponenciális (szaggatott, fekete vonal) és különböző  $\alpha$  paraméterű Pareto-eloszlásokkal (színes, folytonos vonalak) jellemezhető eseményközi idejű, stacionárius dinamika eredményezte fertőzési görbék  $10^4$  csúcsból álló BA hálózaton.  $\langle k \rangle = 4$ , és az elsőnek megfertőzött csúcs  $k_0$  fokszáma 25 és 125 közé esett. Minden görbe  $10^3$  futás eredménye. Az  $\alpha = 1.111, 1.154, 1.25, 1.4$  és  $2$  exponensek  $t_{min} = 0.1, 0.133, 0.2, 0.286$  és  $0.5$  ( $\tau$  egységekben mért) alsó levágásoknak felelnek meg. A kis ábrán a kis időig történő fertőződés követhető nyomon.

Az előző, 3. ábrán látható, hogy az  $\alpha$  kitevő értékétől függően, a poissoni- dinamikával összevetve a hatványfüggvényes eseményközi idők okozhatnak gyors és lassú terjedést. Az ábrán lévő 2, 1.4 és 1.25-ös exponensek esetén a kezdeti fertőzés gyorsan indul, majd valamilyen fertőzöttségi értéktől az exponenciális folyamatok okozta terjedés válik gyorsabbá. Kellően kis Pareto-kitevőket használva a fertőzési görbék végig az exponenciális dinamikájú görbe alatt maradnak azonban, amit a következő, különböző  $\alpha$ -khoz tartozó első néhány fertőzési időt feltüntető táblázatban bemutatunk.

Fertőzés sorszáma	$p_{exp}, \lambda = 1$	$p_{pow}, \alpha = 1.154$	$p_{pow}, \alpha = 1.111$	$p_{pow}, \alpha = 1.1053$
1	0.02845	0.02598	0.02838	0.03674
2	0.05317	0.04939	0.05262	0.08603
3	0.07464	0.07057	0.07628	0.16674
4	0.09605	0.09156	0.09918	0.24006
5	0.11504	0.11092	0.11985	0.32089
6	0.13244	0.12946	0.1415	0.40059
7	0.1486	0.14607	0.16137	0.48712
8	0.1625	0.16188	0.1802	0.56904
9	0.17591	0.17768	0.19851	0.65487
10	0.1872	0.19196	0.21519	0.73876

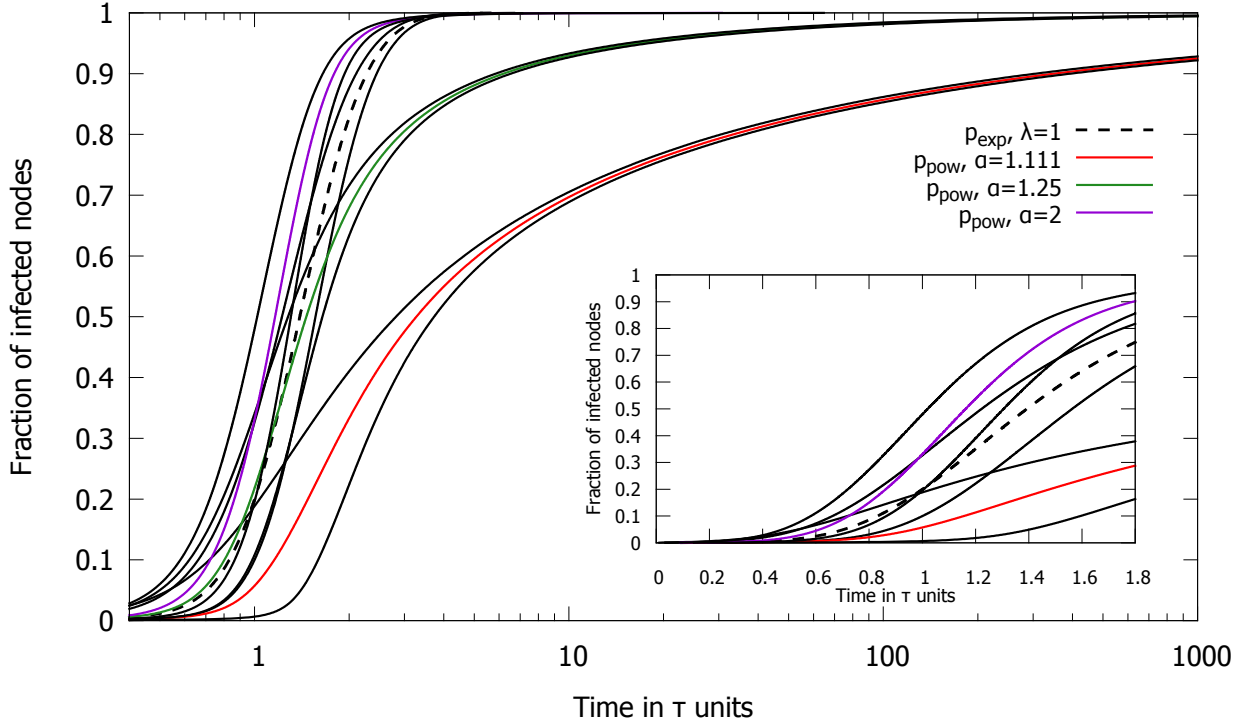
*Az első, második, stb. fertőzés ideje  $\tau$  egységekben különböző eloszlásokkal jellemezhető eseményközi dinamika esetén, stacionárius állapotban.*

A táblázatból látható, hogy a 3. ábrán az  $\alpha = 1.111$ -hez tartozó görbe esetén az első két fertőzés hamarabb következik be, mint az exponenciális esetben, de  $\alpha = 1.053$ -nál már minden fertőzés később történik. Valóban igaz tehát, hogy kellően kis  $\alpha$ -ra a hatványfüggvényes stacionárius dinamika már a kezdetektől lassabb fertőzést idéz elő. Ezt a megfigyelést kicsit precízebbé tehetjük, jöllehet a következő érvelés csak az első fertőzés idejéről mond bármit is. A *Függelék*ben kiszámoljuk az első fertőzés várható idejét a hatványfüggvényes és exponenciális eseményközidő-eloszlásoknak megfelelő várakozásiidő-eloszlások segítségével. Ha a külsőleg megfertőzött csúcshoz  $n$  szomszéd van, akkor a hatványfüggvényes esetre vonatkozó (4.7a) és (4.8a), valamint az exponenciális esetre vonatkozó (4.3a) képletek és 13. ábra szerint található olyan  $\alpha$  melyre a hatványfüggvényes első fertőzés várható ideje kisebb, mint az exponenciális esetben, és olyan kis, de egynél nagyobb  $\alpha$  is, melyre nagyobb ez az idő; de mindez természetesen csak az első fertőzésre vonatkozik. Ugyanakkor úgy tűnik a táblázat és 3. ábra alapján, ha már az első fertőzés is később történik meg, mint az exponenciális esetben, akkor a többi fertőzés is.

Fontos észrevenni, ugyancsak 3. ábra alapján, hogy kis  $\alpha$ -kra a lassulás mértéke is megdöbbenő, hiszen a 3. ábrán a x-tengely logaritmikus léptékű.

A 3. ábrán szimulációs eredmények átlaga látható, melyeknek természetesen szórása is van, melyeket a következő ábrán fel is tüntetünk. A könnyebb átláthatóság kedvéért a 4. ábrán nem tüntetünk fel minden a 3. ábrán látható exponenshez tartozó fertőzési görbét

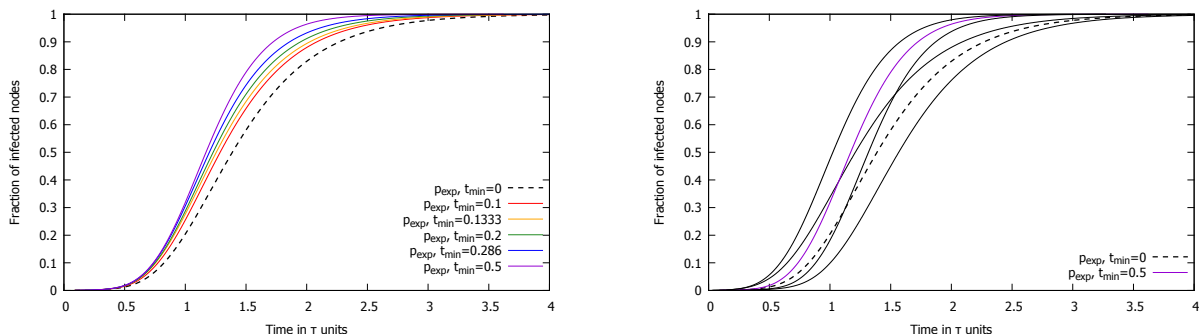
és annak szórását, a szórások jellege és nagysága között nincs lényeges eltérés a különböző (így 4. ábrán szereplő, de 3. ábrán nem szereplő) kitevők tekintetében.



4. ábra. Levágás nélküli exponenciális (szaggatott, fekete vonal) és különböző  $\alpha$  paraméterű Pareto-eloszlásokkal (színes, folytonos vonalak) jellemezhető eseményközi idejű, stacionárius dinamika eredményezte fertőzési görbék szórásaikkal (folytonos, fekete vonalak)  $10^4$  csúcsból álló BA hálózaton.  $\langle k \rangle = 4$ , és az elsőnek megfertőzött csúcs  $k_0$  fokszáma 25 és 125 közé esett. Minden görbe  $10^3$  futás eredménye. Az  $\alpha = 1.111, 1.25$  és 2 exponensek  $t_{min} = 0.1, 0.2$  és  $0.5$  ( $\tau$  egységekben mért) alsó levágásoknak felelnek meg. A kis ábrán a kis időig történő fertőződés követhető nyomon.

