Redes de Filas
Algoritmo de Convolução

1 Distribuição do Número de Usuários no Sistema

Quando se necessita obter mais que os valores médios das grandezas do sistema, o algoritmo AVM não é capaz de fazê-lo como é o caso da distribuição do número de usuários no sistema.

Inicialmente, vamos considerar que os dispositivos do sistema sejam todos de capacidade fixa (CF). O estado deste tipo de rede é dado pelo seguinte vetor:

$$n = \{n_1, n_2, \ldots ,n_M\}$$

onde n_i representa o número de usuários na estação i.

Gordon e Newell mostraram que a probabilidade do sistema estar no estado n (com n usuários) é dada por:

$$P(n_1,n_2,\ldots,n_M) = \frac{(D_1)^{n_1}.(D_2)^{n_2}\ldots(D_M)^{n_M}}{G(N)}$$

D_i é a demanda total de serviço por usuário na estação i;

$N = \sum_{i=1}^{M} n_i$ é o número total de usuários na rede;

$G(N)$ é uma constante de normalização que garante que as probabilidades somem 1 e que pode ser calculada pela seguinte expressão:

$$G(N) = \sum_{n} [(D_1)^{n_1}.(D_2)^{n_2}\ldots(D_M)^{n_M}]$$

Um possível problema com esta fórmula é que o valor de $G(N)$ pode se tornar muito grande ou muito pequeno, provocando ou “overflow” ou “underflow”. Este problema pode ser evitado escalando-se os valores de D_i por um fator α, ou seja:

$$y_i = \alpha * D_i \quad \text{onde} \quad \alpha = \frac{1}{(\frac{1}{M}).\sum_{i=1}^{M} D_i}$$

As probabilidades de estado podem ser calculadas como:

$$P(n_1,n_2,\ldots,n_M) = \frac{(y_1)^{n_1}.(y_2)^{n_2}\ldots(y_M)^{n_M}}{G(N)}$$

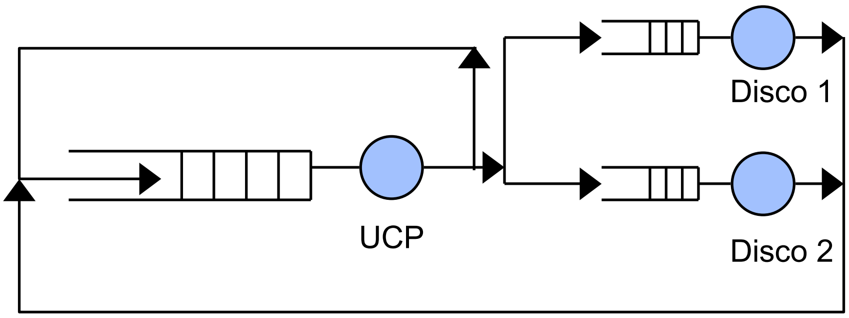
onde $G(N) = \sum_{n} [(y_1)^{n_1} . (y_2)^{n_2} ... (y_M)^{n_M}]$

Exemplo 1 : Sistema "Batch"

Considere um computador "Batch" com um processador e dois discos. Ele pode ser modelado pelo modelo do servidor central.

- O tempo de serviço por visita à UCP é de 39 ms;
- O tempo de serviço por visita ao disco A é de 180ms;
- O tempo de serviço por visita ao disco B é de 260ms;
- Cada "job" faz 13 requisições ao disco A e 6 ao disco B;

Deseja-se calcular as probabilidades de estado quando o grau de multiprogramação é 3.



Resolução:

Tempos de serviço:	S_{UCP} = 0,039 seg.	S_A = 0,180 seg.	S_B = 0,260 seg.
Taxas de Visitas:	V_{UCP} = 13 + 6 + 1 = 20	V_A = 13	V_B = 6
Demanda total de serviço:	D_{UCP} = 20x0,039 = 0,78	D_A = 13x0,180 = 2,34	D_B = 6x0,260 = 1,56

Como fator de escala escolhe-se α=(1/0,78) o que resulta em:

