DETERMINACIÓ DE LA CÀRREGA D'UN CENTRE DE CÀLCUL PER MÈTODES D'AGRUPAMENT

RAMON PUIGJANER, JOSEP LLUIS COLOMER, (Dep. Arquitectura de Computadors)
JAUME BARCELÓ (Dep. d'Estadística).
Facultat d'Informàtica. (Barcelona).

RESUM

Un dels més importants problemes relacionats amb el rendiment i/o planificació d'un sistema és el desenvolupament d'una representació concreta i precisa de la càrrega.

Si ens fixem en el tipus de treballs fets en un periode de temps (un dia, una setmana), podrem veure que representen tant la càrrega d'aplicacions, com la càrrega de posta a punt de nous programes, com la càrrega deguda a funcions de manteniment del sistema.

La seva representació i classificació és, per tant, complexa i treballosa degut a la quantitat de dades a avaluar. L'ús de tècniques estadístiques sembla justificada. En la comunicació es presenta i discuteix la tècnica d'agrupament ("clustering") aplicada al problema de la classificació de la càrrega.

INTRODUCCIÓ

El concepte de <u>càrrega de treball</u> apareix de forma natural dins del llenguatge de Procés de Dades : l'ordinador realitza un <u>treball</u> determinat en executant un <u>programa</u> que utilitza els <u>recursos</u> del sistema.

El conjunt de programes que s'executen durant un interval fixat de temps constitueixen la càrrega de treball del sistema en aquest temps. La importància de conèixer aquesta càrrega permet :

- dissenyar bancs d'assaig (benchmark) acurats.
- planificar futures ampliacions d'acord amb els requeriments dels treballs a fer.
- Conèixer el nivell d'utilització i comportament del sistema.

Quasi tots els sistemes mantenen un arxiu de comptabilitat (accounting) amb nombroses dades significatives de la utilització dels recursos del sistema. Cada subsistema (software) pot mantenir el seu propi : treball batch, sessions de temps-compartit i sistema de bases de dades. Generalment, però, els tamanys d'aquests arxius són excesius per a poguer fer l'anàlisi sense ajuda d'eines estadístiques potents. Les tècniques d'<u>agrupament</u> (clustering) ajuden a fer la necessària reducció de dades per a poguer-les manegar i analitzar.

En aquest treball es presenta l'experiència d'aplicació d'aquesta tècnica a un cas real, posant més èmfasi en la discussió dels resultats que en el mètode estadístic en sí.

MESURES

Un conjunt habitual de mesures que defineixen les demandes de recursos fent treball "batch" és el format per mesures de :

el temps d'unitat central (segons)
la memòria ocupada (Kbytes)
el nombre de línies escrites
el nombre d'operacions d'entrada/sortida sobre cintes.
el nombre d'operacions d'entrada/sortida sobre discs.

Altres seleccions de conjunts de mesures inclouen el nombre de tarjes llegides i el nombre d'arxius en disc (Mamrak 1977).

Independentment del conjunt seleccionat, es pot definir un vector de recursos (Artis 1976) :

$\bar{R}_i$ = <$variable_1$, $variable_2$ ------ $variable_p$ >

representant cada variable un recurs o una activitat del programa sobre un recurs; per exemple :

$\bar{R}_i$ = <temps UCP, memòria, línies, op. E/S cintes, op E/discs > .

Si considerem que per a fer un determinat treball sobre el sistema s'utilitzen els seus recursos, llavors la càrrega d'un treball pot ésser definida en termes del vector de recursos $\bar{R}_i$.

Les components del vector per a definir una càrrega de treballs "batch" és diferent del vector que defineix la càrrega de sessions de temps-compartit o de sistemes de bases de dades. En el primer cas, s'hi afegirien les variables que mesuren l'activitat del terminal. En el segon, la informació sobre els accesos a bases de dades, tamany de les cues, etc.

Les variables triades com a components del vector $\bar{R}_i$ són depenents del volum de dades a tractar pel programa. El cas fora diferent si es tractés d'estudiar el comportament del programa : podria triar el nombre d'operacions E/S per segon de UCP, etc. . Per contra, volem analitzar el total de treball que el sistema realitzi : ens és relevant el volum de dades a tractar en un periode de temps.

