Glevitzky Béla

Operációkutatás II.

mobiDIÁK könyvtár

Glevitzky Béla

Operációkutatás II.

mobiDIÁK könyvtár

SOROZATSZERKESZTŐ

Fazekas István

Glevitzky Béla

Operációkutatás II.

mobiDIÁK könyvtár
Debreceni Egyetem
Informatikai Intézet

mobiDIÁK könyvtár
Debreceni Egyetem
Informatikai Intézet
4010 Debrecen, Pf. 12
http://mobidiak.unideb.hu

A mű a *A mobiDIÁK önszervező mobil portál* (IKTA, OMFB-00373/2003) és a *GNU Iterátor, a legújabb generációs portál szoftver* (ITEM, 50/2003) projektek keretében készült.

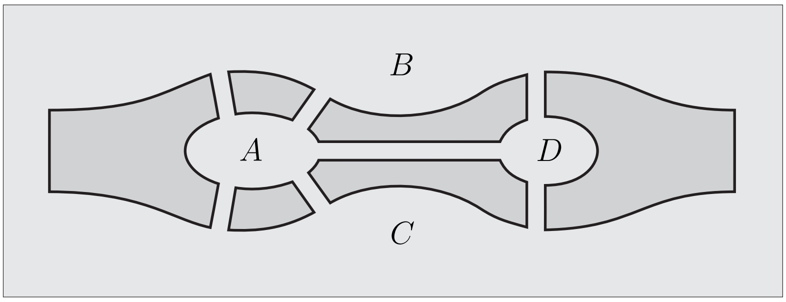
Tartalomjegyzék

I. Hálózati folyamok elmélete

Bevezetés

Bizonyos rendszerek működése nemcsak a rendszert alkotó elemektől, hanem azok egymáshoz viszonyított helyzetétől is függhet. Mindezek miatt ezeket a rendszereket célszerű úgy ábrázolni, hogy az alkotó elemek egymáshoz viszonyított helyzete is áttekinthető legyen. Az ábrában a rendszert alkotó elemeket pontokkal, a közöttük lévő kapcsolatokat pedig szakaszokkal (irányított szakaszokkal) szemléltethetjük. Egy ilyen ábrát általában gráfnak szokás nevezni, bár az egyes alkalmazási területeken más elnevezések is használatosak (például hálózat, diagram, struktúra, stb).

A gráfelmélet a matematikának egyik fiatal, gyorsan fejlődő ága. Az első gráfelméleti problémák az 1700-as években vetődtek fel. Első ismert eredmény Euler-től származik (1736), aki a königsbergi hidak problémájának megoldásához vezette be a gráf fogalmát. Königsberg egy folyó két partján és két szigeten terült el. A városrészeket hidak kötötték össze. Felvetődött az a kérdés, hogy valaki a lakásából elindulva tehet-e egy olyan sétát, amelynek során minden városrészt érint úgy, hogy minden hídon pontosan egyszer megy át és a séta végén hazaérkezik.



Ezen problémában a hidakon való egyszeri áthaladás, az egyes városrészek érintése, valamint a kiinduló pontba való visszatérés a leglényegesebb. Euler olyan ábrát készített, melyben a városrészeknek pontok, a hidaknak a pontokat összekötő vonalak felelnek meg. A kérdés ezek után úgy fogalmazható meg, hogy valamelyik csúcsból kiindulva bejárhatjuk-e a gráf éleit úgy, hogy minden élen pontosan egyszer haladjunk át és végül a kiindulási pontba érjünk vissza. Egy ilyen tulajdonságú útvonalnak meg tudnánk-e adni valamilyen jellemzőjét? Ha valaki egy ilyen útvonalat jár be, akkor ha valamelyik csúcsba eljut, akkor onnan ki is kell lépnie, s ezután más útvonalszakaszon kell folytatnia a sétáját. Így tehát minden pontba (a kezdőpontba is) a sétálónak ugyanannyiszor kell belépnie, mint kilépnie, és ezeket különböző utakon kell megtennie.

Ezért a feladat megoldhatóságához szükséges, hogy minden pontban páros számú él találkozzon. Ez a königsbergi hidak problémájában nem teljesül, ezért a feladat nem oldható meg.

Kétszáz év múltán 1936-ban jelent meg König Dénes munkája, amely ezen tudományág első összefoglaló művének tekinthető.

Egyszerűsége és jól kezelhetősége miatt alkalmazási köre meglehetősen tágas, szerepe a matematikán kívül – többek között – jelentős a műszaki és természettudományokban is.

A gráfokat a gyakorlati élet számos feladatának megoldásánál fel lehet használni. A konkrét feladatokat gráf alakjában írva az áttekinthetőség jelentősen növelhető; a gráfok belső szerkezetére vonatkozó matematikai tételek az adott feladat mély elemzését, fontos tulajdonságok felismerését teszik lehetővé.

I.1. Gráfelméleti alapfogalmak

Gyakran megesik, hogy a mindennapi élet különböző dolgait alkalmas módon sematizálva, azok olyan hasonlóságot mutatnak, amelyek ugyanazt az elméleti fogalmat sugallják. Például egy vasúti hálózat, vagy egy város utcahálózatának térképe. Ezeknek a sémáknak a mindennapi jelentése esetenként teljesen különböző, de van egy olyan formális jellegzetességük, amellyel mindegyikük rendelkezik. Minden esetben csomópontokat, kereszteződéseket, vagy állomásokat találunk, amelyeket pontokkal (csúcsokkal) és az azokat összekötő vonalakkal (élekkel) ábrázolhatjuk. Ez a gráf intuitív fogalma.

Egy irányítatlan gráf a csúcsok N és az élek A halmazából, továbbá az A halmazt az N nem rendezett párjai halmazába átvivő valamilyen F leképezésből áll. Amennyiben ez az F leképezés az A halmazt az N rendszerezett párjai halmazába viszi át, irányított gráfról beszélünk, amit szokás digráfnak is nevezni. Egy digráf jelölésére az $[N, A]$ szimbólumot szokás használni.

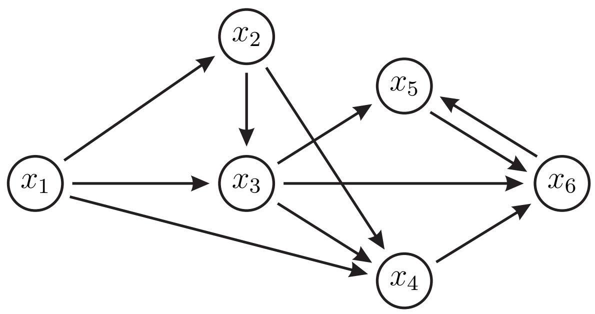
Egy gráf minden élét két végpontja jellemez. Ha például egy x csúcsot él köti össze az y csúccsal, akkor ezt az élt (x, y)-nal jelölhetjük. Irányíthatatlan gáfok esetében az (x, y) és az (y, x) élek ugyanazt jelentik, irányított esetben azonban ezek már különbözők. Előfordulhat olyan él is, amelynek mindkét végpontja ugyanaz a pont. Az ilyen élt *hurokél*nek nevezzük. Két csúcs között több élt is húzhatunk, az ilyen élt *többszörös él*nek nevezzük. Egy csúcsból kiinduló élek számát a csúcs fokszámának nevezzük, amely természetesen 0 is lehet. Ilyenkor a csúcsot *izolált pont*nak nevezzük. A továbbiakban csak olyan gráfokat tekintünk, amelyekben nincs hurokél és nincs többszörös él. Az ilyen gráfokat *egyszerű gráf* oknak nevezzük.

Tekintsük az $[N, A]$ digráfot. Ezt megadhatjuk úgy, hogy felsoroljuk az N ponthalmaz és az A élhalmaz elemeit, vagy ábrát készítünk a digráfról, vagy az úgynevezett *struktúramátrix*ot használjuk. Ez utóbbi egy $N \times N$-es mátrix, amelyben olyan helyen szerepel jelölés, amely a sorjelző csúcsból az oszlopjelző csúcsba vezető élnek felel meg. Minthogy ebben a mátrixban általában sok jelöletlen hely van, szokás még a digráfot megadni az úgynevezett „Honnan-hová” táblázattal is, amelynek a „honnan” sora azon csúcsok azonosítóit tartalmazza, amelyekből élek indulnak ki, a „hová” sora pedig azokat a csúcsokat, amelyekbe befutnak az élek. Ezen táblázattal együtt meg kell adni az N csúcshalmazt is, hiszen a gráfban lehetnek izolált pontok is.

Példa. Legyen

$$N = \{x_1, x_2, x_3, x_4, x_5, x_6\}$$
$$A = \{(x_1, x_2), (x_1, x_3), (x_1, x_4), (x_2, x_3), (x_2, x_4), (x_3, x_4),$$
$$(x_3, x_5), (x_3, x_6), (x_4, x_6), (x_5, x_6), (x_6, x_5)\}.$$

Most tehát megadtuk a gráf csúcsainak N és éleinek A halmazát. Ezen digráfot a következő ábrával is megadhatjuk:



Ennek a digráfnak megfelelő struktúramátrix a következő lesz:

	x_1	x_2	x_3	x_4	x_5	x_6
x_1		X	X	X		
x_2			X	X		
x_3				X	X	X
x_4						X
x_5						X
x_6					X	

Végül ezen digráfot megadhatjuk a „Honnan-hová” táblázattal is:

$$N = \{x_1, x_2, x_3, x_4, x_5, x_6\}$$

Honnan	x_1	x_1	x_1	x_2	x_2	x_3	x_3	x_3	x_4	x_5	x_6
Hová	x_2	x_3	x_4	x_3	x_4	x_4	x_5	x_6	x_6	x_6	x_5

Ha az N pontból álló digráf minden élt tartalmaz, akkor a digráfot *teljes*nek nevezzük és $[N]$-nel jelöljük. Egy teljes digráf éleinek száma (minthogy hurokél nem lehet a digráfban) $N \cdot (N-1)$.

I.2. Út és vágás a digráfban. Cimkézési technika

Legyen $[N, A]$ egy digráf és legyen az N ponthalmaz az S és S' diszjunkt, nemüres részhalmazokra bontva. Jelöljük (S, S')-vel azon élek összességét, amelyek S-ből indulnak és S'-be érkeznek. Ekkor az (S, S') élhalmazt az $[N, A]$ digráf vágásának nevezzük.

Legyenek $s, s' \in N$. Ha $s \in S$ és $s' \in S'$, akkor azt mondjuk, hogy az (S, S') vágás az s és s' pontokat elválasztja.

Legyen $[N, A]$ egy digráf és legyenek $s, s' \in N$. Legyen az $x_0, x_1, \ldots, x_m \in N$ egy olyan pontsorozat, amelyre $x_0 = s, x_m = s'$ és $(x_{i-1}, x_i) \in A$ teljesül minden $i = 1, 2, \ldots, m$ esetén. Ekkor azt mondjuk, hogy az $x_0, x_1, x_2, \ldots, x_m$ egy s-ből s'-be vezető út. Ennek jelölésére a

$$P = \{s = x_0, x_1, \ldots, x_m = s'\}$$

szimbólumot szokás használni.

Egy vágást üresnek nevezünk, ha az (S, S') élhalmaz üres, azaz nincs olyan él, amelyik S-ből indul és S'-be érkezik. (Az élek irányítottsága miatt ekkor az (S', S) vágás természetesen nem szükségképpen lesz üres, sőt általában nem üres, ellenkező esetben a gráf olyan részgráfokból álló lenne, amelyek egymással egyetlen éllel sincsenek összekötve.)

Egy út létezése illetve az üres vágás közötti kapcsolatot megadó tétel Minty-től származik, és a következőképpen fogalmazható meg:

Tétel. *Legyen $[N, A]$ egy digráf és legyenek $s, s' \in N$. Ekkor vagy létezik ebben a digráfban s-ből s'-be vezető út, vagy pedig létezik egy olyan üres vágás, amely s-et és s'-t elválasztja.*

Bizonyítás. Ezen tétel bizonyítása kostruktív lesz, azaz vagy előállítjuk az s-ből s'-be vezető utat, vagy pedig meghatározzuk azt az (S, S') üres vágást, amely egyúttal az s-et és s'-t is elválasztja. Az S halmaz előállítására szolgáló eljárást (technikát) cimkézési technikának nevezzük.

Rendeljünk most az $[N, A]$ digráf minden pontjához egy, a pontot azonosító címet (például a pont sorszámát, vagy a jelét) és egy üres „rekeszt”, amelybe majd az eljárás során beírjuk az adott pont „cimkéjét”. Kiinduláskor legyen $s \in S$ és lássuk el ezt a kezdőpontot a $-s$ cimkével. Ezt követően legyen x_j eleme S-nek, ha $(s, x_j) \in A$. Az s cimkéjét változtassuk pozitívra, majd az S-be tartozó x_j-knek adjuk a $(-s)$-et cimkeként. Amennyiben $s' \in S$, akkor vége van az eljárásnak. Különben tekintsünk egy negatív cimkéjű S-beli x_i pontot, a cimkéjét változtassuk pozitívra és tekintsük S-belinek mindazon $x_j \in N$ csúcsokat, amelyekre $(x_i, x_j) \in A$. Ezen x_j-k cimkeként kapják meg a $(-x_i)$-t, azaz az x_i azonosítójának negatívját. Ha egy x_j-nek már volt cimkéje, akkor annak már nem kell újabbat adnunk. Ha $s' \in S$, akkor az eljárásnak vége, különben az előzőekhez hasonlóan folytassuk. Ezt mindaddig tehetjük, amíg vagy $s' \in S$ adódik, vagy pedig minden S-beli pont cimkéje már pozitív. Ez utóbbi esetben a cimkézett csúcsok lesznek azok, amelyeket s-ből útvonal mentén elérhetünk. Az $S' = N \setminus S$ definícióval megkapjuk az S' halmazt, amelyre igaz, hogy $s' \in S'$. Az (S, S') vágás üres és elválasztja az s és s' pontokat. □

Ha $s' \in S$, akkor tehát vezet s-ből s'-be útvonal. Az s' pont cimkéje alapján meghatározhatjuk, hogy él mentén melyik csúcsból értük el s'-t.

Ezen csúcs cimkéje viszont azt mutatja meg, hogy melyik másik csúcsból juthattunk el él mentén ezen csúcsba. Ha ezt így folytatjuk tovább, akkor eljuthatunk az s pontig és így rekonstruálhatjuk az s-ből s'-be vezető utat.

A vágást a struktúra mátrixon is szemléltethetjük. Egy-egy vonallal fedjük le a struktúramátrix azon oszlopait, amelyek S-beliek, és azon sorait, amelyek S'-beli ponthoz tartoznak. A fedetlen helyen az (S, S')-beli élek találhatók. Üres vágás esetén természetesen nem lesz ilyen fedetlen él.

Belátható, hogy a kétszer fedett helyeken az (S', S) vágásbeli élek találhatók. Az egyszer fedett helyeken vagy (S, S)-beli, vagy pedig (S', S')-beli élek vannak.

I.3. Minimális út, maximális potenciál

Legyen adott egy $[N, A]$ digráf. Ezen digráf minden $(x, y) \in A$ éléhez rendeljünk hozzá egy $\tau(x, y) \geq 0$ egész számot. Egy ilyen konfigurációt hálózatnak nevezünk és $[N, A, \tau]$-val jelölünk.

Legyen most $[N, A, \tau]$ egy hálózat, amelynek legyen s és s' két kitüntetett pontja. Meghatározandó az s-ből s'-be vezető olyan

$$P = \{s = x_0, x_1, \ldots, x_m = s'\}$$

út, amely mentén a

$$\tau(P) = \sum_{j=1}^{m} \tau(x_{j-1}, x_j)$$

útvonalhossz minimális. Ezt a problémát minimális út problémának nevezzük.

Ez esetben tehát nem az a kérdés, hogy vezet-e útvonal az s pontból az s' pontba, hanem, hogy az s-et s'-vel összekötő útvonalak közül melyik a legrövidebb. Ha például a $\tau(x, y)$ az $(x, y) \in P$ él mentén a szállítási költséget jelenti, akkor az előbbi probléma olyan útvonal meghatározását jelenti, amelynél az össz-szállítási költség minimális és az útvonal s-ből s'-be vezet.

Minden hálózattal kapcsolatban megfogalmazható egy másik probléma is: a hálózat minden $x \in N$ csúcspontjához meghatározandó egy $\mu(x) \geq 0$ egész szám úgy, hogy a

$$\begin{aligned} \mu(s) &= 0 \\ \mu(y) - \mu(x) &\leq \tau(x, y) \qquad \forall (x, y) \in A \end{aligned}$$

feltételek teljesüljenek és a $\mu(s')$ maximális legyen. Ezt a $\mu(x)$ egész számot potenciálnak, a feladatot pedig maximális potenciál feladatnak nevezzük.

Ezt a problémát a következőképpen értelmezhetjük az előbbi minimális összköltségű útvonallal kapcsolatban. Egy vállalkozó felajánlja, hogy a szállítást elvégzi. Minden $x \in N$ csúcspontra megad egy $\mu(x)$ értéket, amely kifejezi, hogy ő mennyiért hajlandó s-ből x-be szállítani. Ezen $\mu(x)$ ajánlatnak olyannak kell lennie, hogy az elfogadható legyen, különben a termelő nem veszi igénybe a vállalkozót a szállításhoz. Így tehát ha $x = s$, akkor

$\mu(s) = 0$. Továbbá, ha a vállalkozó x-be szállít, majd utána az (x, y) él mentén y-ba, akkor az y-ba történő vállalkozói szállítás esetén fenn kell állnia a

$$\mu(y) \leq \mu(x) + \tau(x, y)$$

egyenlőtlenségnek. Ennek az útvonal mentén lévő minden csúcs esetében teljesülnie kell. Ilyen feltételek mellett a válalkozó célja olyan $\mu(x), x \in N$ árajánlat megadása, amely a legnagyobb bevételt biztosítja a számára, azaz amelyre $\mu(s')$ a lehető legnagyobb.

Ezen két feladat között szoros kapcsolat van, amit a következőképpen fogalmazhatunk meg:

Tetszőleges s-ből s'-be vezető P út és tetszőleges, a feltételeket kielégítő $\mu(x)$ potenciálrendszer esetén a $\tau(P)$ úthossz nem lehet kisebb, mint a $\mu(s')$ potenciálérték, azaz

$$\tau(P) \geq \mu(s').$$

Ezt az állítást könnyen beláthatjuk, ha felhasználjuk a

$$\mu(y) - \mu(x) \leq \tau(x, y) \qquad \forall (x, y) \in A$$

egyenlőtlenségeket:

$$\tau(P) = \sum_{j=1}^{m} \tau(x_{j-1}, x_j) \geq \sum_{j=1}^{m} \big[\mu(x_j) - \mu(x_{j-1})\big] = \mu(s').$$

Ezt az állítást felhasználva, ha létezik s-ből s'-be vezető olyan P út, valamint egy olyan μ potenciálrendszer, hogy $\tau(P) = \mu(s')$ teljesül, akkor az úthossz minimális értékű, a $\mu(s')$ potenciál értéke pedig maximális, azaz az út és a potenciálrendszer optimális.

A következő, Ford-tól származó tétel az optimális úttal, illetve potenciálrendszerrel kapcsolatban fogalmaz meg egy állítást, amelyet konstruktív módon fogunk bizonyítani.

Ford tétele:

Tétel. *Ha van s-et s'-vel összekötő út, akkor létezik olyan P út és $\mu(x), x \in N$ potenciálrendszer, hogy*

$$\min_P \tau(P) = \max_{\mu(x)} \mu(s').$$

Bizonyítás. A bizonyítás kostruktív jellegű, amely az optimális megoldás meghatározásának algoritmusát is szolgáltatja.

Kiinduláskor legyen $\mu(x) = 0, \forall x \in N$ esetén.

Nyilvánvaló, hogy ez a potenciálrendszer eleget tesz a maximális potenciál probléma feltételrendszerének. Ezek után a hálózat éleiből konstruáljunk egy $[N, \beta]$ digráfot, amelynek β élhalmaza azon $(x, y) \in A$ élekből fog állni, amelyekre a $\tau(x,y)=\mu(y)-\mu(x)$ egyenlőség teljesül. (Minthogy általában igaz, hogy minden $(x,y)\in A$ él esetében $\tau(x,y)>0$, ezért kiinduláskor általában $\beta=\emptyset$.)

Ezt követően az $[N,\beta]$ digráfban keressünk egy s-ből s'-be vezető utat. Minty tétele értelmében két eset lehetséges:

- Van út s-ből s'-be az $[N,\beta]$ digráfban. Minthogy ezen út minden élére $\tau(x,y)=\mu(y)-\mu(x)$ teljesül, ezért ez az út lesz a minimális útvonal, a $\mu(s')$ pedig a maximális potenciál értéke.
- Nincs s-ből s'-be vezető útvonal az $[N,\beta]$ digráfban. Az $[N,\beta]$ digráfban az (S,S') vágás ekkor üres lesz és el fogja választani az s és s' pontokat. Természetesen nem lesz üres az $[N,A]$ digráfban az (S,S') vágás, hiszen feltettük, hogy vezet s-ből s'-be útvonal. Minthogy az $[N,\beta]$ digráfbeli (S,S') vágás üres, de nem üres $[N,A]$-ban, ezért az (S,S') vágásban lévő A-beli (x,y) élekre igaz, hogy

$$\tau(x,y)>\mu(y)-\mu(x), \qquad \text{ahol } x\in S, y\in S'.$$

Képezzük a vágásban lévő éleken az

$$\varepsilon=\min_{\substack{(x,y)\,\in\,(S,S')\\(x,y)\,\in\,A}}\{\tau(x,y)+\mu(x)-\mu(y)\}>0$$

értéket. Ezt felhasználva a következőképpen definiálhatjuk a $\overline{\mu}(x)$ új potenciálrendszert:

$$\overline{\mu}(x)=\begin{cases}\mu(x), & \text{ha } x\in S\\ \mu(x)+\varepsilon, & \text{ha } x\in S'.\end{cases}$$

Ez az új potenciálrendszer is eleget tesz a feltételeknek, továbbá a

$$\overline{\mu}(s')=\mu(s')+\varepsilon>\mu(s')$$

miatt jobb, mint a korábbi. Ezt követően az új potenciálrendszerre megismételhetjük az eddigi lépéseket. Mivel $\varepsilon>0$ egész szám és a $\mu(s')$ felülről korlátos, az eljárás véges sok lépésben véget ér, tehát a potenciálrendszer fokozatos javításával eljuthatunk egy optimális megoldáshoz.

$\square$

A bizonyításból kiderül, hogy ezen probléma megoldása útkeresések sorozatából áll. Olyan éleken keresünk utat, amelyeken $\tau(x,y)+\mu(x)-\mu(y)=0$ teljesül. (Az ilyen éleket az $[N,\beta]$ digráf tartalmazza.) Mindezek miatt az előbbi algoritmus megszervezhető úgy is, hogy nem számítjuk ki a $\mu(x), x\in$

N potenciálrendszert, hanem az élekre számított $\varepsilon(x, y)$ értékekkel dolgozunk, amelyeket az

$$\varepsilon(x, y) = \tau(x, y) + \mu(x) - \mu(y)$$

egyenlőséggel definiálunk. Ekkor a hálózat $\varepsilon(x, y) = 0$ tipusú élein kell utat keresnünk s-ből s'-be. Az ε kiszámítása is könnyen megvalósítható a struktúramátrix sorainak és oszlopainak korábban ismertetett lefedésével. A fedetlen értékek minimuma lesz az ε értéke és az új $\overline{\varepsilon}(x, y)$-ra a következő adódik:

$$\overline{\varepsilon}(x, y) = \begin{cases} \varepsilon(x, y), & \text{ha } (x, y) \text{ egyszer fedett} \\ \varepsilon(x, y) + \varepsilon, & \text{ha } (x, y) \text{ kétszeresen fedett} \\ \varepsilon(x, y) - \varepsilon, & \text{ha } (x, y) \text{ fedetlen.} \end{cases}$$

Minthogy kiinduláskor a $\mu(x) = 0, \forall x \in N$, potenciálrendszert választottuk, ezért induláskor

$$\varepsilon(x, y) = \tau(x, y)$$

teljesül minden $(x, y) \in A$ él mentén. Az eljárás befejeztével a $\mu(x)$ potenciálok optimális értékét az

$$\begin{aligned} \varepsilon(x, y) &= \tau(x, y) + \mu(x) - \mu(y) \\ \mu(s) &= 0 \end{aligned}$$

egyenletek segítségével határozhatjuk meg, ahol célszerű az $\varepsilon(x, y) = 0$ tipusú éleket figyelembe venni.

I.4. Időtervezési feladatok: CPM, PERT

Egy valamilyen termelési folyamat, vagy egy gépsor beállítása, szétszedése, javítása, összeszerelése, stb. egymás után és egymással párhuzamosan végezhető tevékenységekből áll. A tevékenységek logikai kapcsolatát egy speciális digráffal lehet szemléltetni. Ezen digráf élei a tevékenységeket, a csúcspontok pedig az eseményeket jelölik. Ezek az események bizonyos tevékenységek befejezését, illetve belépő tevékenységek kezdését jelölik. Egy ilyen digráfot tervütemhálónak fogunk nevezni, amelynek pontos definiciója a következő:

Definíció. *A tervütemháló olyan digráf, amelynek létezik $s \in N$ kezdőpontja és $s' \in N$ végpontja úgy, hogy bármelyik $x \in N, x \neq s, x \neq s'$ esetén vezet út s-ből x-be és x-ből s'-be; továbbá körútmentes (hurokmentes), azaz nincs önmagába visszatérő út.*

Az eseményeket (csomópontokat), amelyek a tevékenységeket fűzik össze, sorszámmal szokás ellátni, ezért a tevékenységeket a kezdő eseménye és a befejező eseménye sorszámával jelöljük. Így például az x sorszámú eseményt az y sorszámú eseménnyel az (x, y) él köti össze és ehhez a tevékenységhez a $\tau(x, y)$ tevékenységidő tartozik. Az események sorszámozását minden tervütemháló esetén mindig el lehet úgy végezni, hogy minden tevékenységre igaz, hogy a kezdő események sorszáma kisebb, mint a befejezőé.

Egy komplex munkafolyamat megtervezésénél egyik fontos kérdés, hogy mennyi annak az „átfutási" ideje, azaz mennyi idő alatt teljesíthető az összes, a munkafolyamathoz tartozó tevékenység. Látni fogjuk, hogy erre a kérdésre a választ a komplex munkafolyamatot modellező tervütemhálóban lévő leghosszabb útvonal hossza adja. Egy másik fontos kérdés az, hogy hogyan kell ütemezni az egyes tevékenységek elvégzését. Ezen szempontból fontos mozzanat az egyes események bekövetkezése, hiszen ez a feltétele annak, hogy egy adott eseményből kiinduló tevékenységek elkezdődhessenek. Az egyes események bekövetkezését eseményidővel fogjuk jellemezni, és egy x sorszámú csomópont esetében $\mu(x)$-szel fogjuk jelölni.

Ezek után a feladat matematikai modellje a következő lesz:

Meghatározandó az $[N, A, \tau]$ tervütemhálóban az s-ből s'-be vezető

$$P = \{s = x_0, x_1, x_2, \dots, x_m = s'\}$$

utak közül az, amelynél a

$$\tau(P) = \sum_{i=1}^{m} \tau(x_{i-1}, x_i)$$

érték (útvonalhossz) maximális. Ezt a maximális utat kritikus útnak nevezzük. A kritikus út elnevezést azért kapta, mert a hozzá tartozó tevékenységek elvégzési idejei szabják meg az egész komplex munkafolyamat elvégzési idejét. Ha ezekben bármilyen „csúszás" következik be, akkor az az egész munkafolyamat átfutási idejére is kihat. Természetesen az is igaz, hogy ha ezen tevékenységek elvégzési idejét lerövidítjük, akkor az egész munkafolyamat átfutási ideje is lerövidülhet. (Bár nem kötelezően, hiszen több kritikus útvonal is létezhet a tervütemhálóban, illetve valamely kritikus tevékenység elvégzési idejének lerövidítésével más tevékenység válhat kritikussá.)

Az előbbi modellben az átfutási időt a $\tau(P)$ jelenti. Az előbbi modellt a probléma primál megfogalmazásának nevezzük. Megfogalmazható ugyanakkor a probléma duál feladata is:

Meghatározandó minden $x \in N$ ponthoz egy $\mu(x)$ eseményidő, amelyre a

$$\mu(y) - \mu(x) \geq \tau(x, y) \qquad \forall (x, y) \in A$$

feltételek fennállnak és a

$$\mu(s') - \mu(s)$$

célfüggvényérték minimális.

A $\mu(x), x \in N$ eseményidők együttesét ütemezésnek, vagy időpolitikának, a $\mu(s') - \mu(s)$ értéket pedig az ütemezés értékének nevezzük.