$y_{UCP} = 1$ $\quad$ $y_A = 3$ $\quad$ $y_B = 2$

Tabela de Probabilidades de Estado

Número de usuários			Numerador	Probabilidade
UCP	Disco A	Disco B	Πy_i^n	
0	0	3	8	0.089
0	1	2	12	0.133
0	2	1	18	0.200
0	3	0	27	0.300
1	0	2	4	0.044
1	1	1	6	0.067
1	2	0	9	0.100
2	0	1	2	0.022
2	1	0	3	0.033
3	0	0	1	0.011
			G(N) = 90	

O sistema possui 10 estados diferentes. A probabilidade destes estados pode ser calculada pela expressão de Gordon e Newell, calculando-se a constante de normalização G(N) somando-se os produtos y_i^{ni} em todos os possíveis estados.

2 Algoritmo de Convolução

O problema com o método de solução proposto por Gordon e Newell é que se deve enumerar todos os estados do sistema. Esta tarefa pode se tornar muito trabalhosa pois o número de estados do sistema é dado por $\binom{N+M-1}{M-1}$

Buzen resolveu este problema descobrindo um método para calcular G(N) chamado "Método da Convolução", e é baseado nas definições a seguir.

A constante de normalização G(N) é dada por

G(N) = g(N,M) sendo a função auxiliar $g(n,k) = \sum_n \prod_{i=1}^{k} (y_i)^{n_i}$

G(N) pode ser calculada pela seguinte expressão iterativa:

g(n,k) = g(n,k-1) + y_K*g(n-1,k)

Os valores iniciais são:

g(n,0) = 0, para n = 1,2,...,N e

g(0,k) = 1, para k = 1,2,...,M

Uma forma de organizar o cálculo é através da seguinte tabela com os valores de g(n,k):

n	k=0	k=1 y_1	k=2 y_2	…	k=M y_M	
0		1	1		1	= G(0)
1	0	$0 + y_1*g(0,1)$	$g(1,1) + y_2*g(0,2)$		$g(1,N-1) + y_M*g(0,M)$	= G(1)
2	0	$0 + y_1*g(1,1)$	$g(2,1) + y_2*g(1,2)$		$g(2,N-1) + y_M*g(1,M)$	= G(2)
…	…	…	…	…	…	…
N	0	$0 + y_1*g(N-1,1)$	$g(N,1) + y_2*g(N-1,2)$		$g(N,N-1) + y_M*g(N-1,M)$	= G(N)

Exemplo 2 : Sistema "Batch" pelo método da Convolução

Continuando o exemplo anterior, o cálculo da constante de normalização pelo método da Convolução é feito da seguinte forma :

n		y_{UCP}=1	y_A=3	y_B=2	
0	0	1	1	1	= G(0)
1	0	1	4	6	= G(1)
2	0	1	13	25	= G(2)
3	0	1	40	90	= G(3)

Portanto G(N) = 90 para N=3.

3 Parâmetros de Desempenho usando G(N)

A distribuição do número de usuários no sistema é definida por:

$$P(n_i \geq j) = y_i^j \, G(N-j) / G(N)$$

A probabilidade de se ter j usuários no i-ésimo dispositivo é:

$$P(n_i = j) = P(n_i \geq j) - P(n_i \geq j+1) =$$
$$P(n_i = j) = y_i^j * [\, G(N-j) - y_i \, G(N-j-1)\,] / G(N)$$

O número médio de usuários na i-ésima estação é dado por:

$$Q_i = E[n_i] = \sum_{j=1}^{N} P(n_i \geq j) = \sum_{j=1}^{N} y_i^j G(N-j) / G(N)$$

A probabilidade conjunta de se ter j ou mais usuários na estação i e l ou mais usuários na estação k é dada por:

$$P(n_i \geq j,\, n_k \geq l) = y_i^j \; y_k^l \; G(N-j-l) / G(N)$$

O fator de utilização representa a probabilidade de se ter um ou mais usuários na estação i:

$U_i = y_i \ G(N\text{-}1) / G(N)$

A vazão dos dispositivos é dada pela lei da utilização:

$X_i = U_i/S_i$

A vazão do sistema é dada pela lei do fluxo, ou seja:

$X = X_i/V_i = U_i/D_i = \alpha \ G(N\text{-}1)/G(N)$

O tempo de resposta é dado pela lei de Little:

$R_i = Q_i/X_i = Q_i/XV_i$

Tempo de resposta do sistema: $R = \sum_{i=1}^{M} R_i V_i$

Exemplo 3: Parâmetros de desempenho do sistema "Batch"

A probabilidade de se ter 2 ou mais usuários no disco A é dada por:

$P(n_A \geq 2) = y_A^2 \ G(N\text{-}2) / G(N) = 3^2 * (6/90) = 0{,}6$

Este mesmo valor pode ser alcançado se somarmos as probabilidades dos estados em que n seja 2 ou 3. Existem 3 estados nesta situação e a soma de suas probabilidades é :

0,2 + 0,3 + 0,1 = 0,6

Probabilidade de se ter exatamente um usuário no disco A é:

$P(n_A = 1) = y_A [G(N\text{-}1) - y_A * G(N\text{-}2)] / G(N)$

$P(n_A = 1) = 3*[25 - 3x6] /90 = 21/90 = 0{,}233$

Novamente, este valor pode ser calculado pela soma das probabilidades dos estados com n = 1, ou seja,

0,133 + 0,067 + 0,033 = 0,233

A probabilidade de se ter 0, 2 e 3 usuários no disco pode ser obtida de forma similar:

$P(n_A = 0) = y_A^0[G(N\text{-}0) - y_A G(N\text{-}1)] /G(N) = (90\text{-}3x25)/90 = 0{,}166$

$P(n_A = 2) = 32x(1\text{-}3x0)/90 = 0{,}3$

$P(n_A = 3) = 33x(1\text{-}3x0)/90 = 0{,}3$

Média de usuários no disco A:

$$E[n_A] = \sum_{j=1}^{N} jP(n_A = j) = 1x0{,}233 + 2x0{,}3 + 3x0{,}3 = 1{,}733$$

Variância do número de usuários no disco A:

$$E[n_A^2] = \sum_{j=1}^{N} j^2P(n_A = j) = 1^2x0{,}233 + 2^2x0{,}3 + 3^2x0{,}3 = 4{,}133$$

$Var[n_A] = E[n_A^2] - (E[n_A])^2 = 4{,}133 - (1{,}733)^2 = 1{,}13$

Probabilidade conjunta do disco A e o disco B ter 1 ou mais usuários:

$P(n_A \geq 1, n_B \geq 1) = y_Ay_B(G(N\text{-}2)/G(N)) = (3x2x6)/90 = 36/90 = 0{,}4$

Número médio de usuários para cada estação:

$$Q_{UCP} = \sum_{j=1}^{N} y_{UCP}^{j} G(N-j)/G(N) = (1^1x25 + 1^2x6 + 1^3x1)/90 = 32/90 = 0{,}356$$

$$Q_A = \sum_{j=1}^{N} y_A^{j} G(N-j)/G(N) = (3^1x25 + 3^2x6 + 3^3x1)/90 = 156/90 = 1{,}733$$

$$Q_B = \sum_{j=1}^{N} y_B^{j} G(N-j)/G(N) = (2^1x25 + 2^2x6 + 2^3x1)/90 = 82/90 = 0{,}911$$

A vazão do Sistema é dada por:

$X = \alpha\ G(N\text{-}1) / G(N) = (1/0{,}78)*(25/90) = 0{,}356$

A vazão da UCP é:

$X_{UCP} = XV_{UCP} = 0{,}356x20 = 7{,}12$ jobs/segundo;

A utilização da UCP é dada por:

$U_{UCP} = P(n_{UCP} \geq 1) = y_{UCP}\ G(N\text{-}1)/G(N) = 25/90 = 0{,}278$

Tempo de resposta dos dispositivos é dado por:

$R_{UCP} = Q_{UCP}/(XV_{UCP}) = 0{,}356/(0{,}356x20) = 0{,}05$ segundo

$R_A = Q_A/(XV_A) = 1{,}733/(0{,}356x13) = 0{,}37$ segundo

$R_B = Q_B/(XV_B) = 0,911/(0,356x6) = 0,43$ segundo

Tempo de Resposta do Sistema:

$$R = \sum_{i=1}^{3} R_i V_i = \mathbf{0,05x20+0,37x13+0,43x6 = 8,42\ segundos}$$

4 Método da Convolução: Servidores sem Espera

Assumindo que os servidores sem espera (Infinitos Servidores) estejam na estação 0 e modificando a condição de inicialização para:

$$g(n,0) = y_0^n / n!$$

onde $y_0 = \alpha Z$ é o tempo que um usuário que esteja em um terminal fica pensando.