A szórásokról, melyek egy adott fertőzöttséghez tartozó idők empirikus szórásnégyzetéből vannak számolva, ebben az esetben is elmondható, hogy a fertőzés során egy széles, fertőzési hányadra vonatkozó intervallumon értékük alig változik, de mint ahogy várható,  $\alpha$  csökkentésével ez az intervallum egyre rövidebb, azaz ahogy szélesedik az eseményköziidő-eloszlás, úgy válik a fertőzés lefolyása egyre hamarabb kissé bizonytalanabbá.  $\alpha = 2$ -re az utolsó néhány tíz (és természetesen az első néhány) fertőzéstől figyelmen kívül hagyva a szórás kb. állandó,  $\alpha = 1.111$ -re már 60 %-os fertőzöttségtől megfigyelhető lassú, majd egyre gyorsuló növekedés a fluktuációkban. A szórások viselkedése egyébként mellett, hogy érdekes, rengeteg szempontból hasznos is, hiszen a fertőzés alakulásáról nagyon fontos információt szolgáltat. Ennek ellenére nem ezen a vonalon haladunk tovább, alapvetően ugyanis a skálafüggetlen eseményeloszlások gyorsító/lassító hatásaira vagyunk most kíváncsiak. Ahogy 3. és 4. ábrák magyarázó szövegében is szerepel, egy adott  $\alpha$  exponenshez

tartozik valamilyen alsó levágás is, így felmerül az igény, hogy olyan eseteket is vizsgáljunk, ahol exponenciális a dinamika, de az eseményközi idők sűrűségfüggvénye az eltoló exponenciális. Ilyen szimulációk eredményeit mutatja be a következő ábraegyüttes.



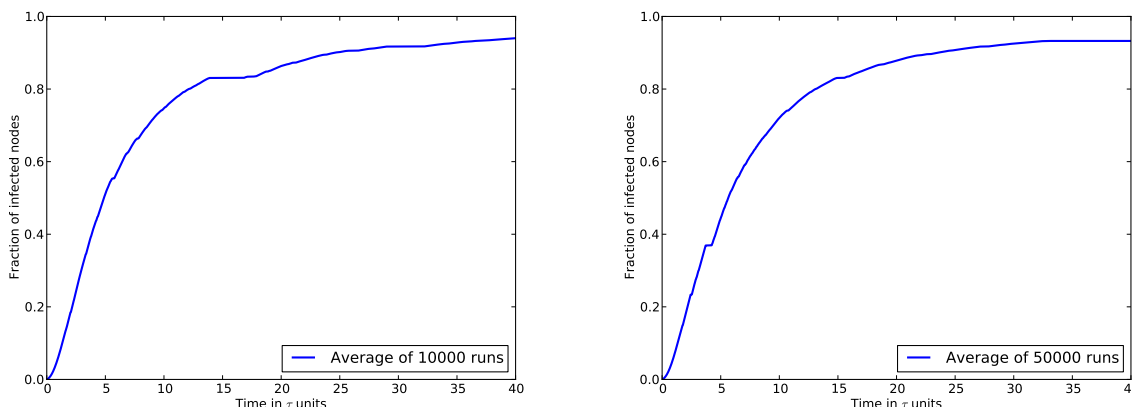
5. ábra. *Levágás nélküli exponenciális (szaggatott, fekete vonal) és különböző alsó levágású exponenciális eloszlásokkal (színes, folytonos vonalak) jellemezhető eseményközi idejű, stacionárius dinamika eredményezte fertőzési görbék, a jobb ábrán a szórásaikkal (folytonos, fekete vonalak)  $10^4$  csúcsból álló BA hálózaton.  $\langle k \rangle = 4$ , és az elsőnek megfertőzött csúcs  $k_0$  fokszáma 25 és 125 közé esett. Minden görbe  $10^3$  futás eredménye. A  $t_{min} = 0.1, 0.133, 0.2, 0.286$  és  $0.5$  ( $\tau$  egységekben mért) alsó levágások  $\lambda = 1.111, 1.154, 1.25, 1.4$  és  $2$  időállandóknak felelnek meg.*

A 3. és 5. ábrákon az azonos színek azonos alsó levágásnak felelnek meg. Az 5. ábráról azt állapíthatjuk meg, hogy exponenciális eseményköziidő-függvényeket használva (pontosabban a belőlük származtatott várakozásiidő-eloszlásokat) egyáltalán nem adódnak olyan markáns effektusok, melyeket a Pareto-eloszlásoknál láthattunk. Az exponenciális alsó levágásának függvényében láthatunk szisztematikus jelenségeket, de ezek erőssége csekély a skálafüggetlen eloszlásoknál tapasztaltakhoz képest. Érdekes, hogy a tiszta exponenciális eseményközi idő esetén kapjuk a leglassabb terjedést, míg a nagy levágás bevezetése gyorsulást okoz, s jóllehet ekkor két esemény egy hálózatbeli élen nem követheti egymást  $t_{min}$ -nél rövidebb időn belül, enyhe gyorsulást tapasztalunk a fertőzésben.

### 2.1.2. Terjedés Barabási-Albert-fákon

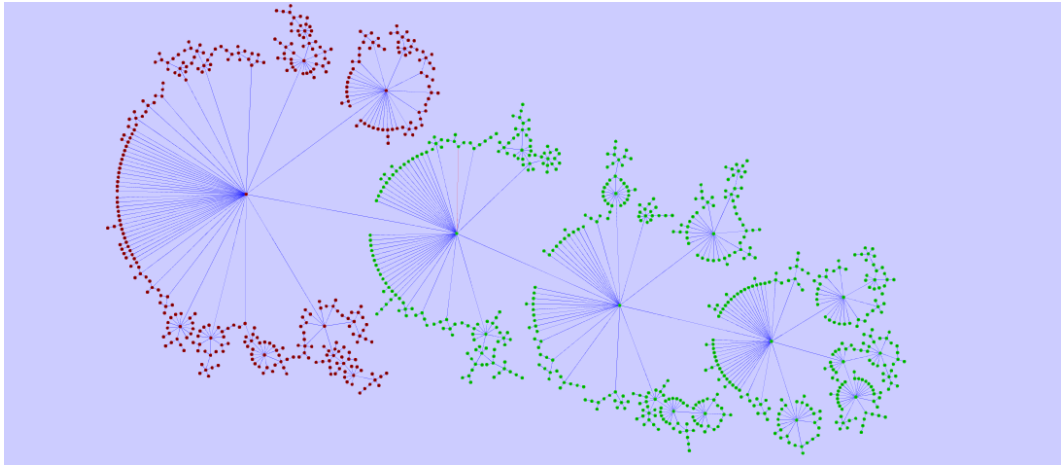
Az előző alfejezetben Barabási-Albert-hálózatokon ( $m = 2$ ) való terjedést vizsgáltunk, és arra jutottunk, hogy összevetve a [7]-beli teljes gráfon kapott és a mi modellünkben látható eredményeket, a hatványfüggvényes eseményköziidő-eloszlás eltérő hatással van a terjedés lefolyásának sebességére különböző modellekben, mely mögött nagyrészt a topológia különbözősége sejthető. A továbbiakban, szintén a topológia fontosságát hangsúlyozva, BA fákon fogjuk megvizsgálni a hatványfüggvényes dinamika néhány sajátosságát. Azt fogjuk látni, hogy lesznek eltérések a BA hálózatoknál látottakhoz képest, ami több szempontból is érdekes. Egyfelől a BA hálózatokra szokás úgy tekinteni, mint lokálisan fára emlékeztető hálózatokra ([11], [12]). Igaz például, hogy  $M$  méretű BA hálózatokban a  $h$  nagyságú

hurkok  $N_h$  száma  $\propto \log(M)^h$  ([13]). Ugyanakkor, mint ahogy fák esetén ez így van, ha két szomszédos csúcs között egyetlen út van, akkor várható, hogy egy hatványfüggvényes, lassan lecsengő eloszlás furcsa hatással lesz a terjedés alakulására. Ezt az állítást támasztja alá a következő ábraegyüttes, a rajta látható eredmények  $\alpha = 2$  exponensű Pareto-eloszlások esetén adódtak, ami azt jelenti, hogy az eseményközi időnek várható értéke véges, de szórása már végtelen, s ennek megfelelően a várakozási időnek nincs véges várható értéke.



6. ábra. Fertőződés  $10^4$  csúcsból álló BA fákön ( $m = 1$ ),  $\alpha = 2$  esetén.  $10^4$  majd további  $4 \cdot 10^4$  (tehát összesen  $5 \cdot 10^4$ ) futás eredménye. Minden futáshoz új fa generálódott, az elsőnek megfertőzött csúcs a legnagyobb fokszámú csúcs volt minden esetben.

A 6. ábrán az látható, hogy a fertőzési görbe nem sima, ugrások vannak benne. Ezzel kapcsolatban két kérdés merülhet fel, nevezetesen, hogy mi ennek az oka, és hogy az átlagolások számának növelésével eltüntethetők-e ezek a lépcsők. A 6. ábrán látható két fertőzési görbe ugyanahhoz a futássorozathoz tartozik, de a bal oldali 10000, a jobb oldali 50000 futás átlagolásával adódik, s megfigyelhetjük, hogy a 10000 átlagolás görbéjén látható kb. 0.8-as fertőzöttségi szintnél lévő ugrás ugyan kisebb mértékben van jelen az 50000-es átlagban, de itt megjelenik máshol egy másik lépcső, kb. 0.4-es fertőzöttségnél. Várható tehát, és ezt később valamivel precízebben is megindokoljuk, hogy a lépcsők nem tűnnek, vagyis igazából nincs jól definiált átlag, legalábbis  $\alpha = 2$ -re. A lépcsők jelenléte részben azzal magyarázható, hogy fákön vizsgáljuk a terjedést, ahol tetszőleges két szomszéd között egyetlen út lehetséges, s mivel a várakozási időnek a várható értéke végtelen (ez a másik összetevője a jelenségnek), előfordulhat, hogy olyan sokat kell várni a következő fertőzésre és adott esetben ezzel a fertőzés folytatódására, hogy nem adódnak értelmes átlagok. A szimulációk segítségével természetesen megmondható, hogy mi az oka az ugrásoknak, s valóban az derül ki, hogy mindezt egyedi futások okozzák, melyekben a fertőzés nagyon jellegzetesen alakul, amit a következő, a könnyebb megjeleníthetőség végett  $10^3$  csúcsból álló BA hálózaton történő, e jellegzetességet hordozó futásból származtatott ábrán mutatunk be.