A primál és duál feladatok kapcsolatát a következő állításban fogalmazhatjuk meg:

Tetszőleges s-ből s'-be vezető P út és a $\mu(x)$ ütemezés esetén a célfüggvények értékei között a

$$\tau(P) \leq \mu(s') - \mu(s)$$

egyenlőtlenség áll fenn.

Ezen egyenlőtlenség következménye, hogy ha az s-ből s'-be vezető P út és a $\mu(x)$ ütemezés olyan, hogy a $\tau(P) = \mu(s') - \mu(s)$ egyenlőség teljesül, akkor a P út maximális hosszúságú, az ütemezés pedig minimális.

Könnyen belátható, hogy az előbbi egyenlőség pontosan akkor áll fenn, ha az út minden élén

$$\tau(x, y) = \mu(y) - \mu(x),$$

azaz az eseményidők különbsége egyenlő a tevékenységek idejével.

Ezek után a kérdés úgy fogalmazható meg, hogy egy adott problémát modellező tervütemháló esetén létezik-e olyan P út és $\mu(x)$ ütemezés, amelyek esetén teljesül az előbbi egyenlőség. A választ erre a kérdésre a következő tétel fogalmazza meg.

Tétel. *Egy adott tervütemháló esetén létezik olyan s-ből s'-be vezető P út és $\mu(x)$ ütemezés, hogy ezekkel fennáll a*

$$\tau(P) = \mu(s') - \mu(s)$$

egyenlőség.

Bizonyítás. Ezen tétel bizonyítása konstruktív jellegű lesz, azaz előállítjuk azt az s-ből s'-be vezető P útvonalat és meghatározzuk azt a $\mu(x)$ ütemezést, amely mellett fennáll az egyenlőség.

Legyen az S halmaz úgy megkonstruálva, hogy

- tartalmazza az s kezdő eseményt, azaz $s \in S$ teljesüljön
- minden S-beli ponthoz vezessen úgynevezett kritikus út a kezdő eseményből.

Az S halmaz az események sorszámozása miatt az

$$S = \{x_i | \ i = 0, 1, \ldots, k-1\}$$

legyen, azaz álljon a tervütemháló első k pontjából. Jelölje az S-beli pontokhoz tartozó eseményidőket a $\underline{\mu}(x_i), i = 0, 1, \ldots, k-1$. A $\underline{\mu}(x_i)$ az az adott időpont, aminél korábban nem kezdhetők az x_i-ből kiinduló tevékenységek. Ezt a $\underline{\mu}(x_i)$ időt a legkorábbi kezdési időnek nevezzük. Ennél korábban azért nem kezdhetjük el a tevékenységet, mert x_i-hez vezet kritikus út, ami azt jelenti, hogy az x_i-be vezető valamely tevékenység éppen $\underline{\mu}(x_i)$-ben fejeződik be.

Ha $S = N$, akkor ez azt jelenti, hogy az s' befejező eseményhez is vezet kritikus út, azaz megtaláltuk az optimális megoldást.

Amennyiben $S \neq N$, akkor tekintsük a soron következő x_k pontot. Ahhoz, hogy x_k is S-be tartozzon, kritikus úton kell azt s-ből elérni, amihez a $\underline{\mu}(x_k)$ eseményidőt a

$$\underline{\mu}(x_k) = \max_{\substack{0 \le i \le k-1 \\ (x_i, x_k) \in A}} \{\underline{\mu}(x_i) + \tau(x_i, x_k)\}$$

definiálja, ahol a maximumképzést azon élekre kell elvégezni, amelyeknek x_k a végpontja.

Ezt az eljárást mindaddig folytatjuk, amíg az $S = N$ egyenlőséghez el nem jutunk.

A kritikus útvonal és az ütemezés meghatározására egy másik eljárást is megadhatunk. Ekkor kiinduláskor egy olyan S' halmazt konstruálunk meg, amely tartalmazza az s' pontot, azaz amelyre $s' \in S'$ teljesül, továbbá megköveteljük, hogy minden S'-beli pontból vezessen kritikus út s'-be.

Legyen $S' = \{x_i |\ i = m+1, \dots, n\}$. Jelölje az S'-beli pontokhoz tartozó eseményidőket $\overline{\mu}(x_i)$. Ez az az idő, aminél később már nem fejeződhetnek be az x_i pontba befutó tevékenységek, mivel az x_i-ből kivezető tevékenységek közül valamelyik éppen $\overline{\mu}(x_i)$-ben kezdődik el. A $\overline{\mu}(x_i)$ időt a legkésőbbi befejezési időnek nevezzük.

Ha $S' = N$, akkor már vezet s-ből s'-be kritikus út. Ha $S' \neq N$, akkor tekintsük a soron következő x_m pontot. Ez az x_m pont akkor tartozik S'-be, ha tőle is vezet s'-be kritikus út. Emiatt a $\overline{\mu}(x_m)$ eseményidőt a

$$\overline{\mu}(x_m) = \min_{\substack{m+1 \le i \le n \\ (x_m, x_i) \in A}} \{\overline{\mu}(x_i) - \tau(x_m, x_i)\}$$

összefüggésből számíthatjuk ki. Ezt az eljárást is addig folytatjuk, amíg az $S' = N$ egyenlőséghez el nem jutunk.

A két eljárás során célszerű a $\underline{\mu}(s) = 0$ és a $\overline{\mu}(s') = \underline{\mu}(s')$ kezdőértékkel számolni. Nyilvánvaló, hogy akár az S, akár az S' halmazt konstruáljuk is meg, mindig véges sok lépésben jutunk el az optimális megoldáshoz.

Ha rendelkezésünkre állnak a legkorábbi kezdési $\underline{\mu}(x)$, illetve a legkésőbbi befejezési $\overline{\mu}(x)$ idők, akkor a $\mu(x)$ eseményidőkre a $\underline{\mu}(x) \le \mu(x) \le \overline{\mu}(x), \forall x \in N$ egyenlőtlenségek teljesülnek.

Azokat az eseményeket, amelyekre $\underline{\mu}(x) = \overline{\mu}(x)$ teljesül, kritikus eseményeknek nevezzük. A kritikus utat a kritikus eseményeket összekötő tevékenységek alkotják, amely tevékenységeket kritikus tevékenységeknek nevezzük. □

Egy komplex munkafolyamat kivitelezésének megtervezéséhez és irányításához az előbbieken kívül még ismernünk kell a tevékenységek ütemezését is. Ezt az eseményidőkből a következőképpen határozhatjuk meg:

- Az (x, y) tevékenység legkorábbi kezdési időpontja $\underline{\mu}(x)$.
- Az (x, y) tevékenység legkorábbi befejezési időpontja a $\underline{\mu}(x)+\tau(x, y)$.
- Az (x, y) tevékenység legkésőbbi kezdési időpontja a $\overline{\mu}(y) - \tau(x, y)$.
- Az (x, y) tevékenység legkésőbbi befejezési időpontja $\overline{\mu}(y)$.

Nyilvánvaló, hogy a kritikus tevékenységeknél a legkorábbi és a legkésőbbi kezdési (befejezési) időpontok egybeesnek. Ugyanakkor a többi tevékenység legkorábbi és legkésőbbi kezdési illetve befejezési időpontjai már eltérnek egymástól, így ezek a tevékenységek bizonyos tartalékidőkkel rendelkeznek. A tartalékidők fajtái a következők lesznek:

- Teljes (összes) tartalékidő:

$$\overline{\mu}(y) - \underline{\mu}(x) - \tau(x, y)$$

  Ez a legkedvezőbb körülmények között számított érték, amikoris az (x, y) tevékenység a legkorábban kezdődhet és a legkésőbben fejeződhet be. (Minden esetben nemnegatív értékű, a kritikus tevékenységekre pedig ez 0 értékűnek adódik.)
- Szabad tartalékidő:

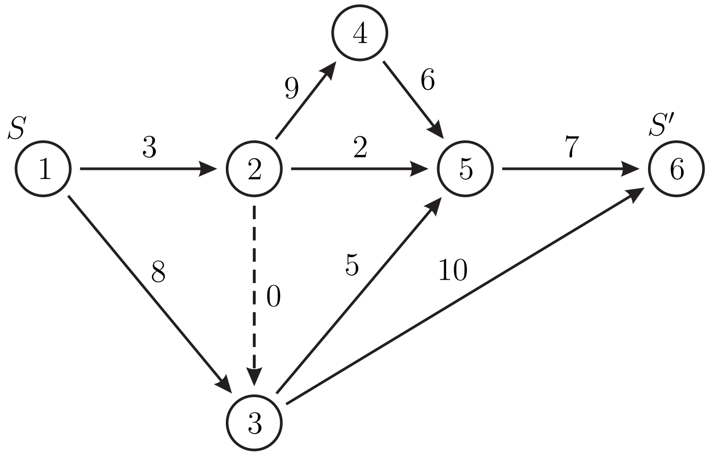
$$\underline{\mu}(y) - \underline{\mu}(x) - \tau(x, y)$$

- Független tartalékidő:

$$\underline{\mu}(y) - \overline{\mu}(x) - \tau(x, y)$$

  Ha ez pozitív, akkor ez az (x, y) tevékenységnél a többi tevékenységtől függetlenül felhasználható anélkül, hogy az bármilyen hatással lenne a munkafolyamat átfutási idejére. Általában azonban ez negatívnak adódik, azaz az (x, y) tevékenységnek nincs független tartalék ideje. Ilyenkor ezt 0-nak kell tekinteni.

Az eddig elmondottakra tekintsük példaként a következő tervütemhálót:



Ezen tervütemhálóban az 1-es csúcs jelöli az s kiindulási, a 6-os csúcs pedig az s' befejeződési csúcspontot, azaz a befejeződést jelentő eseményt.

A bejelölt élek a tevékenységek logikai sorrendjét is kifejezik. Pontosan emiatt volt szükség a $(2,3)$ úgynevezett látszattevékenység bevezetésére is, ugyanis így tudtuk azt kifejezni, hogy a 3-as eseményből kiinduló tevékenységek csak az után kezdődhetnek el, miután a 2-es esemény bekövetkezett.

Az egyes éleknél feltüntetett számok az él által reprezentált tevékenység elvégzési, azaz tevékenységi idejét jelölik. Nyilvánvaló, hogy egy látszattevékenységnél ez 0 értékű.

A $\underline{\mu}(x)$ és $\overline{\mu}(x)$ eseményidőket, illetve a tervütemhálóval kapcsolatos egyéb információkat szokás táblázatba foglalni. Egy ilyen táblázatot Clarke–Weber mátrixnak neveznek. Ha a sorszámozás olyan, hogy nagyobb sorszámú pontba vezet él, akkor ez a táblázat egy felső háromszög mátrix lesz.

A $\underline{\mu}(x)$ legkorábbi eseményidőket az utolsó oszlopban, a $\overline{\mu}(x)$ legkésőbbieket pedig az utolsó sorban számoljuk. Egy x csúcspont $\underline{\mu}(x)$ értékét úgy határozzuk meg, hogy a $\underline{\mu}$ értékek oszlopában és az x csúcspont oszlopában lévő számokat soronként külön-külön összeadjuk, majd ezek közül a maximumot választjuk. Egy x csúcsponthoz a $\overline{\mu}(x)$ értékét pedig úgy határozzuk meg, hogy a $\overline{\mu}$ értékek sorában és az x pont sorában lévő elemeket minden oszlopban külön-külön kivonjuk egymásból, majd vesszük ezen különbségek közül a minimumot.

A $\underline{\mu}(x)$ értékek számolásához a $\underline{\mu}(s) = 0$, a $\overline{\mu}(x)$ értékek kiszámításához pedig a $\overline{\mu}(s') = \underline{\mu}(s)$ kiindulási összefüggéseket használjuk fel.

Ezen megjegyzések után az előbbi tervütemhálóhoz tartozó Clarke–Weber mátrix a következő lesz:

	$S = 1$	2	3	4	5	$S' = 6$	$\underline{\mu}$
$S = 1$		3	8				0
2			0	9	2		3
3					5	10	8
4					6		12
5						7	18
$S' = 6$							25
$\overline{\mu}$	0	3	13	12	18	25	

$\longrightarrow$ a kiindulási feltétel alapján

$\longrightarrow \max\{3 + 0\} = 3$

$\longrightarrow \max\{8 + 0, 0 + 3\} = 8$

$\longrightarrow \max\{9 + 3\} = 12$

$\longrightarrow \max\{2 + 3, 5 + 8, 6 + 12\} = 18$

$\longrightarrow \max\{10 + 8, 7 + 18\} = 25$

a kiindulási feltétel alapján

$\min\{25 - 7\} = 18$

$\min\{18 - 6\} = 12$

$\min\{18 - 5, 25 - 10\} = 13$

$\min\{13 - 0, 12 - 9, 18 - 2\} = 3$

$\min\{3 - 3, 13 - 8\} = 0$

A kritikus út megállapításához meg kell állapítanunk melyek azok a csúcspontok, amelyeknél a $\underline{\mu} = \overline{\mu}$ egyenlőség teljesül. Ennek alapján $1 \to 2 \to 4 \to 5 \to 6$ lesz a kritikus út. Az ütemezés értékére 25 adódott. Ezen feladathoz tartozó tevékenység- és tartalékidőket a következő táblázat tartalmazza:

Tevékenység (x,y)	Tevékenységidő $\tau(x,y)$	$\underline{\mu}(x)$	$\underline{\mu}(x)+\tau(x,y)$	$\overline{\mu}(y)$	$\overline{\mu}(y)-\tau(x,y)$	Teljes tartalékidő $\overline{\mu}(y)-\underline{\mu}(x)-$ $-\tau(x,y)$	Szabad tartalékidő $\underline{\mu}(y)-\underline{\mu}(x)-$ $-\tau(x,y)$	Független tartalékidő $\underline{\mu}(y)-\overline{\mu}(x)-$ $-\tau(x,y)$
$(1,2)$	3	0	3	3	0	0	0	0
$(1,3)$	8	0	8	13	5	5	0	0
$(2,3)$	0	3	3	13	13	10	5	5
$(2,4)$	9	3	12	12	3	0	0	0
$(2,5)$	2	3	5	18	16	13	13	13
$(3,5)$	5	8	13	18	13	5	5	0
$(3,6)$	10	8	18	25	15	7	7	2
$(4,5)$	6	12	18	18	12	0	0	0
$(5,6)$	7	18	25	25	18	0	0	0

Az előbb ismertetett módszert az irodalomban kritikus út módszerének (CPM vagy CPM/time) nevezik.

A kritikus út módszere feltételezte, hogy a tevékenységidők determinisztikusan adottak. A gyakorlati életben ez fordul elő a legkevesebbszer, ezért célszerű volt olyan módszert is kifejleszteni, amelyben a tevékenységidőket valószínűségi változóknak tekinthetjük. Ezt a módszert az irodalomban PERT módszernek nevezik.

A legtöbb gyakorlati problémánál a tevékenységidők eloszlása β-eloszlásnak tekinthető. Az ilyen esetekben a következő három értékkel adhatjuk meg a tevékenységidő becslését.

Minden (x, y) tevékenységhez megadjuk az $a(x, y)$ értéket, amely optimista becslése a tevékenységidőnek. Megadjuk továbbá a $b(x, y)$ mennyiséget is, amely a tevékenységidő legvalószínűbb értéke, végül pedig a $c(x, y)$ számot, amely a tevékenységidő pesszimista becslése. Az optimista becslés a bizonytalanságot okozó akadályokat nem veszi figyelembe, a pesszimista becslés pedig minden lehetséges akadály fellépését számbaveszi. Az előbbi három becslés segítségével a tevékenységidő várható értékére az

$$m(x, y) = a(x, y) + 4 \cdot b(x, y) + c(x, y),$$

szórására pedig az

$$s(x, y) = \frac{c(x, y) - a(x, y)}{6}$$

adható meg becslésként. Az időtervezés ezután a tevékenységidő $m(x, y)$ várható értékével történik a már megismert CPM módszerrel. A modell jellegéből adódóan az eseményidőkre és az átfutási időre is várható értéket kapunk. Tisztáznunk kell még, hogy mekkora lesz az átfutási idő szórása. Minthogy az eseményidőket az egymástól független $m(x, y)$ tevékenységidők összegeként számítjuk a CPM módszerrel, ezért az átfutási idő szórásnégyzetét a kritikus úthoz tartozó tevékenységidők szórásnégyzetei összegeként számíthatjuk ki.

A központi határeloszlástételt felhasználva állíthatjuk, hogy ha a tervütemháló elég sok pontból áll, illetve a kritikus út sok eseményt tartalmaz, akkor az átfutási idő közelítőleg normális eloszlásnak tekinthető. A normális eloszlást felhasználva választ kaphatunk olyan kérdésekre is, hogy mennyi a valószínűsége annak, hogy az átfutási idő egy adott értéken belül illetve kívül maradjon, vagy hogy valamilyen tartományba essen.

Példaként tekintsük a következő táblázattal adott, PERT módszerrel megoldható problémát.

Tevékenység	A tevékenység időtartamának			Tevékenységidő várható értéke	Tevékenységidő szórása
	optimista becslése	legvalószínűbb becslése	pesszimista becslése		
$(1,2)$	7	10	13	10	1
$(1,3)$	8	14	20	14	2
$(2,3)$	2	5	8	5	1
$(2,4)$	14	20	26	20	2
$(3,4)$	13	19	25	19	2

Ezt követően a tevékenységidők várható értékeivel készítsük el a Clarke–Weber mátrixot:

	1	2	3	4	$\underline{\mu}(x)$
1		10	14		0
2			5	20	10
3				19	15
4					34
$\overline{\mu}(x)$	0	10	15	34	

Ebből kiolvasható, hogy az $1 \to 2 \to 3 \to 4$ lesz a kritikus útvonal, 34 lesz az átfutási idő várható értéke és $\sqrt{1^2 + 1^2 + 2^2} = \sqrt{6}$ lesz az átfutási idő szórása.

I.5. Költségtervezés

A CPM (és a PERT) módszernél a költségtervezés a határidőtervezésnek a szerves része. Eddig azt vizsgáltuk, hogy hogyan készül el a kész hálózat időterve, a kritikus út és az időtartalékok számítása. A költségtervezés ennek a szakasznak a lezárása után kezdődik. Az átfutási idő megállapítása után szükségszerűen jelentkezik az a feladat, hogy a kritikus út időtartamát a rendelkezésre álló források, lehetőségek felhasználásával megpróbáljuk csökkenteni. Ez többféleképpen is elképzelhető.

A CPM a költségek figyelembevételével választja ki a kritikus út rövidítésére a legkedvezőbb megoldást. A célja az, hogy a tervütemháló kritikus útjának rövidítését a költségalakulás szempontjából legkedvezőbb módon érje el. Ezen probléma megoldása közben természetes módon felmerülhet az a kérdés, hogy egyáltalán milyen módon csökkenthető a kritikus út időtartama, amikor azt állítottuk róla, hogy az a legrövidebb időtartam, amely alatt a feladat befejezhető? Ezen kérdésre adandó választ megértjük, ha átgondoljuk a következőket. A tevékenységek egész soránál meg lehet változtatni a szükséges időtartamot, mindössze csak a feltételeket kell módosítani. Az időszükséglet változtatására sokféle lehetőséget ismerünk. A munkaerőn és a munkarenden kívül például további lehetőséget ad az addig kézzel végzett bizonyos tevékenységek gépesítése, stb. Ugyanakkor a kritikus út hosszának zsugorítására nemcsak az egyes tevékenységek időtartamának kisebbítése útján van lehetőségünk, hanem a tervütemháló szerkezetének megváltoztatásával is. Az adott gépi kapacitás korlátozott volta miatt például sokszor fordul elő, hogy egyébként párhuzamosan végezhető tevékenységeket is kénytelenek vagyunk egymás után elvégezni. Az ilyen „soros kapcsolás" helyett a kritikus út csökkentésére igen alkalmas eszköz a „párhuzamos kapcsolás" bevezetése. Például bizonyos megmunkálást, vagy esetleg alkatrészek készítését a vállalat bérmunkába adja ki, így mérsékelve a kapacitáshiányt. Ezzel a módszerrel is lényegesen rövidíthető a kritikus út időtartama.

Az átfutási időtartam csökkenésének természetesen korlátai is vannak. Nyilvánvalóan a legnagyobb erőfeszítéssel sem vagyunk képesek egy-egy tevékenység végrehajtási időtartamát minden határon túl csökkenteni. Felmerül tehát a kérdés, hogy meddig csökkentsük hát egy-egy tevékenység időtartamát? Erre választ adni igen nehéz. A vállalatnak számításba kell vennie a rendelkezésre álló összes lehetőséget és azok alapján kell a legkisebb átfutási időt igénylő módot és időtartamot megállapítani, megtervezni. Ezt az időtartamot minimális időtartamnak nevezzük. Mindezekből következik, hogy a tervezéskor minden egyes tevékenységre két időtartamot kell megadni: a normális és a minimális időtartamot. Tudomásul kell azonban azt is venni, hogy minden időrövidítés általában költségtöbblettel jár.

Minthogy az idő-költség arány az egyes tevékenységeknél más és más, az „első menetben" már megtervezett átfutási időtartam csökkentése, attól függően, hogy melyik tevékenység időtartamát rövidítjük, eltérő költségtöbblettel jár. Kézenfekvő, hogy a lehetséges megoldások közül azt kell kiválasztanunk, amelyiknél a határidőrövidítés a költségek legkisebb növekedését okozza.

Egy ilyen feladattal kapcsolatos költségeket két csoportba szokás sorolni: közvetlen és közvetett költségek.

A közvetlen költségek között csak azok a költségek szerepelnek, amelyek az adott tevékenységre közvetlenül elszámolhatók és annak időtartamától részben, vagy egészben függnek. A gazdaságossági számítást kiegészítik olyan nyereségek, bevételek, amelyek a határidőváltoztatásnak a következményei. Például egy tevékenység korábbi befejezése lehetővé teszi az ott használt gépeknek más munkára való beállítását, amely aztán nyereséget eredményezhet.

Az összes tevékenységre egyenként kell meghatározni a normális időtartamhoz, illetve a minimális időtartamhoz tartozó közvetlen költségeket. Magától értetődő, hogy lesznek olyan tevékenységek, ahol az időtartamot nem tudjuk csökkenteni, de elképzelhető olyan is, ahol az időtartam csökkentése nem jár költségnövekedéssel.

A tevékenységekre vonatkozó adatok összeállítása az optimum meghatározásának csupán számítástechnikai alapját képezi. A nehezebb feladat az egész tervre vonatkozó időtartam csökkentésnél optimális megoldást biztosítani.

Az egész tervfeladat időtartamának csökkentése tulajdonképpen az átfutási idő zsugorítását jelenti. Tudjuk, hogy az átfutási időt a kritikus út határozza meg. Ebből következik, hogy az összes lehetséges út közül a kritikus úton kezdjük meg az átfutási idő csökkentését. Ezt úgy érhetjük el, hogy a kritikus utat alkotó egyes tevékenységek időtartamát csökkentjük, amit természetesen úgy kell végrehajtani, hogy az a legkisebb költségnövekedéssel járjon. Ezen algoritmus lépései a következők lesznek.

1. Először meghatározzuk a kritikus utat.
2. Ezt követően a kritikus úton lévő tevékenységek közül kiválasztjuk azt, amelyiknél a tevékenység idejének csökkentése a legkisebb költségnövekedéssel jár.
3. A kiválasztott tevékenység időtartamát mindaddig csökkentjük, amíg lehetséges. Ezt követően a kritikus úton a következő, legkedvezőbb költségfüggvénnyel rendelkező tevékenységeknél folytatjuk az időtartam rövidítését. Amennyiben a minimális időtartamig történő csökkentés előtt már olyan helyzet állna elő, hogy új kritikus út keletkezik, a következő tevékenység kiválasztásánál körültekintően kell eljárnunk.
4. Az eljárást mindaddig folytatjuk, amíg minimális időtartamhoz nem jutunk.

Az egész feladat minimális időtartama nem jelenti szükségszerűen azt, hogy az időtartamot minden tevékenységnél minimálisra kell csökkentenünk. Az egész hálózat minimális időtartamának megállapítása oly módon történik, hogy minden tevékenységnél minimális időtartamot véve figyelembe kiszámítjuk a kritikus utat. Ezen kritikus út hossza lesz az egész tervfeladat végrehajtásának minimális időtartama.

Ez az algoritmus azzal a feltételezéssel él, hogy a normális időtartamhoz rendelt költség a minimális közvetlen költség. Az optimális időtartamcsökkentéshez még ismernünk kell a közvetett költségek alakulását is, továbbá sok esetben nem hanyagolhatók el még más tényezők sem. Csak ezeknek együttes vizsgálata deríthet fényt a tényleges költség-idő összefüggésre, amely a döntés alapját képezheti.

Az állandó költségtételt az egész tervre vonatkozólag állapítjuk meg egy összegben. Ezen költség és nagysága az optimális átfutási időtartam szempontjából közömbös.

A nyereségkiesésből eredő veszteségeknek, mint költségeknek a figyelembe vétele már problémát okoz. A nehézség akkor adódik, amikor a tényleges várható nyereség nagyságát kell eldöntenünk, ez a tényező ugyanis jelentősen befolyásolhatja az optimális idő meghatározását.

Ezek után a költségtervezési feladat (CPM/cost) modellje a következő lesz:

Tekintsünk egy CPM/time fealdatot és tegyük fel, hogy a tevékenységi idők nem fix értékűek, hanem a $\tau(x,y)$ értékek két rögzített korlát között vehetik fel értéküket:

$$a(x,y) \le \tau(x,y) \le b(x,y) \qquad \text{minden } (x,y) \text{ él esetén.}$$

Az $a(x,y)$ értéket szokás „roham”, a $b(x,y)$ értéket pedig „normál” időnek nevezni. Az (x,y) tevékenység költsége az $[a(x,y), b(x,y)]$ intervallumon legyen olyan, hogy normál időnél legyen a legkisebb, a rohamidőnél pedig a legnagyobb értékű és legyen lineáris függvénnyel megoldható. Íly módon, ha valamely $\tau(x,y)$ értéket rögzítünk, akkor az ehhez tartozó költség $k(x,y) - c(x,y) \cdot \tau(x,y)$. Ezek után a feladatunk olyan tervütemezés megadása, amely mellett a terv összköltsége, a

$$\sum_{(x,y)\in A} [k(x,y) - c(x,y) \cdot \tau(x,y)]$$

költség minimális. Minthogy a

$$\sum_{(x,y)\in A} k(x,y)$$

egy konstans érték, így az előbbi minimalizálási probléma ekvivalens a

$$\sum_{(x,y)\in A} c(x,y)\cdot\tau(x,y)$$

maximalizálásával. Ezek után a költségtervezés feladata matematikailag a következő formában írható fel:

Adott az $[N,A]$ tervütemháló, amelynek (x,y) élein adottak az $a(x,y)\le b(x,y)$, valamint a $c(x,y)$ nemnegatív egész értékek. Adott pozitív egész λ esetén keresendő a hálózat csúcsain értelmezett $\mu(x)$, valamint a hálózat élein értelmezett $\tau(x,y)$ egész értékek úgy, hogy a következők teljesüljenek:

$$\begin{gathered} a(x,y)\le\tau(x,y)\le b(x,y)\qquad\forall(x,y)\in A\\ \tau(x,y)\le\mu(y)-\mu(x)\\ \mu(s)=0,\quad\mu(s')=\lambda. \end{gathered}$$

Ezen feltételek mellett maximalizálandó az

$$U(\lambda)=\sum_{(x,y)\in A} c(x,y)\cdot\tau(x,y)$$

haszonfüggvény. (Itt λ az egész terv átfutási idejét jelenti, amit előre megadunk.)

Ez a feladat kapcsolatba hozható egy úgynevezett folyamproblémával. Mielőtt ezt a kapcsolatot vizsgálnánk, a folyamproblémát kell definiálnunk és az azzal kapcsolatos összefüggéseket megadnunk.

I.6. Maximális folyam, minimális vágás feladatok

Tekintsünk egy $[N,A]$ digráfot. Ennek élein értelmezünk egy $k(x,y)$ nemnegatív egész értékű függvényt, amelyet kapacitás függvénynek fogunk nevezni. Az $[N,A,k]$ konfigurációt hálózatnak fogjuk nevezni. Amennyiben a hálózat az összes lehetséges élt tartalmazza, azt teljes hálózatnak fogjuk nevezni és $[N,k]$-val fogjuk jelölni.

Ha a kiindulási hálózatunk, az $[N,A,k]$ hálózat nem teljes, akkor azt teljessé alakíthatjuk úgy, hogy a hiányzó (x,y) éleket is hozzávesszük $k(x,y)=0$ kapacitással.