A tabela com os valores de $g(n,k)$ ficará:

	k=0	k=1	k=2		k=M	
n	Y_0	y_1	y_2	…	y_M	
0		1	1		1	= G(0)
1	$y_0^1 / 1!$	$0 + y_1 * g(0,1)$	$g(1,1) + y_2 * g(0,2)$		$g(1,N-1) + y_M * g(0,M)$	= G(1)
2	$y_0^2 / 2!$	$0 + y_1 * g(1,1)$	$g(2,1) + y_2 * g(1,2)$		$g(2,N-1) + y_M * g(1,M)$	= G(2)
…	…	…	…	…	…	…
N	$y_0^N / N!$	$0 + y_1 * g(N-1,1)$	$g(N,1) + y_2 * g(N-1,2)$		$g(N,N-1) + y_M * g(N-1,M)$	= G(N)

Agora, o vetor de estado possui $(M+1)$ componentes:

$$n = (n_0, n_1, n_2, \ldots, n_M)$$

onde n_i é o número de usuários na i-ésima estação e n_0 é o número de usuários nos terminais (servidor sem espera)

Neste caso as probabilidades de estado são dadas por:

$$P(n_0, n_1, n_2, \ldots, n_M) = (y_0^{n0}\ y_1^{n1}\ y_2^{n2} \ldots y_M^{nM})/(n_0!\ G(N))$$

Se houver mais de um servidor sem espera no sistema, basta adicionar todos os seus atrasos ao valor Z e tratar todos como se fossem uma única estação.

O valor de $G(N)$ calculado desta forma é válido para o sistema e n_0 representa o número médio de usuários em todos os servidores sem espera do sistema.

Para achar o valor de usuários em cada um dos servidores sem espera basta lembrar que estes valores serão proporcionais à demanda de serviço D_i de cada uma das estações com servidores sem espera.

As equações de desempenho do sistema são válidas. A distribuição do número de usuários no sistema só não é valida para os servidores sem espera. Para se calcular o número médio de usuários e a utilização dos servidores sem espera procede-se da seguinte forma:

Número médio de usuários no servidor sem espera:

$$Q_0 = XZ$$

Fator de Utilização do servidor sem espera:

$$U_0 = XZ$$

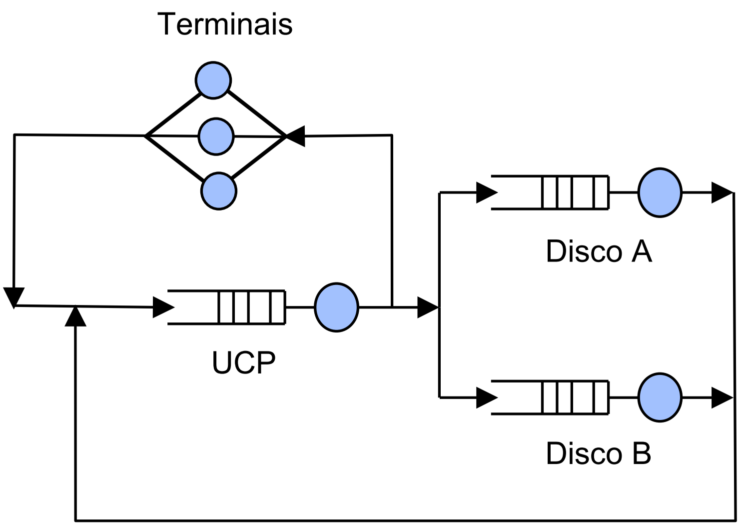
Esta é o Fator de Utilização combinado de N terminais.