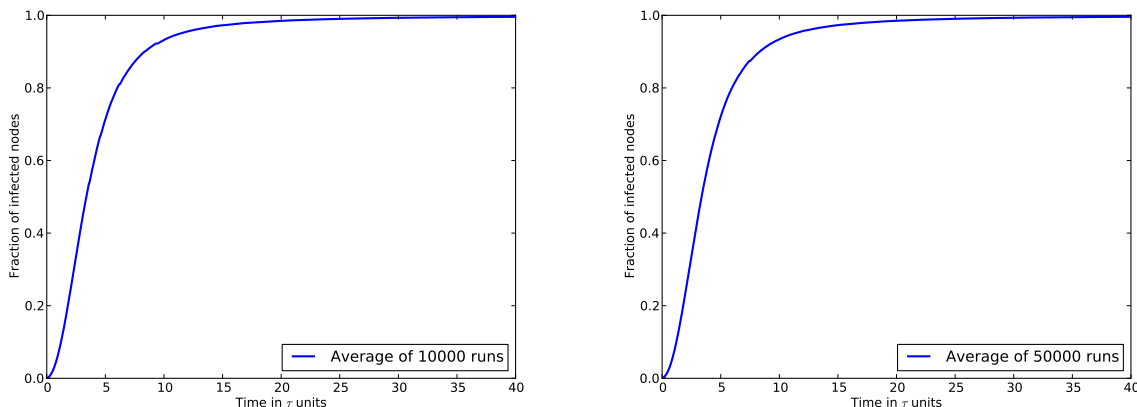
7. ábra.  $10^3$  csúcsból álló BA fa fertőződése.  $t \ll t_{jump}$  ideig megfertőződött csúcsok pirosak, a többi csúcs zöld. A következő fertőzés a piros és zöld csúcsok közötti egyetlen él mentén történik  $t_{jump} = 5129,42 \tau$ -kor.

A 7. ábrán az látható, hogy a fa egy része teljesen megfertőződik valamennyi idő alatt, s ezt követően a fertőzésnek egyetlen él mentén kell továbbhaladnia, az ehhez szükséges idő viszont lehet olyan nagy, hogy ez ugrásokat okoz az átlagokban. Amint az mondtuk, valóban ez a jelenség okozza a lépcsőket a 6. ábrán, erről könnyen meg lehetett győződni, egyszerűen meg kellett keresni a problémás futásokat és az ugráshoz tartozó időket, s minden esetben az adódott, hogy a hálózat megfertőződött és később fertőzendő részét egyetlen él köti össze melyen nagyon nagy várakozási idő sorsolódott, míg a korábban megfertőződött részben a fertőzési idők nagyjából az „átlagnak” megfelelő időkből történtek. Látható tehát, hogy a BA fákon való terjedésben a „szűk keresztmetszet” (bottleneck) effektus révén hogy jelennek meg lépcsőket okozó extrém értékek. A kérdés most már csak az, hogy ezektől az ugrásoktól meg lehet-e szabadulni, ha kellően sok átlagolást végzünk, ezért egy elnagyolt, de a lényegét nagyon jól tükröző érvelést mutatunk be.

Legyen egyelőre az  $\alpha$  értéke továbbra is kettő, s a hálózat mérete legyen egy véges  $M$ , ahogy itt a szimulációban is, ahol  $M = 10000$ . Véges BA fák véges méretű, körszerű modulokból állnak, ahogy azt a 7. ábrán is láthatjuk, melyek mérete legfeljebb  $M$ , de feltesszük, hogy mindig van néhány ilyen modul. (Ez helyénvaló, mivel a legnagyobb fokszámú csúcsnak átlagosan  $\sqrt{M}$  szomszédja van, s mi a legnagyobb fokszámú csúcsból indítottuk a fertőzést, jóllehet ez a tény nem korlátozza a gondolatmenet érvényességét.) Ahhoz, hogy  $N$  futást végezve a szimulációkban megfigyelthez hasonló ugrásokat tapasztaljunk a következő szükséges: fertőződjön meg legalább egy modul minden csúcsa, s a következő moduláris egységben a fertőzés  $\tau$  egységekben mérve  $a \cdot N$  idő múlva történjék, ahol  $a > 1$ , hiszen a fertőzési folyamatok jellemző időskáláját  $\tau$  határozza meg, s ilyen esetben az átlagban pontosan  $\tau$  nagyságrendű ugrást fogunk látni. Legyen a modul mérete, melyből az ugrás származik, mondjuk  $M/2$ , ekkor lényegében akkor látunk ugrást  $N$  átlagolás után, ha  $NM/2$  várakozási időnek megfelelő valószínűségi változóból lesz legalább egy, melynek értéke nagyobb, mint  $aN$ .  $\alpha = 2$ -re ennek valószínűsége  $\frac{1}{4} \cdot \frac{1}{aN}$ ,  $\alpha \neq 2$ -re pedig  $k \cdot \frac{1}{(aN)^{\alpha-1}}$



gyel egyenlő (k valamilyen  $\alpha$ -függő konstans), vagyis annak a valószínűsége, hogy  $N$  futás során nem látunk ugrást okozó futást  $\lim_{n \rightarrow \infty} \left(1 - \frac{k}{N^{\alpha-1}}\right)^{\frac{NM}{2}}$ -gyel egyenlő,  $a = 1$ -re. Ez a határérték rögzített  $M$ -re 0, ha  $\alpha > 2$ ; és 1, ha  $\alpha < 2$ ,  $\alpha = 2$ -re pedig  $e^{-\frac{M}{8}}$ . A gondolatmenetből az következik tehát, hogy  $\alpha > 2$ -re eltüntethetők az ugrások, egyébként nem. Az érvelést némiképp finomíthatjuk, ha most azt is feltételezzük, hogy a korábban (tehát az ugrást okozó él fertőződése előtt) megfertőződött csúcsok biztosan nem okoznak ugrást, legalábbis látható ugrást; ez teljesül akkor, ha a megfertőződésükhöz tartozó idők kisebbek, mint  $b \cdot \frac{N}{C}$ , ahol  $b \in [1, 10]$  és  $C > 10^3$ . Így a lépcsőt okozó futás valószínűsége  $\alpha = 2$ -re  $\frac{1}{4a \cdot N} \left(1 - \frac{C}{4b} \frac{1}{N}\right)^{N \frac{M}{N^2}}$ , ami  $M$  véges volta miatt asszimptotikusan  $\frac{1}{4a \cdot N}$ -nel közelíthető, vagyis visszajutunk az előző gondolatmenethez. Az érvelés szerint  $\alpha > 2$ -re sima görbét kell kapni, a következő ábra ezt a várakozást igazolja,  $\alpha = 2.5$  esetén mutatva, hogy ekkor valóban sima átlag adódik.

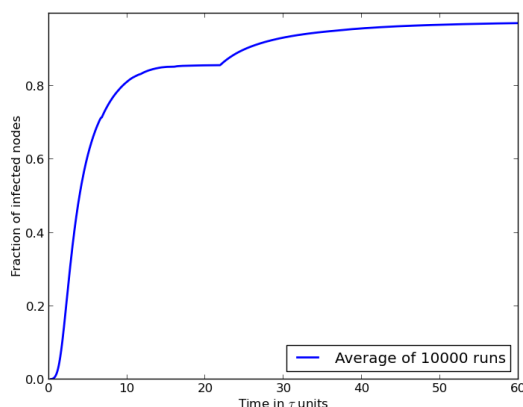


8. ábra. Fertőződés  $10^4$  csúcsból álló BA fákon ( $m = 1$ ),  $\alpha = 2.5$  esetén.  $10^4$  majd további  $4 \cdot 10^4$  (tehát összesen  $5 \cdot 10^4$ ) futás eredménye. Minden futáshoz új fa generálódott, az elsőnek megfertőzött csúcs a legnagyobb fokszámú csúcs volt minden esetben.

Az itt látott jelenség  $\alpha$  exponenstól függő léte vagy nem léte nem csak egy egyszerű modellben tapasztalható érdekesség lehet. A BA fáknak, mint ahogy azt a 7. ábrán is láthatjuk, a BA hálózatoktól eltérően van moduláris felépítése, ami megfelel a fák ágaiból adódó struktúráknak. Jóllehet, itt minden modul legfeljebb egy élen keresztül kapcsolódhat más modulokkal, az átlagban látott ugrások, vagyis a terjedés extrém ideig történő stagnálása, de legalábbis a fertőzés lassulásának lehetősége meg kell hogy jelenjen a valószínű, általában modularitással bíró hálózatokban is. Mi azonban nem ebben a kétségtelenül lényeges irányban haladunk tovább, hanem rátérünk a nemstacionaritás ugyancsak érdekes kérdésére. Előtte azonban megvizsgáljuk a fákon való terjedés egy analitikusan kezelhető esetét is.

### 2.1.3. Terjedés Cayley-fán

A BA fa vizsgálatával láthattuk, hogy a fastruktúra és a kellően lassú lecséngésű valószínűségi függvények okozhatnak extrém terjedést. Ez részben tehát az időbeli folyamatokat jellemző eloszlásból, részben pedig a fa fa-jellegéből következik. Ebben az alfejezetben a Cayley-fát fogjuk röviden tanulmányozni, ellenőrizve, hogy a fa-jelleg ebben az esetben is okozhat-e ugrásokat a terjedési görbe átlagában, a választás oka pedig egy analitikus számolás elvégezhetőségének a lehetősége. A Cayley-fa egy olyan objektum, mely létrehozása során kiindulunk egy csúcsból, amihez  $m$  szomszédos csúcsot rendelünk ( $m$ -et koordinációs számnak is szokás nevezni), majd minden szomszédhoz további  $m - 1$  szomszédot úgy, hogy ezek a hálózat meglévő részéhez kizárólag ezen az egy csúcson keresztül kapcsolódnak. Ezt az eljárást végtelenszer megismételve egy olyan végtelen fát kapunk, melynek minden csúcsának pontosan  $m$  szomszédja van. Természetesen szimulációt csak véges fákon lehet végezni, a következő ábra egy  $m = 7$  paraméterű,  $M = 10886$  méretű Cayley fán,  $\alpha = 2$  esetén, 10000 átlagolás után kapott terjedési görbe látható, a külsőleg megfertőzött csúcs minden esetben a fa legidősebb csúcsa volt.



