

# Időben változó sorbanállási rendszer vizsgálata mátrix-geometrikus eljárások segítségével

ÉLTETŐ TAMÁS, VADERNA PÉTER

{tamas.elteto, peter.vaderna}@ericsson.com

**Kulcsszavak:** sorbanállási rendszer, Markov-érkezési folyamat, mátrix-geometriai módszer

*Ebben a cikkben egy végtelen sok kiszolgálóval rendelkező sorbanállási rendszert vizsgálunk időfüggő esetben, azaz nem a stacionárius megoldását keressük, hanem a rendszerben lévő igények számát követjük nyomon az idő múlásának függvényében. A sorbanállási rendszerbe igények érkeznek egy nem stacionárius pontfolyamat szerint, és az igények kiszolgálása is véletlen ideig tart. Ezt a nem stacionárius érkezési folyamatot egy Markov érkezési folyamatként adjuk meg. Az időben változó sorbanállási rendszer egy inhomogén lineáris differenciálegyenlet-rendszerre vezet, melynek megoldását kiszámítjuk. A számítási módszert egy gyakorlati példára – egy távközlési szolgáltatás vizsgálatára – alkalmazzuk.*

## 1. Bevezetés

Végtelen kiszolgálóval rendelkező sorbanállási modellek először a hagyományos telefonközpontok méretezése során vetődtek fel. A modellben Poisson folyamat szerint érkező telefonhívásokat fogadott és továbbított a telefonközpont, ahol csak véges sok vonal áll rendelkezésre a hívások továbbítására. Ha ismerjük a telefonközponton egy időben átmenő hívások számának stacionárius eloszlását, akkor meg tudjuk tervezni a kimenő vonalak számát úgy, hogy csak nagyon kicsi valószínűsége legyen az érkező hívások kapacitáshiány miatti visszautasításának.

A távközlési hálózatok fejlődésével, a szolgáltatások sokaságának megjelenésével az eredeti feladat általánosítása is szükségessé vált. Az Internet forgalom statisztikai elemzése során kimutatták, hogy a Poisson folyamat sok esetben nem elég jó modell adatahívások, adatátviteli kapcsolatok érkezési folyamatának a leírására [3]. Másrészt, mint az ismertetésre kerülő példa is mutatja, vannak esetek, mikor a stacionárius állapot leírása nem elegendő kérdéseink megválaszolásához.

A cikkben a Poisson folyamat általánosításaként az úgynevezett Markov érkezési folyamatot [1] fogjuk alkalmazni. Az érkezéseket egy Markov lánc állapotai, illetve állapot-átmenetei befolyásolják. Mivel célunk az volt, hogy a végtelen kiszolgálós sorbanállási rendszerek tranziens viselkedésének lehető legáltalánosabb analitikus leírását adjuk, így a kiszolgálási időnek – a hívások hosszának – az eloszlására is kellőképp gazdag eloszláscsaládot választottunk. A cikkben fázis típusú eloszlásokat használunk, melyek elnyelő állapottal rendelkező Markov láncok abszorpcióig eltelő idejét írják le. Ez az eloszláscsalád a gyakorlatban is jól használható, mivel ismert, hogy kellően nagy állapotterű fázis típusú eloszlással bármely eloszlást meg tudunk tetszőlegesen pontosan közelíteni [4].

A Markov érkezési folyamatok és fázis típusú tartási idők segítségével a feladat, bár nagyon általános, tel-

jes mértékben markovi maradt. Ez lehetővé teszi a rendszer analitikus vizsgálatát nem csak stacionárius esetben, hanem követni lehet a rendszer időbeli változását is, ami nagy előnyt jelent más – inkább a stacionárius állapot leírására törekvő – módszerekkel szemben. A rendszerben tartózkodó hívások számának momentumaira megadott időfüggő differenciálegyenlet-rendszert megoldjuk, és megmutatjuk, hogy a megoldás paraméterei hogyan függenek a bemeneti paraméterektől. Megjegyezzük, hogy a vizsgálatunk tárgyát képező rendszerrel [5]-ben egy általánosabb esetet elemeznek, mi egy eltérő, alkalmazás-központúbb megközelítést mutatunk.

A számítási módszer segítségével egy gyakorlatban felmerülő problémát oldunk meg. A feladat egy internetes hírszolgáltató működéséhez kapcsolódik. Azt feltételezzük, hogy a hírszolgáltató a különböző operációs rendszerekkel és sokféle internetes böngészővel rendelkező felhasználóknak egységes külalakkal kívánja a híreket megjeleníteni. Ezt úgy éri el, hogy amikor egy felhasználó egy hírt kér a kiszolgálótól, akkor a kiszolgáló érzékeli a felhasználó beállítását. A rendelkezésre álló információk alapján egy konvertáló segítségével elkészíti, majd elküldi a hírek egy olyan változatát, amelyet a beállításoknak megfelelően formázott. A hírek konvertálásának folyamatát írjuk le az ismertetett számítások segítségével.

A cikk következő fejezetében definiáljuk a használt matematikai eszközöket, majd ismertetjük a számítási eljárás lényeges lépéseit. Ezek után bemutatjuk a gyakorlati problémát, végül összefoglaljuk a cikket. Az írásunk végén található függelékben pedig elméleti állításaink bizonyításai olvashatók.

## 2. Matematikai háttér

Az alábbiakban röviden összefoglaljuk a Markov érkezési folyamatot (angol betűszóval: MAP). A téma részletesebb tárgyalása [1]-ben található angol nyelven.

