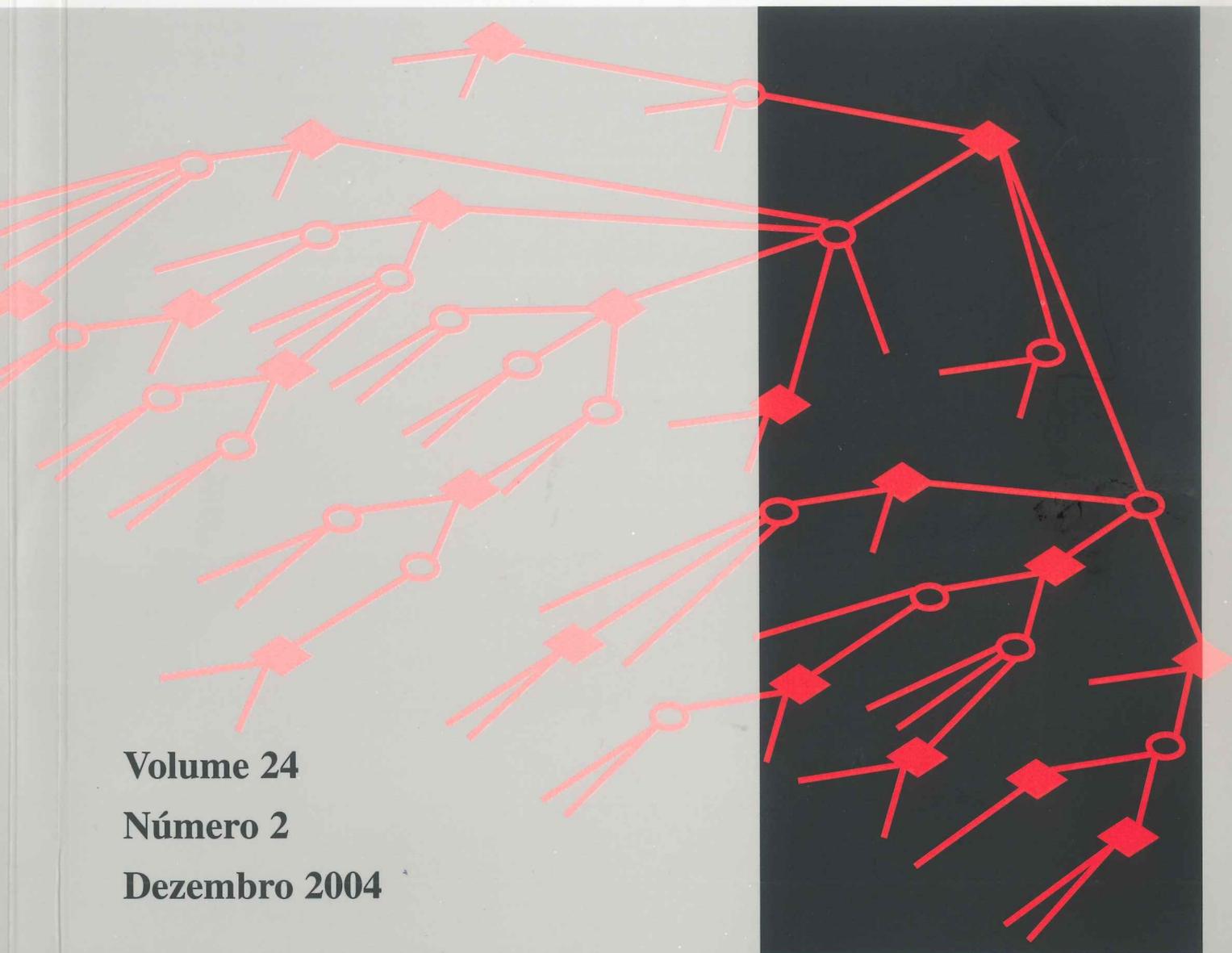


Investigação Operacional



Volume 24

Número 2

Dezembro 2004

INVESTIGAÇÃO OPERACIONAL

Volume 24 — nº 2 — Dezembro 2004

Publicação Semestral

Editor Principal: Joaquim J. Júdice
Universidade de Coimbra
Editor Adjunto: José F. Oliveira
Universidade do Porto

Comissão Editorial

M. Teresa Almeida Inst. Sup. Economia e Gestão	J. Rodrigues Dias Univ. de Évora	Rui Oliveira Inst. Superior Técnico
C. Henggeler Antunes Univ. de Coimbra	Laureano Escudero IBM, Espanha	J. Pinto Paixão Univ. de Lisboa
Marcos Arenales Univ. de São Paulo	Edite Fernandes Univ. do Minho	M. Vaz Pato Inst. Sup. Economia e Gestão
Jaime Barceló Univ. de Barcelona	J. Soeiro Ferreira Univ. do Porto	Mauricio G. Resende AT&T Labs Research
Eberhard E. Bischoff University of Wales, Swansea	J. Fernando Gonçalves Univ. do Porto	A. Guimarães Rodrigues Univ. do Minho
C. Bana e Costa Inst. Superior Técnico	Luís Gouveia Univ. de Lisboa	António J. L. Rodrigues Univ. de Lisboa
M. Eugénia Captivo Univ. de Lisboa	Rui C. Guimarães Univ. do Porto	J. Pinho de Sousa Univ. do Porto
Domingos M. Cardoso Univ. de Aveiro	J. Assis Lopes Inst. Superior Técnico	Reinaldo Sousa Univ. Católica, Rio Janeiro
João Clímaco Univ. de Coimbra	Carlos J. Luz Inst. Polit. Setúbal	L. Valadares Tavares Inst. Superior Técnico
J. Dias Coelho Univ. Nova de Lisboa	Virgílio P. Machado Univ. Nova de Lisboa	B. Calafate Vasconcelos Univ. do Porto
João P. Costa Univ. de Coimbra	Manuel Matos Univ. do Porto	Luís N. Vicente Univ. de Coimbra
Ruy Costa Univ. Nova de Lisboa	N. Maculan Univ. Fed., Rio Janeiro	Victor V. Vidal Technical Univ. of Denmark

Modelo multiobjectivo para apoio à localização de condensadores em redes radiais de distribuição

Dulce F. Pires * A. Gomes Martins † C. Henggeler Antunes †

* Escola Superior de Tecnologia de Setúbal
Instituto Politécnico de Setúbal
Portugal
dcosta@est.ips.pt

† Departamento de Engenharia Electrotécnica e Computadores
Faculdade de Ciências e Tecnologia
Universidade de Coimbra
Portugal

† INESC Coimbra
Portugal
{amartins,ch}@deec.uc.pt

Abstract

The installation of shunt capacitors in electrical distribution networks can effectively reduce energy and peak power losses, while improving quality of service particularly promoting a better voltage profile. Economic and operational benefits depend mainly on the number, location and sizes of the capacitors installed.

In this problem, multiple and conflicting evaluation aspects are generally at stake, such as minimizing capacitor installation cost and reducing system losses. Multiple objective models explicitly address these distinct concerns, and they entail analyzing trade-offs between the different objectives to select a satisfactory compromise solution from the set of potentially non-dominated solutions which can be accepted as the output of the decision process.

In this paper, a tabu search based approach is presented, which provides decision support in the problem of capacitor location in radial distribution networks. The model explicitly considers two conflicting and incommensurate objective functions, related to cost and operation aspects of evaluation.

Resumo

A instalação de condensadores em redes eléctricas de distribuição reduz efectivamente as perdas de energia e a potência de pico, para além de contribuir para uma melhor

qualidade de serviço, nomeadamente na melhoria do perfil de tensões apresentado pela rede. Os benefícios económicos e operacionais obtidos com a colocação de condensadores estão intimamente ligados ao número, localização e dimensão dos dispositivos instalados.

Neste problema existem vários aspectos de avaliação de natureza conflituosa, tais como a redução de perdas e a minimização do custo de instalação das baterias de condensadores. Enveredou-se, por isso, por uma abordagem multiobjectivo, com o desenvolvimento de um modelo que contempla explicitamente os diferentes aspectos de avaliação e permite a selecção de uma solução de compromisso satisfatória, a partir do conjunto de soluções potencialmente não dominadas resultante do processo de pesquisa.

Nesta abordagem, baseada em pesquisa tabu, propõe-se um modelo multiobjectivo de apoio à decisão para o problema da colocação de condensadores em redes de distribuição. O modelo considera explicitamente duas funções objectivo, relacionadas com os aspectos de avaliação referentes ao custo e à redução de perdas.

Keywords: multiobjective models, tabu search, distribution electrical networks.

Title: A multiobjective model for the location of capacitors in radial distribution networks.

1 Introdução

O problema da colocação de condensadores em redes eléctricas de distribuição para compensação de energia reactiva envolve a determinação dos locais, dimensão e número de condensadores a instalar na rede, com o objectivo de atingir o máximo de benefícios, assegurando que as restrições operacionais do sistema, com níveis de carga diferentes, sejam satisfeitas.

Este problema tem sido objecto de estudo por parte de diversos autores, sendo alvo de abordagens diferentes conforme o tipo de metodologias de resolução e/ou simplificações utilizadas. As inúmeras abordagens que foram sendo propostas acompanharam a evolução dos métodos matemáticos existentes para o tratamento deste tipo de problemas.

Os primeiros trabalhos realizados para otimizar a colocação de condensadores usam métodos analíticos para calcular o extremo de uma “função custo” representativa do somatório dos custos a minimizar. Estes métodos permitem obter facilmente soluções para o problema. No entanto, os autores utilizam simplificações irrealistas, tais como a utilização de um único alimentador de secção e características idênticas em todo o seu comprimento, e a carga distribuída uniformemente ao longo do alimentador. Encontram-se neste tipo de abordagens os trabalhos [1, 2, 3, 4, 5, 6, 7], entre outros.

Os trabalhos que utilizam métodos de programação matemática começaram a surgir quando a utilização de recursos computacionais se tornou mais generalizada. Dentro destes métodos, [8, 9, 10] entre outros, utilizam programação linear. Para tal linearizam o problema formulado, uma vez que este é intrinsecamente de programação matemática com função objectivo não linear, variáveis inteiras e contínuas, e restrições não lineares. [11, 12, 13], entre outros, utilizam programação inteira mista e, apesar de aproximarem o modelo matemático do problema por um modelo linear, consideram a natureza discreta de algumas variáveis.

O reconhecimento da natureza combinatoria do problema de colocação de condensadores levou a que, mais recentemente, com o desenvolvimento e a generalização dos métodos

heurísticos, estes tenham vindo a ter uma aplicação crescente na sua resolução. Por exemplo, [14, 15, 16, 17] utilizam algoritmos genéticos na resolução do problema da localização de condensadores. [18, 19, 20, 21], entre outros, recorrem a abordagens baseadas em *Simulated Annealing*. A pesquisa tabu é utilizada por [22, 23, 24, 25], entre outros.

Os métodos baseados em técnicas oriundas da inteligência artificial são dos mais recentes, tendo começado a ser aplicados ao problema da colocação de condensadores apenas na última década. No entanto, existem na literatura científica diversos trabalhos, particularmente os que utilizam teoria dos conjuntos difusos, [26, 27], entre outros, e redes neuronais, [28, 29], entre outros.

2 Modelo Matemático

2.1 Descrição do Problema

O problema de optimização de energia reactiva em redes de distribuição pode formular-se, em linguagem corrente, da seguinte forma. Pretende-se, dada uma rede de distribuição, com um perfil de carga variável no tempo, obter a localização e dimensão adequadas das baterias de condensadores a colocar nos barramentos, com o objectivo de atingir as condições de exploração com menores perdas ao custo mais baixo.

Na formulação matemática do problema, é necessário considerar as características de operação da rede de distribuição e das baterias de condensadores a serem instaladas, que se referem brevemente a seguir.

A rede de distribuição radial pode ser modelada fisicamente com a estrutura que se apresenta na Figura 1 e pormenorizada na Figura 2. A rede tem o seu início na subestação (SE), da qual parte um alimentador principal. Ao longo desse alimentador há pontos de derivação (nós ou barramentos), nos quais se podem encontrar: cargas activas e reactivas alimentadas directamente, ramos laterais, postos de transformação e elementos de controlo da potência reactiva. Estes últimos caracterizam-se por fornecerem energia reactiva à rede (o caso dos condensadores), ou por consumirem energia reactiva da rede (o caso das indutâncias). Os ramos laterais têm a mesma estrutura do alimentador principal, com as modificações devidas à sua menor dimensão.

Neste trabalho consideraram-se as seguintes simplificações:

- A rede trifásica de distribuição é equilibrada e por isso pode ser representada pelo seu diagrama equivalente de uma só fase; todos os estudos feitos sobre uma fase se podem generalizar às outras duas.
- Com os comprimentos típicos dos alimentadores da rede de distribuição radial e com esta a operar aos níveis de tensão utilizados, considera-se que as capacidades *shunt* das linhas são desprezáveis.
- Na resolução do problema da colocação de condensadores optou-se por considerar o instante de tempo em que o trânsito de energia reactiva é mais desfavorável para a operação do sistema, com uma situação simulada de ponta do sistema.

Considera-se como condição do problema que as baterias de condensadores são unidades organizadas em escalões, controladas automaticamente por um relé varimétrico, colocadas no barramento de baixa tensão dos Postos de Transformação escolhidos. De notar que se op-

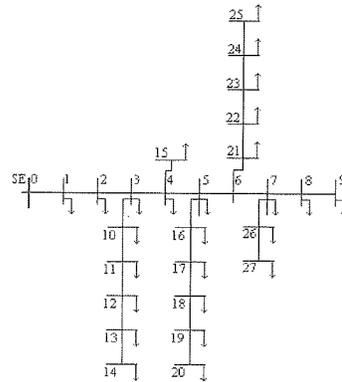


Figura 1: Exemplo de uma rede de distribuição radial, considerando a existência de ramos laterais.

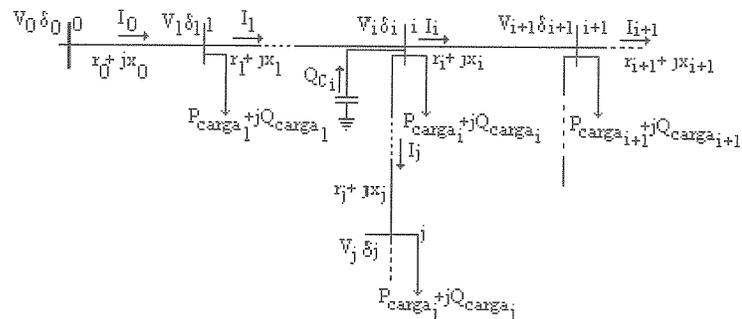


Figura 2: Pormenor da rede de distribuição com as grandezas associadas a cada trecho.

Tabela 1: Relação entre a capacidade e custo das baterias de condensadores.

Dimensão das Baterias de Condensadores (Q_{Fi}) (em KVA)	Custo (c_i) (em unidades monetárias)
Q_{F0}	$c_0 = 0$
Q_{F1}	c_1
(...)	(...)
Q_{Fj}	c_j
(...)	(...)
Q_{FY}	c_Y

tou por colocar as baterias de condensadores no lado de baixa tensão, por vários motivos. Por razões económicas: o custo do equipamento a ser instalado para fazer a compensação é muito superior se se colocar a compensação do lado de média tensão. Por razões técnicas: não se encontraram, em nenhuma das marcas de fabricantes de material eléctrico, baterias de capacidade suficientemente baixa para fazer compensação em postos de transformação no lado de média tensão e a instalação de um relé varimétrico, do lado de baixa tensão, suprime a necessidade de telecontrolo, além de permitir controlar as baterias de condensadores por escalões.

A utilização de condensadores com escalões justifica-se para garantir que, em situações de carga diferentes da escolhida para otimizar a colocação de condensadores, não haja sobrecompensação, e as necessidades de potência reactiva se aproximem tanto quanto possível das exigidas pela rede.

Utilizaram-se condensadores de valor padrão para aproximar o mais possível os dados do problema das características dos equipamentos disponíveis no mercado.

Usa-se uma tabela, como por exemplo a Tabela 1, para relacionar o custo das baterias de condensadores com a sua dimensão.

No custo das baterias apenas está considerado o custo de aquisição.

2.2 Modelo Multiobjectivo

Na formulação matemática do problema recorreu-se à seguinte notação:

- SE – Subestação.
- m – Índice que referencia os barramentos dos ramos laterais.
- n – Índice que referencia os ramos laterais.
- k – Índice que referencia os barramentos do alimentador principal.
- L – Índice que indica que a variável é de uma carga.
- C – Índice que indica que a variável é de compensação.
- F – Índice que indica que a variável é de um condensador padrão (tabela).
- c_j – Custo associado ao condensador Q_{Fj} .
- V_{nm}^k – Tensão no barramento m (lateral n que deriva do barramento principal k).
- P_{Lnm}^k – Carga (potência activa) alimentada a partir do barramento m (lateral n que deriva do barramento principal k).
- Q_{Lnm}^k – Carga (potência reactiva) alimentada a partir do barramento m (lateral n que deriva do barramento principal k).
- P_{nm}^k – Potência activa que flui no troço m , i.e. sai do barramento m para o barramento $m + 1$ (lateral n que deriva do barramento principal k).
- Q_{nm}^k – Potência reactiva que flui no troço m , i.e. sai do barramento m para o barramento $m + 1$ (lateral n que deriva do barramento principal k).
- Q_{Cnm}^k – Potência reactiva de compensação existente no barramento m (lateral n que deriva do barramento principal k).
- B_{nm}^k – Barramento m do lateral n , o qual tem origem no barramento principal k .
- T_{nm}^k – Troço m do lateral n , o qual tem origem no barramento principal k , que liga o barramento B_{nm}^k ao barramento seguinte $B_{n(m+1)}^k$.
- Q_{Fj} – Capacidade do condensador padrão j (tabela).
- M_n^k – Número de barramentos do lateral n que deriva do barramento k do alimentador principal.
- N^k – Número de laterais que derivam do barramento k do alimentador principal.
- K – Número de barramentos do alimentador principal.
- Y – Número de opções de baterias de condensadores consideradas para colocação.

O modelo multiobjectivo criado para o problema da localização de condensadores destina-se a apoiar a tomada de decisão no planeamento da compensação de energia reactiva. Este modelo multiobjectivo fornece ao decisor a possibilidade de escolher uma solução (não dominada) de compromisso entre dois objectivos:

- minimizar as perdas do sistema,
 - minimizar o custo da colocação dos condensadores,
- segundo as prioridades de planeamento.

Destá forma, o modelo respeita a natureza diferente de cada um dos objectivos, um de natureza operacional e outro de natureza económica.

As perdas resistivas de um troço T_{nm}^k da rede são dadas pela expressão (2.1). Verifica-se que alterações nas potências de compensação, variáveis Q_{Cnj}^k , ($j=m+1, \dots, M_n^k$) modificam o valor das perdas no troço.

$$r_{nm}^k \frac{P_{nm}^{k\ 2} + Q_{nm}^{k\ 2}}{V_{nm}^{k\ 2}} \tag{2.1}$$

Considerando todos os troços da rede de distribuição teremos a expressão (2.2).

$$\sum_{k=0}^K \sum_{n=0}^{N^k} \sum_{m=0}^{M_n^k} r_{nm}^k \frac{P_{nm}^{k\ 2} + Q_{nm}^{k\ 2}}{V_{nm}^{k\ 2}} \tag{2.2}$$

Esta expressão, correspondente às perdas resistivas do sistema, é a função objectivo a minimizar, apresentada em (2.3).

$$Min \sum_{k=0}^K \sum_{n=0}^{N^k} \sum_{m=0}^{M_n^k} r_{nm}^k \frac{P_{nm}^{k\ 2} + Q_{nm}^{k\ 2}}{V_{nm}^{k\ 2}} \tag{2.3}$$

Os custos de aquisição dos condensadores consideram-se dados por uma tabela tal como se exemplifica na Tabela 1. A opção de colocar o condensador Q_{Fj} no barramento B_{nm}^k incrementa o custo de compensação de c_j unidades. Assim, a função custo pode escrever-se através da expressão (2.4).

$$\sum_{k=0}^K \sum_{n=0}^{N^k} \sum_{m=0}^{M_n^k} \sum_{j=1}^Y j a_{nm}^k c_j \tag{2.4}$$

em que c_j é custo da bateria de condensadores Q_{Fj} ($j = 1, \dots, y$).

A variável inteira $j a_{nm}^k$ toma o valor 1 no caso de existir um condensador do tipo Q_{Fj} no barramento B_{nm}^k como se indica em (2.5).

$$j a_{nm}^k = \begin{cases} 1 & \text{se é colocado um condensador do tipo } Q_{Fj} \text{ em } B_{nm}^k \\ 0 & \text{caso contrário} \end{cases} \tag{2.5}$$

A expressão (2.6) é a função objectivo a minimizar, relativa ao custo.

$$Min \sum_{k=0}^K \sum_{n=0}^{N^k} \sum_{m=0}^{M_n^k} \sum_{j=1}^Y j a_{nm}^k c_j \tag{2.6}$$

Cada solução do problema tem que respeitar um conjunto de restrições que permitem garantir o funcionamento do sistema e asseguram que as condições, nomeadamente de carácter técnico, definidas previamente para o problema sejam respeitadas.

As restrições de igualdade, equações (2.7) a (2.11), garantem a operação do sistema, e correspondem ao conjunto de equações relativas ao trânsito de potência na rede radial.