Egy teljes hálózaton definiálhatjuk a folyam fogalmát, mint egy, az (x, y) éleken értelmezett olyan $f(x, y)$ egész értékű függvényt, amelyre a következők teljesülnek:

$$f(x, y) \leq k(x, y) \qquad \forall (x, y) \text{ élre (korlátos)}$$
$$f(x, y) = -f(y, x) \qquad \forall (x, y) \text{ élre (ferdén szimmetrikus).}$$

Nyilvánvaló, hogy ha egy (x, y) él nem eleme A-nak, akkor $f(x, y) = 0$ kell, hogy legyen, vagyis x-ből y-ba valójában nem folyhat semmi sem.

A rövidítés céljából vezessük be a következő jelöléseket:

- Ha $A \subset N, B \subset N$ és $h(x, y)$ az éleken értelmezett függvény (például kapacitás, vagy folyamfüggvény), akkor legyen
$$h(A, B) = \sum_{x \in A, y \in B} h(x, y).$$
Ezzel a $h(x, y)$ függvénnyel kapcsolatban érvényesek a következő összefüggések:
$$h(A \cup B, C) = h(A, C) + h(B, C)$$
$$h(C, A \cup B) = h(C, A) + h(C, B).$$
- A folyam- és kapacitás függvényekkel kapcsolatban, az előbbi rövidítést szolgáló jelölést felhasználva kapjuk, hogy
$$f(A, B) \leq k(A, B) \qquad \forall A, B \subset N \text{ esetén}$$
$$f(A, A) = 0 \qquad \forall A \subset N$$
$$f(x, N) = \sum_{y \in N} f(x, y).$$

Ezek után definiálhatjuk az s forrás és az s' nyelő pontokat úgy, mint olyan csúcspontjai a hálózatnak, amelyekre az f folyamfüggvény esetén az $f(s, N) > 0$, illetve az $f(s', N) < 0$ teljesül.

Az $[N, k]$ hálózat egy f folyamát s-ből s'-be irányuló folyamnak nevezzük, ha $f(s, N) > 0, f(s', N) < 0$ és minden $x \neq s, s'$ közbülső pont esetén $f(x, N) = 0$ teljesül.

Szemléletesen nyilvánvaló, de algebrailag is igazolható, hogy s-ből s'-be irányuló folyam esetén $f(s, N) = f(N, s')$.

Ugyanis
$$0 = f(N, N) = f(s, N) + f(s', N) + f(N \setminus \{s, s'\}, N).$$
Minthogy $f(N \setminus \{s, s'\}, N) = 0$, ezért nyilván
$$f(s, N) + f(s', N) = 0,$$

amiből

$$f(s,N) = -f(s',N) = f(N,s')$$

adódik.

Az s-ből s'-be irányuló folyamnál az $f(s,N)$ számot a folyam értékének nevezzük. A legnagyobb értékű folyamot maximális folyamnak nevezzük.

Ezek után fogalmazzuk meg a maximális folyam problémát.

Keressük azt az s-ből s'-be irányuló folyamot, amelyre az

$$\begin{aligned} f(s,N) &> 0 \\ f(s',N) &< 0 \\ f(x,N) &= 0, \qquad \text{ha } x \neq s, s' \\ f(x,y) &\leq k(x,y) \\ f(x,y) &= -f(y,x) \end{aligned}$$

összefüggések teljesülnek és amelyre $f(s,N)$ maximális.

A minimális vágás feladat a következőképpen foglamazható meg.

Legyen az $[N,A,k]$ hálózaton egy (S,S'), s-et s'-től elválasztó vágás. Ekkor a

$$k(S,S') = \sum_{x \in S, y \in S'} k(x,y)$$

értéket az (S,S') vágás átbocsátóképességének nevezzük. Az összes, az s-et s'-től elválasztó (S,S') vágás közül a legkisebb átbocsátóképességűt az s, s' pontokra vonatkozó minimális vágásnak nevezzük.

Azt az (S,S') vágást keressük, amelynek az $[N,k]$ hálózaton a $k(S,S')$ átbocsátóképessége minimális. (Nyilvánvaló, hogy az eredeti $[N,A,k]$ hálózat minimális vágása ugyanaz, mint az $[N,k]$ teljes hálózaté.)

Könnyen igazolható, hogy tetszőleges s-ből s'-be irányuló folyam értéke nem nagyobb, mint az s-et s'-től elválasztó tetszőleges (S,S') vágás átbocsátóképessége, azaz $f(s,N) \leq k(S,S')$. Ezen egyenlőtlenség belátásához induljunk ki az $f(s,N)$ folyamértékből:

$$\begin{aligned} f(s,N) = f(s,N) + f(S \backslash s, N) = f(S,N) = f(S, S \cup S') = \\ = f(S,S) + f(S,S') = f(S,S') \leq k(S,S'). \end{aligned}$$

Itt kihasználtuk, hogy $f(S \setminus s, N) = 0$ és $f(S,S) = 0$, valamint, hogy $N = S \cup S'$.

Ezen egyenlőtlenség következménye, hogy ha az f folyam és az (S,S') vágás olyanok, hogy egyenlőség áll fenn az az előbbi egyenlőtlenségnél, akkor az f folyam maximális, az (S,S') vágás pedig minimális. A kérdés mostmár az, hogy létezik-e ilyen f folyamfüggvény és ilyen (S, S') vágás. A kérdésre a választ a Ford–Fulkerson tétel adja meg.

Tétel. *Létezik olyan f folyamfüggvény és (S, S') vágás, hogy*

$$\max_f f(s, N) = \min_{(S,S')} k(S, S'),$$

azaz létezik maximális folyam és minimális vágás, és a maximális folyam értéke megegyezik a minimális vágás átbocsátó képességével.

Bizonyítás. Legyen f egy tetszőleges s-ből s'-be irányuló folyam. (Kiinduláskor az $f(x, y) \equiv 0$ folyamfüggvényt választjuk.) Konstruáljuk meg azt az $[N, \beta]$ digráfot, amelynek β élhalmaza a „telítetlen élek” halmaza, azaz olyan (x, y) éleké, amelyekre $f(x, y) < k(x, y)$. Az $[N, \beta]$ digráfban vagy van út s-ből s'-be, vagy van s-et és s'-t elválasztó (S, S') vágás.

Ha nincs út, akkor az (S, S') vágás ebben a digráfban üres. Nem üres azonban az $[N, A, k]$ hálózatban, és ott az összes (S, S')-beli él telített él lesz, így ez lesz az optimális megoldás.

Ha viszont vezet út az $[N, \beta]$ digráfban s-ből s'-be, akkor ezen P út minden éle telítetlen, azaz ezek az élek rendelkeznek a $k(x, y) - f(x, y) > 0$ szabad kapacitással. Legyen

$$\delta = \min_{(x,y)\in P} \{k(x, y) - f(x, y) > 0\}, \qquad \text{azaz}$$

δ egyezzen meg a P út mentén adódó legkisebb szabad kapacitással. Ezzel a δ értékkel a folyam növelhető minden, a P úthoz tartozó élen. A ferdeszimmetrikusság miatt viszont a visszaút mentén csökkenteni kell a folyamot:

$$\overline{f}(x, y) = \begin{cases} f(x, y) + \delta, & \text{ha } (x, y) \in P \text{ (az út mentén)} \\ f(x, y) - \delta, & \text{ha } (x, y) \in P \text{ (a visszaút mentén)} \\ f(x, y) & \text{az egyéb éleken.} \end{cases}$$

Az új folyam értéke $\overline{f}(s, N) = f(s, N) + \delta > f(s, N)$. Ezen új folyamfüggvényre megismételjük az előbbi lépéseket. Mivel $\delta > 0$ egész és $f(s, N)$ felülről korlátos, így véges sok lépésben a folyamnövelések sorozatában eljutunk az optimális megoldáshoz. □

(A folyamfüggvény megváltoztatása tulajdonképpen egyenértékű azzal, hogy a szabadkapacitás értéket változtatjuk meg: az út élein δ-val csökkentjük, a visszaút élein pedig δ-val növeljük a szabadkapacitás értékeket.)

I.7. A CPM/cost megoldási algoritmusa

Tekintsük most ismét az

$$\begin{aligned} a(x,y) \le \tau(x,y) \le b(x,y), &\qquad (x,y) \in A \\ \tau(x,y) \le \mu(y) - \mu(x) & \\ \mu(s) = 0, \quad \mu(s') = \lambda & \\ \max U(\lambda) = \max \left\{ \sum_{(x,y)\in A} c(x,y)\tau(x,y) \right\} & \end{aligned}$$

CPM/cost problémát, amellyel kapcsolatban egy olyan $[N,k]$ hálózatot definiálhatunk, amelynél

$$k(x,y) = \begin{cases} \infty, & \text{ha } (x,y) \in A \\ 0, & \text{egyébként.} \end{cases}$$

A következő állítás megadja, hogy ez a probléma milyen kapcsolatba hozható egy folyamproblémával.

Tétel. *Legyenek a $\mu(x)$ és a $\tau(x,y)$ függvények tetszőlegesek, de a definiált CMP/cost feladat feltételeinek eleget tevők. Legyen F az $[N,k]$ hálózaton egy tetszőleges s-ből s'-be irányuló folyam, melynek értéke legyen ν. Ekkor*

$$\begin{aligned} \sum_{(x,y)\in A} c(x,y)\tau(x,y) \le \lambda\mu + \sum_{\substack{(x,y)\in A \\ f<c}} \left[c(x,y) - f(x,y)\right] b(x,y) - & \\ - \sum_{\substack{(x,y)\in A \\ f>c}} \left[f(x,y) - c(x,y)\right] \cdot a(x,y). & \end{aligned} \tag{*}$$

Minthogy ez az egyenlőtlenség fennáll minden $\tau(x,y)$ és $\mu(x)$ értékre, valamint s-ből s'-be irányuló, az $[N,k]$ hálózaton értelmezett folyamra, ezért amennyiben itt egyenlőség áll fenn, akkor a bal oldal maximális, a jobb oldal pedig minimális. Ahhoz pedig, hogy egyenlőség álljon fenn, azaz a $\tau(x,y)$ és a $\mu(x)$ optimalizálják a haszonfüggvényt, elégséges feltétel, hogy létezzék olyan f folyamfüggvény, amelyre

(i) $\tau(x,y) < \mu(y) - \mu(x)$ esetén $f(x,y) = 0$
(ii) $\tau(x,y) < b(x,y)$ esetén $f(x,y) \ge c(x,y)$
(iii) $\tau(x,y) > a(x,y)$ esetén $f(x,y) \le c(x,y)$

teljesül. Az $U(\lambda)$ haszonfüggvény maximalizálása miatt optimális megoldás esetén elég a

$$\tau(x,y) = \min\{\mu(y) - \mu(x), b(x,y)\}$$

értékre szorítkozni. Az optimalitás előbbi kritériumának megfelelően az A-beli éleket egyértelműen besorolhatjuk a következő öt csoport valamelyikébe:

I. ha az (i) (esetleg (iii)) fennáll
II. az (ii) és (iii) egyszerre teljesülnek
III. csak az (ii) teljesül
IV. csak az (iii) teljesül
V. az (i), (ii) és (iii) közül egyik sem teljesül, ami csak úgy lehetséges, ha $a(x,y) = b(x,y) = \mu(y) - \mu(x)$.

A következő tétel a CPM/cost feladat megoldását szolgáló algoritmus gerincét adja.

Tétel. *Ha valamely λ-ra létezik olyan $\tau(x,y), \mu(x), f(x,y)$ megoldás, amely az optimalitási kritériumnak eleget tesz, akkor vagy létezik $\lambda^* < \lambda$ érték, amelyre megadható az optimalitási kritériumnak eleget tevő megoldás, vagy pedig λ az a legkisebb érték, amelyre a feladat megoldható.*

Bizonyítás. Készítsünk el egy $[N, \overline{k}]$, úgynevezett szabadkapacitás hálózatot a következőképpen:

Ha $(x,y) \in A_{I.}$, akkor $\overline{k}(x,y) = 0, \overline{k}(y,x) = 0$.
Ha $(x,y) \in A_{II.}$, akkor $\overline{k}(x,y) = 0, \overline{k}(y,x) = 0$.
Ha $(x,y) \in A_{III.}$, akkor $\overline{k}(x,y) = \infty, \overline{k}(y,x) = f(x,y) - c(x,y)$.
Ha $(x,y) \in A_{IV.}$, akkor $\overline{k}(x,y) = c(x,y) - f(x,y), \overline{k}(y,x) = f(x,y)$.
Ha $(x,y) \in A_{V.}$, akkor $\overline{k}(x,y) = \infty, \overline{k}(y,x) = f(x,y)$.

A többi él kapacitása legyen 0.

Legyen ezen $[N, \overline{k}]$ hálózaton az s-ből s'-be menő maximális folyam g, a minimális vágás pedig (S, S'). (Itt a $g = 0$ is lehetséges!)

Legyen $f^* = f + g$. Ezen f^* folyam a $\overline{k}$ szabadkapacitások megválasztása miatt szintén eleget tesz az optimalitási kritériumnak a $\mu(x), \tau(x,y)$ értékekkel, valamint a ν^* folyamértékkel. (Itt ν^* az f^* folyamfüggvény maximális értékét jelöli.)

Amennyiben f^*, azaz a g értéke ∞, ez azt jelenti, hogy van olyan út s-ből s'-be, hogy az út mentén $(x,y) \in A_{III.} \cup A_{IV.}$. Ez azonban azt jelenti, hogy az út mentén $\tau(x,y) = a(x,y)$, vagyis λ az a minimális szám, amelyre a feladat megoldható.

Amikor $f^* < \infty$, az (S, S') vágásban lévő élek olyanok lesznek, hogy vagy $f^* = 0$, vagy $f^* = c$ teljesül.

Jelöljük δ_1-gyel a

$$\delta_1 = \min_{\substack{(x,y)\in A\\ x\in S, y\in S'\\ f^*(x,y)=0}} \{\mu(y) - \mu(x) - b(x,y)\},$$

δ_2-vel pedig a

$$\delta_2 = \min_{\substack{(x,y)\in A\\ x\in S, y\in S'\\ f^*(x,y)=c(x,y)}} \{\mu(y) - \mu(x) - a(x,y)\}$$

értéket. Legyen továbbá $\delta = \min\{\delta_1, \delta_2\}$. Figyelembe véve az A-beli élek kategorizálását, könnyen belátható, hogy $\delta > 0$.

Legyenek

$$\begin{aligned} \mu^*(x) &= \mu(x), && \text{ha } x \in S\\ \mu^*(x) &= \mu(x) - \nu, && \text{ha } x \in S' \text{ és így}\\ \lambda^* &= \lambda - \nu, && \text{ahol } \nu = 0, 1, 2, \ldots, \delta. \end{aligned}$$

A $\tau^*(x,y)$ értékeket a $\tau^*(x,y) = \min\{\mu^*(y) - \mu^*(x), b(x,y)\}$ összefüggésből határozzuk meg. Belátható, hogy a $\mu^*(x)$, a $\tau^*(x,y)$ és az $f^*(x,y)$ értékek kielégítik az optimalitási kritériumot. □

Jelöljük $\lambda_{\max}$-szal – a $b(x,y)$ tevékenységi időket használva – a CPM/time módszerrel meghatározott befejezési időt, $\lambda_{\min}$-nel pedig az $s(x,y)$ tevékenységi időkkel meghatározottat. Így $\lambda_{\max}$-ból, azaz az azonosan 0 folyamból elindulva minden λ-ra (egészen $\lambda_{\min}$-ig) megoldjuk a CPM-cost feladatot. Ezt felhasználva kimondhatjuk a dualitási tételt:

Tétel. *Tetszőleges $\lambda_{\min} \le \lambda \le \lambda_{\max}$ esetén léteznek olyan $\mu(x), \tau(x,y)$ és $f(x,y)$ függvények, hogy a (*) egyenlőtlenségben egyenlőség áll fenn.*

I.8. A Kőnig feladatok

Tekintsük a következő problémát.

Adottak az $I_1, I_2, \ldots, I_m$ munkások és a $J_1, J_2, \ldots, J_n$ feladatok (munkahelyek). Minthogy nem kell szükségképpen képesnek lennie minden munkásnak minden feladat elvégzésére, ezért megadják azt is, hogy melyik munkás milyen feladatok elvégzésére alkalmas. Ezt szokás egy úgynevezett kvalifikációs táblázatban megadni, amelyben ha az (i,j) pozicióban egy x jelölés van, akkor az azt jelenti, hogy az I_i személy a J_j munkát el tudja végezni. Ezek után a feladat: úgy rendeljünk a munkások mindegyikéhez munkahelyet, hogy

- a munkás értsen ahhoz a munkához
- egy munkás egy munkahelyen dolgozzon
- egy munkahelyen legfeljebb egy munkás dolgozzon.

Ezt a feladatot szokás „házasság” problémának is nevezni. Ez esetben a „munkások” a férfiak lesznek, a hölgyek pedig átveszik a „munkahelyek” szerepét. A kvalifikációs táblázatban adjuk meg azt, hogy melyik férfi melyik hölggyel „szimpatizál”. Ezután a férfiakhoz úgy kell hozzárendelni a hölgyeket, hogy

- minden férfihez olyan hölgyet kell rendelnünk, akivel szimpatizál
- egy férfihez csak egy hölgyet rendelhetünk (bigámia kizárva)
- egy hölgyhöz legfeljebb egy férfit rendelhetünk hozzá (azaz egy hölgy egyszerre csak egy férfinek lehet a felesége).

Mielőtt ezen probléma megoldhatóságára vonatkozó állítást kimondanánk, vezessük be a következő jelöléseket.

Jelölje $I = \{I_1, I_2, \ldots, I_m\}$ a munkások, $J = \{J_1, J_2, \ldots, J_n\}$ pedig a munkahelyek halmazát. Jelölje $P \subset I$ a munkások egy tetszőleges részhalmazát és jelölje $R = J(P) \subset J$ a munkák azon halmazát, amelyekhez a P-beli munkások összességében értenek. Jelölje $||P||$ a P halmaz, $||R|| = ||J(P)||$ pedig a $J(P)$ halmaz elemeinek számát.

Kőnig Dénestől származik a következő tétel.

Tétel. *Adott I, J és kvalifikációs tábla esetén vagy van megoldása a feladatnak, vagy van a munkásoknak olyan $P \subset I$ részhalmaza, hogy*

$$||P|| > ||J(P)||.$$

Más szavakkal: vagy megoldható a feladat, vagy megadható a munkások olyan részhalmaza, amelynek nagyobb az elemszáma, mint azon munkák száma, amennyit ezek a munkások együttesen el tudnak látni.

Ez a probléma általánosabban is megfogalmazható. Jelöljék $T_1, T_2, \ldots, T_m$ a termelőket (termelő helyeket), amelyeknek egy bizonyos termékből legyen $\alpha_1, \alpha_2, \ldots, \alpha_m$ a kínálatuk. Jelöljék továbbá $F_1, F_2, \ldots, F_n$ a fogyasztókat (fogyasztóhelyeket), amelyeknek $\beta_1, \beta_2, \ldots, \beta_n$ keresletük van az adott termékből. Feltesszük, hogy az $\alpha_i \geq 0,\ i = 1, \ldots, m$ és $\beta_j \geq 0,\ j = 1, \ldots, n$ egészek. Adottak továbbá, hogy mely termelőtől mely fogyasztóhoz történhet szállítás; ezt egy kvalifikációs táblázattal adjuk meg.

Ezek után a feladat egy olyan szállítás meghatározása, amelynek során minden termelőtől minden felkínált terméket elszállítunk mégpedig olyan fogyasztókhoz, akikkel azok kapcsolatban állnak ügyelve arra, hogy senkinek se szállítsunk többet, mint amennyi annak az igénye volt.

Ezen probléma esetében is be kell vezetnünk néhány jelölést.

Jelölje $T = \{T_1, \ldots, T_m\}$ a termelők, $F = \{F_1, \ldots, F_n\}$ pedig a fogyasztók halmazát. Jelölje $P \subset T$ a termelők egy tetszőleges részhalmazát, $R = F(P) \subset F$ pedig azon fogyasztókét, akikkel a P-beli termelők kapcsolatban állhatnak, azaz akikhez a P-beli termelők szállíthatnak. Jelölje

$$||P||_\alpha = \sum_{T_i \in P} \alpha_i$$

a P-beli termelők összkínálatát,

$$||R||_\beta = ||F(P)||_\beta = \sum_{F_j \in R} \beta_j$$

pedig az R-beli fogyasztók összkeresletét.

Ezen jelöléseket felhasználva megfogalmazhatjuk Kőnig Dénes tételét.

Tétel. *Legyen adott a termelők T, a fogyasztók F halmaza. Legyen adott továbbá a szállítási kapcsolatokat kifejező kvalifikációs tábla. Végül legyenek adva az $\alpha_i, i = 1, \ldots, m$ és $\beta_j, j = 1, \ldots, n$ mennyiségek, amelyek a termelők kínálatát, illetve a fogyasztók igényét adják meg egy bizonyos termékre vonatkozóan. Ekkor vagy létezik megoldása azon feladatnak, hogy minden termelőtől minden felkínált terméket olyan fogyasztóhoz szállítsunk, akikkel azok kapcsolatban állnak, ügyelve arra, hogy mindenkihez legfeljebb az igényelt mennyiséget szállítsuk, vagy pedig létezik a termelőknek olyan P részhalmaza, hogy az általuk felkínált termékek $||P||_\alpha$ összmennyisége több, mint azon $F(P)$ fogyasztók $||F(P)||_\beta$ összigénye, akikkel ezek a termelők szállítási kapcsolatban állnak.*

Ezen általános feladatból az $\alpha_i = 1$ és $\beta_j = 1, i = 1, 2, \ldots, m; j = 1, 2, \ldots, n$ feltételek mellett adódik a „házasság" probléma, amennyiben az egységnyi termékek már tovább nem osztható mennyiségeket jelentenek. (Ez azt jelenti, hogy egy ilyen egységnyi termék csak egyetlen fogyasztóhoz szállítható el kielégítve annak egységnyi igényét.)

A bizonyítás a Ford-Fulkerson maximális folyam – minimális vágás tételen alapszik. Konstruáljunk meg egy hálózatot az alábbiak szerint:

Vegyünk fel egy s forrás és egy s' nyelőpontot.

Legyen

$$
\begin{aligned}
k(s, T_i) &= \alpha_i \\
k(F_j, s') &= \beta_j \\
k(T_i, F_j) &= \infty, && \text{ha } T_i\text{-ből lehet } F_j\text{-be szállítani,} \\
& && \text{minden más él kapacitása legyen } 0.
\end{aligned}
$$

A Kőnig-feladatot ezáltal egy folyamfeladatra vezettük vissza.

Vizsgáljuk meg, hogy milyen kapcsolat van a Kőnig feladat és a most definiált folyamfeladat között!

Ha az előbb definiált folyamproblémánál meghatározzuk az f folyamfüggvény egyes élekhez tartozó értékét, akkor azok közül az

- $f(s, T_i)$ a T_i-ből összesen elszállított mennyiséget,
- $f(T_i, F_j)$ a T_i termelőtől az F_j fogyasztóhoz szállított mennyiséget,
- $f(F_j, s')$ pedig az F_j-be szállított összmennyiséget

fogja jelenteni. A Kőnig feladat megoldását tehát az $f(T_i, F_j)$ értékek szolgáltatják, amelyek megadják, hogy mennyit lehet a T_i termelőtől az F_j fogyasztóhoz elszállítani.

Az $f(s, N) = \sum_{i=1}^{m} f(s, T_i)$ folyamérték maximális értékére a folyamprobléma megoldása során a következők adódhatnak:

1. eset: $\max f(s, N) = \sum_{i=1}^{m} \alpha_i$, azaz $f(s, T_i) = \alpha_i$ minden $i = 1, 2, \ldots, m$ értékre. Ez esetben minden termelőtől a kínálatnak megfelelő árumennyiséget szállítottuk el, azaz a feladatot megoldottuk.

2, eset: $\max f(s, N) < \sum_{i=1}^{m} \alpha_i$. Ekkor a Kőnig feladat nem oldható meg.

Ez utóbbi esetben határozzuk meg az s-et és s'-t elválasztó (S, S') minimális vágást, amely a vágásban lévő élek telítettsége miatt nyilván nem tartalmazhat ∞ kapacitású élt. Jelöljük P-vel az S-be tartozó termelőket, R-rel pedig az S'-be tartozó fogyasztókat. Ekkor tehát P-ből csak R-be mehet él ebben a vágásban, tehát $R = F(P)$, azaz R-be csak olyan fogyasztók tartozhatnak, amelyekhez P-beli termelők szállíthatnak.

Írjuk fel a minimális vágás értékét:

$$k(S, S') = \sum_{T_i \notin P} \alpha_i + \sum_{F_j \in R} \beta_j.$$

A Ford-Fulkerson tétel értelmében

$$\max f(s, N) = \min k(S, S'),$$

így tehát

$$\sum_{i=1}^{m} \alpha_i > \max f(s,N) = \min k(S,S') = \sum_{T_i \notin P} \alpha_i + \sum_{F_j \in R} \beta_j .$$

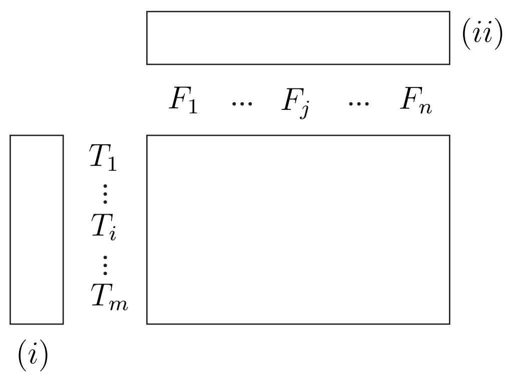
Ezt az egyenlőtlenséget rendezve kapjuk, hogy

$$\sum_{T_i \in P} \alpha_i > \sum_{F_j \in R} \beta_j$$

vagyis

$$||P||_\alpha > ||R||_\beta = ||F(P)||_\beta .$$

Az előbbiek alapján a Kőnig feladat megoldásához használhatjuk a maximális folyam – minimális vágás meghatározására szolgáló algoritmust. Ennél általában nem az azonosan zérus folyamból (szállításból) indulunk ki, hanem egy ún. kezdeti szállításból, amelyet É–NY sarok módszerrel határozhatunk meg. Az algoritmus során nem szükséges a folyamprobléma tábláján dolgozni, mivel a szabad kapacitást az eredeti $m \times n$-es táblában is tudjuk tárolni a következőképpen:



A folyamproblémánál az (s, T_i) él szabad kapacitását az (i) oszlopban, az (F_j, s') él szabad kapacitását pedig az (ii) sorban tároljuk. Ha T_i-ből lehet F_j-be szállítani, akkor a (T_i, F_j) él szabad kapacitása végtelen marad, így a tábla belsejében a tényleges szállítást (folyamot) célszerű tárolni. Végül az (F_j, T_i) él szabad kapacitását a (T_i, F_j) él pozitív folyama adja meg. (Ez nyilvánvalóan azt fejezi ki, hogy az F_j-ből a T_i-be legfeljebb annyit szállíthatunk, amennyit már T_i-ből F_j-be szállítottunk.)

Összefoglalva az előbbieket, a cimkézést a következőképpen kell végrehajtani:

- s-sel azokat a T_i termelőket cimkézzük, akiknél az (i) oszlopban pozitív szám található. (Ezeknek a termelőknek van szabad kapacitása, tehát a hálózatban s-ből ezekbe vezet pozitív kapacitású él.) Természetesen az s cimkét minden ilyen termelőhöz hozzá kell rendelni.
- T_i-ből F_j-be mehetünk akkor, ha az előbbi táblázatban az (i, j) pozícióban pozitív, vagy 0 érték van. A tiltott szállításokat „–" -szal jelöljük.
- F_j-ből T_i-be akkor mehetünk, ha az (i, j) pozícióban pozitív szám található.
- Az útkeresés végpontjai azon F_j fogyasztók lehetnek, amelyeknél az (ii)-ben pozitív szám áll.

Ezek után alkalmazzuk a maximális folyam – minimális vágás probléma megoldási algoritmusát.

I.9. Szűkkeresztmetszet feladatok

Egyszerű szűkkeresztmetszet feladat

Legyenek adottak egy „futószalagon" a $J_1, \ldots, J_n$ munkahelyek és legyenek $I_1, I_2, \ldots, I_n$ azok a munkások, akiket az előbbi munkahelyekhez akarunk hozzárendelni. Minden munkás minden munkahelyen lévő munkát képes ellátni, csak nem egyforma hatékonysággal. Jelölje τ_{ij} azt az időtartamot, amennyi idő alatt az I_i munkás a J_j munkahelyen lévő feladatot képes ellátni. $(i = 1, \ldots, n; j = 1, \ldots, n.)$

Ezek után a feladat: a futószalagon elhelyezett munkahelyekhez úgy rendeljük hozzá a munkásokat, hogy a futószalag ütemideje minimális legyen. (Egy futószalag ütemidején a futószalagról egymás után lekerülő két termék lekerülése között eltelt időt értjük.)