A Markov érkezési folyamat egy olyan egyszerű pontfolyamat, amelynek érkezéseit egy folytonos idejű Markov folyamat befolyásolja. Tekintsünk egy sztochasztikus folyamatot:

$$\{(N(t), J(t)) : t \geq 0\},$$

ahol  $N(t)$  a  $(0, t)$  intervallumban történő érkeзések számát, míg  $J(t)$  a  $t$  időpontban a folyamat „fázisát” jelöli. Itt  $J(t)$  egy folytonos idejű Markov folyamat egy véges,  $M$  elemű állapotterén,  $N(t)$  pedig egy  $J(t)$ -től függő sztochasztikus folyamat. Az  $(N(\cdot), J(\cdot))$  együttes sztochasztikus folyamat egy folytonos idejű Markov folyamat lesz az  $\{(n, j) : n \geq 0, 1 \leq j \leq M\}$  állapotterén, az alábbi definíció szerint.

Osszuk fel az  $\{(n, j) : n \geq 0, 1 \leq j \leq M\}$  állapotteret rész-halmazokra:

$$I(k) = \{(k, 1), (k, 2), \dots, (k, M)\},$$

ahol  $k$  nemnegatív egész. Ekkor legyen  $Q$  mátrix az  $(N(\cdot), J(\cdot))$  folyamat infinitezimális generátor mátrixa az alábbi definícióval:

$$Q = \begin{bmatrix} D_0 & D_1 & 0 & 0 & \dots \\ 0 & D_0 & D_1 & 0 & \dots \\ 0 & 0 & D_0 & D_1 & \dots \\ 0 & 0 & 0 & D_0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

ahol

- $D_0$  és  $D_1$   $M \times M$ -es mátrixok,
- $D_1 \geq 0$ ,
- $[D_0]_{i,j} \geq 0, 1 \leq i \neq j \leq M$ , és  $[D_0]_{i,i} \leq 0, 1 \leq i \leq M$ ,
- $D = D_0 + D_1$  mátrix sztochasztikus, vagyis minden sorösszege 0.

Szemléletesen arról van szó, hogy a  $J(\cdot)$  folytonos idejű Markov folyamat infinitezimális generátor mátrixa  $D$ . Ez a mátrix bomlik fel két mátrix összegére. Ezek közül a  $D_0$  mátrix tartalmazza azokat az állapotátmeneteket, ahol nem történik érkezés, míg  $D_1$  mátrix írja le azokat az állapotátmeneteket, ahol történik érkezés, illetve szintén  $D_1$  mátrix adja meg a feltételes Poisson folyamatok érkezési intenzitását, amikor is nem vált állapotot a  $J(\cdot)$  folyamat, de érkezés történik. (A Markov érkezési folyamatokat általánosíthatjuk nem egyszerű pontfolyamatokra is, azaz megengedhetünk egyszerre több érkeзést is. Ekkor további  $D_n$  mátrixok írják le a csoportos érkeзéseket.)

A Markov érkezési folyamatok egy speciális osztályát Markov-modulált Poisson folyamatoknak nevezzük. A Markov-modulált Poisson folyamatok esetén a  $D_1$  mátrixnak csak a diagonálisában vannak nem nulla elemek, azaz érkezés nem fordulhat elő, amikor a  $J(\cdot)$  Markov folyamat állapotot vált. A Markov-modulált Poisson folyamatok részletes összefoglalása [2]-ben található.

Egy másik speciális osztály az úgynevezett PH felújítási folyamatok osztálya. Ezek olyan felújítási folyamatok, melyek felújítási időközének eloszlása PH (fázis típusú) eloszlás. A PH eloszlások tulajdonképpen különböző paraméterű Erlang eloszlások keverékei, de inkább

egy elnyelő állapottal rendelkező Markov folyamatnak az indítás és az elnyelődés között eltelt idejeként definiálják őket.

Ezen második definíciónak megfelelően, a PH eloszlásokat egy  $(\alpha, T)$  párral reprezentálhatjuk. Az  $\alpha$  vektor adja meg a tranziens állapotok eloszlását a Markov lánc indításakor. A  $T$  mátrix pedig egy szubsztocasztikus mátrix, azaz

- $T$  invertálható mátrix,
- $[T]_{i,j} \geq 0, 1 \leq i \neq j \leq M$ , és  $[T]_{i,i} \leq 0, 1 \leq i \leq M$ ,
- $T$  mátrix sorösszege nem pozitív.

A  $T$  mátrix a lényeges része az elnyelő Markov folyamatunk generátor mátrixának, ami a következőképpen írható le:

$$\begin{bmatrix} T & \tau \\ 0 & 0 \end{bmatrix},$$

ahol

$$\tau = -T \mathbb{1}, \quad \mathbb{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

A PH eloszlás eloszlásfüggvényét a következő módon fejezhetjük ki az  $(\alpha, T)$  pár segítségével:

$$F(t) = 1 - \alpha e^{Tt} \mathbb{1},$$

ahol az  $\alpha$  vektorral fejezzük ki az elnyelő Markov lánc kezdeti eloszlását.

Ekkor  $D_0 = T$  és  $D_1 = \tau \alpha$ , vagyis a  $J(\cdot)$  Markov folyamat állapottere megegyezik az elnyelő Markov folyamat tranziens állapotaival. Amikor az elnyelődés következik, akkor a Markov érkezési folyamatban érkezés történik, és a  $J(\cdot)$  folyamat újraindul  $\alpha$  kezdeti eloszlással.

