

TDK dolgozat

Kommunikációs dinamika modellezése

Vajna Szabolcs

Fizikus MSc szak

témavezető

Dr. Kertész János

egyetemi tanár

BME Elméleti Fizika Tanszék

Budapesti Műszaki és Gazdaságtudományi Egyetem

2011

Tartalomjegyzék

1. Bevezetés	3
1.1. Barabási-modell	3
1.2. Adathalmaz és vizsgált mennyiségek	4
1.2.1. Mérési eredmények	5
1.3. Dolgozat felépítése	8
2. Prioritásrendezett dinamikus lista modell	9
2.1. Stacionárius megoldás	10
2.2. Mester-egyenlet	10
2.3. Egyensúlyi autokorrelációs függvény	11
2.3.1. Az autokorrelációs függvény időfüggése	12
2.3.2. Numerikus számítások ($w_i \sim i^{-\sigma}$ modell)	13
2.4. Várakozási idők és az események között eltelt idő eloszlása	16
2.5. $P(E)$ eloszlás	20
2.6. Diskusszió	20
3. Módosított modell	22
3.1. Autokorrelációs függvény	22
3.2. Visszatérési idő	23
3.3. $P(E)$ eloszlás	25
4. Független események közötti idővel jellemzett folyamatok	27
5. Összefoglalás	31
6. Függelék	32
6.1. Mester-egyenlet származtatása a Markov-lánc átmeneti mátrixából	32
6.2. Az autokorrelációs függvény L -függetlensége	34
6.3. A módosított modell, mint véletlen bolyongás	34

1. Bevezetés

A társadalomban, ill. a gazdaságban megfigyelhető számos jelenséget az egyéni emberi cselekvések mozgatnak. Többek között ennek is köszönhetően vált központi kérdéssé az emberi viselkedés tanulmányozása az alkalmazott tudományok terén. Napjainkban hozzáférhetővé váltak olyan elektronikus adatbázisok, amelyek részletes adatokat tartalmaznak az emberi tevékenységek időzítéséről, például telefonálási [1] és email-küldési [2, 3] szokásaikról. Ezek vizsgálata új utat nyitott a szociális hálózat tanulmányozása előtt, valamint lehetővé tette az emberi viselkedés statisztikus vizsgálatát. Az egyik elsőként vizsgált mennyiség az egymást követő ugyanolyan típusú cselekvések között eltelt idő (események közötti idő, az angol irodalomban inter-event time) eloszlása. Kiderült, hogy sok tevékenység esetén az események között eltelt idő eloszlása hatványfüggvény-szerű lecsengést mutat. Ez azt jelenti, hogy széles skálán megfigyelhetők inaktív és kiemelten aktív periódusok a tevékenységek végrehajtása során. Ez utóbbiakat villanásos (angolul bursty) periódusnak, az ilyen mintázatot tartalmazó idősort pedig villanásos viselkedésnek nevezzük. A villanásos viselkedést jellemzi az események közötti idő eloszlás lecsengésének kitevője. Hatványfüggvény-szerű események közötti idő eloszlás jellemzi például az egyének könyvtári kölcsönzési [4], az email-küldési [2, 4], valamint a telefonhívási [5] szokásait. Szintén hatványfüggvény-szerű várakozási idő eloszlást találtak Einstein, Darwin, ill. Freud hagyományos, postai levelezéseiben [6].

A villanásos viselkedés jelentősen eltér az emberi cselekvések modellezésére korábban használt Poisson folyamattól. A Poisson folyamat homogén idősort eredményez, és exponenciálisan ritkák a nagy események között eltelt idők. A villanásos viselkedés leírására szolgáló első modellt Barabási Albert László javasolta 2005-ben [7], azóta a humán dinamika iránt jelentős tudományos érdeklődés mutatkozik.

1.1. Barabási-modell

Barabási az embereket úgy modellezte, hogy van egy L (rögzített) hosszúságú listájuk, ami a teendőiket ($i = 1..L$) tartalmazza, és mindegyik teendőhöz tartozik egy x_i prioritás, amit egy $\rho(x)$ sűrűségfüggvényből választunk. A vizsgált alany minden lépésben p valószínűséggel kiválasztja a legnagyobb prioritású feladatát, vagy $1 - p$ valószínűséggel egyenletes eloszlással választ feladatot a listából. A kiválasztott feladatot végrehajtja, majd egy új teendő kerül a listába a $\rho(x)$ sűrűségfüggvénnyel jellemzett véletlen prioritással. A modellben tehát egy teendő várakozási ideje (τ_{wt} , waiting time) definiálható és számolható, amiről megmutatható [7, 8], hogy hatványlecsengést mutat $\beta = 1$ kitevővel: $P_{\tau_{wt}}(t) \sim t^{-1}$. Barabási úgy érvel ezután, hogy az események között eltelt idő eloszlásának hatványlecsengéséért a várakozási idők inhomogenitása felel. Az email kommunikációban, amit Barabási modellezni próbált, a várakozási idő becsülhető az adatsorból: a teendő itt válasz írása egy levélre, a listába kerülés ideje az e-mail beérkezési ideje, az ettől a válasz küldéséig eltelt idő pedig a várakozási idő. A fenti modell összhangban van az empirikus eredményekkel, mind a várakozási idő, mind pedig az egymás után küldött e-mailek közt eltelt idő eloszlása jól közelíthető -1 -es kitevőjű hatványfüggvénnyel. A modellben a lista hosszának rögzítését általában az emberi memória végeességével szokták indokolni. Megjegyzendő, hogy amennyiben a lista hosszát nem rögzítjük, más eredményt kapunk. Ha feltételezzük, hogy a teendők λ rátával érkeznek és μ rátával hajtódnak végre, akkor $\lambda = \mu$ esetén a várakozási időre $\beta = 3/2$ -es exponenst

kapunk. Ez az exponens jellemzi az Einstein, Darwin és Freud levelezéseiben mért várakozási idő eloszlásokat [6]. A lista hosszát ez esetekben a fiókban gyűlő levelek mennyisége adja, ami széles skálán ingadozhat (és nem kell őket fejben tartani), így érthető (legalábbis magyarázható), hogy a hagyományos, postai levelezés miért különbözik jelentősen a fenti példáktól.

Vázquez 2006-os cikkében [4] amellett érvel, hogy a különböző teendők a fentieknek megfelelő diszkrét univerzalitási osztályokba sorolhatók. Például a $\beta = 1$ univerzalitási osztályba tartozik az e-mailezés, a web-böngészés, valamint a könyvtári kölcsönzések, a $\beta = 3/2$ -be pedig a postai levelezés. Ezzel szemben egyre több kísérleti megfigyelés szól az univerzalitási osztályok ellen. Például Zhou és társai egy internetes videóközvetítő oldal, a Netflix felhasználói körében azt tapasztalták, hogy a különböző aktivitású (különböző átlagos események közötti idővel jellemzett) felhasználói csoportok hatványkitevője és átlagos aktivitása közt monoton kapcsolat figyelhető meg [9]. Hasonló eredményre jutott Karsai Márton a telefonálási szokások vizsgálata során [10]. A Barabási-modellnek van olyan általánosítása, amiben hangolható az exponens (ezt az eredeti cikkben a számolás könnyítése érdekében vezette be a szerző, de úgy tűnik lehet releváns vonatkozása is). Ez csak annyiban különbözik az eredeti modelltől, hogy az alany minden lépésben $\Pi(x) \sim x^s$ valószínűséggel választ egy x prioritású feladatot a végrehajtásra. A modellben a várakozási idő eloszlása: $P_{\tau_{wt}}(t) \sim t^{-(1+1/s)}$, és a modell $s \rightarrow \infty$ esetén az eredeti modell $p \rightarrow 1$ határesetébe megy át.

Összefoglalva a Barabási-modell fő hátránya az, hogy nehezen lehet benne kezelni az események között eltelt idő eloszlását, csak a várakozási időkre tud valamit mondani. Ez utóbbi sok esetben nem mérhető, például a telekommunikáció esetében általában nem mondhatjuk, hogy a másik fél hívásakor bekerülne az a feladat a listánkba, hogy visszahívjuk őt, mert gondolatainkat az első hívás során is megoszthatjuk vele. Az előző problémával szoros összefüggésben áll a modell egy másik hiányossága: az idősorok egy másik jellemző mennyiségét, az autokorrelációs függvényt sem lehet belőle természetes módon származtatni.

Munkám első célja, hogy találjak egy olyan modellt, ami képes kezelni az események között eltelt idő eloszlását, az autokorrelációs függvényt, valamint reprodukálni tudja a kutatócsoport által valódi adatokon végzett mérési eredményeket.

1.2. Adathalmaz és vizsgált mennyiségek

Egy anonim európai mobiltelefon szolgáltató a kutatócsoport rendelkezésére bocsájtotta egy ország 18 hetes mobilhívási (MPC) és sms-küldési (SM) adatait (2007. január 2-től 119 nap adatai, a szolgáltató piaci részesedése kb. 20% az országban). Az adatsor anonim, nevek helyett minden egyént egy azonosító jellemez. Az adatsor tartalmazza a rendelkezésre bocsájtott időtartamban indított összes hálózaton belüli hívás kezdő és végpontját másodperces felbontással, valamint a hívások irányát. Az adatsorból rekonstruálható az ország telekommunikációs hálózatának egy jelentős hányada. Az MPC adatsor összesen $N = 5.2 \times 10^6$ egyént, $L = 10.6 \times 10^6$ élt és $C = 322 \times 10^6$ hívást, az SM adatsor pedig $L = 8.5 \times 10^6$ élt és $C = 114 \times 10^6$ üzenetet tartalmaz.

Események között eltelt idő eloszlása

Az inhomogén (Poisson folyamattól eltérő) idősorok vizsgálatára az irodalomban leggyakrabban használt mennyiség az események között eltelt idő eloszlása. Az események a hívások indításai, ezek időpontjait egy személy esetében megszámozzuk: t_i az adott egyén i -edik indított hívásának kezdőpontja. Az események között eltelt idő: $\tilde{t}_{ie} = t_{i+1} - t_i$. Ebből általában ki szokták vonni a hívások hosszát, és így az események közötti inaktív időt mérik. A kutatócsoport méréseiben tehát az események között eltelt idő: $t_{ie} = t_{i+1} - (t_i + d_i)$, ahol d_i az i -edik hívás hossza. Az események között eltelt idő eloszlásának sűrűségfüggvényét a későbbiekben $P(t_{ie})$ -vel jelöljük. Ez a mennyiség jellemzi a hívások inhomogenitását, azonban nem mond semmit a hívások közötti korrelációkról.

Autokorrelációs függvény

A korrelációk vizsgálatának egyik legalapvetőbb módszere az $A(t)$ autokorrelációs függvény tanulmányozása. Az egyszerűség kedvéért tekintsünk egy diszkrét idejű adatsort (minden digitális adatbázis ilyen). Vezessük be az $X(t)$ indikátorváltozót: $X(t_i) = 1$, ha t_i időpontban történt esemény, és $X(t_i) = 0$ ha nem. Ezzel az autokorrelációs függvény definíciója

$$A(t) = \frac{\langle X(t+\tau)X(\tau) \rangle_\tau - \langle X(\tau) \rangle_\tau^2}{\langle X(\tau) \rangle_\tau - \langle X(\tau) \rangle_\tau^2}, \quad (1)$$

ahol a nevezőben kihasználtuk, hogy $X(t)^2 = X(t)$. A mérések során a kutatócsoport összehúzza a hívások időtartamát, hogy kiküszöböljék azt a hatást, hogy a vizsgált alany nem indíthat addig új hívást, amíg valaki mással beszél.

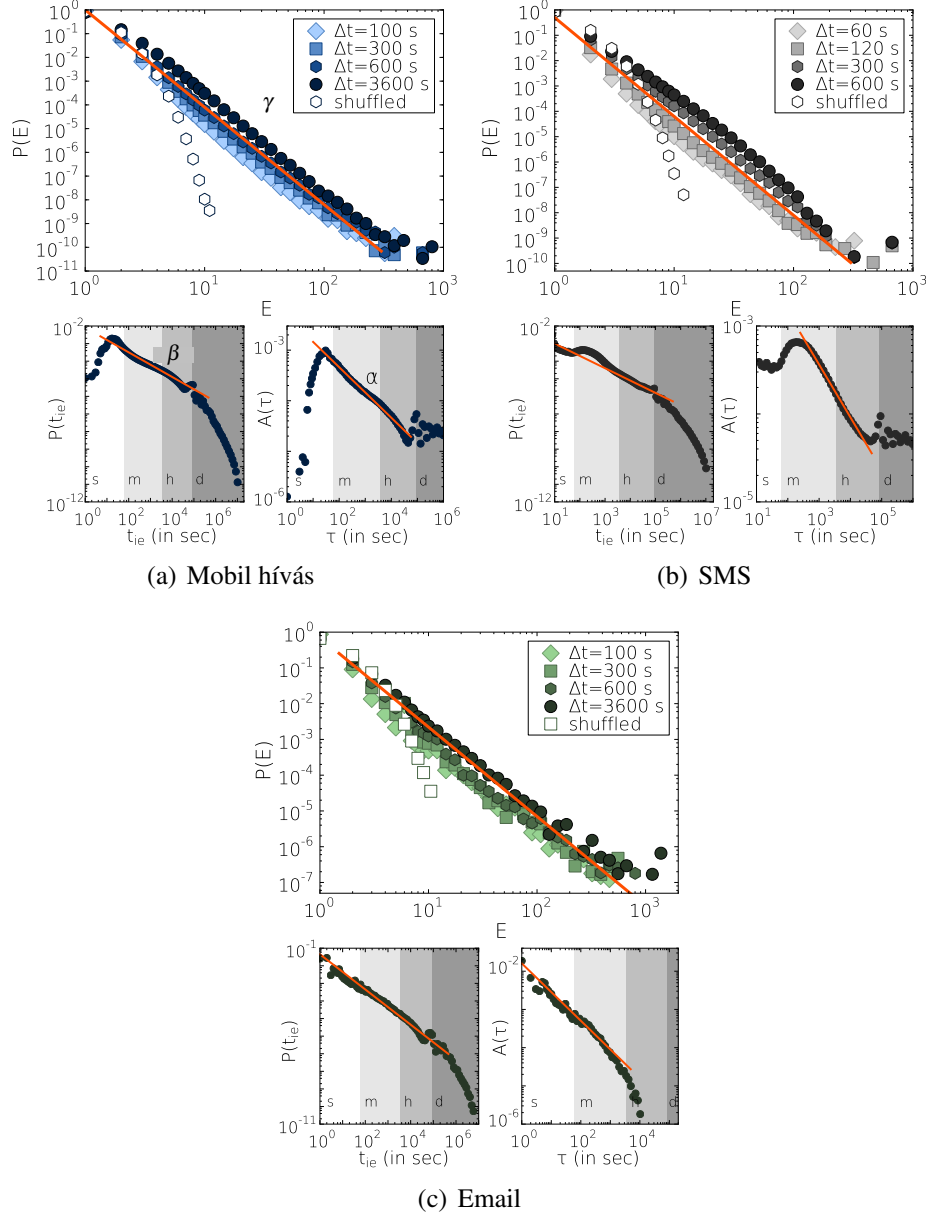
P(E) eloszlás

A kutatócsoport bevezetett a villanásos viselkedés elemzésére egy új jellemző mennyiséget, a $P(E)$ eloszlást [10]. Ez a mennyiség megadja egy Δt időablakkal definiált villanásos periódusban az események számának súlyfüggvényét. A hívások kezdőpontjait továbbra is t_i -vel jelölve két egymást követő esemény ugyanabba a villanásos periódusba tartozik, ha $t_{i+1} - (t_i + d_i) \leq \Delta t$, ahol d_i az i -edik hívás hossza. Egyszerűen megmutatható, hogy független események közötti idővel jellemzett (felújítási) folyamatok esetén a $P(E)$ eloszlás exponenciálisan cseng le:

$$\mathbb{P}(E = n) = \left(\int_0^{\Delta t} P(t_{ie}) dt_{ie} \right)^{n-1} \left(1 - \int_0^{\Delta t} P(t_{ie}) dt_{ie} \right) \quad (2)$$

