

ESTRATÈGIA DE COMANDA D'UN PAS ENDAVANT PER A
ALGORISMES D'ASSIGNACIÓ DE TASQUES EN UN MONO
PROCESSADOR TREBALLANT EN TEMPS REAL CRÍTIC

D. JULVÉ

Aquest article presenta una discussió teòrica sobre una estratègia de comanda subòptima aplicada a un procés de Markov per a una funció de cost cumulativa associada a canvis d'estat en el procés.

Es demostra que partint de certes hipòtesis sobre la regularitat del procés i sobre la comandabilitat de les probabilitats de transició, una estratègia de comanda a $n+1$ passos endavant és millor que una estratègia a n passos endavant, per a la classe de funció de cost que es considera.

S'apliquen els resultats a un procés de cua concret de demandes que arriben a un sistema monoprocessador treballant en temps real crític. Cal que les demandes siguin completades abans de dates límit específiques a cada demanda i aquelles demandes per a les quals això no es compleix degut a congestió del processador, es consideren no satisfetes.

Es mostra aleshores que una estratègia d'un pas endavant redueix més l'esperança matemàtica de la suma de demandes no satisfetes que no pas les estratègies de 0 passos endavant emprades en temps real, les quals ignoren en general aquelles demandes que causen la saturació del temps del processador.

1. INTRODUCCIÓ

S'ha esmerçat molta recerca en diversos problemes que tracten l'assignació¹ de demandes asíncrones que arriben a una certa unitat central (U.C). En general però, la literatura sobre assignació no sembla interessar-se gaire per les situacions de temps real crític¹, en les quals si més no un dels paràmetres explícits associats a cada demanda està constrenyit críticament /2/.

Encara s'ha atorgat menys atenció a aquella assignació en temps real crític en la que -- cal associar una certa forma de pèrdua als -- diversos graus d'eficiència emprables /1/. -- Aquelles demandes per a les quals les restriccions no són satisfetes, es consideren -- no reeixides, i una certa funció del nombre de demandes no satisfetes constitueix una -- pèrdua d'eficiència de la U.C. En una tal -- contextura és ben clar que pot no ser satisfactòria una eficiència mitjana de les demandes de servei.

El fi d'aquest article és de mostrar que si l'eficiència de la U.C. es mesura com a funció del nombre de demandes no reeixides, --

aleshores en situacions de congestió de la unitat central, les estratègies que tenen en compte la dinàmica de la contextura de temps real crític actuen millor que les que la ignoren.

2. DISCUSSIÓ TEÒRICA

La seqüència de demandes aleatòries que arriben a un monoprocessador defineix un procés estocàstic que sota certes condicions d'independència de les variables d'entrada constitueix un procés de Markov.

Hom pretén de demostrar que per a un tal procés de Markov i una certa funció de cost relacionada amb el nombre de pèrdues de demandes, una estratègia d'horitzó $n+1$ és millor que una estratègia d'horitzó n . La demostració a fer es basa en la propietat de la difusió d'un procés de Markov.

2.1 Hipòtesis de partença i propietat de la difusió

Sigui $\{S_t, t \geq 0\}$ un procés de Markov d'estat

- D. Julvé i Valls. Consultor independent. c/ Nàpols, 266. Barcelona - 25.
- Article rebut el Març del 1979.

discret en temps continu i ergòdic.

Formularem les tres hipòtesis següents:

H1) $\{X_t\}$ és completament definit, és a dir, la probabilitat de l'estat inicial i la matriu de probabilitats de transició P_t són conegudes al temps t_0 i per a tot t respectivament.

H2) Per a tot t els elements $p_{\alpha\beta}$ de P_t són tots comandables per a una funció de cost C donada:

$$p_{\alpha\beta}(t) = p_{\alpha\beta}(u_t) \quad (1)$$

on α i β són enters que defineixen estats de X_t , i u_t és una comanda del conjunt U_t de comandes ($u_t \in U_t$) al temps t per a la funció de cost C .