As equações (2.7) e (2.8) garantem a mesma condição de operação, mas no primeiro caso a relação foi estabelecida para os laterais, e no segundo para o alimentador principal.

$$P_{n(i+1)}^k = P_{ni}^k - r_{ni}^k \frac{P_{ni}^{k\ 2} + Q_{ni}^{k\ 2}}{V_{ni}^{k\ 2}} - P_{Ln(i+1)}^k, \forall n \neq 0 \text{ e } k \neq 0 \tag{2.7}$$

$$P_{0(i+1)}^0 = P_{0i}^0 - r_{0i}^0 \frac{P_{0i}^{0^2} + Q_{0i}^{0^2}}{V_{0i}^{0^2}} - P_{L0(i+1)}^0 - \sum_{n=1}^{N^{i+1}} P_{n0}^{(i+1)} \quad (2.8)$$

Acontece o mesmo para as equações (2.9) e (2.10), em que a primeira se destina a assegurar a condição da potência reactiva para os laterais e a segunda para o alimentador principal.

$$Q_{n(i+1)}^k = Q_{ni}^k - x_{ni}^k \frac{P_{ni}^{k^2} + Q_{ni}^{k^2}}{V_{ni}^{k^2}} - Q_{Ln(i+1)}^k + Q_{Cn(i+1)}^k, \quad \forall n \neq 0 \text{ e } k \neq 0 \quad (2.9)$$

$$Q_{0(i+1)}^0 = Q_{0i}^0 - x_{0i}^0 \frac{P_{0i}^{0^2} + Q_{0i}^{0^2}}{|V_{0i}^0|^2} - Q_{L0(i+1)}^0 - \sum_{n=1}^{N^{i+1}} Q_{n0}^{i+1} + Q_{C0(i+1)}^0 \quad (2.10)$$

A relação (2.11) é a equação relativa à tensão nos nós do sistema.

$$V_{n(i+1)}^k{}^2 = V_{ni}^k{}^2 - 2 \times (r_{ni}^k P_{ni}^k + x_{ni}^k Q_{ni}^k) + (r_{ni}^k{}^2 + x_{ni}^k{}^2) \times \left(\frac{P_{ni}^{k^2} + Q_{ni}^{k^2}}{V_{ni}^{k^2}} \right) \quad (2.11)$$

É ainda necessário que a tensão nos nós do sistema permaneça dentro de valores que assegurem a qualidade de serviço. A condição (2.12) garante que o perfil de tensões das soluções admissíveis esteja no intervalo $[V_{nm_{min}}^k, V_{nm_{max}}^k] \forall m, n, k$.

$$V_{nm_{min}}^k \leq V_{nm}^k \leq V_{nm_{max}}^k \quad \forall m, n, k \quad (2.12)$$

A restrição (2.13) impõe que apenas pode ser colocado um condensador em cada nó.

$$\sum_{j=1}^Y j a_{nm}^k \leq 1 \quad \forall m, n, k \quad (2.13)$$

A potência reactiva de compensação em cada barramento é dada pela equação (2.14) e a sua expressão pode ser substituída directamente nas equações em que aparece.

$$Q_{Cnm}^k = b_{nm}^k \sum_{j=1}^Y j a_{nm}^k Q_{Fj} \quad \forall m, n, k \quad (2.14)$$

A variável b_{nm}^k , expressão (2.15), é uma variável binária que indica se é possível ou não colocar condensador no nó B_{nm}^k , uma vez que existem barramentos onde, por razões físicas ou operacionais do sistema, não é possível colocar as baterias de condensadores.

$$b_{nm}^k = \begin{cases} 1 & \text{se é possível colocar condensadores em } B_{nm}^k \\ 0 & \text{caso contrário} \end{cases} \quad (2.15)$$

3 Ferramenta Desenvolvida

Com o objectivo de resolver o problema de colocação e operação de condensadores em redes de distribuição de energia alguns autores desenvolveram ferramentas computacionais. SAMVAR

- *Simulated Annealing based computer package for Multi-objective VAR planning* [30], baseado em *simulated annealing*, e GANOS – *Genetic Algorithm Network Optimization System* [31], baseado em algoritmos genéticos, são exemplos de duas aplicações computacionais dedicadas a este problema.

A necessidade de desenvolver uma aplicação computacional específica que permita obter soluções para o problema deve-se à sua complexidade e à inexistência de ferramentas computacionais genéricas para problemas com estas características. O problema é multiobjectivo, com variáveis reais, inteiras e binárias, e apresenta restrições e funções objectivo não lineares, o que dificulta a utilização de ferramentas de optimização (comerciais ou não) já existentes. Além disso a opção de utilizar uma meta-heurística, pesquisa tabu, algoritmos genéticos, *simulated annealing*, ou outra, obriga necessariamente à criação de um ambiente de simulação e teste, em que se tenha o cuidado de adaptar a metodologia de resolução à especificidade do problema.

A aplicação computacional criada especialmente para apoio à decisão na colocação de condensadores em redes de distribuição radiais é essencialmente composta por quatro processos fundamentais: o de análise do trânsito de potência da rede, o de geração de cenários de compensação de forma aleatória, o de criação de cenários de compensação através da pesquisa local baseada na pesquisa tabu e ainda o de selecção das soluções potencialmente não dominadas (p.n.d) do problema.

Cada cenário de compensação gerado, de forma aleatória ou por pesquisa de vizinhança, é analisado sob o ponto de vista do trânsito de potência da rede, verificando-se desta forma a admissibilidade das soluções geradas relativamente às equações de operação do sistema e limites de tensão dos barramentos. No fim de cada iteração, em que são geradas várias soluções admissíveis, o processo de selecção das soluções potencialmente não dominadas do problema actualiza o ficheiro onde se guardam estas soluções.

Quando termina o processo de geração aleatória de soluções, passa-se para o de optimização (com a meta-heurística pesquisa tabu). Neste processo escolhe-se para *semente* da pesquisa uma das soluções potencialmente não dominadas geradas aleatoriamente. Esta solução inicial é um “compromisso” entre perdas resistivas e custo, ou seja é escolhido um cenário de compensação de valor intermédio relativamente a cada uma destas grandezas.

A solução inicial é depois sujeita a um processo de optimização que permite obter uma solução de compromisso satisfatório num tempo de cálculo computacional baixo, através da pesquisa tabu.

3.1 Análise do Trânsito de Potência

Existem diversas abordagens ao problema do trânsito de potência e várias formas de modelar o sistema. No entanto, para a análise do trânsito de potência as aproximações mais conhecidas não foram desenvolvidas para redes com topologias radiais, mas sim para redes com topologias em anel.

As características distintas das redes de distribuição, relativamente às de transmissão – i.e. são redes radiais, com vários pontos de carga e alimentadores de menor comprimento em que a razão R/X é elevada –, fazem com que a aplicação de metodologias convencionais de resolução do trânsito de potência desenvolvidas para redes de transporte, como, por exemplo,

as baseadas no método de *Newton-Raphson*, não convirjam sem a adopção de técnicas de tratamento de matrizes esparsas. Por estas razões, aplicou-se neste trabalho uma metodologia adaptada às características da rede de distribuição radial, que permite analisar o trânsito de potência de forma expedita e com resultados precisos.

Existem alguns trabalhos que se destinam exactamente a tratar o problema do trânsito de potência em redes radiais, tais como [32, 33, 34, 35]. Destes, apenas [33] cria um algoritmo de resolução substancialmente diferente dos convencionais, adaptado a redes com topologia radial. Além disso, trabalhos como [12, 13], dedicados ao problema da optimização de energia reactiva em redes radiais, apresentam contribuições para o cálculo do fluxo de cargas também adaptadas ao carácter radial destas redes. É nestas três últimas contribuições que se baseou o método de análise do trânsito de potência desenvolvido neste trabalho.

A metodologia de análise do trânsito de potência foi validada por comparação dos resultados obtidos, por simulação da rede fornecida pela EDP Distribuição, na aplicação desenvolvida e na aplicação comercial utilizado pela EDP Distribuição, DINIS (*Distribution Network Information System*). Os resultados reproduziram exactamente aqueles que a aplicação DINIS forneceu.

3.2 Geração Aleatória de Soluções

A geração aleatória de soluções é constituída por duas componentes fundamentais: a primeira gera aleatoriamente uma solução inicial, enquanto que a segunda introduz modificações nesta solução também de forma aleatória. O processo que compreende estas duas componentes é repetido durante um número de iterações especificado.

Para criar um cenário de compensação inicial são gerados aleatoriamente (dentro da gama de valores de capacidade e locais possíveis para colocar as baterias de condensadores):

- os locais onde se coloca a compensação;
- a capacidade das baterias a colocar nos barramentos indicados.

A partir deste cenário de compensação inicial, ou primeira solução, são geradas novas soluções através da introdução de outros pontos de compensação relativamente aos que a primeira solução apresenta. Nesta fase ainda não existe a preocupação de minimizar as duas funções objectivo, até porque a introdução de novos pontos de compensação tem como consequência directa o aumento do valor da função custo. Todas as soluções geradas são posteriormente avaliadas através do processo de selecção de soluções potencialmente não dominadas.

Neste caso, apenas se pretende introduzir modificações na solução inicial de uma forma simples e que consome pouco tempo de cálculo. Assim, cria-se uma base de soluções a partir da qual se vai seleccionar a solução que servirá de semente ao processo de optimização com a pesquisa tabu.

Os movimentos que permitem transformar uma solução x numa solução y são obtidos por alteração de apenas um ponto de compensação. O sistema de geração de novas soluções tem por objectivo adicionar novos pontos de compensação à solução inicial, uma vez que esta apresenta, na maioria dos casos, um número reduzido de locais compensados.

O perfil de compensação da solução y obtém-se a partir da solução x da seguinte forma:

Durante k iterações (com k igual ao número de nós da rede):

- passo 1: gera aleatoriamente um novo local de compensação.
- Passo 2: verifica se o local de compensação indicado já tem alguma bateria de condensadores.
- Passo 3: se sim, volta ao passo 1; caso contrário, gera aleatoriamente um valor de capacidade dentro da gama de baterias existentes, e compensa com esse valor o nó indicado no passo 1.
- Passo 4: faz o cálculo do trânsito de potência e grandezas da rede.
- Passo 5: se a bateria de condensadores colocada tem o valor máximo possível de capacidade, termina; caso contrário, incrementa para o valor de capacidade imediatamente superior e vai para o passo 4.

3.3 Melhoria de Soluções com Pesquisa Tabu

Descrevem-se aqui as opções principais do algoritmo da pesquisa tabu desenvolvido para o problema de colocação de condensadores em redes de distribuição. Este algoritmo começou por uma versão simples desta meta-heurística que apenas utilizava memória de curto termo. Posteriormente introduziu-se uma fase de diversificação, com memória baseada na frequência, com a finalidade de procurar soluções com atributos diversos daqueles encontrados até ao momento. É esta última versão, mais completa, que aqui se apresenta.

Numa pesquisa de vizinhança cada solução x do espaço de soluções \mathbf{X} tem uma vizinhança $\mathbf{N}(x)$, definida como um subconjunto do espaço de soluções, $\mathbf{N}(x) \subset \mathbf{X}$, e cada solução $x' \in \mathbf{N}(x)$ pode ser alcançada a partir de x através de uma operação a que se chama movimento. Na versão da pesquisa tabu desenvolvida, as soluções de $\mathbf{N}(x)$ são todas aquelas que se conseguem obter a partir de uma solução qualquer x de acordo com as seguintes alterações:

- por redução da capacidade do condensador instalado no nó i ,
- por eliminação do condensador instalado no nó i ,
- por deslocação da capacidade instalada no nó i para o nó $i+1$,
- por deslocação da capacidade instalada no nó i para o nó $i-1$.

Com a finalidade de clarificar a definição de vizinhança apresentada, consideremos uma rede com 6 nós e a possibilidade de colocar dois tipos de condensadores, a e b (sendo a capacidade de a inferior à capacidade de b). Consideremos também que o seguinte vector representa a posição dos condensadores nos nós da rede (zero significa a não existência de condensador), de uma solução qualquer x' : 00a0bb. Segundo a definição de vizinhança adoptada tem-se $\mathbf{N}(x') = \{0000bb; 0a00bb; 000abb; 00a0ab; 00a00b; 00ab0b; 00a0ba; 00a0b0\}$.

Com a adopção deste tipo de vizinhança obteve-se um espaço de pesquisa bastante alargado. No entanto, verificou-se que com uma vizinhança mais reduzida a pesquisa não era tão eficaz. De facto, como há restrições à colocação de condensadores em determinados nós, isso tem como consequência uma diminuição das soluções de $N(x)$, i.e. a não admissibilidade das soluções reduz por si o espaço de pesquisa.

Em cada iteração do algoritmo da pesquisa tabu avalia-se um conjunto de soluções candidatas da vizinhança da solução actual, e não apenas uma solução. A partir deste conjunto de soluções, e considerando a respectiva admissibilidade, escolhe-se a melhor solução para uma dada função de avaliação.

O algoritmo desenvolvido baseado na pesquisa tabu usa uma *lista tabu* do tipo “primeiro a entrar, primeiro a sair” (FIFO – *first in, first out*) de tamanho L . Esta lista contém os índices das últimas L posições que trocaram de valor. Experimentaram-se alguns valores para L , que tipicamente deve variar entre 5 e 11, mas como este valor não revelou ter muita influência no resultado, estabeleceu-se um valor intermédio e os testes apresentados neste texto foram realizados com $L=7$.

Quanto à pesquisa, esta divide-se nas seguintes duas fases:

- uma fase inicial que utiliza, como único registo de memória, a lista tabu.
- Uma segunda fase de *diversificação*. Nesta fase, além da lista tabu é também imposta outra condição baseada na frequência dos valores das variáveis, que consiste no seguinte: uma variável pode trocar de valor se o número de vezes que isso aconteceu em iterações passadas não ultrapassa um determinado limiar.

A introdução de uma fase de diversificação na pesquisa surge da observação das soluções obtidas com o algoritmo simples. Estas apresentavam ainda alguma semelhança com a solução inicial, mesmo após um grande número de iterações. Com a fase de diversificação pretendeu-se encorajar a pesquisa de soluções significativamente diferentes daquela que deu origem à pesquisa.

Esta fase de diversificação usa um registo de memória de frequência dos valores das variáveis. O registo, designemo-lo por *frequência1*, é um vector de dimensão n (número de variáveis) que conta o número de vezes que cada variável trocou de valor em soluções admissíveis. O registo *frequência1* é utilizado na definição de vizinhança da segunda fase: uma solução vizinha de x é candidata a avaliação se resulta da comutação de uma variável x_j não tabu que verifica:

$$\begin{cases} frequ\u00eancia1 < filtro1 & \text{se } x_j \text{ passa a ter compensa\u00e7\u00e3o (i.e. passa de zero a um valor que} \\ & \text{indica o tipo de condensador)} \\ frequ\u00eancia1 < filtro0 & \text{se } x_j \text{ deixa de ter condensador (i.e. passa a ter o valor zero)} \end{cases} \quad (3.16)$$

No in\u00edcio da segunda fase *filtro1* e *filtro0* s\u00e3o iguais \u00e0 m\u00e9dia dos valores de *frequ\u00eancia1*.

Os dados come\u00e7am a ser registados em *frequ\u00eancia1* desde o in\u00edcio da primeira fase e s\u00e3o actualizados ao longo da segunda fase.

3.4 Selecção das Soluções Potencialmente não Dominadas

Uma dada solução admissível é não dominada se e só se não existe outra solução admissível que melhore simultaneamente todas as funções objectivo, i.e. a melhoria num objectivo só pode ser alcançada à custa da degradação do valor de pelo menos um dos outros objectivos.

Em problemas de programação matemática multiobjectivo, o cálculo de soluções não dominadas pode ser feito através de funções escalares substitutas cuja solução óptima é solução não dominada do problema multiobjectivo.

Quando a técnica utilizada para a resolução do problema multiobjectivo é uma meta-heurística, não há garantia de que a solução pertença à fronteira eficiente do problema. Podemos então caracterizar estas soluções como potencialmente não dominadas, dado que apenas se garante que o são relativamente ao conjunto de soluções encontradas durante o processo de optimização.

Nas abordagens multiobjectivo com meta-heurísticas em que se gera um conjunto de soluções potencialmente não dominadas, representativo de todo o conjunto de soluções não dominadas, utiliza-se habitualmente uma estrutura de dados que armazena e actualiza as soluções potencialmente não dominadas eficientes encontradas ao longo do processo de pesquisa. O processo de selecção de soluções potencialmente não dominadas desenvolvido neste trabalho baseia-se também na utilização de uma base de dados que facilita a avaliação do conjunto de soluções geradas.

O processo de selecção de soluções potencialmente não dominadas é pela primeira vez chamado no fim da geração aleatória de soluções. De entre todas as soluções admissíveis geradas, são então seleccionadas as soluções potencialmente não dominadas. A solução inicial da pesquisa com a meta-heurística pesquisa tabu é seleccionada a partir deste conjunto.

O processo de selecção de soluções potencialmente não dominadas é novamente chamado no fim de cada pesquisa de vizinhança, para actualizar o conjunto de soluções potencialmente não dominadas do problema. No final de uma corrida da aplicação, este conjunto contém as soluções que permitem traçar a “fronteira” de soluções não dominadas do problema.

4 Aplicação da Ferramenta Desenvolvida a uma Rede de Distribuição da EDP

A ferramenta computacional desenvolvida foi aplicada a três redes reais, duas retiradas da bibliografia [33] e outra cujos dados foram fornecidos pela EDP Distribuição. Os resultados que se apresentam neste artigo foram obtidos com esta última rede.

Em relação aos parâmetros da aplicação que podem ser facilmente modificados, considerou-se o seguinte:

- número de iterações para obter a solução inicial: 20;
- número de iterações da pesquisa tabu: 50;
- tamanho da lista tabu: 7;

Tabela 2: Perdas na rede quando não existe compensação.

	perdas, potência activa (MW)
rede sem compensação	0.2463

Tabela 3: Custo e perdas para a solução inicial da pesquisa tabu.

	custo (milhares escudos)	perdas, potência activa (MW)
solução inicial	8200	0.189

- número de vezes que se reduzem os níveis de reserva (valores máximos que o decisor está disposto a aceitar em cada função objectivo): 2;
- percentagem de redução dos níveis de reserva: 1%;
- posição da solução inicial: ponto intermédio na curva de soluções potencialmente não dominadas.

A rede eléctrica do sistema de distribuição português apresenta 94 nós, distribuídos por um alimentador principal extenso e por 23 laterais. Esta rede caracteriza-se por, em períodos de sobrecarga, como é o que está em estudo, apresentar um perfil de tensões em que a maioria dos nós apresenta um valor abaixo dos 0.93 p.u., que é o limite inferior do intervalo admissível para o valor da tensão.

Na Tabela 2 encontram-se os valores das perdas apresentadas pela rede sem compensação, para o cenário de carga escolhido.

Após a aplicação da abordagem proposta a esta rede, obtiveram-se os resultados ilustrados nas Figuras 3 e 4.

Na figura 3a apresenta-se o número total de soluções geradas e na Figura 3b o subconjunto de soluções admissíveis, enquanto na Figura 4 se apresentam apenas as soluções não dominadas antes e depois da aplicação da pesquisa tabu.

Na Figura 4a assinala-se a solução inicial utilizada para a fase de aplicação da pesquisa tabu, e apresentam-se na Tabela 3 as características desta solução.

Como exemplo de soluções de compensação para a rede em estudo, foram escolhidas as soluções 1, 2 e 3 indicadas na Figura 4b. Na Tabela 4 apresentam-se as características destas soluções.

Os valores da redução global das perdas em potência activa, em relação ao cenário em que não existe qualquer dispositivo de compensação, são apresentados na Tabela 5.

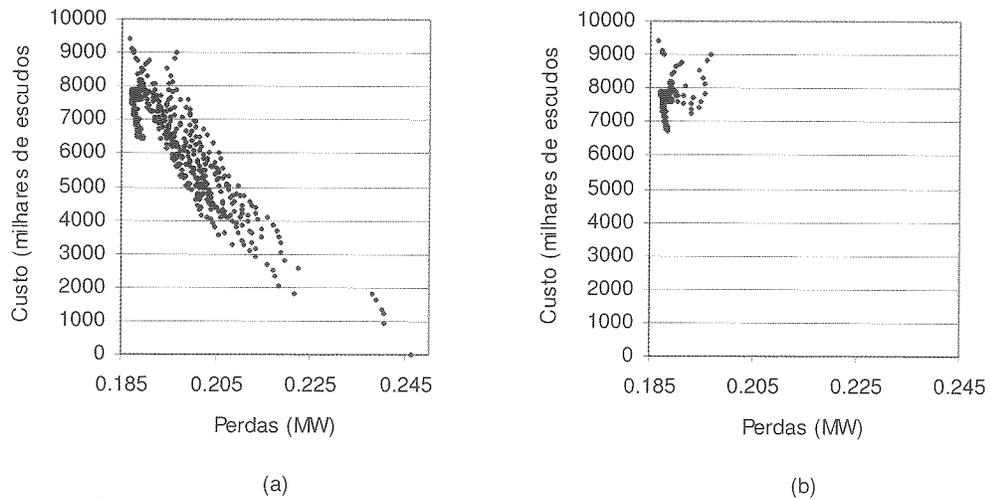


Figura 3: Representação das soluções geradas: total (a), admissíveis (b) para a rede portuguesa de 94 nós.