1.2.1. Mérési eredmények

Az adatbázisokon végzett mérések azt mutatják, hogy a három vizsgált függvény a különböző kommunikációs csatornáknak (telefonálás, sms, e-mail) hasonló (lásd az 1. ábrán). Ez alapján úgy gondoljuk, hogy ezen függvények ilyen értelemben jó mennyiségek az emberi cselekvések időzítésének leírására. Másrészt ez arra utal, hogy az emberi viselkedésben van valami karakterisztikus



1. ábra. Karsai Márton mérései különböző kommunikációs adatbázisokon [10]. A skálázó szakaszokhoz tartozó exponensek az 1. táblázatban találhatók.

vonás, ami ezeket a mintázatokat produkálja, és ez esetleg megfogható egy alkalmas modellben. Minket most elsősorban a telefonálási mintázatok érdekelnek. Az idősor nagyon bonyolult, és rengeteg információt tartalmaz, ezt megpróbáljuk első közelítésben minél jobban leegyszerűsíteni, hogy megérthessük a különböző összetevők hatását a teljes rendszerre vonatkozóan. Ezért a vizsgált mennyiségek ($A(t)$, $P(t_{ie})$ és $P(E)$) nem foglalkoznak a hívások hosszával. A hívások hosszának hatása az idősorra a későbbiekben még elemzendő probléma lehet.

Az autokorrelációs függvényben és az események között eltelt idő eloszlásában két természetes időskála megjelenését látjuk (1. ábra), ami az események közötti idő sűrűségfüggvényének levá-

gásaként jelenik meg. Az alsó levágás kb. 20 másodpercnél kezdődik, ez a hívások utáni tipikus reakcióidőhöz kapcsolódik. A másik levágás az egy napnál nagyobb időtartamoknál jelentkezik. Az irodalomban afelől sincs egyetértés, hogy a levágás exponenciális, vagy pedig egy meredekebb hatványfüggvény lenne. A két skála közt azonban (20s-12h) az események között eltelt idő eloszlásában és az autokorrelációs függvényben is egy kb. 2-3 dekád nagyságú, log-log skálán egyenessel közelíthető szakasz jelenik meg. Ezen skálázó szakaszokat az irodalomban valóban hatványfüggvényként kezelik, mint arról a bevezetőben szó volt. Az események között eltelt idő eloszlásához tartozó exponenst β -val, az autokorrelációs függvény kitevőjét pedig α -val jelöljük. Az említett szakaszokon tehát $P(t_{ie}) \sim t^{-\beta}$, ill. $A(t) \sim t^{-\alpha}$. A különböző kommunikációs csatornák esetén mért exponenseket az 1. táblázat tartalmazza.

	α	β	γ
MPC	0.5	0.7	4.7
SMS	0.6	0.7	3.9
E-mail	0.75	1.0	2.75

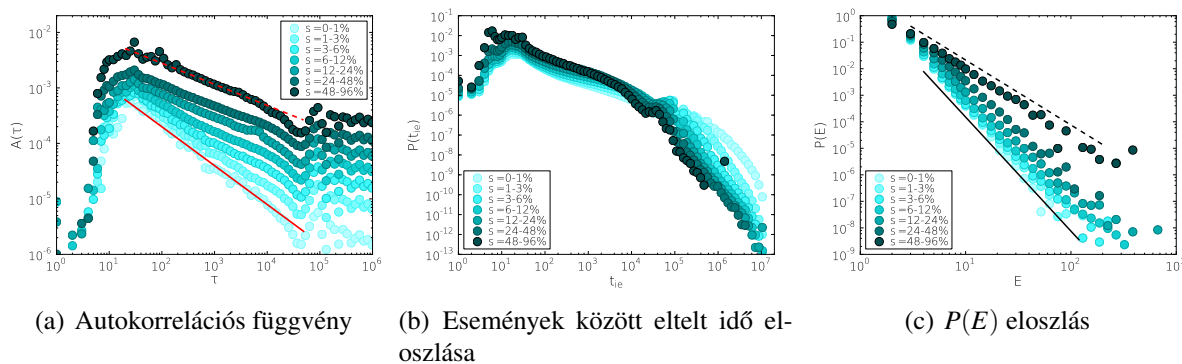
1. táblázat. Az autokorrelációs függvény (α), az események között eltelt idő eloszlás (β), és a $P(E)$ eloszlás (γ) exponensei [10].

A $P(E)$ eloszlás a villanásos periódusok hosszának eloszlását méri az események számában. Ez a mennyiség definíció szerint függ a választott Δt időablaktól, amin belüli két eseményt azonos villanásos periódushoz tartozónak tekintünk. Amennyiben az időablak nagyon rövid (pl. az adatsor felbontásával összemérhető), minden eseményt különálló villanásnak fogunk látni, így a $P(E)$ eloszlás az $E = 1$ értékre lesz centrálva. Növelve az időablak méretét az események egy része összeolvad villanásos periódussá, és már nem lesz triviális a $P(E)$ eloszlás. A másik véget az, amikor az időablak összemérhető az adatsor méretével. Ekkor minden cselekvés egy óriási villanásos periódusba olvad össze, és a $P(E)$ súlyfüggvény az adott egyénre jellemző összes hívásszám esetén vesz fel 1 értéket, egyébként zérus. A mérések szerint ha az időablakot a két levágási idő közötti szakaszból választjuk ($\Delta t \in (20s, 12h)$), a $P(E)$ súlyfüggvény nem függ jelentősen az időablak konkrét értékétől, és hatványfüggvény alakú lecsengést látunk (1. ábra).

Az 1. ábrán látható mennyiségek a teljes hálózatra vonatkozó átlagot jelenítik meg. Az átlagolás azért szükséges, hogy csökkentse a zajt, azonban információt is veszítünk vele. Ezért kérésre Karsai Márton megvizsgálta, hogy milyen jellegű görbéket kapunk, ha a különböző aktivitású (összes hívásszámú) csoportokba osztjuk az embereket, és csak ezeken belül átlagolunk. Eredményül azt kaptuk, hogy az egyes csoportokra vonatkozó görbék kvalitatívan hasonlóak, azonban a skálázó szakaszok exponense monoton függést mutat a csoport aktivitásával (2. ábra).

Napi és heti ritmus

A telefonálási idősorban természetes módon megjelenik a napi, ill. a heti ritmus hatása: éjszaka és munkaidőben az emberek telefonálási lehetőségei korlátozottak, így hosszú események közötti idők megjelenését okozzák. Vannak, akik az események között eltelt idők széles eloszlását egyedül az életben megfigyelhető ciklikus jelenségekkel magyarázzák [11]. Az emberek napi vagy heti ritmusa jellemezhető egy időfüggő átlagos telefonálási rátával. Ez a ráta elegendően nagy adatbázisból becsülhető, és a ráta ismeretében az idősor átskálázható a rátának megfelelően [11, 12], ami



2. ábra. A karakterisztikus függvények függése a felhasználócsoportokra jellemző hívási gyakoriságtól. Karsai Márton mérései az MPC adatbázison ([10], Supplementary Information).

így jelentős mértékben csökkenti a napi ritmus hatását. A kutatócsoportunk elvégezte a napi ritmus kiskálázását, és végeredményben az eredetihez hasonló függvényeket kapott [12]. Sőt, az eljárás lefolytatása után valamelyest megnőtt a skálázó szakasz hossza. Így kijelenthetjük, hogy a hosszú farkú események közötti idő eloszlásért nem kizárólag a napi ritmus a felelős, hanem valamilyen mély, belső emberi tulajdonság is rejlik mögötte.

1.3. Dolgozat felépítése

A mérési eredményeket összefoglalva tehát olyan modellt keresek, amire teljesül hogy

- definiálható benne az események között eltelt idő
- tud produkálni hatványfüggvény viselkedést az autokorrelációs függvényben, valamint az események között eltelt idő eloszlásában
- a villanásos periódusok hosszára ($P(E)$) a méréseknek megfelelő eloszlást mutat
- a megjelenő exponensek hangolhatók a modellben

TDK dolgozatomban prioritás alapján rendezett dinamikus lista-modelleket tanulmányozok analitikus és numerikus módszerekkel. A vizsgált két modellben – a Barabási-modellhez hasonlóan – különböző tevékenységek szerepelnek egy listában, és versenyeznek a végrehajtásért. Numerikus számításaim eredményei alapján mindkét modell hatvány-lecsengésű autokorrelációs függvényt és események közötti idő eloszlást produkál, hangolható exponensekkel. Az exponensek eleget tesznek egy skálatörvénynek, ezt a skálatörvényt bizonyítom, és kiterjesztem a modellek szélesebb körére. Mindkét modell autokorrelációs függvényére végesméret-skálázás végezhető. Ezen tulajdonsága a modelleknek lehetővé teszi, hogy meghatározzam az autokorrelációs függvény lecsengésének exponensét a modellek bemenő paramétereinek függvényeként.

2. Prioritásrendezett dinamikus lista modell

Feltesszük, hogy minden embernek van egy listája a teendőiről, és a következő végrehajtott cselekedetét a listából választja valamilyen súlyfüggvény által megadott valószínűséggel. A lista rendezett olyan értelemben, hogy a teendők kiválasztásának valószínűsége monoton csökken a listában elfoglalt hely szerint. A kiválasztott cselekvést a személy a következő időlépésben elvégzi, és ezzel párhuzamosan a teendő a lista elejére kerül. A modellben ez a mechanizmus fogja szolgáltatni a villanásos periódusok megjelenését. Heurisztikusan ez azt jelenti, hogy ha a vizsgált alanyunk eszébe jut a telefonálás, akkor ez napirenden is marad egy darabig. Egy olyan modell nem tudna villanásos viselkedést produkálni, amiben a telefonálás a végrehajtás után a lista végére kerül.

A modell konstrukciója olyan, hogy elvileg bármilyen jellegű, és akár egyszerre több cselekvést is lehetne vele modellezni. Mi a listában szereplő teendőket két csoportra osztjuk: A -val jelöljük a telefonálási, B -vel pedig az összes többi teendőt. A lista hosszát N -nel, a listában szereplő B elemek számát pedig L -lel jelöljük (A -ból tehát $N-L$ db. van). Egy személy listája, és két átmenet a következőképp nézhet ki (az egyes lépésekben választott teendők színessel vannak szedve):

ABAABBAAB...

AABAAABAAAB...

BAABAABAAAB...

Ebben a példában tehát a vizsgált alany előbb telefonált, majd valami mással töltötte az idejét. Megfigyelhetjük, hogy az elvégzett teendők a lista első elemének időfejlődését jelentik.

A modellben a következő mennyiségeket vizsgálhatjuk meg:

- a cselekvések idősorában A aránya hosszú távon: $\frac{N_A}{N_A+N_B}$, ahol N_A és N_B a vizsgált időszakban rögzített A és B események darabszáma
- az A cselekvéshez kapcsolódó autokorrelációs függvény
- egy A esemény után mikor következik be ismét A (események között eltelt idő)
- a modell által produkált $P(E)$ eloszlás

Kapcsolat a kártyakeverések elméletével

Definiálhatjuk a modell időtükrözött párját, amikor minden lépésben kisorsolunk egy számot ugyanolyan eloszlással, mint az eredeti modellben, majd a lista első elemét kivesszük, és beszurjuk a kisorsolt helyre. Az így kapott modell ugyanolyan tulajdonságokkal rendelkezik, mint az eredeti, például ugyanaz lesz az autokorrelációs függvény, és az események között eltelt idő eloszlása is. Ha most a listára úgy tekintünk, mint egy csomag (pakli) kártyára, akkor az időtükrözött modell egyfajta általánosítása a *top-in shuffle* keverésnek. Ez utóbbi a kártyakeverés egy primitív modellje, leemeljük a legfelső kártyát a csomagból és egyenletes eloszlással beszurjuk a csomag valamely pozíciójába.

A fejezet további részében a modell eredeti változatát vizsgáljuk.

2.1. Stacionárius megoldás

A fent bevezetett modell egy Markov-láncot definiál. Az állapottér az N hosszúságú, L db. B -t tartalmazó $AAA...BBB...$ listák ismétléses permutációinak a tere, tehát $\binom{N}{L}$ elemű. A következő állapot valószínűsége így csak a jelenlegi állapottól függ, az átmeneti mátrixot pedig a lista i -edik elemének kiragadásának valószínűsége, w_i adja meg. Egyelőre csak annyit feltételezünk w_i -ről, hogy egyik elemre sem nulla (később a modellezés során hatvány-lecsengő alakot fogunk választani). Az egyes állapotok valószínűségét sorvektor formájában írva az átmeneti mátrixszal jobbról szorozva kapjuk meg az időfejlődést.