H3) Existeix un cost constant $c_{\alpha\beta}$ associat a cada transició de cada estat α a un altre β . Aleshores, donat $X_t = \alpha$ la transició a l'estat $X_{t+\tau} = \alpha$ té un cost donat per la variable aleatòria $L_t/X_t = \alpha$, l'esperança matemàtica de la qual és

$$E(L_t/X_t = \alpha) = E(L(u_t)/X_t = \alpha) \quad (2)$$

degut a H2).

La propietat de la difusió expressa el fet de que donats tres instants ordenats en el temps t_i, t_j, t_k ($t_i < t_j < t_k$) les esperances, del cost de les transicions a l'estat X_{t_k} donat X_{t_j} , i el de les transicions a X_{t_k} donat X_{t_i} , estan relacionades per

$$E(L(u_{t_k})/X_{t_j} = \beta) \leq E(L(u_{t_k})/X_{t_i} = \alpha) \quad (3)$$

2.2 Anàlisi de la estratègia de comanda d'un pas endavant aplicada al procés de Markov $\{X_t\}$ per a la funció de cost

$$C = E \sum_{i=0}^N (L(u_{t_i})/X_{t_0} = \alpha)$$

Considerem el procés de Markov $\{X_t\}$ introduït a 2.1 i la funció de cost

$$C = E \sum_{i=0}^N (L(u_{t_i})/X_{t_0} = \alpha) \quad (4)$$

Sigui L el conjunt de les estratègies de comandes per a C i suposem que L no sigui buit, de manera que existeixi si més no una seqüència L_{tN} d'accions de comanda u_{t_i} .

Anomenarem u'_i l'acció de control $\in U_{t_i}$ que no considera el futur de $X_m, m > i$ del procés de Markov i anomenem u^*_i aquella acció de control $\in U_{t_i}$ que dona la probabilitat mínima i que satisfà:

$$p_{\alpha\beta}(u^*_{t_i}) \leq p_{\alpha\beta}(u_{t_i}) \quad (5)$$

Anomenant respectivament les estratègies de comanda

$$L^*_N = \{u^*_{t_0}, u^*_{t_1}, \dots, u^*_{t_N}\},$$

i

$$L'_N = \{u'_{t_0}, u'_{t_1}, \dots, u'_{t_N}\},$$

estratègia de comanda a un pas endavant, i a zero passos endavant, volem demostrar el següent:

Teorema

L'estratègia L^*_N és millor que l'estratègia L'_N per a la funció de cost C expressada a (4) quan s'aplica al procés de Markov $\{X_t\}$ definit a 2.1, tal com expressa:

$$E \sum_{i=0}^N (L(u^*_{t_i})/X_{t_0} = \alpha) \leq E \sum_{i=0}^N (L(u'_{t_i})/X_{t_0} = \alpha) \quad (6)$$

Demostració

De l'expressió (5) i del fet que

$$E(L(u_{t_i})/X_{t_i} = \alpha) = \sum_{\alpha\beta} c_{\alpha\beta} p_{\alpha\beta}(u_{t_i}),$$

hom pot deduir² per a tot $t_i \leq t_j$

$$E(L(u^*_{t_j})/X_{t_i} = \alpha) \leq E(L(u_{t_j})/X_{t_i} = \alpha) \quad (7)$$

Per a demostrar (6) utilitzarem l'expressió (3) de la propietat de la difusió i (7), emprant un raonament recurrent recursiu.

Suposem que (6) és veritat a partir de $t_i = t(n+1)$, és a dir

$$E \sum_{i=n+1}^N (L(u^*_{t_i})/X_{t(n+1)} = \alpha) \leq E \sum_{i=n+1}^N (L(u'_{t_i})/X_{t(n+1)} = \alpha) \quad (8)$$

aplicant la propietat de la difusió (3) a -
 cadascun dels membres de (8) podem trobar -
 un majorant de cada un: el del membre més -
 petit de (8)

$$E \sum_{i=n+1}^N (L(u_{ti}^*)/X_{t(n+1)}=\alpha) \leq E \sum_{i=n+1}^N (L(u_{ti}^*)/X_{tn}=\alpha) \quad (9)$$

i el del membre més gran de (8)

$$E \sum_{i=n+1}^N (L(u_{ti}')/X_{t(n+1)}=\alpha) \leq E \sum_{i=n+1}^N (L(u_{ti}')/X_{tn}=\alpha) \quad (10)$$