Tabela 4: Custos e perdas associados a cada uma das soluções exemplo (Figura 4b).

	custo (milhares de escudos)	potência activa de perdas (MW)
solução 1	6711	0.188
solução 2	7245	0.1874
solução 3	7867	0.187

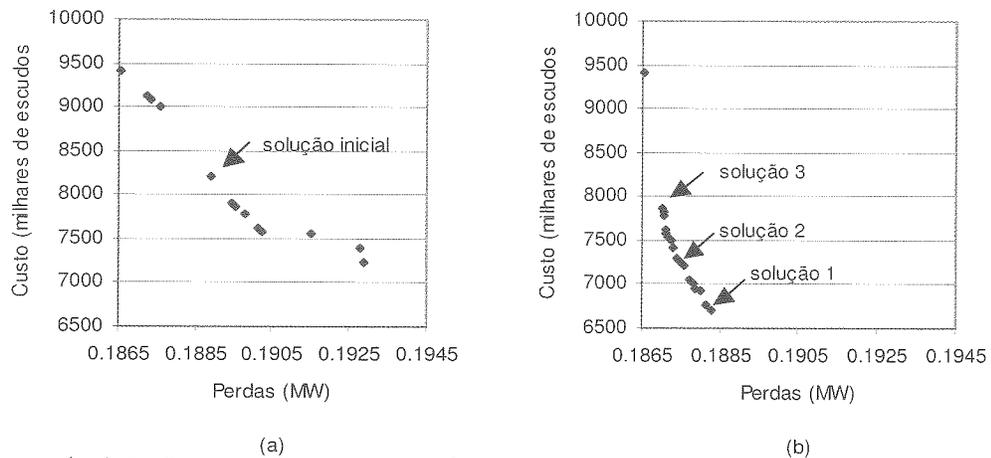


Figura 4: Soluções potencialmente não dominadas: antes (a) e depois (b) da aplicação da pesquisa tabu.

Tabela 5: Redução de perdas, em percentagem, para as soluções exemplo(Figura 4b).

	solução 1	solução 2	solução 3
redução de perdas potência activa (%)	23.5	23.9	24

	sem compensação	solução 1	solução 2	solução 3
distância euclideana	0.783	0.491	0.485	0.477
maior distância	0.11	0.07	0.069	0.068

Tabela 6: Indicadores das melhorias obtidas na tensão.

O perfil de tensões apresentado pela rede deve encontrar-se dentro do intervalo $[0.93, 1.07]$ p.u.. No caso da rede em estudo, para o cenário escolhido, as tensões dos nós, principalmente nas extremidades dos ramos, encontram-se afastadas do limite inferior do intervalo. A compensação de energia reactiva permite melhorar o perfil de tensões da rede, trazendo-o para o intervalo pretendido. De notar que as soluções em que a tensão de algum nó se encontre fora deste intervalo são consideradas não admissíveis, e verifica-se que o intervalo de soluções admissíveis é bastante menor que o total de soluções geradas (Figura 3). Em relação ao perfil de tensões, as soluções que servem de exemplo representam uma melhoria significativa.

O valor da tensão mais afastado do limite inferior do intervalo admissível é o do extremo do alimentador principal, no nó 32. Este valor apresenta um desvio de 0.11 antes da compensação, enquanto que depois da compensação, para a solução exemplo 2, é de apenas 0.069. Como indicador da melhoria global do perfil de tensões do sistema, pode considerar-se a distância euclideana (Tabela 6) ao perfil “ideal” em que se consideram todos os nós com a tensão de 1 p.u..

5 Conclusões

O problema da colocação de condensadores em redes de energia, que envolve a localização e dimensionamento de baterias de condensadores para compensação da energia reactiva, apresenta aspectos conflituosos. A redução de custos de investimento, aquisição, operação e manutenção das baterias, está em conflito com a redução das perdas do sistema por compensação da energia reactiva. Por outro lado, a melhoria do perfil de tensões do sistema pode também estar em conflito com qualquer um dos outros aspectos de avaliação. Assim, o problema é por natureza um problema multiobjectivo, envolvendo aspectos económicos, de operação e de qualidade de serviço.

No modelo matemático desenvolvido foram tidos explicitamente em conta os dois primeiros aspectos, de redução do custo e das perdas do sistema. O último aspecto, perfil de tensões, foi considerado como restrição, embora se tenha criado, na fase de análise de soluções, um indicador para que o agente de decisão possa aferir a qualidade das soluções sob esse aspecto.

Com o auxílio do modelo multiobjectivo criado para o problema de colocação de condensadores, o gestor da rede tem possibilidade de escolher soluções de compromisso satisfatórias entre os dois objectivos que estão em jogo e lhe interessam em simultâneo. O modelo matemático tem ainda a vantagem de respeitar a natureza discreta das variáveis, relativamente à capacidade e local de instalação. Permite também escolher previamente os locais onde não convém colocar compensação, seja por motivos técnicos, seja por motivos operacionais.

A aplicação da ferramenta desenvolvida a casos reais, nomeadamente da rede da EDP Distribuição, comprova que através desta metodologia é possível atingir um conjunto de soluções

potencialmente não dominadas com boas características, ou seja, que reduzem consideravelmente o volume de perdas e apresentam baixos custos.

Agradecimentos Este trabalho foi parcialmente financiado pela Fundação para a Ciência e Tecnologia, através do projecto POCTI/38422/ESE/2001.

Os autores agradecem ao revisor os comentários construtivos e a discussão suscitada no processo de revisão científica deste texto, que contribuirá quer para a melhoria e a clarificação de aspectos do artigo, quer para futuros desenvolvimentos.

6 Referências

- [1] Cook, R.F., (1964), "Calculating Loss Reduction Afforded by Shunt Capacitor Application", IEEE Transactions on Power Apparatus and Systems, PAS-83, 1227-1230.
- [2] Schmill, J.V., (1965), "Optimum Size And Location Of Shunt Capacitors On Distribution Feeders", IEEE Transactions on Power Apparatus and Systems, PAS-84 (9), 825 - 832.
- [3] Chang, Nelson E. (1969), "Locating Shunt Capacitors On Primary Feeder For Voltage Control And Loss Reduction", IEEE Transactions on Power Apparatus and Systems, PAS-88 (10), 1574-1577.
- [4] Chang, Nelson E. (1971), "Generalized Equations On Loss Reduction With Shunt Capacitors", IEEE Transactions on Power Apparatus and Systems, PAS-91, 2189-2195.
- [5] Bae, Y.G., (1978), "Analytical Method Of Capacitor Allocation On Distribution Primary Feeders" IEEE Transactions on Power Apparatus and Systems, PAS-97 (11), 1232 - 1238.
- [6] Grainger, J. J., S.H. Lee (1981), "Optimum Size and Location of Shunt Capacitors for Reduction of Losses on Distribution Feeders", IEEE Transactions on Power Apparatus and Systems, PAS-100 (3), 1105 - 1118.
- [7] Grainger, J. J., S.H. Lee (1982), "Capacity Release by Shunt Capacitor Placement on Distribution Feeders: a New Voltage-Dependent Model", IEEE Transactions on Power Apparatus and Systems, PAS-101 (5), 1236 - 1244.
- [8] Pretelt, A. M. (1971), "Automatic allocation of network capacitors", IEEE Transactions on Power Apparatus and Systems, PAS-90 (1), 54 - 61.
- [9] Fernandes, R.A., F. Lange, R.C. Burchett, H.H. Happ, K.A. Wirgau (1983), "Large Scale Reactive Power Planning", IEEE Transactions on Power Apparatus and Systems, PAS-102 (5), 1083 - 1088.
- [10] Chung, T. S., Ge Shaoyun (1997), "A Recursive LP-Based Approach for Optimal Capacitor Allocation with Cost-Benefit Consideration", Electric Power Systems Research, 39(1997), 129-136.
- [11] Aoki, K, M. Fan, A. Nishikori, (1988), "Optimal VAR Planning by Approximation Method for Recursive Mixed-Integer Linear Programming", IEEE Transactions on Power Systems, PWRS-3 (4), 1741 - 1747.
- [12] Baran, M. E., F.F. Wu (1989a), "Optimal Capacitor Placement on Radial Distribution System", IEEE Transactions on Power Delivery, 4 (1), 725-734.
- [13] Baran, M. E., F.F. Wu (1989b), "Optimal Sizing of Capacitors Placed on a Radial Distribution System", IEEE Transactions on Power Delivery, 4 (1), 735-743.

- [14] Iba, Kenji (1994), "Reactive Power Optimization by Genetic Algorithm", IEEE Transactions on Power Systems, 9(2), 685-692.
- [15] Lee, K.Y., Xiaomin Bai, Young-Moon Park (1995), "Optimization Method for Reactive Power Planning by Using a Modified Simple Genetic Algorithm", IEEE Transactions on Power Systems, 10(4), 1843-1850.
- [16] Kim, K.-H., S.-K. You (1999), "Voltage Profile Improvement by Capacitor Placement and Control in Unbalanced Distribution Systems Using GA", Proceedings of IEEE Power Engineering Society Summer Meeting, 2, 800-805.
- [17] Levitin, G., A. Kalyuhny, A. Shenkman, M. Chertkov (2000) "Optimal Capacitor Allocation in Distribution Systems Using a Genetic Algorithm and a Fast Energy Loss Computation Technique", IEEE Transactions on Power Delivery, 15(2), 623-628.
- [18] Chiang, H.D., J.C. Wang, O. Cockings (1990a), "Optimal Capacitor Placements in Distribution Systems Part I: A New Formulation and the Overall Problem", IEEE Transactions on Power Delivery, 5 (2), 634-642.
- [19] Chiang, H. D., J.C. Wang, O. Cockings (1990b), "Optimal Capacitor Placements in Distribution Systems Part II: Solution Algorithms and Numerical Results", IEEE Transactions on Power Delivery, 5 (2), 643-649.
- [20] Hsiao, Y.-T, C.-C. Liu, H.-D. Hsiao, Y.-L. Chen (1993), "A New Approach for Optimal VAR Sources Planning in Large Scale Electric Power Systems", IEEE Transactions on Power Systems, 8(3), 988-996.
- [21] Chang, C. S., J.S. Huang (1997), "Optimal SVC Placement for Voltage Stability Reinforcement", Electric Power Systems Research, 42(1997), 165-172.
- [22] Yang, H.-T., Y.-C. Huang, C.-L. Huang (1995), "Solution to Capacitor Placement Problem in a radial Distribution System Using Tabu Search Method", Proceedings of EMPD '95, International Conference on Energy Management and Power Delivery, 1, 388-393.
- [23] Mori, H., Y. Ogita (1999), "Capacitor Placement Using Parallel Tabu Search in Distribution Systems", Proceedings of SMC '99, International Conference on Systems, Man and Cybernetics, 6, 521-526.
- [24] Mori, H., Y. Ogita (2000), "Parallel Tabu Search for Capacitor Placement in Radial Distribution Systems", Proceedings of IEEE Power Engineering Society Winter Meeting, 4, 2334-2339.
- [25] Chang, C.S., L.P. Lern (2000), "Application of Tabu Search Strategy in Solving Non-Differentiable Savings Function for the Calculation of Optimum Savings Due to Shunt Capacitor Installation in a Radial Distribution System", Proceedings of IEEE Power Engineering Society Winter Meeting, 4, 2323-2328.
- [26] Chin, H.-C., W.-M. Lin (1994), "Capacitor Placements for Distribution Systems with Fuzzy Algorithm", TENCON '94 IEEE Region 10's Ninth Annual International Conference. The Frontiers of Computer Technology, 2, 1025-1029.
- [27] Abdul-Rahman, K. H., S.M. Shahidepour, M. Daneshdoost (1995), "AI Approach to Optimal VAR Control With Fuzzy Reactive Loads", IEEE Transactions on Power Systems, 10(1), 88-97.
- [28] Dash, P.K., S. Saha, P.K. Nanda (1991), "Artificial Neural Net Approach for Capacitor Placement in Power System", Proceedings of the First International Forum on Applications of Neural Networks to Power Systems, 1, 247-250.
- [29] Zhu, J. Z., C.S. Chang, W. Yan, G.Y. Xu (1998), "Reactive Power Optimisation Using an Analytic Hierarchical Process and a Nonlinear Optimisation Neural Network Approach", IEE Proceedings - Generation, Transmission and Distribution, 145(1), 89-97.

- [30] Hsiao, Y.-T, H.-D. Hsiao, C.-C. Liu, Y.-L. Chen (1994), "A Computer Package for Optimal Multi-Objective VAR Planning in Large Scale Power Systems", IEEE Transactions on Power Systems, 9(2), 668-676.
- [31] Glamocanin, V., D. Andonov, D. Trajanov, B. Stojkoska (1998), "Optimal Power System VAR Planning by AI Algorithm", MELECON 98, 9th Mediterranean Electrotechnical Conference, 2, 1066-1070.
- [32] Chen, T.-H., M.-S. Chen, K.-J. Hwang, , P. Koats , E.A. Chebli, (1991), "Distribution Sstem Power Flow Analysis - A Rigid Approach", IEEE Transactions on Power Delivery, 6(3), 1146-1152.
- [33] Das, D., H. S. Nagi, D.P. Kothari, (1994), "Novel Method for Solving Radial Distribution Networks", IEE Proceedings - Generation, Transmission and Distribution, 141(4), 291-298.
- [34] Moon, Y.-H., B.-K. Choi, B.-H. Cho, S.-H. Kim, B.-N. Ha, J.-H. Lee (1999), "Fast And Reliable Distribution System Load Flow Algorithm Based On The Y/Sub Bus/ Formulation", IEEE Power Engineering Society Summer Meeting 1999., (1), 238-242.
- [35] Expósito, A.G., E. R. Ramos (1999), "Reliable Load Flow Technique for Radial Distribution Networks", IEEE Transactions on Power Systems, 14 (3), 1063 - 1069.

Abordagem hierárquica para avaliação da fiabilidade de sistemas de produção complexos com comportamento não-markoviano

Eusébio Nunes * José Faria † Manuel Matos †, ‡

* Departamento de Produção e Sistemas,
Escola de Engenharia
Universidade do Minho
enunes@dps.uminho.pt

† FEUP – Faculdade de Engenharia
Universidade do Porto
jfaria@fe.up.pt

‡ INESC – Instituto de Engenharia de Sistemas e Computadores
Porto
mam@fe.up.pt

Abstract

In the actual context of global, intense competitiveness, the performance of the production systems is a key competitiveness factor for manufacturing companies. Factors such as the availability of the manufacturing cells and the frequency of production shortages are very important indexes for system designers and managers, both at the planning and at the operational stages of the production systems life cycle.

The reliability, maintainability and redundancy of the manufacturing equipment are well known performance drivers of the production system, but intermediate work-in-process (wip) buffers are another fundamental driver. Despite the endless inventory reduction efforts made by production managers, very often, wip buffers still play a central role in the operation of the production systems. They smooth the unbalance of manufacturing cells and limit the consequences of events such as equipment failures and non conforming lots, preventing them to propagate to the downstream production units.

This paper will present a set of tools and methods for the assessment of performance indexes of production systems. Particular attention will be paid to the assessment of the impact of the dimension and the localization of the wip buffers on the global performance of the systems.

The assessment of availability, frequency and duration related indexes is a complex task for production systems, due to the dimension of the systems under study, and the fact that their behaviour, normally, will not be markovian. To overcome these difficulties, a hierarchical, non-markovian approach was developed. Its originality lays on the fact that it

allows evaluating the performance indexes from a global system model that was obtained from the integration of a set of equivalent models derived at the subsystem level.

In the final part of the paper, a numerical example concerning a production system made by three cells and two buffers will be presented.

Resumo

No actual contexto de forte concorrência e de crescente exigência dos mercados, no projecto e exploração dos sistemas de produção, é fundamental ter em conta os aspectos que afectam a sua disponibilidade, isto é, a probabilidade do sistema estar operacional quando é solicitado a produzir. Também, a frequência de ocorrência dos estados de avaria e a duração média desses estados são índices muito úteis para o planeamento dos sistemas de produção e para a análise dos custos industriais.

Sabe-se da importância da fiabilidade e da manutabilidade dos equipamentos, das redundâncias e os recursos de manutenção sobre estes índices. Reconhece-se, também, o papel que os *buffers* existentes entre secções do sistema de produção desempenham, ao impedir ou atenuar a propagação, para jusante, de efeitos indesejáveis provocados por avaria de equipamentos, absentismo de operadores ou não conformidade de produtos. Neste artigo, será apresentado um método de avaliação do desempenho de sistemas de produção, em especial de análise do efeito que a dimensão e a localização dos *buffers* tem sobre esse desempenho.

A avaliação de índices como a disponibilidade, ou a frequência e a duração das avarias é, normalmente, uma tarefa complexa devido a duas ordens de factores principais: a dimensão dos sistemas em análise e o carácter não markoviano do seu comportamento. Para contornar estas dificuldades, será explorada uma abordagem hierárquica e não-markoviana. A sua originalidade reside no facto de possibilitar a obtenção de um modelo global do sistema, por integração de modelos simplificados ao nível dos subsistemas, com base no qual são determinados os índices de fiabilidade e as análises de sensibilidade.

Na parte final do artigo, apresentar-se-á a aplicação da metodologia a um sistema de produção constituído por três células de fabrico e dois *buffers*.

Keywords: reliability, hierarchical approach, non-markovian systems, manufacturing systems

Title: Reliability assessment of non-markovian complex production systems using a hierarchical approach.

1 Introdução

Para tornar os modelos utilizados na avaliação do desempenho de sistemas de produção sujeitos a processos tratáveis de forma analítica, adopta-se frequentemente a hipótese markoviana [1,2,3,4]. Esta hipótese, que consiste em admitir que todos os processos estocásticos (e.g., falha, reparação e reconfiguração) que determinam o comportamento dos sistemas são exponencialmente distribuídos, introduz uma grande simplificação na análise e obtenção de resultados.

Contudo, no caso da avaliação de sistemas de produção, a hipótese markoviana pode introduzir erros muito significativos. Isso é assim porque muitos dos processos relevantes do comportamento dos sistemas de produção são temporizados ou quasi temporizados e, portanto,

apresentam distribuições muito afastadas da distribuição exponencial. Além disso, há processos que se apresentam concorrentes com outros processos (exponenciais ou não) em diferentes estados dos sistemas. São os casos de certos processos de reparação, de processos de atraso na propagação das falhas, ou mesmo, de processos de reconfiguração. Sobretudo nestas situações, como se demonstra em [5], a hipótese markoviana introduz erros muito significativos pelo que terá de ser evitada, devendo recorrer-se a metodologias de avaliação alternativas bastante mais complexas [6,7,8,9]. Acontece, no entanto, que a dimensão dos sistemas de produção torna, por vezes, impraticável a construção de um modelo global não markoviano para o sistema em análise, pois o seu tratamento requereria uma capacidade de cálculo incomputável.

Neste artigo apresenta-se uma nova abordagem hierárquica em que os índices de fiabilidade do sistema são obtidos a partir de um modelo simplificado/reduzido do sistema, modelo este construído a partir do modelo reduzido de cada subsistema. Para a simplificação/redução dos modelos, exploram-se diferentes métodos, conforme a natureza do modelo do subsistema em análise. Segundo esta abordagem, o problema da avaliação dos índices de fiabilidade é decomposto em três fases:

1. Análise funcional do sistema, onde são analisadas as dependências entre os equipamentos do sistema de produção induzidas pelo fluxo de materiais e pela partilha de recursos de manutenção, e se procede à decomposição em subsistemas. Esta decomposição tem como linhas orientadoras a obtenção de subsistemas (i) independentes do ponto de vista da fiabilidade, i.e., subsistemas cujo comportamento (falhas, reparações, reconfigurações...) não é condicionado pelo estado dos outros subsistemas; (ii) com afinidades em termos de fiabilidade, nomeadamente, em termos de dependência entre equipamentos e/ou partilha de recursos de reparação ou outros.
2. Obtenção do modelo reduzido de cada subsistema por agregação de estados com idêntico comportamento em termos operacionais (falha ou funcionamento). A redução destes modelos pode ser implementada pela via analítica, por simulação de Monte Carlo ou pelo método de frequência/duração [10]. Fundamentalmente, são os processos de falha e reparação dos subsistemas e os processos que caracterizam os *buffers* entre células de fabrico que determinam qual a via mais adequada a seguir para a redução dos modelos.
3. Construção do modelo global do sistema a partir dos modelos reduzidos dos subsistemas, com base no qual são obtidos os índices pretendidos. De referir que o modelo global do sistema, obtido deste modo pode, também ele, ser objecto de redução/simplificação.

Esta nova abordagem hierárquica será apresentada com base no sistema de produção descrito no anexo 1. Este sistema é constituído, basicamente, por uma linha de montagem (LM), duas células de fabrico (CF1 e CF2), um transportador automático (AGV) e dois *buffers* de componentes/produtos semi-acabados (b_1 e b_2). Estes *buffers* permitem desacoplar o sistema de produção tornando o funcionamento da linha de montagem independente do funcionamento das secções de fabrico de componentes, por períodos de tempo mais ou menos curtos (função da dimensão dos referidos *buffers*). Este papel desempenhado pelos *buffers* verifica-se, de um modo geral, em sistemas de produção flexíveis [11]. A decomposição destes sistemas em subsistemas pode ser implementada de tal modo que cada secção, assim como cada linha de montagem, seja considerado um subsistema e os *buffers*, os “elos” de ligação entre os vários subsistemas. Tem-se, nestes casos, em vez de um modelo complexo do sistema, vários modelos,

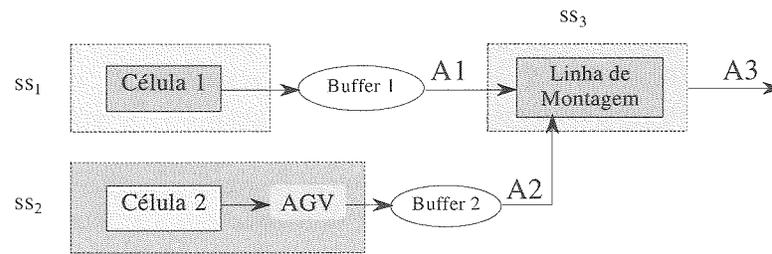


Figura 1: Sistema de produção decomposto em subsistemas.

bastante mais simples, tantos quantos os subsistemas considerados. Saliente-se que as medidas de desempenho a avaliar podem influenciar as possibilidades de decomposição dum sistema. Na Figura 1 mostra-se a decomposição (em subsistemas) do sistema de produção em estudo, tomando em consideração as linhas orientadoras apresentadas acima.