Például az $N = 4$, $L = 3$ modellben az állapottér 4 elemű, és a következőképp számozzuk az állapotokat: 1 : $ABBB$, 2 : $BABB$, 3 : $BBAB$, 4 : $BBBA$

Így az átmeneti mátrix:

$$W = \begin{pmatrix} w_1 & w_2 + w_3 + w_4 & 0 & 0 \\ w_2 & w_1 & w_3 + w_4 & 0 \\ w_3 & 0 & w_1 + w_2 & w_4 \\ w_4 & 0 & 0 & w_1 + w_2 + w_3 \end{pmatrix} \quad (3)$$

Állítás. Az átmeneti mátrix duplán sztochasztikus (azaz minden sor-, és oszlopösszeg is 1).

Bizonyítás. Először térjünk át az ismétlés nélküli permutációk terén felírt Markov-láncre, aminek összehúzottjaként tekinthetünk a fenti konstrukcióra. Itt az állapottér az $123...N$ számok permutációi ($N!$ elemű tér), és a számok közül $N - L$ db. A -ra, L db. pedig B -re vonatkozik. Megállapíthatjuk, hogy a Markov-lánc ergodikus, hiszen bármilyen permutációt előállíthatunk úgy, hogy egy tetszőleges elemet kiemelve a sor elejére rakunk (hasonlóan igaz az $AA...B...$ ismétléses permutációk terén felírt Markov-láncre is). Az átmeneti mátrix sorösszege definíció szerint 1, mert valamilyen állapotban biztosan lesz a lista az időlépés után is.

Egy $p_1 p_2 p_3 ... p_N$ permutáció előállítható w_1 valószínűséggel önmagából, w_2 valószínűséggel $p_2 p_1 p_3 ... p_N$ -ből, $...$ w_N valószínűséggel $p_2 p_3 ... p_N p_1$ -ből, és másképp nem. Egy tetszőleges állapotba ugrást az átmeneti mátrix oszlopai tartalmazzák, tehát $\sum w_i = 1$ lesz az összes oszlopösszeg.

Egy duplán sztochasztikus mátrixnak sajátvektora az $(111...1)$ vektor 1 sajátértékkel, és ez lesz a stacionárius eloszlás (ami egyértelmű, mert ergodikus a lánc). Tehát a stacionárius eloszlás egyenletes, azaz minden permutáció egyenlően valószínű hosszú idő után. Innen leolvashatjuk, hogy az $A...B...$ állapotok ismétléses permutációinak terén is az egyenletes eloszlás a stacionárius, mert minden ilyen ismétléses permutáció $L!(N - L)!$ féle $12...N$ permutációból áll elő. Ebből az is következik, hogy az ismétléses permutációkat léptető mátrix is szükségszerűen duplán sztochasztikus. $\square$

2.2. Mester-egyenlet

A permutációk vagy az ismétléses permutációk terén nézve a Markov-láncot az állapottér $N!$ vagy $\binom{N}{L}$ elemű. Ha csak az érdekel minket, hogy n lépés múlva milyen valószínűséggel lesz telefonálás a lista i -edik helyén, akkor azt egy N hosszúságú vektor írja le. Jelöljük $P_{A,i}(t_n)$ -nel annak a valószínűségét, hogy a lista i -edik eleme A típusú a t_n -edik időpontban. A $P_{A,i}(t_n)$ mennyiségre úgy is gondolhatunk, mint egy $\underline{P}_A(t_n)$ sorvektor i -edik komponensére. Ezen vektor időfejlődését leíró

egyenletet mester-egyenletnek nevezzük. A mester-egyenletet úgy kaphatjuk meg, ha a $\underline{P}_A(t_{n+1})$ valószínűsége alkalmazzuk a teljes valószínűség tételét, feltételezve a legutóbbi lépésre:

$$P_{A,1}(t_{n+1}) = \sum_{k=1}^N P_{A,k}(t_n) w_k \quad (4)$$

$$P_{A,k}(t_{n+1}) = P_{A,k}(t_n) \sum_{i=1}^{k-1} w_i + P_{A,k-1}(t_n) \sum_{i=k}^N w_i, \quad (5)$$

ami azt fejezi ki, hogy az 1. helyre tetszőleges k helyről beugorhat egy A teendő w_k valószínűséggel, amennyiben ott A volt; a k -adik helyre pedig úgy kerülhet A , ha ott eleve az volt, és tőle balról választottunk (ilyenkor nem tolódik a listának ezen része), vagy pedig a $k-1$ -edik helyen A állt, és tőle jobbról választunk (jobbra tolódik a lista ezen része). A mester-egyenlet vektoriális formában:

$$\underline{P}_A(t_{n+1}) = \underline{P}_A(t_n) A \quad (6)$$

ahol az A léptető mátrix tehát a következő alakú:

$$A = \begin{pmatrix} w_1 & \sum_{i=2}^N w_i & 0 & \cdots & 0 \\ w_2 & w_1 & \sum_{i=3}^N w_i & 0 & \vdots \\ w_3 & 0 & w_1 + w_2 & \sum_{i=4}^N w_i & 0 \\ \vdots & \vdots & 0 & \ddots & w_N \\ w_N & 0 & \cdots & 0 & \sum_{i=1}^{N-1} w_i \end{pmatrix} \quad (7)$$

Megjegyzés. Az $L = N - 1$ modellben a Markov-lánc W átmeneti mátrixa és mester-egyenlet A léptető mátrixa megegyezik, ha ugyanúgy számozzuk (rendezzük) az állapotokat.

Megjegyzés. A mester-egyenlet közvetlenül is származtatható a Markov-lánc átmeneti mátrixából (lásd a függelékben).

A mester-egyenlet alkalmazása jelentős dimenziószám-csökkenést eredményez, ezáltal bizonyos kérdések megválaszolása egyszerűbb lehet a mester-egyenlettel, mint a Markov-lánc átmeneti mátrixával dolgozva.

2.3. Egyensúlyi autokorrelációs függvény

A lista időfejlődéséből az elsődlegesen vizsgálni kívánt mennyiség a lista első eleme, ami az elvégzett teendők idősorát adja meg. Persze, ha csak az elvégzett cselekedetek idősorát nézzük, az önmagában nem lenne markovi folyamat.

Vizsgáljuk annak a valószínűségnek az időfejlődését, hogy a lista első eleme A típusú cselekvés, vagyis a $P_A(1, t)$ függvényt! Mivel a mester-egyenlet átmeneti mátrixa szintén duplán sztochasztikus és ergodikus, így tudjuk, hogy stacionárius megoldás az egyenletes lesz,

$$P_A(1, t \rightarrow \infty) = \frac{N-L}{N}, \quad (8)$$

ahol N továbbra is a lista hosszát, L pedig az érdektelen tevékenységek számát jelöli a listában. Ha arra vagyunk kíváncsiak, hogy az egyensúlyi állapotból kitérítve hogyan relaxál a lista, akkor a következő kezdeti feltételt kell választanunk:

$$\underline{P}_A^{(L)}(t_0) = \left(1, \frac{N-L-1}{N-1}, \dots, \frac{N-L-1}{N-1}\right) \quad (9)$$

Az autokorrelációs függvényt pedig a következőképp definiáljuk:

$$A^{(L)}(t) = \frac{P_A(1, t) - \frac{N-L}{N}}{1 - \frac{N-L}{N}} \quad (10)$$

Állítás. Az egyensúlyi autokorrelációs függvény nem függ az L paramétertől, vagyis a listában szereplő A és B tevékenységek számától, csak a lista hosszától.

Bizonyítás. A probléma linearitásából következik (függelékben részletezve). $\square$

2.3.1. Az autokorrelációs függvény időfüggése

Az előzőek értelmében elegendő pl. az N és $L = N - 1$ paraméterekkel jellemzett modellt vizsgálni.

Megfigyelés. A numerikusan vizsgált esetek mindegyikében az A átmeneti mátrix diagonalizálhatónak bizonyult

Állítás. Ha az A átmeneti mátrix diagonalizálható, akkor az autokorrelációs függvény előáll, mint N db. exponenciális függvény összege

Mivel A (duplán) sztochasztikus mátrix, tudjuk, hogy a sajátértékek az egység sugarú körlapon helyezkednek el a komplex számsíkon. Mivel ergodikus a mátrix, azt is tudjuk, hogy $\lambda_1 = 1$ nem degenerált sajátérték, a többi sajátértéket jelöljük λ_i -vel, $|\lambda_i| < 1$ $i = 2..N$ esetén. Feltettük, hogy az A átmeneti mátrix diagonalizálható.

$$A = V^{-1} D V, \quad (11)$$

ahol V sorait az A mátrix baloldali sajátvektorai alkotják (v_i). Ezzel

$$\underline{P}_A(t_0) = \sum_{i=1}^N a_i v_i \quad (12)$$

$$\underline{P}_A(t_n) = \sum_{i=1}^N a_i v_i A^n = \sum_{i=1}^N a_i v_i \lambda_i^n \quad (13)$$

$$P_A(1, t_n) = \sum_{i=1}^N a_i v_{i,1} \lambda_i^n \quad (14)$$

Az időfüggés konkrét alakjához szükségünk lenne a sajátértékek pontos ismeretére, ami már függ a w_i átmeneti valószínűségek explicit alakjától. Mivel a listát az aktuális prioritásoknak megfelelően rendezettnek képzeljük, ezért feltesszük, hogy w_i monoton csökken i -ben. Több példát is vizsgáltam, de a konkrét célunknak megfelelő autokorrelációs függvényt (ami tartalmaz hatványfüggvényszerű tartományt is) csak az a modell nyújtott, amikor w_i hatványfüggvény: $w_i \sim i^{-\sigma}$. Például amikor w_i exponenciális i -ben, akkor az autokorrelációs függvényre hatványfüggvénynél lassabb lecsengést kaptam, végén exponenciális levágással.

2.3.2. Numerikus számítások ($w_i \sim i^{-\sigma}$ modell)

MATLAB program segítségével numerikus számításokat végeztem az egyensúlyi autokorrelációs függvényre különböző hosszú listák (N) esetére a w_i súlyfüggvény különböző σ paramétere mellett.

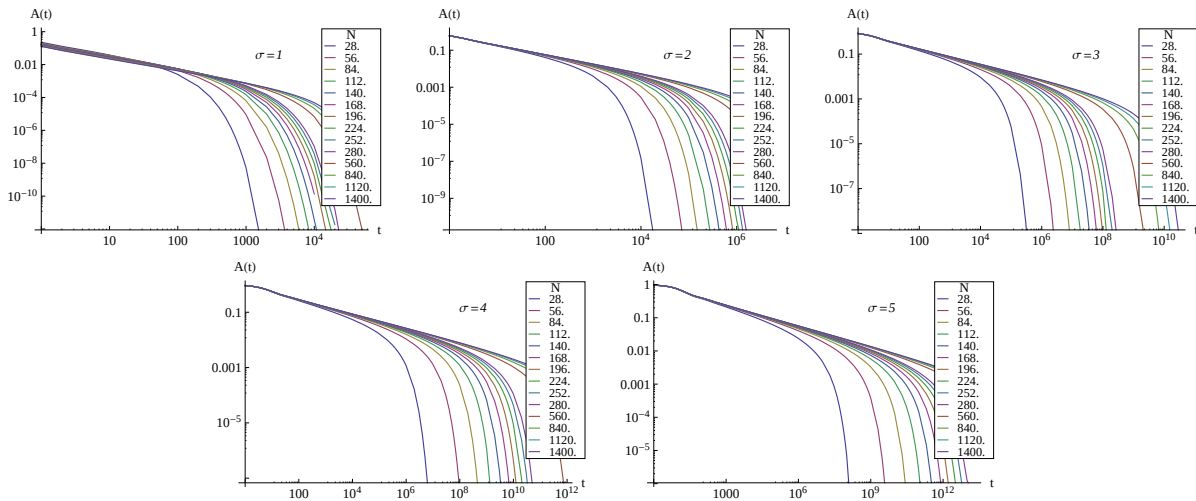
$$w_i = \frac{i^{-\sigma}}{\sum_{i=1}^N i^{-\sigma}} \quad (15)$$

Mivel a modellben a távoli elemek választásának valószínűsége egyre kisebb, azt várjuk hogy hosszú listák esetén az autokorrelációs függvény kezdetben nem függ a lista hosszától (legalábbis ha σ elég nagy).

A nagyobb σ értékek esetén az első pár elem választásának valószínűsége nagyobb, így az autokorrelációs függvény lassabban cseng le.

Programkód ♦ A program az autokorrelációs függvényt a (7) léptetőmátrix egymás utáni használatával számítja ki. Minden dekádban 10 mérési pontot végez, amit úgy ér el, hogy minden 10. lépés után az aktuális léptető mátrix 10. hatányával folytatja az időlépést (ami az előző időlépés 10-szerese).

Eredmények ♦ A szimuláció eredményeit a 3. ábra tartalmazza.



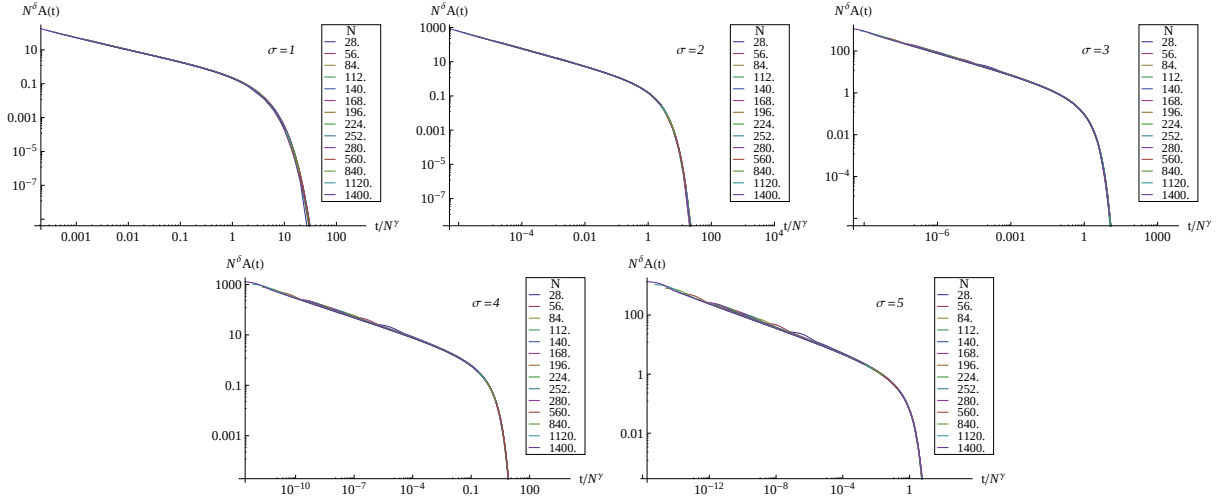
3. ábra. Egyensúlyi autokorrelációs függvények különböző hosszúságú listákra, a σ paraméter 1..5 értékei mellett

A kapott autokorrelációs függvények egy viszonylag széles tartományon hatványfüggvényszerűek, végül exponenciálisan levágnak. Az exponenciális levágás érthető, nagy t -re a (14) egyenletnek megfelelően a második legnagyobb sajátértékkel exponenciálisan lecseng az autokorrelációs függvény. Az exponenciális levágás a lista hosszával egyre kijebb tolódik, úgy tűnik, hogy az

$N \rightarrow \infty$ limeszben az egyensúlyi autokorrelációs függvény hatványfüggvényszerű lecsengést mutat. Ezt a viselkedést végtelen sok exponenciális összegeként kaptuk meg.

