

Forgalmi adatsorok illesztése Markov érkezési folyamattal

Mészáros András

October 28, 2011

Összefoglaló

A telekommunikációs hálózatok megfelelő tervezésének és működtetésének elengedhetetlen feltétele, hogy forgalmukat jól kezelhető modellekkel tudjuk jellemezni. Ezek egyik széleskörűen használt csoportja a Markov modellek, ezen belül pedig a Markov érkezési folyamatok (Markov Arrival Process, MAP) osztálya, melyek számítási szempontból jól kezelhetők, kellően rugalmasak, és alkalmasak különböző forgalmi sajátosságok visszaadására. A Markov érkezési folyamatok illesztése nem triviális feladat. Speciális struktúrák és heurisztikus módszerek alkalmazása csökkenti a rugalmasságot, más, például a várható érték maximalizáláson alapú módszerek általánosabban használhatóak, de számításigényesek. Kidolgoztak egyéb eljárásokat, többek között olyan két lépcsős algoritmusokat[8, 3] is, melyek elsőként egy fázis típusú eloszlást (Phase Type Distribution, PH) illesztnek statikus paraméterekre (például az érkezési időközök eloszlására vagy momentumaira), majd ezt a második lépésben dinamikus jellemzők (autokorreláció vagy együttes momentumok) figyelembevételével MAP-pá egészítik ki. Ez a felbontás nagyban csökkenti a feladat számítási igényét, azonban problémája, hogy a dinamikus jellemzők lehetséges intervallumát az első lépés gyakran drasztikusan behatárolja. Munkám során egy olyan optimalizációs lépést dolgoztam ki, mely a két lépcsős eljárások fent említett hibáját célozza. A módszer lényege, hogy az első lépés eredményeképp létrejövő PH ekvivalens átalakításával a második lépés számára kedvező kiindulópontot biztosít. Az eljárás fontos előnye, hogy az illesztésben használt konkrét algoritmusoktól függetlenül mindig alkalmazható, valamint hogy a benne szereplő célfüggvény az adott feladatnak megfelelő változtatásával rugalmassá tehető. Dolgozatomban leírom a PH-k és MAP-ok legfontosabb tulajdonságait, és áttekintem a korábban vázolt két lépcsős MAP illesztési eljárások alapvető jellemzőit. Bemutatok egy korábban kidolgozott, specifikus esetben használt PH transzformációs eljárást[3], valamint ismertetem

az általam megvalósított eljárást, illetve alkalmazási példákon keresztül szemléltetem a különböző változatok jellemvonásait.

Kulcsszavak: forgalmi modell, illesztés, Markov érkezési folyamat.

Tartalomjegyzék

1	Fázis típusú eloszlások és Markov érkezési folyamatok	4
1.1	Markov-folyamatok	4
1.2	Fázis-típusú eloszlások	5
1.3	Markov érkezési folyamatok	6
2	Két lépcsős MAP illesztési módszerek	7
2.1	PH illesztése statikus paraméterekre	8
2.2	PH kiegészítése MAP struktúrává	8
3	A PH reprezentációjának optimalizálása	9
3.1	A Buchholz-Kriege-féle reprezentáció optimalizálási eljárás	9
3.2	A javasolt reprezentáció optimalizálási eljárás	10
4	Numerikus vizsgálatok	13
4.1	PH reprezentáció transzformációja	13
4.1.1	Első példa	13
4.1.2	Második példa	18
4.2	MAP illesztése MAP segítségével	22
4.3	Valós adatsor illesztése	26
5	Értékelés	27

1 Fázis típusú eloszlások és Markov érkezési folyamatok

1.1 Markov-folyamatok

Azokat a sztochasztikus folyamatokat, melyekre igaz, hogy jövőbeli állapotaik valószínűsége a jelenbeli állapotuk ismeretében egyértelműen meghatározható a múltbeli állapotoktól függetlenül, Markov-folyamatoknak nevezzük. $X(t)$ sztochasztikus folyamat tehát markovi, ha:

$$Pr(X(t_n) = x_n | X(t_{n-1}) = x_{n-1}, \dots, X(t_0) = x_0) = Pr(X(t_n) = x_n | X(t_{n-1}) = x_{n-1})$$

bármely $n \geq 1, 0 < t_0 < t_1 < \dots < t_n$ esetén. Ezt a tulajdonságot memóriamentességnek vagy Markov-tulajdonságnak nevezzük. Megkülönböztetünk diszkrét és folytonos idejű Markov-folyamatokat. A továbbiakban csak folytonos idejű, véges állapotterű Markov-folyamatokat, vagyis folytonos idejű véges Markov-lánccokat fogok vizsgálni. Markov-lánccok alatt a következőkben mindig azoknak ezt a speciális csoportját fogom érteni, kivéve ha ezt külön jelzem.

Egy n állapotú Markov-lánccot jellemezhetünk egy olyan $n \times n$ méretű D generátor mátrix segítségével, melyre

$$D_{i,j} = \lim_{\Delta \rightarrow 0} \frac{Pr(X(t + \Delta) = j | X(t) = i)}{\Delta}, \forall 1 \leq i, j \leq n, i \neq j$$

vagyis $D_{i,j}$ az i -ből j állapotba történő átmenet átmeneti intenzitása minden (i, j) , $i \neq j$ állapotpár esetén, a főátlóbeli elemek pedig az azon kívüliek összegének mínusz egyszeresével egyenlők, tehát a mátrix sorösszege nulla. (A folyamat átlépéskor definíció szerint nem léphet vissza az átlépés előtti állapotba.)

Ebből levezethető, hogy az i állapotból való átlépési idő hossza egyenlő $\min_{1 \leq j \leq n} (Y_{i,j})$, $i \neq j$ -vel, ahol $Y_{i,j}$ $D_{i,j}$ paraméterű exponenciális eloszlású valószínűségi változó, az átlépés után pedig abban a j állapotban vagyunk, amelyre $Y_{i,j}$ minimálisnak adódott. Mivel két exponenciális eloszlású, λ_1 és λ_2 paraméterű valószínűségi változó minimuma egy $\lambda_1 + \lambda_2$ eloszlású valószínűségi változót határoz meg, ezért a D mátrix főátlóbeli elemei az adott állapotokból való átlépési intenzitás mínusz egyszeresével egyenlők.

Egy Markov-lánc stabil, ha minden olyan i állapota elérhető bármely állapotból $N < \infty$ állapotátmenettel, melyre $\lim_{t \rightarrow \infty} p_i(t) > 0$, vagyis az állapotban tartózkodás valószínűsége, ha $t \rightarrow \infty$, nagyobb, mint 0 valamely $p(0)$ -ra, valamint aperiodikus, vagyis bármely (i, j) állapotpárra létezik olyan $N_{i,j}$ szám, mely esetén minden $N \geq N_{i,j}$ -re található olyan állapot átmeneti sorozat, mely i -ben kezdődik, j -ben végződik és pontosan N darab állapotátmenetből áll. Ha

egy Markov-lánc stabil, akkor létezik pontosan egy olyan, a folyamat kezdeti állapotától független $\underline{\alpha}$ vektor melyre $\lim_{t \rightarrow \infty} \underline{p}(t) = \underline{\alpha}$ [?], ahol $\underline{p}(t)$ a Markov-lánc állapotvalószínűségeinek vektora a t pillanatban, melyet röviden a Markov-lánc t időpontbeli állapotának nevezünk. Ezt a $\underline{\alpha}$ valószínűségi vektort a Markov-lánc egyensúlyi eloszlásának nevezzük, és kiszámíthatjuk az $\underline{\alpha} \mathbf{D} = \underline{0}$, $\underline{\alpha} \mathbf{1} = 1$ lineáris egyenletrendszer megoldásával.

Minden folytonos idejű Markov-lánc meghatároz egy diszkrét idejű Markov-láncot is, melyet az eredeti folyamat beágyazott folyamatának nevezünk. Diszkrét idejű Markov-láncok esetén az állapotátmenet előtt az állapotban töltött időt egységnyiinek vesszük, és a \mathbf{P} diszkrét generátor mátrix csak azt mutatja meg, hogy milyen valószínűséggel milyen következő állapotba megyünk, vagyis $P_{i,j} = Pr(X(k+1) = j | X(k) = i)$. A diszkrét generátor mátrix meghatározható a $\mathbf{P} = \Lambda^{-1} \mathbf{D} - \mathbf{I}$ egyenlőség segítségével, ahol Λ diagonális mátrix, melynek főátlóbeli elemei megegyeznek a megfelelő állapotokból való kilépési intenzitásokkal, \mathbf{I} pedig a megfelelő méretű identitás mátrix. Diszkrét Markov-láncokra a folytonos Markov-láncokhoz hasonló stabilitási kritériumokkal definiálható az egyensúlyi eloszlás: minden stabil diszkrét Markov-láncre létezik pontosan egy olyan $\underline{\pi}$, melyre $\underline{\pi} \mathbf{P} = \underline{\pi}$. Ezt a diszkrét Markov-lánc egyensúlyi eloszlásának nevezzük, és kiszámíthatjuk a $\underline{\pi} \mathbf{P} = \underline{\pi}$, $\mathbf{P} \mathbf{1} = \mathbf{1}$ lineáris egyenletrendszer megoldásával.

Ha tehát csak a Markov-lánc egyensúlyi viselkedése érdekel minket, akkor a kezdeti állapotot leíró π_0 valószínűségi vektor nem lényeges, a kezdeti tranzien viselkedés szempontjából azonban igen. Az előzőekből levezethető, hogy

$$p(t) = \pi_0 e^{\mathbf{D}t}$$

1.2 Fázis-típusú eloszlások

Vegyünk egy nyelővel, vagyis olyan állapottal rendelkező Markov-láncot, melyből a folyamat nem lép tovább. Ekkor a folyamat kezdeti állapotából a nyelőbe való eljutásának ideje úgynevezett fázis-típusú eloszlást (phase-type distribution, PH) mutat. A pontos eloszlás természetesen függ a Markov-lánc $\underline{\pi}_0$ kezdeti valószínűségi vektorától is. Ha nyelőnek a Markov-lánc n -edik állapotát választjuk, és megköjtjük, hogy a folyamat nem indulhat a nyelőből, akkor a kezdeti valószínűségi vektor $\underline{\pi}_0 = (\underline{\pi}, 0)$, a Markov-lánc generátora pedig:

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_0 & d_1 \\ \underline{0} & 0 \end{bmatrix}$$

$\mathbf{D}_0 \mathbf{1} = -d_1 \mathbf{1}$ és $\mathbf{1} \mathbf{1} = 1$, $\mathbf{1}$ pedig a megfelelő méretű, egyesekből álló vektort jelöli. Ebből látható, hogy a PH leírásához elegendő a \mathbf{D}_0 mátrix és a π vektor. $p(t)$ képletéből levezethető, hogy a PH eloszlásfüggvénye

$$F(t) = 1 - \pi e^{\mathbf{D}_0 t} \mathbf{1}$$

sűrűségfüggvénye

$$f(t) = \pi e^{\mathbf{D}_0 t} \underline{d}_1 = -\pi e^{\mathbf{D}_0} \mathbf{D}_0 \mathbf{1}$$

k -adik momentumának várható értéke

$$E(X^k) = k! \pi (-\mathbf{D}_0)^{-k} \mathbf{1}$$

A PH azon állapotait, melyek kezdeti valószínűsége nem nulla, bemeneti, azon állapotait, melyekből a nyelőbe jutás intenzitás pozitív, kimeneti állapotnak nevezzük.

