T-79.179 Rinnakkaiset ja hajautetut digitaaliset järjestelmät

Stokastinen analyysi

Marko Mäkelä

15. maaliskuuta 2004

Mihin tarvitaan stokastista analyysiä?

Saavutettavuusanalyysissä ja mallintarkastuksessa jätetään usein ajan tarkka eteneminen mallintamatta. Huomio kiinnittyy siihen, voivatko asiat tapahtua sellaisessa järjestyksessä, ettei järjestelmä toimi oikein.

Käytännössä tapahtumat kuluttavat myös aikaa, ja esimerkiksi reaktiiviselle ohjausjärjestelmälle voi olla kohtalokasta, jos se ei toimi odotetulla nopeudella.

Arkipäiväisemmältäkin järjestelmältä voidaan edellyttää, että esimerkiksi odottavia töitä tai asiakkaita on keskimäärin järkevä määrä tai ettei keskimääräinen odotusaika veny kohtuuttoman pitkäksi.

Stokastisen analyysin voidaan ajatella perustuvan tilakoneisiin tai saavutettavuusgraafeihin, joiden siirtymiin on liitetty tieto siirtymän suoritusajasta tai valintatodennäköisyydestä.

Stokastiset järjestelmät ja Markovin ketjut

Stokastinen prosessi on joukko tila-avaruuden S satunnaisia tiloja $\{\theta^{(t)} \in S | t \in T\}$ jollekin (ajanhetkiä esittävälle) indeksijoukolle T.

Markovin ketju on sellainen stokastinen prosessi, jossa uusi tila $\theta^{(n+1)}$ riippuu vain edellisestä tilasta $\theta^{(n)}$.

Jatkossa tarkastellaan *stationaarisia* ja *jatkuva-aikaisia* Markovin ketjuja, joiden siirtymätodennäköisyys ei riipu ajasta ja jotka voivat siirtyä tilasta toiseen mielivaltaisina ajanhetkinä. Niitä voidaan kuvata *siirtymänopeusmatriisilla*, joka määrittelee eri tilojen välisten siirtymien todennäköisyydet aikayksikköä kohden.

Markovin ketjua voidaan ajatella polkuna jossakin sellaisessa saavutettavuusgraafissa, jonka siirtymiin liittyy laukeamistodennäköisyys tai satunnaisesti jakautunut kesto.

Stokastiset Petri-verkot (1/2)

Stokastiset Petri-verkot (*Stochastic Petri Nets,* SPN) ovat paikka–siirtymä-järjestelmien laajennus, jonka siirtymät ovat ajastettuja.

Vireessä olevat siirtymät *eivät* voi mielivaltaisesti laueta tai olla laukeamatta. Viive siirtymän virittymisestä laukeamiseen noudattaa negatiivista eksponentiaalijakaumaa eli siirtymä laukeaa aikavälillä $0 \leq X \leq x$ todennäköisyydellä $P\{X \leq x\} = \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}$, missä λ on keskimääräinen laukeamisnopeus.

Koska eksponentiaalijakauma on *muistiton* eli $P\{X > t + h | X > t\} = P\{X > h\}$, järjestelmässä mahdollisesti esiintyvistä kilpatilanteista ei tarvitse välittää.

Tavalliset stokastiset Petri-verkot vastaavat jatkuva-aikaisia Markovin ketjuja.

Stokastiset Petri-verkot (2/2)

Yleistetyissä stokastisissa Petri-verkoissa (*Generalised Stochastic Petri Nets,* GSPN) on myös ajastamattomia eli *välittömästi laukeavia* siirtymiä, joihin liittyy painokerroin.

Jos merkinnässä on vireessä yksi ajastamaton siirtymä, se laukeaa välittömästi tuottaen uuden merkinnän. Jos vireessä on monta ajastamatonta siirtymää, niiden laukeamistodennäköisyydet määräytyvät painokertoimista. Ajastetut siirtymät voivat laueta vain, jos ajastamattomia siirtymiä ei ole vireessä. GSPN:t vastaavat semi-Markov-prosesseja. Niillä on sama stationaarinen jakauma kuin vastaavilla Markov-prosesseilla.