Szintén megfigyelhető, hogy a kezdeti tranziens a különböző rendszerméretekre fedi egymást. Ezt a viselkedést a legkisebb sajátértékek tulajdonságai határozzák meg.

Megfigyelés. Adott σ paraméter mellett a különböző rendszerméretekre vonatkozó autokorrelációs függvények összeskálázhatók (4. ábra, (16) egyenlet).



4. ábra. Egyensúlyi autokorrelációs függvények tranziens követő lecsengő részei különböző hosszúságú listákra összeskálázhatók adott σ mellett (a skálázáshoz használt γ és δ paraméterek a 2. táblázatban találhatóak). Az ábrán látható egybeskálázódott görbék menetét az $f_\sigma(t)$ függvény írja le ((16) egyenlet)

$$A(t, N, \sigma) = N^{-\delta} f_\sigma\left(\frac{t}{N^\gamma}\right) \quad (16)$$

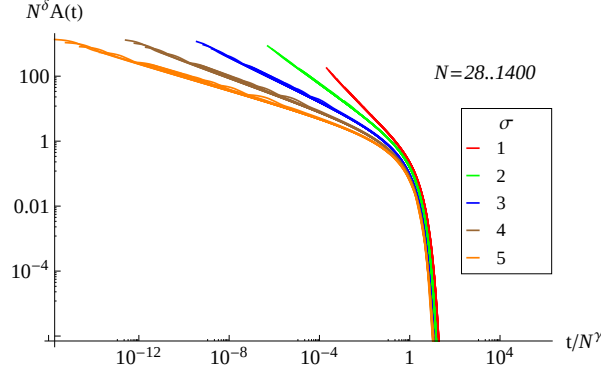
Az egyes σ -knak megfelelő δ és γ paramétereket a 2. táblázat tartalmazza. Az 5. ábrán az N -ben

σ	0.5	0.8	1	1.5	1.8	2	3	4	5
γ	1.07	1.15	1.17	1.57	1.8	2	3	4	5
δ	1	1	1	1	1	1	1	1	1

2. táblázat. γ és δ skálázó paraméterek különböző σ értékek mellett

összeskálázott autokorrelációs függvények láthatók egy ábrán a különböző σ értékekre. A (16) egyenletből következik, hogy az egyensúlyi autokorrelációs függvény (tranziens követő része) általánosított homogén függvénye a lista hosszának, és az időnek:

$$A(b^{\gamma_t} t, \frac{N}{b}, \sigma) = b^\delta N^{-\delta} f_\sigma\left(\frac{b^{\gamma_t} t}{N^\gamma}\right), \quad (17)$$



5. ábra. N -ben összeskálázott autokorrelációs függvények különböző σ paraméterek mellett.

amiből $y_t = -\gamma$ választással,

$$A(t, N, \sigma) = b^{-\delta} A(b^{-\gamma} t, \frac{N}{b}, \sigma) \quad (18)$$

Innen leolvasható ($b = N$ esetén), hogy

$$f_{\sigma}\left(\frac{t}{N^{\gamma}}\right) = A\left(\frac{t}{N^{\gamma}}, 1, \sigma\right) \quad (19)$$

másrészt $b = t^{\frac{1}{\gamma}}$ választással (16) átírható a következő formába:

$$A(t, N, \sigma) = t^{-\frac{\delta}{\gamma}} \tilde{f}_{\sigma}\left(\frac{t}{N^{\gamma}}\right), \quad (20)$$

ahol

$$\tilde{f}_{\sigma}\left(\frac{t}{N^{\gamma}}\right) = A\left(1, \frac{N}{t^{1/\gamma}}, \sigma\right). \quad (21)$$

Az 5. ábrán az összeskálázott görbéket, tehát lényegében az $f_{\sigma}(t)$ függvényt láthatjuk. Ezen görbék kezdeti, log-log skálán egyenes szakaszára hatványfüggvényt illesztettem, a hatványfüggvény kitevőit a 3. táblázat tartalmazza. Az exponenseket úgy határoztam meg, hogy az adatsor logaritmusára illesztettem egyenest a legkisebb négyzetek módszerével, és annak a meredekségét vettem. Az illesztés hibáit azért nem tüntettem fel a táblázatban, mert az kisebb volt, mint az utolsó tizedesjegy. Azonban a meredekség értékét befolyásolja az is, hogy milyen tartományt jelölök ki az illesztésre, ennek optimalizálása pedig nem törekedtem, csak a szememre hagytam (tehát a legkisebb négyzetek módszerével kapott hiba alulbecsüli az exponens becslésének valódi hibáját). Mivel az autokorrelációs függvény időfüggését f_{σ} hordozza, ennek a kitevője megegyezik az autokorrelációs függvény kitevőjével (ezekkel összezsengő kitevőket kaptam, amikor közvetlenül a legnagyobb méretű listával előállított autokorrelációs függvényekre is elvégeztem az illesztést).

Tehát a tranzienszt követő és az exponenciálisan levágó tartományok között

$$A(t) \sim t^{-\alpha}. \quad (22)$$

σ	1	2	3	4	5
α	0,755	0,535	0,356	0,274	0,222

3. táblázat. Az autokorrelációs függvény, ill. $f_\sigma(t)$ hatványlecsengésének kitevője

A 3. táblázat alapján az a sejtés fogalmazódott meg, hogy $N \rightarrow \infty$ esetben fennáll, hogy

$$\alpha \stackrel{?}{=} \frac{1}{\sigma}, \quad (23)$$

az ettől való eltérést pedig a véges méretből származó korrekciók okozhatják. A (23) sejtés összhangban van a (20) skálázó összefüggéssel, ugyanis $N \rightarrow \infty$ esetben

$$A(t, N, \sigma) = t^{-\frac{\delta}{\gamma}} \tilde{f}_\sigma\left(\frac{t}{N^\gamma}\right) \rightarrow t^{-\frac{\delta}{\gamma}} \tilde{f}_\sigma(0), \quad (24)$$

tehát $\alpha = \frac{\delta}{\gamma}$, ami a 2. táblázat alapján elég nagy σ értékekre a (23) összefüggést adja.

2.4. Várakozási idők és az események között eltelt idő eloszlása

Várható értékek

Vizsgáljuk meg, hogy a lista i -edik eleme várhatóan mikor kerül sorra! Legyen τ_i az a valószínűségi változó, ami megadja, hogy a kezdetben az i -edik helyen ülő elem mikor kerül a lista elejére. Először határozzuk meg τ_i várható értékét! A várható érték kiszámításakor feltételezhetünk az első lépés eredményére: ha először a lista első $i-1$ eleméből választunk (ezt az eseményt jelöljük $E_{1..i-1}$ -vel), akkor az i -edik elem helyben marad, ha pedig $i+1$ -től felfelé választunk ($E_{i+1..N}$), akkor a vizsgált elem innentől az $i+1$ -edik helyen fog ülni. Végül, ha pont az i -edik elemet választjuk (E_i), akkor az a lista elejére ugrik. Tehát

$$\mathbb{E}(\tau_i) = \mathbb{E}(\tau_i | E_{1..i-1})\mathbb{P}(E_{1..i-1}) + \mathbb{E}(\tau_i | E_i)\mathbb{P}(E_i) + \mathbb{E}(\tau_i | E_{i+1..N})\mathbb{P}(E_{i+1..N}) \quad (25)$$

$$= [\mathbb{E}(\tau_i) + 1] \sum_{j=1}^{i-1} w_j + w_i + [\mathbb{E}(\tau_{i+1}) + 1] \sum_{j=i+1}^N w_j \quad (26)$$

Így ki tudjuk fejezni az i -edik elem lista elejére kerülésének várható értékét az $i+1$ -edik elem várható várakozási idejével:

$$\mathbb{E}(\tau_i) = \frac{1 + \mathbb{E}(\tau_{i+1}) \sum_{j=i+1}^N w_j}{1 - \sum_{j=1}^{i-1} w_j} \quad (27)$$

A lista utolsó elemére tudjuk a várakozási idő eloszlását, egyszerű (optimista) geometriai eloszlás w_n paraméterrel, ugyanis minden lépésben mindentől függetlenül w_N valószínűséggel választjuk ki az N -edik elemet.

$$\mathbb{P}(\tau_N = k) = (1 - w_N)^{k-1} w_N \quad (28)$$

$$\mathbb{E}(\tau_N) = \frac{1}{w_N} \quad (29)$$

Ezt felhasználva visszafejthetjük a rekurziót, és a következő eredményt kapjuk:

$$\mathbb{E}(\tau_k) = \frac{1 + N - k}{\sum_{j=k}^N w_j} \quad (30)$$

Speciálisan a lista első elemére a visszatérési idő várható értéke N , függetlenül attól, hogy milyen a listából való választás súlyfüggvénye (legalábbis amíg az ergodikus bejárást biztosítja).

Rekurzió a generátorfüggvényre (várakozási idők)

A fentihez hasonló gondolatmenettel a τ_l -ek eloszlására is tudunk egy rekurziós összefüggést adni.

$$\mathbb{P}(\tau_l = k) = \mathbb{P}(\tau_l = k | E_{1..l-1})\mathbb{P}(E_{1..l-1}) + \mathbb{P}(\tau_l = k | E_l)\mathbb{P}(E_l) + \mathbb{P}(\tau_l = k | E_{l+1..N})\mathbb{P}(E_{l+1..N}) \quad (31)$$

$$= \mathbb{P}(\tau_l = k - 1) \sum_{i=1}^{l-1} w_i + \delta_{k,1} w_l + \mathbb{P}(\tau_{l+1} = k - 1) \sum_{i=l+1}^N w_i \quad (32)$$

Ebből az egyenletből felírhatunk egy rekurziós összefüggést a τ_l eloszlásának generátorfüggvényére (megszorozzuk a fenti egyenletet z^k -nal, és összeadjuk $k = 1$ -től ∞ -ig). Jelölje $G_{\tau_l}(z)$ a τ_l generátorfüggvényét:

$$G_{\tau_l}(z) \equiv \sum_{k=1}^{\infty} \mathbb{P}(\tau_l = k) z^k \quad (33)$$

A rekurziós összefüggés pedig:

$$G_{\tau_l}(z) = \frac{z w_l + z G_{\tau_{l+1}}(z) \sum_{i=l+1}^N w_i}{1 - z \sum_{i=1}^{l-1} w_i} \quad (34)$$

τ_N -re ismerjük a generátorfüggvényt, helyettesítsük be a (28) sűrűségfüggvényt a (33) definícióba:

$$G_{\tau_N}(z) = \sum_{k=1}^{\infty} (1 - w_N)^{k-1} w_N z^k = z w_N \sum_{k=0}^{\infty} [z(1 - w_N)]^k = \frac{z w_N}{1 - z(1 - w_N)} \quad (35)$$

A $G_{\tau_N}(z)$ generátorfüggvény csak $z < \frac{1}{1-w_N}$ -re van értelmezve. A rekurziót nem sikerült a várható értékhez hasonlóan zárt formában visszafejtenem, viszont numerikusan egyszerűen kiértékelhető.

Megjegyzés. A várható értékre a (27) rekurzió egyszerűen levezethető a (34) generátorfüggvényekre vonatkozó rekurzióból, ha ez utóbbit deriváljuk z szerint az 1 helyen.

Események között eltelt idő az $L = N - 1$ modellben

Ha az összesen 1 db A -t tartalmazó modellt vizsgáljuk, akkor a τ_1 valószínűségi változó éppen az A teendőre vonatkozó események között eltelt időt adja meg. A τ valószínűségi változók súlyfüggvényeit a generátorfüggvény különböző rendű deriváltjaiként kaphatjuk meg:

$$P(\tau_l = k) = \frac{1}{k!} G^{(k)}(0) \quad (36)$$

A τ_1 súlyfüggvényére vonatkozó numerikus számítások eredményeit különböző σ paraméterek mellett a 6. ábra tartalmazza. A várakozási idő széles tartományon hatványfüggvénnyel közelíthető:

$$P(t_{ie}) \sim t_{ie}^{-\beta}, \quad (37)$$

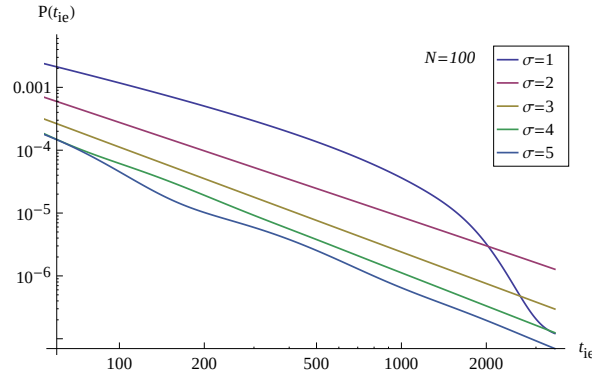
amit exponenciális lecsengés követ. Az exponenciális levágás $(1 - w_N)^t$ alakú, ami abból származik, ha az A teendő a lista végére kerül (nagyon hosszú várakozási idők esetén ez valószínűleg bekövetkezett). Ugyanez az exponenciális lecsengés jellemzi a lista távolabbi elemeinek várakozási idejét is. A hatványfüggvény szerint viselkedő tartományon illesztéssel meghatároztam hatványfüggés kitevőjét különböző σ értékekre (4. táblázat). Az illesztési eredmények azt sugallják, hogy a hatványkitevő eleget tesz a $\frac{2\sigma-1}{\sigma}$ összefüggésnek.

$$\beta \stackrel{?}{=} 2 - \frac{1}{\sigma} \quad (38)$$

Megfigyelhető, hogy a σ paraméter növekedésével kis hullámok jelennek az események között eltelt idő súlyfüggvényében. Ez kapcsolatban áll az autokorrelációs függvényben megjelenő egyre nagyobb tranziens szakasszal, amit pedig az okoz, hogy az alacsony sajátértékeknél megjelenik egy „lyuk” a mester-egyenlet (7) átmeneti mátrixának spektrumában.