Efectuada a decomposição do sistema, é construído um modelo para cada subsistema. Por agregação de estados, cada um destes modelos é simplificado/reduzido obtendo-se um modelo equivalente com apenas dois estados (funcionamento e falha). Nesta abordagem, a principal dificuldade reside na obtenção das funções de distribuição dos processos dos modelos (simplificado/reduzido) equivalentes. Como se verá, são as distribuições dos processos de falha, de reparação e dos *buffers* a determinar a necessidade de se conhecer a forma das distribuições dos processos do diagrama de estados reduzido ou, apenas, os respectivos valores médios. No ponto seguinte apresentamos em detalhe os procedimentos para a obtenção dos modelos reduzidos equivalentes.

Refira-se, ainda, a respeito dos subsistemas, que estes possuem uma disponibilidade própria - função da fiabilidade dos equipamentos, da qualidade do projecto, do nível de redundância, do plano de manutenção, etc - e uma disponibilidade induzida - função dos mecanismos de tolerância a falhas externos ao subsistema, como é o caso dos *buffers*. Um subsistema diz-se independente se a sua disponibilidade própria não for afectada pelo comportamento dos outros subsistemas.

2 Obtenção do modelo equivalente

Uma vez efectuada a decomposição do sistema, trata-se de obter os modelos equivalentes dos respectivos subsistemas. A metodologia apresentada neste artigo é baseada na ideia de que o modelo de um dado subsistema pode, do ponto de vista da análise do sistema a jusante, ser substituído por um modelo reduzido equivalente com apenas dois estados: um estado de funcionamento e um estado de falha. A questão central na obtenção desse modelo reside na caracterização dos processos de falha e de reparação equivalentes. De seguida, serão introduzidos três vias para a obtenção do modelo reduzido equivalente: a via analítica, aplicável nos casos mais simples, a simulação de Monte Carlo, aplicável nos casos complexos, e o método Frequência/Duração, quando é legítima a aplicação da hipótese markoviana.

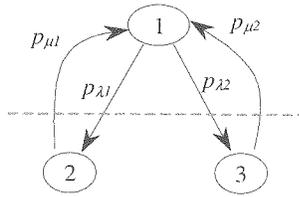


Figura 2: Diagrama de estados (M1, M2).

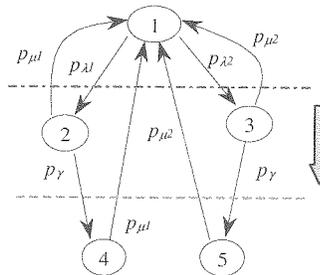


Figura 3: Diagrama de estados (M1, M2, buffer).

2.1 Via analítica

A obtenção de modelos reduzidos equivalentes pela via analítica será ilustrada através do subsistema ss_1 . De acordo com as condições de operacionalidade descritas no anexo 1, os modelos correspondentes a ss_1 e ao conjunto ss_1 / b_1 são os reproduzidos nas Figuras 2 e 3. O processo p_{γ} refere-se, neste caso, ao processo relativo ao *buffer* 1. Conforme já mencionado, o objectivo da simplificação consiste em obter modelos equivalentes onde o comportamento do subsistema seja descrito pelo menor número de estados (idealmente dois estados, falha e funcionamento), tal como representado nas Figuras 4 e 5.

O modelo reduzido equivalente poderá, ou não, ser markoviano, dependendo das distribuições dos processos de reparação $p_{\mu 1}$ e $p_{\mu 2}$, conforme os parágrafos seguintes o demonstram.

Situação 1

Admita-se que os processos de falha e de reparação das máquinas M1 e M2 que constituem ss_1 são exponencialmente distribuídos e que as respectivas taxas (constantes) verificam as relações seguintes:

$$\lambda_1 \neq \lambda_2 \text{ e } \mu_1 = \mu_2$$

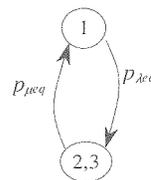


Figura 4: Diagrama de estados reduzido de ss_1 .

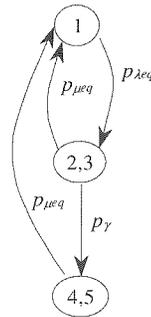


Figura 5: Diagrama de estados reduzido de ss_1 mais b_1 .

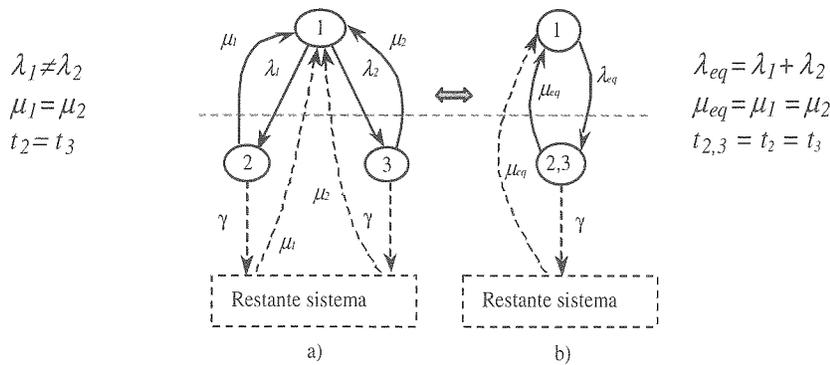


Figura 6: Diagramas de estados de ss_1

Neste caso, embora as frequências e as probabilidades dos estados 2 e 3 (e_2 e e_3) sejam diferentes, os tempos de ocupação destes estados são idênticos (quer em termos de valor médio quer em termos de distribuição), pelo facto dos equipamentos apresentarem processos de reparação idênticos. Agrupando os estados e_2 e e_3 num único estado - $e_{2,3}$, o tempo de ocupação desse estado, também terá distribuição exponencial com valor médio, $t_{2,3} = t_2 = t_3$. Assim, da simplificação do modelo de ss_1 (Figura 6-a), resulta um modelo equivalente (Figura 6-b), cujo processo de falha equivalente se caracteriza por uma taxa constante, $\lambda_{eq} = \lambda_1 + \lambda_2$ e o processo de reparação equivalente, por uma taxa de reparação também constante, $\mu_{eq} = \mu_1 = \mu_2$. Por conseguinte, as funções densidade de probabilidade destes processos são:

$$f p_{\lambda eq} = (\lambda_1 + \lambda_2) e^{-(\lambda_1 + \lambda_2) t} \tag{1}$$

$$f p_{\mu eq} = \mu_1 e^{-\mu_1 t} = \mu_2 e^{-\mu_2 t} \tag{2}$$

Note-se que esta simplificação é válida independentemente da forma da distribuição do processo p_{γ} , pelo qual o subsistema se liga ao restante sistema. No caso do processo p_{γ} ser exponencial, a probabilidade desta “ligação” (probabilidade de transição) é obtida simplesmente por: $\frac{\gamma}{\gamma + \mu}$, com $\mu = \mu_1 = \mu_2$.

Situação 2

Veja-se, agora, a situação em que os processos de ss_1 continuam exponenciais com as seguintes relações entre taxas (constantes): $\lambda_1 = \lambda_2$ e $\mu_1 \neq \mu_2$

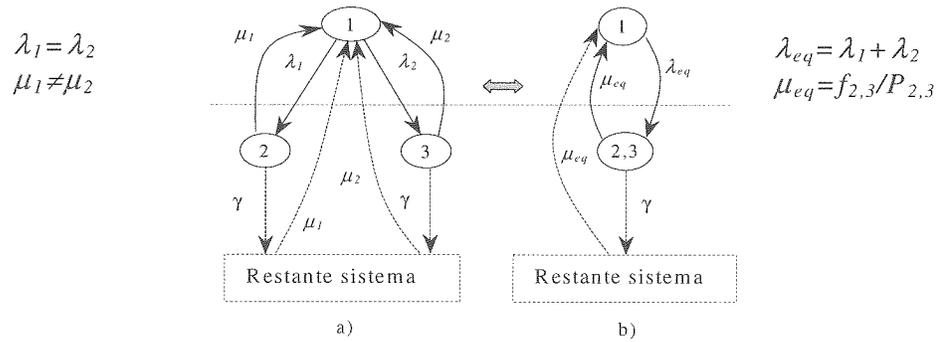


Figura 7: Diagramas de estados equivalentes de ss_1 .

Nestas circunstâncias, o modelo de ss_1 (Figura 7-a) mantém-se com comportamento markoviano. As frequências dos estados e_2 e e_3 são idênticas mas as probabilidades são diferentes. Quanto aos tempos de permanência nesses estados, são exponencialmente distribuídos com valores médios, também, diferentes. Na Figura 7-b apresenta-se o diagrama reduzido do sub-sistema ss_1 que resulta de agruparmos os estados de falha e_2 e e_3 , num único estado, o estado $e_{2,3}$. A frequência e a probabilidade deste estado são aqui designadas por $f_{2,3}$ e $P_{2,3}$, respectivamente.

Apesar das distribuições dos tempos de permanência em e_2 e e_3 serem exponenciais, a distribuição do tempo de permanência em $e_{2,3}$ não é exponencial. Quanto ao seu valor médio, é calculado pela média dos tempos de permanências em e_2 e e_3 . Independentemente da distribuição do processo p_γ , a determinação, apenas, das taxas λ_{eq} e μ_{eq} não permite estabelecer a equivalência entre os modelos a e b da Figura 7. Torna-se necessário, neste caso, conhecer-se a função densidade de probabilidade do processo $p_{\mu_{eq}}$.

Designando $f_{p_{\mu_i}}$ a função densidade de probabilidade do processo de reparação p_{μ_i} , então a função densidade de probabilidade equivalente, $f_{p_{\mu_{eq}}}$ será:

$$f_{p_{\mu_{eq}}} = \frac{\lambda_1}{\lambda_1 + \lambda_2} f_{p_{\mu_1}} + \frac{\lambda_2}{\lambda_1 + \lambda_2} f_{p_{\mu_2}} \quad (3)$$

Para a situação particular em que $f_{p_{\mu_1}} = \mu_1 e^{-\mu_1 t}$ e $f_{p_{\mu_2}} = \mu_2 e^{-\mu_2 t}$, ter-se-á:

$$f_{p_{\mu_{eq}}} = \frac{\lambda_1}{\lambda_1 + \lambda_2} \mu_1 e^{-\mu_1 t} + \frac{\lambda_2}{\lambda_1 + \lambda_2} \mu_2 e^{-\mu_2 t} \quad (4)$$

De realçar que, apesar dos processos de reparação p_{μ_1} e p_{μ_2} serem exponencialmente distribuídos, o processo $p_{\mu_{eq}}$ é não exponencial. Já o processo de falha equivalente $p_{\lambda_{eq}}$, continua exponencialmente distribuído com a seguinte função densidade de probabilidade:

$$f_{p_{\lambda_{eq}}} = (\lambda_1 + \lambda_2) e^{-(\lambda_1 + \lambda_2) t} \quad (5)$$

Incluindo o processo p_γ no diagrama da Figura 7-b, obtém-se o diagrama da Figura 8. Neste caso, dado que o processo de reparação equivalente, $p_{\mu_{eq}}$, é não exponencial, ter-se-á de recorrer à expressão geral (independente das distribuições de p_γ e de $p_{\mu_{eq}}$):

$$t_{2,3} = \int_0^\infty t \cdot \left[f_{p_\gamma}(t) \left(\int_t^\infty f_{p_{\mu_{eq}}}(\tau_1) d\tau_1 \right) + f_{p_{\mu_{eq}}}(t) \left(\int_t^\infty f_{p_\gamma}(\tau_2) d\tau_2 \right) \right] dt \quad (6)$$

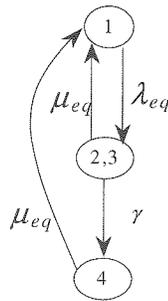


Figura 8: Diagrama de estados reduzido de ss_1 mais b_1 .

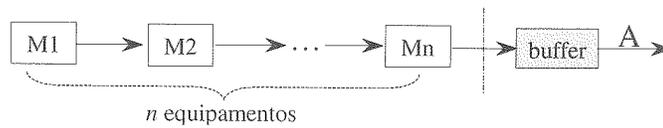


Figura 9: Diagrama funcional de n equipamentos em série e um *buffer* a jusante.

$$P_{2,3 \rightarrow 4} = \int_0^\infty f_{p_\gamma}(t) \left(\int_t^\infty f_{p_{\mu_{eq}}}(\tau) d\tau \right) dt \tag{7}$$

Conforme se mostra em [5], a probabilidade de transição do estado $e_{2,3}$ para e_4 , $P_{2,3 \rightarrow 4}$ (7) é sensível ao tempo médio de permanência em $e_{2,3}$ (6) e à forma da distribuição do processo concorrente, $p_{\mu_{eq}}$. Mostra-se, também, que se os processos concorrentes em $e_{2,3}$ apresentarem constantes de tempo muito diferentes ($\rho \gg 1$ ou $\rho \ll 1$, sendo ρ a razão entre as médias desses processos), os erros introduzidos pela hipótese markoviana são pouco significativos, pelo que nestes casos, esta simplificação pode ser adoptada, com todas as vantagens daí decorrentes.

Situação 3

Considere-se, por fim, uma situação mais abrangente em que um subsistema é constituído por n equipamentos em série e um *buffer* a jusante do $n^{\text{ésimo}}$ equipamento (Figura 9). Todos os equipamentos têm processos de falha exponencialmente distribuídos, p_{λ_i} , e processos de reparação, p_{μ_i} , independentes com funções densidade de probabilidades quaisquer, f_{p_i} .

Na Figura 10-a mostra-se o diagrama de estados deste subsistema. As funções densidade de probabilidade do processo de falha e do processo de reparação equivalentes do diagrama de estados simplificado (Figura 10-b), serão obtidas pelas expressões (8) e (9), respectivamente.

$$f_{p_{\lambda_{eq}}} = \sum_{i=1}^n \lambda_i e^{-\sum_{i=1}^n \lambda_i t} \tag{8}$$

$$f_{p_{\mu_{eq}}} = \frac{1}{\sum_{i=1}^n \lambda_i} \left(\sum_{i=1}^n \lambda_i f_{p_i} \right) \tag{9}$$

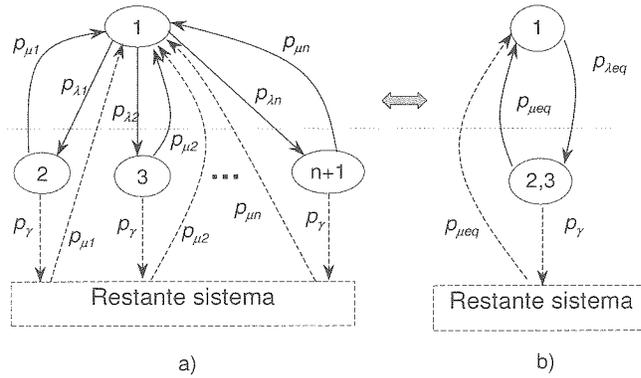


Figura 10: Diagrama de estados de n componentes em série.

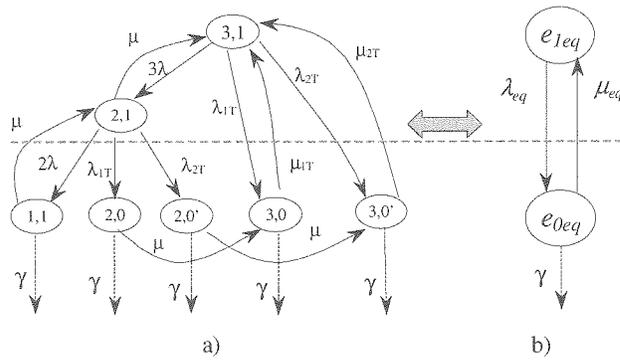


Figura 11: Diagramas de estados de ss_2 .

2.2 Simulação de Monte Carlo

Nem sempre as funções densidade de probabilidade dos processos do diagrama de estados equivalente podem ser obtidas de forma simples como no ponto anterior. Em muitos casos, a obtenção analítica destas funções revela-se difícil. Então, pode recorrer-se à simulação de Monte Carlo para estimar experimentalmente essas distribuições, tal como se discute neste ponto. Considere-se para o efeito o diagrama de estados correspondente ao subsistema ss_2 (Figura11-a) e suponha-se que se pretende determinar as distribuições dos processos $p_{\lambda_{eq}}$ e $p_{\mu_{eq}}$ do modelo equivalente (o subsistema ss_2 apresenta-se descrito em detalhe no anexo 1).

Relativamente à distribuição do processo $p_{\lambda_{eq}}$ é necessário, apenas, conhecer o seu valor médio já que não existe nenhum outro processo concorrente com $p_{\lambda_{eq}}$. Já no caso do processo $p_{\mu_{eq}}$, torna-se necessário conhecer o valor médio e a forma da distribuição, uma vez que este processo é concorrente com o processo p_{γ} (note-se que os processos dos diagramas de estados da Figura 11 são referenciados pelas respectivas taxas médias).

Recorrendo à simulação de Monte Carlo e, tendo em conta a informação que consta na Tabela 2-A (anexo 1), relativa aos processos, calculou-se os tempos de permanência nos estados de falha de ss_2 (Figura 11-a). Com estes tempos construiu-se o histograma de frequências da Figura 12. Este histograma caracteriza o tempo de permanência no estado de falha e_{0eq} (Figura

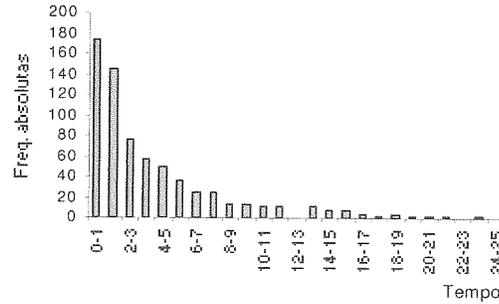


Figura 12: Histograma dos tempos de falha.

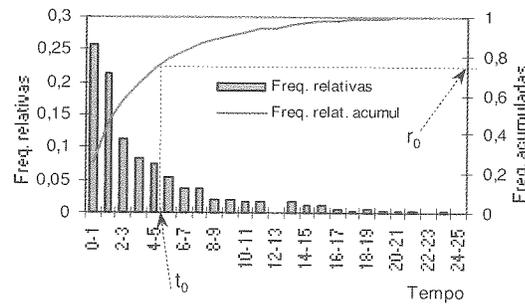


Figura 13: Gráfico de frequências de falhas.

11-b) resultante da agregação todos os estados de falha de ss_2 .

Na Figura 13 representam-se as frequências relativas e as frequências relativas acumuladas dos mesmos tempos de falha. As frequências relativas acumuladas definem a forma da distribuição cumulativa dos tempos de falha $F(t)$ e, por conseguinte, a distribuição cumulativa de $p_{\mu eq}$. Conhecida a distribuição de $p_{\mu eq}$ pode, por simulação, obter-se os tempos de reparação recorrendo ao método de geração de variáveis aleatórias. Este método consiste em gerar uma variável aleatória (uniforme), no domínio $[0,1]$, r_0 e para este valor determinar o correspondente valor, t_0 no contra-domínio de $F(t)$.

$$t_0 = F^{-1}(r_0) \quad (10)$$

A distribuição $F(t)$ será utilizada, mais adiante, na simulação do sistema de produção para gerar o tempo de permanência no estado e_{0eq} . Quanto à frequência de transição, λ_{eq} obtém-se pelo quociente da soma das frequências dos estados de falha ($e_{3,0}$, $e_{3,0'}$, $e_{1,1}$, $e_{2,0}$ e $e_{2,0'}$) pela soma das probabilidades dos estados $e_{3,1}$ e $e_{2,1}$. Por seu lado, o recíproco de λ_{eq} fornece o tempo médio de permanência no estado e_{1eq} .

2.3 Método Frequência/Duração

Finalmente, situações há em que o modelo equivalente pode ser obtido através de um procedimento baseado no método Frequência/Duração [10], conforme se exemplifica de seguida com base no subsistema ss_3 . O respectivo diagrama de estados é apresentado na Figura 14 (o

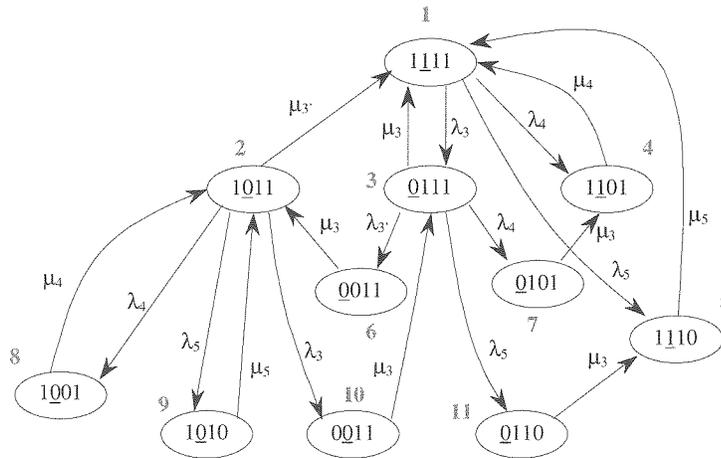


Figura 14: Diagrama de estados de ss_3 .