Deterministisissä ja stokastisissa Petri-verkoissa (*Deterministic and Stochastic Petri Nets,* DSPN) on lisäksi vakioajassa laukeavia siirtymiä. Jos vireessä voi olla monta tällaista siirtymää samanaikaisesti, analyyttiset keinot ovat kovilla ja on turvauduttava simulointiin.

Marko Mäkelä

Stokastisten järjestelmien analysointi

Stokastisten järjestelmien mallintamiseen ja analysointiin on työkaluja, kuten GreatSPN, DSPNExpress, TimeNET ja SMART. Valitettavasti emme tunne yhtään vapaata työkalua.

Yksinkertaisten järjestelmien analysointiin riittää saavutettavuusanalysaattori ja matriisilaskin, esimerkiksi Matlab tai GNU Octave. Analyysi tapahtuu kolmessa vaiheessa:

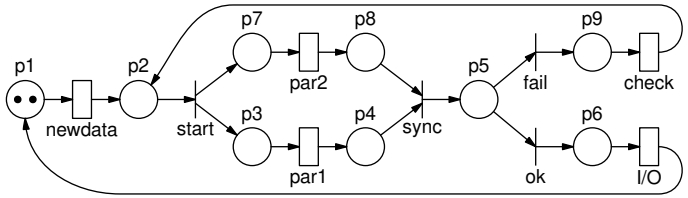
1. Muodostetaan järjestelmän saavutettavuusgraafi. GSPN:n tapauksessa yhdistetään välittömästi laukeavat siirtymät ajastettujen siirtymien kaariin.

2. Merkitään saavutettavuusgraafiin kunkin siirtymän laukeamisfrekvenssi ja muodostetaan siitä siirtymänopeusmatriisi Q.

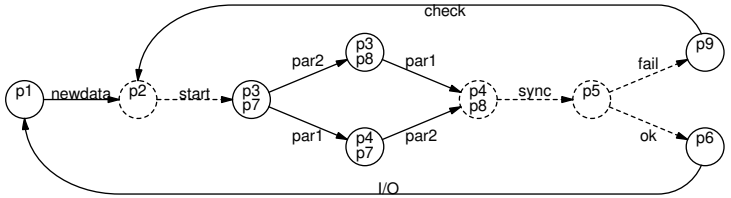
3. Ratkaistaan tasapainoyhtälö ja lasketaan prosessin tunnusluvut.

Marko Mäkelä

Esimerkkijärjestelmä ja saavutettavuusgraafi



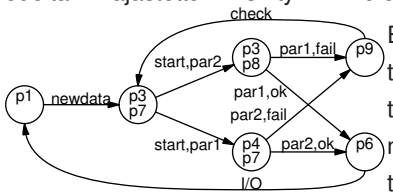
Saavutettavia tiloja on 38 ja kaaria 62. Kun $M_0(\text{p1}) = 1$, tila-avaruus kutistuu:



Marko Mäkelä

Välittömien siirtymien poistaminen

Saavutettavuusgraafia voi pienentää yhdistämällä välittömästi laukeavat siirtymät niitä edeltäviin ajastettuihin siirtymiin. Tuloksena on tavallinen Markovin ketju:



Emme vielä liittäneet siirtymiin painoja emmekä kestoja. Oletetaan siirtymien "fail" ja "ok" laukeamissuhteeksi $5:95$ eli asetetaan niille painot 5 ja 95. Siirtymien "sync" ja "start" painoilla ei tässä ole merkitystä, sillä ne eivät esiinny missään kilpatilanteessa.

Asetetaan nopeudet $\lambda_{\text{newdata}} = 37, \lambda_{\text{par1}} = 40, \lambda_{\text{par2}} = 60, \lambda_{\text{I/O}} = 15, \lambda_{\text{check}} = 2.$

Selvästi $\lambda_{\text{start,par1}} = \lambda_{\text{par1}}$ ja $\lambda_{\text{start,par2}} = \lambda_{\text{par2}}$, koska "start" on yksinään vireessä. Vastaavasti voidaan unohtaa "sync".

Kilpatilanteessa olevat välittömät siirtymät "fail" ja "ok" on otettava huomioon: $\lambda_{\text{par1,ok}} = \frac{95}{95+5}\lambda_{\text{par1}} = 38$, $\lambda_{\text{par1,fail}} = \frac{5}{95+5}\lambda_{\text{par1}} = 2$, $\lambda_{\text{par2,ok}} = 57$ ja $\lambda_{\text{par2,fail}} = 3$.