Könnyen belátható, hogy egy futószalag ütemideje megegyezik a futószalagon lévő egyes munkák elvégzéséhez szükséges időtartamok maximumával. A feladat tehát az, hogy úgy rendeljük az egyes munkásokat az egyes munkákhoz, hogy ez a maximális érték a lehető legkisebb legyen.

Ha χ egy hozzárendelést jelöl, akkor a problémánk a

$$\min_{\chi} \left\{ \max\{\tau_{i,j} \mid (I_i, I_j) \in \chi\} \right\}$$

formában írható le.

Egy ilyen problémát egyszerű Kőnig feladatok sorozatával oldhatjuk meg. Első lépésben meghatározunk egy χ kezdeti hozzárendelést, amelyet célszerű É–NY sarok módszerrel a legkisebb τ_{ij}-ken keresztül megadni. Ezt követően a második lépésben meghatározzuk a maximális műveleti időt, amit $\tau = \max\{\tau_{ij} \mid (I_i, J_j) \in \chi\}$-val fogunk jelölni. A harmadik lépésben egy olyan egyszerű Kőnig feladatot konstruálunk, amelynek a kvalifikációs táblázatában olyan kapcsolatok lesznek megengedettek, amelyekre $\tau_{ij} < \tau$ teljesül.

Ezután megpróbáljuk megoldani az így adódó egyszerű Kőnig-feladatot. Ha nem oldható meg, akkor a futószalag feladat optimális megoldása az előző χ hozzárendelés lesz. Ha viszont megoldható, akkor meghatározzuk az új χ' hozzárendelést, amelynél már nyílvánvalóan kisebb lesz az ütemidő, mint amilyen a χ esetében volt, majd a második lépésnél folytatjuk az eljárást.

Általános szűkkeresztmetszet feladat

Legyenek $T_1, T_2, \ldots, T_i, \ldots, T_m$ a termelőhelyek, amelyeknek egy bizonyos fajta termékből legyenek $\alpha_1, \ldots, \alpha_i, \ldots, \alpha_m$ a kínálatuk. Legyenek továbbá $F_1, \ldots, F_j, \ldots, F_n$ a fogyasztóhelyek, amelyeknek az említett termékből $\beta_1, \ldots, \beta_j, \ldots, \beta_n$ a keresletük. Feltesszük, hogy a $\sum_{i=1}^{m} \alpha_i = \sum_{j=1}^{n} \beta_j$. Legyen továbbá adott az alábbi táblázat:

$$\begin{array}{ccccccc}
 & & \beta_1 & \cdots & \beta_j & \cdots & \beta_n \\
 & & F_1 & \cdots & F_j & \cdots & F_n \\
\alpha_1 & T_1 & & & & & \\
\vdots & \vdots & & & & & \\
\alpha_i & T_i & & & \tau_{ij} & & \\
\vdots & \vdots & & & & & \\
\alpha_m & T_m & & & & &
\end{array}$$

Itt τ_{ij} a T_i-ből az F_j-be történő szállítás idejét jelöli, amely független a szállított mennyiségtől.

Ezek után a feladatunk olyan, a kereslet–kínálatot kielégítő szállítás meghatározása, amelynél a szállítási idő minimális, feltéve, hogy a szállítást minden útvonalon egy időben kezdik.

Jelölje ξ_{ij} a T_i-ből F_j-be szállított mennyiséget. Ekkor tehát meghatározandók a $\xi_{ij} \geq 0$ értékek úgy, hogy

$$\sum_{j=1}^{n} \xi_{ij} = \alpha_i, \qquad i = 1, 2, \ldots, m$$

$$\sum_{i=1}^{m} \xi_{ij} = \beta_j, \qquad j = 1, 2, \ldots, n$$

továbbá a

$$\max\{\tau_{ij} \mid \xi_{ij} > 0\}$$

minimális legyen.

Ezt a feladatot általános Kőnig feladatok sorozatán keresztül oldhatjuk meg. Az első lépésben É–NY sarok módszerrel meghatározunk egy kezdeti ξ_{ij} szállítást (mindent elszállítva, minden igényt kielégítve) a legkisebb τ_{ij} értékek felhasználásával. A második lépésben meghatározzuk a

$$\tau = \max\{\tau_{ij} \mid \xi_{ij} > 0\}$$

szállítási időt. Ezt követően a 3. lépésben konstruálunk egy általános Kőnig feladatot, amelynek kvalifikációs táblázatában csak olyan (i,j) pozíció lesz lehetséges, ahol $\tau_{ij} < \tau$. Ezt követően megpróbáljuk megoldani ezt az általános Kőnig feladatot. Ha nem oldható meg, akkor a szűkkeresztmetszetű szállítási feladat megoldása az előző szállítás lesz. Ha viszont megoldható, akkor meghatározzuk az új szállítást, amely nyilván rövidebb idővel bír, mint az előző, majd a második lépésnél folytatjuk.

I.10. Szállítási feladat

Ezzel a problémával már találkoztunk a lineáris programozás témakörben. Most ezt a problémát általános Kőnig feladatok sorozatának segítségével próbáljuk megoldani.

Legyenek $T_1, T_2, \ldots, T_i, \ldots, T_m$ a termelőhelyek, $F_1, \ldots, F_j \ldots, F_n$ pedig a fogyasztók. A T_i termelőhely kínálata legyen $\alpha_i \geq 0$ (egész), az F_j fogyasztó kereslete pedig $\beta_j \geq 0$ (egész) számmal megadva. A T_i-ből F_j-be történő, egységnyi termékre vonatkozó szállítás költségét jelölje a $\gamma_{ij} \geq 0$ (egész) szám. Az általánosság megszorítása nélkül most is feltesszük, hogy az összkínálat egyenlő az összkereslettel, azaz

$$\sum_{i=1}^{m} \alpha_i = \sum_{j=1}^{n} \beta_j.$$

Ezek után az ún. primál szállítási feladatot a következőképpen fogalmazhatjuk meg:

A keresletet és a kínálatot kielégítő olyan szállítást kell meghatározni, amelynél a szállítás összköltsége minimális. A feladat matematikai modellje:

Meghatározandó a

$$\begin{aligned} \xi_{ij} &\geq 0, && \forall i,j \text{ -re} \\ \sum_{j=1}^{n} \xi_{ij} &= \alpha_i, && \forall i \text{ -re} \\ \sum_{i=1}^{m} \xi_{ij} &= \beta_j, && \forall j \text{ -re} \end{aligned}$$

feltételek mellett a

$$\sum_{i=1}^{m} \sum_{j=1}^{n} \gamma_{ij} \xi_{ij}$$

minimuma, ahol ξ_{ij} a T_i termelőtől az F_j fogyasztóhoz szállítandó mennyiséget jelöli.

A probléma duálisa:

Míg az előbbi (primál) szállítási feladat esetében a termelőknek kell az árut a fogyasztókhoz elszállítaniuk, addig a duál feladat esetében ezt egy vállalkozó végzi. Ez a vállalkozó a T_i termelőtől μ_i értékért felvásárolja a kínált áru egységét, majd elszállítja az F_j fogyasztóhoz, és ott ν_j értékért eladja. Nyílvánvalóan a μ_i és a ν_j felvásárlási, illetve eladási egységáraknak olyannak kell lenniük, hogy a termelőknek megérje a vállalkozóval való szállíttatás.

A T_i termelőnek az F_j fogyasztóhoz való szállításnál akkor éri meg igénybe venni a vállalkozó szolgáltatását, ha így nem kerül többe a szállítás, mint amikor azt maga végzi, azaz

$$\nu_j - \mu_i \leq \gamma_{ij}.$$

Ezek után a duál feladat: a μ_i és a ν_j árakat úgy kell meghatároznunk, hogy a termelők a szállítást a vállalkozóval végeztessék és a vállalkozó bevétele maximális legyen. Ezen probléma matematikai modellje:

Meghatározandók a $\mu_1, \ldots, \mu_m$, $\nu_1, \ldots, \nu_n$ nemnegatív egész számok úgy, hogy a

$$\nu_j - \mu_i \leq \gamma_{ij} \qquad \forall i, j\text{-re}$$

feltételek mellett a

$$\sum_{j=1}^{n} \beta_j \nu_j - \sum_{i=1}^{m} \alpha_i \mu_i$$

maximális legyen.

A két feladat megoldásaival kapcsolatban megfogalmazhatjuk a következő állítást.

Ha ξ_{ij} kielégíti a primál, μ_i és ν_j pedig a duál feladat feltételeit, akkor a célfüggvényértékek között fennáll a

$$\sum_{i=1}^{m} \sum_{j=1}^{n} \gamma_{ij} \xi_{ij} \geq \sum_{j=1}^{n} \beta_j \nu_j - \sum_{i=1}^{m} \alpha_i \mu_i$$

egyenlőtlenség.

Ugyanis a $\nu_j - \mu_i \leq \gamma_{ij}, i = 1, \ldots, m; j = 1, \ldots, n$ egyenlőtlenségek alapján az

$$\sum_{i=1}^{m} \sum_{j=1}^{n} \gamma_{ij} \xi_{ij} \geq \sum_{i=1}^{m} \sum_{j=1}^{n} (\nu_j - \mu_i)\, \xi_{ij}$$

egyenlőtlenség adódik. Ebből beszorzás, két összegre való bontás, az összegzés sorrendjének felcserélése és kiemelés felhasználásával a

$$\sum_{i=1}^{m}\sum_{j=1}^{n}(\nu_j-\mu_i)\,\xi_{ij}=\sum_{i=1}^{m}\sum_{j=1}^{n}\nu_i\xi_{ij}-\sum_{i=1}^{m}\sum_{j=1}^{n}\mu_i\xi_{ij}=$$
$$=\sum_{j=1}^{n}\nu_j\sum_{i=1}^{m}\xi_{ij}-\sum_{i=1}^{m}\mu_i\sum_{j=1}^{n}\xi_{ij}$$

adódik. Felhasználva az $\sum_{i=1}^{m}\xi_{ij}=\beta_j$ és a $\sum_{j=1}^{n}\xi_{ij}=\alpha_i$ egyenlőségeket kapjuk, hogy

$$\sum_{j=1}^{n}\nu_j\sum_{i=1}^{m}\xi_{ij}-\sum_{i=1}^{m}\mu_i\sum_{j=1}^{n}\xi_{ij}=\sum_{j=1}^{n}\nu_j\beta_j-\sum_{i=1}^{m}\mu_i\alpha_i,$$

tehát

$$\sum_{i=1}^{m}\sum_{j=1}^{n}\gamma_{ij}\xi_{ij}\geq\sum_{j=1}^{n}\nu_j\beta_j-\sum_{i=1}^{m}\mu_i\alpha_i.$$

Mindebből az következik, hogy ha meghatározzuk a primál és duál feladat olyan megoldását, amely mellett az előbbi egyenlőtlenségnél éppen az egyenlőség teljesül, akkor azok optimális megoldásai lesznek az adott feladatnak.

A kérdés tehát az, hogy milyen megoldáspár esetében teljesül az előbbi egyenlőtlenség esetében az egyenlőség, azaz mi lesz az optimalitás kritériuma?

Ha ξ_{ij}, illetve a μ_i és $\nu_j, i=1,\ldots,m; j=1,\ldots,n$ olyan megoldások, amelyekre minden i és j érték mellett teljesülnek a

$$(\gamma_{ij}+\mu_i-\nu_j)\cdot\xi_{ij}=0$$

egyenlőségek, akkor könnyen belátható, hogy az előbbi egyenlőtlenségnél az egyenlőség teljesül, vagyis az ilyen megoldáspár optimális megoldása lesz a primál–duál feladatpárnak.

Kuhn tétele az ilyen megoldáspár létezésére mond ki állítást.

Tétel. *A primál–duál szállítási feladatpár esetében létezik olyan ξ_{ij}, illetve μ_i és $\nu_j, i=1,\ldots,m; j=1,\ldots,n$, hogy a*

$$\sum_{i=1}^{m}\sum_{j=1}^{n}\gamma_{ij}\xi_{ij}=\sum_{j=1}^{n}\nu_j\beta_j-\sum_{i=1}^{m}\mu_i\alpha_i$$

egyenlőség teljesül, azaz mindkét feladatnak létezik optimális megoldása és az optimum értékek egybeesnek.

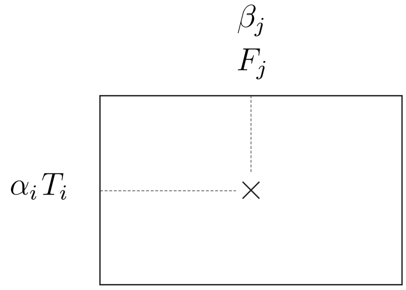
Bizonyítás.

Konstruktív, amely a megoldás meghatározásának menetét is szolgáltatja.

Legyen μ_i és $\nu_j, i = 1, \ldots, m; j = 1, \ldots, n$, a duál feladat egy megoldása. Ekkor minden i-re és j-re meghatározhatók az

$$\varepsilon_{ij} = \gamma_{ij} + \mu_i - \nu_j \geq 0$$

mennyiségek, amelyek felhasználásával megkonstruálhatunk egy általános Kőnig feladatot:



Ebben a feladatban egy (i, j) pozició akkor lehetséges (ezt fejezi ki a táblázatban lévő x szimbólum), ha $\varepsilon_{ij} = 0$.

Ezek megadása után megpróbáljuk megoldani ezt az általános Kőnig feladatot. Ha megoldható, akkor a megoldásával megkaptuk a primál feladat egy lehetséges megoldását, amely egyben optimális is.

Ha viszont nem oldható meg az előbbi feladat, akkor Kőnig tétele szerint kapunk olyan P és R halmazokat, amelyekkel a

$$||P||_\alpha > ||R||_\beta$$

egyenlőtlenség teljesül. Határozzuk meg az

$$\varepsilon = \min_{\substack{T_i \in P \\ F_j \notin R}} \{\varepsilon_{ij}\} > 0$$

számot. Ezen ε segítségével a duálváltozók $\overline{\mu}_i$ és $\overline{\nu}_j$ új értékeit a

$$\overline{\mu}_i = \begin{cases} \mu_i - \varepsilon, & \text{ha } T_i \in P \\ \mu_i, & \text{ha } T_i \notin P \end{cases}$$

$$\overline{\nu}_j = \begin{cases} \nu_j - \varepsilon, & \text{ha } F_j \in R \\ \nu_j, & \text{ha } F_j \notin R \end{cases}$$

összefüggések felhasználásával definiálhatjuk. Belátható, hogy ezek szintén kielégítik a duál feladat feltételeit, azaz

$$\overline{\nu}_j - \overline{\mu}_i \leq \gamma_{ij}, \quad \text{vagy} \quad \overline{\varepsilon}_{ij} = \gamma_{ij} + \overline{\mu}_i - \overline{\nu}_j \geq 0$$

teljesül minden i-re és j-re. Ugyanis:

ha $T_i \in P$ és $F_j \in R$, akkor $\overline{\varepsilon}_{ij} = \gamma_{ij} + \mu_i - \varepsilon - (\nu_j - \varepsilon) = \varepsilon_{ij} = 0$,
ha $T_i \in P$ és $F_j \notin R$, akkor $\overline{\varepsilon}_{ij} = \gamma_{ij} + \mu_i - \varepsilon - \nu_j = \varepsilon_{ij} - \varepsilon \geq 0$ (az ε definiciója miatt),
ha $T_i \notin P$ és $F_j \in R$, akkor $\overline{\varepsilon}_{ij} = \gamma_{ij} + \mu_i - (\nu_i - \varepsilon) = \varepsilon_{ij} + \varepsilon > 0$,
ha pedig $T_i \notin P$ és $F_j \notin R$, akkor $\overline{\varepsilon}_{ij} = \gamma_{ij} + \mu_i - \nu_j = \varepsilon_{ij} \geq 0$.

Az új duálváltozókhoz tartozó célfüggvény értéke növekedett:

$$\sum_{j=1}^{n} \beta_j \overline{\nu}_j - \sum_{i=1}^{m} \alpha_i \overline{\mu}_i =$$

$$= \sum_{j:F_j \in R} \beta_j(\nu_j - \varepsilon) + \sum_{j:F_j \notin R} \beta_j \nu_j - \sum_{i:T_i \in P} \alpha_i(\mu_i - \varepsilon) - \sum_{i:T_i \notin P} \alpha_i \mu_i =$$

$$= \sum_{j=1}^{n} \beta_j \nu_j - \sum_{i=1}^{m} \alpha_i \mu_i + \varepsilon \left(\sum_{i:T_i \in P} \alpha_i - \sum_{j:F_j \in R} \beta_j \right) >$$

$$> \sum_{j=1}^{n} \beta_j \nu_j - \sum_{i=1}^{m} \alpha_i \mu_i,$$

ugyanis $\varepsilon > 0$ és $||P||_\alpha > ||R||_\beta$, azaz

$$\sum_{i:T_i \in P} \alpha_i > \sum_{j:F_j \in R} \beta_j.$$

Ezekkel az új duálváltozókkal megismételjük az előbbi eljárást. Mivel $\varepsilon > 0$ egész és a duál célfüggvény felülről korlátos, ezért az eljárás véges sok lépésben az optimális megoldások meghatározásával véget ér.

Ezen eljárás első lépésében a kezdeti duál-változókat kellett meghatároznunk. Leggyakrabban a

$$\mu_i = \min_j \gamma_{ij}$$
$$\nu_j = \min_i \{\gamma_{ij} + \mu_i\}$$

egyenlőségek alapján történik ezek megválasztása.

Az $\varepsilon_{ij} = \gamma_{ij} + \mu_i - \nu_j$ értékek az előbbi duálváltozók meghatározása nélkül is közvetlenül nyerhetők a γ_{ij} mátrixból sor–oszlop redukcióval. (A sorredukció során a táblázat minden sorából kivonjuk a sor legkisebb elemét, majd az így adódó táblázatban oszlopredukciót hajtunk végre, azaz minden oszlopból kivonjuk az oszlop legkisebb elemét is.

Erre az ε_{ij}-ket tartalmazó táblázatra többek között a Kőnig feladat megkonstruálásakor lesz szükségünk.)

Ezt követte aztán az általános Kőnig feladat megkonstruálása és megoldása. Ha a Kőnig feladat megoldható volt, akkor vége lett az eljárásnak, különben a P és az R halmazok alapján meg kellett határoznunk az új duál változókat. Ezt megtehetjük úgy is, hogy a P és az R halmazok meghatározása után elkészítjük az ε_{ij}-ket tartalmazó táblázat „lefedését" ($\overline{P}$-beli sorokat és az R-beli oszlopokat fedjük le), majd meghatározzuk a le nem fedett elemek minimumát, ami éppen az ε értéke lesz.

Ezután az új ε_{ij}-ket tartalmazó táblázatot határozzuk meg úgy, hogy a régiben a fedetlen helyeken lévő értékeket ε-nal csökkentjük, a kétszeresen fedett helyeken pedig ε-nal növelünk minden értéket; az egyszeresen fedett értékeket változatlanul hagyjuk. Az új $\overline{\varepsilon}_{ij}$ táblázatból újabb általános Kőnig feladatot konstruálunk és ezzel megismételjük az eljárást. □

Minden egyes Kőnig feladat megoldását egy kiinduló szállítással kell kezdeni, amit É–NY sarok módszerrel határozhatunk meg. Megjegyezzük ugyanakkor azt is, hogy az algoritmus során meghatározott új Kőnig feladat induló szállításaként az előző feladat maximális szállítását is használhatjuk, ugyanis a kvalifikációs táblában csak olyan helyen történt változás, ahol az előző Kőnig feladatban nem volt folyam.

II. Készletgazdálkodási modellek

II.1. Bevezetés

Készleteket rendszerint azért tartunk, hogy valamely szükségletet, igényt kielégítsünk. A szóban forgó anyag, cikk iránti igény, kereslet a készlet fogyását idézi elő. Gondoskodnunk kell tehát időben a raktárkészlet pótlásáról, feltöltéséről. A megrendelés feladásától a megrendelt mennyiség raktárba érkezéséig egy bizonyos idő eltelik, amit utánpótlási, vagy röviden pótlási időnek nevezünk. A pótlási idő alatt is történhet kivét a raktárból, emiatt a megrendelési időpontokat tehát úgy kell megválasztani, hogy a raktárkészlet az utánpótlási idő alatt is fedezze a szükségletet.

Az árubeérkezés és az árukivét, más szóval a beáramlás és a kiáramlás együttesen meghatározzák a raktárkészlet időbeli alakulását, amelyet egy koordináta rendszerben is ábrázolhatunk. A vízszintes tengelyen a t időt, a függőlegesen pedig az $y(t)$ raktárkészletet szokás ábrázolni, mégpedig ez utóbbit abban az egységben kifejezve, amely a vizsgált árucikk természetéből következik.

A készletgazdálkodás az újratermelési folyamathoz kapcsolódó keresleti, megrendelési, kivitelezési, stb. elemekből álló olyan szabályozási folyamat, ahol a gazdálkodó arra törekszik, hogy se több, se kevesebb készlet ne keletkezzen, mint amire a termelés és forgalom zavartalan működéséhez szükség van. Más szavakkal a készletgazdálkodás egy olyan szabályozási folyamat, ahol a különböző készletformák mennyiségének és egymáshoz való arányának szabályozásával foglalkoznak. A készletgazdálkodási, vagyis a készletet alakító gazdasági tevékenység során a legfontosabb kérdések egyike az, hogy adott körülmények között, adott időpillanatban, vagy időszakban mennyi az a minimális készlet, amely a termelés és forgalom zavartalan működéséhez szükséges.

A készlettartás, a készlet pótlása, a beszerzési lehetőségek mérlegelése idő- és költségigényes, de ugyanúgy gazdasági konzekvenciái vannak az igények ki nem elégítésének is. A készletgazdálkodással kapcsolatos költségek jelentős szerepet játszanak a készletmodellekben. Ezek három alapvető csoportba sorolhatók:

a, Az utánpótlással, a raktárfeltöltéssel kapcsolatos költségek, amiket beszerzési, illetve előállítási költségnek szokás nevezni.
b, A raktár fenntartásának a költsége, a raktárkészletben lekötött eszközök költsége, a kamat, az eszközlekötési járulék, a készlet elévüléssel, romlásával járó veszteségek, stb., amelyeket gyűjtőnéven raktározási költségnek nevezünk.
c, A raktárhiány okozta termeléskiesés, pótlólagos beszerzéssel együttjáró többletköltség, vagy a hiány miatti nyereségkiesés, stb., ami gyűjtőnéven a hiányköltség.

Ezen költségek számszerű nagyságának, függvényalakjának a meghatározása nem egyszerű feladat, tény azonban, hogy e meggondolások hiányában optimális készletezési eljárásról csak nagyon kivételes és leszűkített esetekben beszélhetünk. A figyelembe veendő költségeket az is meghatározza, hogy milyen gazdasági célt, gazdasági eredményt kívánunk megvalósítani a készletezéssel. Előfordulhat például, hogy mindenképpen 100%-os szükségletkielégítést kell elérnünk. Nyilvánvaló, hogy ebben az esetben nem kell foglalkoznunk sem a hiányköltség számszerű nagyságával, sem azzal, hogy a raktárkészlet időbeli alakulásával az milyen függvényszerű kapcsolatban áll. Bár meggondolásainkban a hiányköltség ekkor látszólag nem szerepel, de látni fogjuk, hogy bizonyos modellekben valójában nagyságrendileg nagyobb minden más tekintetbe vett költségfajtánál.

A beszerzési költségek két egyszerű fajtája a tétel nagyságától független és a tétel nagyságával egyenes arányban lévő költség. Az előbbire példa a tétel átvizsgálási költsége, vagy a sorozatgyártásnál a sorozat beindítási költsége, stb. Az utóbbi egy termék anyagköltsége, vagy az áru egy egységének beszerzési ára lehet.

A raktározási költséget többnyire a készlet raktáron eltöltött idejével arányos költségként definiáljuk, mégpedig a készlet egységnyi mennyiségének, vagy az egységnyi értékének időegységre eső költségét adjuk meg. Ekkor egy $[0,T]$ időintervallum (pl. negyedév, félév, stb.) raktározási költségét úgy számoljuk ki, hogy a raktárkészlet (illetve a raktárkészlet értékének) időbeli alakulását megadó függvénynek az időtengely $[0,T]$ intervalluma által meghatározott területét kiszámítjuk és megszorozzuk az idő és áruegységre (idő és értékegységre) eső raktározási költséggel. Hogy ezt valóban így célszerű kiszámítani, azt a következőképpen láthatjuk be. A vizsgált $[0,T]$ időtartam alatti raktározási időt megkapjuk, ha a készlet (illetve értékének) minden egységéről megállapítjuk, hogy mennyi ideig tartózkodott a raktárban, azaz megállapítjuk, hogy mennyi idő telt el a raktárba érkezése időpontjától a T időpontig, illetve a kivét időpontjáig. Ezeket az időtartamokat összegezve megkapjuk az összes raktározási időt. Ennél az eljárásnál célszerűbb a következőképp okoskodnunk. Osszuk be az időtengelyt egységekre és nézzük meg, hogy mindegyik időegység alatt mennyi áru volt a raktáron. E mennyiségeknek az időegységgel való szorzata megadja az adott időegységre eső raktározási időt, amelynek összege lesz a T időpontig a raktározási összidő. Ha az időtengely felbontását minden határon túl fimomítjuk és $y(t)$ jelenti a raktárkészlet pillanatnyi nagyságát, akkor a T időszakra eső összes raktározási idő nem más, mint az $\int_0^T y(t)dt$ integrál értéke.

Hasonlóképpen határozható meg a hiányköltség is, ha az az egységnyi hiány időegységre eső költségeként van megadva. Valójában a hiányt leíró függvény a raktárkészlet alakulását kifejező függvénynek az időtengely alatti folytatása. Így a készletgörbe negatív ágának az időtengellyel bezárt területét megszorozva a hiányköltség idő– és áruegységre eső értékével megkapjuk a hiányból származó összköltséget.

A raktározás tárgyát képező anyag, cikk iránti kereslet, szükséglet, tehát az output folyamata, valamint az elérendő cél – például a teljes, vagy csak a részleges igénykielégítés – és az utánpótlási, feltöltési lehetőségek, tehát az input folyamata együttesen határozzák meg a készletáramlás lehetséges alakulását, amelyeket a készletkezelési modellek írják le. Ezen készletezési modellek alapján – a kitűzött cél szem előtt tartásával – meghozzuk azokat a döntéseket, amelyek a lehetséges készletezési politikák, készletezési eljárások (például mikor és mennyit rendeljünk) közül azt a készletezési politikát alakítják ki, amelyik a tekintetbe vett körülmények között a célt a legjobban megvalósítja. Ezt az eljárást optimális készletezési eljárásnak fogjuk nevezni.

A magyarországi termelési gyakorlatban meglévő ún. előszállításos eljárás meglehetősen nagy készletek tartását igényli. Ez még néhány évvel ezelőtt is jóval meghaladta az éves nemzeti jövedelem felét. A modern gazdaságok esetében már jóval kisebb, de még mindig nagy ez az arány, hiszen még ezekben az országokban is az éves nemzeti jövedelem harmada közelében van a készletek összértéke. Ez egy óriási „holt tőkét” jelent egy gazdaság számára, ezért nem érdektelen, ha az optimális készletezési eljárások megvalósításával bármilyen mértékű, akár csak néhány százalékos csökkenést is el tudunk érni.

A készletezési modellek osztályozási szempontjai igen különbözőek. Egyik alapvető osztályozási elv az, hogy mind a beáramlással, mind a kiáramlással, valamint a költségekkel kapcsolatos minden információ megadható-e előre teljes bizonyossággal, vagy pedig ezek között szerepelnek olyanok is, amelyekre csupán statisztikai törvényszerűségek állnak fenn. Az előbbi esetben determinisztikus, míg az utóbbiban sztochasztikus modellről beszélünk. A véletlen (sztochasztikus) jelenségekre vonatkozó ismereteink, ítéleteink csupán valószínűségi jellegűek. Nagyszámú megfigyelés, kísérleti tapasztalat s a kísérletnek a matematikai statisztika és a valószínűségszámítás törvényein alapuló értékelése teszi lehetővé a törvényszerűségek megismerését, feltárását.

Az osztályozás egy másik szempontja lehet az, hogy az optimális készletezési politika egy vagy több időszakasz egymásutánjára vonatkozik-e. Az egy időszakaszt átfogó modellt statikus modellnek, a döntés egymásutánjaira vonatkozót pedig dinamikus modellnek nevezzük.