### 3. Egy korlátlan számú kiszolgálóval rendelkező sorbanállási rendszer momentumai

Tekintsünk egy olyan sorbanállási rendszert, melybe az igények MAP folyamat szerint érkeзnek, és végtelen sok kiszolgáló egymástól függetlenül, párhuzamosan szolgálja ki az igényeket PH eloszlású kiszolgálási idővel. A cél a rendszerben egyidejűleg kiszolgálás alatt álló igények számának meghatározása.

Jelölje  $X(t)$  a kiszolgálás alatt lévő igények számát  $t$  időpontban és  $J(t)$  a MAP érkezési folyamat fázisát  $t$ -időpontban! Jelölje továbbá  $\mu^{(K)}(t)$  azt az  $M$  dimenziós vektort, melynek  $i$ -ik koordinátája  $\mu_i^{(K)}(t)$ , ahol

$$\begin{aligned} \mu_i^{(K)}(t) &= E[X^{(K)}(t) | X(0) = 0, J(0) = i], \\ K &\geq 1, 1 \leq i \leq M, \end{aligned} \quad (1)$$

$$X^{(K)}(t) = X(t)[X(t) - 1] \cdots [X(t) - K + 1]$$

Azaz, itt  $\mu_i^{(K)}(t)$  a kiszolgálás alatt lévő igények számának  $K$ -dik faktoriális momentuma, ha a rendszert az  $i$ -dik fázisból indítjuk.

A cél az, hogy számítsuk ki a fenti sorbanállási rendszer momentumainak időbeli változásait leíró függvé-

nyeket. Ezeket a függvényeket egy differenciálegyenlet-rendszer határozza meg. A momentum függvények kiszámításának módját a Függelékben ismertetett (10)-(16) egyenletekből olvashatjuk le. A momentum függvények lényegében

$$t^j e^{-\gamma_i t}, \text{ illetve } t^j e^{\gamma_m t}$$

alakú függvények lineáris kombinációi. Minél magasabb momentum függvényt kívánunk kiszámítani, annál több tagja van a függvényt megadó összegnek. Az exponensek és a lineáris kombináció együttthatóinak a kiszámítását a Függelékben ismertetjük.

#### 4. Gyakorlati példa

Ebben a fejezetben egy példa segítségével vizsgáljuk a momentumok gyakorlati kiszámításának lehetőségeit. A példa egy hírszolgáltató működését vizsgálja. A hírszolgáltatás egy honlapon keresztül működik. A hírszolgáltató szervere érzékeli a felhasználók böngészőinek típusát és az operációs rendszereket. A cél az, hogy a lehető legtöbb típusú operációs rendszeren futó különféle böngészőkkel a lehető legjobb minőségben jelenjenek meg a cikkek. (Tehát a képek mérete, felbontása, a szöveg tördelése, a honlapon esetleg előforduló animációk hasonlóképpen mutassanak a különféle típusú számítógépeken.)

A szolgáltatás működését az 1. ábrán illusztráljuk.

Célunk a konvertálási funkcionalitás vizsgálata egy új hír megjelenésétől egészen addig, amíg már az összes lehetséges formátumban létezik a cikk. A felhasználók gyors kiszolgálása érdekében egyszerre több konvertálás is előfordulhat. Ha a konvertálók száma nem elég nagy, akkor előfordulhat az az eset, amikor az összes konvertáló foglalt, és ezért a kiszolgáló nem képes egy új beállításokkal rendelkező felhasználó kérését teljesíteni. A konvertálók száma viszont nem lehet nagyon nagy sem, például azért, mert egy ilyen berendezés drága. A feladat az, hogy megállapítsunk a foglalt konvertálók számára egy olyan felső korlátot, melynél csak nagyon kicsi valószínűséggel lesz egyszerre több foglalt.

Az egyszerűség kedvéért feltételezzük, hogy a felhasználók kérései Poisson folyamat szerint érkeznek, bár ennél általánosabb eseteket is tudnánk kezelni. A példában használt adatok az illusztrációt szolgálják, a gyakorlatban nem vizsgáltuk, hogy hányféle böngészővel olvasnak cikkeket a felhasználók, illetve azt sem, hogy mennyi ideig tarthat egy konvertálás elvégzése egy valós számítógépen. Feltesszük tehát, hogy összesen tízféle böngészőre van konvertáló eljárás, és átlagosan másodpercenként 8 kérés érkezik. A konvertáláshoz szükséges idő exponenciális eloszlású, és egy konvertálás átlagosan 5 másodpercig tart.

$$H(t) = 1 - e^{-\frac{t}{5}}$$

A tízféle böngészőtípus előfordulási gyakoriságait és a másodpercenként előforduló kérések számát (érkezési intenzitást) az 1. táblázat foglalja össze:

Típus	Gyakoriság	Intenzitás
1. típus	21%	1.68
2. típus	20%	1.6
3. típus	13%	1.04
4. típus	13%	1.04
5. típus	13%	1.04
6. típus	4%	0.32
7. típus	4%	0.32
8. típus	4%	0.32
9. típus	4%	0.32
10. típus	4%	0.32

1. táblázat

A 2. táblázatban látható, hogy a Markov érkezési folyamat állapotterének leírása számnégyesek segítségével lehetséges. Ezekkel a számnégyesekkel körülbelül azt írjuk le, hogy melyik típusú felhasználó számára van már a tárolóban megfelelő változat.