Marko Mäkelä

Siirtymänopeusmatriisi

Numeroidaan kutistetun saavutettavuusgraafin tilat ja esitetään se siirtymänopeusmatriisina Q, jonka alkiot $q_{i,j}, i \neq j$ ilmoittavat todennäköisyyden aikayksikköä kohden siirtyä tilasta i tilaan j eli siirtymänopeuden:

$$\mathbf{Q} = \begin{array}{c} \begin{array}{cccccc} \{p_1\} & \{p_3,p_7\} & \{p_4,p_7\} & \{p_3,p_8\} & \{p_6\} & \{p_9\} \end{array} \\ \begin{pmatrix} -37 & 37 & 0 & 0 & 0 & 0 \\ 0 & -100 & 40 & 60 & 0 & 0 \\ 0 & 0 & -60 & 0 & 57 & 3 \\ 0 & 0 & 0 & -40 & 38 & 2 \\ 15 & 0 & 0 & 0 & -15 & 0 \\ 0 & 2 & 0 & 0 & 0 & -2 \end{pmatrix} \end{array} \begin{array}{c} \\ \{p_1\} \\ \{p_3,p_7\} \\ \{p_4,p_7\} \\ \{p_3,p_8\} \\ \{p_6\} \\ \{p_9\} \end{array}$$

Jos tilojen välillä ei ole siirtymää, siirtymänopeus on 0. Lävistäjäalkiot asetetaan siten, että rivisummat ovat nollia: $q_{i,i} = -\sum_{j \neq i} q_{i,j}$.

Marko Mäkelä

Tasapainojakauman laskenta (1/2)

Voidaan määritellä todennäköisyysvektori $\pi(t)$, joka ilmoittaa järjestelmän kunkin tilan todennäköisyyden ajan funktiona. Raja-arvoa $\pi = \lim_{t\to\infty} \pi(t)$ sanotaan stationaariseksi ratkaisuksi, sillä se toteuttaa yhtälön $\pi \cdot \mathbf{Q} = \mathbf{0}$. Lisäksi sen on toteutettava todennäköisyyden normiehto $\pi \cdot \mathbf{e}^{\mathrm{T}} = 1$ eli $\pi_1 + \pi_2 + \cdots + \pi_n = 1$. Normiehto voidaan kirjoittaa matriisimuotoon kirjoittamalla sama yhtälö moneen kertaan: $\pi \cdot \mathbf{E} = \mathbf{e}$. (Tässä $\mathbf{e}$ on pelkkiä ykkösiä sisältävä vaakavektori ja $\mathbf{E}$ ykkösneliömatriisi.)

Laskemalla saadut matriisiyhtälöt

$$\begin{aligned} \pi \cdot \mathbf{Q} &= \mathbf{0} \\ \pi \cdot \mathbf{E} &= \mathbf{e} \end{aligned}$$

yhteen saadaan $\pi \cdot (\mathbf{Q} + \mathbf{E}) = \mathbf{e}$, josta voidaan ratkaista $\pi = \mathbf{e} \cdot (\mathbf{Q} + \mathbf{E})^{-1}$.

Marko Mäkelä

Tasapainojakauman laskenta (2/2)

Lasketaan kutistetun saavutettavuusgraafin tilojen todennäköisyydet GNU Octavella:

```
octave:1> Q=[-37,   37,   0,   0,   0,  0;
               0, -100,  40,  60,   0,  0;
               0,    0, -60,   0,  57,  3;
               0,    0,   0, -40,  38,  2;
              15,    0,   0,   0, -15,  0;
               0,    2,   0,   0,   0, -2]

octave:2> pi=ones(1,6)*inverse(Q+ones(6))*
pi =

0.176252 0.068646 0.045764 0.102968 0.434756 0.171614
```

Käsin laskemalla olisi saatu tarkempi tulos, mutta käytännössä likiarvot riittävät hyvin kuvaajien piirtämiseen (tasapainojakauma esimerkiksi λ_{newdata}:n funktiona).

*Lauseke `pi=((Q'+ones(6))\ones(6,1))'` olisi hieman tehokkaampi.