Szokásos osztályozási elv még az is, amikor aszerint teszünk különbséget az egyes modellek között, hogy az idő, illetve a döntési változó (amely rendszerint a raktárkészlettel kapcsolatos) folytonos vagy diszkrét értéket vesz-e fel. Ilyen értelemben beszélhetünk folytonos időparaméterű diszkrét, folytonos időparaméterű folytonos készletmodellről, illetve diszkrét időparaméterű folytonos és diszkrét időparaméterű diszkrét modellről.

Az előbbieken kívül még sok más osztályozási elv is ismeretes a szakirodalomban. Ilyen például a költségtipusok, vagy pedig a hiány kezelése szerinti csoportosítás. Készlethiány esetén például kétféle modell lehetséges. Az egyiknél a beérkező készletből pótolják a korábban keletkezett hiányt, a másiknál viszont a kielégítetlen kereslet elvész, tehát a tervezettnél nagyobb zárókészlettel kell számolni.

Az előbbiek miatt nem meglepő, hogy a készletezési modellek sok fajtája és tipusa alakult ki, már csak a fellépő véletlen törvényszerűségek miatt is, nem beszélve a gyakorlati élet bonyolult szituációiról, a költségek, célok különbözőségeiről, stb. Tudnunk kell azonban, hogy a készletgazdálkodási modell is – mint minden modell – a sokrétű valóságnak csak néhány jellemzőjét ragadhatja meg, hogy aztán a matematikai módszerek és logikai következtetések útján olyan mennyiségi összefüggéseket tárjon fel, amelyek elemzése, értékelése alapján a döntésre sor kerülhet.

A következőkben néhány alapvető statikus és sztochasztikus modellt mutatunk be, és ezek felhasználásával megadunk néhány optimális eljárást is az adott modellel modellezhető készletezési - raktározási probléma megoldására.

II.2. Determinisztikus modellek

II.2.1. Optimális tételnagyság (sorozatnagyság) modell

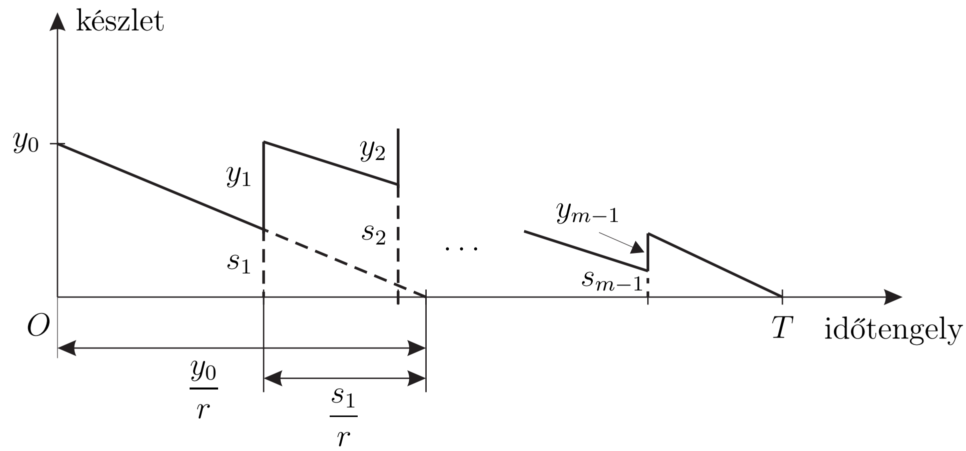
Valamely árucikkből, anyagból egy T időszak alatt összesen R egységnyire van szükség, mégpedig időegységenként mindig r-re. A kivét, azaz a raktárból való kiáramlás tehát időben egyenletes, hiány pedig nem engedett meg.

A modellben a következő költségeket vesszük figyelembe:

- Egy tétel beszerzésének (előállításának) tételben foglalt mennyiségtől független költsége, amit a c_1 (Ft) fog jelölni.
- A szóban forgó anyag, cikk egy egységének időegységre eső raktározási költsége; ezt jelölje c_2 (Ft/db/idő).

Célunk olyan raktározási politika kialakítása, amely egyfelől biztosítja az időegységenkénti r árumennyiség meglétét, másfelől a beszerzéssel és raktározással kapcsolatos költségeket minimalizálja.

Induljunk ki abból, hogy a feltöltések szabálytalan időközönként, szabálytalan mennyiségekben történnek. Tegyük fel a T idő alatt összesen m feltöltést hajtanak végre. Emiatt a rendelési költség $m \cdot c_1$ lesz, az egyes tételek nagyságától függetlenül. Ehhez fog még hozzáadódni a raktározási költség, amely megegyezik az összraktározási idő és a c_2 szorzatával. A 0 időpontban álljon y_0 készlet a rendelkezésünkre, továbbá még $m-1$ alkalommal $y_1, y_2, \ldots, y_{m-1}$ készlet érkezzen a raktárba. Az egyes készletek megérkezésekor legyenek $s_1, s_2, \ldots, s_{m-1}$ a raktáron lévő mennyiségek ($s_0 = 0, s_m = 0$). Az idő függvényében a készletszintet egy „fűrészfoggörbe” jelzi, a raktározási időt az ez alatti terület adja meg:



Az „egyenes” darabok meredeksége állandó (r), ami az egységnyi idő alatt felhasznált mennyiségeket jelenti. A görbe alatti területet úgy kapjuk meg, hogy az $y_i + s_i$, $\frac{y_i+s_i}{r}$ befogókkal rendelkező derékszögű háromszögek területösszegéből kivonjuk az s_{i+1}, $\frac{s_{i+1}}{r}$ befogókkal rendelkező derékszögű háromszögek területösszegét. Itt $i = 0, 1, \ldots, m-1$. A teljes terület:

$$t = \frac{y_0^2}{2r} - \frac{s_1^2}{2r} + \frac{(s_1+y_1)^2}{2r} - \frac{s_2^2}{2r} + \cdots + \frac{(s_{m-1}+y_{m-1})^2}{2r} = \\ = \frac{1}{2r}\sum_{i=0}^{m-1} y_i^2 + \frac{1}{r}\sum_{i=1}^{m-1} s_i y_i.$$

Így a teljes költség:

$$K_T(y_0, y_1, \ldots, y_{m-1}; s_1, s_2, \ldots, s_{m-1}) = m \cdot c_1 + t \cdot c_2 = \\ = m \cdot c_1 + c_2 \cdot \frac{1}{2r}\sum_{i=0}^{m-1} y_i^2 + c_2 \cdot \frac{1}{r}\sum_{i=1}^{m-1} s_i y_i.$$

Ez a költség csökkenthető, ha akkor történik az utánpótlás, amikor a raktárkészlet 0-vá válik, azaz ha $s_i = 0, i = 1, 2, \ldots, m-1$. Így

$$K_T(y_0, y_1, \ldots, y_{m-1}; s_1, s_2, \ldots, s_{m-1}) = m \cdot c_1 + c_2 \frac{1}{2r}\sum_{i=0}^{m-1} y_i^2.$$

Minthogy $\sum_{i=0}^{m-1} y_i = R$, ezért az $y_i - k$ számtani közepe megegyezik $\frac{R}{m}$-mel. Legyen most $x_i = y_i - \frac{R}{m}$, így $y_i = x_i + \frac{R}{m}$. Az $R = \sum_{i=0}^{m-1} y_i$ egyenlőségből adódik, hogy $\sum_{i=0}^{m-1} x_i = 0$. Ezt felhasználva kapjuk, hogy

$$\sum_{i=0}^{m-1} y_i^2 = \sum_{i=0}^{m-1}\left(\frac{R}{m} + x_i\right)^2 = \frac{R^2}{m} + \sum_{i=0}^{m-1} x_i^2.$$

Ebből azt kapjuk, hogy a $\sum_{i=0}^{m-1} y_i^2$ akkor minimális, ha $\sum_{i=0}^{m-1} x_i^2 = 0$, vagyis ha $x_i = 0$ minden $i = 0, \ldots, m-1$ esetén. Ez azonban egyúttal azt is jelenti, hogy minden raktárfeltöltés alkalmával $y_i = \frac{R}{m} = q, i = 0, \ldots, m-1$ mennyiséggel növeljük a raktárkészletet. Így azt kaptuk, hogy ha éppen akkor pótoljuk a raktárkészletet, amikor az 0-vá vált és minden egyes pótlás alkalmával éppen $\frac{R}{m} = q$ mennyiséggel növeljük azt, akkor a költség

$$K_T(m) = m \cdot c_1 + \frac{c_2}{2r} \cdot \frac{R^2}{m}$$

lesz, amely kisebb lesz, mint bármelyik másik raktárfeltöltési eljárásé.

Ez a $K_T(m)$ költségfüggvény mostmár csak a beszerzések számának (m) lesz a függvénye, így a problémánkat mint egy egyváltozós függvény minimumát adó m értékének meghatározási problémáját fogalmazhatjuk meg. A szélsőértékszámításnál tanultak alapján deriválnunk kell a $K_T(m)$ függvényt, majd azt 0-val kell egyenlővé tennünk, és meg kell oldanunk az úgy adódó egyenletet:

$$\frac{d}{d_m} K_T(m) = c_1 - \frac{c_2}{2r} R^2 \cdot \frac{1}{m^2}$$

$$c_1 - \frac{c_2}{2r} R^2 \cdot \frac{1}{m^2} = 0 \quad \Longrightarrow \quad m_0 = R\sqrt{\frac{c_2}{2rc_1}}$$

(A

$$K_T''(m_0) = \left.\frac{c_2 \cdot R^2}{rm^3}\right|_{m=m_0} = \frac{c_2 R^2}{r \cdot R^3 \frac{c_2}{2rc_1}\sqrt{\frac{c_2}{2rc_1}}} > 0$$

miatt m_0 valóban a minimumot adó érték lesz.)

Így az optimális tételnagyság a

$$q_0 = \frac{R}{m_0} = \sqrt{\frac{2rc_1}{c_2}}$$

lesz. Ebből az optimális rendelési időközre a

$$t_0 = \frac{q_0}{r} = \sqrt{\frac{2c_1}{rc_2}},$$

míg a minimális költségre a

$$K_0 = \sqrt{2RTc_1c_2} = \sqrt{2rT^2c_1c_2}$$

adódik.

Összefoglalva eddigi eredményeinket megállapíthatjuk, hogy ha ismert egy T időtartam össz-szükséglete, az igénykielégítés konstans intenzitású, két költségtényezőt veszünk figyelembe (beszerzés, raktározás), a beáramlási folyamat determinisztikus, akkor az optimális raktárfeltöltési eljárás az, ha szabályos időközönként az optimális tételnagyságra töltjük fel az éppen 0-vá váló készletet.

Megjegyezzük még, hogy ha a beszerzési költséget c_1 helyett $by + c_1$-nek választottuk volna, akkor a régi költségfüggvény $b \cdot R$ értékkel változott volna, így az optimális beszerzési politika változatlan maradt volna.

Végül belátható az is, hogy ha a nyereség maximalizálását tűztük volna ki célul, akkor az optimum érték éppen ott adódott volna, ahol a költségünk minimális.

Az optimális eljárás megkeresése az esetek túlnyomó részében nem ilyen egyszerű és sokszor csak közelítő algoritmussal határozható meg. A közelítő megoldások is hasznosak, de lényegesen mélyebb matematikai módszereket és meggondolásokat igényelnek, mint az előbbi eljárás. Ugyanakkor az előbbi módszer sok szempontból jellegzetes. Nevezetesen:

(i) Meg kell ismerkedni a beáramlás és a kiáramlás sajátosságaival, a vizsgálatba vonható költségekkel.
(ii) Meg kell határoznunk az elérni kívánt célt, vagy célokat.
(iii) Matematikai összefüggések segítségével felírjuk a feltételeket és a célfüggvényt.
(iv) A lehetséges eljárások közül kiválasztjuk az optimálisat.
(v) Meghatározzuk az optimális eljárás azon paramétereit, amelyek az eljáráshoz tartozó célfüggvényt optimalizálják.

A készletgazdálkodási modellek az esetek túlnyomó részében valamely optimumszámítási feladatra vezetnek, gyakran lineáris illetve nemlineáris programozási feladatra.

II.2.2. Optimális tételnagyság modell az önköltségi beszerzési árral arányos raktározási költséggel

Legyen most is a $[0, T]$ időszakban R az össz-szükséglet, amelyből időegységenként r mennyiséget használnak fel. Egy y tétel beszerzési (előállítási) költsége legyen $by + c_1$, ahol b a tétel egy egységének a beszerzési ára, c_1 pedig a tételben foglalt mennyiségtől független költség. A raktározási költséget jelölje w, amely a raktárkészlet 1 Ft értékének egy időegységre eső költsége lesz Ft/Ft/idő mértékegységgel.

A rendelési és utánrendelési politika tőlünk függ. Nem nehéz belátni, hogy az optimális készletezési eljárás most is az, hogy szabályos időközönként mindig ugyanarra a szintre töltjük fel a raktárkészletet. Ehhez az optimális eljáráshoz tartozó $K_T(q)$ összköltség kiszámítási módja azonban módosul, minthogy a raktározási költség most nem áru- és időegységre, hanem a beszerzési költség egy egységének időegységre eső részeként van megadva.

Egy q nagyságú tétel beszerzési költsége $bq + c_1$ Ft, ami $\frac{q}{r}$ idő alatt 0 Ft értékűre fog lecsökkenni. Így ehhez a tételnagysághoz $\frac{bq+c_1+0}{2} \cdot \frac{q}{r}$ átlagos raktározási érték tartozik, aminek a w-szerese adja meg egy tétel raktározási költségét. Abban az esetben, ha m alkalommal kerül sor raktárfeltöltésre, azaz $R = m \cdot q$, akkor az optimális eljáráshoz tartozó összköltség a

$$K_T(q) = mc_1 + m \cdot \frac{q}{r} \cdot \frac{bq + c_1}{2} w + bR$$

minimuma lesz. Az $m = \frac{rT}{q}$ helyettesítés és a T-vel való osztás után megkapjuk az időegységre eső költséget:

$$k(q) = \frac{K_T(q)}{T} = \frac{rc_1}{q} + \frac{1}{2}bwq + \frac{1}{2}c_1 w + br.$$

Könnyen belátható, hogy a $k(q)$ függvény akkor veszi fel a minimumát, ha $q_0 = \sqrt{2r\frac{c_1}{bw}} = \sqrt{2 \cdot \frac{R}{T} \cdot \frac{c_1}{bw}}$. Ennek megfelelően:

$$t_0 = \frac{q_0}{r} = \sqrt{2\frac{c_1}{rbw}} = \sqrt{2\frac{T}{R}\frac{c_1}{bw}}$$

$$k(q_0) = \sqrt{2rbc_1 w} + br + \frac{1}{2}c_1 w, \quad \text{vagy}$$

$$K_T(q_0) = \sqrt{2RTbc_1 w} + bR + \frac{T}{2}c_1 w.$$

II.2.3. Optimális tételnagyság modell árengedmény esetén

Ismét tételezzük fel, hogy a T idő alatti össz-szükséglet R, amelyet $r = \frac{R}{T}$ időegységre eső intenzitással elégítünk ki az egész T időtartam alatt. Jelölje Q azt a mennyiséget, amely felett árengedményt adnak, vagyis egy y tétel beszerzési költsége

$$b_1 y + c_1, \quad \text{ha } 0 < y < Q$$

$$b_2 y + c_1, \quad \text{ha } Q \le y, \quad \text{ahol } b_1 > b_2.$$

A készlettartás időegységre és forintra eső költsége legyen ismét w. A feladat most is a legkisebb összköltséggel járó beszerzési-raktározási politika meghatározása.

Ez esetben a következőképpen fogunk eljárni.

Az előző modell esetében már megállapítottuk az optimális eljárást, amely szerint $t_0 = \sqrt{2\frac{c_1}{rbw}} = \sqrt{2\frac{T}{R}\frac{c_1}{bw}}$ időközönként kell $q_0 = \sqrt{2r\frac{c_1}{bw}} = \sqrt{2 \cdot \frac{R}{T} \cdot \frac{c_1}{bw}}$ nagyságú tételt beszereznünk, és ekkor $K_T(q_0) = \sqrt{2RTbc_1 w} + bR + \frac{T}{2}c_1 w$ lesz a minimális összköltségünk. Határozzuk meg ezért a b_2 egységárhoz tartozó optimális tételnagyságot, amit jelöljön q_2. Ha $q_2 \ge Q$ teljesül, akkor a q_2 valóban az optimális tételnagyság lesz. Amennyiben $q_2 < Q$, akkor ez nem lehet az optimális tételnagyság, hiszen árengedményt ekkor nem kapunk. Ezért ki kell számítanunk a b_1 árhoz tartozó q_1 tételnagyságot, amelyre nyilvánvalóan igaz, hogy $q_1 < q_2 < Q$. Ezt követően kiszámoljuk a q_1-hez és a Q-hoz tartozó időegységre jutó összköltséget, azaz a

$$k(q_1) = \frac{rc_1}{q_1} + \frac{1}{2}c_1 w + \frac{1}{2}b_1 w q_1 + b_1 r,$$

illetve a

$$k(Q) = \frac{rc_1}{Q} + \frac{1}{2}c_1 w + \frac{1}{2}b_2 w Q + b_2 r$$

mennyiséget. A $b_1 < b_2$, $q_1 < q_2 < Q$ relációk miatt

$$\frac{rc_1}{q_1} > \frac{rc_1}{Q}, \qquad b_1 r > b_2 r.$$

Ugyanakkor azonban az $\frac{1}{2}b_1 w q_1$ és az $\frac{1}{2}b_2 w Q$ mennyiségek között bármelyik reláció fennálhat. Emiatt újból két esetet fogunk megkülönböztetni: ha $k(q_1) \geq k(Q)$, akkor Q nagyságú tételt fogunk beszerezni, ha viszont $k(q_1) < k(Q)$, akkor q_1 lesz az optimális tételnagyság.

II.2.4. Optimális tételnagyság modell hiány megengedésével

Tekintsük ismét az optimális tételnagyság modellt azzal a különbséggel, hogy most nem törekszünk az $R = rT$ szükséglet T idő alatti teljes kielégítésére, vagyis a raktárhiány megengedett. Jelöljük a hiány árumennyiség és időegységre eső költségét c_3-mal, amelynek Ft/db/idő lesz az egysége. Foglaljuk most össze az adott feltételeket:

- a $[0, T]$ időszakasz össz-szükséglete R
- minden időegység alatt pontosan r egységre van szükségünk
- c_1 Ft jelöli a tétel nagyságától független beszerzési költséget
- a készlet egységének időegységre eső raktározási költsége c_2 Ft.
- a hiány időegységre és áruegységre eső költsége c_3 Ft.

Ezen modell esetén ha az időegységenkénti r intenzitású felhasználás során elfogy a raktárkészlet és nem érkezik pótlás, akkor hiány lép fel, amely addig áll fenn, amíg a beérkező készlet meg nem haladja a beérkezés időpontjában meglévő hiányt. Ez esetben is csak a beérkező készlet és a raktárhiány különbségének megfelelő mértékűre emelkedik a raktárkészlet. Nem nehéz belátni, hogy az optimális készletezési eljárás most is a szabályos „fűrészfog"-görbével ábrázolható raktárfeltöltési politika lesz, csakhogy most nem az egész R igényt elégítjük ki, hanem annak csak egy részét.

Az x tengely alatt elhelyezkedő derékszögű háromszögek területösszege a hiánnyal kapcsolatos raktározási időt jelenti. Az optimális készletezési eljárás tehát most is az, hogy egyenlő időközönként ugyanarra a szintre töltjük fel a raktárkészletet, csakhogy most nem q, hanem valamely q-nál kisebb s értékre. Az s és q aránya a költségtényezők egymáshoz való arányától függ. Meghatározandó az s és q azon mennyisége, amely mellett az összköltség, mint e két mennyiség függvénye minimális. Jelöljük a minimumot adó értékeket s_0-lal és q_0-lal. A költségfüggvényt a következő módon számolhatjuk ki. Ha q egységet szerzünk be, akkor $\frac{q}{r}$ időközönként $\frac{rT}{q}$ számú beszerzéssel a teljes igényt ki tudjuk elégíteni. Most azonban $\frac{q}{r}$ időközönként csupán $s < q$ mennyiséget szerzünk be, amely csak $\frac{s}{r}$ ideig fedezi a szükségletet. Ebből nyilvánvaló, hogy készlet az $\frac{s}{r}$ időtartamig, hiány pedig az $\left(\frac{s}{r}, \frac{q}{r}\right)$ időintervallumban lesz. Ezek figyelembevételével

$$K(q,s) = \frac{rT}{q}\left[c_1 + \frac{s^2}{2r}c_2 + \frac{(q-s)^2}{2r}c_3\right]$$

lesz a költségfüggvény. Meg kell tehát határozni ezen kétváltozós függvény minimum helyét. Belátható, hogy

$$q_0 = \sqrt{2r\frac{c_1}{c_2}} \cdot \sqrt{\frac{c_2+c_3}{c_3}} = \sqrt{2\frac{Rc_1}{Tc_2}} \cdot \sqrt{\frac{c_2+c_3}{c_3}}$$

$$s_0 = \sqrt{2r\frac{c_1}{c_2}} \cdot \sqrt{\frac{c_3}{c_2+c_3}} = \sqrt{2\frac{Rc_1}{Tc_2}} \cdot \sqrt{\frac{c_3}{c_2+c_3}}$$

szolgáltatja az optimális megoldást. Raktárfeltöltésre $\frac{q_0}{r}$ időközönként kerül sor; ezt az időtartamot jelölje $t_0(q_0, s_0)$, amelynek értéke

$$t_0(q_0,s_0) = \sqrt{\frac{2c_1}{rc_2}}\sqrt{\frac{c_2+c_3}{c_3}} = \sqrt{2\frac{Tc_1}{Rc_2}} \cdot \sqrt{\frac{c_2+c_3}{c_3}}.$$

Az egész időtartamhoz tartozó összköltség:

$$K(q_0,s_0) = \sqrt{2RTc_1c_2}\sqrt{\frac{c_3}{c_2+c_3}} = T \cdot \sqrt{2rc_1c_2} \cdot \sqrt{\frac{c_3}{c_2+c_3}}.$$

Ha a c_2 raktározási költségnél lényegesen nagyobb a c_3 hiányköltség, akkor a $\sqrt{\frac{c_3}{c_2+c_3}} \approx 1$, így lényegében az első modellünket nyerjük vissza. Ilyenkor tehát hiány nem engedhető meg, vagyis $s_0 = q_0$.

II.3. Sztochasztikus készletmodellek

II.3.1. A sztochasztikus modellekről általában

A készletezési problémák túlnyomó többsége sztochasztikus modellre vezet. A kiáramlási folyamat például sok esetben a kereslet függvénye, ami pedig legtöbbször a véletlentől függ. Az alkatrészek tartalékolási problémái is különböző valószínűségszámítási meggondolásokat igényelnek. A szállítási késedelmek és maga az ún. előszállításos rendszer, amikor a megrendelt cikk egy meghatározott időintervallumon belül meg nem adott időpontokban és résznagyságokban érkezik be, sztochasztikus törvényszerűségeket követ.

A valószínűségszámítás és a matematikai statisztika nagy segítséget nyújt olyan esetekben is, amikor valójában sem kísérletre, sem többszörös megfigyelésre nincs lehetőség, de a vizsgált jelenségekkel azonos tipusú jelenségeket korábban már megfigyeltük.

Sztochasztikus modellekkel gyakran lehet bonyolult, bár alapjában véve determinisztikus jelenségeket is modellezni. Például egy nagy kikötő hajóforgalmát – noha az szigorú menetrend szerint bonyolódik le – igen jól lehet Poisson-folyamattal leírni. Talán ez a tény is megnyugtathatja azokat – a valójában érthetően – aggodalmaskodókat, akik a sztochasztikus modellek gyakorlati alkalmazhatóságában kételkednek, mondván, hogy soha nincsen elég tapasztalatunk, a jelenségek sohasem ismételhetők meg akárhányszor ugyanolyan körülnények között, vagy hogy miért mindig csak azzal a néhány valószínűségeloszlással dolgozunk, amelyek meglehetősen egyszerűen kezelhetők, holott a valóság sokkal bonyolultabb. A mérési adatok esetében a pontosabb műszerekkel újabb és újabb pontatlanságok mutathatók ki. A gyakorlat azonban igen jól boldogul a „durvább” mérési adatokkal is. Az elkövetett hiba becsülhető, s ez gyakorlatilag kielégítő.

A sztochasztikus modellek felépítése, logikája a legtöbb esetben ugyanaz, mint a determinisztikusoké. Ezeknél is meg kell ismernünk a beáramlás és kiáramlás törvényszerűségeit, meg kell határoznunk azt a célt, vagy célokat, amelyeket el akarunk érni, ha pedig költségoptimumra, vagy nyereségoptimumra törekszünk, akkor a megfelelő költségtényezőt is meg kell adnunk. A beáramlási és kiáramlási folyamatokban fellépő véletlen tényezők valószínűségeloszlásainak meghatározása a megfigyelési adatok alapján matematikai statisztikai módszerekkel valósítható meg.

Ebben a fejezetben a leggyakrabban előforduló sztochasztikus modellekkel foglalkozunk, ügyelve a valószínűségszámítási módszerek és tételek alkalmazhatóságára. Előbb az egyszerű megbízhatósági tipusú sztochasztikus készletmodelleket mutatjuk be, majd a véletlen ütemezésű részszállítmányok modelljét tárgyaljuk, végül az olyan statikus sztochasztikus készletmodelleket vizsgáljuk, amelyekben költségtényezők is szerepelnek.

II.3.2. Egyszerű megbízhatósági tipusú statikus sztochasztikus készletmodellek

Ezeknél a készletmodelleknél nem szerepelnek költségtényezők, csupán egy előre adott biztonság, s a különböző véletlen jelenségek valószínűségi törvényeinek ismeretében olyan raktárfeltöltési eljárást kell követnünk, amely az előre megadott biztonsággal kielégíti az igényt, a szükségletet.

1. modell. Egy vállalat (nevezzük B-nek) időegységenkénti felhasználása valamely anyagból r egység. A $[0, T]$ időszak össz-szükségletét megrendeli egy másik (nevezzük A-nak) vállalattól. Az A vállalat azt vállalja, hogy a megrendelt mennyiséget egyszerre beszállítja a $[0, T]$ időintervallumban, azon belül azonban véletlenszerű időpontban. A beérkezési időpont tehát valószínűségi változó, amelyről feltesszük, hogy a megfigyelések alapján ismerjük például az $F(x)$ eloszlásfüggvényét. Ennek ismeretében kell meghatároznunk egy M_0 kezdőrészletet, amely $1-\varepsilon$ valószínűséggel fogja az egész $[0, T]$ időintervallumra biztosítani az időegységenkénti r felhasználást. A feladat tehát az

$$F\left(\frac{M_0}{r}\right) = P\left(\xi < \frac{M_0}{r}\right) = 1 - \varepsilon$$

egyenlet megoldása M_0-ra, ahol $0 < \varepsilon < 1$ adott.

Megjegyzzük, hogy a B vállalat szerepét átveheti egy raktár, így a raktárnak kell az időegységenkénti r kiáramlást biztosítania a $[0, T]$ időintervallumban. Így a kérdést úgy fogalmazhatjuk meg, hogy milyen kezdő raktárkészlete legyen, hogy az említett kiáramlást $1 - \varepsilon$ adott valószínűséggel biztosítani tudja. (Itt ξ a „beáramlás” időpontját jelölő valószínűségi változó, azaz a megrendelt mennyiség beérkezési időpontját jelöli.)

Ha például a $[0, T]$-ben egyenletes eloszlású a ξ valószínűségi változó, akkor

$$\frac{\frac{M_0}{r}}{T} = 1 - \varepsilon, \quad \text{amiből}$$
$$M_0 = rT \cdot (1 - \varepsilon) = R \cdot (1 - \varepsilon).$$

Ez tehát azt jelenti, hogy ekkor az M_0 kezdőkészlet megegyezik a teljes igény $(1 - \varepsilon)$-ad részével.

2. modell. Az A vállalat a B vállalattal olyan szerződést köt, hogy a megrendelt rT mennyiséget a $[0, T]$ időintervallumon belül egy fix, előre meghatározott t_1 időpontban fogja szállítani. A B vállalatnak számolnia kell azonban azzal, hogy előre nem látható véletlen okok következtében előfordulhat az is, hogy a $t_1 + \xi, \xi \geq 0$ időpontban érkezik meg a szállítmány.