Gyakoriság	Típusok	Intenzitás
21%	1.	1.68
20%	2.	1.6
13%	3.,4.,5.	1.04
4%	6.,7.,8.,9.,10.	0.32

2. táblázat

1. ábra A hírszolgáltatás működésének vázlata



A számnégyes tagjainak a jelentései sorrendben:

- 1) Ha még nem történt konvertálás az 1. típusú felhasználók részére, akkor értéke 0, különben 1.
- 2) Ha még nem történt konvertálás a 2. típusú felhasználók részére, akkor értéke 0, különben 1.
- 3) Itt számoljuk a 3.-5. típusú felhasználók számára elvégzett konvertálásokat. Ez a tag 0 és 3 közötti értékeket vehet fel.  
Itt fontos megjegyezni, hogy nem számít az, hogy milyen sorrendben érkeztek a 3.-5. típusú felhasználók, mivel gyakoriságaik egyformák.
- 4) Itt számoljuk a 6.-10. típusú felhasználók számára elvégzett konvertálásokat.  
Ez a tag 0 és 5 közötti értékeket vehet fel.  
Itt szintén azért elegendő egy számláló, mivel a 6.-10. típusú felhasználók gyakoriságai megegyeznek, és ezért itt sem számít, hogy milyen sorrendben érkeztek.

Az állapotok száma  $2 \cdot 2 \cdot 4 \cdot 6 = 96$ .

Az állapot átmenet csak olyan számnégyesek között lehetséges, ahol csak egyetlen tag különbözik eggyel. A kiinduló állapot a  $(0,0,0,0)$ , vagyis ekkor még a tároló üres. Az állapotátmeneteket a  $D_1$  96x96-os mátrixban adjuk meg. A  $D_1$  mátrix elemeit az alábbi módon adjuk meg.

Ha a mátrixelemnek megfelelő állapotátmenet az 1. típusú felhasználók részére történő konvertálást ír le, akkor a mátrixelem értéke  $8 \cdot 0.21 = 1.68$ , ahogyan azt az 1. táblázatban is láthatjuk az „Intenzitás” oszlopban. A 2. típusú konvertálási igények beérkezéséhez tartozó állapotátmenetek intenzitása 1.6.

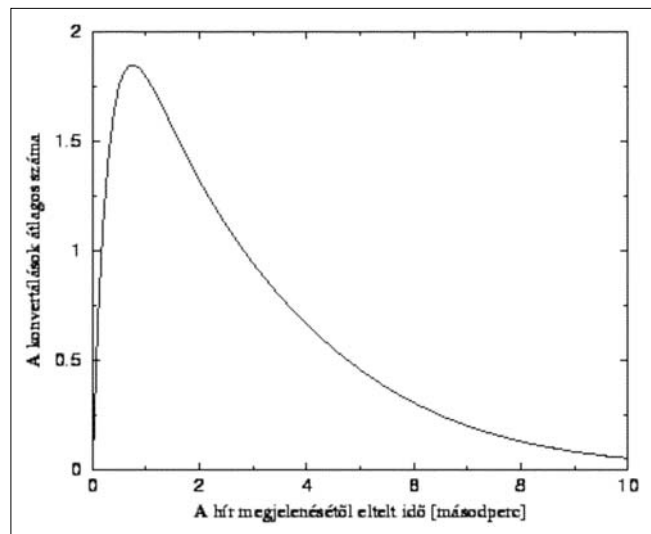
A 3.-5. típusú felhasználók esete egy kicsit különbözik az előzőektől. Ezek érkezési intenzitása egyforma, ezért az olyan átmenetknél, ahol az állapotleíró számnégyes 3. tagja változik, elegendő csupán annyit számon tartani, hogy eddig hány konvertálást végeztünk el a 3.-5. típusok számára. Az induláskor ugyanis  $3 \cdot 1.04 = 3.12$  intenzitással jöhet kérés a 3.-5. típusok valamelyikétől. Az első kérés után azonban már csak a megmaradó két típustól jöhet kérés, tehát az intenzitás lecsökken  $2 \cdot 1.04 = 2.08$ -ra, míg ha már kettő típust kiszolgáltunk, akkor a megmaradó 3.-5. típusú kérés érkezési intenzitása 1.04 lesz.

A 6.-10. típusú kéréseket leíró állapot átmenetekre ugyanazok vonatkoznak, mint a 3.-5. típusúak kéréseire azzal a különbséggel, hogy a kezdeti intenzitás  $5 \cdot 0.32 = 1.6$  lesz, és ez csökken le 0.32-re, amikor már csak 1 hiányzó típus marad.

Ekkor a hír megjelenésétől kezdve folyamatosan érkeznek a kérések a kiszolgálóhoz. Az első másodpercekben sorra érkeznek be olyan böngészőkről kérések, amelyekre a hírek még nem létezik konvertált változata. Ekkor van a konvertálóknak a legtöbb dolguk. Amikor a leggyakrabban előforduló beállításoknak megfelelő változatok már elkészültek, akkor egyre kevesebb konvertálásra van szükség. Ha 10 konvertáló áll rendelkezésre, akkor azonnal ki tudunk szolgálni minden igényt. Azonban látni fogjuk, hogy ennél kevesebb konvertáló is elegendő a feladat elvégzéséhez.

Miután megkonstruáltuk a  $D_0$  és  $D_1$  mátrixokat, kiszámítjuk a (6) egyenlethez szükséges  $V$  mátrixot, és ezzel megkapjuk a differenciálegyenlet-rendszer homogén részének a megoldását is. Megjegyezzük, hogy a  $D = D_0 + D_1$  mátrix minden sajátértéke valós és egyszeres. A partikuláris megoldás felírásához először kiszámítjuk a  $H(t)$  függvényből kiindulva a (7) egyenletben szereplő  $f_k(t)$  függvényeket, melyek (8)-as alakúak. Ezek után felírhatjuk a  $\hat{c}_m(t)$  függvényeket (10) alapján. A jelen példában a  $\bar{c}_m(t)$  függvények felírására nincs szükség. Végül pedig a momentum függvények a homogén rész és a partikuláris megoldás összegeként állnak elő (15)-ben.

