

Tömegkiszolgálási modellek

1 Sorbanállási modellek

A sorbanállási (tömegkiszolgálási) modellek alkotják a születési és halálozási folyamatok egyik legfontosabb alkalmazási területét. Ezekben a modellekben a rendszer állapotát a rendszerben tartózkodó igények száma adja. Az egyes modellek leírásához sokféle jellemzőre van szükségünk: hogyan érkeznek az igények, milyen folyamat szerint történik a kiszolgálás, hány kiszolgálóegység van a rendszerben, mekkora a rendszer befogadóképessége, milyen protokoll szerint történik a kiszolgálás. Általában feltételezzük, hogy az egyes igények beérkezései között eltelt idő független és azonos eloszlású, valamint az igények kiszolgálási idői egymástól és az érkezési folyamattól is függetlenek.

2 A tömegkiszolgálási rendszerek osztályozása

Egy tömegkiszolgálási rendszert egy, az alábbi szerkezetű kóddal jellemezhetünk¹:

$$A/B/m/k/n/P,$$

ahol

- A az egymást követő igények beérkezése között eltelt idő. Ez lehet
 - D determinisztikus (konstans),

¹Ez az ún. Kendall-féle jelölés

- M exponenciális (vagyis az igények Poisson-folyamat szerint érkeznek, tehát a beérkezés markovi),
 - G általános (tetszőleges).
- B a kiszolgálási idő. Ez lehet
 - D determinisztikus (konstans),
 - M exponenciális (markovi),
 - G általános (tetszőleges).
 - m a kiszolgáló egységek (szerverek) száma.
 - k a rendszer befogadó képessége (legfeljebb mennyi igény lehet jelen a rendszerben). Ha ezt külön nem tüntetjük fel, akkor azt feltételezzük, hogy a rendszer befogadó képessége végtelen.
 - n az igényforrások száma. Ha ezt külön nem tüntetjük fel, akkor végtelennek tekintjük.
 - P a kiszolgálási protokoll. Például
 - az első beérkezőt szolgálják ki először (FIFO: first in, first out; FCFS: first customer, first service),
 - az utolsó beérkezőt szolgálják ki először (LIFO: last in, first out; LCFS: last customer, first service),
 - véletlenszerű kiszolgálás,
 - az igényekhez rendelt prioritás alapján történő kiszolgálás.

Ha ezt külön nem tüntetjük fel, akkor FIFO rendszert tételezünk fel.

3 A Little–formula

A következőkben levezetünk egy egyszerű de fontos összefüggést a rendszerben tartózkodó igények átlagos száma és a rendszerben eltöltött átlagos idő között. Ehhez

be kell vezetnünk néhány jelölést.

$\alpha(t)$ = a $(0, t)$ időintervallumban beérkezett igények száma ,

$\delta(t)$ = a $(0, t)$ időintervallumban távozott igények száma .

Ekkor a rendszerben tartózkodó igények száma (feltéve, hogy a $t = 0$ időpontban nem tartózkodott igény a rendszerben) nyilvánvalóan

$$N(t) = \alpha(t) - \delta(t) .$$

Legyen $\lambda(t)$ az igények érkezési intenzitása (igények száma/időegység) a $(0, t)$ időintervallumban, tehát

$$\lambda(t) = \frac{\alpha(t)}{t} .$$

Tekintsük a $(0, t)$ időintervallum során az igények által a rendszerben eltöltött összes időt, jelölje ezt $\gamma(t)$. Legyen továbbá $T(t)$ az egy igényre eső rendszerbeli idő átlaga a $(0, t)$ időintervallumban (összes idő/összes igény). Ekkor

$$T(t) = \frac{\gamma(t)}{\alpha(t)} .$$

Legyen végül $\bar{N}(t)$ a rendszerben tartózkodó igények átlagos száma a $(0, t)$ időintervallumban. Ezt megkaphatjuk, ha a rendszerben eltöltött összes időt elosztjuk az intervallum hosszával (t) :

$$\bar{N}(t) = \frac{\gamma(t)}{t} .$$

Ezt átalakítva:

$$\bar{N}(t) = \frac{\gamma(t)}{t} = \frac{\alpha(t)}{t} \frac{\gamma(t)}{\alpha(t)} = \lambda(t)T(t) .$$