Legyen a ξ változó $F(x)$ eloszlásfüggvénye ismert. Ekkor meghatározandó a kiinduló M_0 raktárkészletnek az a része, amely csak a véletlen okozta késés fedezésére szolgál. Ha a termelés folyamatosságát $1-\varepsilon$ biztonsággal akarjuk biztosítani, és M jelöli a késés fedezésére szolgáló készletet, akkor ez a

$$P(r\xi < M) = P\left(\xi < \frac{M}{r}\right) = F\left(\frac{M}{r}\right) = 1 - \varepsilon$$

összefüggésből határozható meg. Így

$$M_0 = rt_1 + M$$

lesz a kiinduló készlet.

3. modell. Tegyük fel, hogy a $[0, T]$ időköz n számú egyenlő hosszú olyan részidőintervallumra tagozódik, amelyek mindegyikében a megengedett $rT = R$ mennyiség n-ed része biztosan megérkezik, csupán az a bizonytalan, hogy a részintervallumon belül melyik napon érkezik a szállítás. A szállítások időpontjait egymástól független, azonos eloszlású valószínűségi változóknak tekintjük. Jelölje ismét M_0 azt a kezdeti készletet, amely már a 0 időpontban a fennálló bizonytalanságok okozta esetleges termelési kieséseket hivatott adott biztonsággal fedezni. Nyilvánvalóan bármely $k = 1, 2, \ldots, n$-re

$$\left[(k-1)\frac{T}{n} + \xi_k\right] \cdot r < (k-1) \cdot \frac{R}{n} + M_0$$

$$R = rT \qquad k = 1, 2, \ldots, n$$

kell, hogy teljesüljön, ahol ξ_k jelöli a k-adik intervallumban az áru beérkezési idejét. Így

$$P(\xi_k \cdot r < M_0, k = 1, 2, \ldots, n) \leq 1 - \varepsilon$$

kell, hogy fennálljon. Mivel a $\xi_k - k$ függetlenek, azonos eloszlásúak, ezért az

$$F^n\left(\frac{M_0}{r}\right) = 1 - \varepsilon$$

relációt kapjuk, amelyből

$$F\left(\frac{M_0}{r}\right) = \sqrt[n]{1-\varepsilon}.$$

Ha például $F(x)$ egyenletes eloszlású a $\left[0, \frac{T}{n}\right]$ intervallumon, akkor

$$\frac{\frac{M_0}{r}}{\frac{T}{n}} = \sqrt[n]{1-\varepsilon}.$$

Ebből az

$$M_0 = \frac{R}{n} \cdot \sqrt[n]{1-\varepsilon}$$

adódik.

II.3.3. Véletlen ütemezésű rész-szállítmányok modellje

A hazai tapasztalatok azt bizonyítják, hogy az árucikkek, anyagok jelentős részénél a megrendelés teljesítésére az ún. előszállításos rendszert használják, azaz a megrendelt R mennyiség egy T időintervallumon belül kizárólag a megrendelést teljesítő féltől függően érkezik meg. A szállítmány előre meg nem határozható időpontokban és részletekben érkezik meg, úgy azonban, hogy a T időpontig az egész megrendelt R mennyiség beérkezik. Ha a több éves tapasztalat azt mutatja, hogy a megrendelt mennyiségek időszakról időszakra többnyire n ($n > 2$) alkalommal és nagyjából egyenlő részletekben érkezik be, akkor ennek az utánpótlási rendszernek a leírására a véletlen ütemezésű (Prékopa-Ziermann) modell alkalmas.

1. modell: véletlen ütemezésű, egyenlő nagyságú rész-szállítmányok modellje. Tekintsünk egy $[0, T]$ intervallumot, amelyre véletlenszerűen „dobjunk” rá n pontot. E pontok bármely lehetséges elhelyezkedését úgy tekintjük, mint n számú szállítási időpont egy lehetséges realizációját. Feltesszük, hogy minden egyes lehetséges realizáció ugyanolyan valószínűségű, ami azt jelenti, hogy a pontok egyenletes eloszlást követnek a $[0, T]$ időszakaszon. Így tehát a beérkezési időpontok egymástól független, egyenletes eloszlású valószínűségi változóknak tekinthetők, amelyeket nagyságszerint rendezve jelöljünk $t_1, t_2, \ldots, t_m$-mel. Ezekben az időpontokban a megrendelt mennyiség n-ed része érkezik be a raktárba. Meghatározandó az a legkisebb M_0 kezdő raktárkészlet, amely az egész időtartam minden egységében $1-\varepsilon$ valószínűséggel biztosítja az r intenzitású raktárkészlet felhasználást.

Jelöljük tehát $[0, T]$-vel a vizsgált időtartamot és $M_0(n, \varepsilon)$-nal a keresett kezdőkészletet. Minthogy időegységenként r, ezért az egész időtartam alatt $r \cdot T = R$ lesz a felhasználás. Emiatt R mennyiséget fogunk a rendelési szokásoknak megfelelő időben megrendelni. Jelölje y_t a t időpontig összesen a raktárba érkezett anyag, cikk mennyiségét, z_t pedig a t időpontig összesen felhasznált, illetve a raktárból kivett anyagot, cikket. Nyilvánvalóan teljesül a $z_t = r \cdot t$ egyenlőség, azaz a felhasználást az origón átmenő r iránytangensű egyenes reprezentációja. Ezzel szemben az y_t egy lépcsős függvény, amelynek ugrásai a $t_1, \ldots, t_n$ időpontokban vannak. Tegyük fel, hogy a 0 kezdőpontban rendelkezésünkre áll az $M_0 = M_0(n, \varepsilon)$ kezdőkészlet. Így ha az

$$M_0 + y_t > r \cdot t$$

egyenlőtlenség minden T időpontban legalább $1 - \varepsilon$ valószínűséggel teljesül, akkor az időegységenkénti r felhasználás, raktári kivét az egész $[0, T]$ időintervallumban ε kockázattal van biztosítva. Az előbbi egyenlőtlenség ekvivalens az

$$r \cdot t - y_t < M_0$$

relációval. Mi azonban még ennél is többet fogunk megkívánni, nevezetesen a

$$\sup_{0 \le t < T} \{r \cdot t - y_t\} < M_0$$

teljesülését, azaz megköveteljük, hogy a

$$P\left(\sup_{0 \le t < T} \{r \cdot t - y_t\} < M_0\right) = 1 - \varepsilon$$

teljesüljön, amiből a

$$P(rt - y_t < M_0) \ge 1 - \varepsilon$$

egyenlőtlenség adódik. Prékopa András és Ziermann Margit 1962-ben határoztak meg közelítő értéket az M_0 kezdőkészletre. Az $M_0(n, \varepsilon)$ egzakt értékeit Szmirnov, illetve tőle függetlenül Birnbaum és Tingey határozták meg.

Prékopa András és Ziermann Margit tétele a következőképpen fogalmazható meg:

Tétel. *Ha $n \ge 20$, akkor az egész $[0, T]$ időtartam alatti időegységre eső r konstans felhasználást $1 - \varepsilon$ szinten biztosító kezdőrészletre az*

$$M_0(n, \varepsilon) \approx r \cdot T \cdot \sqrt{\frac{1}{2n} \ln \frac{1}{\varepsilon}}$$

közelítő érték használható.

2. modell: Véletlen ütemezésű, véletlen nagyságú részszállítmányok modellje. A megfigyelt anyagok, cikkek egy jelentős részénél nem teljesül az a feltétel, hogy a véletlen időpontokban beérkező részmennyiségek közel állandóak, egyenlőek. A tapasztalat inkább azt mutatja, hogy a részszállítmányok nagysága is jelentős ingadozást mutat. Tegyük fel, hogy az egy-egy alkalommal beérkező mennyiségek között megállpítható egy olyan legkisebb α mennyiség, amelyet ha szállítás történik, akkor biztosan elszállítanak. Látni fogjuk azt is, hogy ez a megkötés is feloldható, azaz az $\alpha = 0$ is lehetséges lesz.

Mint minden véletlentől függő mennyiség esetében, így most is vagy empirikus úton, vagy a vizsgált jelenség természetéből adódó elméleti hipotézisek alapján feltevéssel kell élnünk a vizsgált jelenséget generáló, azt leíró törvényszerűségekre.

Ezen meggondolások alapján modellünk a következő lesz.

Legyenek adottak a $n \in N$, $0 \leq \alpha \leq \frac{rT}{n}$, $0 < \varepsilon < 1$ számok. A $[0, T]$ időintervallumban összesen $R = r \cdot T$ mennyiséget szállítanak a megrendelt cikkből, anyagból. A szállítások n alkalommal, véletlenszerűen választott időpontban következnek be. Az egyes alkalmakkor szállított mennyiségek nagysága is valószínűségi változó, de minden egyes alkalommal legalább α mennyiséget elszállítanak. Feltesszük azt is, hogy a raktárfeltöltési időpontok bármely lehetséges realizációja a $[0, T]$ intervallumon belül egyenlően valószínű. Végül feltesszük azt is, hogy az $n\alpha$ mennyiség feletti $rT - n\alpha$ mennyiség bármely n részre történő véletlen felosztása egyenlően valószínű. Ezt a modellt is Prékopa András és Ziermann Margit dolgozták ki 1973-ban.

Ezen modell esetén az $M_0 = M_0(n, \varepsilon, \alpha)$ kezdőrészletre az

$$M_0 \approx rT\sqrt{\frac{\ln\dfrac{1}{\varepsilon}}{2n}K_n(\alpha)}$$

közelítő értéket határozták meg, ahol

$$K_n(\alpha) = \sqrt{1 + \left(1 - n\frac{\alpha}{rT}\right)^2}.$$

Látható, hogy az $\alpha = 0$ esetén $M_0 \approx \sqrt{2}M_1$, ahol M_1 az egyenlő részszállítások esetén adódó kezdőkészletet jelöli. Így az általános esetre, vagyis a véletlen részszállítmányok esetére adódó M_0 kezdőrészlet az

$$M_1(n, \varepsilon) \leq M_0(n, \varepsilon, \alpha) < \sqrt{2} \cdot M_1(n, \varepsilon)$$

egyenlőtlenség teljesül.

A véletlen ütemezésű modellek gyakorlati alkalmazásakor kiderült, hogy még olyan esetekben is jó közelítését adják a minimális M_0-nak, amikor a feltételek más készletmodell felállítását indokolják.

II.3.4. Statikus sztochasztikus készletmodellek költségtényezőkkel

1. modell: Kezdő költség és kezdő (kiindulási) raktárkészlet nélüli modell. A $[0, T]$ időintervallumban egy árucikkből raktárról kell biztosítanunk a felhasználást. A szükséges ξ mennyiség a véletlentől függ. A ξ lehetséges értékei a nemnegatív egész számok, amelyekhez tartozó valószínűségeket előzetes tapasztalatból ismertnek tételezzük fel. Ekkor a ξ eloszlásfüggvénye a $p_r = P(\xi = r)$ jelölést felhasználva

$$F(x) = P(\xi < x) = \sum_{i=0}^{r-1} p_i, \quad \text{ahol } r - 1 < x \leq r.$$

Jelölje S a raktári készletet. Ha az S készlet kisebb, mint a T idő alatt ténylegesen fellépő r igény, akkor a hiány minden egysége után h Ft veszteség éri a raktárt (például elesik bizonyos nyereségtől). Ezt nevezzük hiányköltségnek. Ha viszont a raktáron eladatlan készlet marad, azaz $S > r$, akkor a többlet minden egysége után u Ft veszteség ér bennünket (például amiatt, hogy az áru „öregszik", vagy lejár a szavatossági ideje, stb.). Ez a fenntartási költség. Ezenkívül jelölje $c < h$ az egységenkénti vételárat. Az összköltségben szereplő mindhárom költség csak a darabszámtól függ.

Ezek után a kérdés az, hogy mekkora legyen az az S_0 raktárkészlet, amely várható értékben a minimális veszteséget eredményezi. Jelölje $C(S)$ az S raktárkészlet melletti veszteséget (költséget). Minthogy a szükséglet valószínűségi változó, így $C(S)$ is az, ezért meg kell határoznunk azt az S_0 raktárkészletet, amely mellett a $C(S)$ várható értéke minimális.

Írjuk fel az $E\big(C(S)\big)$ várható értéket. Ha a ξ értéke, a r kisebb S-nél, akkor $S - r$ feleslegünk van. Az ehhez tartozó várható költség $u \cdot \sum_{r=0}^{S}(S - r)p_r$. Ha viszont a r nagyobb az S-nél, akkor $r - S$ a hiányunk, aminek a költsége várhatóan

$$h \sum_{r=s+1}^{\infty} (r - s)p_r.$$

S darab árucikk vételára $c \cdot S$, így a költségünk várható értéke

$$E\big(C(S)\big) = u \cdot \sum_{r=0}^{S}(S - r)p_r + h \sum_{r=s+1}^{\infty} (r - s)p_r + c \cdot S.$$

Határozzuk most meg az $E\big(C(S)\big)$ minimumát adó S_0 értéket. Ehhez előbb számítsuk ki az $E\big(C(S+1)\big)$, $E\big(C(S-1)\big)$ értékeket:

$$E\big(C(S+1)\big) = u\cdot\sum_{r=0}^{S+1}(S+1-r)p_r + h\sum_{r=S+2}^{\infty}(r-S-1)p_r + c\cdot(S+1) =$$

$$= u\cdot\sum_{r=0}^{S}(S+1-r)p_r + h\sum_{r=S+1}^{\infty}(r-S-1)p_r + c\cdot(S+1) =$$

$$= u\cdot\sum_{r=0}^{S}(S-1)p_r + u\sum_{r=0}^{S}p_r + h\sum_{r=S+1}^{\infty}(r-S)p_r - h\sum_{r=S+1}^{\infty}p_r +$$

$$+ c\cdot(S+1) = u\sum_{r=0}^{S}(S-r)p_r - h\sum_{r=S+1}^{\infty}(r-S)p_r + cS +$$

$$+ u\sum_{r=0}^{S}p_r - h\sum_{r=S+1}^{\infty}p_r + c = E\big(C(S)\big) + u\sum_{r=0}^{S}p_r -$$

$$- h\left(1-\sum_{r=0}^{S}p_r\right) + c = E\big(C(S)\big) + (u+h)\sum_{r=0}^{S}p_r + c - h =$$

$$= E\big(C(S)\big) + (u+h)P(\xi\le S) + c - h.$$

Hasonlóan kapjuk, hogy

$$E\big(C(S-1)\big) = E\big(C(S)\big) - (u+h)P(\xi\le S-1) + h - c.$$

Ha S_0 a várható érték minimumát adó raktárkészlet, akkor az

$$E\big(C(S_0-1)\big) > E\big(C(S_0)\big)$$
$$E\big(C(S_0+1)\big) > E\big(C(S_0)\big)$$

egyenlőtlenségek teljesülnek. Felhasználva az előbbi eredményeket, az

$$E\big(C(S_0+1)\big) - E\big(C(S_0)\big) = (u+h)P(\xi\le S_0) + c - h > 0$$

és az

$$E\big(C(S_0-1)\big) - E\big(C(S_0)\big) = -(u+h)P(\xi\le S_0-1) + h - c > 0$$

egyenlőtlenségek adódnak. Így végül a

$$\frac{h-c}{u+h} < P(\xi\le S_0)$$

és a

$$\frac{h-c}{u+h} > P(\xi \le S_0 - 1)$$

egyenlőtlenségekhez jutunk.

Ha $P(\xi \le S_0 - 1) < \frac{h-c}{u+h} = P(\xi \le S_0)$, akkor $E\big(C(S_0+1)\big) = E\big(C(S_0)\big)$, vagyis S_0 és $S_0 + 1$ egyaránt optimumhelyek. Ha pedig $P(\xi \le S_0 - 1) = \frac{h-c}{u+h} < P(\xi \le S_0)$, akkor $S_0 - 1$ és S_0 optimumhelyek.

Folytonos esetre is megfogalmazhatjuk az előbbi problémát: Legyen ξ egy folytonos, nemnegatív értéket felvevő valószínűségi változó, amelynek $F(x)$ jelölje az eloszlásfüggvényét, $f(x)$ pedig a sűrűségfüggvényét. Meghatározandó az az S_0 raktárkészlet, amelyre az

$$E\big(C(S)\big) = u\int_0^S (S-x)f(x)dx + h\int_S^\infty (x-S)f(x)dx + c \cdot S$$

felveszi a minimumát.

A szélsőértékszámítást felhasználva megoldandó a

$$\frac{dE\big(C(S)\big)}{dS} = 0$$

egyenlet. Belátható, hogy az előbbi egyenlet megoldásaként az $F(S) = \frac{h-c}{u+h}$ egyenlőség adódik. Megmutatható, hogy az az S_0, amelyre $F(S_0) = \frac{h-c}{u+h}$ teljesül, valóban minimumhelye az $E\big(C(S)\big)$-nek, tehát nemcsak a szükséges, de az elegendő feltételnek is leget tesz.

2. modell: Kezdő (kiindulási) raktárkészlettel rendelkező, de kezdő költség nélküli modell. Legyen x_0 az a kezdőkészlet, amely már a rendelés feladása előtt rendelkezésünkre áll. Ekkor a várható költségre

$$\begin{cases} c(S-x_0) + L(S), & \text{ha } S > x_0 \\ L(x_0), & \text{ha } S = x_0, \end{cases}$$

ahol

$$L(S) = h\int_S^\infty (x-S)f(x)dx + u\int_0^S (S-x)f(x)dx.$$

Itt a $c \cdot S + L(S)$ ugyanaz a várható költség, mint ami az előző modell esetében volt. Legyen most is S_0 az a készlet, amely minimalizálja a $c \cdot S + L(S)$ függvényt, azaz az $F(S_0) = \frac{h-c}{h+u}$ egyenlőség teljesüljön.

Ha $x_0 \ge S_0$, akkor minden $S \ge x_0$ esetén

$$c \cdot S + L(S) \ge c \cdot x_0 + L(x_0),$$

így

$$c \cdot (S - x_0) + L(S) \geq L(x_0), \quad \text{vagyis}$$
$$\min_{S \geq x_0} \{c \cdot (S - x_0) + L(S)\} \geq L(x_0).$$

Ez tehát azt jelenti, hogy ebben az esetben nem érdemes rendelni. Ha viszont $S_0 > x_0$, akkor ntilvánvalóan

$$\min_{S \geq x_0} \{c \cdot (S - x_0) + L(S)\} = \min_{S \geq x_0} \{L(S) + c \cdot S - cx_0\} =$$
$$= L(S_0) + c \cdot S_0 - cx_0 = L(S_0) + c(S_0 - x_0).$$

Így tehát az optimális stratégia a következő lesz:

- ha $S_0 \leq x_0$, akkor nem rendelünk,
- ha pedig $S_0 > x_0$, akkor $S_0 - x_0$-at rendelünk.

3. modell: Kezdő (kiindulási) raktárkészlettel és kezdő költséggel rendelkező modell. Tegyük fel most, hogy a rendelés megkezdésekor K összeget kell fizetnünk (pélául ilyen a rendelési díj).

Mint az előző modellben, legyen most is

$$L(S) = h \int_S^\infty (x - S) f(x) dx + u \int_0^S (S - x) f(x) dx$$

a várható hiány és fenntartási költség összege. Így x_0 kezdőkészlet esetén a várható költség

$$\begin{cases} K + c \cdot (S - x_0) + L(S), & \text{ha } S > x_0 \\ L(x_0), & \text{ha } S = x_0, \end{cases}$$

Minthogy $c \cdot S + L(S)$ ugyanaz a várható költség, mint amit már az 1. modellnél használtunk, ezért a minimumát adó S_0 készletre az $F(S_0) = \frac{h-c}{h-u}$ egyenlőség kell, hogy teljesüljön.

Ha $x_0 \geq S_0$, akkor

$$K + c \cdot S + L(S) > cx_0 + L(x_0) \qquad \forall S > x_0 \text{ esetén teljesül,}$$

s mivel

$$c \cdot S_0 + L(S_0) < c \cdot S + L(S) \qquad \forall S \neq S_0,$$

ezért

$$cx_0 + L(x_0) < c \cdot S + L(S) < c \cdot S + L(S) + K.$$

Innen

$$K + c \cdot (S - x_0) + L(S) > L(x_0)$$

adódik, ahol a baloldal az x_0 kezdőrészlet S szintre történő feltöltés költsége és $L(x_0)$ az átlagos költség, ha nem töténik rendelés. Ebből tehát az adódik, hogy ha $x_0 > S_0$, akkor az optimális stratégia az, ha nem rendelünk.

Legyen s_0 az a legkisebb értéke S-nek, amelyre

$$c \cdot s_0 + L(s_0) = K + c \cdot S_0 + L(S_0), \qquad s_0 < S_0.$$

Abban az esetben, ha $s_0 \leq x_0 < S_0$ és $x_0 < S$, akkor

$$K + c \cdot S + L(S) \geq cx_0 + L(x_0), \text{ s mivel}$$
$$c \cdot S + L(S) > cS_0 + L(x_0), \text{ így}$$
$$K + cS + L(S) \geq K + c \cdot S_0 + L(S_0) = c \cdot s_0 + L(s_0) \geq cx_0 + L(x_0).$$

Ebből a

$$K + c(S - x_0) + L(S) \geq L(x_0)$$

adódik, ami azt jelenti, hogy a nemrendelés ismét olcsóbb, mint a rendelés.

Végül, ha $x_0 < s_0 (< S_0)$, akkor

$$\min_{S \geq x_0} \{K + cS + L(S)\} = K + cS_0 + L(S_0) = c \cdot s_0 + L(s_0) \leq$$
$$\leq cx_0 + L(x_0),$$

amiből

$$\min_{S \geq x_0} \{K + c(S - x_0) + L(S)\} = K + c(S_0 - x_0) + L(S_0) \leq L(x_0)$$

adódik. Így $x_0 < s_0$ esetén a minimális költség

$$K + c \cdot (S_0 - x_0) + L(S_0)$$

lesz. Ezek után az optimális eljárás:

$x_0 < s_0$ esetén $s_0 - x_0$ mennyiséget rendelünk,
$x_0 \geq s_0$ esetén nem rendelünk.

III. Sorbanállási rendszerek

Ez a fejezet az élet egyik „leghaszontalanabb” tevékenységével, a várakozással, a sorbanállással foglalkozik. Sorok bármikor keletkezhetnek, amikor egy adott kérés kiszolgálása meghaladja a kiszolgáló egység kapacitását. Ilyennel a gyakorlati életben is sokszor találkozunk. Például egy áruházban sorbanállunk a pénztárnál, várakozunk az útkereszteződésben a piros lámpánál, tartjuk a telefont, amíg kicseng, várjuk, amíg a CPU kiszolgálja a job-ot, stb. A várakozás, sorbanállás tanulmányozására azért van szükség, mert ha rájövünk az összefüggéseire, törvényszerűségeire, akkor csökkenthetjük annak időtartamát és az egyes tevékenységeinket jobban tervezhetjük.

III.1. Sztochasztikus folyamatok elméletének rövid áttekintése

A sorbanállási rendszereket a dinamikai rendszerek egy tágabb osztályába, a folyamatok közé soroljuk.

A folyamatok két nagy csoportja a determinisztikus és a sztochasztikus folyamatok. A determinisztikus folyamatok olyan rendszerek, amelyeknél előre pontosan ismert a „folyamat mennyisége” és az többnyire állandó a vizsgált időtartam alatt. Ismert továbbá az az időpont is, amikor a folyamat megjelenik a „csatorna” bemenetén. Végül a folyamatnak a „csatornával” szemben támasztott igénye is ismert és állandó. (A „folyamat mennyisége” és a „csatorna” fogalmakat később definiálni fogjuk.)

A sztochasztikus folyamatok osztályába tartozóknál bizonytalanok és előrejelezhetetlenek azok az időpontok, amikor az igények beérkeznek és a csatornával szemben támasztott igények nagyságát sem lehet előre megmondani. A gyakorlatban található rendszerek nagy része ebbe a kategóriába tartozik.

A sztochasztikus folyamatokkal kapcsolatban számos kérdést felvethetünk, amelyekre értelmes és határozott választ szeretnénk kapni. Ilyenek például a következők:

- Várhatóan mennyi ideig áll sorban egy igény, mielőtt azt kiszolgálják?
- Hány igényt fognak kiszolgálni, mielőtt az újonnan érkezett igény sorra kerül?
- A munkaidő mekkora hányadában lesz foglalt a CPU egység?
- Mekkorák lesznek a CPU-ban folyamatosan lefoglalt időintervallumok?

Ezek és a hasonló kérdések bizonyos valószínűségekre illetve várható értékekre vonatkoznak.

A sztochasztikus folyamat az $X(t)$ valószínűségi változók egy családja, ahol t a időparamétert jelenti. Például egy moziban a nézők száma az idő függvényében sztochasztikus folyamatnak tekinthető.

A sztochasztikus folyamatok osztályozása három tényezőtől függ: az állapottértől, az indextől (azaz a t időparamétertől) és a t index különböző értékeihez tartozó $X(t)$ valószínűségi változók közötti statisztikai összefüggésektől. Vizsgáljuk most meg ezeket a tényzőket.

Először tekintsük az állapotteret. Állapottérnek nevezzük az $X(t)$ által felvehető lehetséges értékek, másként állapotok halmazát. Diszkrét állapoterű folyamatról, más elnevezéssel láncról beszélhetünk, ha a folyamat csak véges vagy legfeljebb megszámlálhatóan végtelen sok különböző állapot valamelyikében lehet. Egy lánc állapottere gyakran az egész számok halmaza.

Ha a megengedett állapotok egy véges vagy végtelen intervallumot határoznak meg, akkor folytonos állapotterű folyamatról beszélünk.

Tekintsük most az időparamétert (indexet). Ha a t időparaméter csak véges, vagy legfeljebb megszámlálhatóan végtelen sok különböző olyan értéket vehet fel, amelyben az állapotváltozás végbemehet, akkor diszkrét paraméterű folyamatról beszélünk. Ha viszont az állapotváltozások akárhol előfordulhatnak az időtengely egy véges vagy végtelen intervallumán, akkor folytonos paraméterű folyamatról beszélünk. A diszkrét időparaméter esetében az $X(t)$ jelölés helyett az X_k jelölést használjuk és ekkor véletlen (sztochasztikus) sorozatról beszélünk.

Az egyes sztochasztikus folyamatok (sorozatok) meghatározó jellemzője az a viszony, amely a különböző $X(t)$ (X_k) valószínűségi változók között fennáll. A folyamat leírásához szükséges lesz az $X(t)$ (X_k) valószínűségi változók együttes eloszlására: tekintsük a különböző $t_1, t_2, \ldots, t_n$ időpontokat és határozzuk meg az

$$F(x_1, x_2, \ldots, x_n) = P\big(X(t_1) < x_1,\ X(t_2) < x_2, \ldots, X(t_n) < X_n\big)$$

együttes eloszlásfüggvényt minden $n, (x_1, \ldots, x_n)^T \in R^n, (t_1, \ldots, t_n)$ értékre, ahol $t_i \geq 0,\ i = 1, \ldots, n$.

A sztochasztikus folyamatoknak több tipusa is ismeretes: stacionárius, független növekményű, Markov, születési – halálozási, szemi-Markov és felújítási folyamatok. Részletesen csak a Markov, és azon belül is a születési – halálozási folyamatokkal foglalkozunk, mivel az elemi sorbanállási rendszerek ezen folyamatokra épülnek.

Egy $\{X_k\}$ diszkrét állapotterű és diszkrét idejű sztochasztikus folyamatot Markov-láncnak nevezünk, ha az X_{k+1} érték csak az X_k értéktől függ, a korábbiaktól nem. Ez azt is jelenti, hogy a múlt befolyását a folyamat jövőjére vonatkozóan a pillanatnyi állapot teljesen tartalmazza.

Szemléletesen azt is mondhatjuk, hogy a rendszer (a sztochasztikus folyamat) „emlékezet nélküli”, hiszen a múltbéli értékek (az emlékezet) nem befolyásolják a rendszer jövőbeli viselkedését.

Folytonos idejű Markov-lánc esetében bármelyik időpillanatban bekövetkezhet az állapotváltozás.

A Markov-tulajdonságot analitikusan a következőképpen lehet felírni:

$$P\big(X(t_{k+1}) = x_{k+1} \big| X(t_k) = x_k,\ X(t_{k-1}) = x_{k-1}, \ldots, X(t_1) = x_1\big) =$$
$$= P\big(X(t_{k+1}) = X_{k+1} \big| X(t_k) = x_k\big).$$

A Markov-folyamatok egy igen fontos speciális osztálya a születési – halálozási folyamatok. Ezt a folyamatot definiáló feltétel: minden állapotból csak „szomszédos” állapotba mehet végbe átmenet. Állapottérnek ekkor az egész számok halmazát választjuk. Így az $x_k = i$ esetében x_{k+1} vagy $i-1$, vagy i, vagy pedig $i+1$ lehet. Ezeknek a folyamatoknak nagy szerepük van a sorbanállási rendszerek vizsgálatában.