A 2. ábrán az egyszerre foglalt konvertálók átlagos számának időbeli változását láthatjuk. Az is látható az ábrán, hogy a legnagyobb foglaltság a hír megjelenése után körülbelül 0.75 másodperccel következik be.



2. ábra  
A foglalt konvertálók átlagos számának alakulása röviddel egy új hír megjelenése után

Annak érdekében, hogy egy olyan korlátot állapíthassunk meg, melynél több foglalt konvertáló előfordulásának a valószínűsége kicsi, kiértékeljük a párhuzamosan futó konvertálások számának momentumait az új hír megjelenése után 0.75 másodperccel. Ekkor azt feltételezzük, hogy tetszőleges számú konvertáló áll a rendelkezésünkre. A 3. táblázat tartalmazza az eredményeket:

	Faktorális momentum	Momentum
1.	1.848	1.848
2.	3.063	4.911
3.	4.499	15.537
4.	5.763	56.048
5.	6.307	224.208
6.	5.734	976.641
7.	4.157	4573.540
8.	2.253	22805.388
9.	0.812	120181.397
10.	0.146	665327.954

3. táblázat

Mivel legfeljebb 10 párhuzamos konvertálás fordulhat elő, így az első 10 momentum segítségével egyszerűen kiszámítható a párhuzamos konvertálások számának az eloszlása egy Vandermonde típusú mátrix segítségével [10]. A 4. táblázatban azt láthatjuk, hogy milyen valószínűséggel lép túl a párhuzamos konvertálások száma ( $N$ ) különböző korlátokat, amelyeket  $n$ -el jelöltünk.

4. táblázat

A párhuzamosan futó konvertálások számának eloszlása a hír megjelenése után 0.75 másodperccel

$n$	$P(N > n)$
0.	0.871
1.	0.577
2.	0.276
3.	0.095
4.	0.023
5.	0.004
6.	$5 \cdot 10^{-4}$
7.	$4 \cdot 10^{-5}$
8.	$2 \cdot 10^{-6}$
9.	$3 \cdot 10^{-8}$
10.	0

Ha a tervezés során 1 ezrelék alá szeretnénk csökkenteni annak a valószínűségét, hogy a foglalt konvertálók, illetve párhuzamosan futó programszálak száma meghaladja a rendelkezésre álló kapacitásokat, a táblázatból megállapíthatjuk, hogy 10 konvertáló helyett elegendő csupán 5 konvertáló, illetve ennek megfelelő processzorteljesítmény használata, mivel annak valószínűsége, hogy több, mint 5-re lenne szükség  $5 \cdot 10^{-4}$ .

### 5. Összefoglalás

Jelen cikkben egy végtelen számú kiszolgálóval rendelkező (Kendall jelölésével MAP/PH/ $\infty$ ) sorbanállási rendszer tranzien viselkedését vizsgáltuk analitikus eszközök segítségével. A sorbanállási rendszer momentumaira egy differenciálegyenlet-rendszert adtunk meg, melynek megtaláltuk az időfüggő megoldását. A sorban tartózkodó igények számának átlagát, szórásnégyzetét illetve magasabb momentumait felhasználva megkaptuk a sorban tartózkodó igények számának az eloszlását is, szintén az eltelt idő függvényében.

A számítási módszer hasznosságát egy gyakorlati példán keresztül mutattuk be. A példa egy internetes hírszolgáltatás lehetséges működéséről szól. A hírszolgáltató minden hírt testre szab a különböző beállításokkal rendelkező előfizetői csoportok számára. Ez a folyamat azonban időigényes, és előfordulhat, hogy egy időben több előfizetőnek is konvertálni kell a hírt. Az analitikus modell segítségével kiszámítjuk az egy időben szükséges konvertálások számát, és egy felső korlátot adunk erre, amelyet a rendszer nagy valószínűséggel nem lép túl. A korlát pedig megadja a szolgáltatás zavartalan működéséhez szükséges számítási kapacitás nagyságát.

Mind az elméleti, mind a gyakorlati példához kapcsolódóan vannak további kutatási lehetőségek.

A momentumok kiszámításakor feltételeztük, hogy a rendszer az indulás pillanatában nem szolgál ki egyetlen igényt sem. Ezt a (2) illetve (3) egyenletekkel leírt differenciálegyenlet-rendszer kezdeti feltételeinél rögzítettük. Feltételezhetjük azonban azt is, hogy induláskor a rendszerben véletlen számú igény van. Az igények

számának eloszlása ismeretében meghatározhatjuk a faktoriális momentumokat, és ezek a momentumok lesznek a differenciálegyenlet-rendszer kezdeti feltételei.

A gyakorlati példában felmerül az a probléma, hogy nagyszámú és sokféle lehetséges beállítási típusnál a Markov érkezési folyamat állapottere igen nagy lesz, és emiatt a számítás futási ideje jelentős mértékben megnő. A példában megjelenő Markov érkezési folyamat azonban tulajdonképpen hasonló struktúrájú független Markov érkezési folyamatok szuperponáltja, így a számítások egy elemi esetre egyszerűsíthetők le, mely szintén analitikusan kezelhető, és ezzel a számítási igény csökkenthető.