índice junto a cada estado é uma identificação compacta do respectivo estado). Conforme foi já salientado, a adopção da hipótese markoviana introduz erros significativos quando existem no sistema processos concorrentes não exponenciais. Da análise do diagrama de estados verifica-se que o processo não exponencial $p_{\mu_{3'}}$ (ver distribuição de $p_{\mu_{3'}}$ no anexo 1) está activo no estado (0011), mas aí sem nenhum processo concorrente. No estado (1011), o processo é concorrente com outros processos que apresentam constantes de tempo muito superiores. Sendo assim, ss_3 pode ser tratado como uma cadeia de Markov [12], cuja matriz Q se apresenta a seguir.

$$Q = \begin{bmatrix} -(\lambda_3 + \lambda_4 + \lambda_5) & 0 & \lambda_3 & \lambda_4 & \lambda_5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mu_{31} & -(\lambda_3 + \lambda_4 + \lambda_5 + \mu_{31}) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mu_3 & 0 & -(\lambda_{31} + \lambda_4 + \lambda_5 + \mu_3) & 0 & 0 & 0 & \lambda_{31} & \lambda_4 & 0 & 0 & 0 & \lambda_3 \\ \mu_4 & 0 & 0 & -\mu_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mu_5 & 0 & 0 & 0 & -\mu_5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \mu_3 & 0 & 0 & 0 & -\mu_3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu_3 & 0 & 0 & -\mu_3 & 0 & 0 & 0 & 0 & 0 \\ 0 & \mu_4 & 0 & 0 & 0 & 0 & 0 & -\mu_4 & 0 & 0 & 0 & 0 \\ 0 & \mu_5 & 0 & 0 & 0 & 0 & 0 & 0 & -\mu_5 & 0 & 0 & 0 \\ 0 & 0 & \mu_{31} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\mu_{31} & 0 \\ 0 & 0 & 0 & 0 & \mu_3 & 0 & 0 & 0 & 0 & 0 & 0 & -\mu_3 \end{bmatrix}$$

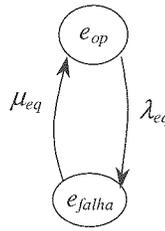
As probabilidades de estado em regime estacionário da Cadeia de Markov podem ser obtidas através da resolução do seguinte sistemas de equações:

$$\begin{cases} P \cdot Q = 0 \\ P \cdot h = 1 \end{cases} \text{ com } h = [1 \ 1 \ 1 \ \dots]^T \quad (11)$$

Conhecidas as probabilidades dos estados de falha e as taxas de transição entre si, pode, pelo método Frequência-Duração, reduzir-se o diagrama da Figura 14 a um diagrama equivalente, com apenas dois estados, um de funcionamento e outro de falha (Figura 15).

A probabilidade do estado operacional e_{op} obtém-se pelo somatório das probabilidades dos estados operacionais do diagrama da Figura 14. O valor resultante é a disponibilidade de ss_3 . O cálculo da probabilidade do estado de falha e_{falha} é efectuado de forma idêntica mas agora com os estados de falha.

Relativamente à frequência dum estado agrupado obtém-se pelo somatório das frequências dos estados individuais que deram origem ao estado agrupado, retirando as transições existentes entre esses mesmos estados. Assim, a frequência do estado de falha, $f_{e_{falha}}$, é determinada

Figura 15: Diagrama de estados reduzido de ss_3 .

do seguinte modo:

$$f_{e_{falha}} = \sum_{i=4}^{11} f_i - (P_7 \times \mu_3 + P_{11} \times \mu_3) \quad (12)$$

Uma vez que o diagrama reduzido tem apenas dois estados, $f_{e_{op}} = f_{e_{falha}}$. Tem-se, ainda, para as taxas de transição,

$$\lambda_{eq} = \frac{f_{e_{op}}}{P_{e_{op}}} \quad (13)$$

$$\mu_{eq} = \frac{f_{e_{falha}}}{P_{e_{falha}}} \quad (14)$$

Estas taxas tornam equivalentes os diagramas de estados das Figuras 14 e 15.

2.4 Modelo global do sistema

Depois de efectuada a decomposição do sistema em subsistemas, e realizada a simplificação do modelo de cada subsistema, surge o terceiro estágio da metodologia que consiste na construção do modelo global do sistema, o qual poderá ser, também, objecto de simplificação. É com base nesse modelo que são obtidos os índices de fiabilidade globais do sistema.

Note-se que foi o trabalho de simplificação efectuada ao nível dos subsistemas que possibilitou a construção do modelo global do sistema de produção pois, sem esse trabalho, seria impraticável construir o modelo global (para o caso em estudo, o número de estados desse modelo seria da ordem de 10^3).

Na Figura 16 apresentam-se os diagramas de estado dos três subsistemas anteriormente analisados. De realçar as alterações introduzidas na designação dos estados e dos processos (aqui representados pelas respectivas taxas médias) dos subsistemas com o objectivo de tornar mais fácil a leitura do modelo global. Com base nestes diagramas e, tendo em conta as condições de operacionalidade do sistema, foi construído o modelo da Figura 17. Cada estado deste modelo é representado por um vector (x_1, x_2, x_3) em que a variável x_i representa os estados do diagrama reduzido do subsistema ss_i . Os índices de fiabilidade do sistema serão obtidos por simulação a partir deste modelo, conforme se mostra na aplicação numérica apresentada no ponto seguinte.

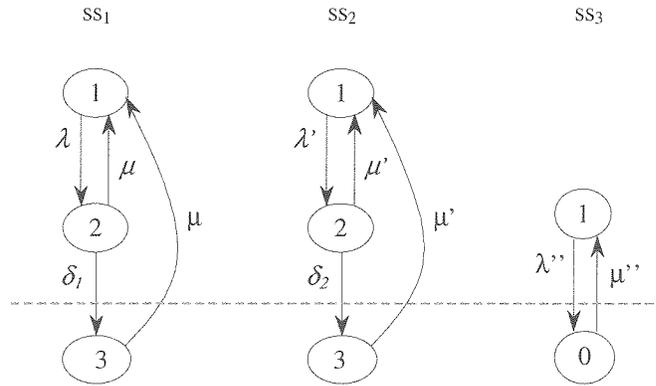


Figura 16: Diagrama de estados reduzidos dos 3 subsistemas.

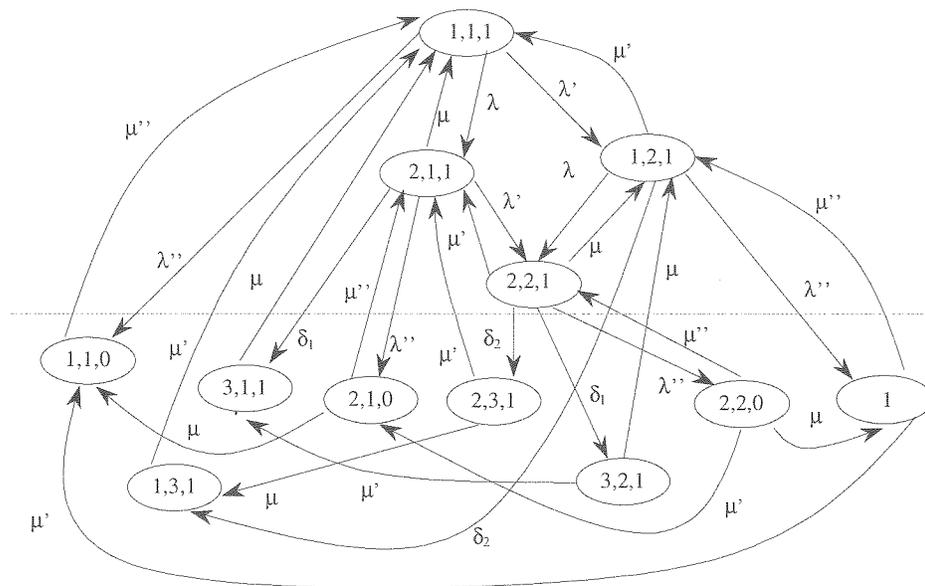


Figura 17: Diagrama de estados do sistema de produção.

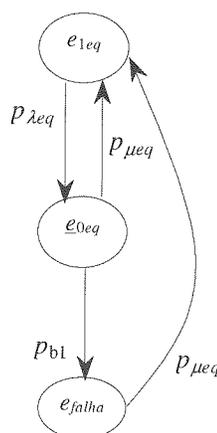


Figura 18: Diagrama de estados reduzido de ss_1 mais b_1 .

3 Aplicação numérica

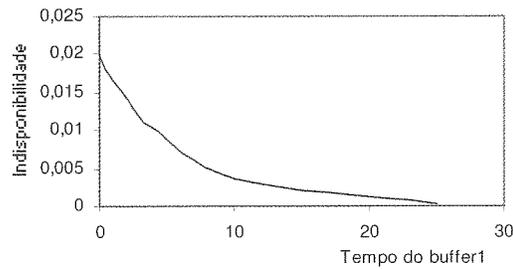
Neste ponto, são calculados os índices de fiabilidade (disponibilidade, frequência e duração dos estados de avaria) do sistema de produção apresentado no anexo 1, com base na abordagem hierárquica e nos modelos desenvolvidos nos pontos anteriores. Estuda-se, ainda, a forma como os *buffers* afectam a disponibilidade do sistema.

3.1 Análise de ss_1

O modelo deste subsistema serviu de caso de estudo em 2.3.1, pelo que nos limitamos, agora, a apresentar os resultados numéricos para os índices de fiabilidade deste subsistema e a analisar o efeito do b_1 sobre esses índices. Os processos que determinam o comportamento deste subsistema, assim como os valores dos parâmetros, constam da Tabela 1A (anexo 1).

Na Figura 18 mostra-se o diagrama de estados simplificado de ss_1 juntamente com o processo p_{b_1} associado a b_1 . Com base neste diagrama de estados, pode avaliar-se a disponibilidade a montante e a jusante de b_1 . Tendo em conta a natureza não-markoviano do modelo e o facto de se dispor da expressão analítica para $p_{\mu_{eq}}$ (4) pode calcular-se os índices de fiabilidade do subsistema recorrendo a (6) e (7) ou, em alternativa, à metodologia DepCim [8]. Na Tabela 1 apresentam-se os valores obtidos para ss_1 considerando que o processo p_{b_1} é determinístico com duração de 1 hora.

Na Figura 19 mostra-se a evolução da indisponibilidade à saída de b_1 em função da sua duração (também determinística). Com o aumento da dimensão do *buffer* verifica-se uma redução da indisponibilidade do sistema, marginalmente menor à medida que a dimensão do *buffer* aumenta.

Figura 19: Indisponibilidade em função de b_1 Tabela 1: Índices de fiabilidade a jusante de b_1

<i>Estados</i>	<i>Prob. Estado</i>	<i>Disponib.</i>	<i>Indisponib.</i>
e_{1eq}	0.978478	0.981207	
e_{0eq}	0.0027337		
e_{falha}	0.0187927		0.0187927

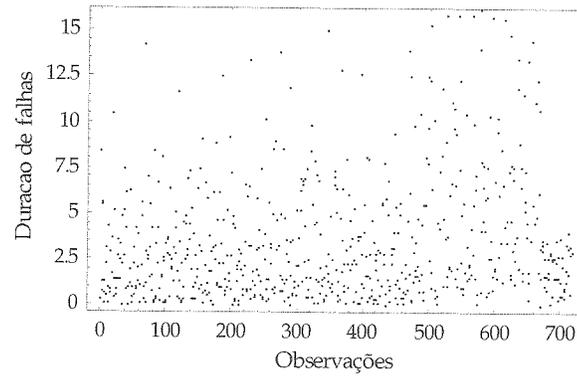
3.2 Análise de ss_2

Este sistema foi analisado na secção 2.3.2. Pretende-se agora obter os índices de fiabilidade de ss_2 e analisar em detalhe o papel desempenhado pelo *buffer* b_2 sobre estes índices, tomando em consideração a informação constante da Tabela 2A (anexo 1) relativa à caracterização dos processos. Na Tabela 2 apresentam-se os resultados obtidos por simulação - valores médios de oito replicações com durações de 5E-05 horas cada. Apesar destes valores resultarem de várias replicações independentes, não devem ser tomados como informação absoluta. É importante dispor de uma medida de confiança sobre estas estimativas calculadas, expressas, por exemplo, na forma de intervalos de confiança [13, 14]. Para tal é necessário dispor de uma medida de dispersão dos resultados (por exemplo da variância), conforme se mostra no Anexo 2.

Na Figura 20 mostra-se um gráfico com os tempos de permanência nos estados falha ($e_{3,0}$, $e_{3,0'}$, $e_{1,1}$, $e_{2,0}$ e $e_{2,0'}$) de ss_2 . Pela análise deste gráfico verifica-se que grande parte dos tempos de falha tem durações inferiores a 3 h. Como tal, a existência de um *buffer* com esta duração seria uma forma eficaz de tolerar as falhas do ss_2 , evitando a propagação das consequências da sua falha a linha de montagem a jusante. Na Figura 21, é bem visível a importância de b_2 , pelo deslocamento da linha de corte que separa os estados operacionais dos estados de falha a jusante de ss_2 .

A não redução do diagrama de estados do subsistema ss_2 (Figura 11-a) obrigaria a tomar o diagrama de estados da Figura 21-a como modelo para o cálculo da disponibilidade à saída do b_2 . No entanto, a simplificação do modelo de ss_2 permite avaliar essa disponibilidade com base no diagrama da Figura 21-b. A partir dos valores da Tabela 2 obtém-se o valor de $0.00743 h^{-1}$ para a taxa de falhas equivalente, λ_{eq} , e o valor de 135 h para o tempo médio de permanência em e_{1eq} .

Neste momento temos condições para avaliar os índices de fiabilidade à saída de b_2 e as

Figura 20: Tempos de falhas de ss_2 (amostra de 689 obs.)Tabela 2: Índices de fiabilidade de ss_2

<i>estado</i>	<i>Prob. estado</i> (P_i)	<i>Frequência</i> ($horas^{-1}$)	<i>Disponibilidade.</i>		<i>Indisponibilidade</i>	
			<i>Média</i>	<i>Var</i>	<i>Média</i>	<i>Var</i>
1 (3,1)	0.931613	0.015856	0.968885	4.6554E-08	0.031115	4.708E-08
2 (2,1)	0.0372721	0.009572				
3 (3,0)	0.01453643	0.004843				
4 (3,0')	0.01556614	0.001944				
5 (1,1)	0.00050285	0.0002476				
6 (2,0)	0.00036674	0.00018277				
7 (2,0')	0.00014772	0.00007352				

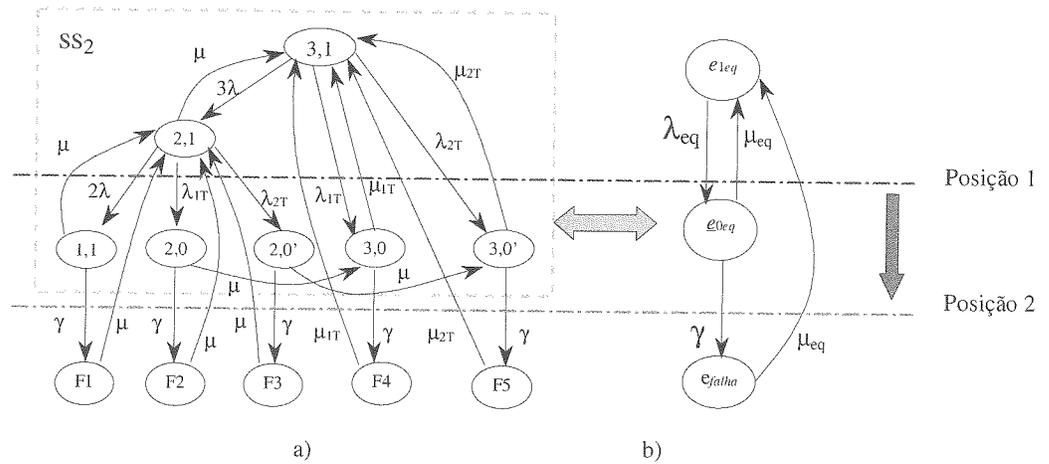


Figura 21: Diagrama de estados reduzido de ss_2 mais b_2 .

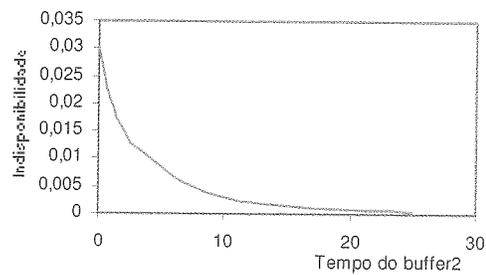


Figura 22: Indisponibilidade em função de b_2

alterações provocadas nestes índices por variações na dimensão deste *buffer*. Para realizar este estudo recorreu-se à Simulação Monte Carlo devido ao facto do processo $p_{\mu eq}$ ser caracterizado por um histograma (Figura 13). Na Tabela 3 apresentam-se os valores obtidos, considerando o processo p_γ (que aqui representa o *buffer* b_2) determinístico com duração de 1 h e um tempo de simulação de 10^7 h.

Na Figura 22 podemos ver a evolução da indisponibilidade à saída do b_2 em função da sua dimensão. Como seria de esperar a indisponibilidade diminui à medida que aumenta a dimensão do *buffer* (caindo para cerca de 1/3 do seu valor quando a duração do *buffer* passa de 1 para 3 horas) verificando-se, também, uma diminuição marginal nos ganhos de disponibilidade.

Tabela 3: Índices de fiabilidade a jusante de b_2

Estados	Prob. Estado	Freq.	Disponib.	Indisponib.
e_{1eq}	0.97326	0.0069333	0.9793	
e_{0eq}	0.00605465	0.0069332		
e_{falha}	0.0207119	0.0051761		0.0207

Tabela 4: Índices de fiabilidade de ss_3

<i>estado</i>	<i>Prob. Estado</i> (P_i)	<i>Frequência</i> ($horas^{-1}$)	<i>Disp.</i>	<i>Indisp.</i>
1 (<u>1</u> 111)	0.984161	0.00295248	0.98996	0.010038
2 (1 <u>0</u> 11)	0.0000345	0.00001160		
3 (0 <u>1</u> 11)	0.0057668	0.000984195		
4 (1 <u>1</u> 01)	0.00494964	0.00098993		
5 (1 <u>1</u> 10)	0.00494964	0.00098993		
6 (0 <u>0</u> 11)	0.00006920	0.0000115335		
7 (0 <u>1</u> 01)	0.00003460	5.76677×10^{-6}		
8 (1 <u>0</u> 01)	1.72486×10^{-7}	3.44971×10^{-8}		
9 (1 <u>0</u> 11)	1.72486×10^{-7}	3.44971×10^{-8}		
10 (0 <u>0</u> 11)	1.03491×10^{-7}	3.44971×10^{-8}		
11 (0 <u>1</u> 10)	0.00003460	5.76677×10^{-6}		

3.3 Análise de ss_3

A resolução do sistema de equações (11) utilizando a informação relativa às taxas de transição que consta na Tabela 3A (anexo 1), permite a obtenção das probabilidades dos estados individuais do subsistema ss_3 (Tabela 4). A probabilidade do estado agrupado e_{op} (Figura 15) obtém-se pelo somatório das probabilidades dos estados operacionais do diagrama da Figura 14, enquanto que o cálculo da probabilidade do estado agrupado de falha e_{falha} é efectuado de forma idêntica mas agora com as probabilidades dos estados individuais de falha. Temos deste modo:

$$P_{e_{op}} = 0.989962 \text{ e } P_{e_{falha}} = 0.0100381$$

Pela expressão (12) calcula-se a frequência do estado agrupado de falha, $f_{e_{falha}} = 0.00199149$. Como o diagrama reduzido tem apenas dois estados, $f_{e_{falha}} = f_{e_{op}}$. Através das expressões (13) e (14) obtém-se os valores das taxas de falha e de reparação equivalentes, $\lambda_{eq} = 0.00201169 \text{ h}^{-1}$ e $\mu_{eq} = 0.198393 \text{ h}^{-1}$.

3.4 Análise global do sistema

Em 2.4 desenvolveu-se o modelo do sistema global, apresentando-se na Figura 17 o respectivo diagrama de estados e, dada a natureza dos processos envolvidos, indicou-se a simulação como sendo a técnica mais adequada para a obtenção dos índices de fiabilidade.

Neste ponto são apresentados os resultados obtidos para os índices de fiabilidade do sistema de produção utilizando as distribuições equivalentes dos processos de falha e de reparação obtidos nos pontos anteriores, as quais estão reproduzidas na Tabela 5. Valerá a pena referir que todos os resultados apresentados quer para os subsistemas quer para o sistema foram obtidos recorrendo a uma ferramenta de cálculo matemático [15].

A Tabela 6 fornece os valores médios dos índices de fiabilidade do sistema e valores das

Tabela 5: Caracterização dos processos do sistema global.