σ	2	3	4	5
β	1.501	1.668	1.753	1.7998

4. táblázat. $P(\tau_1)$ hatványlecsengésének paramétere különböző σ értékek mellett



6. ábra. Események között eltelt idő eloszlásának súlyfüggvénye az $L = N - 1$ modellben különböző σ paraméterértékek mellett.

Események között eltelt idő az $L < N - 1$ modellben

A következő kérdés az, hogy ha több A teendő is szerepel a listában, akkor milyen lesz az A teendőre vonatkozó τ_A események közötti idő. Ha az A teendők a lista $i_1 \dots i_{N-L}$ helyein találhatók, ahol $i_1 = 1$, akkor

$$\tau_A = \min_k \{\tau_{i_k}\} \quad (39)$$

τ_A eloszlásának számolásához ismernünk kellene a τ_{i_k} valószínűségi változók együttes eloszlását. Ez általánosan nehéz probléma, speciális esetként megoldható az $L = 1$ esete, amikor csak egy B teendő szerepel a listában.

Példa. τ_A eloszlása $L=1$ esetben

Ha tudjuk, hogy a B teendő pontosan hol helyezkedik el a listában, akkor fel tudjuk írni τ_A súlyfüggvényét. A stacionárius eloszlás egyenletes, tehát egyensúlyban B a $2 \dots N$ helyek bármelyikén azonos valószínűséggel helyezkedik el. Az egyensúlyi súlyfüggvény számításához alkalmazzuk a teljes valószínűség tételét!

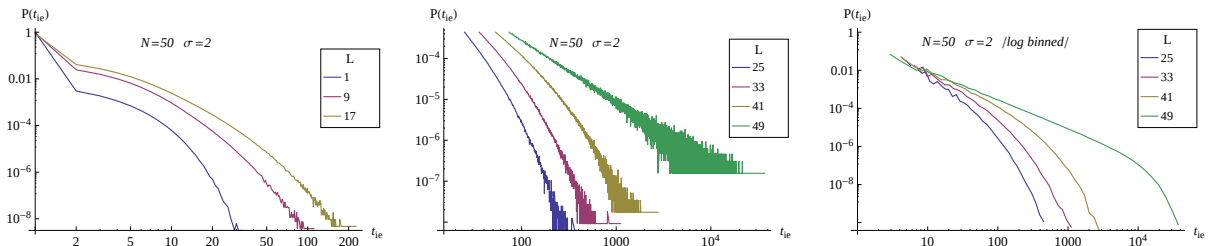
$$\mathbb{P}(\tau_A = 1) = \sum_{k=2}^N \mathbb{P}(\tau_A = 1 | \{B \text{ a } k\text{-adik helyen van}\}) \mathbb{P}(\{B \text{ a } k\text{-adik helyen van}\}) \quad (40)$$

$$= \sum_{k=2}^N (1 - w_k) \frac{1}{N-1} = 1 - \frac{1 - w_1}{N-1} \quad (41)$$

$$\mathbb{P}(\tau_A = l) = \sum_{k=2}^N \mathbb{P}(\tau_A = l | \{B \text{ a } k\text{-adik helyen van}\}) \mathbb{P}(\{B \text{ a } k\text{-adik helyen van}\}) \quad (42)$$

$$= \sum_{k=2}^N w_k (1 - w_1) w_1^{l-2} \frac{1}{N-1} = (1 - w_1)^2 w_1^{l-2} \frac{1}{N-1} \quad (43)$$

Azt láthatjuk tehát, hogy ha csak egy darab A szerepel a listában, akkor az események közötti idő (hatványfüggvényt követő) exponenciális levágása $(1 - w_N)^l$ -nel megy, $N - 1$ darab A esetén pedig w_1^l -nel cseng le. Nem triviális tehát, hogy a listában az A tevékenységek arányát növelve hogyan változik az események közötti idő súlyfüggvénye. Ecélből szimulációkat futtattam rögzített N mellett különböző L paraméterekkel, és mértem a τ_A események közötti idő hisztogramját. A szimulációs eredmények a 7. ábrán láthatók. Megfigyelhető, hogy a hatványfüggvénnel jellemezhető szakasz eltűnik, ha növeljük az A tevékenységek számát a listában (L csökken), és valamilyen hatványfüggvénnél gyorsabb, de exponenciálisnál lassabb lecsengést láthatunk.



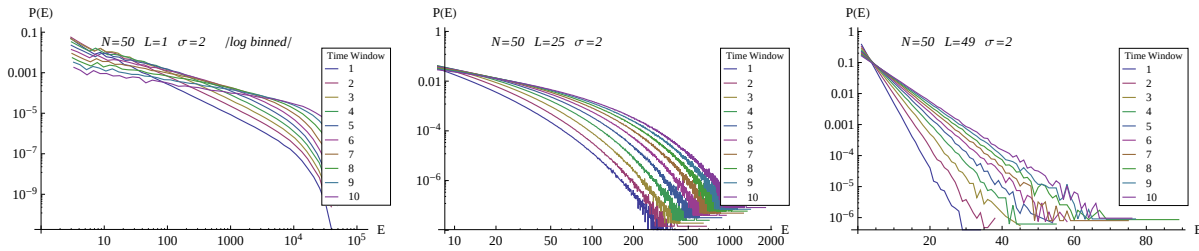
7. ábra. Események között eltelt idő súlyfüggvénye különböző összetételű listák esetén (szimulációs eredmények)

2.5. $P(E)$ eloszlás

Az autokorrelációs függvény és az események között eltelt idő eloszlása után már csak azt kell megnéznünk, milyen $P(E)$ eloszlást produkál a modell. A villanásos periódusokat most úgy definiáljuk, mint egy időablak által rögzített hosszúságú B sorozatok közötti események sorozatát. Az ezekben előforduló A típusú események számát jelöljük itt E -vel. Például ha az időablakot 4-nek választjuk, akkor egy villanásos periódus a következő idősorhoz hasonlít:

...BBBBABAABBBAAAAABAABBBABBBB...

A $P(E)$ eloszlás mérésére szimulációkat futtattam (8. ábra). A szimulációs eredmények szerint ebben a modellben a $P(E)$ súlyfüggvény erősen függ az időablak szélességétől, és nem triviális L függést mutat. Vizsgáljunk két speciális esetet:



8. ábra. $P(E)$ eloszlás súlyfüggvénye (szimulációs eredmények)

Példa. $L = N - 1$ eset

Amikor a listában csak 1db. A esemény található, akkor exponenciális a $P(E)$ súlyfüggvény (log-lin skálán egyenes), viszont az időállandó függ az időablak méretétől (8. ábra). Ez következik a (2) egyenletből, diszkrét esetre vonatkoztatva, ugyanis ilyenkor az A cselekvés egymást követő visszatérési idejei függetlenek.

Példa. $L = 1$ eset

A másik határeset az $L = 1$ esete. Ekkor ha az időablakot 1-nek választjuk, a $P(E)$ súlyfüggvény az első elemre (ami ebben az esetben az egyetlen B típusú elem) vonatkozó várakozási időnek felel meg:

$$P(E = k) = \mathbb{P}(\tau_1 = k + 1 | \tau_1 \geq 2) = \frac{\mathbb{P}(\tau_1 = k + 1)}{1 - \mathbb{P}(\tau_1 = 1)} \mathbb{I}(k \geq 2), \quad (44)$$

amiről pedig korábban láttuk, hogy hatványlecsengésű.

2.6. Diszkusszió

Összegzésképpen megállapíthatjuk, hogy a modell képes produkálni hatványfüggvény szakaszt tartalmazó autokorrelációs függvényt, és a hatványfüggvény kitevője hangolható a σ paraméter változtatásával. A telefonálási cselekvések aránya az összes cselekvésből hangolható az $\frac{L}{N}$ arány változtatásával, N pedig az exponenciális levágás helyét tolja el. A modell szempontjából azonban az

események között eltelt idő eloszlása és a $P(E)$ eloszlás komplementer mennyiségekként viselkednek: $L = N - 1$ esetén az események közötti idő eloszlása lesz hatványfüggvény és a $P(E)$ eloszlás exponenciális, míg az $L = 1$ esetben fordítva.

Az $L = N - 1$ modellben sejtésként megfogalmazódott két skálatörvény:

$$\alpha = \frac{1}{\sigma} \quad (45)$$

$$\beta = 2 - \frac{1}{\sigma}, \quad (46)$$

illetve ezek összegére

$$\alpha + \beta = 2. \quad (47)$$

Ez utóbbi összefüggést a 4. fejezetben bizonyítom, és kiderül, hogy a modellek egy szélesebb családjára is általánosítható.

A 2. fejezet elején, a modell bevezetésénél megjegyzésként szerepelt, hogy a vizsgált modell időtükrözöttje a *top-in shuffle* keverési módszer egyfajta általánosítása. Kártyakeverés esetén az egyik érdekes mennyiség a keverés n lépése után kapott valószínűségi eloszlás variációs távolsága az egyenletes eloszlástól (a permutációk terén). Ez a mennyiség a standard keverési eljárások esetén egy viszonylag jól meghatározott számú keverés után hirtelen lecsökken, a *top-in shuffle* esetén ez az érték kb. $N \log N$ [13], ahol N a kártyák száma (a témakörhöz kapcsolódó leghíresebb eredmény, hogy a *riffle shuffle* keverési eljárással egy csomag francia kártya esetén 7 keverés elég [14]).

Az általam vizsgált modellben a hatványfüggvény szerint csökkenő beszúrási valószínűség lényegesen csökkenti a keveredés mértékét. Érdekes kérdés lehet, hogy ebben az esetben is látnánk-e egy meredek „törést” a variációs távolságban, és ha igen, akkor az hogyan skálázik a csomag kártya méretével a modell különböző σ paraméterei esetén.

3. Módosított modell

Ebben a fejezetben az előző modell egy kicsit módosított változatát vizsgálom. Az egyéneket itt is egy N hosszúságú, $N - L$ db. A és L db. B tevékenységet tartalmazó listával modellezzük, és minden lépésben kiválasztjuk a lista valamely elemét a w_i súlyfüggvény által meghatározott valószínűséggel. Itt azonban a kiválasztott tevékenység nem kerül a lista elejére, csak egyetel feljebb ugrik a prioritási listán. Ha a választott elem a lista első eleme, akkor az helyben marad. Az elvégzett cselekvés minden lépésben a lista első helyén szereplő tevékenység. A lista dinamikája itt is egy ergodik Markov-láncot definiál, a stacionárius megoldás az egyenletes. Az autokorrelációs függvény számolásához itt is felírható egy mester-egyenlet:

$$\underline{P}_A(t_{n+1}) = \underline{P}_A(t_n)A, \quad (48)$$

ahol a mester-egyenlet átmeneti mátrixa a következő alakú:

$$A = \begin{pmatrix} 1 - w_2 & w_2 & 0 & \cdots & 0 \\ w_2 & 1 - w_2 - w_3 & w_3 & 0 & \vdots \\ 0 & w_3 & 1 - w_3 - w_4 & w_4 & 0 \\ \vdots & \vdots & \ddots & \ddots & w_N \\ 0 & 0 & \cdots & w_N & 1 - w_N \end{pmatrix} \quad (49)$$

3.1. Autokorrelációs függvény

Az autokorrelációs függvény:

$$A^{(L)}(t) = \frac{P(1, t) - \frac{N-L}{N}}{1 - \frac{N-L}{N}} \quad (50)$$

Ez a mennyiség nem függ L értékétől (kivéve a triviális $L = 0$ és $L = N$ eseteket), ahogy az eredeti modellnél is láttuk (bizonyítás a függelékben). Az eredeti modellhez hasonlóan a különböző méretű listákra kapott autokorrelációs függvények összeskálázhatók:

$$A(t, N, \sigma) = N^{-\delta} f_{\sigma} \left(\frac{t}{N^{\gamma}} \right) \quad (51)$$

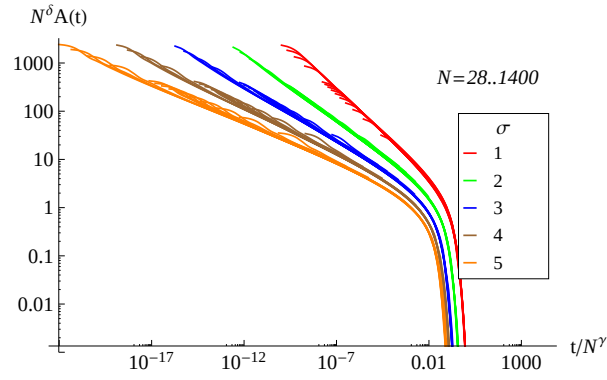
A különböző σ értékekhez tartozó γ és δ skálázó paraméterek az 5. táblázatban szerepelnek, az összeskálázott autokorrelációs függvények pedig a 9. ábrán láthatók. Egyszerű átalakításokkal a

σ	1	2	3	4	5
γ	3.17	4	5	6	7
δ	1.08	1.08	1.08	1.08	1.08

5. táblázat. A különböző rendszermérehez tartozó autokorrelációs függvények összeskálázásához használt γ és δ paraméterek σ különböző értékeire

skálatulajdonság a következő alakba is írható,

$$A(t, N, \sigma) = t^{-\frac{\delta}{\gamma}} \tilde{f}_{\sigma} \left(\frac{t}{N^{\gamma}} \right). \quad (52)$$



9. ábra. A σ paraméter azonos értékeihez tartozó autokorrelációs függvények összeskálázhatók

Ebből az alakból $N \rightarrow \infty$ esetre kapjuk, hogy $A(t, N, \sigma) \rightarrow t^{-\frac{\delta}{\gamma}} \tilde{f}_\sigma(0)$, tehát $\alpha = \frac{\delta}{\gamma}$. Úgy tűnik, hogy σ elég nagy értékeire $\gamma = \sigma + 2$, valamint $\delta = 1.08$, így

$$\alpha = \frac{1.08}{\sigma + 2}. \quad (53)$$

Ebből a γ paraméter valószínűleg meghatározható a modell analitikus vizsgálatából, viszont a δ paraméter 1-től való eltérését egyelőre nem látom miből származhat.