9. ábra. Fertőződés 10886 csúcsból álló Cayley-fán ( $m = 7$ ),  $\alpha = 2$  esetén.  $10^4$  futás eredménye. Az elsőnek megfertőzött csúcs a legidősebb csúcs volt minden esetben.

A várakozásainknak megfelelően itt is adódik ugrás az átlagban, mely olyan futásból származhat, amikor egy él mentén nagyon nagy várakozási idő sorsolódott és közben a hálózat elérhető része megfertőződött. Nyilvánvalóan  $\alpha$ -tól az előbb leírt módon függ, hogy lesz-e értelmes átlag vagy nem, a következőkben azonban egy egzakt számolás során megmutatjuk, hogy exponenciális dinamika esetén valóban sima átlag kell, hogy adódjon, egyúttal meg is határozzuk a fertőzési görbe átlagát. A számolás alapja a Függelékben kiszámolt mennyiség, amire a BA hálózatok tanulmányozása során már hivatkoztunk. A Függelékben kiszámoljuk, hogy egy  $k_0 = n$  szomszédal rendelkező csúcs esetén mi az első fertőzés várható ideje, ami a csúcs és szomszédai között történik. Ha az eseményközi, s így a várakozási idők eloszlása tiszta exponenciális, akkor az eredmény (4.3a) szerint  $\frac{\tau}{n}$ . Képzeljük most el, hogy van egy hálózatunk, ahol  $t = 0$ -ban külsőleg megfertőződik egy A

csúcs, melynek  $n$  szomszédja van. Ekkor az előbbiek szerint a következő fertőzés várható ideje  $\frac{\tau}{n}$  exponenciális esetben. Ha valami oknál fogva biztosan tudjuk, hogy miután A megfertőz egy B csúcsot s B további  $n'$  csúcs fertőzhet meg, akkor meg tudjuk mondani a következő fertőzés várható idejét is. Ha B  $t$  időpontban fertőződik meg, akkor annak a valószínűsége, hogy A és egy nem B csúcs között fertőződési lehetőség, azaz „telefonhívás” történik  $t + t'$  időpontban, feltéve, hogy  $t$ -ig nem történt esemény, az exponenciális eloszlás örökifjúsága miatt ugyanannyi, mint annak a valószínűsége, hogy  $t'$  múlva lesz esemény. Vagyis amikor B megfertőződik, a folyamat újjászületik, azaz  $n + n' - 1$  független (a függetlenség könnyen ellenőrizhető), exponenciális eloszlású valószínűségi változó minimuma határozza meg a következő fertőzés várható idejét (ami egy feltételes várható érték, de ez funkcionálisan nem függ a feltételt képező előző fertőzési időponttól), ami ennek megfelelően  $\frac{\tau}{n+n'-1}$ . Mindezt általánosítva azt mondhatjuk, hogy a következő fertőzés várható idejét minden lépésben az határozza meg, hogy hány olyan él van, mely fertőzött és nem fertőzött csúcsok között található.

Létezik olyan hálózat, melyben egy ilyen fertőzési lánc kialakulása közben a megfelelő élek száma egyértelműen megmondható. A Cayley-fa pontosan ilyen tulajdonságokkal bír, s bár ez a fa az előző mondat alapján túlságosan triviális topológiájúnak tűnhet, sok szempontból egyáltalán nem az, ugyanis fraktálszerű képződménynek, perkolációt tekintve pedig végtelen dimenziós rácsnak tekintendő [9]. Most nézzük meg, hogy az általunk vizsgált fertőzés esetén hogyan alakul a fertőzőképes élek száma Cayley-fán. Amikor kiválasztunk egy csúcsot, melyet kívülről megfertőzünk, akkor először  $m$  fertőzőképes él lesz. Ezt követően minden fertőzéssel  $m - 1$  új él fertőzhet és minden fertőzéssel pontosan egy él levonódik; egy, hiszen ha egy csúcs megfertőz egy másikat a köztük lévő élen, akkor ezen az élen több fertőződés már nem történhet, és pontosan egy, mert a hálózat egy fa, vagyis minden szomszéd között egyetlen út van. A következő táblázat a fertőzni képes élek számát mutatja.

i. fertőzés	élek száma az i. fertőzés előtt
0	0
1	$m - 1 + 1$
2	$2m - 2 + 0$
3	$3m - 3 - 1$
4	$4m - 4 - 2$
5	$5m - 5 - 3$

*Fertőzni képes élek száma az  $i$ . fertőzésig,  $m$  koordinációs számú Cayley-fán. A nulladik fertőzés a külsőleg megbetegített csúcsra utal.*

Azt mondhatjuk tehát, hogy az  $i$ . fertőzés előtt az aktív élek száma  $im - 2(i - 1)$ ,  $i \geq 1$ . A  $t(i)$   $i$ . fertőzés várható idejére így a következő összefüggés írható fel.

$$t(i) = \sum_{j=1}^i \frac{1}{jm - 2(j-1)} \quad (2.1)$$

melynek megoldása

$$t(i) = \frac{\psi^{(0)}\left(\frac{i(m-2)+m}{m-2}\right) - \psi^{(0)}\left(\frac{m}{m-2}\right)}{m-2} \quad (2.2)$$

ahol  $\psi^{(0)}(x) = \frac{d}{dx} \ln(\Gamma(x))$ . Megkísérelhetünk egy differenciálegyenletet is felírni a fertőzési időkre, jóllehet ennek a közelítésnek nyilvánvaló korlátai lesznek. Legyen ehhez egy véges Cayley-fánk, mely  $M$  csúcsból áll és jelölje  $n$  a fertőzött csúcsok és összes csúcsok hányadosát. Ekkor  $dn \frac{1}{M}$  lépésekben változhat,  $dt$  pedig  $\frac{\tau}{nMm-2(Mn-1)}$ -ként írható. Így

$$\frac{dn}{dt} = \frac{1}{\tau} \left( n(m-2) + \frac{2}{M} \right) \quad (2.3)$$

írható fel, melynek megoldása  $t = 0$ -ban külsőleg megfertőzve egy csúcsot, azaz  $n(0) = \frac{1}{M}$  esetén

$$n(t) = \frac{m}{m-2} \frac{1}{M} e^{\frac{m-2}{\tau} t} - \frac{2}{m-2} \frac{1}{M}. \quad (2.4)$$

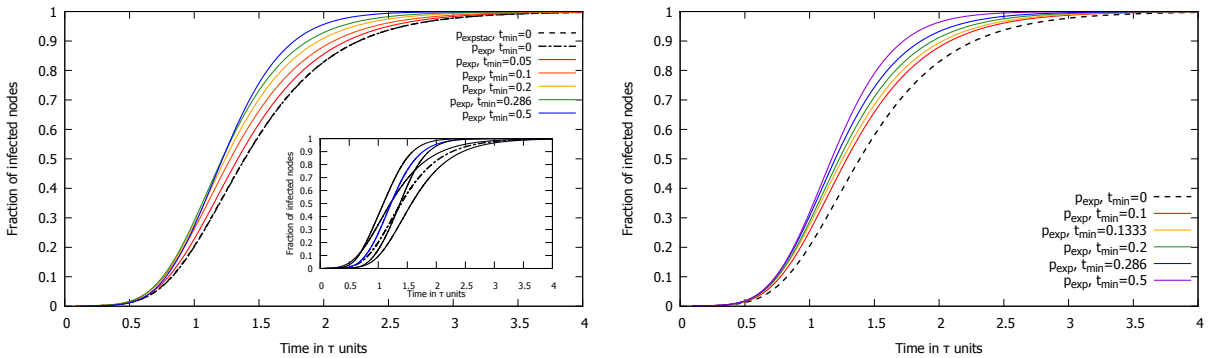
$M \rightarrow \infty$ -re az az eredmény azonban nem értelmes, így csak véges méretekre alkalmazható. A Cayley-fa egyik sajátossága, hogy véges méreteknél a csúcsok makroszkopikus hányada levél, levelek esetén pedig nem igaz az élek számának alakulására felírt összefüggés, vagyis (2.4) csak  $n \ll 1$  esetben lehet érvényes, míg (2.2) általános érvényű, végtelen fákra. Amiért érdemes megvizsgálni (2.4)-t, annak az az oka, hogy így nyilvánvalóan folytonos görbe adódik; van azonban egy másik érdekesség is, ugyanis az angol irodalomban degree-based approximation-nek nevezett közelítő módszerrel lényegében egyező eredményt kaptunk [14]. E közelítés azt feltételezi, hogy terjedés és fertőzés szempontjából egy csúcs fertőzési képessége kizárólag annak fokszámától függ. Ez a feltevés végtelen Cayley-fára természetesen helytálló. A módszer a kezdeti, kis időkhöz tartozó fertőzésekre ugyancsak exponenciális növekedést jósol, ahol a növekedés ütemét egy a mi modellünkben  $\frac{1}{\tau}$ -nak megfeleltethető tag és a  $\frac{c_2}{c_1} - 1$  kifejezés szorzata határozza meg, ahol  $c_1$  az átlagos fokszám,  $c_2$  pedig a másodsomszédok átlagos száma. Cayley-fára  $\frac{c_2}{c_1} - 1$  pontosan  $m - 2$ -nek adódik, ahogy (2.4)-ben is. A Cayley-fára vonatkozó számolással azt igazoltuk, hogy kellően gyorsan lecsengő eloszlás esetén a fán való terjedés jól átlagolható fertőzési hányadot ad, míg numerikusan itt is láttuk, hogy a lassú lecsengés el nem tüntethető lépcsőkhöz vezet.