Tegyük fel, hogy az adott sorbanállási rendszerben léteznek a következő határértékek:

$$\lambda = \lim_{t \rightarrow \infty} \lambda(t) ,$$

$$E(T) = \lim_{t \rightarrow \infty} T(t) = \lim_{t \rightarrow \infty} \frac{\gamma(t)}{\alpha(t)} .$$

Ekkor $\bar{N}(t)$ határértéke is létezik, és

$$E(N) = \lim_{t \rightarrow \infty} \bar{N}(t) .$$

Mindezeket összevetve kaptuk tehát a következőt (Little-formula):

$$E(N) = \lambda E(T) ,$$

azaz

$$\begin{aligned} & (\text{a rendszerben levő igények átlagos száma}) = \\ & = (\text{az igények érkezési intenzitása}) \times (\text{a rendszerben eltöltött átlagos idő}) . \end{aligned}$$

4 Az $M/M/1$ rendszer

Az ismerttetett jelölések szerint most az igények Poisson-folyamat szerint érkeznek (a paraméter legyen λ), a kiszolgálási idő pedig exponenciális eloszlású (az exponenciális eloszlás paramétere legyen μ). Tételezzük fel azt is, hogy $\mu > \lambda$, vagyis egységnyi idő alatt átlagosan több igényt tudnak kiszolgálni, mint amennyi érkezik (különben a sor hossza a végtelenségig növekedne, a rendszer összeomlana). Csak egyetlen kiszolgáló egység van, a rendszer kapacitása végtelen (akármennyi lehet a várakozók száma), az igények forrása végtelen, a kiszolgálási protokoll FIFO.

A rendszer állapota legyen a rendszerben levő igények száma: $S = \{0, 1, 2, \dots\}$. Ekkor a folyamat infinitezimális generátora:

$$A = \begin{bmatrix} * & \lambda & 0 & \dots & \dots & \dots \\ \mu & * & \lambda & 0 & 0 & \dots \\ 0 & \mu & * & \lambda & 0 & \dots \\ \vdots & & \ddots & \ddots & \ddots & \ddots \end{bmatrix} .$$

Az invariáns eloszlás meghatározásához a $\underline{0} = \underline{\pi}A$ egyenletet kell megoldanunk. Koordinátánként kiírva:

$$\begin{aligned} 0 &= -\lambda\pi_0 + \mu\pi_1 & \implies \pi_1 &= \frac{\lambda}{\mu}\pi_0 \\ 0 &= \lambda\pi_0 - (\lambda + \mu)\pi_1 + \mu\pi_2 & \implies \pi_2 &= \frac{\lambda}{\mu}\pi_1 = \left(\frac{\lambda}{\mu}\right)^2 \pi_0 \\ 0 &= \lambda\pi_1 - (\lambda + \mu)\pi_2 + \mu\pi_3 & \implies \pi_3 &= \frac{\lambda}{\mu}\pi_2 = \left(\frac{\lambda}{\mu}\right)^3 \pi_0 \\ &\vdots & & \vdots \\ 0 &= \lambda\pi_{k-2} - (\lambda + \mu)\pi_{k-1} + \mu\pi_k & \implies \pi_k &= \frac{\lambda}{\mu}\pi_{k-1} = \left(\frac{\lambda}{\mu}\right)^k \pi_0 \\ &\vdots & & \vdots \end{aligned}$$

Mivel π valószínűség eloszlás, ezért az elemeinek összege 1, tehát

$$\begin{aligned}
 \sum_{k=0}^{\infty} \pi_k &= 1 = \\
 &= \pi_0 + \pi_1 + \pi_2 + \pi_3 + \dots = \pi_0 + \frac{\lambda}{\mu} \pi_0 + \left(\frac{\lambda}{\mu}\right)^2 \pi_0 + \left(\frac{\lambda}{\mu}\right)^3 \pi_0 + \dots = \\
 &= \pi_0 \underbrace{\left(1 + \frac{\lambda}{\mu} + \left(\frac{\lambda}{\mu}\right)^2 + \left(\frac{\lambda}{\mu}\right)^3 + \dots\right)}_{\frac{1}{1-\frac{\lambda}{\mu}}} = \pi_0 \frac{1}{1-\frac{\lambda}{\mu}} \implies \pi_0 = 1 - \frac{\lambda}{\mu}.
 \end{aligned}$$