Volem ara demostrar que la relació de desi--
 gualtat es manté per als majorants de (9) i
 de (10). Per a fer-ho explicitem els majo--
 rants de (9) i (10). El de (9) es desenvolupa en

$$E \left[(L(u_{t(n+1)}^*)/X_{tn}) + (L(u_{t(n+2)}^*)/X_{tn}) + \dots \dots + (L(u_{tN}^*)/X_{tn}) \right] \quad (11)$$

i el de (10) en

$$E \left[(L(u_{t(n+1)}')/X_{tn}) + (L(u_{t(n+2)}')/X_{tn}) + \dots \dots + (L(u_{tN}')/X_{tn}) \right] \quad (12)$$

comparant membre a membre els components de
 les sumes (11) i (12) i tenint en compte (7)
 hom dedueix que l'expressió (11) és més peti--
 ta que la (12), i per consegüent, que els ma--
 jorants esmentats satisfan:

$$E \sum_{i=n+1}^N (L(u_{ti}^*)/X_{tn}=\alpha) \leq E \sum_{i=n+1}^N (L(u_{ti}')/X_{tn}=\alpha) \quad (13)$$

Afegint a cada membre de (13) el terme

$$E(L(u_{tn}^*)/X_{tn}=\alpha)$$

no variaria el sentit de la desigualtat. --
 L'expressió (7) ens dóna un majorant d'aquest
 terme, el qual és

$$E(L(u_{tn}')/X_{tn}=\alpha)$$

Afegint aquest a la dreta de (13) i el pri--
 mer a l'esquerra, encara amb més raó es man--
 té el sentit de la desigualtat, quedant així:

$$E \sum_{i=n}^N (L(u_{ti}^*)/X_{tn}=\alpha) \leq E \sum_{i=n}^N (L(u_{ti}')/X_{tn}=\alpha) \quad (14)$$

QÜESTIÓ - v.3, nº1 (març 1979)

amb el qual queda demostrat que si (8) és ve--
 ritat, també ho és (14). Repetint aquest ma--
 teix raonament als temps $t(n-1)$, $t(n-2)$, ...
 ..., t_1 , t_0 , queda demostrada completament -
 l'expressió (6).

3. L'ESTRATÈGIA DE COMANDA L^* APLICADA A UN ALGORISME D'ASSIGNACIÓ DE TASQUES

Considerem la seqüència de demandes que arri--
 ben a un monoprocessador, cada una de elles
 amb la configuració aleatòria

$$(\Delta, PT, RT)_{tn}$$

on:

n es refereix a la n -èsima arribada

Δ és la variable aleatòria (v.a.) de temps -
 entre arribades, $\Delta = t(n+1) - t_n$

PT és la v.a. del temps de servei a la n -ès--
 sima demanda

RT és la v.a. del temps de resposta crític -
 per a n -èsima demanda relativament a t_n .

En situació de congestió aquestes demandes -
 han de fer cua per a ser completades per la
 U.C. Amb la configuració aleatòria esmentada
 aquesta cua constitueix un procés d'acumula--
 ció (o procés de cua) estocàstic, l'estat --
 del qual està donat per:

$$Q_{tn} = \{N(tn), PT_1(tn), RT_1(tn), \dots, PT_2(tn), RT_2(tn), \dots, PT_j(tn), RT_j(tn), \dots, PT_N(tn), RT_N(tn)\} \quad (15)$$

on $j=1, 2, \dots, N(tn)$ indica la posició de l'e--
 lement corresponent dins la cua, tenint per
 al cap de cua $j=1$, i sent $N(tn)$ el nombre --
 d'elements de la cua.

L'espera de quada al procés de cua podria cau--
 sar l'acabament si més no d'un dels seus ele--
 ments, passada la seva data límit; una tal -
 demanda serà anomenada insatisfeta. A $1/$ --
 hom demostra que és possible de saber a qual
 sevol data d'arribada t_n les demandes insa--
 tisfetes de la cua, i llur nombre i que -
 aquest nombre sempre pot ser reduït a 1.