Ennek az alfejezetnek a legfontosabb mondanivalója a BA hálózatok esetén látott hatványfüggvényes dinamikához köthető lassulás, és fák tekintetében a megfelelő exponensű, ugyancsak hatványfüggvények jelenléte esetén fellépő, a folyamatok időbeliségéből és strukturális sajátosságokból adódó extrém terjedési jelenségek fellépése. Látható tehát, hogy a topológiai és időbeli sajátosságok együttesen milyen hatásokkal bírnak. A következő részben a nemstacionárius jellegből adódó jelenségeket fogjuk vizsgálni.

## 2.2. Nemstacionárius dinamika esete

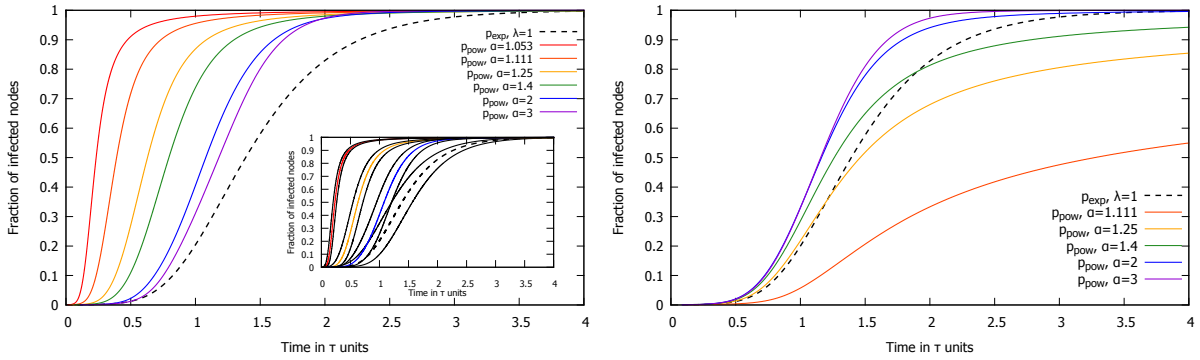
A *Bevezető*ben említettük, hogy a fertőzés terjedését lehetővé tevő folyamatok egyik legfontosabb jellemzője az eseményközi idők eloszlása, mely eloszlásból származtattuk a várakozási idők eloszlását. Eddig kizárólag a stacionárius esettel foglalkoztunk, amikor a várakozásiidő-eloszlást a (1.9) határozza meg. Ezt úgy is fel lehet fogni, hogy az éleken lévő kis órák, melyek a továbbadás lehetőségét vagy másképp fogalmazva a villanásokat megmondják, már végtelen ideje működnek. Egy valóságos telefonhálózat esetén e feltételezés alkalmazhatósága már a napi periodicitás miatt is megkérdőjelezhető, ezért most megvizsgálunk olyan eseteket, amikor a hálózat nincs stacionárius állapotban, és így a várakozásiidő-eloszlásra vonatkozó formulát nem használhatjuk.

Képzeljük most el, hogy a külsőleg megfertőzött csúc fertőzése ezúttal is  $t = 0$ -kor történik, de a hálózat élein lévő órák ki voltak kapcsolva, és bekapcsolásuk, mely jelentősen számunkra egy esemény létrejöttét, a  $[-\xi - t_0, -t_0]$  időintervallumban történt, ahol  $\xi$  egyenletes eloszlású valószínűségi változó a  $[0, \tau]$  intervallumon,  $t_0$  pedig egy rögzített, nemnegatív szám. Ha tehát tudjuk, hogy egy élen mikor volt az első esemény, mely tehát minden él esetén egy véletlen és egy rögzített szám összege, akkor az eseményközi idők sorsolásával és összeadásával megmondhatjuk, hogy egy adott  $t$  időt követően mikor lesz a következő esemény, vagyis nincs szükségünk a várakozási időkre. Nézzük tehát meg először különböző exponenciális, majd Pareto-eloszlások esetén, hogy hogyan alakulnak a terjedési görbék, ha  $t_0$  értékét 0-nak választjuk meg, vagyis az órák beüzemelését követően rövid időn belül megindul a fertőzés. A következő ábrán az előbb leírt módon (bal ábra) és a hagyományos, várakozási idős módszerrel (jobb ábra) kapott szimulációs eredmények láthatók levágással rendelkező és tiszta exponenciális eloszlások esetén.



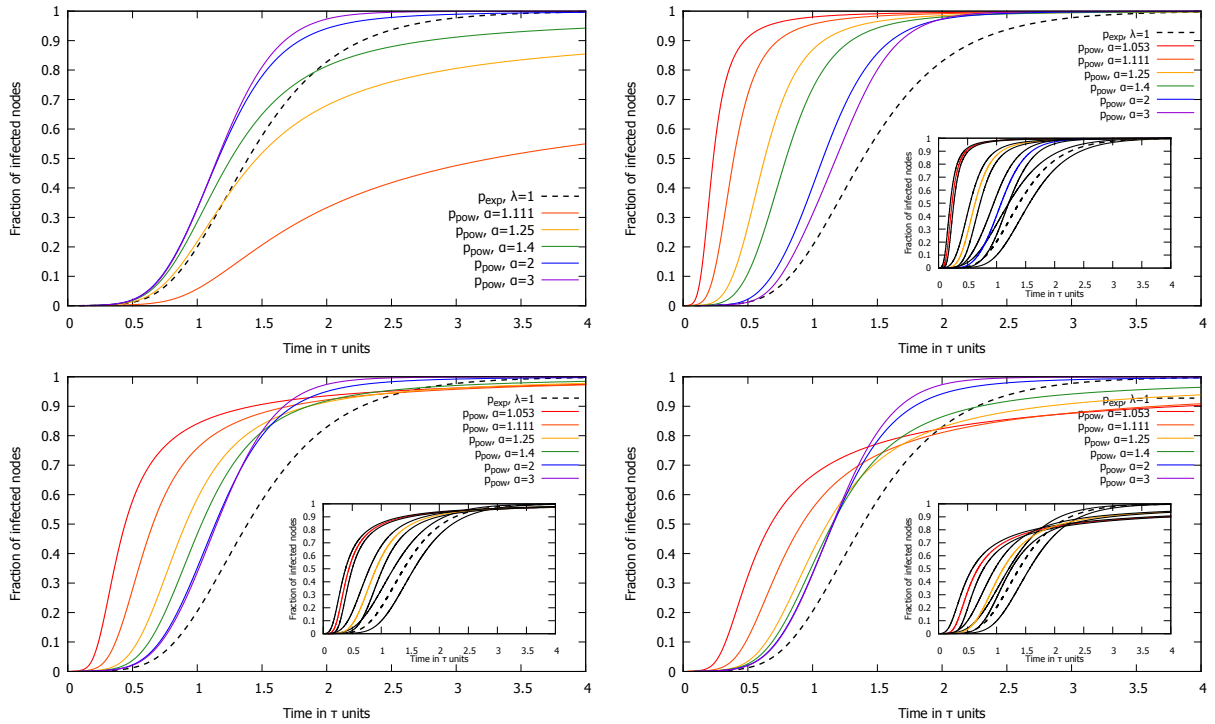
10. ábra. Levágás nélküli exponenciális (szaggatott, fekete vonal) és különböző alsó levágású exponenciális eloszlású eseményközi idejű dinamika (színes, folytonos vonalak) eredményezte fertőzési görbék nemstacionárius (bal alábra) és stacionárius (jobb alábra) esetben, (a bal alábra sarkában néhány görbe a folytonos, fekete vonalakkal ábrázolt szórásával látható)  $10^4$  csúcsból álló BA hálózaton.  $\langle k \rangle = 4$ , és az elsőnek megfertőzött csúc  $k_0$  fokszáma 25 és 125 közé esett. Minden görbe  $10^3$  futás eredménye. A  $t_{\text{min}} = 0.05, 0.1, 0.133, 0.2, 0.286$  és  $0.5$  ( $\tau$  egységekben mért) alsó levágások  $\lambda = 1.053, 1.111, 1.154, 1.25, 1.4$  és  $2$  időállandóknak felelnek meg.

A fenti, 10. ábra bal alábráján az exponenciális kommunikációs folyamatokhoz két görbe is tartozik, egyik esetben a várakozási idő eloszlását használva ( $p_{expstac}$ ), ami a jobb alábrán is szerepel; a másik esetben pedig az előbb bemutatott nemstacionárius esetben végezve a szimulációt ( $p_{exp}$ ), láthatóan nincs érdemi különbség a fertőzési görbék között tiszta exponenciális esetben, ahogy az várható is, hiszen a tiszta exponenciális eloszlásnak nincs memóriája. Az is leolvasható a 10. ábráról, hogy a levágással rendelkező exponenciális eloszlások esetén sincs lényegi szerepe annak, hogy várakozási időket használunk, vagy pedig minden órát a  $[-\tau, 0]$  időben kapcsolunk be és így származtatjuk a folyamatot; ugyanaz, a levágás növelésével kissé gyorsuló terjedés figyelhető meg. Egészen más a helyzet azonban, ha a dinamikához köthető eseményközi idők Pareto-eloszlásúak, amint az a következő ábrán is látható, ahol a nemstacionárius esetekben  $t_0$  értéke ismét nulla volt.



11. ábra. Bal ábra: Levágás nélküli exponenciális (szaggatott, fekete vonal) eseményközi idejű dinamika stacionárius (várakozási idős) állapotban, és különböző  $\alpha$  paraméterű Pareto-eloszlású eseményközi dinamika (színes, folytonos vonalak) nemstacionárius állapotban eredményezte fertőzési görbék  $10^4$  csúcsból álló BA hálózaton. A bal ábra jobb alsó sarkában néhány görbe szórása is látható fekete, folytonos vonallal. Jobb ábra: fertőzési görbék a várakozási időket használva hatványfüggvényes dinamikánál ugyancsak  $10^4$  csúcsból álló BA hálózaton.  $\langle k \rangle = 4$ , és az elsőnek megfertőzött csúcs  $k_0$  fokszáma 25 és 125 közé esett. Minden görbe  $10^3$  futás eredménye. Az  $\alpha = 1.053, 1.111, 1.25, 1.4$  és  $2$  exponensek  $t_{min} = 0.05, 0.1, 0.2, 0.286$  és  $0.5$  ( $\tau$  egységekben mért) alsó levágásoknak felelnek meg. A  $t_0$  rögzített paraméter értéke  $0$  volt.