További problémát kaphatunk a gyakorlati példa módosításával. Tegyük fel, hogy egyetlen konvertáló eszközünk van, és a konvertálási kéréseket egy FIFO sorban tároljuk, amíg kiszolgálásra nem kerülnek. A probléma ekkor egy véges számú kiszolgálóval rendelkező sorbanállási rendszerre vezethető vissza. Ebben az esetben más matematikai eszközöket felhasználó analitikus vizsgálatokra van lehetőségünk.

### 6. Függelék

[6] alapján a következő differenciálegyenlet-rendszert adhatjuk meg a  $\mu^{(k)}(t)$  vektorokra:

$$\frac{d}{dt} \mu^{(1)}(t) = \mathbf{D} \mu^{(1)}(t) + \{1 - H(t)\} \mathbf{D}_1 \mathbf{1}, \quad \mu^{(1)}(0) = 0, \quad (2)$$

és  $K \geq 2$  esetén (3)

$$\frac{d}{dt} \mu^{(k)}(t) = \mathbf{D} \mu^{(k)}(t) + K \{1 - H(t)\} \mathbf{D}_1 \mu^{(k-1)}(t), \quad \mu^{(k)}(0) = 0.$$

Itt a kiszolgálási idő eloszlásfüggvénye  $H(t)$ , valamint  $\mathbf{1}$  az 1-ekből álló  $M$  dimenziós oszlopvektor. A bevezetőben megjegyeztük, hogy a vizsgálatunk tárgyát képező rendszerrel [5]-ben egy általánosabb esetet tárgyalnak. [5]-ben a rendszer időfüggő momentum-generáló függvényét közvetlenül adják meg, míg mi a differenciálegyenlet-rendszert megoldva kapjuk meg momentumok függvényeit az időfüggő generátorfüggvény felhasználása nélkül.

Látható, hogy a (2) egyenlet megoldásából a magasabb momentumok iteratív módon számolhatóak a (3) egyenlet felhasználásával.

Bár nem tartozik ezen cikk szűken vett témakörébe, megjegyezzük, hogy faktoriális momentumok aszimptotikus viselkedésére a következő tételt mondják ki [6]-ban:

**1. Tétel:** *Ha az igények kiszolgálási idejének az átlaga véges ( $\mu = \int_0^\infty (1 - H(t)) dt < \infty$ ), akkor a  $\mu^{(k)}(t)$  vektor egy  $c_k e = (c_k, \dots, c_k)$  vektorhoz konvergál  $K \geq 1, t \rightarrow \infty$  esetén.*

A tételből látható tehát, hogy véges várható értékű kiszolgálási idő esetén a faktoriális momentumok  $t \rightarrow \infty$  határértéke független a kezdeti állapottól.

[6] cikk szerzői a (2) és (3) differenciálegyenletekre numerikus megoldást javasoltak. Megkaphatjuk azonban az egzakt megoldást is, ha feltételezzük, hogy az

igények kiszolgálási ideje PH eloszlású, vagyis felírható különböző paraméterű Erlang eloszlások kevertjeként. Azaz feltételezzük, hogy  $H(t)$  a következő alakban írható fel:

$$H(t) = 1 - \sum_{i=1}^I e^{-\beta_i t} \sum_{j=0}^{J_i} \alpha_{i,j} t^j, \quad (4)$$

ahol  $I \geq 1, J_i \geq 0, \beta_i \geq 0$ .

A megoldás az elsőrendű inhomogén lineáris differenciálegyenletek szokásos megoldási menetét követi (részletek: [7,8]). Az (1) egyenlettel definiált faktoriális momentumok felírhatóak az egyenlet homogén részének általános megoldása és az inhomogén rész partikuláris megoldása összegeként. Mivel a  $\mathbf{D}_0$  és  $\mathbf{D}_1$  mátrixok konstansok, a homogén rész megoldása az összes momentum esetében ugyanaz. A  $K$ -ik momentumra vonatkozó differenciálegyenlet inhomogén részének megoldása viszont a  $(K-1)$ -ik momentumtól függ ( $K \geq 2$ ). Először kiszámítjuk a megoldást  $\mu^{(1)}(t)$ -re, majd a bemutatott módszert alkalmazva kaphatjuk meg a megoldást a magasabb momentumokra.

A (2) egyenlet megoldásának első lépése a homogén rész megoldása:

$$\frac{d}{dt} \mu^{(1 \times H)}(t) = \mathbf{D} \mu^{(1 \times H)}(t). \quad (5)$$

A megoldás exponenciális függvények lineáris kombinációjaként adható meg:

$$\mu_n^{(1 \times H)}(t) = \sum_{m=1}^M [\mathbf{V}]_{n,m} c_m e^{r_m t}. \quad (6)$$

Itt az  $M \times M$ -es  $\mathbf{V}$  mátrix a  $\mathbf{D}$  mátrixot Jordan alakúra hozó hasonlósági transzformáció mátrixa (oszlopai az  $\mathbf{D}$  mátrix általánosított jobboldali sajátvektorai). Ha a  $\mathbf{D}$  mátrix minden sajátértéke egyszeres, akkor a  $\mathbf{V}$  mátrix oszlopai a  $\mathbf{D}$  mátrix sajátvektorai, az  $r_m$  számok a sajátvektorokhoz tartozó sajátértékek. A  $c_m$ -ek pedig meghatározandó ismeretlenek.

A  $\mathbf{D}$  mátrix sztochasztikusságából azonnal következik az alábbi tétel:

**2. Tétel:** *A  $\mathbf{D}$  mátrixnak a 0 az egyik sajátértéke és az összes sajátértékének a valós része nem pozitív.*

Mivel a 0 is egy sajátértéke  $\mathbf{D}$ -nek, így a homogén rész megoldásában van egy tag, amelyik nem függ  $t$ -től. Ez a tag marad meg  $t \rightarrow \infty$  esetén. Az általánosság megtartása mellett tegyük fel, hogy  $r_1 = 0$ .