Azokat a PH-kat, melyek \mathbf{D}_0 mátrixa alsó (vagy felső) háromszög mátrix, aciklikus fázis típusú eloszlásoknak (acyclic phase type distribution, APH) nevezzük. Minden APH-nak végtelen egyenértékű reprezentációja létezik, kanonikus reprezentációjának azt nevezzük, melynél

$$\mathbf{D}_0 = \begin{bmatrix} -\lambda_1 & \lambda_1 & & & & \\ & -\lambda_2 & \lambda_2 & & & \\ & & \ddots & \ddots & & \\ & & & -\lambda_{n-1} & \lambda_{n-1} & \\ & & & & -\lambda_n & \end{bmatrix}$$

ahol $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$. Az APH-k kanonikus formára alakítása részletesen le van írva a [4] forrásban. A továbbiakban a π és \mathbf{D}_0 által meghatározott PH jelölésére a $PH(\pi, \mathbf{D}_0)$ jelölést fogjuk használni.

Bármely fázis-típusú eloszlás segítségével fázis-típusú felújítási folyamatot definiálhatunk. Ez értelmezhető két egymásba ágyazott folyamatként is. A belső folyamat a PH eloszlású időt generáló Markov-lánc, a külső folyamat pedig az eddigi nyelőbe jutások száma. Amikor a belső folyamat a nyelőbe érkezik, újraindul, a külső folyamat értéke pedig eggyel növekszik. Ezt az eseményt érkezésnek nevezzük.

1.3 Markov érkezési folyamatok

A Markov érkezési folyamatok (Markov arrival process, MAP) a PH felújítási folyamatok általánosításaként értelmezhetők. Utóbbiak problémája modellezési szempontból, hogy az egymás utáni időközök eloszlása független, így semmilyen dinamikus jellemző nem ragadható meg velük. A MAP-okat két mátrixszal írhatjuk le: Az első a \mathbf{D}_0 mátrix, megfelel a PH-t jellemző mátrixnak, a második a \mathbf{D}_1 mátrix, mely az érkezéssel járó állapotátmenetek intenzitásait tartalmazza. A két mátrix összege az állapotátmeneteket jellemző \mathbf{D} generátor mátrix, ennek

megfelelően $D_0 \mathbf{1} + D_1 \mathbf{1} = \mathbf{0}$. Bármely PH felújítási folyamatot leírhatunk MAP-ként is: A MAP D_0 mátrixa az eredeti marad, és $D_1 = d_1 \pi$.

MAP-ok esetén a beágyazott diszkrét Markov-lánc P mátrixát megkaphatjuk a $P = (-D_0)^{-1} D_1$ egyenlőség segítségével. A MAP érkezési intenzitása

$$\lambda = \frac{1}{E(X)} = \frac{1}{\pi - D_0^{-1} \mathbf{1}} = \alpha D_1 \mathbf{1}$$

együttes momentumainak várható értéke

$$E[X_{i_1}^{k_1} X_{i_2}^{k_2}] = k_1! k_2! \pi (-D_0)^{-k_1} P^{i_2 - i_1} (-D_0)^{-k_2} \mathbf{1}$$

az érkezési időközök lag-k korrelációja $(\frac{E[X_i X_{i+k}] - E^2[X]}{E[X^2] - E^2[X]})$ pedig

$$\rho_k = \frac{\lambda^2 \pi (-D_0)^{-1} P^k (-D_0)^{-1} \mathbf{1} - 1}{2\lambda^2 \pi (-D_0)^{-1} (-D_0)^{-1} \mathbf{1} - 1}.$$

Minden MAP érkezési időköz eloszlása jellemezhető a π , D_0 által megadott PH-val, ahol π a beágyazott folyamat egyensúlyi állapoteloszlása, D_0 pedig a MAP azonosan jelölt mátrixa. Hasonlóképp minden PH felújítási folyamat kiegészíthető egy azonos eloszlású MAP-pá. Ehhez a D_0 mátrix változatlanul hagyása mellett a PH kezdeti állapoteloszlásának meg kell egyeznie a MAP beágyazott folyamatának egyensúlyi eloszlásával. Fontos megjegyezni, hogy a PH felújítási folyamat alternatív leírása MAP-ként nem egyenértékű a PH felújítási kiterjesztésével, előbbi esetén ugyanis független marad az egymás utáni érkezési időközök időtartama, míg utóbbi esetében nem feltétlenül.

2 Két lépcsős MAP illesztési módszerek

Mint korábban említettem, a két (vagy több) lépcsős MAP illesztési eljárások egyik fontos előnye, hogy segítségükkel az illesztés bonyolult problémája számítási szempontból jóval kezelhetőbb lépésekre bontható. Az előző fejezetben már kitértem rá, hogy egy MAP eloszlása egy azonos rendű PH segítségével definiálható, illetve hogy bármely PH felújítási folyamat kiegészíthető azonos érkezési időköz eloszlást mutató MAP-pá. Ezek megfelelő motivációt szolgáltatnak olyan illesztési módszerekhez, melyek két lépcsőből állnak, első lépésük eredménye egy PH, mely a statikus paraméterek (érkezési időközök momentumai, vagy eloszlása) megragadására szolgál. Ezt második lépésben MAP struktúrává egészítik ki oly módon, hogy az eredmény minél jobban illeszkedjen a választott dinamikus jellemzőkre (autokorreláció, együttes momentumok).

2.1 PH illesztése statikus paraméterekre

Fázis típusú eloszlások illesztésére számos módszer létezik[1, 2, 6, 7, 5, 9], melyek különböző bemeneti adatok, például az érkezési időközök sűrűség- vagy eloszlásfüggvénye esetleg minták vagy momentumok alapján végzik az illesztést. Az általam használt PhFit[7] nevű program eloszlásfüggvény, komplementer eloszlásfüggvény és minták feldolgozására is képes, míg a későbbiekben még említett Buchholz-Kriege-féle illesztési eljárás bemenetül az érkezési időközök momentumai szolgálnak.

2.2 PH kiegészítése MAP struktúrává

A PH MAP-pá való kiegészítéskor a $D_1 \mathbf{1} = -D_0$ és a $\underline{\pi}(-D_0)^{-1} D_1$ egyenlőségeknek minden esetben teljesülnie kell. Ezen egyenlőségek a következő lineáris egyenletrendszerre alakíthatók:

$$\underbrace{\begin{bmatrix} \begin{bmatrix} \mathbf{I}_{m \times m} \\ \underline{\gamma} \end{bmatrix} & \begin{bmatrix} \mathbf{I}_{m \times m} \\ \underline{\gamma} \end{bmatrix} & \cdots & \begin{bmatrix} \mathbf{I}_{m \times m} \\ \underline{\gamma} \end{bmatrix} \end{bmatrix}}_{\mathcal{A}_{2m \times m^2}} \cdot \underline{x} = \underbrace{\begin{bmatrix} d_1 \\ \vdots \\ \pi \end{bmatrix}}_{b_{2m}}, \quad (1)$$

ahol $\mathbf{I}_{m \times m}$ az m méretű identitás mátrix, $\underline{\gamma} = \underline{\pi}(-D_0)^{-1}$, $d_1 = -D_0 \mathbf{1}$, \underline{x} pedig a D_1 mátrix oszlopvektoraiból képzett oszlopvektor, vagyis

$$\underline{x} = \begin{bmatrix} \{D_1\}_1 \\ \{D_1\}_2 \\ \vdots \\ \{D_1\}_m \end{bmatrix}$$

ahol $\{D_1\}_i$ a D_1 mátrix i -edik oszlopvektora. További korlát, hogy a D_1 elemeinek nemnegatívnak kell lennie, vagyis $\underline{x} \geq \underline{0}$.

Ezek mellett a korlátok mellett határozhatók meg további, a dinamikus viselkedésre vonatkozó egyenletek és célfüggvények. Mivel a dinamikus paraméterek legtöbbször a D_1 mátrix elemeinek nemlineáris függvényei, ezért a pontos illesztés általában túl számításigényes feladat, valamint gyakran nem is lehetséges (elképzeltető például, hogy az illesztendő paraméterek értékei kívül

esnek a MAP osztály által elérhető tartományon), így elsődlegesen célfüggvények indokolt. Számítási komplexitás szempontjából kivétel az $E[X_i X_{i+1}]$ együttes momentum, és az ezzel nagyon szorosan összefüggő lag-1 korreláció, melyek a D_1 mátrixszal lineáris kapcsolatban vannak.

3 A PH reprezentációjának optimalizálása

A két lépcsős illesztési módszereknek van egy lényeges hátránya: Az első lépés eredményeképp kapott PH reprezentáció meghatározza, hogy a dinamikus jellemzők mekkora értéktartományát tudjuk elérni a második lépésben. Ha a második lépés bemenete egy tetszőleges PH kanonikus reprezentációja például, az eleve meghatározza, hogy az autokorreláció csak nulla értéket vehet fel, az együttes momentumokra pedig az $E[X_{i1}^{k1} X_{i2}^{k2}] = E[X^{k1+k2}]$ egyenlőség teljesül a D_1 mátrixtól függetlenül, vagyis a dinamikus viselkedés semmilyen módon nem modellezhető, a MAP egy PH felújítási folyamat alternatív leírása lesz csupán. (Ez valójában minden olyan esetben igaz, melyben a PH-nak csak egy lehetséges bemeneti vagy kimeneti állapota van.) Az olyan PH illesztési módszerek tehát, melyek a PH kanonikus reprezentációját adják kimenetül - ilyen eljárást alkalmaz a PhFit program is - önmagukban nem képezhetik a két lépcsős módszerek első lépését, egy megfelelő rugalmasságot biztosító reprezentáció találása azonban kritikus kérdés minden más esetben is.