Az ábra legfontosabb mondanivalója az  $\alpha$  exponens csökkentésével tapasztalt szisztematikus gyorsulás a nemstacionárius állapotbeli terjedésben, ami pontosan az ellenkezője annak, amit stacionárius esetben a 11. ábra jobb részén, vagy a 3. ábrán láttunk. Mindez természetesen roppant érdekes, hiszen az órák előbb látott szinkronizálásával a modellünk szerint alapvetően lehet befolyásolni a terjedés gyorsaságát hatványfüggvényes eseményközi időknél; ráadásul, ugyancsak a 11. ábrán láthatóan, alapvetően a hálózat legnagyobb részét igen gyorsan eléri a fertőzés, legalábbis  $t_0 = 0$ -nál. Nézzük most meg, hogy milyen hatással van a fertőződésre, ha a  $t_0$  időt növeljük. A következő ábraegyüttes az előbbi,  $t_0 = 0$ -ás eset megismételt ábrája mellett  $t_0 = 10$  és  $100$  ( $\tau$  egységekben gondolkodva) valamint a stacionárius állapot fennállása esetén mutatja a fertőzés alakulását.



12. ábra. Levágás nélküli exponenciális (szaggatott, fekete vonal) dinamika stacionárius (várakozási idő) állapotban, és különböző  $\alpha$  paraméterű Pareto-eloszlású eseményközi idejű dinamika (színes, folytonos vonalak) stacionárius (bal felső ábra) és nemstacionárius (többi három ábra) állapotban eredményezte fertőzési görbék  $10^4$  csúcsból álló BA hálózaton. Három alábra jobb alsó sarkában néhány görbe szórása is látható fekete, folytonos vonallal.  $\langle k \rangle = 4$ , és az elsőnek megfertőzött csúcs  $k_0$  fokszáma 25 és 125 között. Minden görbe  $10^3$  futás eredménye. Az  $\alpha = 1.053, 1.111, 1.25, 1.4, 2$  és  $3$  exponensek  $t_{min} = 0.05, 0.1, 0.2, 0.286, 0.5$  és  $0.666$  ( $\tau$  egységekben mért) alsó levágásoknak felelnek meg. A  $t_0$  rögzített paraméter értéke  $0$  a jobb felső,  $10$  a bal alsó és  $100$  a jobb alsó ábrán.

A 12. ábrát figyelve érdekes dolgot állapíthatunk meg,  $t_0$  növelésével adódik egyfajta átmenet gyors és lassú hatványfüggvényes terjedés között.  $t_0 = 10$ -re és valóban már  $t_0 = 0$ -ra is lényegében minden  $\alpha$  exponens esetén a nagyon kései fertőzés egyértelműen lassú, ám a terjedési görbék korai szakaszán megfigyelhető egyfajta átcsapás. Az  $\alpha = 2$ -höz és  $3$ -hoz tartozó görbe helyzete ugyan keveset változik, de változik,  $t_0$  növelésével ez a két görbe helyet cserél.  $t_0 = 0$  esetén, eltekintve a legutolsó fertőzésektől, az  $\alpha = 3$ -hoz tartozó görbe a leglassabb, de  $t_0 = 10$  és  $100$ -ra már itt mutatkozik gyorsabb terjedés, ami, akárcsak a két görbe ekkor látható helyzete, a stacionárius esetnek megfelel. A zöld ( $\alpha = 1.4$ ) és sárga ( $\alpha = 1.25$ ) görbéket figyelve, itt is láthatjuk, hogy megindul egy gyors-lassú átmenet  $t_0$  növelésével, s valóban ez az átmenet az ennél kisebb  $\alpha$  értékek esetén is megfigyelhető, ám különböző mértékben. Azt lehet az ábrákról leszűrni, hogy minél kisebb az  $\alpha$  értéke, annál több idő szükséges ahhoz, hogy a rendszer kellőképpen megközelítse a stacionárius állapotot, amikor is nincs lényeges eltérés a várakozási idő

terjedéstől. Valószínű tehát, hogy van valamilyen a stacionárius állapot elérését jellemző karakterisztikus idő, mely az  $\alpha$  exponens csökkenő függvénye. Egy ilyen függvénykapcsolat megállapítása fontos feladat, ami az eddig elvégzett vizsgálatok folytatásának egyik fontos irányát jelenti.

Mint hogy ilyen vizsgálatokra nem került még sor, egyúttal a fejezet zárásául is, foglaljuk össze néhány mondatban, mit is láttunk az előbbi néhány ábrán. Megkülönböztettük a fertőzést lehetővé tévő folyamatok stacionárius és nemstacionárius állapotát, s azt figyeltük meg, hogy exponenciális eseményközi-idők esetén a stacionaritástól függetlenül a levágás bevezetése és növelése enyhe gyorsulást okoz a terjedésben, míg hatványfüggvényes esetben a terjedés lefolyása nagyban függ a folyamatok életkorától, vagyis az  $\alpha$  exponens/alsó levágás csökkentésével egyre lassuló vagy egyre gyorsuló terjedés is létrejöhet.

Az exponenciális függvényeknél látott enyhe gyorsulást magyarázhatja, hogy nagyobb levágásnál némiképp csökkennek a szórások az eseményközi és várakozási időkben is, aminek lehet szerepe a terjedés sebességében, miként azt [15]-ben is olvashatjuk, melyben a szerzők többek között a várakozási idő szórását (pontosabban a burstiness-nek nevezett mennyiséget) és annak hatását vizsgálják. Amellett is érvelhetünk, hogy, stacionárius esetben, a Pareto-dinamika miatt okoz alapvetően lassú terjedést, hiszen a hatványfüggvényes lecsengés miatt mindig adódhatnak hatalmas várakozási idők, amint azt a fákon való terjedésnél is láttuk, s noha hálózatokban a fákkal ellentétben vannak alternatív utak, hurkok, ezek lehetnek hosszúak, vagyis sok időbe telhet a megfertőződés. Ezek után persze nem világos, hogy nemstacionárius esetben ez az érvelés miért nem működik, de a kérdés megválaszolásában fontos szerepe lehet a nem-poissoni folyamatok memóriájának, amit a nemstacionárius folyamatoknál fellépő karakterisztikus idők is alátámaszthatnak. Sok tehát a megválaszolatlan kérdés, a megfigyelt jelenségek azonban amellet, hogy nagyon érdekesek, rendkívül fontosak is, ezért mindenképpen további kutatások irányába ösztönöznek.

### 3. Összefoglalás

Ennek a TDK-dolgozatnak a célja a sok valós hálózaton megfigyelhető, időben inhomogén, villanásos dinamika okozta terjedés vizsgálata volt, mely napjainkban fontos és aktívan vizsgált terület, számos, nem részben ellentmondásos eredménnyel. Az ellentmondó eredmények mögött a probléma összetettsége áll, hiszen mind az időbeli folyamatok inhomogenitása, mind pedig a hálózatok topológiai sajátosságai befolyással vannak a terjedés alakulására. Éppen ezért igyekeztünk ebben a dolgozatban is figyelembe venni fontos hálózattopológiai elemeket (skálafüggetlen foksámeloszlás, kisvilág-tulajdonság) és a villanásos dinamika sajátosságait, mely utóbbiaknál az inhomogén folyamatok stacionárius vagy nemstacionárius állapota is szerepet kapott.

A dolgozat első részében BA hálózatokon vizsgáltunk hatványfüggvényes dinamikát állandósult állapotban, s vizsgálatainkból az derült ki, hogy a terjedés ebben esetben az exponenciális dinamikával összevetve alapvetően lassú és az eseményköziidő-eloszlás lecsengését változtatva szisztematikus effektus tapasztalható, azaz kisebb exponensekre lassabb terjedéssel válaszol a hálózat, s a terjedés lassúsága megmutatkozhat mind a kezdeti, mind



pedig az asszimptotikus fertőződésben is. Ez az eredmény ellentmondásban van egy a miénknél triviálisabb topológiájú, teljes gráfon vizsgált és analitikusan számolható eredménnyel [7]. Az eltérés léte tehát a különböző topológiákra vezethető vissza és így eredményünk aláhúzza a topológiai elemek fontosságát.

A BA fákon való terjedés tanulmányozásával egyrészt bemutattuk, hogy a fákban meglévő „szűk keresztmetszetek” milyen hatással vannak a terjedésre, ha a dinamika hatványfüggvényes; e két lényeges elem (megfelelő exponensek esetén) a terjedést nagyon egyenetlenné, kiszámíthatatlanná teszi, a fertőzés elakadhat, s adott esetben csak nagyon sok idő elteltével képes tovaterjedni. E megfigyelés rámutat arra, hogy bár a BA hálózatok sok szempontból lokálisan fára emlékeztetnek, az esetükben tapasztalt eltérő terjedési jellegek miatt ennek a képnek nyilvánvaló korlátai vannak, legalábbis inhomogén terjedések esetén. Mindemellett a BA fák vizsgálatával mi is rávilágítottunk arra, hogy a valóságos, s ezért általában modularitással rendelkező hálózatokon milyen lényeges effektusokat okozhat a moduláris tagozódás, bár ilyenkor a fáknál tapasztaltnál enyhébb hatásokra lehet számítani.

A dolgozat utolsó részében megvizsgáltuk azt, hogy milyen hatása van a terjedésre a fertőzési folyamatok stacionárius vagy nemstacionárius jellege. A stacionaritás kérdésének nyilvánvaló relevanciája van sok, a biológiai ritmust figyelembe vevő hálózat esetén, ennek ellenére eddig az irodalomban ezt a kérdést figyelmen kívül hagyták. A dolgozatban bemutatott vizsgálódásból az derült ki, hogy exponenciális dinamika esetén nem, de hatványfüggvényes eseményközidő-eloszlásokkal modellezhető dinamika esetén a stacionaritásnak lényeges szerepe van a terjedés sebességében. A stacionárius esetben tapasztalt alapvetően lassú és a terjedés gyorsaságának  $\alpha$ -függését tekintve szisztematikus fertőzés ( $\alpha$  csökkentésével lassul a fertőzés), gyors és ugyancsak szisztematikus, de  $\alpha$  változtatására pontosan ellentétesen viselkedő fertőződésbe mehet át nemstacionárius esetekben; kiderült továbbá az is, hogy az állandósult állapot elérésének ideje ugyancsak függ a Pareto-eloszlás exponensétől. A leírtak fontossága nyilvánvaló, hiszen az eredmények szerint egy adott időben bejuttatott fertőzés elterjedésének gyorsaságát az időpont megválasztásával is jelentősen lehet befolyásolni, a hálózat fertőzésre való fogékonysága tehát erősen időfüggő.