Ahhoz, hogy egy $X(t)$ Markov-lánc születési – halálozási folyamat legyen, ki kell elégítenie az alábbi feltételeket:

1. $P\big(X(t+h) = k+1 \big| X(t) = k\big) = \lambda_k \cdot h + o(h), \quad \lambda_k > 0$
2. $P\big(X(t+h) = k-1 \big| X(t) = k\big) = \mu_k \cdot h + o(h), \quad \mu_k > 0$
3. $P\big(X(t+h) = k \big| X(t) = k\big) = 1 - (\lambda_k + \mu_k)h + o(h)$
4. $P\big(X(t+h) = m \big| X(t) = k\big) = o(h), \quad \text{ha } |m-k| > 1,$

ahol h egy tetszőlegesen kis intervallumot jelent, $o(h)$-ra pedig igaz, hogy $\frac{o(h)}{h} \to 0$, ha $h \to 0$. A λ_k és μ_k értékek függetlenek az időtől. A λ_k értékeket születési, a μ_k-kat pedig halálozási intenzitásnak nevezzük.

Jelöljük $P_k(t)$-vel annak a valószínűségét, hogy a folyamat a t időpillanatban a k állapotban van, vagyis $P_k(t) = P\big(X(t) = k\big)$. Ezen (abszolut) valószínűségeknek a kiszámításához a következőket kell figyelembe venni: a $t+h$ időpillanatban az $X(t+h)$ akkor és csakis akkor lesz a k állapotban, ha a következő feltételek valamelyike teljesül:

1. a t időpillanatban a folyamat a k állapotban van és a $(t,\ t+h)$ időintervallumban nem következik be változás;
2. a t időpillanatban a folyamat a $k-1$ állapotban volt, és a $(t,\ t+h)$ időintervallumban a k állapotba történt az átmenet;
3. a t időpillanatban a folyamat a $k+1$ állapotban volt, és a $(t,\ t+h)$ időintervallumban a k állapotba történt az átmenet;
4. a t időpillanatban valamilyen állapotban volt a folyamat, a $(t,\ t+h)$ időintervallumban legalább 2 átmenet történt úgy, hogy $(t+h)$-ban a k állapotba kerül a folyamat.

Nyilvánvalóan igaz, hogy $\sum_{k=0}^{\infty} P_k(t) = 1$.
Most írjuk fel a $P_k(t+h)$ valószínűséget:

$$\begin{aligned} P_k(t+h) = P_k(t) \cdot \big(1 - \lambda_k h - \mu_k h + o(h)\big) + \\ + P_{k-1}(t) \cdot \big(\lambda_{k-1} \cdot h + o(h)\big) + \\ + P_{k+1}(t) \cdot \big(\mu_{k+1} \cdot h + o(h)\big) + o(h), \end{aligned}$$

ahol $k \geq 1$. Ha mindkét oldalból kivonjuk a $P_k(t)$ értéket, majd az egyenlet mindkét oldalát elosztjuk h-val, akkor $h \to 0$ esetén a következő differenciálegyenleteket kapjuk (amit a folyamat állapotegyenleteinek is nevezünk):

$$\begin{aligned} P_0'(t) &= -\ P_0(t) + \mu_1 P_1(t) \\ P_k'(t) &= -\ (\lambda_k + \mu_k) P_k(t) + \lambda_{k-1} P_{k-1}(t) + \mu_{k+1} P_{k+1}(t), \end{aligned}$$

ahol $k = 1, 2, 3, \ldots$. Amennyiben megadjuk a $P_k(0), k = 0, 1, 2, \ldots$ értékeket és megköveteljük, hogy $\sum_{k=0}^{\infty} P_k(t) = 1$ legyen, akkor az előbbi differenciálegyenletrendszer megoldható.

Az általános, időtől függő megoldást nehéz meghatározni, ezért megelégszünk az ún. egyensúlyi (vagy stacionárius) megoldással is, ami a legtöbb esetben elegendőnek bizonyul.

A stacionárius megoldást úgy definiáljuk, mint egy olyan p_k valószínűségeloszlást, amelyre a $P_k(t) = p_k$ teljesül minden k értékre. Ha egy ilyen eloszlás létezik, akkor az egyértelműen meghatározott és minden k-ra teljesül, hogy

$$\lim_{t \to \infty} P_k(t) = p_k.$$

Mivel minket csak a folyamat időtől független tulajdonságai érdekelnek, ezért az előbbi differenciálegyenletrendszerben vegyük a $t \to \infty$ határértéket. Ekkor a $k = 1, 2, \ldots$ esetén a baloldal 0, míg a jobboldal $-(\lambda_k + \mu_k)p_k + \lambda_{k-1}p_{k-1} + \mu_{k+1}p_{k+1}$ lesz. Ebből átalakítással a

$$\lambda_k p_k - \mu_{k+1} p_{k+1} = \lambda_{k-1} p_{k-1} - \mu_k p_k \qquad k = 1, 2, \ldots$$

adódik. Ebből viszont azt kapjuk, hogy $\lambda_{k-1} \cdot p_{k-1} - \mu_k p_k$ konstans értékű minden $k = 1, 2, \ldots$ értékre. Minthogy pedig

$$\lambda_0 p_0 - \mu_1 p_1 = 0,$$

ezért az előbbi konstans értéke is 0 lesz, így minden $k = 1, 2, \ldots$ értékre igaz lesz a $\mu_k p_k = \lambda_{k-1} p_{k-1}$ egyenlőség. Az egyensúlyi állapot valószínűségeire a

$$p_k = \frac{\lambda_{k-1}}{\mu_k} p_{k-1} = \frac{\Lambda(k)}{M(k)} \cdot p_0 \text{ adódik,}$$

ahol

$$\Lambda(k) = \prod_{i=1}^{k} \lambda_{i-1} \quad \text{és} \quad M(k) = \prod_{i=k}^{k} \mu_i .$$

A p_0 valószínűséget a $\sum_{k=1}^{\infty} p_k = 1$ összefüggésből határozhatjuk meg: ha az

$$1 + \sum_{k=1}^{\infty} \frac{\Lambda(k)}{M(k)} = 1 + \sum_{k=1}^{\infty} \prod_{i=1}^{k} \frac{\lambda_{i-1}}{\mu_i}$$

sor konvergens és összege φ, akkor $p_0 = \frac{1}{\varphi}$.

Ez a feltétel egyben stacionárius eloszlás létezésének elégséges feltétele is.

A vizsgált rendszerek közül a legegyszerűbbek a tiszta születési folyamatok. Ilyenkor feltesszük, hogy a $\mu_k = 0$ teljesül minden k esetén. Tovább egyszerűsítve a problémát, tegyük fel, hogy $\lambda_k = \lambda, k = 0, 1, 2, \ldots$. Ekkor a differenciálegyenletrendszerünk a

$$\begin{aligned} P_k'(t) &= -\lambda P_k(t) + \lambda P_{k-1}(t) \qquad && k \geq 1 \\ P_0'(t) &= -\lambda P_0(t) && k = 0 \end{aligned}$$

formájú lesz. Az egyszerűség kedvéért feltesszük, hogy a rendszer a 0 állapotból indul a 0 időpillanatban, azaz

$$P_k(0) = \begin{cases} 1, & \text{ha } k = 0 \\ 0, & \text{ha } k \neq 0. \end{cases}$$

Megoldásként a

$$
\begin{aligned}
P_0(t) &= e^{-\lambda t} \\
P_1(t) &= \lambda t e^{-\lambda t} \\
&\vdots \\
P_k(t) &= \frac{(\lambda t)^k}{k!} e^{-\lambda t} \qquad k \geq 0, t \geq 0 \\
&\vdots
\end{aligned}
$$

adódik. A konstans λ születési intenzitású tiszta születési folyamatban előforduló születések sorozatát nevezik Poisson-folyamatnak. A Poisson-folyamat központi szerepet tölt be a sorbanállási elméletben. Ezt a folyamatot mint az igények beérkezési folyamatát fogjuk tekinteni valamilyen kiszolgálási rendszerben, ezért a jelöléseket is ennek megfelelően fogjuk használni. A λ most az igénybeérkezés átlagos intenzitását fogja jelenteni. A kezdeti feltétellel együtt a $P_k(t)$ megadja annak a valószínűségét, hogy a $(0, t)$ intervallumban k igény érkezzen be.

A születési folyamat tulajdonságaiból adódik, hogy a Poisson-folyamat homogén: ha $X(s, s+t)$ jelöli a t hosszúságú $(s, s+t)$ intervallum alatti beérkezések számát, akkor

$$
P\big(X(s, s+t) = k\big) = \frac{(\lambda t)^k e^{-\lambda t}}{k!},
$$

függetlenül attól, hogy hol helyezkedik el az intervallum, azaz független az intervallum s kezdőpontjától.

Legyen most a beérkezési folyamat Poisson-folyamat és vizsgáljuk meg, hogy két beérkezés között eltelt idő milyen eloszlást követ. Legyen a $\tilde{t}$ valószínűségi változó a két egymás utáni beérkezés között eltelt idő, amelynek eloszlás- és sűrűségfüggvényét jelöljék $F(t)$ illetve $f(t)$. Ekkor $f(t) \cdot \Delta t + o(\Delta t)$ annak a valószínűsége, hogy a legutolsó beérkezéstől a soron következő beérkezésig eltelt idő legalább t, de legfeljebb $(t + \Delta t)$. Mivel $F(t)$ a valószínűsége annak, hogy a beérkezések között eltelt idő t-nél kisebb legyen, ezért

$$
F(t) = 1 - P(\tilde{t} > t).
$$

Minthogy pedig $P(\tilde{t} > t)$ annak a valószínűsége, hogy a $(0, t)$ intervallumban egyetlen beérkezés sem következik be, ezért $F(t) = 1 - P_0(t) = 1 - e^{-\lambda t}$. Ebből $f(t) = \lambda e^{-\lambda t}$, ha $t \geq 0$. Ez tehát azt jelenti, hogy Poisson-folyamat esetén a beérkezések időköze exponenciális eloszlást követ λ paraméterrel. Az exponenciális eloszlás legfontosabb jellemzője az, hogy emlékezet nélküli, azaz a valószínűségi változó múltja nem játszik szerepet jövőjének meghatározásában.

A $\lim_{t\to 0} \frac{1-e^{-\lambda t}}{t} = \lambda$ határértéket felhasználva láthatjuk be, hogy

$$1 - e^{\lambda t} = \lambda t + o(t).$$

Ez utóbbi egyenlőség a sorbanállási problémáknál fontos lesz számunkra.

A születési – halálozási folyamatok időtől függő megoldása gyakorlatilag kezelhetetlenné válik, amint a λ_k, μ_k intenzitásokat tekintjük. Mindezek miatt természetes kérdés az is, hogy $t \to \infty$ esetén lesz-e, s ha igen, akkor mi lesz az állapotegyenletek (azaz a differenciálegyenletek) stacionárius megoldása.

Feltételezve, hogy a határérték létezik, a születési – halálozási folyamatokra felírt differenciálegyenletekben a $\lim_{t\to\infty} P_k'(t)$ mennyiségeket 0-val tehetjük egyenlővé.

Ebből adódik, hogy

$$\begin{aligned} 0 &= -(\lambda_k + \mu_k)p_k + \lambda_{k-1}p_{k-1} + \mu_{k+1} \cdot p_{k+1}, && \text{ha } k \geq 1 \\ 0 &= \lambda_0 p_0 + \mu_1 p_1, && \text{ha } k = 0. \end{aligned}$$

Megköveteljük a teljes eseményrendszerre vonatkozó $\sum_{k=0}^{\infty} p_k = 1$ összefüggést is, amelyre normalizáló feltételként fogunk hivatkozni.

A stacionárius megoldásra

$$p_k = \frac{\lambda_0 \cdot \lambda_1 \cdot \ldots \cdot \lambda_{k-1}}{\mu_1 \cdot \mu_2 \cdot \ldots \cdot \mu_k} p_0, \quad k = 1, 2, \ldots$$

adódik. A normalizáló feltételt felhasználva p_0-ra a

$$p_0 = \frac{1}{1 + \sum\limits_{k=1}^{\infty} \prod\limits_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}}}$$

értéket kapjuk. Ez a megoldás az egyik legfontosabb összefüggése a sorbanállási elméletnek.

Meg kell vizsgálnunk még, hogy a $p_k, k = 0, 1, 2, \ldots$ stacionárius eloszlás milyen feltétel mellett létezik. Ahhoz, hogy a $p_k, k = 0, 1, 2, \ldots$ értékek valószínűségeloszlást alkothassanak a $p_0 > 0$ szükségképpen kell, hogy teljesüljön. A $p_0 > 0$ feltétel pedig az egyenletekben szereplő születési és halálozási együtthatókra ró ki feltételt.

Legyenek

$$S_1 = \sum_{k=0}^{\infty} \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}}$$

$$S_2 = \sum_{k=0}^{\infty} \frac{1}{\prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}}}.$$

Azt mondjuk, hogy a születési – halálozási folyamat minden egyes állapota ergodikus, ha $S_1 < \infty$, $S_2 = \infty$. Ergodikus esetben $p_0 > 0$. Ahhoz pedig, hogy a folyamat ergodikus legyen elegendő feltétel, hogy a $\frac{\lambda_k}{\mu_k}$ sorozat egy bizonyos k értéktől kezdve 1-nél kisebb értékű legyen. A legtöbb sorbanállási rendszerban teljesül ez a feltétel.

III.2. A sorbanállási rendszerek jellemzői

Ahhoz, hogy teljesen jellemezzünk egy sorbanállási rendszert, azonosítanunk kell azt a sztochasztikus folyamatot, amely a beérkező igényeket írja le, továbbá le kell írnunk a kiszolgálás szabályait és struktúráját.

A beérkező folyamatot általában az egymás után beérkező igények közötti időintervallumok valószínűségeloszlása segítségével írjuk le. Jelöljük ezt az $A(t)$ szimbólummal, tehát

$$A(t) = P(\text{két egymás utáni beérkezés közötti időköz} < t).$$

A sorbanállás elméletében többnyire feltesszük, hogy az egymás utáni beérkezések közötti időközök (röviden: beérkezési időközök) azonos eloszlású, független valószínűségi változók. (A beérkezési folyamat tehát ún. felújítási folyamat.)

Egy másik sztochasztikus mennyiség (amit szintén meg kell adnunk) a beérkező igények által a csatornával szemben támasztott követelmények nagysága. Ezt kiszolgálási időnek nevezzük és a valószínűségeloszlását $B(x)$-szel jelöljük, azaz

$$B(x) = P(\text{a kiszolgálási idő} < x).$$

A kiszolgálás ideje annak az időintervallumnak a hosszát jelenti, amelyet az igény a kiszolgáló egységben eltölt.

A kiszolgálás szabályára és strukturájára vonatkozóan további mennyiségeket kell még meghatározni.

Az egyik ilyen a befogadóképesség, ami nem más, mint a kiszolgálásra várakozók számának maximális nagysága. Ezt rendszerint K-val jelöljük és értékét gyakran végtelennek tekintjük. Egy további jellemző a rendelkezésre álló kiszolgáló egységek (állomások), azaz a csatornák száma. A kiszolgálási sorrend írja le azt a szabályt, amely szerint a várakozók sorra kerülnek a kiszolgáláskor. Leggyakrabban használt kiszolgálási elvek:

- FIFO (First In – First Out), azaz érkezési sorrendben történő kiszolgálás;
- LIFO (Last In –First Out), azaz a fordított sorrendben történő kiszolgálás.

Ha a beérkező igényeket bizonyos csoportokba való tartozás szerint meg lehet különböztetni, akkor a csoportok között prioritást lehet megállapítani, ami majd a kiszolgálás sorrendjét fogja meghatározni. Ez az egyik legalkalmasabb ütemezési elv, mivel így az igények közötti fontossági sorrendet felállítva történik a kiszolgálás. A prioritásos sorbanállási elvnek két fő tipusa van: abszolút és relatív. Az előbbi azt jelenti, hogy ha egy igény kiszolgálása folyamatban van és érkezik egy magasabb prioritású igény, akkor az éppen kiszolgálás alatt álló igény kiszolgálása megszakad, újra beáll a várakozók sorába és a magasabb prioritású igény kerül kiszolgálásra.

A sorbanállási rendszerek hatékonyságának és teljesítményének vizsgálatához a következő mérőszámokat fogjuk meghatározni:

- az igények várakozási idejét,
- a rendszerben lévő igények számát,
- a foglaltsági intervallum hosszát (vagyis azt az időintervallumot, amelyben a kiszolgáló egység folyamatosan foglalt),
- az üresjárati időszakasz hosszát,
- a pillanatnyi munkahátralék eloszlását.

Ezek mindegyike valószínűségi változó, így általában az eloszlásfüggvényeiket szeretnénk meghatározni, amit általában nehéz megadni, ezért sokszor megelégszünk ezek átlagos értékeivel is.

A sorbanállási rendszerek teljesítményének mérésére legalkalmasabb a torlódás vizsgálata: egycsatornás modellben

$$\varrho = \text{forgalmi intenzitás} = \frac{\text{átlagos kiszolgálási idő}}{\text{átlagos beérkezési időköz}}.$$

Jelöljük egy végtelen populációjú modell esetében λ-val a beérkezési intenzitást, $\frac{1}{\mu}$-vel pedig az átlagos kiszolgálási időt. Ekkor

$$\varrho = \text{érkezési intenzitás} * \text{átlagos kiszolgálási idő} = \frac{\lambda}{\mu}.$$

Az 1-nél nagyobb forgalmi intenzitás azt mutatja, hogy az igények gyorsabban érkeznek, mint ahogy azokat egy szerver (kiszolgáló egység, csatorna) képes kiszolgálni.

Egy m párhuzamos szerverből álló rendszerben átlagosan $\lambda T|m$ igény érkezik szerverenként, feltéve, hogy a forgalom egyenletes eloszlású az m kiszolgáló egység között. Ha minden beérkezett kérés kiszolgálása átlagosan $\frac{1}{\mu}$ ideig tart, akkor a szerver teljes foglaltsági idejének várható értéke $\frac{\lambda \cdot T}{m \cdot \mu}$. Ezt T-vel elosztva $\varrho = \frac{\lambda}{m\mu}$ adódik. Mivel a kihasználtság maximum 1 lehet, így az m szerveres rendszer kihasználtsági tényezőjére vonatkozó korrekt kifejezés:

$$\varrho = \min\left\{\frac{\lambda}{m\mu}, 1\right\}.$$

Jelölje $X(t) = 0$ azt az eseményt, hogy a kiszolgáló tétlen a t-edik időpillanatban. Ekkor a szerver (csatorna) időegységre eső kihasználtságát az

$$U_s = \lim_{T\to\infty} \frac{1}{T} \int_0^T \chi\big(X(t) \neq 0\big) dt = 1 - p_0 = \frac{E\delta}{E\delta + E_i}$$

határérték adja meg, feltéve, hogy az létezik.

(Itt p_0 azon (stacionárius) valószínűség, hogy a szerver tétlen, $E\delta$ a kiszolgáló egység átlagos foglaltsági periódushossza, E_i pedig az átlagos tétlenségi periódushossz. Végül a χ függvény az $X(t) \neq 0$ esemény karakterisztikus függvénye.)

Egy másik gyakran használt teljesítménymérő a rendszer átbocsátó képessége. Ezt a mennyiséget úgy definiálhatjuk, mint az időegységenkénti kiszolgált igények átlagos számát. Egy m szerveres rendszerben minden időegység alatt $m \cdot \varrho \cdot \mu$ igény kiszolgálása fejeződik be, így az

$$\text{átbocsátó képesség} = m\varrho\mu = \min\{\lambda, m_\lambda\}.$$

A legfontosabb teljesítménymérő az igények szempontjából az az idő, amit az igény a várakozási sorban, vagy a rendszerben tölt. Jelölje W_j a várakozási időt, mint a j-edik igénynek a várakozási sorban eltöltött idejét, T_j pedig a válaszidőt, mint az igény által ezen rendszerben eltöltött teljes idejét. Ezen jelöléseket használva kapjuk, hogy

$$T_j = W_j + S_j,$$

ahol S_j a j-edik igény kiszolgálási ideje.

A rendszer teljesítményének vizsgálata történhet a várakozási sor hosszának mérésével is. Jelölje $Q(t)$ (valószínűségi változó) a t időpillanatban a rendszerben található igények számát. Egy rendszerben lévő igény vagy a várakozási sorban van, vagy éppen kiszolgálás altt áll. Így m szerveres rendszer esetén:

$$Q(t) = \max\{0,\ N(t) - m\}.$$

Mielőtt rátérnénk a konkrét elemi sorbanállási rendszerek vizsgálatára, bevezetjük azt a Kendalltól származó jelölésrendszert, amelynek segítségével osztályozhatjuk a sorbanállási rendszereket. Ez a következő lesz:

$$N|A|B|m|K,$$

ahol

N az igényforrás halmazának számossága,
A a beérkezési időközök eloszlásfüggvénye,
B a kiszolgálási idő eloszlásfüggvénye,
m a kiszolgáló egységek (csatornák) száma,
K a várakozó igények számának maximuma.

Ha az említett eloszlások exponenciálisak, akkor az M jelölést használjuk. Továbbá, ha a rendszer befogadóképessége és az igényforrás végtelen, akkor ezeket a jelöléseket elhagyjuk. Így például az $M|M|1$ rendszer egy 1 kiszolgálós rendszer Poisson beérkezéssel és exponenciális kiszolgálási idővel; az igényforrás számossága és a rendszer befogadó képessége végtelen. Vagy például az $M|G|m$ rendszernél a beérkezések Poisson-folyamat szerint történnek, a kiszolgálási idők általános eloszlásúak, m szerver áll rendelkezésünkre. Az igényforrás számossága és a rendszer befogadó képessége ez esetben is végtelen. Harmadik példaként tekintsük az $N|M|M|r$ rendszert, ahol az igények egy N elemű forrásból származnak, a beérkezési folyamat Poisson-folyamat, a kiszolgálási idő exponenciális eloszlást követ, a rendszerben r kiszolgáló egység van. Ugyanakkor ezen rendszer befogadóképessége végtelen.

III.3. Konkrét sorbanállási modellek

Előbb olyan modellekkel fogunk foglalkozni, ahol a beérkezési folyamat Poisson-folyamat, azaz a forrásunk végtelennek tekinthető. Ezt követően néhány véges forrású modellt is bemutatunk, minthogy ilyen modellre vezető problémák a gyakorlatban is előfordulhatnak.

1. modell: $M|M|1$ tipusú klasszikus modell. Az $M|M|1$ modell a legegyszerűbb nemtriviális modell. A beérkezési folyamat λ paraméterű Poisson-folyamat, a beérkezési időközök pedig λ paraméterű exponenciális eloszlású valószínűségi változók. A modellben egy kiszolgáló egység (csatorna) van.

A kiszolgálási idők μ paraméterű exponenciális eloszlású valószínűségi változók. Jelölje most $X(t)$ a t-edik időpillanatban a rendszerben tartózkodó igények számát és legyen a rendszer a k állapotban. Mivel a fellépő valószínűségi változók exponenciális eloszlásúak, ezért az $X(t)$ folyamat folytonos idejű Markov-lánc.

Vizsgáljuk meg a rendszer állapotváltozásainak valószínűségeit egy adott h időtartam alatt:

$$P_{k,k+1}(h) = \big(\lambda h + o(h)\big) \cdot \Big(1 - \big(\mu h + o(h)\big)\Big) + \\ + \sum_{k=2}^{\infty} \big(\lambda h + o(h)\big)^k \big(\mu h + o(h)\big)^{k-1} \qquad k = 0, 1, 2, \ldots$$

Az összeg első tagja annak a valószínűsége, hogy a rendszerbe egy igény érkezett és egyet sem szolgáltak ki. Az összeg második tagja pedig annak a valószínűségét adja, hogy a rendszerbe 2 vagy több igény érkezett és a beérkezettnél eggyel kevesebb került kiszolgálásra. Ez a valószínűség éppen $o(h)$-val egyenlő, így

$$P_{k,k+1}(h) = \lambda h + o(h).$$

Ehhez hasonlóan írható fel annak a valószínűsége, hogy a rendszer a k állapotban volt és a h időtartam után a $k-1$ állapotba került:

$$P_{k,k-1}(h) = \big(\mu h + o(h)\big) \cdot \Big(1 - \big(\lambda h + o(h)\big)\Big) + \\ + \sum_{k=2}^{\infty} \big(\lambda h + o(h)\big)^{k-1} \big(\mu h + o(h)\big)^{k} = \mu h + o(h).$$

Könnyen belátható továbbá az is, hogy

$$P_{k,j} = o(h), \quad \text{ha } |k-j| \geq 2.$$

Ezek szerint egy olyan születési-halálozási folyamattal van dolgunk, amit a születési és halálozási intenzitások következő megválasztásával lehet jellemezni:

$$\lambda_k = \lambda \qquad k = 0, 1, 2, \ldots$$
$$\mu_k = \mu \qquad k = 1, 2, 3, \ldots,$$

vagyis az összes születési intenzitás λ, az összes halálozási intenzitás pedig μ.

Ebben a modellben feltesszük azt is, hogy végtelen hosszúságú sorok is létrejöhetnek, és hogy az igények kiszolgálása a FIFO elv alapján történik.

A stacionárius megoldásra a

$$p_k = p_0 \left(\frac{\lambda}{\mu}\right)^k, \qquad k \geq 0$$

adódik. Az ergodikusság feltétele most az

$$S_1 = \sum_{k=0}^{\infty} \left(\frac{\lambda}{\mu}\right)^k < \infty, \qquad S_2 = \sum_{k=0}^{\infty} \frac{1}{\lambda} \left(\frac{\mu}{\lambda}\right)^k = \infty$$

formában adható meg, ami pontosan akkor fog teljesülni, ha $\lambda < \mu$. A normalizáló feltétel felhasználásával meghatározhatjuk a p_0 értékét is:

$$p_0 = \frac{1}{1 + \sum\limits_{k=1}^{\infty} \left(\frac{\lambda}{\mu}\right)^k} = \frac{1}{1 + \frac{\frac{\lambda}{\mu}}{1 - \frac{\lambda}{\mu}}} = 1 - \frac{\lambda}{\mu}.$$

A kihasználtsági tényezőre $\varrho = \frac{\lambda}{\mu}$ fog adódni. A $p_0 > 0$ biztosítva van, hiszen a $\lambda < \mu$ teljesül. Így a stacionárius megoldásra

$$p_k = (1 - \varrho)\varrho^k = \left(1 - \frac{\lambda}{\mu}\right) \cdot \left(\frac{\lambda}{\mu}\right)^k, \qquad k = 0, 1, 2, \ldots$$

adódik, ami valóban valószínűségi eloszlás, nevezetesen a geometriai eloszlás.

Az $M|M|1$ modellel megadható sorbanállási rendszer jellemzői:

- A rendszerben tartózkodó igények általános száma:

$$\overline{N} = \sum_{k=0}^{\infty} k p_k = \frac{\varrho}{1 - \varrho} = \frac{\frac{\lambda}{\mu}}{1 - \frac{\lambda}{\mu}}.$$

- A rendszerben tartózkodó igények számának szórásnégyzete:

$$\sigma_N^2 = \sum_{k=0}^{\infty} \left(k - \overline{N}\right)^2 p_k = \frac{\varrho}{(1 - \varrho)^2} = \frac{\frac{\lambda}{\mu}}{\left(1 - \frac{\lambda}{\mu}\right)^2}.$$

- A várakozó igények átlagos száma (átlagos sorhossz):

$$\overline{Q} = \sum_{k=1}^{\infty}(k-1)p_k = \overline{N} - (1-p_0) = \overline{N} - \varrho = \frac{\varrho^2}{1-\varrho} = \frac{\left(\dfrac{\lambda}{\mu}\right)^2}{1-\dfrac{\lambda}{\mu}}.$$

- Az átlagos sorhossz szórásnégyzete:

$$\sigma_Q^2 = \sum_{k=1}^{\infty}(k-1)^2 p_k - \overline{Q}^2 = \frac{\varrho^2(1+\varrho-\varrho^2)}{(1-\varrho)^2}.$$

- A szerver (kiszolgáló egység) kihasználtsága:

$$U_s = 1 - p_0 = \frac{\lambda}{\mu} = \varrho.$$

- A kiszolgáló átlagos foglaltsági periódushossza:

$$E\delta = \frac{1}{\mu-\lambda}$$

(Ezt a $p_0 = \dfrac{\dfrac{1}{\lambda}}{\dfrac{1}{\lambda} + E\delta}$ összefüggésből lehet meghatározni.)

- Egy igény várakozási idejének eloszlásával kapcsolatban kimutatható, hogy ha az igények a Poisson-folyamat szerint érkeznek be, akkor $P_k(t) = R_k(t)$, azaz annak a valószínűsége, hogy a rendszer a k állapotban legyen, megegyezik annak a valószínűségével, hogy a beérkező igény a rendszert a k állapotban találja.