A utilização de cada terminal individualmente é:

$$UZ = U_0/N = XZ/N = Z/(R+Z)$$

Exemplo 4: Sistema "Timesharing"

Considere um sistema "Timesharing" similar ao sistema "Batch", ou seja, um sistema com um processador, dois discos e 3 terminais. Um usuário médio faz 13 acessos ao disco A e 6 acessos ao disco B. Os tempos de serviço da UCP, Disco A e Disco B são respectivamente 39, 180 e 260 ms. O tempo médio que um usuário pensa enquanto está no terminal é de 4,68 segundos.



Resolução:

Tempos de serviço:	S_{UCP} = 0,039 seg.	S_A = 0,180 seg.	S_B = 0,260 seg.
Taxas de Visitas:	V_{UCP} = 13 + 6 + 1 = 20	V_A = 13	V_B = 6
Demanda total de serviço:	D_{UCP} = 20x0,039 = 0,78	D_A = 13x0,180 = 2,34	D_B = 6x0,260 = 1,56

Escolhe-se para o fator de escala α = 1/0,78 o que resulta em:

$y_0 = y_T = 6$ $\quad y_{UCP} = 1$ $\quad y_A = 3$ $\quad$ e $\quad y_B = 2$

O número total de estados é 20 e as suas probabilidades são calculadas a seguir:

Tabela para cálculo da constante de normalização G(N)

n	k=0 Y_T=6	k=1 Y_{UCP}=1	k=2 Y_A=3	k=3 Y_B=2	
0		1	1	1	= G(0)
1	6	7	10	12	= G(1)
2	18	25	55	79	= G(2)
3	36	61	226	384	= G(3)

Podem-se calcular outros índices de desempenho através de G(N), que no caso é G(3)=384.
Probabilidade de se ter 1 ou mais usuários no disco A:

$P(n_A \geq 1) = y_A \, G(N\text{-}1)/G(N) = 31x79/384 = 0{,}617$
$P(n_A \geq 2) = y_A^2 \, G(N\text{-}2)/G(N) = 32x12/384 = 0{,}281$

A probabilidade de o sistema estar no estado com 1 usuário no disco A, é dada por:

$P(n_A = 1) = P(n_A \geq 1) - P(n_A \geq 2) = 0{,}617 - 0{,}281 = 0{,}336$

Tabela de Probabilidades de Estado

	Número de usuários			Probabilidade
Terminais	UCP	Disco A	Disco B	$\dfrac{y_0^{n_0}\, y_1^{n_1}\, y_2^{n_2} \ldots y_M^{n_M}}{n_0!\, G(N)}$
0	0	0	3	0.021
0	0	1	2	0.031
...	...	...	...	...
1	0	0	2	0.063
1	0	1	1	0.094
...	...	...	...	...
2	0	0	1	0.094
2	0	1	0	0.141
2	1	0	0	0.047
3	0	0	0	0.094
			Soma=	1,000

A vazão do sistema é:

$X = \alpha\, G(N\text{-}1)/G(N) = (1/0{,}78) * (79/384) = 0{,}264$

O fator de utilização dos dispositivos é:

$U_{UCP} = XD_{UCP} = 0{,}264x0{,}78 = 0{,}206$

$U_A = X\, D_A = 0{,}264x2{,}34 = 0{,}618$

$U_B = X\, D_B = 0{,}264x1{,}56 = 0{,}412$

Número médio de usuários nas estações:

$$Q_{UCP} = \sum_{j=1}^{N} y_{UCP}^{j} G(N-j)/G(N) = (1^1 x79 + 1^2 x12 + 1^3 x1)/384 = 92/384 = 0{,}240$$

$$Q_A = \sum_{j=1}^{N} y_A^j G(N-j)/G(N) = (3^1 x79 + 3^2 x12 + 3^3 x1)/384 = 372/384 = 0{,}969$$

$$Q_B = \sum_{j=1}^{N} y_B^j G(N-j)/G(N) = (2^1 x79 + 2^2 x12 + 2^3 x1)/384 = 214/384 = 0{,}557$$

Q_{Term} = N - (Q_{UCP}+Q_A+Q_B) = 3 - (0,240 + 0,969 + 0,557) = 1,234

Tempo de resposta dos dispositivos:

R_{UCP} = Q_{UCP}/(XV_{UCP}) = 0,240/(0,264x20) = 0,045 seg.

R_A + Q_A/(XV_A) = 0,969/(0,264x13) = 0,283 seg.