Az autokorrelációs függvény log-log skálán lineáris szakaszára hatványfüggvényt illesztettem a legnagyobb listaméretek esetén. Az exponensek a 6. táblázatban szerepelnek, és közelítőleg teljesítik az (53) összefüggést, a kis (kb. 0.03 nagyságrendű) eltérés származhat a δ skálázó exponens és az illesztés hibájából, valamint a véges méret korrekciók is szerepet játszhatnak. Az illesztés hibájáról itt is elmondható, hogy az egyenes szakasz meghatározása miatt itt is nagyobb, mint a legkisebb négyzetek módszerével kapott hiba, ami kisebb, mint az táblázatban feltüntetett utolsó tizedesjegy.

σ	1	2	3	4	5
α	0.359	0.271	0.219	0.184	0.16
β	1.639	1.734	1.787	1.817	(?)
$\alpha + \beta$	1.998	2.005	2.006	2.001	(?)

6. táblázat. Az autokorrelációs függvény (α) és az események közötti idő (β) exponensei

3.2. Visszatérési idő

Ebben a fejezetben a lista első elemének a visszatérési idejét szeretnénk kiszámolni, ami az $L = N - 1$ modellben megegyezik az események között eltelt idővel. Felírtam egy rekurziót a visszatérési időre, amit véges hosszú listák esetén numerikusan értékeltem ki.

Legyen β_k (a $t = 0$ időpontban) a k -adik listaelem várakozási ideje, amíg a $k - 1$ pozíciót először

eléri (balra lépési idő). Legyen ξ Bernoulli eloszlású valószínűségi változó: $\mathbb{P}(\xi = 1) = w_2$ és $\mathbb{P}(\xi = 0) = 1 - w_2$. Ekkor a lista első elemének visszatérési ideje:

$$\tau = 1 + \xi \beta_2. \quad (54)$$

Ez azt jelenti, hogy $1 - w_2$ valószínűséggel $\tau = 1$, és w_2 valószínűséggel a vizsgált teendő a második helyre ugrik, ahonnan β_2 időt kell várni a visszatérésig. A β_k -k generátorfüggvényére viszont fel tudunk írni egy rekurziót.

$$\mathbb{P}(\beta_k = 1) = w_k \quad (55)$$

$$\begin{aligned} \mathbb{P}(\beta_k = s) = & \mathbb{P}(\beta_k = s | \{ 1. \text{ választott elem } < k \text{ vagy } > k + 1 \}) (1 - w_k - w_{k+1}) + \\ & + \mathbb{P}(\beta_k = s | \{ 1. \text{ választott elem } = k \}) w_k + \\ & + \mathbb{P}(\beta_k = s | \{ 1. \text{ választott elem } = k + 1 \}) w_{k+1} \end{aligned}$$

$$\mathbb{P}(\beta_k = s) = \mathbb{P}(\beta_k = s - 1) (1 - w_k - w_{k+1}) + \delta_{1s} w_k + \mathbb{P}(\beta_{k+1} + \beta_k = s - 1) w_{k+1}$$

Bevezetjük a $b_k(s) = \mathbb{P}(\beta_k = s)$ jelölést, és kifejtjük a jobb oldali tagot, mint konvolúciót:

$$b_k(s) = b_k(s - 1) (1 - w_k - w_{k+1}) + \delta_{1s} w_k + \sum_{l=1}^{\infty} b_k(s - l - 1) b_{k+1}(l) w_{k+1} \quad (56)$$

Vezessük be a β_k valószínűségi változók generátorfüggvényét: $G_k(z) = \sum_{s=1}^{\infty} b_k(s) z^s$. Szorozzuk be az (56) egyenletet z^s -nel és adjuk össze az így nyert egyenleteket $s = 1$ -től ∞ -ig!

$$G_k(z) = z G_k(z) (1 - w_k - w_{k+1}) + z w_k + G_k(z) G_{k+1}(z) \quad (57)$$

$$G_k(z) = \frac{z w_k}{1 - z(1 - w_k - w_{k+1}) - z G_{k+1}(z) w_{k+1}} \quad (58)$$

Véges listák (N) esetén tudjuk, hogy az utolsó elemre a balra lépés ideje geometriai eloszlású w_N paraméterrel.

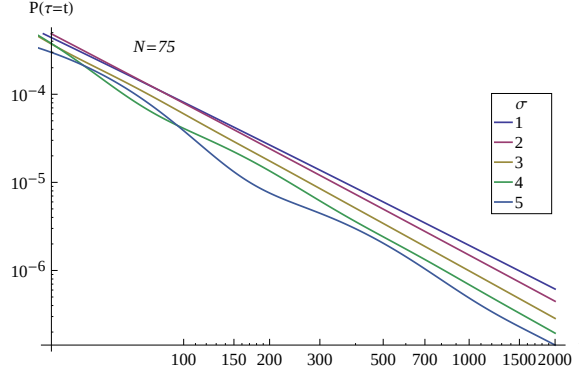
$$G_N(z) = \frac{z w_N}{1 - z(1 - w_N)} \quad (59)$$

Innen $G_2(z)$ visszafejthető számítógéppel, ahonnan a $b_2(s)$ -ek kifejthetők a generátorfüggvény

$$b_k(s) = \frac{1}{s!} G_k^{(s)}(0) \quad (60)$$

tulajdonsága segítségével.

Megjegyzés. A várható értékekre vonatkozó rekurzió analitikusan is visszafejthető: $\mathbb{E}[\beta_k] = \frac{N-k+1}{w_k}$, ahonnan az első elem várakozási idejére az eredeti modellhez hasonlóan az $\mathbb{E}[\tau] = N$ értéket kapjuk.



10. ábra. A lista első elemének visszatérési idő súlyfüggvénye különböző σ értékek mellett. Ez az $L = N - 1$ modellben az események között eltelt idő eloszlása

Numerikusan meghatároztam a $b_2(s)$ függvényt $N = 75$ -ös listaméret mellett a σ paraméter 1 – 5 értékeire. Az (54) egyenlet alapján a τ visszatérési idő eloszlása:

$$\mathbb{P}(\tau = 1) = 1 - w_2 \quad (61)$$

$$\mathbb{P}(\tau = s) = w_2 b_2(s - 1) \quad s \geq 2\text{-re} \quad (62)$$

Az így kapott visszatérési idő eloszlás a 10. ábrán, az illesztett exponensek pedig a 6. táblázatban láthatók.

Az eredeti modellhez hasonlóan itt is egyre hullámosabbá válik az események közötti idő eloszlása növekvő σ paraméterek mellett. Ezért a $\sigma = 5$ esetben már nem is tudtam a numerikusan kiértékelt szakaszból meghatározni a lecsengés exponensét. A 6. táblázatban szereplő exponensek jó közelítéssel teljesítik az eredeti modellnél megfogalmazott $\alpha + \beta = 2$ sejtést. Ezt összevetve az autokorrelációs függvény exponensére vonatkozó (53) kifejezéssel,

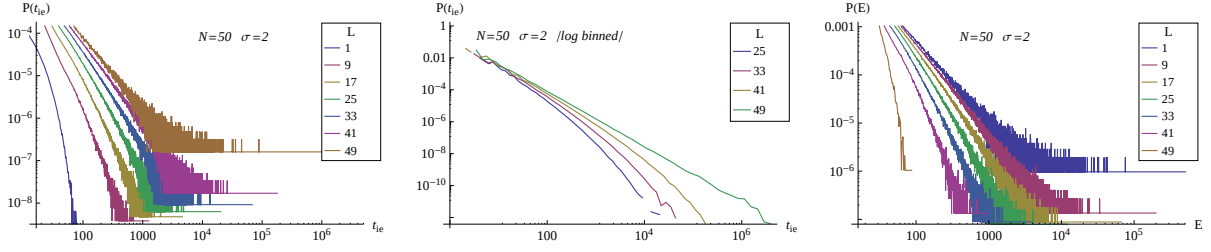
$$\beta = 2 - \frac{\delta}{\sigma + 2} \quad \sigma \geq 2\text{-re} \quad (63)$$

Események között eltelt idők L -függése

Szimulációkat futtattam az események közötti idők eloszlásának meghatározására rögzített N mellett az L paraméter különböző értékeire. Az eredeti modellhez hasonlóan itt is azt tapasztaltam, hogy az $\frac{L}{N}$ arány csökkentésével (telefonálási cselekvések számának növelésével) az események közötti idő eloszlása hatványfüggvéynél gyorsabban cseng le, az $L = 1$ esetben pedig exponenciális (11. ábra).

3.3. $P(E)$ eloszlás

A $P(E)$ eloszlásra is kvalitatívan ugyanazt az eredményt kaptam, mint az eredeti modell esetén. $L = 1$ esetén hatványfüggvény szerint cseng le, $L = N - 1$ esetén pedig exponenciális (ennek



11. ábra. Események közötti idő és a $P(E)$ eloszlás súlyfüggvényeinek függése az L paramétertől (szimulációs eredmények)

magyarázatát lásd az eredeti modellnél). A két szélső eset közt pedig interpolál a $P(E)$ eloszlás (11. ábra).

Autokorrelációs függvény skálázása és véletlen bolyongás

A 2. és a 3. fejezetben tárgyalt (eredeti, ill. módosított) modellek lényegében ugyanúgy viselkednek a vizsgált mennyiségek szempontjából, az egyetlen különbség az autokorrelációs függvény összeskálázásánál használt γ és δ exponensek értékében jelentkezik. Az eredeti modell esetén $\gamma = \sigma$, a módosított modellben pedig $\gamma = \sigma + 2$ választással érhetjük el, hogy a különböző méretekre számolt autokorrelációs függvények fedjék egymást. Érdekes megfigyelés, hogy a lista elemeinek bolyongását vizsgálva szintén előkerülnek a σ , ill. a $\sigma + 2$ exponensek.

A módosított modellben ha kiválasztunk egy tetszőleges tevékenységet, és annak útját követjük, akkor egy elsőszomszéd véletlen bolyongást látunk $[1, N] \cap \mathbb{Z}$ -n. Ebben a modellben analitikusan kiszámolható (függelékben részletezve) annak az eloszlása, hogy a lista első helyéről indított tevékenység milyen maximális távolságra jut a visszatérés előtt. Jelöljük ezt a távolságot ξ -vel, ekkor:

$$\mathbb{P}(\xi = x) \sim x^{-(\sigma+2)} \quad (64)$$

Így annak valószínűsége, hogy az első helyről induló elem bolyongása során eléri a lista végét $\sim N^{-(\sigma+2)}$. Ez a mennyiség kiszámolható az eredeti modellben is, ehhez célszerű az időtükrözött változatát vizsgálni. Itt már nem elsőszomszéd-bolyongást végeznek az elemek: a lista elejéről indulva egy tevékenység először a w_i valószínűségek által megadott helyre ugrik, onnan viszont már csak a lista eleje felé tud lépdelni (ha nála távolabbra dobjuk az első elemet), vagy helyben marad. A véletlen-bolyongó elem visszatérés előtt elért maximális távolságának eloszlását (az időtükrözött modellben) definíció szerint $w_x \sim x^{-\sigma}$ adja meg. Annak valószínűsége pedig, hogy a teendő bolyongása során eléri a lista végét $\sim N^{-\sigma}$. Az eredeti, ill. a módosított modellben a kezdetben első helyen álló elem várhatóan $\sim N^\sigma$, ill. $\sim N^{\sigma+2}$ véletlen bolyongás után éri el először a lista végét. Ezek után az autokorrelációs függvény

$$A(t, N, \sigma) = N^{-\delta} A\left(\frac{t}{N^\gamma}, 1, \sigma\right) = N^{-\delta} f_\sigma\left(\frac{t}{N^\gamma}\right) \quad (65)$$

skálázásában megjelenő N^γ tényezőnek szemléletes jelentést adhatunk. Egyelőre nyitott kérdés, hogyan kapcsolódik közvetlenül a véletlen bolyongás az autokorrelációs függvény skálázásához.

4. Független események közötti idővel jellemzett folyamatok

A 2. és a 3. fejezetben tárgyalt modellek $L = N - 1$ határesetében hatványlecsengésű autokorrelációs függvényt és események közötti idő eloszlást tapasztaltunk, melyek exponensei eleget tettek az

$$\alpha + \beta = 2 \quad (66)$$

skálátörvénynek. A két modell egyik közös vonása, hogy a kezdetben a lista elején ülő cselekvés egymást követő visszatérési idejei függetlenek, és azonos eloszlásúak. Ebből már következett, hogy a $P(E)$ eloszlás ezekre a modellekre exponenciálisan lecsengő ((2). egyenlet).

Most megvizsgáljuk az autokorrelációs függvény és az események között eltelt idő eloszlásának kapcsolatát független, azonos eloszlású események közötti idővel jellemzett folyamatok esetére. Legyen $X(t_n)$ az események indikátora: $X(t_n) = 1$, ha a t_n időpontban történt esemény (a vizsgált modellekben ha az A típusú elem áll a lista élén) és $X(t_n) = 0$ egyébként. Ilyenkor az autokorrelációs függvény időfüggő része:

$$\tilde{A}(t_n) = \mathbb{E}[X(t_0)X(t_n)], \quad (67)$$

az $X(t_0) = 1$ kezdőfeltétellel. Ennek lineáris függvénye az (1)-ben definiált autokorrelációs függvény.

Tétel. *Ha független események hatványfüggvény lecsengésű események közötti idő eloszlással jellemezhetők:*

$$P(t_{ie}) \sim t_{ie}^{-\beta} \quad \text{ahogy } t_{ie} \rightarrow \infty,$$

ahol $\beta \in (1, 2)$, akkor az autokorrelációs függvény is hatványlecsengésű:

$$\tilde{A}(t) \sim t^{-\alpha} \quad \text{ahogy } t \rightarrow \infty$$

és a kitevők összegei teljesítik az $\alpha + \beta = 2$ összefüggést.

Bizonyítás. Jelölje τ_m a visszatérési időket: $\tau_0 = 0$, $\tau_{m+1} = \inf\{t_n > \tau_m | X(t_n) = 1\}$, T pedig a visszatérési idők halmazát: $T = \{\tau_m : m = 0, 1, 2, \dots\}$. $X(t_n)$ tehát csak azokra a t_n időpontokra nem nulla, amik elemei a T halmaznak.