3.1 A Buchholz-Kriege-féle reprezentáció optimalizálási eljárás

Peter Buchholz és Jan Kriege egyik cikkükben [3] leírnak egy olyan reprezentáció optimalizációs módszert, melynek bemenete egy APH kanonikus formája, kimenete pedig ennek egy olyan reprezentáció, melynek minden állapota kimeneti állapot. Az alábbiakban röviden bemutatom az eljárást.

Hajtsuk végre az alábbi lépéseket az összes (i, j) , $i < j$ állapotpárra lexikografikus sorrendben:

- Definiáljuk $\delta \leq \delta_{max}$ értéket, ahol

$$\delta_{max} = \min \left(\pi(j), \frac{\pi(i) D_0(i, j)}{\lambda_j - \lambda_i}, \min_{k < i, D_0(k, i) > 0} \left(\pi(i) \frac{D_0(k, j)}{D_0(k, i)} \right) \right) \quad (2)$$

Ha $\lambda_i = \lambda_j$, a zárójelben szereplő második kifejezést végtelennek tekintjük. Ha $\delta_{max} > 0$, δ -t szabadon megválaszthatjuk a δ_{max} és a 0 érték között.

- A δ értékének kiválasztása után hajtsuk végre a következő transzformációs lépéseket:

$$\pi'(k) = \begin{cases} \pi(i) + \delta & \text{for } k = i \\ \pi(j) - \delta & \text{for } k = j \\ \pi(k) & \text{egyébként} \end{cases} \quad (3)$$

$$D_0(k, l) = \begin{cases} D_0(i, j) \frac{\pi(i)}{\pi(i)+\delta} - \frac{(\lambda_j - \lambda_i)\delta}{\pi(i)+\delta} & \text{for } k = i \text{ and } l = j \\ D_0(i, l) \frac{\pi(i)}{\pi(i)+\delta} + D_0(j, l) \frac{\delta}{\pi(i)+\delta} & \text{for } k = i \text{ and } l \neq j \\ D_0(k, i) \frac{\pi(i)+\delta}{\pi(i)} & \text{for } k < j \text{ and } l = i \\ D_0(k, j) + D_0(k, i) \frac{\delta}{\pi(i)} & \text{for } k < j \text{ and } l = j \\ D_0(k, l) & \text{egyébként} \end{cases} \quad (4)$$

3.2 A javasolt reprezentáció optimalizálási eljárás

A fent bemutatott eljárás számítási szempontból kedvező, de van néhány problémája:

- Nincs világos kapcsolat a dinamikus paraméterek elérhető értéktartománya és a módszer között
- Az eljárás eredményeképp létrejövő reprezentáció függ a δ értéktől, melyre a $\delta = 0.9\delta_{max}$ értéket javasolják a szerzők, de ez az ajánlás pusztán kísérleti eredményeken alapul.
- Az eljárás bemenete csak egy APH kanonikus formája lehet.
- A kimenet APH, mely szintén korlátozó tényező lehet.
- A π_{out} kimeneti valószínűségvektor (melynek i -edik eleme azt mutatja meg, hogy mekkora a valószínűsége, hogy a folyamat az i állapotból megy a nyelőbe) alig változik, ez pedig nyilvánvalóan behatárolja a dinamikus paraméterek elérhető értékeit.

Az általam kidolgozott eljárás figyelembe veszi ezeket a tényezőket, és olyan alternatívát kínál, mely nagyobb rugalmasságot biztosít. A kifejlesztett módszer a következő tételre épül: Vegyünk egy $PH(\underline{\pi}, \mathbf{D}_0)$ eloszlást. Ugyanennek az eloszlásnak egy másik reprezentációját kapjuk, ha alkalmazzuk a következő transzformációt:

$$\mathbf{D}'_0 = \mathbf{T}^{-1} \mathbf{D}_0 \mathbf{T}$$

$$\underline{\pi}' = \underline{\pi}T$$

ahol T olyan nonszinguláris, nemnegatív mátrix, melyre $T\mathbf{1} = \mathbf{1}$. A tétel belátható a következő lépések segítségével:

$$\begin{aligned} f(t) &= -\underline{\pi}e^{D_0 t} D_0 \mathbf{1} = -\underline{\pi} \sum_{i=0}^{\infty} \frac{t^i}{i!} D_0^i D_0 \mathbf{1} = -\underline{\pi} \sum_{i=0}^{\infty} \frac{t^i}{i!} D_0^i D_0 T \mathbf{1} = \\ &= -\underline{\pi} \sum_{i=0}^{\infty} \frac{t^i}{i!} D_0^i T D_0' \mathbf{1} = -\underline{\pi} T \sum_{i=0}^{\infty} \frac{t^i}{i!} D_0'^i D_0' \mathbf{1} = -\underline{\pi}' e^{D_0' t} D_0' \mathbf{1} \end{aligned}$$

Ezt kihasználva végrehajthatjuk a következő lépéseket:

- Kiszámítjuk az eredeti PH reprezentáció egy választott mértékét (később megadunk néhány lehetőséget), melynek alapján optimalizálni szeretnénk.
- Minden (i, j) , $i \neq j$ -re elvégezzük a fent leírt transzformációt a $T^{(i, j)} = I + K^{i, j}$ transzformációs mátrixot alkalmazva, ahol $K_{i, i}^{(i, j)} = -a$ vagy $K_{i, j}^{(i, j)} = a$, $K_{k \neq i, l}^{(i, j)} = 0$ és $a = 0.5$
- Megvizsgáljuk, hogy az új reprezentáció teljesíti-e a markovi PH reprezentáció feltételeit, vagyis a $\underline{\pi}'$ vektor valódi valószínűségi vektor (minden eleme 1-nél kisebb nemnegatív szám), a D_0' mátrixnak pedig minden főátlóbeli eleme negatív, azon kívüli elemei nemnegatívak, minden sorösszege nempozitív.
- Ha az új reprezentáció markovi, kiszámítjuk rá az első pontbeli mértéket.
- Ha találtunk a mérték szempontjából jobb reprezentációt, a legjobb új reprezentációra megismételjük az előző lépéseket. Ha nem találtunk ilyet, az a értékét a régi felére változtatjuk, és a korábbi reprezentációval hajtjuk végre a lépéseket.
- Az algoritmus akkor áll le, ha az a elér egy előre meghatározott, megfelelően kis értéket.

Ahogy azt említettem, többféle különböző mértéket alkalmazhatunk. Munkám során két típust teszteltem. Az elsőbe a PH kezdeti és kimeneti valószínűségvektorának entrópiájának (a továbbiakban bemeneti és kimeneti entrópia), vagyis $H_{in} = \sum \underline{\pi}(i) \log(\underline{\pi}(i))$ és $H_{out} = \sum \underline{\pi}_{out}(i) \log(\underline{\pi}_{out}(i))$, ahol $\underline{\pi}_{out} = (\underline{\pi} D_0^{-1}) \cdot (-D_0 \mathbf{1})$ függvényei tartoznak, a másodikba a két lépcsős eljárás második lépésében elvileg elérhető lag-1 autokorrelációval kapcsolatos mennyiségek.

Az entrópia típusú mértékek közül a bemeneti és kimeneti entrópiát, valamint ezek összegét és szorzatát vizsgáltam, az autokorrelációs mértékek közül pedig a lag-1 korreláció minimális és maximális elérhető értékét, valamint ezek különbségét. A lag-1 korreláció képlete:

$$\rho_1 = \frac{\lambda^2 \underline{\pi} (-D_0)^{-2} D_1 (-D_0)^{-1} \mathbf{1} - 1}{2\lambda^2 \underline{\pi} (-D_0)^{-2} \mathbf{1} - 1},$$

Ennek segítségével, a PH reprezentáció kiegészítésével elérhető korreláció kiszámítása egy lineáris programozási feladat, melynek korlátait a korábban leírt lineáris egyenletrendszer adja meg, a maximális elérhető korreláció maximalizálandó célfüggvénye pedig $\underline{c}^T \underline{x}$, ahol \underline{x} a korábbi lineáris egyenletrendszerben is szereplő, változókat tartalmazó oszlopvektor, és

$$\underline{c} = \begin{bmatrix} \{C\}_1 \\ \{C\}_2 \\ \vdots \\ \{C\}_m \end{bmatrix}$$

ahol

$$C = \frac{1}{2\lambda \underline{\alpha} (-D_0)^{-1} \mathbf{1} - 1} (\lambda^2 (-D_0)^{-1} \mathbf{1} \underline{\pi} (-D_0)^{-2}) - 1)$$

Ezen kívül bármilyen egyéb mérték típus alkalmazható, azonban figyelembe kell venni, hogy az algoritmus során ezt sokszor ki kell számolnunk. (A mátrixok rangjának növekedésével négyzetesen nő a második lépés mértékszámításainak mennyisége, valamint emelkedik a leállás előtt végrehajtott iterációk átlagos száma is.)

Ennek az eljárásnak van egy nagyon fontos problémája: A megfelelő PH reprezentáció megtalálására mohó algoritmust használ, ennek következtében pedig az eljárás könnyen megállhat egy lokális optimumban. A vizsgálataim alapján a PH-k kanonikus reprezentációjára jellemző, hogy mindkét mérték esetén található olyan, hozzá nagyon közel eső lokális optimum, melyben az algoritmus megakad, ennek következtében az eredmény rossz reprezentáció lesz.

Figyelembe véve a két bemutatott optimalizációs megközelítést, logikus lépésként kínálja magát a két módszer kombinálása: Először a Buchholz-Kriege-féle eljárás segítségével kilépünk a kanonikus reprezentáció környezetéből, majd az előbb leírt algoritmust alkalmazzuk, ezáltal kiküszöböljük a korábban felsorolt hiányosságokat.

4 Numerikus vizsgálatok

A következőkben néhány példán keresztül szemléltetem a leírt eljárás jellemzőit. Elsőként egy PH reprezentáció transzformációját mutatom be, egyúttal ismertetve a Buchholz-Kriege módszer néhány sajátosságát. Ezután egy három állapotú MAP-ot illeszték egy másik három állapotú MAP segítségével, megmutatva, hogy az elérhető korreláció hogyan befolyásolja az illesztés jóságát, ezek után pedig valós forgalmi adatsorok illesztésének eredményeit fogom ismertetni.