Marko Mäkelä

Tasapainojakauman sovelluksia (1/2)

Äskeisen perusteella tiedämme muun muassa, että järjestelmä viettää suurimman osan ajastaan tilassa $\{p_6\}$ (se on siellä $43,4\%$:n todennäköisyydellä) eli siitä lähtevä siirtymä "I/O" on mahdollinen pullonkaula.

Stokastisissa Petri-verkoissa voidaan tarkastella paikkojen keskimääräistä merkkimäärää. Äskeisen järjestelmän paikan p_1 merkintä on 1 tilassa $\{p_1\}$ ja muualla 0. Niinpä sen keskimääräinen merkintä on suoraan kyseisen tilan todennäköisyys, $0{,}176252$.

Yleisemmin paikan keskimääräinen merkintä on tilojen todennäköisyyksillä painotettu summa paikan merkinnästä kussakin tilassa. Esimerkiksi paikan p_3 merkintä saavutettavissa tiloissa voidaan esittää pystyvektorilla $\mathbf{m}_{p_3} = (0\ 1\ 0\ 1\ 0\ 0)^{\mathrm{T}}$. GNU Octaven mukaan paikan p_3 keskimääräinen merkintä `pi*[0,1,0,1,0,0]'` on $0{,}17161$.

Marko Mäkelä

Tasapainojakauman sovelluksia (2/2)

Vastaavasti voidaan laskea mutkikkaampienkin ehtojen todennäköisyyksiä. Määritellään pystyvektori $\mathbf{r}$, joka saa arvon 1 niissä tiloissa, joissa haluttu ominaisuus pätee, ja muualla arvon 0. Esimerkki: ”paikassa p_7 tai p_8 on merkki” eli $\mathbf{r} = (0\ 1\ 1\ 1\ 0\ 0)^{\mathrm{T}}$ toteutuu mielivaltaisella ajanhetkellä todennäköisyydellä $0,21738$.

Voidaan tarkastella myös siirtymän odotusarvoista laukeamistiheyttä. Määritellään $\mathbf{r}$ siten, että se on siirtymän laukeamisnopeus niissä tiloissa, joissa kyseinen siirtymä on vireessä, ja muualla 0. Keskimääräinen laukeamistiheys on $\pi \cdot \mathbf{r}$.

Esimerkkimme siirtymälle ”par1” saadaan `pi*[0,40,0,40,0,0]'` eli $6{,}8646$. Koska tämän siirtymän ainoa esipaikka on p_3, jossa voi olla enintään yksi merkki, sama tulos olisi tässä tapauksessa saatu myös kertomalla p_3:n merkinnän todennäköisyys siirtymänopeudella eli laskemalla $40 \cdot 0{,}17161$.

Marko Mäkelä

Yhteenveto

Suorituskykyanalyysi perustuu saavutettavuusgraafeihin. Jos järjestelmässä on n saavutettavaa tilaa, on käsiteltävä $n \times n$-neliömatriiseja. Koska se voi olla hyvin raskasta, mallia on usein yksinkertaistettava tilojen vähentämiseksi. Esitimme yhden tavan, välittömästi laukeavien siirtymien poistamisen. Muut tavat vaativat mallinnettavan järjestelmän tarkempaa tuntemusta.

Tämä luento on vain pintaraapaisu stokastisten järjestelmien analyysiin. Emme esitelleet erilaisia jonotus- ja palvelukäytäntöjä (montako asiakasta voidaan palvella kerrallaan, alkaako odotusaika alusta siirtymän vireessäolon poistuttua jne.) emmekä muunlaisia ajastettuja siirtymiä kuin negatiivista eksponentiaalijakaumaa noudattavia.

Tunnuslukuja voidaan saada simuloimalla järjestelmää riittävän kauan vaikkapa MCMC-menetelmällä (Monte Carlo Markov Chain). Simulointi on monesti ainoa vaihtoehto, jos järjestelmässä on monenlaisia jakaumia noudattavia siirtymiä tai sillä on hyvin monia saavutettavia tiloja.

Marko Mäkelä

Lisätietoja

1. Gianfranco Balbo: *An Introduction to Generalized Stochastic Petri Nets*, pages 217–266 in *Petri Nets 2000: Introductory Tutorial* (http://www.daimi.au.dk/PetriNets/introductions/pn2000_introtut.pdf)

2. Mat-2.111 Stokastiset prosessit (3 ov)

3. S-38.143 Jonoteoria (3 ov) L

Marko Mäkelä