Állítás. $\tilde{A}(t_n) = \mathbb{P}(t_n \in T)$

Bizonyítás.

$$\tilde{A}(t_n) = \mathbb{E}[X(t_0)X(t_n)] = \mathbb{E}[X(t_n)] = \sum_{\tilde{T}} \mathbb{E}[X(t_n) | \tilde{T}] \mathbb{P}(T = \tilde{T}) \quad (68)$$

$$= \sum_{\tilde{T}} \mathbb{I}(t_n \in \tilde{T}) \mathbb{P}(T = \tilde{T}) = \mathbb{P}(t_n \in T) = \sum_{m=0}^{\infty} \mathbb{P}(\tau_m = t_n) \quad (69)$$

□

Állítás. *Az autokorrelációs függvény Laplace-transzformáltja kifejezhető az események között eltelt idő súlyfüggvényének Laplace-transzformáltjával.*

Bizonyítás.

$$g(\lambda) \equiv \sum_{n=0}^{\infty} \tilde{A}(t_n) e^{-\lambda t_n} \quad (70)$$

$$\begin{aligned} g(\lambda) &= \sum_{n=0}^{\infty} e^{-\lambda t_n} \mathbb{P}(t_n \in T) = \sum_{n=0}^{\infty} e^{-\lambda t_n} \sum_{m=0}^{\infty} \mathbb{P}(\tau_m = t_n) = \sum_{m=0}^{\infty} \mathbb{E} \left(e^{-\lambda \tau_m} \right) \\ &= \sum_{m=0}^{\infty} \left[\mathbb{E} \left(e^{-\lambda \tau_1} \right) \right]^m = \left(1 - \mathbb{E} \left[e^{-\lambda \tau_1} \right] \right)^{-1}, \end{aligned} \quad (71)$$

ugyanis τ_m előáll, mint m darab független τ_1 eloszlású valószínűségi változó összege. Innentől τ_1 -ből leghagyjuk az indexet, és egyszerűen csak τ -val jelöljük. $\square$

Definíció. Egy u függvény monoton végű, ha $\exists x_0 > 0$, hogy u monoton az (x_0, ∞) intervallumban

Tétel (Tauber,[15]). Ha u monoton végű függvény ω Laplace-transzformálttal, valamint ha L a végtelenben lassan változó függvény és $0 \leq \rho < \infty$, akkor a következő állítások ekvivalensek:

$$u(x) \sim \frac{1}{\Gamma(\rho)} x^{\rho-1} L(x) \quad t \rightarrow \infty \quad (72)$$

$$\int_0^{\infty} e^{-\lambda x} u(x) dx \equiv \omega(\lambda) \sim \lambda^{-\rho} L\left(\frac{1}{\lambda}\right) \quad \lambda \rightarrow 0 \quad (73)$$

Az irodalomban szokásosan a (72) $\Rightarrow$ (73) irányt Abel-tételnek, a (73) $\Rightarrow$ (72) irányt pedig Tauber-tételnek nevezik.

Mi az Abel-tételt az események között eltelt idő eloszlására szeretnénk használni. Az események közötti idő sűrűség- vagy súlyfüggvénye normált, tehát $\beta \geq 1$ a mi esetünkben. Ez a $\rho \leq 0$ tartománynak felel meg, amire már nem vonatkozik a tétel.

Tétel (Abel, kiterjesztett változat). Ha u monoton végű, a $(0, \infty)$ tartományon integrálható függvény ω Laplace-transzformálttal, és $xu(x)$ is monoton végű, valamint ha L a végtelenben lassan változó függvény és $-1 < \rho < 0$, akkor az u függvény végtelenbeli és a Laplace-transzformáltjának 0 körüli aszimptotikája a következő kapcsolatban állnak: (74) $\Rightarrow$ (75)

$$u(x) \sim \frac{1}{\Gamma(\rho)} x^{\rho-1} L(x) \quad t \rightarrow \infty \quad (74)$$

$$\omega(\lambda) - \int_0^{\infty} u(x) dx \sim \lambda^{-\rho} L\left(\frac{1}{\lambda}\right) \quad \lambda \rightarrow 0 \quad (75)$$

Bizonyítás. Visszavezetjük a Tauber-tétel fenti változatára: legyen $-1 < \rho < 0$, ekkor a $v(x) = \frac{1}{\rho} xu(x)$ függvény teljesíti a Tauber-tétel feltételeit, a Laplace-transzformáltját jelöljük η -val.

$$\int_0^{\infty} e^{-\lambda x} u(x) dx = \int_0^{\infty} \frac{e^{-\lambda x}}{x} \rho v(x) dx = \int_0^{\infty} \int_{\lambda}^{\infty} e^{-vx} dv \rho v(x) dx = \quad (76)$$

$$= \int_{\lambda}^{\infty} \int_0^{\infty} e^{-vx} \rho v(x) dx dv = \rho \int_{\lambda}^{\infty} \eta(v) dv = \quad (77)$$

$$= \rho \int_0^{\infty} \eta(v) dv - \rho \int_0^{\lambda} \eta(v) dv = \int_0^{\infty} u(x) dx + \lambda^{-\rho} L\left(\frac{1}{\lambda}\right) \quad (78)$$

λ -ban vezető rendben. $\square$

Ezek után már ki tudjuk értékelni a (71) egyenletben szereplő Laplace-transzformáltakat. A (74) $\Rightarrow$ (75) Abel-tételt alkalmazva (71) jobb oldalára kapjuk, hogy

$$g(\lambda) = \left(1 - \mathbb{E} \left[e^{-\lambda \tau} \right] \right)^{-1} \sim \lambda^{1-\beta} \quad \lambda \rightarrow 0 \quad (79)$$

A (73) $\Rightarrow$ (72) Tauber-tételt alkalmazva $g(\lambda)$ -ra (ami definíciószerűen az autokorrelációs függvény Laplace-transzformáltja) megkapjuk az autokorrelációs függvény lecsengését:

$$\tilde{A}(t) \sim t^{\beta-2} \quad (80)$$

Ezzel bizonyítottuk a tételt. $\square$

A $\beta \geq 2$ eset kvalitatívan különbözik a $\beta < 2$ esettől, ugyanis ekkor az események között eltelt időnek lesz várható értéke, és ezzel párhuzamosan az eseményeknek lesz véges sűrűsége. Az autokorrelációs függvény $t \rightarrow \infty$ esetén ehhez a sűrűséghez fog tartani:

$$\tilde{A}(t) \rightarrow \frac{1}{\mathbb{E}[\tau]} \quad t \rightarrow \infty \quad (81)$$

Ez az eredmény megkapható a fenti kontextusban is, ha az Abel-tételt megpróbáljuk kibővíteni a $-2 < \rho < 1$ tartományra. A fenti gondolatmenetet használva azt kapjuk, hogy

$$\int_0^\infty e^{-\lambda x} u(x) dx = \rho \int_0^\infty \eta(v) dv - \rho \int_0^\lambda \eta(v) dv = \int_0^\infty u(x) dx - \rho \int_0^\lambda \eta(v) dv, \quad (82)$$

ahol $\eta(\lambda) = \int_0^\infty e^{-\lambda x} v(x) dx$, és $v(x) = \frac{1}{\rho} x u(x)$, amire már teljesül a kiterjesztett Abel-tétel. Bevezetve az $\int_0^\infty u(x) dx = I$ és az $\int_0^\infty \rho v(x) dx = \int_0^\infty x u(x) dx = \tilde{I}$ jelölést a fenti eredmény az alábbi formát önti:

$$\int_0^\infty e^{-\lambda x} u(x) dx = I - \tilde{I} \lambda + \lambda^{-\rho} L \left(\frac{1}{\lambda} \right), \quad (83)$$

tehát megjelenik a Laplace-transzformáltban egy λ -val arányos tag, ami $\lambda \rightarrow 0$ esetben erősebb, mint a $\lambda^{-\rho}$ tényező. Ha ezt a formulát λ -ban vezető rendben alkalmazzuk az események között eltelt időre, és visszahelyettesítjük az autokorrelációs Laplace-transzformáltjába, a

$$g(\lambda) = \left(1 - \mathbb{E} \left[e^{-\lambda \tau} \right] \right)^{-1} \sim (\mathbb{E}[\tau] \lambda)^{-1} \quad \lambda \rightarrow 0 \quad (84)$$

kifejezést kapjuk. Alkalmazzuk erre a (73) $\Rightarrow$ (72) Tauber-tételt:

$$\tilde{A}(t) \sim \frac{1}{\mathbb{E}[\tau]} \quad t \rightarrow \infty \quad (85)$$

Ebben az esetben az érdekes kérdés az, hogyan éri el az aszimptotikus értékét az autokorrelációs függvény. A (hullám nélküli) autokorrelációs függvényt például eleve úgy definiáljuk, hogy levonjuk belőle az aszimptotikus értéket (lásd az (1) egyenletet). A $\beta > 2$ esetben az autokorrelációs függvény aszimptotikája további vizsgálódások tárgya lehet. Ez matematikai szempontból

érdekes, a modellezés szempontjából azonban nem releváns, ugyanis mind az eredeti, mind a módosított modellben az események között eltelt idő eloszlásának β exponense az $(1,2)$ intervallumba esik.

Ebben a fejeletben azt láttuk be, hogyan jelenhetnek meg hosszútávú korrelációk egy független eseményekből álló idősorban. Ennek az adatok empirikus kiértékelésénél van jelentősége, ugyanis azt jelenti, hogy a lassan lecsengő korrelációs függvény nem jelent minden esetben az események közötti memóriát. Szintén láttuk, hogy a független azonos eloszlású események közötti idővel jellemzett folyamatok $P(E)$ eloszlása exponenciális. Ez a probléma feloldható [10], ha olyan folyamatot definiálunk, ahol az egyének kétállapotú rendszerként viselkednek: normál állapotban kisebb, aktív állapotban nagyobb rátával indítanak hívást. Az események között eltelt idő pedig független és azonos eloszlású külön-külön az egyik és a másik állapotban. Ebben a modellben az aktív állapot eseményekben mért hosszát kívülről beletehetjük a modellbe, ezzel garantálható, hogy a $P(E)$ eloszlás a megfigyeléseknek megfelelő maradjon. Ebben a modellben az autokorrelációs függvény és az események között eltelt idő eloszlására teljesül az általam levezetett $\alpha + \beta = 2$ skálatörvény [10]. A telekommunikációs hálózatban az 1. táblázat alapján azonban megállapíthatjuk, hogy nem teljesül ez a skálatörvény, ezért ezzel kizárhatjuk a modellezésből a független események közötti idővel rendelkező folyamatokat.

5. Összefoglalás

Tudományos Diákköri dolgozatomban prioritás alapján rendezett dinamikus lista-modellekkel foglalkoztam. A modellek fő motivációja az emberi cselekvések időzítésének modellezése volt, azonban hasznosak lehetnek bármilyen időben erősen inhomogén folyamat vizsgálata során. Az emberi cselekvések időzítéséhez hasonló, ún. villanásos viselkedést figyelhetünk meg például a földrengések, ill. a neuronok tüzelésének idősorában [16, 17]. Ezen folyamatok esetén a három karakterisztikus függvény (az események között eltelt idő eloszlása, az autokorrelációs függvény, valamint a $P(E)$ eloszlás) hatványlecsengést mutat [10]. A dolgozatomban vizsgált modellek egyik fő előnye a Barabási által javasolt modellhez képest, hogy lehet bennük definiálni a telekommunikációs hálózatban mérhető események között eltelt idők eloszlását és az autokorrelációs függvényt. Dolgozatomban részletesen megvizsgáltam, hogy a modellekben mi adódik a három említett függvényre.

Beláttam, hogy a modellek autokorrelációs függvénye nem függ a lista összetételétől. Numerikus számolásaim eredményeként mindkét modellre azt kaptam, hogy az autokorrelációs függvény hatványlecsengésű, amit exponenciális levágás követ. Észrevettem, hogy a különböző hosszúságú listákra számolt autokorrelációs függvények összekálázhatók, és ezt az eredményt felhasználva meghatároztam az autokorrelációs függvény hatvány-szakaszának exponensét a modell bemenő paraméterének függvényében.

Az események között eltelt idő és a $P(E)$ eloszlás függését a lista összetételétől numerikus szimulációkkal vizsgáltam. Kiderült, hogy a fenti két mennyiség komplementer a modell szemszögéből: ha az egyik hatványfüggvény, a másik exponenciálisan cseng le, általános lista-összetétel esetén pedig mindkét függvény hatványfüggvénynél gyorsabb lecsengést mutat. A hatványfüggvényt adó határesetek egyikét (egymásba képezhetők) részletesebben is megvizsgáltam. A modellek események közötti idő eloszlására rekurzív formulát írtam fel, amit numerikusan értékeltem ki. A numerikus számítások eredményeképpen azt kaptam, hogy az események között eltelt idő eloszlása hatványfüggvényt követ. Az autokorrelációs függvény, valamint az események közötti idő eloszlásának numerikusan meghatározott exponensei eleget tesznek egy skálatörvénynek, miszerint a két exponens összege 2. Ezt az eredményt analitikusan is beláttam, és a modellek szélesebb körére kiterjesztettem: minden olyan pontfolyamatra, amit független, hatványfüggvény eloszlású események közötti idők definiálnak.

Munkám közvetlen folytatását jelentené az exponensek analitikus meghatározása, valamint a modellben szereplő lista keveredésének vizsgálata a kártyakeverés elméletében bevezetett fogalmak segítségével. További célom egy olyan modell megtalálása, ami reprodukálni tudja a telekommunikációs hálózatban végzett méréseket. Az ezután következő lépés a kommunikáló egyének kölcsönhatásának és az információ terjedésének vizsgálata lehetne.

Végezetül szeretném kifejezni köszönetemet témavezetőmnek, Dr. Kertész Jánosnak a TDK dolgozatom elkészítésében nyújtott segítségéért, és Dr. Tóth Bálintnak a konzultációkért. Szintén köszönettel tartozom Karsai Mártonnak, aki rendelkezésemre bocsájtotta a telekommunikációs hálózaton végzett méréseit, és szívesen nyújtott segítséget bármilyen jellegű kérdésben. Végül köszönöm Hang-Hyun Jo-nak a gondolatébresztő beszélgetéseket.

6. Függelék

6.1. Mester-egyenlet származtatása a Markov-lánc átmeneti mátrixából

A mester-egyenlet (7) átmeneti mátrixa származtatható közvetlenül a Markov-lánc átmeneti mátrixából is. Ezt a következő példán szemléltetem.