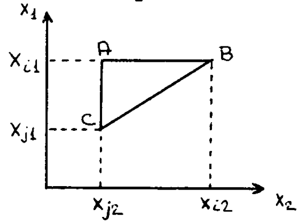
NOTA SOBRE EL MÈTODE D'AGRUPAMENT

El problema que es vol resoldre és el de trobar agrupaments naturals dins de l'espai p-dimensional definit per les p variables del vector de recursos. El fet de que s'utilitzin classes per a fer la planificació (Scheduling) de la càrrega del sistema és un indici de que implícitament s'està fent servir el concepte d'agrupament del treball: així, es classifiquen les feines en grups, depenent del grau de similitud en la utilització dels recursos. A partir de les dades dels arxius de contabilitat, s'intenta formar grups homogènis que representin demandes similars.

La mesura del grau de similitud que correntment es fa servir és la distancia Euclídea definida :

$$d_{ij} = \left(\sum_{K=1}^{p} (x_{ik} - x_{jk}) \right)^{1/2}$$

suposant p variables ortogonals com a base de l'espai on es representen les dades. Si p = 2 es té l'expressió del teorema de Pitàgoras :



$$d_{ij} = \{(x_{i2}-x_{j2})^2 + (x_{i1}-x_{j1})^2\}^{1/2} = \sqrt{\overline{AB}^2 + \overline{AC}^2}$$

En usar la distància Euclídea com a mesura de la similitud s'accepta que :

- les distàncies separant n punts en l'espai de p dimensions son proporcionals a les similituds en les demandes de recursos i hi hà una raó inversa entre elles : a menys distància, més similitud.

- cap observació pot pertanyer simultàniament a dos agrupaments.

La idea base és la d'assignar cada vector a l'agrupament corresponent, de forma que es maximitzi la "qualitat" de la partició de les dades en grups. L'algorisme que fa servir està estudiat a (Duda, 1973).

El criteri més senzill que es pot fer servir és el de tractar de minimitzar globalment la suma dels quadrats de les distàncies dels punts respecte del centre del seu agrupament. La millor partició, és a dir, la que cumpliria millor el criteri esmentat es podria conseguir amb la enumeració exhaustiva de totes les formes de partir un conjunt de n elements en C grups. Això porta a xifres de l'ordre de $C^n/C!$ que obliguen a buscar altres mètodes. Inicialment, es busquen a l'atzar els centres dels agrupaments; els vectors de recursos s'assignen a l'agrupament més proper (en el sentit de la distància Euclídea). El centre de l'agrupament té les coordenades :

$$Y_{rj} = \frac{1}{n_k} \sum_{i=1}^{n_k} \cdot X_{ij} \qquad (j=1,2---P;\ k = 1,2\ ---Cmax.)$$

amb n_k el nombre de punts (vectors) assignats al grup R i Cmax. el nombre inicial de grups.

A continuació es mouen els punts d'un grup a l'altre a fí de satisfer el criteri de minimització esmentat :

$$S_c = \frac{1}{n} \sum_{k=1}^{c} \sum_{i=1}^{n_k} \sum_{j=1}^{P} (X_{ij} - Y_{kj})^2$$ ha d'ésser mínim. En aquests moviments també canvien els centres dels agrupaments.

Un cop acabada aquesta primera iteració es redueix el nombre de grups en 1 (a menys de que c=Cmin). La parella de grups a barreijar es troba buscant la combinació de dos grups que minimitzin l'increment de les sumes dels quadrats de distàncies dels punts als centres. Un cop trobats, es barreigen, calculant-se les coordenades del nou centre. Aquest procediment es repeteix fins a arribar a un nombre de grups (Cmin) determinat. La determinació del nombre d'agrupaments "adequat" es planteja formalment utilitzant les proves de chi-quadrat i de Kolmogorow-Smirnov (Duda 1973); es tracta de trobar el nombre C_n d'agrupaments que és el mateix que els agrupaments naturals de les dades. També intuitivament es pot arribar a determinar el C adequat : S_c ha d'augmentar monòtonament a mida que es redueix C, (Si C=n S_c=0) augmentant lentament en tant s'acosti al nombre C_n d'agrupaments naturals i creixent ràpidament a mida que ens separem de C_n en reduir C. Aquest comportament es veura en l'exemple que segueix.

Altres algorismes es poden trobar a (Agrawala 1976) i Ball (1965). Un punt essencial és l'escalat de les dades. Considerem el cas en que l'escala d'una variable sigui molt més gran que la de l'altra, llavors, un petit canvi en el valor de la primera dominarà completament el valor de la distància d_{ij}.