<i>Processo</i>	p_λ	$p_{\lambda'}$	$p_{\lambda''}$	p_μ	$p_{\mu'}$	$p_{\mu''}$	$p_{\delta 1}$	$p_{\delta 2}$
<i>Valor médio (h)</i>	1000/3	135	497	7.333	4.2969	5.04	1	1
<i>Distribuição</i>	<i>Exp</i>	<i>Exp</i>	<i>Exp</i>	<i>Expressão (10)</i>	<i>Hist. Fig. 13</i>	<i>Exp</i>	<i>Dirac</i>	<i>Dirac</i>

Tabela 6: Resultados obtidos de 7 replicações de duração 10^6 horas cada

<i>estado</i>	<i>Prob. Estado</i> (P_i)	<i>Frequência</i> ($horas^{-1}$)	<i>Disponibilidade</i>		<i>Indisponibilidade</i>	
			<i>Média</i>	<i>Variância</i>	<i>Média</i>	<i>Variância</i>
1 (1,1,1)	0,940257857	0,011844429	0,94919443	5,9255E-07	0,05091297	5,71719E-07
2 (2,1,1)	0,006133496	0,007116714				
3 (1,2,1)	0,00278475	0,002868714				
4 (2,2,1)	1,82609E-05	4,08571E-05				
5 (1,1,0)	0,009507881	0,001895286				
6 (3,1,1)	0,023204214	0,005278857				
7 (2,1,0)	1,76736E-05	9,14286E-06				
8 (2,3,1)	6,29537E-05	1,94286E-05				
9 (2,2,0)	0	0				
10 (1,2,0)	1,67399E-05	5,42857E-06				
11 (1,3,1)	0,018050386	0,002854857				
12 (3,2,1)	5,31439E-05	1,57143E-05				

variâncias da disponibilidade e da indisponibilidade obtidos por simulação.

Com base na expressão (2A), apresentada no anexo 2, é possível estimar os intervalos de confiança para os valores médios da disponibilidade e da indisponibilidade do sistema com um grau de confiança α . Para um valor de $\alpha = 5\%$, obtemos os seguintes intervalos de confiança:

$$\text{Disponibilidade} = [0,947310795; 0,951078062];$$

$$\text{Indisponibilidade} = [0,04906274; 0,0527632]$$

Considere-se, agora, que se pretende obter o valor da disponibilidade do sistema de produção com precisão 0.5 %. Retomando a amostra de resultados de simulação já obtidos (7 replicações de 10^6 h cada), estima-se o valor da variância da disponibilidade em 5,9255E-07. Através da expressão (2A), pode calcular-se a precisão dada por esta amostra. O valor obtido ($\delta \approx 0.0007501$) indica uma precisão significativamente superior à pretendida ($\delta \leq 0.005$), razão pela qual não se torna necessário efectuar simulações adicionais.

Refira-se, a este respeito, que a elevada precisão obtida com uma amostra reduzida prende-se com o facto dos valores da amostra apresentarem diferenças muito pequenas. Para isso terá contribuído os longos tempos de simulação adoptados em cada replicação (10^6 h).

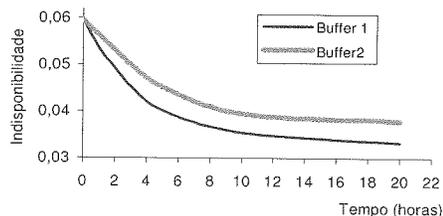


Figura 23: Indisponibilidade do sistema de produção com b_1 e b_2 temporizados

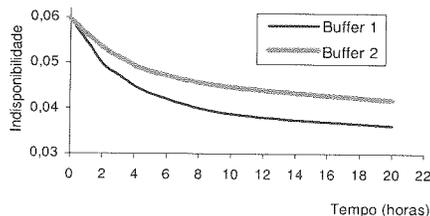


Figura 24: Indisponibilidade do sistema de produção com b_1 e b_2 Exponenciais

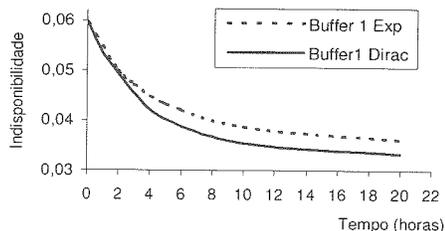


Figura 25: Indisponibilidade do sistema de produção com b_1 Exponencial / b_1 temporizado

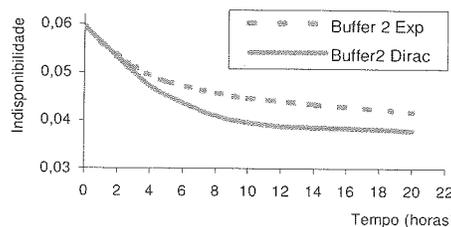


Figura 26: Indisponibilidade do sistema de produção com b_2 Exponencial / b_2 temporizado

3.5 Os buffers e a disponibilidade do sistema

A influência dos *buffers* sobre a disponibilidade de ss_1 e ss_2 foi já analisada. Mais interessante, contudo, é analisar essa influência ao nível do sistema. A Figura 23 mostra a evolução da indisponibilidade do sistema quando (i) se atribuem a b_1 durações temporizadas entre 0 e 20 horas e se considera b_2 inexistente (curva a fino) e (ii) quando se atribuem as mesmas durações a b_2 , considerando b_1 inexistente (curva a grosso). Uma análise idêntica é apresentada na Figura 24, admitindo os *buffers* com tempos de duração exponenciais. Verifica-se através da análise dos gráficos que, para idênticos valores médios, b_1 produz melhores resultados que b_2 em termos de redução da indisponibilidade do sistema.

O gráfico da Figura 25 apresenta a traço contínuo a curva da evolução da indisponibilidade do sistema em função da duração temporizada de b_1 , admitindo a não existências de b_2 e, a traço interrompido, a evolução da mesma grandeza considerando, agora, a duração de b_1 exponencial. A Figura 26 apresenta os resultados de um estudo idêntico para b_2 . Da análise destes gráficos, pode constatar-se que para durações médias dos *buffers* até cerca de 3 horas, a forma das distribuições dos respectivos processos associados (temporizada ou exponencial) não tem efeitos significativos sobre os índices do sistema. No entanto, para durações médias acima deste valor os índices mostram-se visivelmente mais favoráveis com as distribuições temporizadas.

Finalmente, nos gráficos seguintes, mostra-se a evolução da disponibilidade do sistema considerando simultaneamente a existência dos *buffers* 1 e 2, e assumindo durações temporizadas. Com base nestes gráficos e conhecidos os custos de posse dos *buffers*, os engenheiros e gestores industriais podem determinar a dimensão dos *buffers* que conduz a uma otimização dos custos globais do sistema de produção.

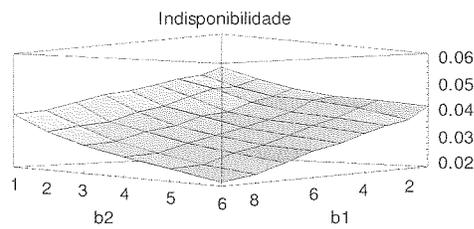


Figura 27: Disponibilidade do sistema (b_1 e b_2 temporizados)

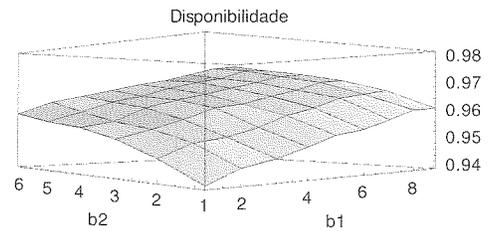


Figura 28: Indisponibilidade do sistema (b_1 e b_2 temporizados)

4 Conclusões

A análise e a avaliação do desempenho dos sistemas de produção são tarefas complexas pelo facto de, normalmente, não ser possível representar num modelo único todos os estados do sistema. A abordagem convencional para ultrapassar esta dificuldade consiste em decompor o sistema em estudo em subsistemas estatisticamente independentes, analisar e calcular os índices de fiabilidade pretendidos ao nível dos subsistemas e, através de uma “fórmula integradora” destes índices, obter os valores para os índices globais ao nível do sistema.

Neste artigo foi apresentada uma abordagem distinta segundo a qual os índices de fiabilidade são obtidos a partir de um modelo global do sistema, construído com base em modelos reduzidos de cada subsistema. Estes modelos podem ser obtidos por via analítica, por simulação de Monte Carlo ou através do método Frequência/Duração. Da análise apresentada ao longo do artigo, constata-se que a via mais adequada a seguir para a obtenção dos modelos dos subsistemas é determinada, fundamentalmente, pelos processos de falha e de reparação e pelos processos que caracterizam o comportamento dos *buffers* existentes entre as células do sistema de produção.

A originalidade desta abordagem reside no facto de possibilitar a obtenção de um modelo do sistema “tratável” do ponto de vista do cálculo numérico, a partir do qual podem ser obtidos os índices de fiabilidade e efectuadas análises de sensibilidade relativamente a parâmetros de dimensionamento do sistema. Em particular, a partir dos modelos simplificados é possível analisar a influência do dimensionamento e localização dos *buffers* intermédios, os quais desempenham um papel importante na disponibilidade do sistema de produção.

À medida que a dimensão destes *buffers* aumenta, assiste-se a uma diminuição da indisponibilidade do sistema, resultando naturalmente em ganhos de produção. Por outro lado, sobem os custos correspondentes aos stocks em curso, e dificulta-se o controlo da produção. Assim, a dimensão óptima dos *buffers* será determinada por aspectos de ordem económica [16]: prejuízos causados pela paragem do sistema de produção, custos de implementação dos *buffers* (posse de inventário, espaço ocupado, dificuldades de circulação no interior das instalações industriais, ...). A expressão geral de custos por unidade de tempo pode escrever-se como:

$$C_T = \sum_{j=1}^k c_{1j} q_j + CI \times \bar{A}(q_1, q_2, \dots, q_k) \quad (15)$$

onde:

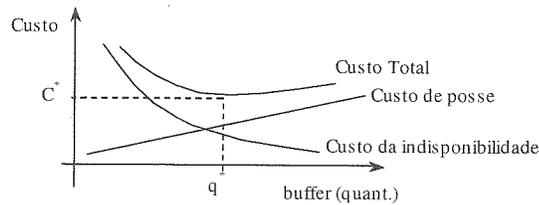


Figura 29: Curvas de custos.

- c_{1j} é o custo unitário de posse do *buffer* j por unidade de tempo;
- q_j é a dimensão (número de unidades) do *buffer* j ;
- CI é o custo da indisponibilidade do sistema por unidade de tempo;
- $\bar{A}(q_1, q_2, \dots, q_k)$ é a indisponibilidade do sistema;
- k é o número de *buffers*.

A Figura 29 representa qualitativamente a evolução dos custos totais. Dada a natureza desta curva, o valor mínimo para o custo total será obtido quando se verificar:

$$\frac{\partial C_T}{\partial q_j} = 0$$

Tal como o caso de estudo apresentado demonstrou, o método introduzido ao longo deste artigo permite determinar a evolução da indisponibilidade do sistema de produção em função da dimensão dos *buffers* e estabelecer quais as dimensões dos *buffers* que asseguram um determinado nível de disponibilidade do sistema, podendo constituir, portanto, um auxiliar precioso no dimensionamento dos sistemas de produção.

5 Referências

- [1] Zakarian, A., Kusiak A., "Modeling Manufacturing Dependability", IEEE Transactions on Robotics and Automation, Vol. 3, No. 2, April, 1997.
- [2] Koenigsberg, E., Mamer, J., "The analysis of production systems", International Journal of Production Research, Vol. 20, No. 1, 1982, pp. 1-16.
- [3] Koichi, T., Shigeru, Y., "Markovian availability modeling for software-intensive systems", International Journal of Quality & Reliability Management, Vol.17, No.2, 2000, pp. 200-212.
- [4] Simeu-Abazi, Z., Daniel, O., Descotes-Genon, B., "Analytical method to evaluate the dependability of manufacturing systems", Reliability Engineering & System Safety, Vol. 55, No. 2, 1997, pp. 125-130.
- [5] Nunes, E., Faria, J. A., Matos, M. A., "A comparative analysis of dependability assessment methodologies" Proceedings of the $\lambda\mu$ 13 ESREL Conference, Lyon, 2002, pp. 228-235.
- [6] Cox, D.R. and Miller, H. D., The Theory of Stochastic Processes, Chapman and Hall, London, UK, 1965.
- [7] Villemeur, A., "Evaluation de la fiabilité, disponibilité et maintenabilité de systèmes réparables: la Méthode de l'Espace des Etats". EDF-DER HT/50/3- juillet 1987.

- [8] Faria, J.A., Matos, M.A., "An analytical methodology for the dependability evaluation of non-Markovian systems with multiple components", *Journal of Reliability Engineering and System Safety*, 74, 2001, pp. 193-210.
- [9] Dubi, A., Gandini, A., Goldfeld, A., Righini, R., and Simonot, H., "Analysis of non-Markovian Systems by a Monte-Carlo Method", *Annales of Nuclear Energy*, Vol.18, n 3, 1991, pp. 125-130.
- [10] Billinton, R., Allan, A., "Reliability evaluation of engineering systems: concepts and techniques", Longman scientific and technical, 1983.
- [11] Mahadevan, B., Narendran, T. T., "Buffer levels and choice of material handling device in Flexible Manufacturing Systems", *European Journal of Operational Research* 69, 1993, pp.166-176.
- [12] Sahner, R. A., Trivedi, K. S. and Puliafito A., *Performance and Reliability Analysis of Computer Systems*, Kluwer Academic Publishers, Norwell, Massachusetts, USA, 1996.
- [13] Rodrigues, A. G., *Simulação*, Universidade do Minho, 1988.
- [14] Averill, M., Law, W., Kelton, D., *Simulation Models and Analysis*, Mc Graw Hill Book Company, 1982.
- [15] Wolfram, S., *Mathematic: A system for doing Mathematics by Computer*, Second Edition, Reading, MA: Addison-Wesley, 1990.
- [16] Malcolm Bradley and Ray Dawson, "The cost of unreliability : a case study", *Journal of Quality in Maintenance Engineering*, Vol. 4, No. 3, 1998, pp. 212-218.

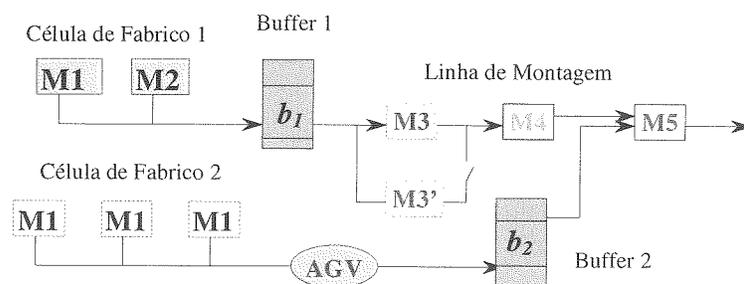


Figura 1A: Sistema de produção.

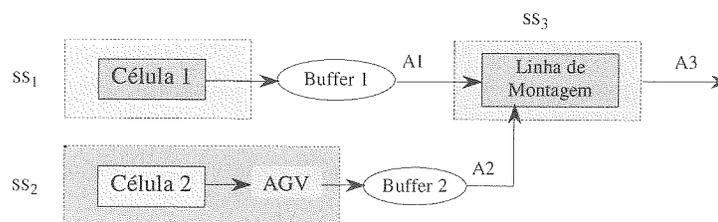


Figura 2A: Decomposição do sistema de produção em subsistemas.

Anexo 1 - Apresentação do sistema de produção em estudo

Muitos sistemas de produção são constituídos por uma ou várias linhas de montagens (LM) de produto final, abastecidas por transportadores de componentes (manuais ou automáticos), produzidos em secções de fabrico de componentes. Frequentemente, alguns dos componentes são adquiridos a outras empresas. Entre as secções de produção de componentes e as linhas de montagens existem, frequentemente, stocks de componentes e stocks de produtos semiacabados – *buffers*. A especificidade e diversidade dos equipamentos requerem equipas de manutenção especializadas afectas às secções e às linhas de montagem.

Na Figura 1A mostra-se o esquema de um sistema de produção com características como as acima referidas, constituído por uma linha de montagem (LM), duas células de fabrico (CF1 e CF2), um transportador automático (AGV) e dois *buffers* (b_1 e b_2). É com base neste sistema que será apresentada a abordagem hierárquica proposta neste artigo.

Na Figura 2A apresenta-se a decomposição deste sistema de produção em três subsistemas, de acordo com as linhas orientadoras introduzidas no ponto 2. Os três subsistemas são descritos nos parágrafos seguintes.

Subsistema ss_1

O esquema apresentado na Figura 3 representa o modelo funcional do ss_1 (constituído por duas máquinas, M1 e M2) juntamente com o *buffer* 1. Quando uma das máquinas falha a outra pára imediatamente e só entra em funcionamento quando reparada a máquina em falha. Deste modo, a missão do ss_1 em conjunto com o *buffer* 1 cumpre-se em qualquer das seguintes situações:

1. M1 e M2 operacionais;
2. M1(2) falha, e o *buffer* tolera a falha de M1(2) durante a sua reparação.

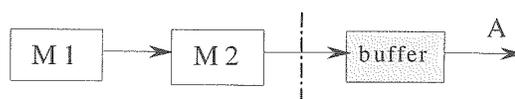


Figura 3A: Diagrama funcional de célula de fabrico.

Tabela 1A: Descrição dos processos de ss_1

<i>Processo</i>	<i>Descrição</i>	<i>f.d.p.</i>	<i>Duração média (h)</i>
$p_{\lambda 1}$	Falha de M1	Exp.	$m_{\lambda 1} = 500$
$p_{\lambda 2}$	Falha de M2	Exp.	$m_{\lambda 2} = 1000$
$p_{\mu 1}$	Rep. de M1	Exp.	$m_{\mu 1} = 6$
$p_{\mu 2}$	Rep. de M2	Exp.	$m_{\mu 2} = 10$

As condições de operacionalidade referidas permitem construir os diagramas de estados apresentados nas Figuras 2 e 3, e aqui reproduzidos nas Figuras 4A e 5A. O estado 1 representa o estado de funcionamento das duas máquinas, M1 e M2; o estado 2 representa o estado em que a máquina M1 está em falha; o estado 3 é idêntico ao estado 2 estando agora M2 em falha em vez de M1. No caso da não existência do *buffer*, do ponto de vista dos subsistemas a jusante, os estados 2 e 3 são estados de falha. Com a inclusão do *buffer* estes estados passam a estados de funcionamento surgindo os estados 4 e 5 como estados de falha, pois só então os subsistemas a jusante serão afectados pelas avarias de M1 ou M2. Naturalmente, o *buffer* tem uma dimensão limitada e, por conseguinte, tolera uma falha a montante durante um período de tempo determinado pela sua dimensão.

Subsistema ss_2

A célula de fabrico 2 (CF2) constitui o subsistema ss_2 . Fazem parte desta célula três máquinas semelhantes e um AGV. Duas das três máquinas são suficientes para assegurar o normal funcionamento da linha de montagem. As máquinas M1 apresentam taxas de falhas λ_1 , λ_2 e λ_3 idênticas, e taxas de reparação μ_1 , μ_2 e μ_3 também idênticas. O transportador automático tem dois modos de falha distintos e independentes, com taxas de falha λ_{1T} e λ_{2T} . As respectivas reparações também apresentam taxas diferentes, μ_{1T} e μ_{2T} . Dispõe-se de uma equipa de manutenção para a célula de fabrico, pelo que, em cada momento, apenas um equipamento se pode encontrar em reparação.

Todas as distribuições dos processos que modelam o comportamento da CF2 são exponencialmente distribuídos ou assumidos como tal, excepto os processos de reparação das máquinas, que são temporizados. O processo γ modela o comportamento do *buffer* existente entre a CF2 e a linha de montagem.

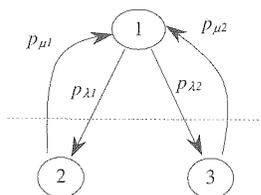


Figura 4A: Diagrama de estados (M1, M2).

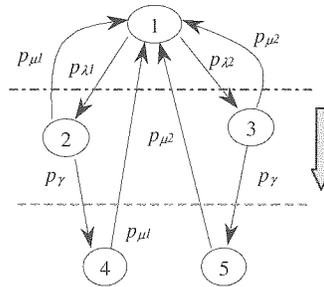


Figura 5A: Diagrama de estados (M1, M2, buffer).

Tabela 2A: Descrição dos processos de ss_2

<i>Processo</i>	<i>Descrição</i>	<i>f.d.p.</i>	<i>Duração (h)</i>
p_{λ}	Falha de uma máquina	Exp.	$m_{\lambda} = 300$
$p_{\lambda 1T}$	Falha de AGV modo1	Exp.	$m_{\lambda 1T} = 200$
$p_{\lambda 2T}$	Falha de AGV modo2	Exp.	$m_{\lambda 2T} = 500$
p_{μ}	Rep. de uma máquina	Dirac	$m_{\delta} = 4$
$p_{\mu 1T}$	Rep. do AGV modo1	Exp.	$m_{\mu 1T} = 3$
$p_{\mu 2T}$	Rep. do AGV modo2	Exp.	$m_{\mu 2T} = 8$

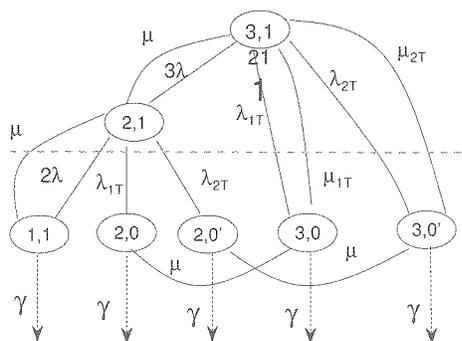


Figura 6A: Diagramas de estados da CF2.