Anomenarem cua factible aquella en la qual - el nombre suara esmentat és zero, i infactible quan és positiu.

Amb la suposició adicional de que una cua - buida és factible, aleshores tota cua és o - factible o infactible. Això permet de defi-- nir un procés estocàstic Z_{tn} en temps conti-- nu de dos estats: factible-infactible (1-0).

Ens proposem ara d'obtenir una comanda sub-- òptima L per al procés

$$\{Z_{t(n+1)}/Q_{tn}\}$$

per a una funció del nombre cumulatiu de de-- mandes insatisfetes a partir d'un instant -- inicial t_0 , i que millori l'eficiència de la comanda L' que ignora, a cada tn , aquella de manda que causa infactibilitat. En altres pa raules: en els sistemes de temps real conven cionals la comanda utilitzada generalment -- consisteix en no deixar entra -i per conse-- güent refusar- tota demanda que en arribar a la data tn causi saturació del processador, és a dir, que provoqui una transició de la -- cua a un estat infactible. Aquesta estratè-- gia L' , no té en consideració cap informació sobre la dinàmica del procés de cua Q_{tn} , i - per consegüent es de 0 passos endevant /2/.

Per a una funció de cost C expressada com -- l'esperança matemàtica de la suma del nombre de demandes refusades a cada transició de -- Q_{tn} a l'estat infactible,

$$C = E \sum_{i=0}^N L(u_{ti})/Q_{t0} \quad (16)$$

on $L(u_{ti})$ és la v.a. de valor 0 ó 1 que re-- presenta la pèrdua associada a la transició esmentada, volem trobar un altre estratègia que millori l'eficiència de L' per a C .

Per a fer això considerem a cada transició - de la cua des d'un estat factible a un estat infactible, el conjunt de demandes de la cua que poden ser refusades, incloent-hi la de-- manda arribada a tn . El fet de refusar una - d'aquestes demandes constitueix una comanda amb cost idèntic al de L' , és a dir 1.

A l'estratègia L^* s'escull la comanda que mi-- nimitzi la probabilitat

$$\text{prob}(Z_{t(n+1)}=0/Q_{tn}, u_{tn}) \quad (17)$$

de tenir la cua $Q_{t(n+1)}$ infactible. Aquesta probabilitat és del tipus (5).

Per definició, l'estratègia L^* és del tipus d'un pas endevant tal com s'ha definit a -- l'apartat 2.2. Per consegüent, L' i L^* satis fan la desigualtat:

$$E \sum_{i=0}^N (L(u_{ti}^*)/Q_{t0}) \leq E \sum_{i=0}^N (L(u_{ti}')/Q_{t0}) \quad (18)$$

on L^* i L' són respectivament les estratè-- gies definides per les comandes u_{ti}^* i u_{ti}' , - veient-se doncs com L^* millora l'eficiència de L' .

4. SIMULACIÓ

Una simulació en FORTRAN del procés de cua $\{Q_t\}$ dóna uns resultats que confirmen la -- teoria quan es comparen L' i L^* per a la fun ció de cost C .

Aquesta simulació³ genera demandes aleatòries que generen un procés de cua $\{Q_t\}$. Repetint les mateixes seqüències, s'apliquen respecti vament les estratègies de comanda L' i L^* i es comparen els nombres de demandes refusa-- des.

L'experiment es repeteix fent variar el coe-- ficient de tràfic $\rho = E(PT)/E(\Delta)$. La Taula 1 - dóna els valors trobats per a valors succes-- sius de $K = E(RT)/E(PT)$ que és la mesura de la urgència mitjana de les tasques.

Veiem com perquè L^* millori sensiblement l'e-- ficiència de L' , cal arribar a valors del -- tràfic $\rho > 1$, i també arribar a valors grans - de K . Això és normal ja que les variants que presenta l'estratègia L^* en relació a L' , en tant que nombre de comandes u^* diferents de u' , creix amb el nombre d'elements de la cua, així com la probabilitat (17), i el nombre - d'elements de la cua creix amb ρ i K .