R_B = Q_B/(XV_B) = 0,557/(0,264x6) = 0,352 seg.

Tempo de resposta do sistema:

$$R = \sum_{i=1}^{3} R_i V_i = 0{,}045x20+0{,}283x13+0{,}352x6 = 6{,}694 \text{ segundos}$$

Número de usuários no sistema:

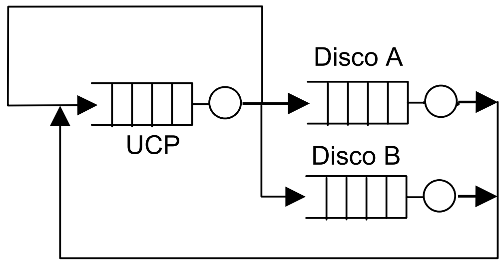
N = X(R+Z) = 0,264(6,694 +4,68) = 3 como deveria de ser.

5 Bibliografia

[1] Jain, R., "The Art of Computer Systems Performance Analysis", John Wiley & Sons Inc, ISBN: 0-471-50336-3, 1991, 685 p.

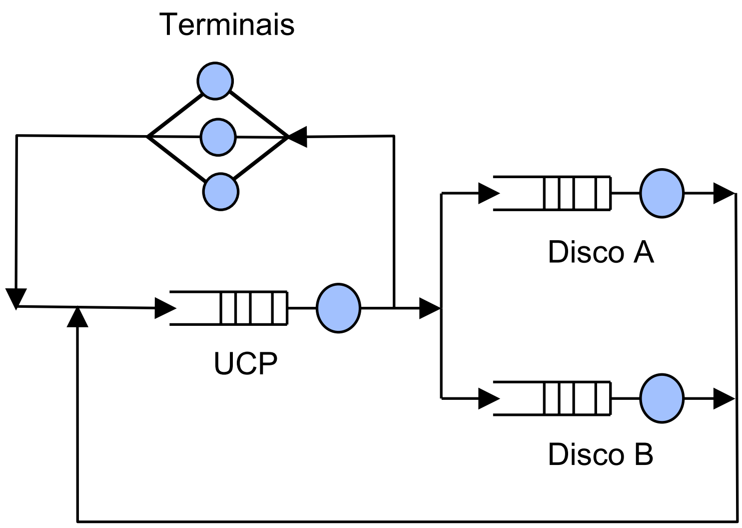
6 Exercícios

1) Considere o sistema em batch da figura com dois discos. Cada programa faz 4 visitas ao disco A e 20 visitas ao disco B. O tempo de serviço por visita à UCP, disco A e disco B são 40, 25 e 30 ms. O grau de multiprogramação é 3. Determine a distribuição do tamanho da fila da UCP. Determine também a vazão do sistema variando o grau de multiprogramação de 1 a 3.



Resp.: $P(Q_{UCP}=n|\ N=3)$ para n=0,1,2,3 são: 0.108, 0.180, 0.293 e 0.419
$X(N)$ para n=1,2,3 são: 0.588, 0.798 e 0.892.

2) Em um sistema de timesharing com 2 discos (para usuários e sistema), após o uso da UCP, a probabilidade de um programa utilizar o disco A é de 0,80, de utilizar o disco B é de 0,16 e de utilizar os terminais é de 0,04. O tempo que o usuário fica pensando é de 5 segundos, o tempo de serviço dos discos A e B é 30 e 25 ms, e o tempo médio de serviço por visita à UCP é 40 mseg. Analise o sistema de timesharing usando o método de convolução com N=3 usuários. Determine a distribuição do tamanho da fila da UCP.



Resp.: $P(Q_{UCP}=n|\ N=3)$ para n=0,1,2,3 são: 0.580, 0.298, 0.103 e 0.018.

3) Seja o seguinte sistema de filas onde a UCP, o disco A e o disco B possuem tempos médios de serviço respectivamente iguais a 0.2, 0.4 e 0.5 seg, taxa de visitas à UCP do disco A e disco B respectivamente 1, 0.5 e 0.5 e grau de multiprogramação 2.

   Analise o sistema pelo método da convolução e calcule X, U_i, Q_i, R_i e R e as probabilidades de ter 0, 1 e 2 programas na UCP.