Ezt felhasználva kapjuk, hogy

$$\pi_k = \left(1 - \frac{\lambda}{\mu}\right) \left(\frac{\lambda}{\mu}\right)^k.$$

Határozzuk meg a rendszerben levő igények számának várható értékét ($E(N)$).

$$\begin{aligned}
 E(N) &= \sum_{k=0}^{\infty} k \pi_k = \sum_{k=0}^{\infty} k \left(1 - \frac{\lambda}{\mu}\right) \left(\frac{\lambda}{\mu}\right)^k = \sum_{k=1}^{\infty} k \left(1 - \frac{\lambda}{\mu}\right) \left(\frac{\lambda}{\mu}\right)^k = \\
 &= \left(1 - \frac{\lambda}{\mu}\right) \frac{\lambda}{\mu} \underbrace{\sum_{k=1}^{\infty} k \left(\frac{\lambda}{\mu}\right)^{k-1}}_{\frac{1}{\left(1-\frac{\lambda}{\mu}\right)^2}} = \left(1 - \frac{\lambda}{\mu}\right) \frac{\lambda}{\mu} \frac{1}{\left(1 - \frac{\lambda}{\mu}\right)^2} = \\
 &= \frac{\lambda}{\mu} \frac{1}{1 - \frac{\lambda}{\mu}} = \frac{\lambda}{\mu} \frac{\mu}{\mu - \lambda} = \frac{\lambda}{\mu - \lambda}.
 \end{aligned}$$

A Little-formula segítségével könnyen meghatározhatjuk a rendszerben eltöltött idő várható értékét:

$$E(N) = \lambda E(T) \implies E(T) = \frac{E(N)}{\lambda} = \frac{\lambda}{\mu - \lambda} \frac{1}{\lambda} = \frac{1}{\mu - \lambda}.$$

A tömegkiszolgálási folyamat egy igényre nézve két fő részből áll. Először az igény arra vár, hogy sorra kerüljön (addig nyilván az előtte lévőeket szolgálják ki), majd az ő kiszolgálása következik. Határozzuk meg most a várakozó igények számának várható értékét ($E(N_w)$). Ha csak egy igény van a rendszerben, akkor az nem várakozik, hanem egyből sorra kerül, ha pedig k igény van a rendszerben, akkor

egyét éppen kiszolgálják, a többi $k - 1$ pedig várakozik. Tehát

$$\begin{aligned} E(N_w) &= \sum_{k=2}^{\infty} (k-1) \pi_k = \sum_{k=2}^{\infty} (k-1) \left(1 - \frac{\lambda}{\mu}\right) \left(\frac{\lambda}{\mu}\right)^k = \left(1 - \frac{\lambda}{\mu}\right) \left(\frac{\lambda}{\mu}\right)^2 \sum_{k=2}^{\infty} (k-1) \left(\frac{\lambda}{\mu}\right)^{k-2} = \\ &= \left(1 - \frac{\lambda}{\mu}\right) \left(\frac{\lambda}{\mu}\right)^2 \sum_{j=1}^{\infty} j \left(\frac{\lambda}{\mu}\right)^{j-1} = \left(1 - \frac{\lambda}{\mu}\right) \left(\frac{\lambda}{\mu}\right)^2 \frac{1}{\left(1 - \frac{\lambda}{\mu}\right)^2} = \left(\frac{\lambda}{\mu}\right)^2 \frac{1}{1 - \frac{\lambda}{\mu}} = \\ &= \left(\frac{\lambda}{\mu}\right)^2 \frac{\mu}{\mu - \lambda} = \frac{\lambda^2}{\mu} \frac{1}{\mu - \lambda}. \end{aligned}$$

A Little-formula segítségével a várakozással eltöltött idő várható értéke ($E(T_w)$):

$$E(N_w) = \lambda E(T_w) \implies E(T_w) = \frac{E(N_w)}{\lambda} = \frac{\lambda^2}{\mu} \frac{1}{\mu - \lambda} \frac{1}{\lambda} = \frac{\lambda}{\mu} \frac{1}{\mu - \lambda}.$$