4.1 PH reprezentáció transzformációja

4.1.1 Első példa

Az első példában egy kanonikus formában adott, ötödrendű APH transzformációját mutatom be, melynek kezdeti reprezentációja:

$$\underline{\pi} = [0.2, 0.3, 0.24, 0.16, 0.1], \quad D_0 = \begin{bmatrix} -8 & 8 & 0 & 0 & 0 \\ 0 & -9 & 9 & 0 & 0 \\ 0 & 0 & -11 & 11 & 0 \\ 0 & 0 & 0 & -13 & 13 \\ 0 & 0 & 0 & 0 & -17 \end{bmatrix}$$

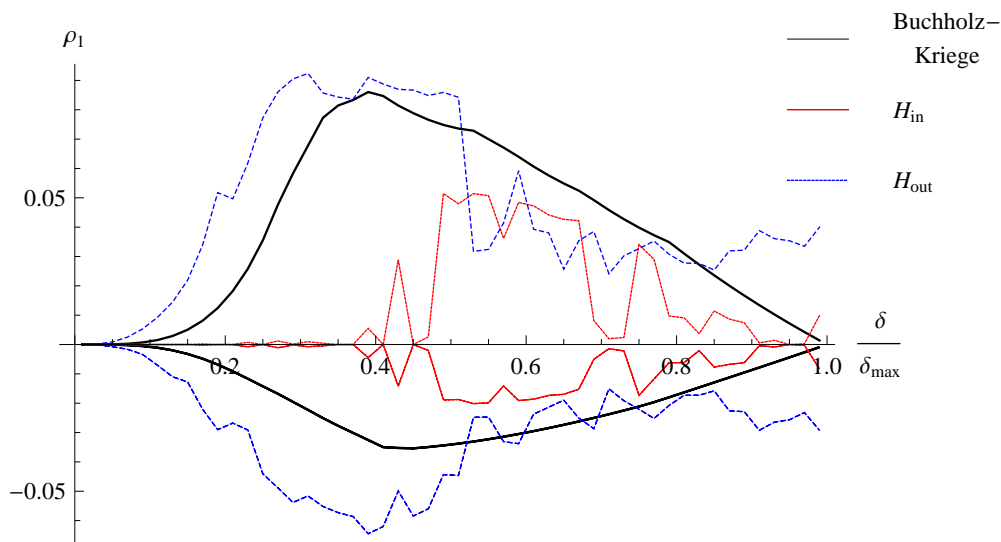
A Buchholz-Kriege-módszert először az ajánlásnak megfelelően 0.9-es δ/δ_{max} aránnyal alkalmaztam. Ezek után a korábban leírt mértékeknek megfelelő optimalizációt végeztem. Az eredményeket az alábbi táblázat mutatja:

Módszer/Korreláció	Minimum	Maximum	Tartomány
Buchholz-Kriege-módszer	-0.00868	0.015	0.0237
Bemeneti entrópia	-0.0063	0.0073	0.0136
Kimeneti entrópia	-0.0229	0.0322	0.0551
Entrópiák összege	-0.0984	0.112	0.210
Entrópiák szorzata	-0.0975	0.112	0.210
Minimális korreláció	-0.100	0.107	0.207
Maximális korreláció	-0.0986	0.114	0.213
Korrelációtartomány	-0.118	0.121	0.240

Jól látható, hogy a Buchholz-Kriege-módszer által produkált értékekhez képest még jelentős tartalék rejlett az elérhető korrelációban, melyet mindkét mértéksaláddal ki is lehetett használni.

A kapott értékek alapján úgy tűnik, hogy csak a bemeneti vagy csak a kimeneti entrópia és a korreláció között nincs szoros összefüggés. Ez nem meglepő, hiszen - ahogy korábban már leírtam - ha csak egy kezdeti, vagy kimeneti állapot van, nem lehet nullától eltérő korrelációt elérni. Ezzel összhangban, ha a bemeneti vagy kimeneti entrópia értéke alacsony marad, nem várhatunk számottevő javulást.

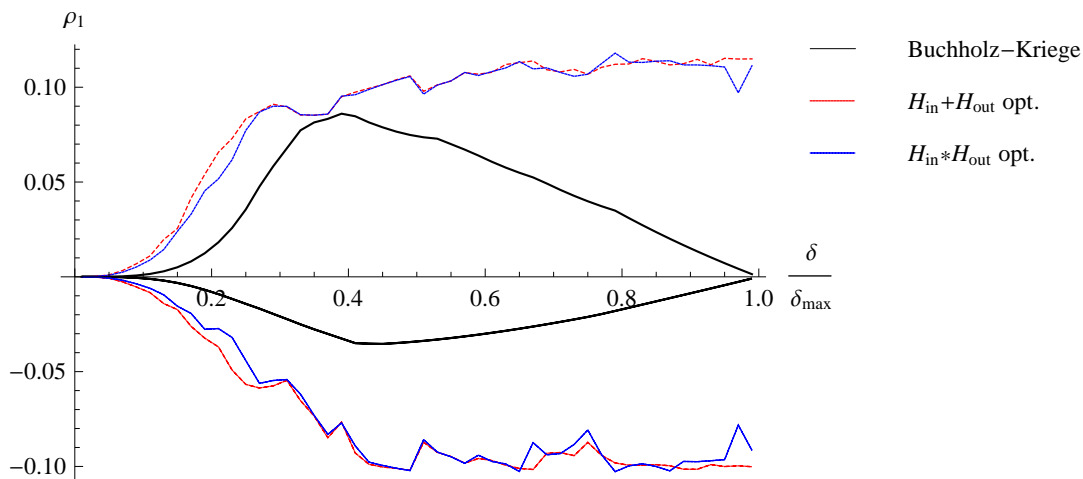
Korábban már említettem, hogy a $\delta_{rate} = \delta/\delta_{max}$ értékétől nagyban függhet a Buchholz-Kriege-módszer eredménye. A függés tanulmányozására a Buchholz-Kriege-módszert, és a saját optimalizációs eljárást is megvizsgáltam a teljes δ/δ_{max} spektrumra.



Ábra 1:

A legszembeötlőbb jelenség az első ábrán az, hogy a Buchholz-Kriege-módszerrel elérhető korreláció nem a $\delta_{rate} = 0.9$ érték esetén maximális, hanem a 0.4 körüli tartományban. Az is jól látszik, hogy a minimálisan és a maximálisan elérhető korreláció nagyon hasonló változást mutat (ellentétes előjellel) a δ_{rate} függvényében minden eljárásra. Harmadik észrevételünk az lehet, hogy a kimeneti entrópia alapú eljárás nagyjából ugyanott éri el a legjobb eredményt, mint a Buchholz-Kriege-módszer. Ezen kívül közepes δ_{rate} esetén valamivel rosszabbul, egyébként jobban teljesít nála. Az is észrevehető, hogy a bemeneti entrópia alapú módszer ugyanott javul, ahol a kimeneti entrópia alapú romlik.

A második ábrán több dolgot is megfigyelhetünk. Elsőként azt, hogy az entrópia összegét és szorzatát optimalizáló eljárásokkal elérhető korrelációs

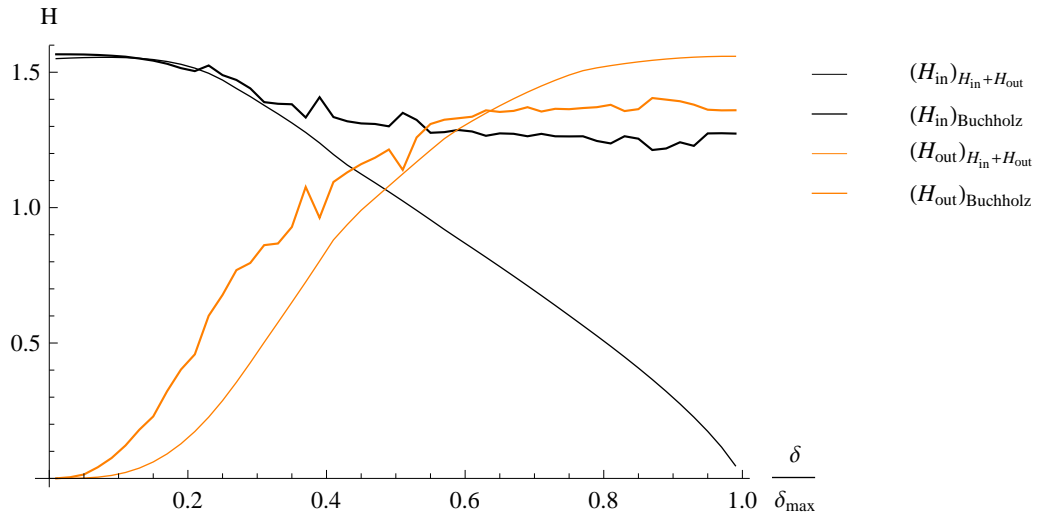


Ábra 2:

értékek jellemzően sokkal jobbák, főleg a negatív korreláció terén, mint amire a Buchholz-Kriege-módszer képes. Ezen kívül az is szembetűnő, hogy a két entrópia alapú eljárás majdnem teljesen megegyező eredményt ad. Ha megvizsgáljuk a minimális és maximális korrelációhoz tartozó mátrixokat, azt tapasztalhatjuk, hogy ezek is majdnem megegyeznek a két esetben. Ezek alapján a két mérték az optimalizálás szempontjából hasonlóan mondható, legalábbis nagyjából ilyen méretű PH-ra. Megjegyzendő még az is, hogy az előző ábrához hasonlóan itt is észrevehető a minimális és a maximális elérhető korreláció mértékenként hasonló jellegű változása.

Látva, hogy az entrópia alapján ilyen jó optimalizáció lehetséges, vagyis a korreláció és az entrópia ilyen jól látható kapcsolatban van, felmerül a kérdés, hogy a Buchholz-Kriege-módszer vajon milyen entrópia tulajdonságokkal rendelkezik.

A harmadik ábra az entrópia összeg alapú optimalizálás és a Buchholz-Kriege-módszer során megfigyelhető bemeneti és kimeneti entrópiákat mutatja. Látható, hogy a két eljárás között az a különbség, hogy előbbi igyekszik a bemeneti és kimeneti entrópiát is magas szinten tartani, míg az utóbbi a bemeneti entrópiát folyamatosan csökkentve növeli a kimeneti entrópiát. Ez könnyen megérthető, ha egy pillantást vetünk a Buchholz-Kriege-féle algoritmusra. Látható ugyanis, hogy a kezdeti valószínűségi vektor korábbi elemeit a későbbiek rovására növeli, ezáltal csökkentve a bemeneti entrópiát. Ha a későbbi állapotokhoz tartozó kezdeti valószínűsége magasabbak, a Buchholz-Kriege-módszer növelheti is a bemeneti entrópiát, azonban egyéb esetekben monoton módon csökkenni fog. A kimeneti entrópia növekedése sem meglepő, ugyanis az egyes állapotokból (az utolsó kivételével) a nyelőbe való átlépés intenzitása a δ_{rate} -tel arányosan nő.