Si (per p = 2) $X_{i1} >> X_{i2}$ i $X_{j1} >> X_{j2}$

$$d_{ij} = \sqrt{(X_{i1} - X_{j1})^2 + (X_{i2} - X_{j2})^2} \simeq X_{i1} - X_{j1}$$

una forma convenient d'escalat és prendre :

$$Z_{ik} = (X_{ik} - \bar{X}_k)/S_k$$

éssent X_{ik} el punt original.
$\bar{X}_k$ la mitja dels punts de la variable k.
S_k la desviació standard de la variable k.

EXEMPLE

S'estudia la càrrega de treball d'una instal.lació amb una barreja de feines de tipus científic. Tant s'hi processen programes de pràctiques d'estudiants com programes relacionats amb treballs d'investigació. El nombre mig de programes passats en 13 hores es de 300 en èpoques de treball normal. El sistema de comptabilitat del sistema operatiu dona les mesures de la memòria utilitzada (en Kbytes), el temps de UCP (segons) i el nombre de línies escrites. No dona, en canvi, el nombre d'operacions de E/S sobre els dispositius cintes, discs i tambor.

R_i s'agafa :

$\overline{R_i}$ = (memòria (KB), temps UCP (segons), línies escrites)

Sobre una mostra de 2800 treballs, els valors mitjos i les desviacions Standard valen :

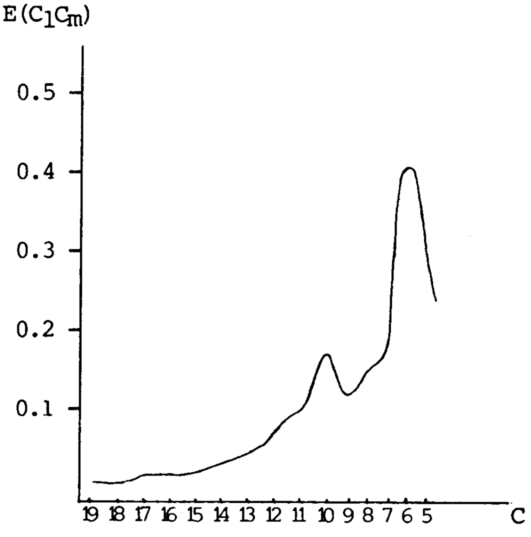
Memòria : 25.98KB i 14.78 KB.
temps UCP : 77.46 segons i 230.84 segons.
línies : 672.50 i 1825.99

El nombre inicial d'agrupaments triats va ésser de 20 i el mínim de 5. Calculant el valor :

$$E\ (C_1\ C_m) = \frac{S_{c1} - S_{cm}}{S_{cm}}$$

és a dir, la variació del valor de S_C en passar de C_m grups a C_1 ($C_1<C_m$) i representant els valors de E (C_1C_m) respecte C s'obtingueren :

C_1	C_m	$E(C_1C_m)$
19	20	0.0173318
18	19	0.0181159
17	18	0.0342014
16	17	0.0377541
15	16	0.0379555
14	15	0.0545678
13	14	0.0642306
12	13	0.0755559
11	12	0.0987478
10	11	0.1780739
9	10	0.1127721
8	9	0.1406126
7	8	0.1624211
6	7	0.4195526
5	6	0.2511872



La variació de S_c en l'interval C =17 a C = 15 és suficientment estable per poguer considerar-lo com el nombre "adequat" de grups. Amb C = 15 es fa el compromís entre la volguda precissió per un costat i la també desitjada reducció de les dades.

Grup	%total	radi del grup	memòria	tUCP	línies
1	15.10	0.057	5.23	0.57	9.57
2	2.17	0.58	27.49	956.34	473.39
3	49.96	0.10	22.20	13.30	150.62
4	0.85	1.60	38.37	347.05	9451.79
5	3.21	0.81	71.21	148.84	1286.82
6	1.92	0.65	24.59	154.85	4795.66
7	0.46	2.18	32.61	439.68	19649.51
8	7.00	0.45	51.37	54.81	490.70
9	2.60	0.72	40.98	157.09	2965.02
10	2.28	0.74	30.32	439.64	612.50
11	0.03	0.0	68.00	7248.13	236.00
12	0.89	1.42	59.55	838.96	2314.51
13	7.75	0.34	33.80	57.37	790.41
14	5.67	0.33	23.38	63.39	1265.07
15	0.03	0.0	32.38	2279.56	376.00