Tehát $M/M/1$ rendszer esetén:

a rendszerben levő igények számának várható értéke:	$E(N) = \frac{\lambda}{\mu - \lambda}$
a rendszerben eltöltött idő várható értéke:	$E(T) = \frac{1}{\mu - \lambda}$
a várakozó igények számának várható értéke:	$E(N_w) = \frac{\lambda}{\mu} \cdot \frac{\lambda}{\mu - \lambda}$
a várakozással eltöltött idő várható értéke:	$E(T_w) = \frac{\lambda}{\mu} \cdot \frac{1}{\mu - \lambda}$

Példa

Egy hivatalban egyetlen ablaknál intézik az ügyeket. Az ügyfelek Poisson-folyamat szerint érkeznek, óránként átlagosan 10, a kiszolgálási idő exponenciális eloszlású 5 perc várható értékkel. Tekintsük a hivatalban tartózkodó ügyfelek számát.

- Átlagosan hány ügyfél tartózkodik a hivatalban?
- Átlagosan mennyi időt tölt egy ügyfél a hivatalban?
- Átlagosan hányan várakoznak?
- Átlagosan mennyi telik várakozással?
- Véletlenszerű időpontban érkezünk meg. Mi a valószínűsége annak, hogy legfeljebb ketten állnak előttünk?

- (f) Mit mondhatnánk a folyamatról, ha a kiszolgálás átlagos ideje 8 perc lenne?

Megoldás

Az érkezési intenzitás 10 ügyfél/óra, a kiszolgálási pedig 12 ügyfél/óra, tehát most $\lambda = 10$ és $\mu = 12$.

- (a) $E(N) = \frac{\lambda}{\mu - \lambda} = \frac{10}{12 - 10} = 5$ ügyfél van átlagosan a hivatalban.
- (b) $E(T) = \frac{1}{\mu - \lambda} = \frac{1}{12 - 10} = 0,5$ órát tölt átlagosan egy ügyfél a hivatalban.
- (c) $E(N_w) = \frac{\lambda^2}{\mu} \frac{1}{\mu - \lambda} = \frac{10^2}{12} \frac{1}{12 - 10} = 4\frac{1}{6}$ ügyfél várakozik átlagosan.
- (d) $E(T_w) = \frac{\lambda}{\mu} \frac{1}{\mu - \lambda} = \frac{10}{12} \frac{1}{12 - 10} = \frac{5}{12}$ óra telik átlagosan várakozással.
- (e) $P(\text{legfeljebb ketten állnak előttünk}) = P(0, 1 \text{ vagy } 2 \text{ ember van a rendszerben}) = \pi_0 + \pi_1 + \pi_2 =$

$$= 1 - \frac{\lambda}{\mu} + \left(1 - \frac{\lambda}{\mu}\right) \frac{\lambda}{\mu} + \left(1 - \frac{\lambda}{\mu}\right) \left(\frac{\lambda}{\mu}\right)^2 = \frac{1}{6} + \frac{5}{36} + \frac{25}{216} = \frac{91}{216} \approx 0,4213.$$
- (f) Ekkor a kiszolgálási intenzitás 7,5 ügyfél/óra, azaz $\mu = 7,5$ lenne. Mivel ekkor $\lambda > \mu$, ezért a rendszer hosszú távon nem tudná kiszolgálni a beérkező igényeket.

Az $M/M/1$ rendszer esetében kapott formulák természetesen nem alkalmazhatóak pl. $M/M/m$ vagy $M/M/1/k$ rendszerre, de a módszer, ahogyan a formulákat levezettük ezekben az esetekben is használható.

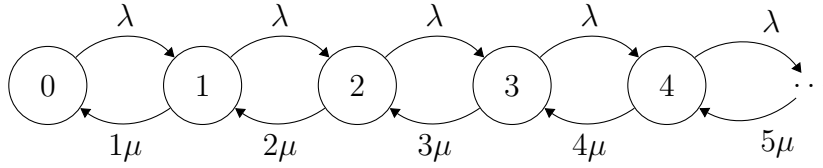
5 Az $M/M/\infty$ rendszer

Tételezzük fel, hogy olyan tömegkiszolgálási rendszerrel van dolgunk, ahol a szerverek száma az igények számához igazodik: ha új igény jelenik meg a rendszerben, akkor egyből hozzárendelődik egy szerver is. Ekkor az igénybe vehető kiszolgálóegységek száma nem korlátozott, hanem végtelen. Legyen az érkezési intenzitás most is λ (igény/időegység), a kiszolgálási intenzitás pedig minden szerver esetében μ . Állapotnak most is a rendszerben levő igények számát tekintjük.