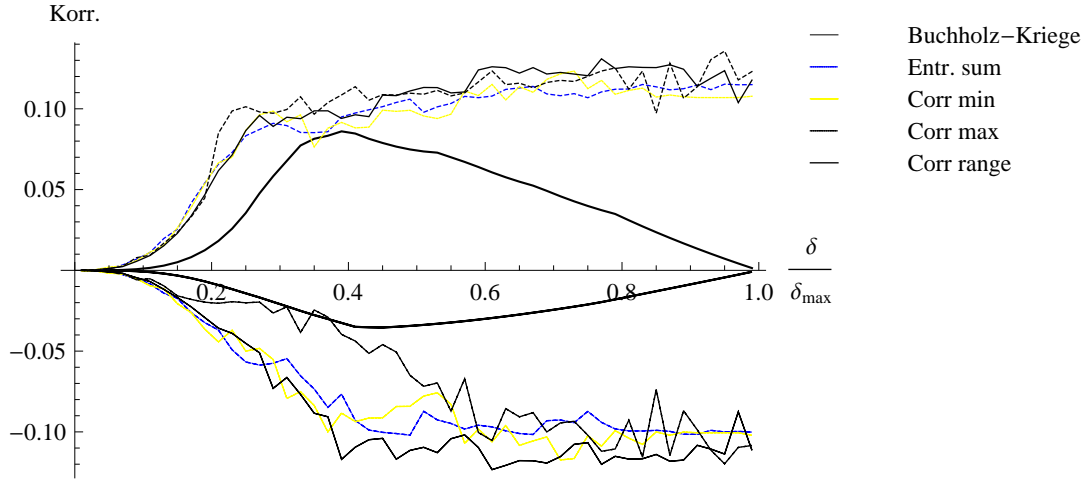
Ábra 3:

A negyedik ábrán láthatjuk a korreláció alapú mértékek függvényeit, kiegészítve a Buchholz-Kriege-módszer, és az entrópia összeg alapú maximalizálás eredményével. A különböző mértékek nagyjából hasonló eredményt adnak, a Buchholz-Kriege-módszernél mindegyik lényegesen jobban teljesít. Itt két dolgot érdemes megjegyezni: Az első az, hogy a maximális korreláció optimalizálásával végzett transzformáció a minimális korreláció szempontjából a $\delta_{rate} = 0.4$ környékén jelentősen rosszabb, mint a többi változat, azonban ez gyakorlati szempontból nem lényeges, hiszen ha negatív korrelációjú MAP-ot szeretnénk generálni, akkor nyilvánvalóan nem ezt a mértéket alkalmazzuk. A második pedig az, hogy az entrópia összeg alapú (és az entrópia szorzat alapú) optimalizálás hasonlóan jó eredményeket produkált, mint a többi, számítási szempontból kedvezőtlenebb mérték.

Mint az ötödik ábrán látható, a korrelációs mértéket alkalmazó optimalizációk entrópiájának és korrelációjának kapcsolatáról általában elmondható, hogy, ha az entrópiák nagyon alacsonyak, a korreláció is nagyon alacsony lesz, majd a kettő együtt nő, azonban nagyobb értékek esetén a korreláció változásait nem tükrözik az entrópia ingadozásai.

Érdemes még egy pillantást vetni a Buchholz-Kriege-módszerrel létrehozott reprezentációkra $\delta_{rate} = 0.4$ és $\delta_{rate} = 0.9$ esetén:

$$\underline{\pi}^{(0.4)} = [0.517, 0.278, 0.12, 0.053, 0.032],$$



Ábra 4:

$$D_0^{(0.4)} = \begin{bmatrix} -8 & 4.41 & 1.04 & 0.804 & 0.521 \\ 0 & -9 & 6.31 & 0.619 & 0.643 \\ 0 & 0 & -11 & 9.76 & 0.3 \\ 0 & 0 & 0 & -13 & 12.65 \\ 0 & 0 & 0 & 0 & -17 \end{bmatrix}$$

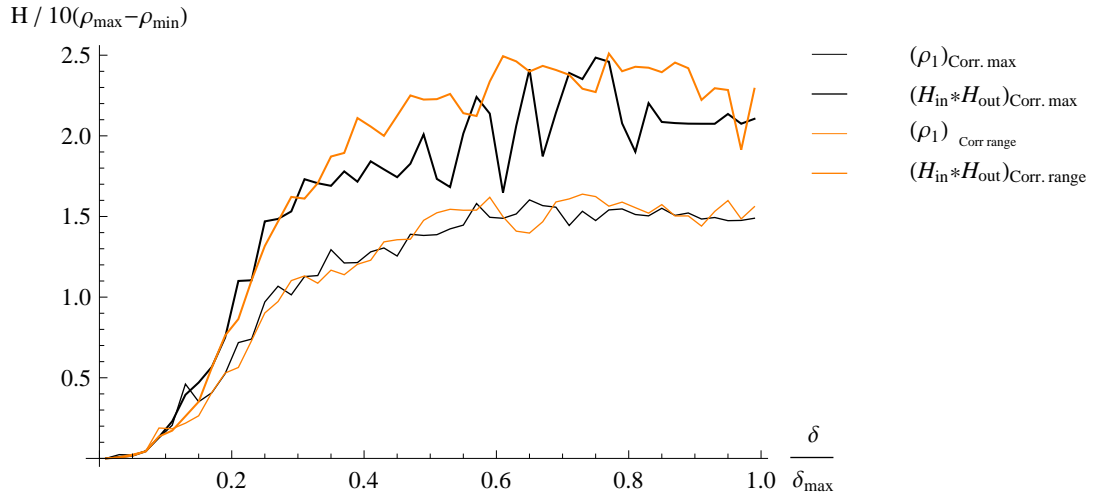
illetve

$$\underline{\pi}^{(0.9)} = [0.92, 0.075, 0.00474, 0.000242, 0.0000184],$$

$$D_0^{(0.9)} = \begin{bmatrix} -8 & 3.61 & 1.77 & 0.858 & 0.0949 \\ 0 & -9 & 5.97 & 0.884 & 0.104 \\ 0 & 0 & -11 & 7.49 & 0.281 \\ 0 & 0 & 0 & -13 & 7.26 \\ 0 & 0 & 0 & 0 & -17 \end{bmatrix}$$

Jól megfigyelhető az, amit az entrópiák alapján, illetve a Buchholz-Kriege-módszer általános hatásának elemzésével is le lehetett szűrni: A bemeneti entrópia a kezdeti valószínűség első állapotba tömörülésével drámaian lecsökken, míg a kimeneti entrópia nő, köszönhetően legfőképp annak, hogy a nagyobb δ_{rate} miatt az $(i, i + 1)$ átmeneti intenzitások nagyobb mértékben csökkennek, és ennek a csökkenésnek a nagyrésze a nyelőbe menő intenzitást növeli, valamint a bemeneti entrópiát csökkentő hatás egyúttal szintén növeli a kimeneti entrópiát.

Hasonlítsuk még ezt össze a $\delta_{rate} = 0.9$ -hez tartozó, a korrelációs tartományt optimalizáló eljárással:



Ábra 5:

$$\underline{\pi}^{(0.9)} = [0.459, 0.414, 0.11, 8.48e - 9, 0.0179],$$

$$D_0^{(0.9)} = \begin{bmatrix} -8 & 7.65 & 5.48e - 6 & 1.54e - 7 & 1.79e - 6 \\ 0 & -9 & 6.65 & 7.87e - 7 & 2.61e - 7 \\ 0 & 0 & -12.9 & 9.51 & 7.22e - 6 \\ 0 & 0 & 6.23e - 6 & -16.5 & 10.7 \\ 0 & 0 & 0.00967 & 0.25 & -11.6 \end{bmatrix}$$

Látható, hogy a D_0 mátrix a hozzá tartozó Buchholz-Kriege-féle mátrixhoz képest, melyből a korreláció optimalizáló kiindult, teljesen kiürült, a főátló és a mellette lévő elemek kivételével. Ezzel egy időben a kezdeti valószínűségek vektora valamelyest kiegyenlítettődött, vagyis a bemeneti korreláció megnőtt. A másik nagy változás az állapotokból való kilépések intenzitásának (főátlóbeli elemek) drasztikus változása, melyeket a Buchholz-Kriege-módszer minden esetben érintetlenül hagy. Ezek az intenzitások mintegy felcserélődtek, a harmadik állapot intenzitása nagyjából megfelel a korábbi negyedikének, a negyediké az ötödikének, az ötödiké a harmadikénak.

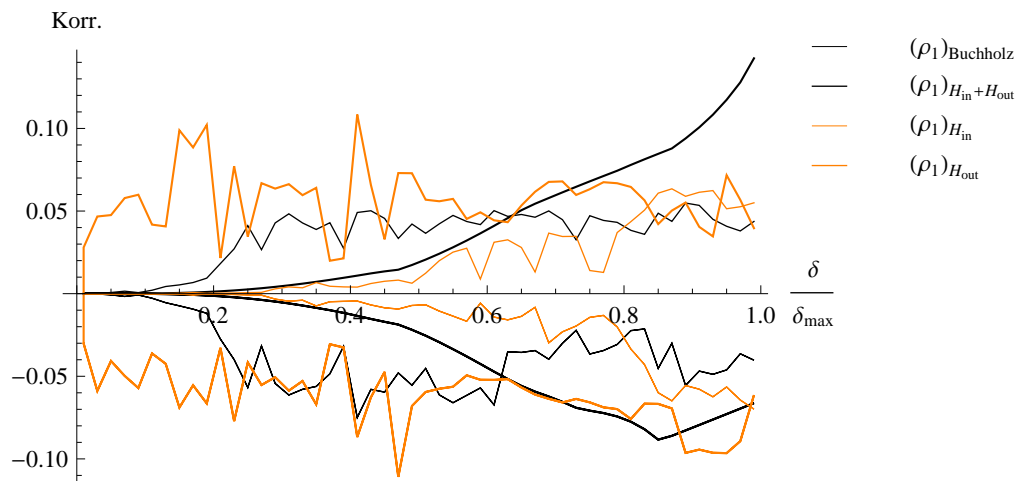
4.1.2 Második példa

A második példában egy másik, az előzővel azonos méretű, kanonikus reprezentációjú, ám természetében különböző PH transzformációját elemzem.

$$\underline{\pi} = [0.4, 0.2, 0.21, 0.09, 0.1],$$

$$D_0 = \begin{bmatrix} -8 & 8 & 0 & 0 & 0 \\ 0 & -29 & 29 & 0 & 0 \\ 0 & 0 & -61 & 61 & 0 \\ 0 & 0 & 0 & -153 & 153 \\ 0 & 0 & 0 & 0 & -170 \end{bmatrix}$$