Notis els valors dels grups per C = 5:

Grup	%total	radi del grup	memòria	tCPU	línies
1	77.64	0.54	20.07	21.21	309.42
2	3.67	1.238	32.81	879.14	834.32
3	17.34	1.163	50.18	121.85	1346.07
4	1.25	3.00	35.25	393.95	13410.90
5	0.03	0.0	68.00	7248.13	236.00

Encara que en classificació de la càrrega de treballs no ens preocupem del comportament dels programes individualment, fou fàcil identificar qualitativament cert tipus de programes que generen aquesta càrrega :

Així, s'identificaren els grups,

Grup 1 : programes d'utilitat del sistema : inicialitzar cintes, descarregar-les, demanar-les, etc. .

Grup 3, 14 : compiladors Fortran i execucions del muntador i execució de programes d'usuaris aprenent a programar.

Grup 11 i 15 : programes especials lligats a projectes d'investigació ben coneguts (amb permís especial per tenir un temps d'execució il.limitat).

representant el 70.79% dels programes passats.

La resta es divideix en programes tant d'alumnes avançats de programació com de projectes.

El sistema operatiu dona possibilitats de treballar amb multiprogramació amb particions de memòria estàtiques un cop definides; és per tant la memòria el recurs al que prioritariament s'ha d'atendre a l'hora de planificar quin programa passa per quina partició.

Sobre 98KB de memòria lliure per un màxim de 3 particions, es tracta de buscar el repartiment òptim de memòria per que surti el màxim de treball possible. Sense utilitzar cap eina matemàtica, es fa el següent raonament

Donat que el programa de 24KB correspon al típic procés:compilar,muntar i executar, i donat que el temps del grup 3, i 14 al que pertany es baix i també donat que tant el compilador com el muntador son programes d'entrada/sortida (I/O bound), es proposa deixar sempre com a mínim una partició de 24 KB amb prioritat la més alta. Així, una configuració típica és :

Area 1	50 KB		Area 1	74 KB
Area 2	24 KB	ó bé,	Area 2	0 KB
Area 3	24 KB		Area 3	24 KB

la primera configuració cobreix el 88.19 % dels treballs a passar.

Els valors dels grups per C = 5 portarien essencialment a les mateixes conclusions encara que amb els valors dels grups molt distorsionats.

NOTES FINALS

La forma de classificar una gran quàntitat de dades es fa en dos passos. En el primer es tracta una mostra (per exemple els nostres 2800 punts) generant-se els centres dels agrupaments. En el segon, s'asigna la resta dels vectors de recursos al grup més proper en el sentit de la distància al centre del grup. L'experiència demostra (Artis, 1976) que la classificació en grups sobre períodes molt llargs de temps (un any, per exemple) no fa canviar apreciablement els valors de les coordenades que defineixen els grups.

CONCLUSIONS

La tècnica de reducció i classificació de la càrrega que s'ha presentat obre la possibilitat de

1e. Analitzar la barreja de treballs que s'executen en cada moment sobre el sistema.

2n. Dissenyar un programa que, executat sobre un sistema a provar, el carregui amb les típiques barreiges de feines del sistema amb el que es compara. La mesura dels temps de servei del nou sistema respecte a l'original es l'objectiu d'aquest mètode.

Agraîments

Al Director del Centre de Càlcul de la Universitat Politècnica de Barcelona, Dr. Martí Vergés i Trías, on pertanyen les dades de l'exemple i al meu company Josep M. Sellés, Cap d'explotació del mateix Centre de Càlcul

REFERÈNCIES

Mamrak, S.A.; Amer, P.D.; (1977) A feature selection tool for workload characterization. Proceedins of the 1977 ACM Simetrics/ACM VIII, 1, 113.

Artis, Pat (1976) A technique for determining the capacity of a Computer Performance Evaluation User's Group Meeting, San Diego 1976.

Duda, R.O.; Hart, P.E.; (1973 Pattern classification and Scene Analysis. John Wiley & Sons 1973.

Agrawala, A.K.; Mohr, J.N.; Bryant, R.M. (1976). An approach to the workload characterization problem. Computer, June 1976 p. 182.

Ball, G.; Hall, D., (1975) ISODATA, a novel method of data analysis and Pattern classification. Standford Research Institute 1965.