## 4. Függelék

Ebben a részben a szimulációkban használt eloszlásokkal kapcsolatos mennyiségeket számolunk ki. Bár a kapott eredmény ennél általánosabb, a számolást motiválhatja az, hogy kíváncsiak vagyunk, ha valamilyen  $k_0 = n$  fokszámú csúcst fertőződik meg legelőször, akkor mi lesz az első igazi fertőzés várható ideje és ennek a szórása. Mivel a következő fertőzést a legkisebb idő határozza meg, matematikailag arra vagyunk kíváncsiak, hogy  $n$  IID  $(Y_1, Y_2, Y_3 \dots Y_n)$  valószínűségi változó minimumának, azaz  $X_n = \min(Y_1, Y_2, Y_3 \dots Y_n)$ -nek mi a várható értéke és szórása. Két  $Y$  valószínűségi változó esetén, azaz  $n = 2$ -re  $E(X_2)$  és  $E(X_2^2)$ -et a következő integrálok adják meg, melyekben  $p(y)$   $Y$  sűrűségfüggvénye.

$$E(X_2) = 2 \int_0^\infty \int_x^\infty p(y) dy xp(x) dx \quad (4.1a)$$

$$E(X_2^2) = 2 \int_0^\infty \int_x^\infty p(y) dy x^2 p(x) dx \quad (4.1b)$$

$n=3$ -ra még könnyen látható, hogy milyen tartományra kell integrálni, ahonnan adódik az általános eredmény is  $\forall n > 1$ -re.

$$E(X_n) = n \int_0^\infty \left( \int_x^\infty p(y) dy \right)^{n-1} xp(x) dx \quad (4.2a)$$

$$E(X_n^2) = n \int_0^\infty \left( \int_x^\infty p(y) dy \right)^{n-1} x^2 p(x) dx \quad (4.2b)$$

Ha  $p(y) = p'_{exp}(t)$ , azaz levágás nélküli, tiszta  $\tau$  várható értékű exponenciális a várakozási idő eloszlása (s ekkor maga az eseményközi idő eloszlása is), akkor lévén, hogy  $\int_x^\infty e^{-y} dy = e^{-x}$

$$E(X_n) = \tau n \int_0^\infty x e^{-nx} dx = \frac{\tau}{n} \quad (4.3a)$$

$$D(X_n) = \tau \sqrt{n \int_0^\infty x^2 e^{-nx} dx - \frac{1}{n^2}} = \tau \sqrt{\frac{2n}{n^3} - \frac{1}{n^2}} = \frac{\tau}{n} \quad (4.3b)$$

Az a szemléletes eredmény adódik, hogy exponenciális várakozási időknél a kezdetben megfertőzött csúcs első szomszédot érintő fertőzéséig várhatóan  $\tau/n$  időt kell várni. A következőkben Pareto-eloszlások esetén számítjuk ki a fentebbi mennyiségeket. Ehhez először is szükség van a  $\int_x^\infty p'_{pow}(y) dy$  integrálra. Emlékeztetőül, (1.10) szerint

$$p'_{pow}(t) = \begin{cases} \frac{1}{\tau}, & \text{ha } 0 \leq t \leq t_{min} \\ \frac{1}{\tau} t_{min}^\alpha \frac{1}{t^\alpha}, & \text{ha } t > t_{min} \end{cases} \quad (4.4)$$

így az integrál

$$\int_x^\infty p'_{pow}(y) dy = \begin{cases} \frac{1}{\alpha-1} \frac{1}{\tau} \frac{t_{min}^\alpha}{x^{\alpha-1}}, & \text{ha } x \geq t_{min} \\ \frac{1}{\alpha-1} \frac{t_{min}}{\tau} + \frac{1}{\tau} (t_{min} - x), & \text{ha } 0 \leq x \leq t_{min} \end{cases} \quad (4.5)$$

Végül tehát az adódik (4.2a) és (4.2b) alapján, hogy

$$E(X_n) = n \left[ \int_0^{t_{min}} \left( \frac{1}{\alpha-1} \frac{t_{min}}{\tau} + \frac{(t_{min} - x)}{\tau} \right)^{n-1} \frac{x}{\tau} dx + \int_{t_{min}}^\infty \left( \frac{1}{\alpha-1} \frac{1}{\tau} \frac{t_{min}^\alpha}{x^{\alpha-1}} \right)^{n-1} \frac{x}{\tau} \frac{t_{min}^\alpha}{x^\alpha} dx \right] \quad (4.6a)$$

$$E(X_n^2) = n \left[ \int_0^{t_{min}} \left( \frac{1}{\alpha-1} \frac{t_{min}}{\tau} + \frac{(t_{min}-x)}{\tau} \right)^{n-1} \frac{x^2}{\tau} dx + \int_{t_{min}}^{\infty} \left( \frac{1}{\alpha-1} \frac{1}{\tau} \frac{t_{min}^\alpha}{x^{\alpha-1}} \right)^{n-1} \frac{x^2}{\tau} \frac{t_{min}^\alpha}{x^\alpha} dx \right] \quad (4.6b)$$

Először (4.6a) első tagját számoljuk ki, ehhez  $\frac{1}{\alpha-1} \frac{t_{min}}{\tau}$ -t elnevezzük  $a$ -nak,  $\frac{1}{\tau}$ -t pedig  $b$ -nek.

$$\begin{aligned} \int_0^{t_{min}} (a + b(t_{min} - x))^{n-1} \frac{x}{\tau} dx &= \left[ \frac{1}{n} \frac{(a + b(t_{min} - x))^n x}{-b} \right]_0^{t_{min}} + \frac{1}{n} \int_0^{t_{min}} \frac{(a + b(t_{min} - x))^n}{b\tau} dx \\ &= \frac{1}{n} \frac{a^n t_{min}}{-b} + \left[ \frac{1}{n(n+1)} \frac{(a + b(t_{min} - x))^{n+1}}{-b^2\tau} \right]_0^{t_{min}} \\ &= \frac{1}{n} \frac{a^n t_{min}}{-b} + \frac{1}{n(n+1)b^2\tau} ((a + bt_{min})^{n+1} - a^{n+1}) \end{aligned}$$

Visszaírva  $a$  és  $b$  értékeit,

$$\begin{aligned} \int_0^{t_{min}} (a + b(t_{min} - x))^{n-1} \frac{x}{\tau} dx &= \left( \frac{t_{min}}{(\alpha-1)\tau} \right)^n \frac{t_{min}}{n} + \frac{\tau}{n(n+1)} \left( \left( \frac{t_{min}}{(\alpha-1)\tau} + \frac{t_{min}}{\tau} \right)^{n+1} - \left( \frac{t_{min}}{(\alpha-1)\tau} \right)^{n+1} \right) \\ &= \frac{\tau}{n(n+1)} \left[ \left( \frac{t_{min}}{(\alpha-1)\tau} + \frac{t_{min}}{\tau} \right)^{n+1} - (1 + (\alpha-1)(n+1)) \left( \frac{t_{min}}{(\alpha-1)\tau} \right)^{n+1} \right] \\ &= \frac{\tau}{n(n+1)} \left[ 1 - (1 + (\alpha-1)(n+1)) \frac{1}{\alpha^{n+1}} \right] \end{aligned}$$

(4.6a) második tagja könnyebben megadható.

$$\begin{aligned} \int_{t_{min}}^{\infty} \left( \frac{1}{\alpha-1} \frac{1}{\tau} \frac{t_{min}^\alpha}{x^{\alpha-1}} \right)^{n-1} \frac{x^2}{\tau} \frac{t_{min}^\alpha}{x^\alpha} dx &= \int_{t_{min}}^{\infty} \left( \frac{1}{\alpha-1} \right)^{n-1} \left( \frac{1}{\tau} \right)^{n-1} t_{min}^{\alpha n} \frac{1}{x^{n(\alpha-1)}} dx \\ &= \frac{1}{n\alpha - (n+1)} \left( \frac{1}{\alpha-1} \right)^{n-1} \left( \frac{1}{\tau} \right)^n t_{min}^{\alpha n} \left( \frac{1}{t_{min}} \right)^{\alpha n - (n+1)} \\ &= \frac{1}{n\alpha - (n+1)} \left( \frac{1}{\alpha-1} \right)^{n-1} \left( \frac{t_{min}}{\tau} \right)^n t_{min} \\ &= \frac{1}{n\alpha - (n+1)} \left( \frac{1}{\alpha-1} \right)^{n-1} \left( \frac{\alpha-1}{\alpha} \right)^{n-1} \frac{\alpha-1}{\alpha} \tau \frac{\alpha-1}{\alpha} \\ &= \frac{\tau}{n\alpha - (n+1)} \frac{1}{\alpha^{n-1}} \frac{(\alpha-1)^2}{\alpha^2} \\ &= \frac{\tau}{n\alpha - (n+1)} \frac{(\alpha-1)^2}{\alpha^{n+1}} \end{aligned}$$

Lényeges, hogy ez a második integrál akkor létezik, ha  $\alpha > \frac{1}{n} + 1$ . Az előbb levezetett számolással ugyanilyen módon  $E(X_n^2)$  is meghatározható. Összeírva az  $E(X_n)$ -nél látott

két tagot, nem elfeledkezve az  $n$ -es szorzóról, továbbá mellőzve az  $E(X_n^2)$ -re vonatkozó számolásokat, a következő végeredményt kapjuk  $\alpha$  exponensű Pareto-eloszlású várakozási időknél.