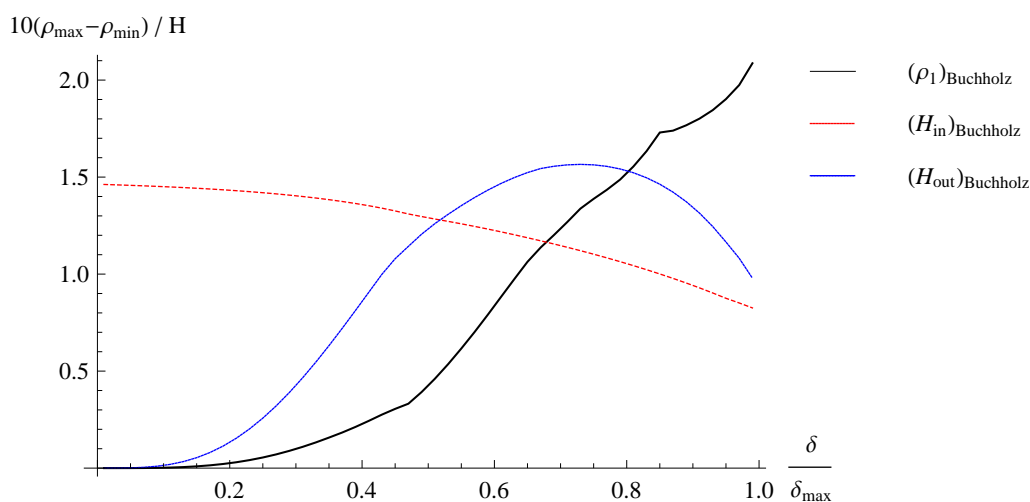
Az előző PH-hoz képest a legnagyobb különbség az, hogy az átmeneti intenzitások aránya nagyobb. Ennek egyik következménye az, hogy a δ_{max} értéke általában csökkenni fog, aminek hatására a bementi korreláció értéke kedvezőbb marad, mint az előző példában, hiszen kisebb lesz a valószínűség korai állapotok felé mozdulása. A kimeneti korreláció változását nehezebb előre megbecsülni, mivel az előbbi hatás csökkenti azt, azonban a Buchholz-Kriege-eljárásban szereplő $D_0(i, j) \frac{\pi(i)}{\pi(i)+\delta} - \frac{(\lambda_j - \lambda_i)\delta}{\pi(i)+\delta}$ kifejezés második tagjának értéke az elsőhöz képest megnő, így nagyobb csökkenés tapasztalható a D_0 mátrix $(i, i + 1)$ elemében, ami a kimeneti entrópiára nézve kedvező.



Ábra 6:

A hatodik ábrán látható, hogy a kimeneti korrelációra történő optimalizálás kis δ_{rate} értékekre ezúttal is jobb, mint a Buchholz-Kriege-módszer esetén, azonban nagyobb δ_{rate} -ek esetén a helyzet megváltozik. A Buchholz-Kriege-módszer előnye ekkor nyilvánvaló, bár az elérhető korreláció minimuma -0.0883 ($\delta_{rate} = 0.85$), míg a kimeneti entrópia optimalizálásánál ez -0.111 ($\delta_{rate} = 0.47$). Kimondható az is, hogy $\delta_{rate} = 0.9$ -es ajánlás csakugyan közel van az optimumhoz, azonban megjegyzendő, hogy a korrelációs tartomány nagysága itt 0.178 a maximális korreláció pedig 0.0972 , míg $\delta_{rate} = 0.999$ -re ezek az értékek 0.218 és 0.153 .

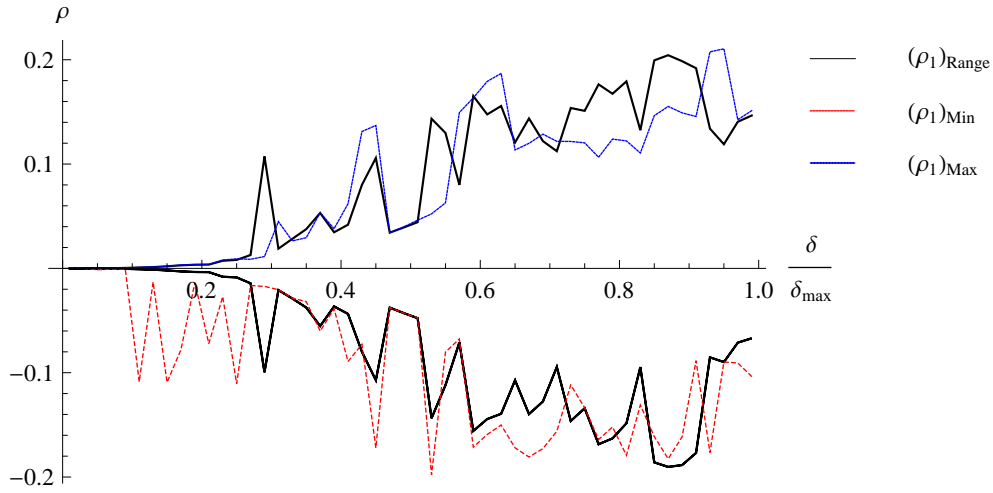
Az igazán nagy változást az entrópia összeg és szorzat alapú optimalizáció mutatja, melynél még a kimeneti entrópia alapú is jobban teljesít. Ez, és a korábbiak alapján azt mondhatjuk, hogy a korreláció összege illetve szorzata csak kisebb értékekre hozható megfelelően kapcsolatba a korrelációval, ekkor vagy a bemeneti vagy kimeneti korreláció nagyon kicsi. Nagyobb értékek esetén nincs szoros összefüggés.



Ábra 7:

Ezt támasztja alá a Buchholz-Kriege-módszer esetén jelentkező kimeneti és bemeneti entrópia, valamint a korrelációs tartomány tízszeresét ábrázoló hetedik ábra is, melyen látható, hogy, bár 0.8 fölött a δ_{rate} növekedésével mind a bemeneti, mind a kimenet entrópia csökken, a korrelációs tartomány mégis nő. A várakozásoknak megfelelően egyébként a bemeneti entrópia értéke lassabban csökken, mint az előző példa esetén, a kimeneti entrópia pedig a korábban leírt ellentétes irányba mozdító hatásoknak köszönhetően nagyjából $\delta_{rate} = 0.8$ -ig nő, majd elkezd csökkenni.

A nyolcadik ábrán a Buchholz-Kriege-módszer és a korrelációs tartományra való optimalizálás által elérhető korrelációs tartomány mellett a minimális korrelációra történő optimalizálással elérhető minimális, a maximális korrelációra történő optimalizálással elérhető korreláció maximális értéke látható. A korreláció alapú optimalizálás, akárcsak az előző példában, most is minden esetben jellemzően jobban teljesít. Megfigyelhető azonban, hogy a δ_{rate} változásával mindhárom függvény sokkal erősebben ingadozik, például a minimális korreláció alapú eljárással 0.95-ös δ_{rate} érték esetén -0.189 -es korreláció érhető el, míg 0.97-es δ_{rate} esetén csak -0.0659 -es, egyébként mind nagyjából hasonló tendenciát



Ábra 8:

mutat, mint a Buchholz-Kriege-féle függvény. Ez azért lényeges, mert így a δ_{rate} értéke kulcsfontosságúvá válik. Az is észrevehető, hogy a korrelációs tartomány szerint optimalizáló eljárás egyes δ_{rate} -ekre szignifikánsan jobb eredményt ad, mint a maximális korreláció szerint optimalizáló. Ez természetesen annak köszönhető, hogy mindkettő mohó optimalizációs eljárás.

Érdeemes most is megnézni Buchholz-Kriege-módszer által létrehozott PH reprezentációt két különböző δ_{rate} -re. Például $\delta_{rate} = 0.4$ és $\delta_{rate} = 0.9$ esetén:

$$\underline{\pi}^{(0.4)} = [0.477, 0.178, 0.202, 0.0643, 0.0774],$$

$$\mathbf{D}_0^{(0.4)} = \begin{bmatrix} -8 & 5.16 & 1.63 & 0.732 & 0.178 \\ 0 & -29 & 20.8 & 4.25 & 1.4 \\ 0 & 0 & -61 & 43.66 & 7.19 \\ 0 & 0 & 0 & -153 & 134.35 \\ 0 & 0 & 0 & 0 & -170 \end{bmatrix}$$

$$\underline{\pi}^{(0.9)} = [0.652, 0.159, 0.176, 0.0052, 0.00763],$$

$$\mathbf{D}_0^{(0.9)} = \begin{bmatrix} -8 & 1.24 & 0.405 & 0.459 & 0.228 \\ 0 & -29 & 2.21 & 1.86 & 1.22 \\ 0 & 0 & -61 & 14.9 & 0.326 \\ 0 & 0 & 0 & -153 & 120 \\ 0 & 0 & 0 & 0 & -170 \end{bmatrix}$$

Látható most is, hogy a kezdeti valószínűségek eltolódnak a korábbi állapotok felé, de a korábbinál kevésbé markánsan. Megfigyelhető az is, hogy az egyes

állapotokból a nyelőbe jutás relatív intenzitása jelentősebben nőtt, mint az előző példánál, ahogy ezt előre meg is lehetett becsülni. Ennek következtében a D_1 mátrix rugalmasabban konstruálható. Hasonlítsuk ezt össze a $\delta_{rate} = 0.9$ értékhez tartozó korrelációs tartományt optimalizáló eljárással:

$$\underline{\pi}^{(0.9)} = [0.507, 2.76e - 8, 0.323, 0.0195, 0.15],$$

$$D_0^{(0.9)} = \begin{bmatrix} -8 & 7.66 & 7.89e - 8 & 0.328 & 1.47e - 8 \\ 0 & -31.6 & 0.991 & 3.08 & 0.136 \\ 0 & 17.4 & -95.4 & 51.5 & 26.6 \\ 0 & 29.66 & 20.6 & -127 & 33.7 \\ 0 & 6.67e - 6 & 30.3 & 20.83 & -159 \end{bmatrix}$$

Az előző példához hasonlóan az első állapotnál itt is megfigyelhető bizonyos mértékű kiegyenlítődés a kezdeti valószínűségvektorban. Ezenkívül az első állapotban ugyanúgy tapasztalható a főátló mellett elem növekedése, ahogy a főátlóbeli elemek jelentős változása is. Egy új jelenség is észrevehető: a transzformáció az APH-t ciklikus PH-vá alakította. (Valójában ez az előző példánál is így volt, azonban ott a főátló alatti elemek egyike sem volt igazán nagy.) Ez is mutatja, hogy az átalakítás rugalmasabb, mint a Buchholz-Kriege-módszer.