5. CONCLUSIONS

Aquest article ha demostrat una propietat ge-- neral d'una classe L d'estratègies de coman-- da d'un procés estocàstic markovià $\{X_t\}$ per

Taula 1.
Resultats de la simulació que compara les estratègies de comanda ' i , per diferents valors del coeficient de tràfic i del coeficient d'urgència K. Els valors donats representen el nombre de demandes refusades sobre un total de 80 arribades.

$\rho=0.490$															
K	1.5	2	3	4	5	6	7	8	9	10	11	12	13	14	15
L^*	7	3	0												
L'	7	3	0												

$\rho=0.986$															
K	1.5	2	3	4	5	6	7	8	9	10	11	12	13	14	15
L^*	10	8	2	1	0										
L'	11	14	2	1	0										

$\rho=1.045$															
K	1.5	2	3	4	5	6	7	8	9	10	11	12	13	14	15
L^*	11	9	2	1	0										
L'	10	10	2	1	0										

$\rho=1.556$															
K	1.5	2	3	4	5	6	7	8	9	10	11	12	13	14	15
L^*	13	10	7	3	1	2	1	1	1	0					
L'	14	19	11	7	1	8	8	1	1	0					

a una funció de cost C donada. Les hipòtesis sobre $\{X_t\}$ suposen bàsicament que les probabilitats de transició són comandables per a tot t, i que el procés és ergòdic. La funció de cost C és l'esperança matemàtica del cost acumulatiu de les transicions del procés $\{X_t\}$.

Amb aquestes suposicions s'ha demostrat primer, que una estratègia de comanda que a tot instant t considera les estimacions dels costos en n+1 instants futurs, és millor per a C que la comanda que només considera les estimacions en n instants futurs per a la funció de cost C.

Aquests resultats s'han aplicat després a un exemple concret constituït per la seqüència de demandes asíncrones que arriben a un mono processador, i que tenen cada una d'elles un paràmetre explícit del temps de resposta crític d'acabament. En situacions de congestió, es crea un procés estocàstic de cua $\{Q_t\}$ markovià del tipus $\{X_t\}$. Prenent com a funció de cost C l'esperança matemàtica de la suma acumulada, de demandes per a les que el temps de resposta crític no és satisfet, s'han comparat dues estratègies de control del procés Q_t . L'estratègia L' refusa les demandes que

forcin la cua Q_{tn} a un estat infactible segons van arribant i sense utilitzar la dinàmica del procés Q_t . (Es tracta d'una comanda de 0-passos endavant). D'altre banda l'estratègia L^* refusa la demanda que a cada arribada que forci la cua Q_{tn} a un estat infactible, minimitzi la probabilitat de tenir $Q_{t(n+1)}$ infactible. (Es una comanda de un pas endavant que com s'ha mostrat als resultats teòrics, és més eficient).

Finalment una simulació en FORTRAN del procés $\{Q_t\}$ ha donat uns resultats que confirmen la teoria presentada quan es comparen L' i L^* per a la funció de cost C.

6. RECONeixEMENTS

L'autor vol agrair en Josep Aguilar i Martin del L.A.A.S. de Toulouse pels seus comentaris que han fet possible la realització d'aquest treball.

7. BIBLIOGRAFIA

/1/ JOHNSON, H.H. "Single server queues with known service times and specified times

of completion of services", I.C.D.C.C. -
Technical Report, 1971, Imperial College
of Science and Technology, London SW7 --
2BZ.

/2/ JULVE, D. "Étude d'un algorithme d'ordon-
nancement en temps réel pour un monopro-
cesseur soumis à des délais critiques".
Thèse de 3ème cycle. Université Paul Sa-
batier, Toulouse, 1976.

8. NOTES

¹El terme anglès "scheduling" s'ha fet co-
rrespondre a assignació. Igualment s'usa --
l'expressió Temps Real Crític corresponent
a l'expressió anglesa Hard Real Time.

²Se n'omet la demostració.

³La disciplina d'ordenació de les demandes -
és la de prioritat de servei a la més curta
data límit (SDFS: Shortest Deadline First -
Served de la terminologia anglesa) i que --
per a una cua estàtica és òptima per al cri-
teri de factibilitat de Q_t . Les propietats
d'aquesta disciplina es poden trobar a /1/
i /2/, i només afecten l'organització està-
tica d'una cua donada.