$$E(X_n) = \tau \left[ \frac{1}{(n+1)} \left[ 1 - (1 + (n+1)(\alpha-1)) \frac{1}{\alpha^{n+1}} \right] + \frac{n}{n\alpha - (n+1)} \frac{(\alpha-1)^2}{\alpha^{n+1}} \right] \quad (4.7a)$$

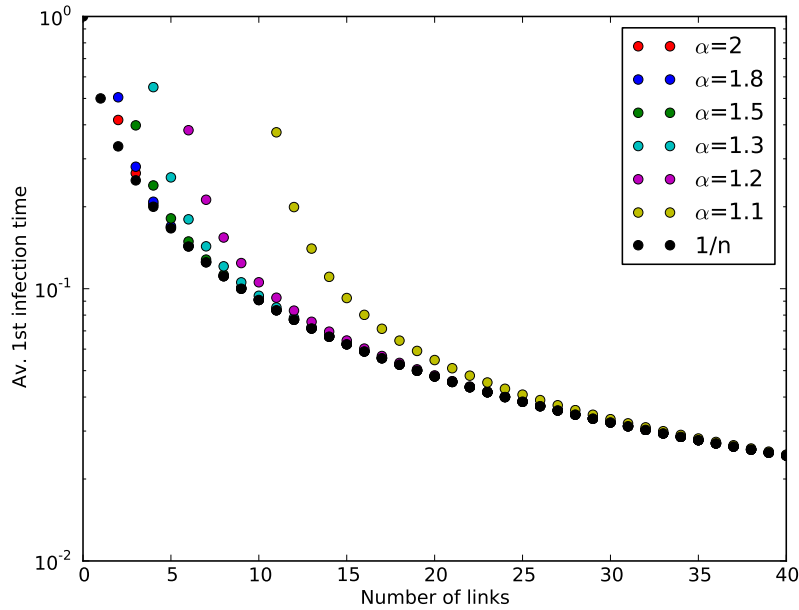
$$E(X_n^2) = \tau^2 \left[ \frac{2}{(n+1)(n+2)} \left[ 1 - \left( 1 + (n+2)(\alpha-1) + \frac{1}{2}(n+2)(n+1)(\alpha-1)^2 \right) \frac{1}{\alpha^{n+2}} \right] + \frac{n}{n\alpha - (n+2)} \frac{(\alpha-1)^3}{\alpha^{n+2}} \right] \quad (4.7b)$$

Ahogy  $E(X_n)$  létezésére az  $\alpha > \frac{1}{n} + 1$  feltétel adódott,  $E(X_n^2)$  akkor véges, ha  $\alpha > \frac{2}{n} + 1$  teljesül. Bár a (4.7a) és (4.7b) formulákat  $n \geq 2$  esetre vezettük le, egyetlen  $Y$ , várakozási időnek megfelelő valószínűségi változó várható értéke és szórása esetében is ez a létezés feltétele. Így tehát teljesen általánosan kijelenthetjük, hogy ha BA fán egy levélről indítjuk a folyamatot, akkor az első nem mesterséges fertőzésig várhatóan végtelen időt kell várni, ha  $\alpha \leq 2$ , míg  $m = 2$  paraméterű BA hálózat esetén levélről való indulás esetén  $\alpha \leq 1.5$  esetén adódik végtelen várható idő. A (4.7a) és (4.7b) formulák viszonylag hosszúak és csúnyák, ezért érdemes megvizsgálni valamilyen határesetet. Ha  $n$  kellően nagy, akkor az  $\frac{1}{\alpha^{n+1}}$  jellegű tagok kicsisége miatt ( $\alpha$  mindig nagyobb 1-nél)  $E(X_n)$ -t és  $D(X_n)$ -t a következő módon közelíthetjük.

$$E(X_n) \approx \frac{\tau}{n+1} \quad (4.8a)$$

$$D(X_n) \approx \frac{\tau}{n+1} \sqrt{\frac{n}{n+2}} \quad (4.8b)$$

Ez azért érdekes, mert lényegében összhangban van a 3. ábrán látottakkal, jöllehet itt most csak az első fertőzési időről tudunk mondani bármit is. (4.8a) szerint kellően nagy  $n$ -re (ami  $\alpha = 2$ -re nem is kell, hogy olyan nagy legyen,  $n = 5$  is elég, hogy a közelítés jó legyen) az első fertőzés átlagosan hamarabb megtörténik Pareto-, mint exponenciális dinamika esetén. Az alábbi ábrán az  $E(X_n)$  várható érték látható ( $\tau$  egységekben) exponenciális és különböző  $\alpha$  paraméterű skálafüggetlen eseményközidő-eloszlásokból adódó várakozási idők esetében, az x-tengelyen pedig a szomszédok száma, vagyis  $n$  szerepel.



13. ábra. Az  $E(X_n)$  várható érték  $\tau$  egységeken exponenciális és különböző  $\alpha$  paraméterű eloszlásokból adódó várakozási idők esetében a szomszédok  $n$  számának függvényében.

A 13. ábrán látható, hogy az  $\frac{1}{n+1}$ -es asszimptotikától távol hatványfüggvényes várakozási időknél többet kell várni az első fertőzésig, mint az exponenciális esetben. Ismét hangsúlyozva, hogy mindez az első (nem mesterséges) fertőzésre vonatkozik, a 3. ábrán a fertőzési görbék kezdetén, tehát nem csak az első fertőzés esetén látható hasonló tendencia, azaz  $\alpha$ -tól függen gyors vagy lassú kezdeti fertőzés, melyet valamilyen mértékben megerősít a most bemutatott számolás. Megkérdezhetjük, hogy ezek után mi akadály van annak, hogy a második, harmadik, stb. fertőzés várható idejét is kiszámoljuk. Ehhez egyrészt tudni kell azt, hogy a különböző fokszámú csúcsok között milyen korreláció áll fenn, mely BA hálózatok esetén ismert, de a sokadik fertőzésre a probléma borzasztóan bonyolulttá válik. Másfelől nem tiszta exponenciális eloszlások esetén, ha egy él mentén történhet fertőzés, de tudjuk, hogy egy adott  $t$  ideig nem történt esemény, akkor annak a megválaszolása, hogy milyen eséllyel lesz esemény ezt követően  $t'$  időn belül, nem olyan egyszerű, mint az örökifjú exponenciális eloszlás esetén. Tiszta exponenciális eloszlással dolgozva, egy BA hálózatnál triviálisabb szerkezetű hálózaton, a Cayley-fán azonban van lehetőség a fertőzési idők várható értékének analitikus kiszámolására, ahogy ebben a dolgozatban ezt be is mutatjuk.

## A szimulációkhoz használt programnyelv

A TDK-dolgozatban bemutatott eredmények nagyrészt számítógépes szimulációkon alapulnak. A szimulációkhoz a Python programozási nyelvet használtuk, mely rendelkezik

tudományos igényű hálózatkutatási (NetworkX) és véletlen számok előállítására alkalmas (Random) modulokkal, melyeket a munka során fel is használtunk. A szimulációk során rendkívül lényeges szerepe volt a használandó adatstruktúrának, mely segítségével sok-sok él közül ki lehet választani az időben következőnek fertőzendőt. A feladathoz először saját készítésű  $\ln(N)$ -es elérési idejű struktúra után a hatékonyabb memóriakezelésű, a Python Heapq moduljából elérhető *heap* adatstruktúrát használtuk.

## Hivatkozások

- [1] A. Barrat, M. Barthélemy, A. Vespignani, *Dynamical Processes on Complex Networks*. Cambridge Univ. Press, 2012.
- [2] Alessandro Vespignani and his group, <http://www.youtube.com/watch?v=ONEOc-MTm1Q>
- [3] Barabási-Albert László, *The origin of bursts and heavy tails in human dynamics*. Nature 435, 207-211 (2005)
- [4] Petter Holme, Jari Saramäki, *Temporal networks*. Physics Reports, Volume 519, Issue 3, October 2012, Pages 97–125
- [5] M. Karsai, M. Kivelä, R. K. Pan, K. Kaski, J. Kertész, A.-L. Barabási, and J. Saramäki, *Small But Slow World: How Network Topology and Burstiness Slow Down Spreading*. Phys. Rev. E 83, 025102 (2011)
- [6] J.L. Iribarren and E. Moro, *Impact of Human Activity Patterns on the Dynamics of Information Diffusion*. Phys. Rev. Lett. 103, 038702 (2009)
- [7] Hang-Hyun Jo, Juan I. Perotti, Kimmo Kaski, János Kertész, *Enhanced Spreading Dynamics by Non-Poissonian Processes*. Preprint
- [8] Hang-Hyun Jo, Márton Karsai, János Kertész and Kimmo Kaski, *Circadian pattern and burstiness in mobile phone communication*. Preprint
- [9] Réka Albert, Albert-László Barabási, *Statistical mechanics of complex networks*. New J. Phys. 14, 013055 (2012)
- [10] William, Feller, *An Introduction to Probability Theory and Its Applications, Vol. 2*. Wiley, New York, 1971
- [11] N. Berger, C. Borgs, J. Chayes and A. Saberi, *Asymptotic behavior and distributional limits of preferential attachment graphs*. Annals of Probability (to appear)
- [12] S. Dereich and P. Morters, *Random networks with sublinear preferential attachment: the giant component*. Electronic Journal of Probability, 14, Article 43 (2009)

- [13] Ginestra Bianconi and Andrea Capocci, *Number of loops of size  $h$  in growing scale-free networks*. Phys. Rev. Lett. 90, 078701 (2003)
- [14] M. Newman, *Networks-An introduction*. Oxford Univ. Press, 2010
- [15] Jean-Charles Delvenne, Renaud Lambiotte, and Luis E C Rocha *Bottlenecks, burstiness, and fat tails regulate mixing times of non-Poissonian random walks*. <http://arxiv.org/pdf/1309.4155.pdf>

## Köszönetnyilvánítás

Szeretnék köszönetet mondani témavezetőmnek, Kertész János tanár úrnak a dolgozat és az eddigi eredmények megszületésében nyújtott mindenféle segítségéért, különösen a témayári művelését lehetővé tévő finnországi tanulmányútért.