4.2 MAP illesztése MAP segítségével

A következő, a valós illesztési feladatoknak jobban megfelelő példa egy MAP illesztése MAP segítségével. Ehhez először szimulációt végeztem egy MAP alapján, majd a szimulált érkezési időköz eloszlásra a PhFit segítségével PH-t illesztettem, végrehajtottam az optimalizációs eljárásokat, majd az eredményt MAP-pá egészítettem ki. Ennél a lépésnél az optimalizációt az együttes momentumoktól való súlyozott négyzetes távolság szerint illesztettem, vagyis a

$$\theta = \sum_{(i,j) \in \mathcal{J}} \beta_{i,j} \left(1 - \frac{\mu_{i,j}}{\nu_{i,j}}\right)^2$$

kifejezést minimalizáltam, ahol $\nu_{i,j}$ az eredeti MAP $[X_k^i X_{k+1}^j]$ együttes momentuma, $\mu_{i,j}$ az illesztett MAP $[X_k^i X_{k+1}^j]$ együttes momentuma, $\beta_{i,j}$ egy súlyozó tényező, \mathcal{J} pedig a figyelembevenni kívánt (i, j) momentum párok halmaza. A [3] cikk is, melyben a Buchholz-Kriege-féle reprezentáció optimalizációs módszer alkalmazása található, ezt a célfüggvényt alkalmazza. (A jelölések megegyeznek, azzal a különbséggel, hogy az itteni β helyett ott β^2 a súlyozó tényező.) A θ minimalizálása nemlineáris optimalizálási feladat, melynek megoldását legkisebb négyzetek szerint illesztettem.

Korábban már említettem, hogy a $\mu_{1,1}$ együttes momentum és a ρ_1 korreláció között szoros kapcsolat van, utóbbi az előbbi szórással normált változata, így

ha nagyobb korrelációs tartományban tudunk illeszteni, várható, hogy a $\mu_{1,1}$ együttes momentumot is jobban tudjuk illeszteni, ez pedig azért fontos, mert elsősorban az alacsonyabb együttes momentumok illesztése a fontos, ezek nagyobb megbízhatósággal számíthatók valós adatok esetén is. Ezzel összhangban a $\beta_{i,j}$ súlyozó tényezőt érdemes az $i + j$ növekedésével csökkenteni.

Az illeszteni kívánt MAP a következő D_0 és D_1 mátrixokkal rendelkezik:

$$D_0 = \begin{bmatrix} -3.1 & 0.78 & 0.02 \\ 0.1 & -1.206 & 0.005 \\ 0.001 & 0.002 & -0.211 \end{bmatrix}$$

$$D_1 = \begin{bmatrix} 0.2 & 2.0 & 0.1 \\ 1 & 0.1 & 0.001 \\ 0.005 & 0.003 & 0.2 \end{bmatrix}$$

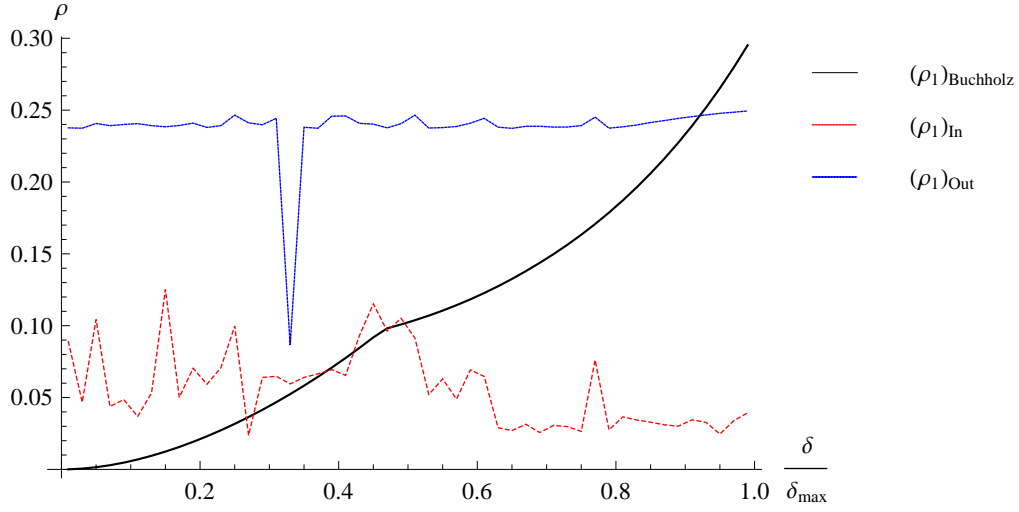
lag-1 korrelációja $\rho_1 = 0.286$. Ez az érték nagyobb, mint amit az eddigiekben sikerül elérni, azonban kisebb állapotszámú PH esetén ez könnyebben elérhető.

Az első lépés tehát az érkezők szimulációja volt, itt 100.000 érkezőt szimuláltam. Ez megfelelően nagy érték ahhoz, hogy a statisztikai jellemzők megfelelő pontossággal beálljanak, azonban az adatsor viszonylag könnyen kezelhető marad.

A szimuláció után a korábban már említett, PhFit nevű program segítségével egy harmadrendű PH-t illesztettem az előbb növekvő sorrendbe rendezett érkezői időközökre. A PhFit többféle bemenetet is képes feldolgozni: sorbarendezett mintákat, és sorbarendezett súlyozott mintákat, valamint megfelelően megadott eloszlás-, és komplementer eloszlásfüggvényt is. Az illesztés során lehet törzs- és farokeloszlást is illeszteni, munkám szempontjából csak a törzseloszlás illesztése lényeges. A PhFit kimenete kanonikus formájú PH, ezen végeztem el a különféle optimalizálási eljárásokat.

A maximális (pozitív) korreláció δ_{rate} függését a kilencedik ábrán megvizsgálva a Buchholz-Kriege-módszer szempontjából hasonló eredményt kapunk, mint a második PH optimalizációs példánál. A Buchholz-Kriege-módszerrel kapott maximális korreláció a δ_{rate} növelésével monoton módon nő. A bemeneti entrópiát optimalizáló eljárás meglehetősen rosszul teljesít, a kimeneti entrópiát optimalizáló viszont megközelítőleg konstans értéket mutat, és nagyjából $\delta_{rate} = 0.9$ -ig jobb korrelációra képes, mint a Buchholz-Kriege-módszer.

A tizedik ábrán látható, hogy az entrópia összeg alapú optimalizációs módszer szinte teljesen azonos, a kimeneti korrelációt optimalizáló eljáráséhoz nagyon közeli értéket ad végig, az eredményképp kapott reprezentáció is csak minimálisan változik. A maximális korrelációt és a korrelációs tartományt optimalizáló eljárások a legjobbak most is, de a $\delta_{rat} = 0.99$ -es érték környékén



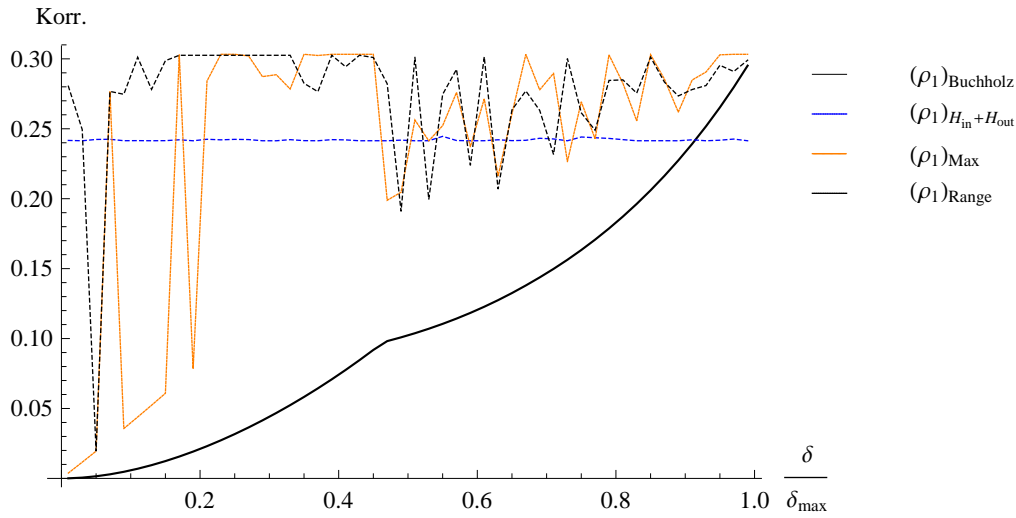
Ábra 9: Képleírás

már alig előnyösebbek a maximális korreláció szempontjából, mint a Buchholz-Kriege-módszer. Az is látható, hogy ezek az eljárások általában képesek elérni az eredeti MAP által meghatározott korrelációt, míg a Buchholz-Kriege-módszerrel ez csak 0.98-nál nagyobb δ_{rate} esetén teljesül. A kapott PH reprezentációk alapján az is látható, hogy a korreláció mértékű eljárások nagyjából hasonló $\underline{\pi}$ -t és \mathbf{D}_0 -t állítanak elő különböző δ_{rate} -ek esetén. A korrelációs tartományt optimalizálva például $\delta_{rate} = 0.8$ és $\delta_{rate} = 0.99$ esetén:

$$\underline{\pi}^{(0.8)} = [0.384, 1.42e - 7, 0.616], \quad \mathbf{D}_0^{(0.8)} = \begin{bmatrix} -0.225 & 0.00369 & 2.30e - 6 \\ 0 & -1.39 & 0.00592 \\ 0 & 1.27 & -3.08 \end{bmatrix}$$

$$\underline{\pi}^{(0.99)} = [0.396, 2.3e - 7, 0.604], \quad \mathbf{D}_0^{(0.99)} = \begin{bmatrix} -0.225 & 0.00136 & 9.04e - 7 \\ 0 & -1.40 & 0.0297 \\ 0 & 1.25 & -3.06 \end{bmatrix}$$

A három állapotú MAP esetén a [3] $(n - 1)^2$, azaz összesen négy együttes momentum illesztését javasolja, egész pontosan a $\mu_{1,1}, \mu_{1,2}, \mu_{2,1}, \mu_{2,2}$ együttes momentumokat, mivel a $\underline{\pi}(-\mathbf{D}_0^{-1}\mathbf{D}_1 = \underline{\pi}$ és $\mathbf{0}\mathbf{1} + \mathbf{D}_1\mathbf{1} = \mathbf{0}$ korlátoknak köszönhetően ezek teljesen behatárolják a MAP-ot olyan esetekben, amikor pontos illesztés lehetséges. Sok esetben ez nincs így, ekkor az $(n - 1)^2$ szám nem annyira kiemelt jelentőségű, mivel azonban MAP alapján generált érkezéseket illesztettem, ez mindenképp ésszerű választás volt. Súlyozó tényezőnek a [3]



Ábra 10: Képleírás

cikkben alkalmazottak négyzetgyökét vettem, vagyis $\beta_{i,j} = 2^{(i-1)(j-1)}$, tehát például a $\mu_{2,2}$ súlya 0.25.