A  $\mathbf{D}$  valós együtthatós mátrix komplex sajátértékei konjugált párokat alkotnak. Emiatt azonban a továbbiakban megjelenő egyenletek komplex együtthatói is konjugált párokba rendezhetőek, így a végeredményben a képzetes részek kiegyenlítik egymást. Ennek részletes ellenőrzésével ebben a cikkben nem foglalkozunk.

Abban az esetben, ha valamelyik sajátérték többszörös, akkor a  $\mu_n^{(1 \times H)}(t)$  függvény felírásában az  $e^{r_m t}$  együtthatói között  $t$  szerinti polinom függvények is megjelennek,

$$\mu_n^{(1 \times H)}(t) = \sum_m [\mathbf{V}]_{n,m} \sum_j c_{m,j} t^j e^{r_m t},$$

de a megoldási módszer lényegében változatlan marad.

A differenciálegyenlet-rendszer általános megoldásának a felírásához szükség van egy partikuláris megoldásra is. Ezt az állandók variálásának módszerével kaphatjuk meg, azaz tekintsük a homogén megoldás  $c_m$  együtthatóit  $t$  szerinti függvényeknek.

$$\sum_{m=1}^M [\mathbf{V}]_{n,m} \frac{d}{dt} c_m(t) e^{r_m t} = f_k(t), \quad (7)$$

ahol

$$f_k(t) = (1 - H(t)) \sum_{i=1}^M [\mathbf{D}_1]_{k,i}.$$

Az  $f_k(t)$  függvény általános formája a következő:

$$f_k(t) = \sum_{i=1}^{I^{(K)}} e^{-\gamma_{i,k} t} \sum_{j=0}^{J_i^{(K)}} \alpha_{i,j}^{(k \times K)} t^j. \quad (8)$$

Ha a (2) egyenlet partikuláris megoldását keressük, akkor az  $f_k(t)$  (8)-beli kifejezésében  $\gamma_{i,k} = \beta_i$  minden  $k$ -ra. Megjegyezzük, hogy a magasabb momentum függvények egyre több tagból állnak. Amint azt látni lehet a (3) egyenlet inhomogén részén, a  $\gamma_{i,k}$  exponensek és a  $\beta_j$  exponensek összegei adják meg a  $\gamma_{i,k+1}$  exponenseket.

A  $c_m(t)$  ( $1 \leq m \leq M$ ) függvények deriváltjait kifejezhetjük az  $f_k(t)$  függvényekkel és a  $\mathbf{V}$  mátrix inverzével. Jelölje  $\mathbf{A} = \mathbf{V}^{-1}$  a  $\mathbf{V}$  mátrix inverzét! Ekkor a (7) egyenletet megoldva  $\frac{d}{dt} c_m(t)$ -re kapjuk a következő kifejezést:

$$\begin{aligned} \frac{d}{dt} c_m(t) &= \sum_{k=1}^M [\mathbf{A}]_{m,k} f_k(t) e^{-r_m t} = \\ &= \sum_{k=1}^M [\mathbf{A}]_{m,k} \sum_{i=1}^{I^{(K)}} e^{-(\gamma_{i,k} + r_m)t} \sum_{j=0}^{J_i^{(K)}} \alpha_{i,j}^{(k \times K)} t^j, \quad 1 \leq m \leq M \end{aligned} \quad (9)$$

A  $c_m(t)$  függvényt a (9) egyenlet integrálásával kaphatjuk meg. Itt az integrálás során az összeg különböző  $t$ -től függő tagjait a következő séma szerint kaphatjuk meg:

$$\int e^{\delta t} t^j dt = -\frac{j!}{\delta^{j+1}} \sum_{l=0}^j \frac{e^{-\delta t} \delta^l t^l}{l!} + C, \quad \text{ha } \delta \neq 0,$$

és

$$\int t^j dt = \frac{t^{j+1}}{j+1} + C, \quad \text{ha } \delta = 0.$$

Egy partikuláris megoldásnál elegendő az integrálás során kapott konstans 0-nak venni, a  $c_m(t)$  függvényeket a következő alakban írhatjuk fel (10):

$$\hat{c}_m(t) e^{r_m t} = \sum_{k=1}^M \sum_{i=1}^{I^{(K)}} \sum_{j=0}^{J_i^{(K)}} [\mathbf{A}]_{m,k} \alpha_{i,j}^{(k \times K)} \frac{j!}{(\delta_{i,k}^{(m)})^{j+1}} \frac{e^{-\gamma_{i,k} t} (\delta_{i,k}^{(m)})^j}{l!}, \quad 1 \leq m \leq M.$$

itt  $\delta_{i,k}^{(m)} = \gamma_{i,k} + r_m$ . Ha valamely  $i$  és  $k$  indexekre  $\delta_{i,k}^{(m)} = 0$ , akkor az ehhez tartozó  $c_m(t)$  függvények a következő alakúak:

$$\bar{c}_m(t) e^{r_m t} = \sum_{k=1}^M \sum_{i=1}^{I^{(K)}} \sum_{j=0}^{J_i^{(K)}} [\mathbf{A}]_{m,k} \alpha_{i,j}^{(k \times K)} \frac{t^{j+1}}{j+1} e^{r_m t}. \quad (11)$$

Ekkor a partikuláris megoldás  $n$ -ik komponensét  $\hat{c}_m(t)$  és  $\bar{c}_m(t)$  függvényekből kaphatjuk meg:

$$\hat{\mu}_n^{(1 \times P)}(t) = \sum_{m=1}^M [\mathbf{V}]_{m,n} \hat{c}_m(t) e^{r_m t}, \quad (12)$$

$$\bar{\mu}_n^{(1 \times P)}(t) = \sum_{m=1}^M [\mathbf{V}]_{m,n} \bar{c}_m(t) e^{r_m t}, \quad (13)$$

és a teljes partikuláris megoldás pedig

$$\mu_n^{(1)(P)}(t) = \hat{\mu}_n^{(1)(P)}(t) + \bar{\mu}_n^{(1)(P)}(t), \quad (14)$$

ahol  $1 \leq n \leq M$ .