Tabela 3A: Caracterização dos processos de ss_3

Processo	Descrição	f.d.p.	Duração (h)
p_{λ_3}	Falha de máq. M3	Exp	$m_{\lambda_3} = 1000$
$p_{\lambda_{3'}}$	Falha de máq. M3'	Exp	$m_{\lambda_{3'}} = 500$
p_{λ_4}	Falha de máq. M4	Exp	$m_{\lambda_4} = 1000$
p_{λ_5}	Falha de máq. M5	Exp	$m_{\lambda_5} = 1000$
p_{μ_3}	Rep. da máq. M3	Exp	$m_{\mu_3} = 6$
$p_{\mu_{3'}}$	Rep. da máq. M3'	LogNormal	$m_{\mu_{3'}} = 3$
p_{μ_4}	Rep. da máq. M4	Exp	$m_{\mu_4} = 5$
p_{μ_5}	Rep. da máq. M5	Exp	$m_{\mu_5} = 5$

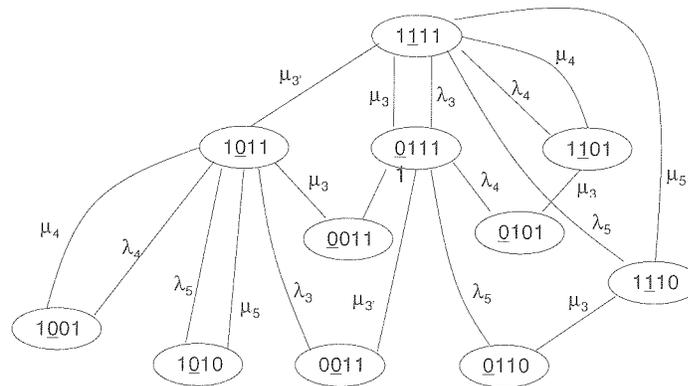


Figura 7A: Diagrama de estado de ss_3 .

Subsistema ss_3

Este subsistema é formado pela linha de montagem (LM) propriamente dita funcionando em fluxo unitário. A máquina M3', constitui uma redundância passiva à máquina M3, tecnicamente mais evoluída e com menores custos de manutenção. A paragem das máquinas M4 ou M5 provoca a paragem da linha de montagem. Dispõe-se apenas de uma equipa de manutenção para a LM, não havendo possibilidade de reparação de mais que uma máquina em simultâneo. Admite-se que o sistema sensor comutador tem fiabilidade unitária e que as máquinas M3, M4 e M5 têm taxas de falha constantes λ_3 , λ_4 e λ_5 , respectivamente. As reparações são também exponenciais com taxas μ_3 , μ_4 e μ_5 , respectivamente. Considera-se que qualquer máquina da linha de montagem tem uma probabilidade de falha nula quando parada. A máquina M3' apresenta um processo de reparação com distribuição logNormal ($m_{\mu_{3'}}$, σ^2) e o respectivo processo de falha é exponencial com taxa $\lambda_{3'}$.

Ao contrário do que acontece com os processo de reparação de todas as outras máquinas de LM, o processo de reparação de $M_{3'}$ poderá ser interrompido e reinicializado mais tarde se a operacionalidade do sistema depender da intervenção da manutenção nas máquinas M4 ou M5.

Na Figura 20 mostra-se o diagrama de estados da LM. Cada estado é representado por um vector $[x_3, x_{3'}, x_4, x_5]$ em que as variáveis x_i representam o estado da máquina M_i . A informação que consta da Tabela 3A permite caracterizar todos os processos do subsistema ss_3 , em termos de distribuições e tempos médios.

Anexo 2 – Medidas de confiança sobre resultados obtidos por simulação

Intervalos de confiança

Sendo os resultados obtidos por simulação, valores experimentais, não devem ser tomados como informação absoluta, mesmo que sejam obtidos através de médias calculadas sobre várias replicações independentes [13]. É importante dispor de uma medida de confiança sobre as estimativas calculadas, expressas, por exemplo, na forma de intervalos de confiança.

O intervalo de confiança para uma estimativa da média, será, para um grau de confiança igual a α dado por:

$$\bar{X} \pm t_{N-1}(\alpha/2) \cdot \sqrt{\frac{\sum_{i=1}^N (X_i - \bar{X})^2}{(N-1)}} \quad (1A)$$

onde $X_1, X_2, X_3, \dots, X_n$ são valores obtidos de várias execuções do modelo de simulação, \bar{X} a média destes valores, N o número de execuções do modelo de simulação (tamanho da amostra) e $t_{N-1}(\alpha/2)$ o valor da distribuição de Student com $N - 1$ graus de liberdade.

Dimensão da amostra para uma precisão pretendida

Por vezes não se pretende obter um intervalo de confiança calculado com base numa amostra extraída de várias replicações independentes (*runs*), mas sim, determinar que dimensão deve ter a amostra de modo a que se possa obter uma estimativa com uma precisão relativa pré-especificada [14]. A partir de um número reduzido de replicações n , pode calcular-se uma estimativa da variância para uma determinada grandeza. Se admitirmos que essa estimativa não se altera significativamente com o número de replicações, podemos determinar a dimensão necessária para a amostra, pela expressão:

$$\frac{t_{r-1, 1-\alpha/2} \sqrt{\frac{S^2}{r}}}{|\bar{X}|} \leq \delta \quad (2A)$$

sendo S^2 a estimativa da variância calculada com base na amostra inicial de dimensão n , δ a precisão relativa pretendida e r a dimensão que a amostra deverá ter para se obter essa precisão com um nível de confiança α . Uma vez determinado o valor de r , deverão realizar-se $(r - n)$ simulações adicionais. A utilização de uma amostra reduzida para estimar S^2 pode resultar numa sobre-estimativa. Um procedimento para reduzir esse efeito consiste em rever a estimativa S^2 após a realização de cada simulação adicional, até se obter a precisão pretendida.

Avaliação do Desempenho de Alunos Considerando Classificações Obtidas e Opiniões dos Docentes

Maria Helena Campos Soares de Mello *
Heitor Luiz Murat de Meirelles Quintella *
João Carlos Correia Baptista Soares de Mello *

* Departamento de Engenharia de Produção
Escola de Engenharia - Universidade Federal Fluminense
Brasil

{mhelenamello@netscape.net, hquintel@unisys.com.br, jcsmello@sapo.pt}

Abstract

This paper intends to evaluate the academic evolution of an engineering students group who entered Universidade Federal Fluminense (UFF) - Brazil in 1997. Both scholarship bulletins, and lecturers opinions are considered in order to understand what “a good student” is. The lectures opinions are evaluated with the multicriteria elementary methods of Borda and Condorcet. A statistical analysis is performed with the scholarship bulletins from 1997 until 2000.

Resumo

Este artigo avalia a situação acadêmica dos alunos dos cursos de Engenharia que ingressaram na Universidade Federal Fluminense (UFF) - Brasil em 1997 pelo acompanhamento do seu rendimento escolar, tentando estabelecer relações de causa e efeito entre suas classificações nas diferentes etapas do exame de admissão, chamado “concurso vestibular”. A amostra é constituída por cerca de cento e cinquenta alunos, com dados referentes às suas classificações em todas as disciplinas frequentadas até ao segundo semestre de 2000. É feita também uma análise da opinião docente sobre as condições que permitem formar um bom aluno e sobre o que se considera um bom aluno. Os resultados dos questionários aplicados são analisados por métodos multidecisor elementares (métodos de Borda e Condorcet) e comparados com a análise anterior.

Keywords: education , student evaluation, multicriteria analysis.

Title: Teachers perception and actual students ranking in performance appraisal.

1 Introdução

O rendimento escolar dos alunos dos cursos de Engenharia tem sido questionado e comparado com o rendimento de alunos de alguns anos passados. Existem mesmo os que afirmam que os alunos de hoje em dia são piores do que os de antigamente, outros rotulam os alunos de "bons" ou "maus" e agem de maneira diferenciada com uns e outros. Entretanto, nem sempre a percepção corresponde à realidade dos números. Este artigo analisa a situação vivida de facto, com alunos que ingressaram no curso de Engenharia da Universidade Federal Fluminense, em 1997. É verificado o seu rendimento no exame de admissão à Universidade (chamado "Concurso Vestibular") e ao longo do curso. Os resultados obtidos são confrontados com o que é esperado que seja um "bom" aluno, na opinião dos professores. Foram utilizadas técnicas estatísticas (correlação linear e análise de agrupamentos) e métodos de análise multidecisor (Métodos de Borda e Condorcet), considerados elementares (Dias et al, 1996), para análise das informações.

2 Processo de admissão na Universidade

No Brasil, para a maioria das Universidades a principal, forma de acesso é o chamado "Concurso Vestibular". Este concurso, realizado por cada instituição, ou por grupos de instituições, pretende ser utilizado primordialmente para dois grandes objectivos: ordenar os estudantes, com o objectivo de preencher as vagas nos cursos oferecidos e avaliar o nível de conhecimentos que eles obtiveram no ensino secundário (que, oficialmente, é denominado ensino médio), bem como sua aptidão para frequentarem um determinado curso superior. Particularmente a Universidade Federal Fluminense (UFF), objecto deste estudo, faz a sua selecção em duas etapas (COSEAC, 2001, Soares de Mello et al 2002) onde a primeira é constituída por provas de "escolha múltipla- em que o candidato deve escolher entre cinco alternativas aquela que contém a resposta correcta à questão apresentada, (com todas as matérias estudadas no nível médio) e a segunda fase é constituída por provas "discursivas- em que o candidato deve responder por escrito às questões apresentadas (uma redacção e uma ou duas matérias específicas para o curso pretendido) visando avaliar os conhecimentos mínimos obtidos no nível médio de estudos. A primeira fase entra como critério de eliminação e de ordenação e seu objectivo é avaliar se foi atingido um nível mínimo de conhecimentos gerais. A segunda fase é apenas critério de ordenação e visa verificar a existência de conhecimentos necessários para seguir o curso escolhido. A primeira fase é constituída por provas de Língua Portuguesa e Literatura Brasileira, Língua Estrangeira (Inglês, Francês ou Espanhol), Química, Física, Matemática, Biologia, Geografia e História. Para os cursos de Engenharia, a segunda fase é composta por provas de Redacção, Matemática e Física. A ordenação dos candidatos não eliminados é feita recorrendo a uma soma ponderada, em que os pesos são atribuídos de tal forma que o conjunto de cada fase tenha igual peso; dentro de cada fase todas as provas tenham igual "importância"; e a pontuação final esteja contida no intervalo]0,100]. Desta forma, as provas de Matemática e Física da segunda fase entrarão na pontuação final com peso inferior a 17%, cada uma.

Ressalte-se que não há classificação mínima para aprovação. Através da já referida soma ponderada, é feita uma ordenação dos estudantes, usada para dividi-los em 3 classes: os que ingressam no primeiro semestre do ano lectivo, os que ingressam no segundo semestre do

ano lectivo e os que não ingressam. Assim, é possível o ingresso de alunos com classificações extremamente baixas, principalmente em cursos de baixa procura.

3 Progressão dos alunos no curso de Engenharia da UFF

Após o ingresso na Universidade, o aluno frequenta, no seu primeiro semestre de estudos, seis disciplinas. Três delas são consideradas "difíceis" (Cálculo I, Física I e Álgebra Linear) e as outras três são consideradas "fáceis" (Introdução à Engenharia, Introdução à Informática e Introdução à Geometria Descritiva).

A "qualidade" dos alunos é medida por intermédio de índices de desempenho. Um índice, por disciplina, é a média final obtida pelo aluno. Na UFF, é adoptada a classificação mínima 6,0 (numa escala entre 0,0 e 10,0) para aprovação. No caso de o aluno ter obtido classificação entre 4,0 e 6,0, pode submeter-se a um exame suplementar, chamado VS, no qual deve ter classificação maior ou igual a 6,0 para ser aprovado. Caso contrário (média menor do que 4,0 e classificação menor do 6,0 na VS), o aluno é reprovado. Outro índice de desempenho, que leva em conta todas as disciplinas cursadas é o Coeficiente de Rendimento (CR).

O CR é obtido pela média das classificações, ponderadas com o número de créditos (valor relacionado com o número de horas de aula semanais) das disciplinas. Assim, os alunos são rotulados pelo seu CR e ordenados por ele, principalmente na escolha dos horários mais concorridos.

Após completar as disciplinas dos três primeiros semestres - chamado de Ciclo Básico, o aluno que ingressou até 1998 no curso de Engenharia da UFF, podia fazer a sua opção por uma das seis especializações oferecidas nesta Universidade (Produção, Mecânica, Civil, Agrícola, Eléctrica e Telecomunicações).

4 Delimitações dos estudos

Este trabalho limita-se a analisar informações referentes aos alunos que ingressaram no primeiro semestre do curso de Engenharia (com posterior opção pela habilitação profissional) da Universidade Federal Fluminense, em 1997. Este foi o penúltimo ano em que a forma de ingresso permitiu o agrupamento de dados. A partir do vestibular de 1999, a UFF decidiu que o ingresso já seria feito para a especialização de Engenharia, deixando de haver a possibilidade de opção no final do Ciclo Básico (Soares de Mello & Soares e Mello, 2000).

Deste universo foram excluídos os alunos que não iniciaram seus estudos no semestre subsequente ao do ingresso e aqueles que já haviam frequentado disciplinas (e obtido aprovação) noutra curso superior, para uniformizar o mais possível o universo estudado.

Assim, do total de 360 alunos que ingressaram no Ciclo Básico dos cursos de Engenharia da UFF em 1997, serão utilizadas apenas informações relativas aos 159 alunos que ingressaram no primeiro semestre lectivo e aos 135 do segundo semestre de 1997.

As informações sobre a percepção dos professores foram obtidas por meio de um questionário com perguntas abertas, em que se pediu que o professor ordenasse as alternativas por

ordem decrescente de concordância, ou seja, em primeiro lugar deveria vir a alternativa com a qual estivesse mais de acordo e por último aquela com a qual concordasse menos.

Este questionário pretende descrever a opinião dos professores, sem conhecimento prévio dos alunos, sobre o conceito de "bom" aluno de Engenharia.

5 Avaliação dos questionários

Foram colocadas três questões aos professores para que, de entre os estereótipos apresentados, fosse caracterizada a sua noção de um bom aluno de Engenharia (Soares de Mello, 2002).

Em primeiro lugar, foi questionado aos professores se consideravam que os seus "melhores" alunos tiveram: a) as melhores *classificações* no vestibular; b) as melhores *classificações na segunda etapa* do vestibular; c) as melhores *classificações nas disciplinas relacionadas* com a disciplina leccionada pelo professor e d) classificações que *não interferem* no rendimento escolar posterior destes alunos.

Em segundo lugar, foi questionado aos professores se o melhor aluno seria aquele que: a) obteve aprovação em todas as disciplinas do primeiro semestre, tendo feito VS em todas elas e obtido um CR baixo; b) obteve aprovação apenas nas três disciplinas consideradas "fáceis" e obtido um CR elevado e c) obteve aprovação apenas nas três disciplinas consideradas "difíceis", com um CR próximo de 6,0.

Em terceiro lugar, foi questionado aos professores se o melhor aluno seria aquele que: a) resolvesse todos os exercícios e reproduzisse o que o professor apresentou em sala; b) se interessasse em pesquisar assuntos relacionados a matéria leccionada e c) fosse capaz de construir o seu conhecimento, compreendendo rapidamente o encaminhamento dado pelo professor.

Para analisar estas respostas, como se trata de ordenações feitas por múltiplos decisores, podem ser utilizados os métodos de Borda e de Condorcet.

Estes métodos são descritos na literatura como sendo responsáveis pelas origens, respectivamente da chamada "Escola Americana" e "Escola Francesa" de técnicas multicritério. Ambos os métodos foram desenvolvidos no século XVIII para resolver problemas referentes a escolhas sociais. (Barba-Romero e Pomerol, 1997).

O Método de Borda baseia-se na pontuação das alternativas, de acordo com a ordenação fornecida pelo decisor, como num campeonato. A alternativa mais preferida ganha um ponto, a segunda melhor ganha dois pontos e assim sucessivamente. No final, os pontos atribuídos pelos decisores a cada alternativa são somados e a alternativa que tiver obtido a menor pontuação será a escolhida. Exemplos correntes de aplicação de variações do método de Borda são encontrados com frequência em competições desportivas, como o campeonato mundial de fórmula 1

O Método de Condorcet baseia-se em relações de preferências. Quando cada decisor ordena as alternativas por ordem de preferência, o analista verifica, em cada par de alternativas, qual delas foi preferida pela maioria dos decisores. Neste caso, diz-se que esta alternativa é preferível em relação à outra. Podem ser traçados grafos representativos destas relações de preferência, em que o arco (u, v) pertence ao grafo se, e só se, o número de decisores que preferiram u a

v é maior ou igual dos que preferiram v a u . Estes são análogos aos que se obteriam com o método ELECTRE I (Roy e Bouyssou, 1993), desde que todos os critérios tivessem o mesmo peso e não houvesse veto ou discordância. Através da representação da relação de preferência por um grafo, a determinação de alternativas dominantes e dominadas (quando existem), fica bastante facilitada. Quanto existe uma, e só uma, alternativa dominante, ela é a escolhida.

O método de Borda apresenta a desvantagem de não ser indiferente às alternativas irrelevantes. Ou seja, a retirada de uma alternativa pode levar a modificações na ordenação relativa de outras alternativas. O método de Condorcet, considerado mais justo, tem a grande desvantagem de conduzir a situações de intransitividade, levando ao célebre "paradoxo de Condorcet". Este ocorre quando A é preferível a B , B é preferível a C e C é preferível a A . Isto significa que o método de Condorcet nem sempre induz uma pré-ordem no conjunto das alternativas. No entanto, em situações onde não ocorram ciclos de intransitividade este método deve ser preferido ao método de Borda.

Em particular, em resposta à primeira pergunta, obteve-se o grafo da Figura 1, considerando as alternativas a , b , c e d apresentadas.

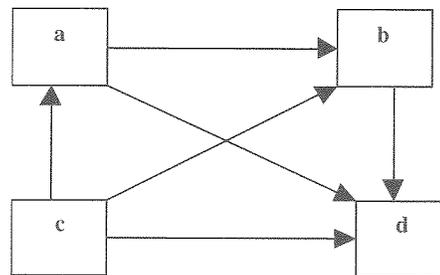


Figura 1: Grafo representativo das respostas à primeira pergunta do questionário.

Verifica-se que a alternativa a é preferível às alternativas b e d , a alternativa b é preferível à alternativa d e a alternativa c é preferível a todas as demais. Desta forma, a alternativa dominante é a alternativa c e a dominada é a alternativa d . Não foram formados ciclos de intransitividade, portanto, pode-se dizer que, dentre as alternativas apresentadas, a preferida é a alternativa c , ou seja, os professores entendem que os melhores alunos devem ser aqueles que obtiveram as melhores classificações nas provas do vestibular com matérias relacionadas com as disciplinas lecionadas por eles. Exemplificando, para um professor de Cálculo, os seus melhores alunos devem ser aqueles com as melhores classificações em Matemática no Vestibular, independentemente de ser a prova de "múltipla escolha" ou "discursiva". São os alunos que possuem "uma boa base", e espera-se que venham a ter pouca dificuldade em compreenderem os assuntos novos que lhes serão apresentados, resultando em um menor esforço para o professor.

Em resposta à segunda pergunta, obteve-se o grafo da Figura 2. Observa-se que a alternativa a é a preferida pelos professores, ou seja, mais uma vez os professores demonstraram sua preferência pelos alunos que são regulares, não importando as suas classificações: ser aprovado em todas as disciplinas, facilitando sua progressão no curso, é preferível a ter classificações altas e ser reprovado em algumas, apesar de não comprometer sua progressão no curso.

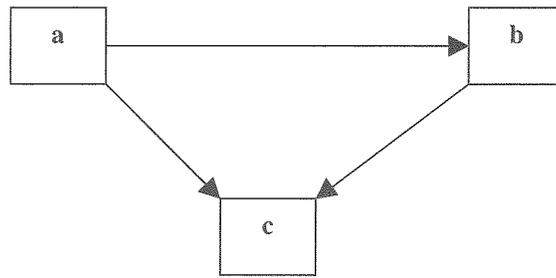


Figura 2: Grafo representativo das respostas à segunda pergunta do questionário.

Em relação à terceira questão, obteve-se o grafo apresentado na Figura 3. Nota-se que a alternativa *a* é a preferida em relação a todas as outras, ou seja, os professores preferem os alunos disciplinados, no sentido de fazerem simplesmente as suas tarefas.

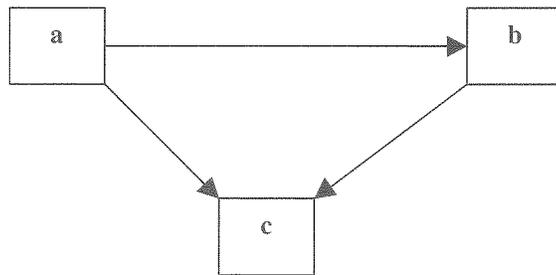


Figura 3: Grafo representativo das respostas à terceira pergunta do questionário.

Ou seja, pode-se imediatamente compreender a interpretação e caracterização que os professores dão aos "melhores" alunos. Estes são os que tiveram as melhores bases para acompanhar as suas disciplinas; obtêm aprovação em todas as disciplinas, independentemente de ter obtido boas classificações e são capazes de reproduzir os conceitos ensinados na sala de aula.

Para terminar este item, ressalte-se que não ocorreram ciclos de intransitividade em nenhuma das análises, o que justifica o facto de não ter sido usado o método de Borda.