A $\delta_{rate} = 0.99$ pontban vizsgálva az eljárásokat, az tapasztalható, hogy, bár a Buchholz-Kriege-módszer eredményeképp létrejövő PH reprezentáció eléri a kívánt maximális korrelációt, az együttes momentumok szerinti illesztés nem talál olyan D_1 mátrixot, mely tökéletes MAP-ot eredményezne, azonban jó közelítést mutat az esetek többségében. Az alábbi táblázat mutatja az eredményeket:

Módszer/Korreláció	θ	$1 - \frac{\nu(1,1)}{\mu_{1,1}}$	ρ_1
Buchholz-Kriege-módszer	0.00391	-0.0448	0.288
Bemeneti entrópia	0.591	0.336	0.0394
Kimeneti entrópia	0.0219	0.0141	0.249
Entrópiák összege	0.0291	0.0264	0.241
Entrópiák szorzata	0.0291	0.0264	0.241
Minimális korreláció	0.00402	-0.0463	0.289
Maximális korreláció	0.00384	-0.0446	0.287
Korrelációtartomány	0.00398	-0.0460	0.289 height

Ezzel szemben azonban $\delta_{rate} = 0.9$ esetén a Buchholz-Kriege-módszerre $\theta = 0.0375$, körülbelül 3,9%-os eltéréssel $\mu_{1,1}$ -re, míg például a korrelációtartományt optimalizáló eljárásra $\theta = 0.00398$ továbbra is. Összességében elmondható, hogy

a korrelációt optimalizáló eljárások stabilan jó eredményt mutatnak, az entrópia alapúak, a bemeneti entrópiát optimalizáló eljárásokon kívül pedig szintén elfogadhatóan teljesítenek.

4.3 Valós adatsor illesztése

Az igazi feladatot a valós adatsorok illesztése jelenti. Az általam vizsgált példa a pAug89 adatsor[10], mely Ethernet csomagok érkezési időközzeit tartalmazza. A trace első százezer sorát vizsgáltam. Az adatsor korrelációja 0.167-nek adódott. Az elvégzendő lépések az előző példához hasonlóak, annyi különbséggel, hogy az adatsor már adva van. Az illesztést itt három és öt állapotú állapotú MAP-pal végeztem. A PhFit által végzett, három állapotú illesztés eredményeképp a következő kanonikus formájú PH jön létre:

$$\underline{\pi} = [0.0384, 0.714, 0.248], \quad \mathbf{D}_0 = \begin{bmatrix} -90.1 & 90.1 & 0 \\ 0 & -867 & 867 \\ 0 & 0 & -884 \end{bmatrix}$$

A korábbiakhoz hasonlóan a Buchholz-Kriege-módszerrel elérhető korreláció itt is a δ_{rate} növekedtével nő, ezért a $\delta_{rate} = 0.99$ értékre végeztem el a vizsgálatokat. Azt tapasztalhatjuk, hogy a Buchholz-Kriege-féle eljárás rosszabb eredményt ad, mint az összes többi, melyek θ értéke alig különbözik. Az illesztés minden esetben valamivel rosszabb, mint az előző példánál, bár a $\mu_{1,1}$ első együttes momentumot minden módszer jobban illeszti, mint a MAP illesztésnél. Megfigyelhető az is, hogy a korrelációk jóval nagyobbak, mint az eredeti adatsor esetén, ami azt jelenti, hogy a szórásnégyzetek kisebbek, ugyanis a várható értékek nagyjából azonosak, és a $\mu_{1,1}$ együttes momentum eltérése is elenyésző ehhez a különbséghez képest.

Módszer/Korreláció	θ	$1 - \frac{\nu_{1,1}}{\mu_{1,1}}$	ρ_1
Buchholz-Kriege-módszer	0.0238	0.00142	0.307
Bemeneti entrópia	0.0193	-0.00546	0.313
Kimeneti entrópia	0.0194	-0.00535	0.312
Entrópiák összege	0.0189	0.0152	0.321
Entrópiák szorzata	0.0192	-0.0119	0.318
Minimális korreláció	0.0186	-0.0159	0.321
Maximális korreláció	0.0186	-0.0160	0.321
Korrelációtartomány	0.0189	-0.0151	0.321 height

A PhFit öt állapotú illesztésének eredménye alább látható:

$$\pi = [0.0433, 0.367, 0.444, 0.000162, 0.145],$$

$$D_0 = \begin{bmatrix} -96.7 & 96.7 & 0 & 0 & 0 \\ 0 & -1030 & 1030 & 0 & 0 \\ 0 & 0 & -1309 & 1309 & 0 \\ 0 & 0 & 0 & -1414 & 1414 \\ 0 & 0 & 0 & 0 & -3021 \end{bmatrix}$$

Az illesztés itt már sokkal rosszabb. A Buchholz-Kriege-módszer esetén a monoton növekvő kimeneti entrópia a $\delta_{rate} = 0.99$ pontban kevesebb, mint 0.17, ami a korábbi példákkal összevetve igen alacsony. Ezt a kimeneti entrópiát egyetlen módszer sem tudja megemelni, ez pedig, a korábbiakkal összehangban, elenyésző korrelációt eredményez. A nagyjából megegyező értékű, igen alacsony korrelációnak köszönhetően az együttes momentumok illesztése is sikertelen, a θ értéke 1.46 körüli értéket vesz fel, ami azt jelenti, hogy már a $\mu_{1,1}$ együttes momentum illesztett értéke is 36% körüli eltérést mutat a valós adatokhoz képest.

5 Értékelés

A fentiekben bemutatam egy olyan módszert, mely alkalmas a PH-k reprezentációjának javítására, majd ezt összehasonlítottam egy meglévő optimalizációs módszer, a Buchholz-Kriege-féle eljárás tulajdonságaival, mellyel szemben többféle elvi előnye is van. Nagyobb rugalmasságot biztosít, mely annak köszönhető, hogy többféle mérték alapján képes optimalizálni, bemeneti reprezentációja pedig a kanonikustól eltérő is lehet. Olyan mérték alapján is lehet optimalizálni, mely a később illesztendő dinamikus paraméterekkel közvetlen kapcsolatban van. A kimeneti reprezentáció szélesebb tartományban mozoghat: nincs az APH-k családjára korlátozva, és az eredeti reprezentációhoz képest megváltozhatnak főátlóbeli elemei.

A gyakorlat ezeket az elvi előnyöket részben igazolta, az elérhető korreláció sok esetben jelentősen megnőtt, az illesztés a mértékek többsége esetén megközelítette, vagy meghaladta a Buchholz-Kriege-módszerrel elérhető. A várakozásoknak megfelelően bebizonyosodott, hogy a csak bemeneti vagy csak kimeneti entrópiát optimalizáló eljárások nem adnak stabilan jó eredményt. Az is nyilvánvalóvá vált, hogy az entrópiák összegén és szorzatán alapuló eljárások jobb eredményt nyújtanak, de ezeknél a mértékeknél sincs szoros kapcsolat a korrelációval nagyobb értékek esetén. A korreláció alapú módszerek bizonyultak a legjobbaknak, ezek hátránya nagyobb számítás igényükben rejlik. Az, hogy az eljárások mohó algoritmust alkalmaznak, továbbra is jelentkezik. Amikor

a Buchholz-Kriege-módszer nem ad kedvező kezdeti reprezentációt, az eljárás ugyanúgy megakadhat.

Ezeket figyelembevéve többféle további javítási lehetőség kínálkozik. A Buchholz-Kriege-módszer helyett megvizsgálható más módszer is az algoritmus kezdeti szakaszába, mely kedvezőbb kiindulási reprezentációt nyújt a további, mérték alapú optimalizálás számára. Kézenfekvő új mértékek keresése, melyek lehetnek akár az entrópia alapú mértékek továbbfejlesztései, vagy alapulhatnak dinamikus paramétereken. A mohó algoritmus valamilyen módosításával a lokális optimumok egy része esetleg kikerülhetővé válik.

Hivatkozásjegyzék

- [1] S. Asmussen and O. Nerman. Fitting Phase-type distributions via the EM algorithm. In *Proceedings: "Symposium i Advent Statistik"*, pages 335–346, Copenhagen, 1991.
- [2] A. Bobbio and A. Cumani. ML estimation of the parameters of a PH distribution in triangular canonical form. In G. Balbo and G. Serazzi, editors, *Computer Performance Evaluation*, pages 33–46. Elsevier Science Publishers, 1992.
- [3] P. Buchholz and J. Kriege. A heuristic approach for fitting maps to moments and joint moments. In *Proc. QEST'09*, pages 53–64, Budapest, Hungary, 2009. CPS press.
- [4] A. Cumani. On the canonical representation of homogeneous Markov processes modelling failure-time distributions. *Microelectronics and Reliability*, 22:583–602, 1982.
- [5] R. El Abdouni Khayari, R. Sadre, and B. Haverkort. Fitting world-wide web request traces with the em-algorithm. In *Proc. of SPIE*, volume 4523, pages 211–220, Denver, USA, 2001.
- [6] A. Horváth and M. Telek. Approximating heavy tailed behaviour with phase type distributions. In G. Latouche and P. Taylor, editors, *Advances in algorithmic methods for stochastic models, MAM3*, pages 191–214. Notable Publications Inc., 2000.
- [7] A. Horváth and M. Telek. PhFit: A general purpose phase type fitting tool. In *Tools 2002*, pages 82–91, London, England, April 2002. Springer, LNCS 2324.

- [8] Buchholz Peter Horváth Gábor, Telek Miklós. A map fitting approach with independent approximation of the inter-arrival time distribution and the lag correlation. In *Proc. QEST'09*, pages 124–133. IEEE CS, 2005.
- [9] A. Thümmler, P. Buchholz, and M. Telek. A novel approach for fitting probability distributions to trace data with the EM algorithm. In *Int. Conf. on Dependable Systems and Networks (DSN)*, Yokohama, Japan, 2005. IEEE CS.
- [10] The internet traffic archive. <http://ita.ee.lbl.gov/index.html>.