Megjegyezzük, hogy mind  $\hat{\mu}_n^{(1)(P)}(t)$ , mind  $\bar{\mu}_n^{(1)(P)}(t)$ , 0-hoz konvergál  $t \rightarrow \infty$  esetén.

Miután megkaptuk az első momentumra vonatkozó differenciálegyenlet homogén részének a megoldását ( $\mu_n^{(1)(H)}(t)$ ) és egy partikuláris megoldását ( $\mu_n^{(1)(P)}(t)$ ) is, így az általános megoldást szintén megkaptuk.

$$\mu_n^{(1)}(t) = \mu_n^{(1)(H)}(t) + \mu_n^{(1)(P)}(t). \quad (15)$$

A (6) egyenletben található  $c_m$  együtthatók meghatározásához a (2) egyenlet kezdeti feltételét használhatjuk, azaz

$$\mu_n^{(1)}(0) = \mu_n^{(1)(P)}(0) + \sum_{m=1}^M [\mathbf{V}]_{n,m} c_m = 0.$$

Jelölje  $c = [c_1, c_2, \dots, c_m]$  az együtthatók vektorát. Ekkor ismételtén az  $\mathbf{A} = \mathbf{V}^{-1}$  jelölést használva a konstansokat a

$$c = -\mathbf{A} \cdot \mu^{(1)(P)}(0) \quad (16)$$

egyenletből kaphatjuk meg.

Mivel az egyenletrendszer általános alakjában az  $(n+2)$ -dik momentumra vonatkozó egyenletet ugyanúgy vezettük vissza az  $(n+1)$ -dik momentum függvényére, mint ahogyan az  $(n+1)$ -dik momentumra vonatkozó egyenletet az  $n$ -dik momentum függvényére, a momentum függvényeket iteratív úton kaphatjuk meg a (3) egyenlet jobb oldalába való behelyettesítéssel.

Megjegyezzük, hogy a momentum függvények tulajdonképpen a faktoriális momentumokat adják meg. Az eredeti momentumokat a másodfajú Stirling számok ( $\Delta_{i,n}$ ) segítségével kaphatjuk meg:

$$k^n = \sum_{i=1}^n \Delta_{i,n} k(k-1)\dots(k-i), k \geq n.$$

$\Delta_{i,n}$  meghatározására pedig találhatunk egy rekurzív kifejezést [9] 5.2 fejezetében.

A gyakorlatban az iterációs lépések során csak szorzást és összeadást végzünk. Ekkor az  $(n+1)$ -ik momentum függvényben szereplő  $\gamma_{i,k}$  és  $\alpha_{i,j}^{(k)}$  konstansokat állítjuk elő az  $n$  momentum függvény ismeretében. Sajnos tetszőlegesen nagy momentum függvényeket nem tudunk meghatározni, mivel az iterációs lépések során a  $\gamma_{i,k}$  és  $\alpha_{i,j}^{(k)}$  konstansok száma nem korlátos.

## Irodalom

- [1] G. Latouche, V. Ramaswami: „Introduction to Matrix Analytic Methods in Stochastic Modeling”, ASA-SIAM Series on Statistics and Applied Probability, 1999.
- [2] W. Fischer, K. Meier-Hellstern: „The Markov-Modulated Poisson process (MMPP) cookbook”, Performance Evaluation 18, pp.149–171., 1993.
- [3] N. Hohn, D. Veitch, P. Abry: „Does fractal scaling at the IP level depend on TCP flow arrival processes?”, ACM SIGCOMM Internet Measurement Workshop (IMW-2002), Marseille, France, 2002.
- [4] A. Horváth, M. Telek: „Approximating heavy tailed behaviour with phase type distributions”, Advances in algorithmic methods for stochastic models, MAM3 (G. Latouche and P. Taylor, eds.), pp.191–214., Notable Publications Inc., 2000.
- [5] Hiroyuki Masuyama, Tetsuya Takine: „Analysis of an Infinite-Server Queue with Batch Markovian Arrival Streams”, Queueing Systems, vol. 42, No.3, pp.269–296., 2002.
- [6] V. Ramaswami, M. F. Neuts: „Some explicit formulas and computational methods for infinite-server queues with phase-type arrival”, Journal of Applied Probability 17, pp.498–514., 1980.
- [7] Ordinary Differential Equation System with Constant Coefficients, MathWorld, <http://mathworld.wolfram.com/>
- [8] I. N. Bronstein, K. A. Semendjajew, G. Musion, H. Mühlig: Taschenbuch der Mathematik, Verlag Harri Deutsch, Frankfurt am Main, 1999.
- [9] D. J. Daley, D. Vere-Jones: „An Introduction to the Theory of Point Processes, Volume I: Elementary Theory and Methods”, Second Edition, Springer, 2003.
- [10] Rózsa Pál: „Lineáris algebra és alkalmazásai”, Műszaki Könyvkiadó, 1974.