A folyamat infinitezimális generátora:

$$A = \begin{bmatrix} * & \lambda & 0 & \dots \\ \mu & * & \lambda & 0 & \dots \\ 0 & 2\mu & * & \lambda & 0 & \dots \\ 0 & 0 & 3\mu & * & \lambda & 0 & \dots \\ \vdots & & & \ddots & \ddots & \ddots \end{bmatrix}.$$

A folyamat gráfja:



Most meghatározzuk az invariáns eloszlást, ebből kiindulva pedig az egyensúlyi állapotot jellemző mennyiségeket ($E(N)$, $E(T)$). Invariáns eloszlás esetén $\underline{0} = \underline{\pi}A$, ebből az egyenletrendszerből meghatározhatjuk a $\pi_0, \pi_1, \pi_2, \dots$ értékeket. Egy másik lehetséges módszer a gráfon alapul: minden egyes állapotban a kimenő és bejövő intenzitásoknak egyenlőeknek kell lenniük (ez persze a $\underline{0} = \underline{\pi}A$ egyenlet átrendezéséből is adódik). Ekkor balról jobbra haladva a gráfon:

$$\begin{array}{ll} \lambda\pi_0 = \mu\pi_1 & \rightarrow \pi_1 = \frac{\lambda}{\mu}\pi_0 \\ \lambda\pi_1 = 2\mu\pi_2 & \rightarrow \pi_2 = \frac{\lambda}{2\mu}\pi_1 = \frac{\lambda^2}{2\mu^2}\pi_0 \\ \lambda\pi_2 = 3\mu\pi_3 & \rightarrow \pi_3 = \frac{\lambda}{3\mu}\pi_2 = \frac{\lambda^3}{6\mu^3}\pi_0 \\ \vdots & \vdots \\ \lambda\pi_{k-1} = k\mu\pi_k & \rightarrow \pi_k = \frac{\lambda}{k\mu}\pi_{k-1} = \frac{\lambda^k}{k!\mu^k}\pi_0 \\ \vdots & \vdots \end{array}$$

Fejezzük ki π_0 -at, majd pedig azzal az összes többi:

$$1 = \pi_0 + \pi_1 + \pi_2 + \pi_3 + \dots = \pi_0 \left(1 + \frac{\lambda}{\mu} + \frac{\lambda^2}{2\mu^2} + \frac{\lambda^3}{6\mu^3} + \dots \right) = \pi_0 \cdot \sum_{k=0}^{\infty} \frac{\lambda^k}{k!\mu^k}$$

Analízisból ismert, hogy

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

Ezt felhasználva:

$$1 = \pi_0 \cdot \sum_{k=0}^{\infty} \frac{\lambda^k}{k! \mu^k} = \pi_0 \cdot e^{\lambda/\mu} \rightarrow \pi_0 = e^{-\lambda/\mu}$$

Az invariáns eloszlás általános tagja pedig

$$\pi_k = \frac{\lambda^k}{k! \mu^k} e^{-\lambda/\mu} = \left(\frac{\lambda}{\mu}\right)^k \frac{1}{k!} e^{-\lambda/\mu}, \quad k = 0, 1, 2, \dots$$

Ez nem más, mint egy λ/μ paraméterű Poisson-eloszlás.

A rendszerben levő igények számának várható értéke a fenti Poisson-eloszlás várható értéke. Ismert, hogy a Poisson-eloszlás várható értéke megegyezik a paraméterével, tehát

$$E(N) = \frac{\lambda}{\mu}$$

A rendszerben eltöltött idő várható értéke a Little-formula alapján:

$$E(T) = \frac{E(N)}{\lambda} = \frac{1}{\mu}$$

Ezt persze másképp is megkaphattuk volna. Mivel a kiszolgálási intenzitás μ , ezért az átlagos kiszolgálási idő $1/\mu$. A rendszerben nincs várakozás, ezért ez lesz a rendszerben eltöltött teljes idő várható értéke is. Mivel minden egyes újonnan érkező igény azonnal sorra kerül, ezért

$$E(N_w) = 0 \quad \text{és} \quad E(T_w) = 0$$

Figyeljük meg, hogy a rendszer tetszőleges λ és μ esetén működőképes (nem szükséges a $\lambda < \mu$ feltétel). Valós rendszerekben persze nem lehetséges végtelen számú szerver jelenléte. Azért fontos ez a példa, mert a többszerveres valós rendszerek valahol az egyszerveres és a végtelen szerveres eset között helyezkednek el.