  Ha egy tetszőleges pillanatban egy igény érkezik, akkor annak a valószínűsége, hogy nem kell várakoznia, p_0. Ha az érkezés pillanatában n igény tartózkodik a rendszerben ($n \geq 1$), akkor az érkező igénynek meg kell várnia, amíg az előtte lévő igényeket ki nem szolgálják és azok elhagyják a rendszert.
  Az átlagos várakozási idő:

$$\overline{W} = \overline{N} \cdot \frac{1}{\mu}.$$

- Az átlagos rendszerbeli tartózkodási idő:

$$\overline{T} = \frac{1}{\mu(1-\varrho)} = \frac{1}{\mu-\lambda}.$$

A két utóbbi összefüggésből kapjuk, hogy

$$\lambda\overline{T} = \overline{N}$$
$$\lambda\overline{W} = \overline{Q}.$$

Ezeket az összefüggéseket Little-formuláknak nevezzük, amelyek az előbbinél általánosabb esetekben is igazak.

2. modell: $M|M|1|K$, azaz véges befogadóképességű sorbanállási rendszer modellje. Olyan sorbanállási rendszert vizsgálunk, amelyben rögzített a várakozó igények számának maximuma. Feltesszük tehát, hogy a rendszerben legfeljebb K igény tartózkodhat, beleértve a kiszolgáló egységben lévő igényt is. Ezen felül egyetlen igény sem léphet be a rendszerbe, hanem azonnal távozik anélkül, hogy kiszolgálnák. Az igények továbbra is Poisson-folyamat szerint érkeznek, azonban csak azok az igények léphetnek be a rendszerbe, amelyek érkezésekor kevesebb, mint K igény van a rendszerben.

Ennek a látszólag bonyolult rendszernek a leírását a következőképpen tudjuk összhangba hozni a születési-halálozási modellel:

$$\lambda_k = \begin{cases} \lambda, & \text{ha } k < K \\ 0, & \text{ha } k \geq K \end{cases}$$
$$\mu_k = \mu, \quad \text{ha} \quad k = 1, 2, \ldots, K.$$

Ez a rendszer mindig ergodikus, hiszen az állapottere véges. Továbbá

$$p_k = p_0 \prod_{j=0}^{k-1} \frac{\lambda}{\mu} = p_0 \left(\frac{\lambda}{\mu}\right)^k, \qquad \text{ha } k \leq K$$
$$p_k = 0, \qquad \text{ha } k > K.$$

A normalizáló feltétel alapján

$$p_0 = \left[1 + \sum_{k=1}^{K} \left(\frac{\lambda}{\mu}\right)^k\right]^{-1} = \frac{1 - \dfrac{\lambda}{\mu}}{1 - \left(\dfrac{\lambda}{\mu}\right)^{K+1}},$$

így

$$p_k = \begin{cases} \dfrac{1 - \frac{\lambda}{\mu}}{1 - \left(\frac{\lambda}{\mu}\right)^{K+1}} \cdot \left(\dfrac{\lambda}{\mu}\right)^k, & \text{ha } 0 \leq k \leq K \\ 0, & \text{egyébként.} \end{cases}$$

A $K = 1$ esetén adódó $M|M|1|1$ modell azt jelenti, hogy egyáltalán nincs várakozás. Ekkor

$$p_k = \begin{cases} \dfrac{1}{1+\frac{\lambda}{\mu}}, & \text{ha } k = 0 \\ \dfrac{\frac{\lambda}{\mu}}{1+\frac{\lambda}{\mu}}, & \text{ha } k = 1 = K \\ 0, & \text{egyébként.} \end{cases}$$

3. modell: az $M|M|n$ tipusú sorbanállási rendszer modellje. Ez ismét olyan λ állandó beérkezési intenzitású sorbanállási rendszer modellje, amelyben korlátlan hosszúságú sor kialakulása is megengedett. Ez a rendszer n darab kiszolgáló egységgel (csatornával, szerverrel) van ellátva. Az igények Poisson-folyamat szerint érkeznek, tehát a beérkezési időközök λ paraméterű exponenciális eloszlású valószínűségi változók. A kiszolgálási idők pedig μ paraméterű exponenciális eloszlást követnek. Ez az eset is leírható születési-halálozási folyamattal.

Annak a valószínűsége, hogy a rendszer h idő alatt a k állapotból $k-1$ állapotba, illetve a k állapotból a $k+1$ állapotba kerül

$$\begin{aligned} P_{k,k-1}(h) =& \mu_k \cdot h + o(h) \\ P_{k,k+1}(h) =& \lambda h + o(h) \end{aligned}$$

lesz, ahol

$$\mu_k = \min\{k\mu, n\mu\} = \begin{cases} k\mu, & \text{ha } 0 \le k \le n \\ n\mu, & \text{ha } n < k. \end{cases}$$

Az ergodikusság feltétele most a $\frac{\lambda}{n\mu} < 1$ egyenlőtlenség.

A stacionárius eloszláshoz tartozó valószínűségekre a következők adódnak:

$$p_k = \begin{cases} p_0 \left(\dfrac{\lambda}{\mu}\right)^k \cdot \dfrac{1}{k!}, & \text{ha } k < n \\ p_0 \left(\dfrac{\lambda}{\mu}\right)^k \cdot \dfrac{1}{n!\; n^{k-n}}, & \text{ha } k \ge n. \end{cases}$$

A $\frac{\lambda}{n\mu} < 1$ értékre szokás használni a ϱ jelölést is, amit kihasználtsági tényezőnek neveznek. A normalizáló feltétel alapján meghatározhatjuk a p_0

értékét is:

$$p_0 = \left(1 + \sum_{k=1}^{n-1} \frac{(n\varrho)^k}{k!} + \sum_{k=n}^{\infty} \frac{(n\varrho)^k}{n!} \cdot \frac{1}{n^{k-n}}\right)^{-1} =$$

$$= \left(\sum_{k=0}^{n-1} \frac{(n\varrho)^k}{k!} + \frac{(n\varrho)^k}{n!} \cdot \frac{1}{1-\varrho}\right)^{-1}.$$

Annak a valószínűsége, hogy egy újonnan érkező igénynek sorba kell állnia:

$$P(\text{sorbanállás}) = \sum_{k=n}^{\infty} p_k = \sum_{k=n}^{\infty} p_0 \frac{(n\varrho)^k}{n!} \frac{1}{n^{k-n}}.$$

Az $M|M|n$ sorbanállási rendszer jellemzői:

- Átlagos sorhossz:

$$\overline{Q} = \sum_{k=n}^{\infty} (k-n) \cdot p_k = \sum_{j=0}^{\infty} j \cdot p_{n+j} = p_0 \frac{\left(\frac{\lambda}{\mu}\right)^n}{n!} \cdot \frac{\varrho}{(1-\varrho)^2}.$$

- A foglalt szerverek átlagos száma:

$$\overline{n} = \sum_{k=0}^{n-1} kp_k + \sum_{k=n}^{\infty} np_k = n \cdot \varrho = n \cdot \frac{\lambda}{n\mu} = \frac{\lambda}{\mu}.$$

- A rendszerben tartózkodó igények átlagos száma:

$$\overline{N} = \sum_{k=0}^{\infty} kp_k = \overline{n} + \overline{Q}.$$

  Ez utóbbi azon megfontolásból is nyerhető, hogy a rendszerben tartózkodó igény vagy várakozik, vagy kiszolgálás alatt van. A kiszolgálás alatt lévő igények száma viszont megegyezik a foglalt kiszolgálóegységek számával.

- Átlagos várakozási idő:

$$\overline{W} = \frac{\left(\frac{\lambda}{\mu}\right)^n}{n} \cdot p_0 \cdot \frac{1}{(1-\varrho)^2 \cdot n \cdot \mu}.$$

- Átlagos tartózkodási idő:

$$\overline{T} = \frac{1}{\mu} + \overline{W}.$$

Stacionárius esetben a rendszerből távozó igények átlagos számának meg kell egyeznie az érkező igények átlagos számával, így a rendszerben tartózkodók átlagos száma állandó. Így

$$\lambda\overline{T} = \overline{N} = \overline{Q} + \overline{n}$$

továbbá

$$\lambda\overline{W} = \overline{Q}.$$

Ezek lesznek a Little-formulák.

- A szerverek összkihasználtsága:

  A rendszerben n darab szerver van. Ezek akkor foglaltak, ha legalább 1, legalább $2, \ldots$, legalább n foglalt. Így

$$U_n = \sum_{k=1}^{n} \sum_{j=k}^{\infty} p_j = n\varrho,$$

  amiből egy szerver kihasználtságára $\frac{n\varrho}{n} = \varrho$ adódik.

- Foglaltsági periódushosszak:

  A sorbanállási rendszert akkor nevezzük tétlennek, ha a rendszerben nem tartózkodik igény, minden más esetben foglaltnak nevezzük. Ha $E\delta^{(n)}$ jelöli a remdszer átlagos foglaltsági periódushosszát, akkor arra az

$$E\delta^{(n)} = \frac{1 - p_0}{\lambda p_0}$$

  adódik. Jelölje $E\delta$ egy szervernek az átlagos foglaltsági periódushosszát. Erre

$$E\delta = \frac{1}{\mu - \lambda}$$

  fog adódni, amit az első modellnél is megkaptunk az egyetlen szerver foglaltsági periódushosszára.

4. modell: az $M|M|n|n$ tipusú modell, azaz az Erlang-féle veszteséges sorbanállási rendszer modellje. Ezen modellben n csatorna (szerver) található. A rendszerbe Poisson-folyamat szerint érkeznek az igények. Ha van üres csatorna (szerver) az igény kiszolgálására, akkor azt μ paraméterű exponenciális eloszlásnak megfelelően fogják kiszolgálni. Ha viszont minden kiszolgáló egység foglalt, akkor az igény elvész, így tehát sorbanállás nem jöhet szóba.

Ezen probléma a tömegkiszolgálás egyik legrégibb problémája, amellyel a század elején a telefonközpontok kihasználtságával kapcsolatban foglalkozott A. K. Erlang és C. Palm. Hasonló a helyzet a parkolóhelyek esetében is.

A feltételek alapján ez is egy születési-halálozási folyamat, amelynek intenzitásai a következők:

$$\lambda_k = \begin{cases} \lambda, & \text{ha } k < n \\ 0, & \text{ha } k \geq n \end{cases}$$

$$\mu_k = k \cdot \mu, \qquad k = 1, 2, \ldots, n.$$

Azt mondjuk, hogy a folyamat k állapotban van, ha k szerver foglalt, azaz ha k igény tartózkodik a rendszerben. Az ergodikus eloszlás is létezik, mivel a folyamat véges állapotterű. A folyamat stacionárius eloszlása:

$$p_k = \begin{cases} p_0 \left(\frac{\lambda}{\mu}\right)^k \frac{1}{k!}, & \text{ha } k \leq n \\ 0, & \text{ha } k > n. \end{cases}$$

A normalizáló feltétel miatt

$$p_0 = \left(\sum_{k=0}^{n} \left(\frac{\lambda}{\mu}\right)^k \frac{1}{k!} \right)^{-1},$$

így

$$p_k = \frac{\left(\frac{\lambda}{\mu}\right)^k \frac{1}{k!}}{\sum\limits_{i=0}^{n} \left(\frac{\lambda}{\mu}\right)^i \frac{1}{i!}} = \frac{\frac{\varrho^k}{k!}}{\sum\limits_{i=0}^{n} \frac{\varrho^i}{i!}}.$$

A rendszer egyik jellemzője a

$$p_n = \frac{\frac{\varrho^n}{n!}}{\sum\limits_{k=0}^{n} \frac{\varrho^k}{k!}}$$

valószínűség, melyet először Erlang vezetett be, s amelyet Erlang-féle veszteségformulának neveznek és $B(n, \frac{\lambda}{\mu})$ szimbólummal jelölnek. Ez a valószínűség annak a valószínűsége, hogy egy újonnan érkező igényt nem fogad a rendszer, azaz az igény elvész.

Kis n-re a p_0 valószínűség könnyen kiszámolható, vagy n értékre és kis ϱ-ra pedig $p_0 \approx e^{-\varrho}$, így

$$p_k \approx \frac{\varrho^k}{k!} e^{-\varrho},$$

ami megegyezik a Poisson-eloszlás $(k+1)$-edik tagjával. Nagy n-re és nagy ϱ-ra általában

$$\sum_{j=0}^{n} \frac{\varrho^j}{j!} \neq e^{\varrho} .$$

Ebben az esetben a nevező a ϱ közepű Poisson-eloszlás első $(n+1)$ tagjának összege. Elegendő nagy ϱ értékre $(\varrho > 15)$ a Poisson-eloszlást közelítjük a ϱ közegű és $\sqrt{\varrho}$ szórású normális eloszlással, így

$$p_n \approx \frac{\frac{\varrho^n}{n!} e^{-\varrho}}{\Phi(s)},$$

ahol

$$\Phi(s) = \int_{-\infty}^{s} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \quad \text{és} \quad s = \frac{n + \frac{1}{2} - \varrho}{\sqrt{\varrho}}.$$

Az $M|M|n|n$ rendszer jellemzői:

- A rendszerben tartózkodó igények száma, amely megegyezik a foglalt szerverek átlagos számával:

$$\overline{N} = \overline{n} = \sum_{j=0}^{n} j p_j = \sum_{j=0}^{n} j \frac{\varrho^j}{j!} p_0 = \varrho \sum_{j=0}^{n-1} \frac{\varrho^i}{i!} p_0 = \varrho(1 - p_n).$$

  Így az 1 szerverre jutó átlagos igényszám:

$$\frac{\varrho}{n}(1 - p_n).$$

- A szerverek kihasználtsága:
  A szerverek akkor foglaltak, ha a rendszerben legalább 1, legalább 2, ..., legalább n igény tartózkodik, így

$$U_n = \sum_{i=1}^{n} \sum_{j=i}^{n} p_j = \sum_{k=0}^{n} j p_j = \overline{n} ,$$

  amiből

$$U_s = \frac{\overline{n}}{n} = \frac{\varrho}{n}(1 - p_n).$$

- Átlagos tétlenségi idő (egy konkrét kiszolgáló esetén):

$$\overline{e} = \frac{n}{\lambda(1 - p_n)} - \frac{1}{\mu}.$$

  Ha az üres szerverek olyan sorrendben kezdik kiszolgálni az érkező igényeket, mint amilyen sorrendben azok megüresedtek, akkor egy szerver működését a következőképpen írhatjuk le. Ha egy üressé vált szerver $(j-1)$ másik üres szervert talál a megüresedése pillanatában, akkor csak a j-edik igény kiszolgálásával kezdődik ismét a foglaltsági periódusa.

- A rendszer átlagos foglaltsági periódushossza:

$$E\delta^{(n)} = \frac{1-p_0}{\lambda p_0} = \frac{\sum\limits_{i=1}^{n} \dfrac{\varrho^i}{i!}}{\lambda\left(1+\sum\limits_{i=1}^{n}\dfrac{\varrho^i}{i!}\right)}.$$

5. modell: az $n|M|M|1$ tipusú sorbanállási rendszer modellje. Feltesszük, hogy az igények egy n elemű véges forrásból érkeznek, ahol a forrásban eltöltött idő minden igény esetén egymástól független λ paraméterű exponenciális eloszlású valószínűségi változó. A kiszolgálási idők szintén exponenciális eloszlású valószínűségi változók μ paraméterrel, amelyek függetlenek a forrásban eltöltött időtől.

Jelölje $x(t)$ a t időpillanatban a kiszolgáló egységnél tartózkodó igények számát. Ez is egy születési-halálozási folyamat, amelynél

$$\lambda_k = \begin{cases} (n-k)\lambda, & \text{ha } 0 \le k \le n \\ 0, & \text{ha } k > n \end{cases}$$

a születési intenzitás és $\mu_k = \mu, k \ge 1$ a kihalási intenzitás. Így

$$p_k = \frac{n!}{(n-k)!} \cdot \varrho^k \cdot p_0 = (n-k+1)\cdot\varrho\cdot p_{k-1},$$

ahol

$$\varrho = \frac{\lambda}{\mu}$$

és

$$p_0 = \frac{1}{1+\sum\limits_{k=1}^{n}\dfrac{n!}{(n-k!)}\varrho^k} = \frac{1}{\sum\limits_{k=0}^{n}\dfrac{n!}{(n-k)!}\varrho^k}.$$

Az ergodikus eloszlás mindig létezik, de ha $\varrho > 1$, akkor az igények torlódnak, ami miatt több kiszolgáló egységre lenne szükség.

Az $n|M|M|1$ rendszer jellemzői:

- A szerver kihasználtsága és a rendszer átbocsátóképessége:
  A szerver kihasználtsága: $U_s = 1 - p_0$.
  A rendszer átbocsátóképessége: $\lambda_t = \mu U_s$.

- A rendszerben tartózkodó igények átlagos száma:

$$\overline{N} = \sum_{k=0}^{n} kp_k = n - \sum_{k=0}^{n}(n-k)p_k = n - \frac{1}{\varrho}\sum_{k=0}^{n}(n-k)\varrho p_k =$$
$$= n - \frac{1}{\varrho}\sum_{k=0}^{n-1} p_{k+1} = n - \frac{1}{\varrho}(1-p_0) = n - \frac{U_s}{\varrho}.$$

- Az átlagos sorhossz:

$$\overline{Q} = \sum_{k=1}^{n}(k-1)p_k = \sum_{k-1}^{n} kp_k - \sum_{k=1}^{n} p_k = n - \frac{\mu}{\lambda}(1-p_0) -$$
$$-(1-p_0) = n - 1 + \frac{1}{\varrho}U_s.$$

- Az igénygenerálásra alkalmas terminálok átlagos száma:

$$\overline{m} = \sum_{k=0}^{n}(n-k)p_k = n - \overline{N} = \frac{\mu}{\lambda}(1-p_0) = \frac{U_s}{\varrho}.$$

- A szerver átlagos foglaltsági periódushossza:

$$E\delta = \frac{1-p_0}{n\lambda p_0} = \frac{U_s}{n\lambda(1-U_s)}.$$

- A várakozás valószínűsége:

$$P(W>0) = \sum_{k=1}^{n} p_k = 1 - p_0 = U_s.$$

Számítógépes alkalmazásoknál gyakran van szükségünk még a következő jellemzőkre is.

- A terminálok kihasználtsága:
  Az i indexű terminál kihasználtságán az

$$U^{(i)} \lim_{T\to\infty} \frac{1}{T}\int_0^T \chi(\text{a } T\text{-edik időpillanatban az } i\text{-edik terminál működik})dt,$$

határértéket értjük, ha létezik. Ekkor tehát

$$U^{(i)} = P(\text{az } i\text{-edik terminál működik}),$$

ahol a P a stacionárius valószínűséget jelöli.

Nyilvánvaló, hogy a terminálok akkor vannak kihasználva, ha működnek, így az összes terminál kihasználtsága:

$$\sum_{i=1}^{n}\sum_{k=1}^{i} p_k = \sum_{k=1}^{n}(n-k)p_k = \overline{m} = \frac{\mu}{\lambda}(1-p_0).$$

Egy tetszőleges terminál kihasználtsága:

$$U_t = \frac{\mu}{n\lambda}(1-p_0) = \frac{\overline{m}}{n}.$$

Mivel a terminálok azonos kihasználtságúak, ezért

$$U_t = U^{(i)}.$$

- A terminálok átlagos várakozási ideje:
  Felhasználva az

$$U_t = \frac{\frac{1}{\lambda}}{\frac{1}{\lambda}+\overline{W}+\frac{1}{\mu}} = \frac{\overline{m}}{n}$$

összefüggést, a

$$\lambda\overline{m}\overline{W} = n - \overline{m}\left(1+\frac{\lambda}{\mu}\right) = n - \frac{U_s}{\varrho}(1+\varrho) = \overline{Q}$$

adódik, ami az átlagos sorhosszra vonatkozó Little-formula. Így $\overline{W} = \frac{\overline{Q}}{\lambda\overline{m}}$.
Az átlagos válaszolási idő:

$$\overline{T} = \overline{W} + \frac{1}{\mu} = \frac{1}{\mu}\left(\frac{n}{U_s} - \frac{1}{\varrho}\right).$$

Egyszerű számolással könnyű bebizonyítani, hogy

$$\overline{m}\lambda\overline{T} = \overline{N},$$

ami szintén egy Little-formula.

6. modell: Az $n|M|M|r$ tipusú sorbanállási rendszer modellje. Az előző modellben megfogalmazott feltevéseinket most csupán annyiban változtatjuk, hogy az n számú terminált r szerver szolgálja ki ($r < n$). Így a $k \leq r$ esetén a k állapot azt jelenti, hogy éppen k darab terminál igénye van kiszolgálás alatt, egyetlen várakozó igény sincs, és $r-k$ szerver tétlen.

A szerverek tevékenységüket egymástól függetlenül végzik. Ekkor is egy születési-halálozási folyamatról van szó a

$$\lambda_k = (n-k)\lambda, \quad 0 \le k \le n-1$$

$$\mu_k = \begin{cases} k\mu, & 1 \le k \le r \\ r\mu, & r < k \le n \end{cases}$$

intenzitásokkal. A stacionárius megoldáshoz tartozó valószínűségekre

$$p_k = \binom{n}{k} \varrho^k p_0, \qquad \text{ha } 0 \le k \le r$$

$$p_k = \frac{k!}{r! r^{k-r}} \binom{n}{k} \varrho^k p_0, \qquad \text{ha } r < k \le n$$

adódik. Minthogy $\sum_{k=0}^{n} p_k = 1$ fenn kell, hogy álljon, ezért ebből ki lehetne számolni a p_0 értékét. Ezen összefüggés bonyolult volta miatt azonban más módszert szokás választani p_0 értékének megállapításához.

Az $n|M|M|r$ rendszer jellemzői:

- A rendszerben tartózkodó igények száma:

$$\overline{N} = \sum_{k=0}^{n} k p_k$$

- A várakozási sor átlagos hossza:

$$\overline{Q} = \sum_{k=r+1}^{n} (k-r) p_k = \frac{r^r p_0}{r!} \sum_{k=r+1}^{n} \frac{(k-r)k!}{r^k} \binom{n}{k} \varrho^k$$

- Az igénygenerálásra alkalmas terminálok átlagos száma:

$$\overline{m} = n - \overline{N}$$

- A rendszer összkihasználtsága:

$$U_n = 1 - p_0$$

így egy kiszolgálóegység kihasználtsága

$$U_s = \frac{U_n}{r}$$

- A rendszer átlagos foglaltsági periódushossza:

$$E\delta^{(n)} = \frac{1-p_0}{n\lambda p_0} = \frac{U_n}{n\lambda p_0}.$$

- A várakozás valószínűsége:

$$P(W > 0) = \sum_{k=r}^{n} p_k$$

- A foglalt kiszolgálóegységek átlagos száma:

$$\begin{aligned}\overline{r} = \sum_{k=1}^{r} kp_k + \sum_{k=r+1}^{n} rp_k = \sum_{k=1}^{r-1} kp_k + r\sum_{k=r}^{n} p_k = \\ = \sum_{k=1}^{r-1} kp_k + rP(W > 0).\end{aligned}$$

Ebből

$$U_s = \frac{\overline{r}}{r}.$$

- A tétlen kiszolgáló egységek átlagos száma:

$$\overline{S} = r - \overline{r}.$$

- A rendszerben tartózkodó igények átlagos száma:

$$\begin{aligned}\overline{N} &= \sum_{k=1}^{r} kp_k + \sum_{k=r+1}^{n} (k-r)p_k + r \cdot \sum_{k=r+1}^{n} p_k = \\ &= \overline{Q} + \overline{r} = \overline{Q} + r - \overline{S} = n - \overline{m}\end{aligned}$$

- A terminálok kihasználtsága:

$$U_t = \sum_{k=1}^{n} \frac{n-k}{n} p_k = \frac{\overline{m}}{n}.$$

- A terminálok átlagos várakozási ideje:

$$\overline{W} = \frac{\overline{N}}{\overline{m}} \cdot \frac{1}{\lambda} - \frac{1}{\mu} = \frac{1}{\mu}\left(\frac{\overline{N}}{\overline{m}\varrho} - 1\right).$$

- Az átlagos válaszolási idő:

$$\overline{T} = \overline{W} + \frac{1}{\mu} = \frac{\overline{N}}{\overline{m}\lambda},$$

amiből

$$\overline{m}\lambda\overline{T} = \overline{N}$$

adódik, ami a már ismert Little-formula, ami szerint az átlagos beérkezési intenzitás és a rendszerben töltött átlagos idő szorzata a rendszerben tartózkodó igények átlagos számával egyenlő.

Az $\overline{N}$ és $\overline{T}$ értékeit felhasználva vezethetjük be a

$$\overline{Q} = \overline{m}\lambda\overline{W}$$

összefüggést, ami szintén egy Little-formula. ($\overline{Q}$ a várakozási sor átlagos hosszát jelöli.)

- A kiszolgáló egységek átlagos tétlenségi periódushossza:

  Ha a tétlen kiszolgálók olyan sorrendben kezdik kiszolgálni az igényeket, mint amilyen sorrendben előzőleg befejezték a foglaltsági periódusukat, akkor egy szerver tevékenységét a következőképpen írhatjuk le. Ha egy tétlenné vált szerver $j-1$ másik tétlen szervert talál a munkabefejeződés pillanatában, akkor csak a j-edik igény kiszolgálásával kezdődik ismét a foglaltsági periódusa.

  Jelölje $\overline{e}$ a szerver átlagos üresjárati periódusának hosszát, e_j pedig a fenti állapotban az átlagos tétlenségi időt. Nyilvánvalóan

$$\overline{e}_j = \frac{j}{\lambda}$$
$$\overline{e} = \sum_{j=1}^{r} \frac{p_{r-j}}{P(e)} \cdot \frac{j}{\lambda} = \frac{\overline{S}}{P(e) \cdot \lambda},$$

  ahol

$$P(e) = \sum_{j=0}^{r-1} p_j = 1 - P(W > 0),$$

  azaz annak a valószínűsége, hogy van tétlen szerver.

- A szerverek átlagos foglaltsági periódushossza:

$$E\delta = \frac{\overline{m}}{\mu \cdot P(e)}.$$

Irodalomjegyzék

[1] Beaumont G. P., *Introductory Applied Probability Ellis Harwood*, Chichester, 1983

[2] Csáth Magdolna, *Operációkutatás*, SZÁMOK, Budapest, 1973

[3] Éltető Ö. – Meszéna Gy. – Ziermann M., *Sztochasztikus módszerek és modellek*, Közgazdasági és Jogi Kiadó, Budapest, 1982.

[4] Ford L. R. Jr. – Fulkerson D. R., *Flow in Networks*, Princeton University Press, 1962.

[5] Gale D., *The Theory of Linear Economic Models*, Mc Graw Hill, 1960.

[6] Glevitzky Béla – Sztrik János, *Az operációkutatás elemei*, KLTE jegyzet, Debrecen, 1992.

[7] Gnedenko B. V. – Kovalenko I. N., *Bevezetés a tömegkiszolgálás elméletébe*

[8] Gross D. – Harris C. M., *Fundamentals of Queueing Theory*, John Wiley and Sons, Inc., New York, 1985.

[9] Klafszki Emil, *Hálózati folyamok*, Bolyai J. Matematikai Társulat kiadványa, Budapest, 1969.

[10] Kleinrock L., *Sorbanállás–kiszolgálás*, Bevezetés a tömegkiszolgálási rendszerek elméletébe. Műszaki Könyvkiadó, Budapest, 1979.

[11] Kobayashi H., *Modelling and Analysis*, An introduction to System Perforance Evaluation Methodology. Addison-Wesley Publishing Company. London, 1978.

[12] Kuhn H. W., *The Hungarian Method for the Assignment Problem*, Naval Res. Logist. Quart.,1., 83–97. 1955.

[13] Lukács O., *Operációkutatás*, Tankönyvkiadó, Budapest, 1982.

[14] Miller F. S. Liebermann G. I., *Introduction to operation research*, Walden-Day, Inc. SanFrancisco, 1974.

[15] Dr. Nagy Tamás, *Matematikai programozás*, Tankönyvkiadó, Budapest, 1985.

[16] Ross. S. M., *Applied Probability Models with Optimalization Applications*, Holden-Day, San Francisco, 1970.

[17] Takács L., *Introduction to the Theory of Queues*, New York, Oxford University Press, 1962.

[18] Tomkó J., *A Markov-folyamatok elemei és néhány operációkutatási vonatkozása*, Bolyai János Matematikai Társulat kiadványa, Budapest, 1968.

[19] Tomkó J., *Tartózkodási idő problémák Markov-láncokra*, Alkalmazott Matematikai Lapok, 1982 (8), 91–106.