Példa. $N = 4, L = 2$ modell

Számozzuk az állapotokat a következőképp:

- 1 AABB
- 2 ABAB
- 3 ABBA
- 4 BAAB
- 5 BABA
- 6 BBAA

Így az átmeneti mátrix,

$$W = \begin{pmatrix} w_1 + w_2 & 0 & 0 & w_3 + w_4 & 0 & 0 \\ w_3 & w_1 & 0 & w_2 & w_4 & 0 \\ w_4 & 0 & w_1 & 0 & w_2 + w_3 & 0 \\ 0 & w_2 + w_3 & 0 & w_1 & 0 & w_4 \\ 0 & w_4 & w_2 & 0 & w_1 & w_3 \\ 0 & 0 & w_3 + w_4 & 0 & 0 & w_1 + w_2 \end{pmatrix} \quad (86)$$

A mester-egyenletben annak a valószínűsége jelenik meg, hogy az i -edik helyen A van ($P_A(i, t) = \mathbb{P}\{i\text{-edik helyen } A \text{ van, } t \text{ időben}\}$). Ezt a következőképp fejezhetjük ki a Markov-lánc állapotainak valószínűségeivel ($\mathbb{P}(i, t)$):

$$P_A(1, t) = \mathbb{P}(1, t) + \mathbb{P}(2, t) + \mathbb{P}(3, t) \quad (87)$$

$$P_A(2, t) = \mathbb{P}(1, t) + \mathbb{P}(4, t) + \mathbb{P}(5, t) \quad (88)$$

$$P_A(3, t) = \mathbb{P}(2, t) + \mathbb{P}(4, t) + \mathbb{P}(6, t) \quad (89)$$

$$P_A(4, t) = \mathbb{P}(3, t) + \mathbb{P}(5, t) + \mathbb{P}(6, t), \quad (90)$$

Ezt mátrixos formában a következőképp fejezhetjük ki:

$$\underline{P}_A(t) = \underline{\mathbb{P}}(t) T, \quad (91)$$

ahol $\underline{P}_A(t) = (P_A(1, t), P_A(2, t), \dots, P_A(N, t))$ és $\underline{\mathbb{P}}(t) = (\mathbb{P}(1, t), \mathbb{P}(2, t), \dots, \mathbb{P}(\binom{N}{L}, t))$ sorvektorok, és

$$T = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \quad (92)$$

$\underline{\mathbb{P}}(t)$ időfejlődését ismerjük:

$$\underline{\mathbb{P}}(t_{n+1}) = \underline{\mathbb{P}}(t_n) W, \quad (93)$$

ebből szeretnénk kifejezni $\underline{P}_A(t)$ időfejlődését:

$$\underline{P}_A(t_{n+1}) = \underline{\mathbb{P}}(t_{n+1}) T = \underline{\mathbb{P}}(t_n) W T = \underline{\mathbb{P}}(t_0) W^{n+1} T \quad (94)$$

A mester-egyenlethez keresünk egy $\tilde{A}$ mátrixot, melyre fennáll a következő összefüggés:

$$W T = T \tilde{A}, \quad (95)$$

ahol tehát W egy $\binom{N}{L} \times \binom{N}{L}$ -es T egy $\binom{N}{L} \times N$ -es, $\tilde{A}$ pedig egy $N \times N$ -es mátrix. Ugyanis, ha a (95) egyenlet teljesül, akkor

$$\underline{P}_A(t_n) = \underline{\mathbb{P}}(t_{n-1}) W T = \underline{\mathbb{P}}(t_{n-1}) T \tilde{A} = \underline{P}_A(t_{n-1}) \tilde{A} \quad (96)$$

$$= \underline{\mathbb{P}}(t_0) W^n T = \underline{\mathbb{P}}(t_0) W^{n-1} T \tilde{A} = \dots = \underline{\mathbb{P}}(t_0) T \tilde{A}^n = \underline{P}_A(t_0) \tilde{A}^n. \quad (97)$$

Ha találunk olyan $N \times \binom{N}{L}$ -es $\tilde{T}$ mátrixot, melyre

$$\tilde{T} T = I_{N \times N} \quad (98)$$

teljesül, akkor kifejezhető $\tilde{A}$, és létezik mester-egyenlet. (98) egy olyan alulhatározott lineáris egyenletrendszerre írható át, ami minden szóba jöhető T esetén megoldható. Ugyanis T soraiban az $N - L$ db. 1 és L db. 0 különböző permutációi kerülnek elő, amik közül bármelyik N lineárisan független. Ebből következik, hogy a T mátrix felső $N \times N$ -es tartománya invertálható, így $\tilde{T}$ -re egy jó megoldást ad, ha a $\tilde{T}$ bal szélső $N \times N$ -es részére az előbbi inverzet írjuk, az ettől jobbra eső elemeket pedig 0-nak választjuk. A jelen példában:

$$\tilde{T} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & -\frac{1}{2} & 0 & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 1 & \frac{1}{2} & 0 & 0 \end{pmatrix} \quad (99)$$

és

$$\tilde{A} = \tilde{T} W T = \begin{pmatrix} w1 & w2 + w3 + w4 & 0 & 0 \\ w2 & w1 & w3 + w4 & 0 \\ w3 & 0 & w1 + w2 & w4 \\ w4 & 0 & 0 & w1 + w2 + w3 \end{pmatrix} \quad (100)$$

Ez épp megegyezik a mester-egyenletben felírt (7) átmeneti mátrixszal.

Állítás. Az $\tilde{A} = \tilde{T} W T$ léptető mátrix egyértelmű és megegyezik a mester-egyenletben felírt (7) léptető mátrixszal (az állapotok azonos permutációit tekintve) tetszőleges N és L értékek esetén. (nem bizonyítjuk)

6.2. Az autokorrelációs függvény L -függetlensége

Az egyensúlyi autokorrelációs függvényhez tartozó kezdeti feltételt a (9), az autokorrelációs függvény definícióját pedig a (10) egyenlet tartalmazza. Használjuk ki a modell linearitását, és írjuk fel mester-egyenlet állapotvektorát két tag összegeként!

$$\underline{P}_A^{(L)}(t_0) = \frac{L}{N-1}(1,0,0,\dots,0) + \frac{N-L-1}{N-1}(1,1,1,\dots,1)$$

$$\underline{P}_A^{(L)}(t_n) = \frac{L}{N-1}(1,0,0,\dots,0)A^n + \frac{N-L-1}{N-1}(1,1,1,\dots,1)A^n \quad (101)$$

$$= \frac{L}{N-1}(1,0,0,\dots,0)A^n + \frac{N-L-1}{N-1}(1,1,1,\dots,1) \quad (102)$$

A lista első elemének időfejlődése tehát kifejezhető az $L = N - 1$ -es modell (tehát ahol csak 1 db. A van) segítségével:

$$P_A^{(L)}(1, t_n) = \frac{L}{N-1}P_A^{(N-1)}(1, t_n) + \frac{N-L-1}{N-1} \quad (103)$$

$$= \frac{L}{N-1}[A^n]_{11} + \frac{N-L-1}{N-1} \quad (104)$$

Innen,

$$A^{(L)}(t_n) = \frac{P_A^{(L)}(1, t) - \frac{N-L}{N}}{1 - \frac{N-L}{N}} = \frac{\frac{L}{N-1}P_A^{(N-1)}(1, t_n) + \frac{N-L-1}{N-1} - \frac{N-L}{N}}{1 - \frac{N-L}{N}} \quad (105)$$

$$= \frac{\frac{L}{N-1}\left(P_A^{(N-1)}(1, t_n) - \frac{1}{N}\right)}{\frac{L}{N-1}\left(1 - \frac{1}{N}\right)} = \frac{P_A^{(N-1)}(1, t_n) - \frac{1}{N}}{1 - \frac{1}{N}} = A(t_n) \quad (106)$$

Az autokorrelációs függvény pedig kifejezhető a léptető mátrix hatványainak bal felső elemével:

$$A(t_n) = \frac{[A^n]_{11} - \frac{1}{N}}{1 - \frac{1}{N}} \quad (107)$$

6.3. A módosított modell, mint véletlen bolyongás

Ha a listából kiszemelünk egy cselekvést, és csak annak az útját követjük, akkor ez egy véletlen bolyongást végez $[1, N] \cap \mathbb{Z}$ -n. Ha a cselekvés aktuális pozíciója k , akkor w_k valószínűséggel balra, w_{k+1} valószínűséggel jobbra ugrik, vagy helyben marad. Mivel $w_k > w_{k+1}$, a vizsgált elemnek lesz egy balra irányuló driftje, ami alól csak az első elem képez kivételt, amit ha választunk, akkor helyben marad.

Ebben a modellben kiszámolható, hogy milyen távolra jut a véletlen bolyongó a kezdőponttól indulva a visszatérés előtt. Legyen q_k^M annak a valószínűsége, hogy a kezdetben k -adik helyen álló

elem a mozgása során előbb éri el a lista elejét, mint az M -edik pozíciót. Definíció szerint $q_1^M = 1$ és $q_M^M = 0$.

$$q_k^M = \frac{w_{k+1}}{w_k + w_{k+1}} q_{k+1}^M + \frac{w_k}{w_k + w_{k+1}} q_{k-1}^M, \quad (108)$$

mivel a balra lépés valószínűsége feltételezve, hogy lép valamerre $\frac{w_k}{w_k + w_{k+1}}$, és hasonlóan a jobbra ugrás esetén $\frac{w_{k+1}}{w_k + w_{k+1}}$. Átrendezéssel kapjuk, hogy

$$\begin{aligned} q_{k+1}^M - q_k^M &= \frac{w_k}{w_{k+1}} (q_k^M - q_{k-1}^M) = \prod_{i=2}^k \frac{w_i}{w_{i+1}} (q_2^M - q_1^M) \\ -1 &= q_M^M - q_1^M = \sum_{k=1}^{M-1} (q_{k+1}^M - q_k^M) = \sum_{k=1}^{M-1} \prod_{i=2}^k \frac{w_i}{w_{i+1}} (q_2^M - q_1^M) \\ &= (q_2^M - q_1^M) \sum_{k=1}^{M-1} \left(\frac{k+1}{2} \right)^\sigma = (q_2^M - q_1^M) \sum_{k=2}^M \left(\frac{k}{2} \right)^\sigma. \end{aligned} \quad (109)$$

Átrendezéssel

$$q_2^M = 1 - \frac{2^\sigma}{\sum_{k=2}^M k^\sigma} \quad (110)$$

$$q_k^M = 1 - \frac{\sum_{i=2}^k i^\sigma}{\sum_{j=2}^M j^\sigma}. \quad (111)$$

Legyen ξ_k a maximális távolság, amíg eljut a kezdetben k -adik helyről induló véletlen bolyongó, mielőtt elérné az első pozíciót. Ekkor q_k^M , mint M -nek a függvénye a ξ_k valószínűségi változó (kumulatív) eloszlásfüggvényét adja $1 < k < M$ -re:

$$q_k^M \equiv \mathbb{P}(\xi_k < M). \quad (112)$$

Ugyanis q_k^M definíció szerint annak a valószínűsége, hogy a bolyongó előbb elnyelődik az első pozícióban, minthogy elérné M -et, amit úgy is fogalmazhatunk, hogy annak a valószínűsége, hogy az elnyelődésig megtett távolság kisebb, mint M . M egész értékeire tehát ($1 < k < M$):

$$\mathbb{P}(\xi_k < M) = 1 - \frac{\sum_{i=2}^k i^\sigma}{\sum_{j=2}^M j^\sigma}$$

Felhasználva, hogy $\sum_{i=2}^M i^\sigma \approx \int_2^M i^\sigma di = \frac{M^{\sigma+1} - 2^{\sigma+1}}{\sigma+1} \approx \frac{M^{\sigma+1}}{\sigma+1}$ amikor M nagy,

$$\mathbb{P}(\xi_k < M) = 1 - \sum_{i=2}^k i^\sigma \frac{\sigma+1}{M^{\sigma+1}}$$

Ebből deriválással megkaphatjuk a ξ_k súlyfüggvényét közelítő sűrűségfüggvényt. Nagy M -re,

$$P_{\xi_k}(M) \approx (\sigma+1)^2 \frac{\sum_{i=2}^k i^\sigma}{M^{\sigma+2}} \quad (113)$$

Hogy megkapjuk a maximális távolságot a lista elejéről indulva, feltételezzünk az első lépésre.

$$\mathbb{P}(\xi_1 = x) = \delta_{x,1}(1 - w_2) + (1 - \delta_{x,1})w_2\mathbb{P}(\xi_2 = x), \quad (114)$$

tehát ugyanazt az aszimptotikát kapjuk, mint ξ_2 -re.

Hivatkozások

- [1] W. Aiello, F. Chung and L. Lu (2000) Proceedings of the 32nd ACM Symposium on the Theory of Computing (ACM, New York), pp. 171-180.
- [2] J.-P. Eckmann, E. Moses, D. Sergi, Proc. Natl Acad. Sci. USA **101**, 14333 (2004).
- [3] H. Ebel, L.-I. Mielsch, and S. Bornholdt, Phys. Rev. E **66**, R35103 (2002).
- [4] A. Vázquez, J. G. Oliveira, Z. Dezső, K.-I. Goh, I. Kondor and A.-L. Barabási, Phys. Rev. E, **73**, 036127 (2006).
- [5] M. Karsai, M. Kivelä, R. K. Pan, K. Kaski, J. Kertész, A. L. Barabási and J. Saramäki, Phys. Rev. E **83** 025102 (2011)
- [6] J. G. Oliveira and A.-L. Barabási, Nature (London) **437**, 1251 (2005).
- [7] A.-L. Barabási, Nature (London) **435**, 207 (2005)
- [8] A. Vazquez, Phys. Rev. Lett. **95**, 248701 (2005).
- [9] T. Zhou, L.-L. Jiang, R.-Q. Su and Y. C. Zhang, Euro. Phys. Lett. **81**, 58004 (2008).
- [10] M. Karsai, K. Kaski, A.-L. Barabási, J. Kertész (*publikálás alatt*)
- [11] R.D. Malmgren, D.B. Stouffer, A.E. Motter, L.A.N. Amaral, Proc. Nat. Acad. Sci. **105**, 18153 (2008).
- [12] H-H. Jo, M. Karsai, J. Kertész, K. Kaski arXiv:1101.0377 (2011).
- [13] D. Aldous, P. Diaconis, Amer. Math. Monthly **93** 333-348 (1986).
- [14] D. Bayer, P. Diaconis, Ann. Appl. Prob. **2**, 294-313 (1992).
- [15] W. Feller, *An Introduction to Probability Theory and its Applications* 2nd ed. Vol. 2., Wiley India Pvt. Ltd. (2008).
- [16] Á. Corral, Phys. Rev. Lett. **92**, 108501 (2004).
- [17] T. Kemuriyama, et al., BioSystems **101**, 144-147 (2010).