Feladatok

1. ($M/M/3$) Egy hivatalban három ablaknál intézik az ügyeket. Az ügyfelek Poisson-folyamat szerint érkeznek, óránként átlagosan 12, a kiszolgálási idő pedig exponenciális eloszlású 4 perc várható értékkel. Tekintsük a hivatalban tartózkodó ügyfelek számát.

- (a) Adjuk meg a folyamat infinitezimális generátorát!
 - (b) Adjuk meg az invariáns eloszlást!
 - (c) Véletlenszerű időpontban érkezünk a hivatalba. Mi a valószínűsége, hogy sorba kell állnunk?
 - (d) Átlagosan hány ügyfél tartózkodik a hivatalban?
 - (e) Átlagosan hányan váraкоznak?
 - (f) Átlagosan mennyi időt tölt egy ügyfél a hivatalban?
 - (g) Ebből mennyi fordítódik váraкоzásra?
 - (h) Mi történne, ha az egyik ablak bezárna? És ha még egy?
2. ($M/M1/4$) Egy lakótelepi kisboltba mindössze négy vásárló fér be. Ha egy potenciális vevő azt tapasztalja, hogy tele van a bolt, akkor be sem megy a boltba. A boltban egyetlen eladó dolgozik, ő exponenciális eloszlású idő alatt szolgálja ki a vevőket, az átlagos kiszolgálási idő hat perc. A potenciális vevők Poisson-folyamat szerint érkeznek, óránként átlagosan nyolc². Tekintsük a boltban tartózkodó vevők számát!
- (a) Adjuk meg a folyamat infinitezimális generátorát!
 - (b) Adjuk meg az invariáns eloszlást!
 - (c) Az idő mekkora hányadában van tele a bolt?
 - (d) A potenciális vevők hány százalékából nem lesz vevő (mert nem tudtak bejutni a boltba)?
 - (e) Véletlenszerű időpontban érkezünk a boltba. Mi a valószínűsége, hogy legfeljebb egy ember van előttünk?
 - (f) Tegyük fel, hogy bejutottunk. Mi a valószínűsége, hogy nem kell váraкоznunk?
 - (g) Átlagosan hány vevő tartózkodik a boltban?
 - (h) Átlagosan hányan váraкоznak?
 - (i) Átlagosan mennyi időt tölt egy vevő a boltban?
 - (j) Ebből mennyi fordítódik váraкоzásra?

²ügyeljünk arra, hogy a potenciális vevők még nem biztos, hogy vevők!

3. ($M/M/2/5$) Egy fodrászatban két mester dolgozik. Az üzlet elég kicsi, mindössze három várakozó ügyfél számára van benne hely. Mindkét mester exponenciális eloszlású idő alatt végez egy klienssel, melynek várható értéke negyedóra. A potenciális ügyfelek Poisson-folyamat szerint érkeznek, óránként átlagosan öt. Tekintsük a fodrászatban tartózkodó kliensek számát!
- (a) Adjuk meg a folyamat infinitezimális generátorát!
 - (b) Adjuk meg az invariáns eloszlást!
 - (c) A potenciális ügyfelek hányadrésze lesz ténylegesen ügyfél (azaz még be tud menni az üzletbe)?
 - (d) Az idő hány százalékában dolgozik mindkét mester?
 - (e) Az idő hány százalékában nincs senki a mestereken kívül a fodrászatban?
 - (f) Tegyük fel, hogy bejutottunk. Mi a valószínűsége, hogy nincs előttünk várakozó ügyfél?
 - (g) Tegyük fel, hogy bejutottunk. Mi a valószínűsége, hogy nem kell várakoznunk?
 - (h) Átlagosan hányan vannak a fodrászatban?
 - (i) Átlagosan hányan várakoznak?
 - (j) Tegyük fel, hogy bejutottunk. Várhatóan hányan vannak előttünk?
 - (k) Tegyük fel, hogy bejutottunk. Várhatóan mennyi idő alatt végzünk?
 - (l) Ebből mennyi idő fordítódik várakozásra?