6 Avaliação das correlações

As opiniões dos professores foram então confrontadas com os dados das classificações dos alunos, no vestibular e em algumas das disciplinas do curso de Engenharia. Para tal foram calculadas correlações lineares entre vários pares de classificações, previamente normalizados de forma que as classificações variassem sempre entre 0 e 10. A Tabela 1 apresenta os valores das correlações investigadas.

Tabela 1: Correlações entre classificações no vestibular e na universidade.

X - Variável obtida do vestibular	Y - Variável obtida no primeiro semestre lectivo	Alunos considerados	Correlação entre X e Y
Classificação final	Coefficiente de rendimento	Todos	0,39
Classificação final	Coefficiente de rendimento	Progressão curricular correcta	0,43
Classificação final	Coefficiente de rendimento	Aprovação total no 1 ^o semestre	0,43
Classificação na 2 ^a fase	Coefficiente de rendimento	Todos	0,34
Classificação na 2 ^a fase	Coefficiente de rendimento	Progressão curricular correcta	0,48
Classificação na 2 ^a fase	Coefficiente de rendimento	Aprovação total no 1 ^o semestre	0,45
Classificação em Matemática	Classificação em Cálculo I	Todos	0,35
Classificação em Matemática	Classificação em Cálculo I	Progressão curricular correcta	0,27
Classificação em Matemática	Classificação em Cálculo I	Aprovação total no 1 ^o semestre	0,26
Classificação em Matemática	Classificação em Alg. Linear	Todos	0,38
Classificação em Matemática	Classificação em Alg. Linear	Progressão curricular correcta	0,18
Classificação em Matemática	Classificação em Alg. Linear	Aprovação total no 1 ^o semestre	0,15
Classificação em Física	Classificação em Física I	Todos	0,45
Classificação em Física	Classificação em Física I	Progressão curricular correcta	0,51
Classificação em Física	Classificação em Física I	Aprovação total no 1 ^o semestre	0,50

Com base nestes números, que mostram correlações extremamente baixas, decidiu-se traçar diagramas de dispersão para uma avaliação "visual" dos dados, tentando compreender se haveria alguma possibilidade de correlação, ou se as hipóteses e percepções de professores experientes seriam totalmente sem propósito. Além da opinião dos professores, a existência de uma relação causa-efeito entre classificações do vestibular e classificações no curso superior, foi apresentada por Soares de Mello et al (2000) ao verificar que cursos com maior procura tinham como consequência classificações no vestibular mais altas e, melhor aproveitamento escolar na universidade.

A avaliação visual faz uso dos gráficos exibidos na Figura 4.

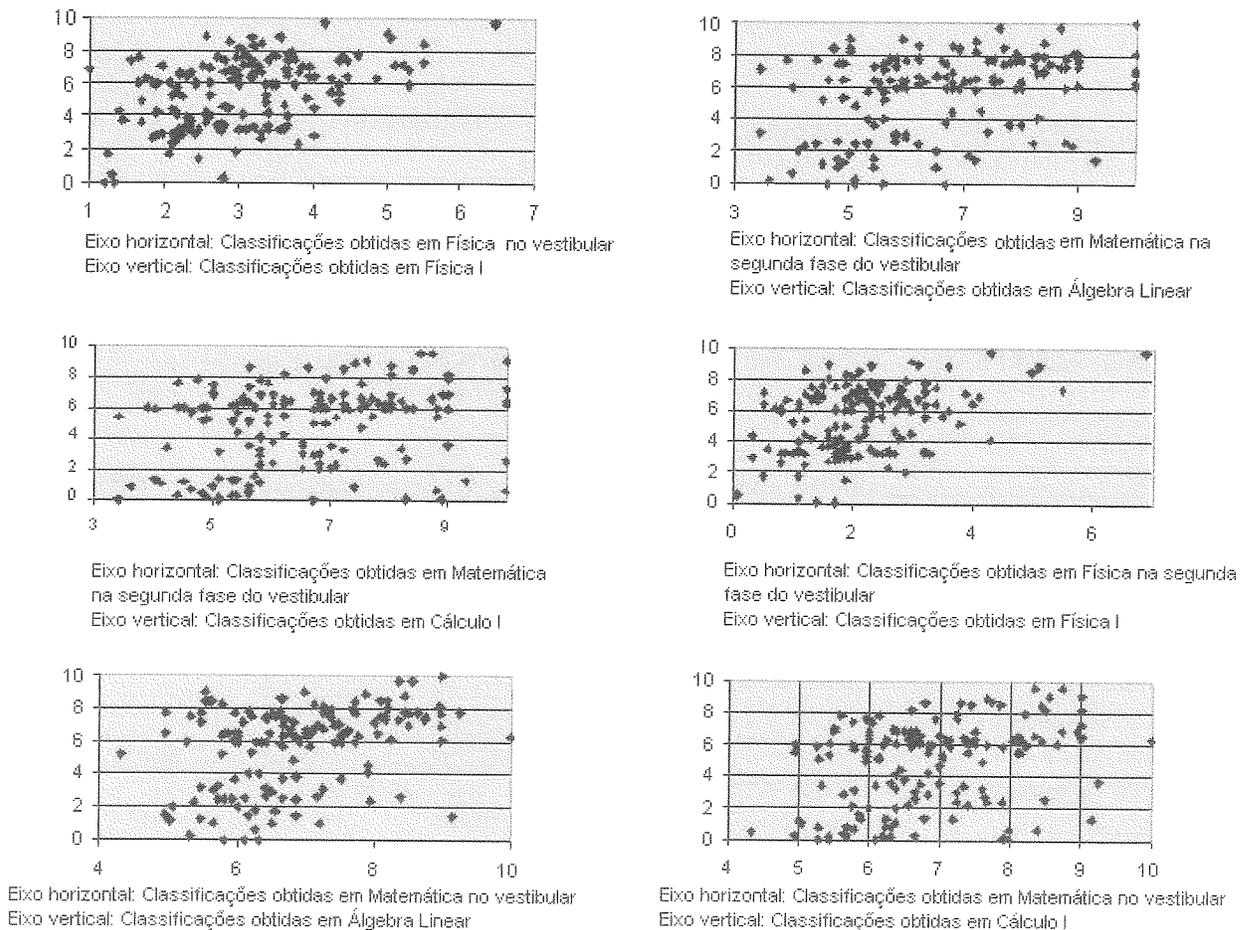


Figura 4: Relações entre classificações no vestibular e no primeiro semestre lectivo.

A observação dos gráficos sugere a existência de agrupamentos de alunos com características semelhantes. Estes agrupamentos, quando considerada a natureza multidimensional do problema, poderiam ser obtidos pelo uso da técnica de Análise de Agrupamentos ou "Cluster Analysis". Esta técnica é utilizada quando se deseja explorar as similaridades entre indivíduos definindo-os em grupos. Dada intenção de fazer este estudo com o uso de métodos elementares, optou-se por fazer uma análise bidimensional entre pares de variáveis. Fez-se inicialmente uma divisão de agrupamentos de forma visual, agrupando dados que, na análise gráfica, indicam a existência de forte correlação. Por exemplo, o primeiro gráfico sugere a existência de alta correlação entre as classificações de Cálculo I, quando situadas em [8,10], e as classificações da segunda fase do vestibular. De facto, ao agrupar estes dados, obtém-se coeficiente de correlação 0,81. A mesma análise foi feita para os demais conjuntos, com agrupamentos semelhantes, e obteve-se o resultado apresentado na Tabela 2.

Tabela 2: Correlações obtidas ao considerar só classificações entre 8,0 e 10,0 na universidade.

Classificações correlacionadas (Universidade - Vestibular)	Coefficiente de Correlação
Cálculo I - Matemática	0,805571
Cálculo I - Matemática na segunda fase	0,675615
Física I - Física	0,537828
Física I - Física na segunda fase	0,636131
CR no primeiro semestre - Classificação final	0,61539

Embora técnicas mais elaboradas de Análise de Agrupamentos possam ser usadas, os resultados aqui obtidos já mostram correlações razoáveis, das quais conclusões importantes podem ser tiradas. Assim, não foi julgado necessário proceder a análises estatísticas mais elaboradas.

7 Conclusões

Os métodos, tanto estatísticos quanto multicritério, usados neste artigo podem ser considerados elementares. Portanto, as conclusões a seguir apresentadas devem ser consideradas como preliminares.

Em primeiro lugar, as opiniões dos professores só são coerentes com as classificações obtidas pelos alunos, se admitir-se que os professores apenas observam com atenção os alunos que têm excelentes classificações na sua disciplina quando chamados a definir qual deve ser o estereótipo do bom aluno. Portanto, pode-se considerar que os professores entendem que "bom aluno" é aquele que tem capacidade para resolver as provas. Note-se que este não é necessariamente um parâmetro que caracterize sucesso profissional nem boa progressão curricular.

As provas de Física do vestibular têm maior correlação com as classificações posteriores nesta matéria do que as de Matemática com Cálculo e Álgebra. Assim, há uma indicação de que as provas de Física do vestibular estão mais adequadas que as de Matemática na avaliação dos conhecimentos necessários ao desempenho nas correspondentes disciplinas dos cursos de Engenharia. Outra alternativa é que a disciplina de Física I não acrescente muito além do ensino secundário, considerando as dificuldades que os estudantes encontram nas disciplinas de Cálculo I e Álgebra Linear.

Existem técnicas de análise de dados que podem enriquecer o estudo efectuado. A título de exemplo a correlação linear múltipla pode evidenciar a existência de relações lineares entre mais de duas variáveis do conjuntos de dados. Há ainda que considerar que as correlações podem não ser lineares, identificando outras correlações, obtidas de diferentes relações funcionais entre as variáveis.

Apesar das técnicas usadas serem consideradas elementares, mostraram-se extremamente úteis para agregar respostas de questionários. O uso de técnicas mais elaboradas, tanto de multicritério, quanto de estatística, embora possam melhorar os resultados obtidos, sem dúvida, dificultam a compreensão (e aceitação) por leigos. Mesmo que se opte por usar técnicas mais avançadas, uma análise preliminar com as técnicas aqui usadas revela-se bastante útil.

Finalmente, um outro agrupamento, de natureza geográfica, poderia ser tentado, caso houvesse disponibilidade de dados. Como os alunos da UFF, devido à característica de dispersão geográfica do seu vestibular (Soares de Mello et al, 2001), são oriundos de diversas regiões do Estado do Rio de Janeiro, onde existem várias desigualdades sócio-económicas (Gomes et al, 2002), a relação de causa efeito entre as classificações do vestibular e o desempenho no curso superior pode ser mascarada pelos referidos factores sociais, económicos e até de adaptação a uma nova cidade.

8 Referências

- [1] Barba-Romero, S., Pomerol, J.C., Decisiones Multicriterio: Fundamentos Teóricos e Utilización Práctica, Colección de Economía, Universidad de Alcalá, (1997.)
- [2] COSEAC, Manual do Vestibular UFF 2002, Universidade Federal Fluminense, Niterói, Brasil, (2001.)
- [3] Dias, L.M.C., Almeida, L.M.A.T., Clímaco, J.C.N., Apoio Multicritério à Decisão, Faculdade de Economia, Universidade de Coimbra, Coimbra, Portugal, (1996.)
- [4] Gomes, E. G.; Lins, M. P. E.; Soares de Mello, J. C. C. B., Seleção do melhor município: Integração SIG-Multicritério, *Investigação Operacional*, v. 22, n. 1, (2002)
- [5] Roy, B., Bouyssou, D. Aide multicritère à la décision: méthodes et cas, *Economica*, Paris (1993).
- [6] Soares de Mello, J. C. C. B.; Gomes, E. G.; Lins, M. P. E.; Vieira, L. A. M., Um caso de estudo de integração SIG-DEA-MCDA: a influência de uma instituição de ensino superior em vários municípios do Estado do Rio de Janeiro, *Investigação Operacional*, v. 21, n. 2, (2001)
- [7] Soares de Mello, J. C. C. B.; Lins, M. P. E.; Soares de Mello, M. H. C.; Gomes, E. G., Evaluating the Performance of Calculus Classes Using Operational Research Tools, *European Journal Of Engineering Education*, Inglaterra, 27, n. 2, (2002.)
- [8] Soares de Mello, M. H. C.; Soares de Mello, J. C. C. B., Ingresso nos cursos de Engenharia: Vestibular junto e separado, XXVIII COBENGE-Congresso Brasileiro de Ensino de Engenharia, Ouro Preto, Brasil, (2000)
- [9] Soares de Mello, M. H. C.; Soares de Mello, J. C. C. B.; Vaz, M. R., Análise exploratória dos dados da separação do vestibular para as habilitações da Engenharia na UFF, . VI EEE-Encontro de Educação em Engenharia, Itaipava - Petrópolis, Brasil, (2000)
- [10] Soares de Mello, M.H.C., Avaliação de Desempenho nas Engenharias: Estudo de Caso UFF, Tese de Mestrado, Coordenação de Pós-Graduação em Engenharia de Produção, Universidade Federal Fluminense, Niterói, Brasil (2002).

Planejamento e Programação da Produção numa Fundição Cativa Automatizada de Grande Porte

Silvio Alexandre de Araujo *

Marcos Nereu Arenales †

* Departamento de Informática
Universidade Estadual de Maringá-PR
Brasil
silvio@din.uem.br

† Instituto de Ciências Matemáticas e de Computação
Departamento de Ciências de Computação e Estatística
Universidade de São Paulo
Brasil
arenales@icmc.usp.br

Abstract

This work consists of a case study in a large foundry, which has three molding lines and two sets of furnaces. The molding lines produce different types of items, which must be made of different alloys produced by the furnaces. In each period, the production programming in this foundry has two important levels of decision making which are linked: 1) which alloys should be produced in the furnace in each period and 2) the quantity of each item to be produced in each molding machine. Initially, the productive process is described and then a mixed integer programming optimization model is proposed in which the lot-sizing of the items and the capacity programming of the furnaces are considered. A commercial package/software was utilized to obtain the computational results. A simulation lasting 5 working days was carried out and the computational results obtained from the model solution were compared with the ones obtained by the company, showing a significant increase in the productivity and reduction concerning delays of the delivery deadline.

Resumo

Este trabalho consiste num estudo de caso em uma fundição cativa de grande porte, a qual tem três linhas de moldagem e dois conjuntos de fornos. As linhas de moldagem produzem diferentes tipos de itens com demandas conhecidas e que devem ser feitos com diferentes ligas produzidas pelos fornos. Em cada período, a programação da produção nessa fundição envolve dois níveis importantes de decisão que estão inter-relacionados: 1) qual liga deve ser produzida no forno em cada período, e 2) a quantidade de cada item a ser produzida em cada máquina de moldagem. Inicialmente, o processo produtivo é

descrito e então um modelo matemático de otimização inteira mista é proposto o qual considera o dimensionamento de lotes para os itens e a programação da capacidade dos fornos. Um pacote comercial foi utilizado para a obtenção dos resultados computacionais. Foi feita uma simulação de 5 dias de trabalho e os resultados obtidos a partir da solução do modelo foram comparados com os resultados obtidos na prática pela indústria, mostrando expressivos ganhos de produtividade e redução de atrasos na entrega dos pedidos.

Keywords: programação inteira mista, dimensionamento de lotes, programação da produção, fundição.

Title: Lot sizing and Scheduling in a Large Automated Foundry

1 Introdução

Atualmente, no Brasil as fundições estão presentes em todas as regiões produzindo desde peças simples de uso doméstico, até as mais sofisticadas, como autopeças de veículos e peças e/ou partes de máquinas e equipamentos de uso industrial. Segundo o relatório da ABIFA- Associação Brasileira de Fundição, em Novembro de 2003 o setor produziu uma média de 10.082 toneladas/dia. Esse número subiu para uma média de 12.338 toneladas/dia em Novembro de 2004 representando um crescimento de 22,3% gerando 55.715 empregos diretos (ABIFA, 2004).

A grande maioria da produção brasileira de fundidos é fabricada por **fundições cativas**, que são departamentos de grandes empresas (por exemplo, algumas montadoras de automóveis) cuja produção é destinada, basicamente, para consumo próprio. Essas fundições cativas estão orientadas para a produção em série, automatizada e fabricam grandes quantidades de cada item. Este trabalho enfoca problemas de planejamento e programação da produção que surgem numa fundição cativa do interior do estado de São Paulo.

Revisões bibliográficas de trabalhos que tratam de problemas de planejamento e programação da produção podem ser encontradas em Billington *et al.* (1983), Bahl *et al.* (1987), Maes e van Wassenhove (1988), Goyal e Gunasekaran (1990), Potts e van Wassenhove (1992), Kuik *et al.* (1994), Drexl e Kimms (1997), Potts e Kovalyov (2000) e Karimi *et al.* (2003). Entretanto, existem poucos trabalhos publicados em revistas que tratam de problemas de planejamento e programação da produção, especificamente orientados para o setor de fundições.

Em Gonçalves-Vianna e Arenales (1995), Santos-Meza *et al.* (2002), Araujo e Arenales (2003), Araujo *et al.* (2003) e Araujo (2003) são realizados estudos orientados para fundições de pequeno e médio porte. Os trabalhos encontrados na literatura orientados para grandes fundições tratam, especificamente, do problema de sequenciamento das tarefas nas máquinas, como em Sounderpandian e Balashanmugam (1991), Gravel *et al.* (2000). A grande maioria dos trabalhos relacionados ao tema trata do problema de programação da produção em grandes siderúrgicas. Este tipo de fábrica é caracterizado pela inexistência do setor de moldagem, pois os itens demandados são chapas de aço de diferentes tamanhos e composições, fabricadas por máquinas de laminação, onde o metal passa por cilindros rolantes distanciados de acordo com as especificações da chapa a ser fabricada. Tang *et al.* (2001) fazem uma revisão bibliográfica de trabalhos de planejamento e programação da produção aplicados à indústria de fabricação de aço. Vários trabalhos tratam de problemas práticos encontrados em grandes siderúrgicas: Petersen *et al.* (1992), Hamada *et al.* (1995), Bowers *et al.* (1995), Hendry *et al.* (1996), Lee

et al. (1996), Lopes *et al.* (1998), Tang *et al.* (2000a), Tang *et al.* (2000b).

2 O Processo Produtivo

Uma fundição é caracterizada pela transformação, numa primeira etapa, de minérios de ferros em ligas metálicas, que por sua vez, ainda em estado líquido, são despejadas dentro de moldes, os quais são normalmente feitos de areia e resina, para a produção de peças fundidas. Assim, dois tipos de equipamentos básicos estão presentes numa fundição: fornos de fusão, para a produção de ligas metálicas e, máquinas de moldagem, para a produção dos moldes (em fundições pequenas, a moldagem é normalmente manual). A Figura 1 mostra um conjunto de nove fornos e três máquinas de moldagem. Nas máquinas de moldagem estão integradas algumas das operações, tais como a produção dos moldes, a injeção da liga metálica nos moldes, o resfriamento da liga, a eliminação da areia e, finalmente a eliminação de rebarbas, que consistem em resíduos de ligas. Por isto, são chamadas *linhas de moldagem*. Um molde, em geral, é composto de várias cavidades e, cada uma das cavidades gera uma peça (produto final). A Tabela 1 ilustra uma carteira de pedidos que na coluna “cavidade” mostra o número de cavidades de cada molde, que varia de 6 a 48. Por exemplo, um único molde pode produzir um conjunto de 48 peças do tipo 24 e esse conjunto de peças é chamado “árvore”.

É comum que grandes fundições organizem seus equipamentos em setores, de modo que certas linhas de moldagem sejam abastecidas por ligas metálicas de certos fornos. Por exemplo, a fundição estudada tem dois setores. O setor 1 consiste em cinco fornos (na Figura 1 são F1, ..., F5) e duas linhas de moldagem (L1 e L2). Os fornos do setor 1 abastecem apenas as linhas de moldagem desse setor, conforme representado pelas setas na figura 1. Dois tipos principais de ligas são produzidos: cinzenta e nodular. A liga cinzenta pode ser subdividida em 5 outras ligas e a nodular em 2 diferentes tipos de ligas. O setor 1 produz apenas peças referentes à liga cinzenta e o setor 2 pode produzir peças das duas ligas (nodular e cinzenta). Entretanto, uma vez que uma liga esteja sendo produzida num determinado período, os fornos do setor 2 devem produzir essa mesma liga e repassá-la para o reservatório (um forno de grande capacidade que mantém a liga em estado líquido).

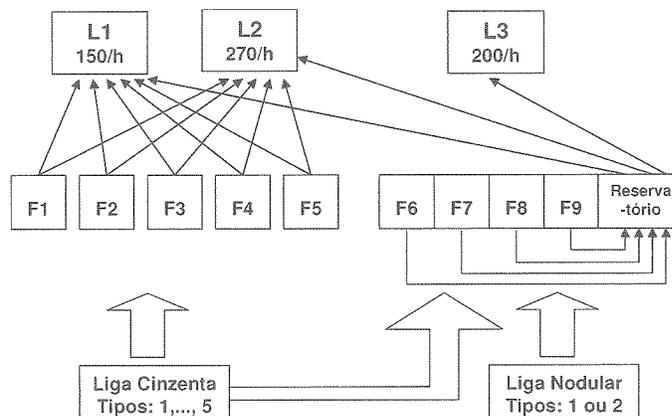


Figura 1: Processo produtivo da empresa.

Veja na Figura 1 que não há transporte de liga dos fornos F6-F9 para as linhas de moldagem.

Assim como os fornos têm uma capacidade limitada para a produção de ligas por hora, as linhas de moldagem têm uma capacidade produtiva limitada. Por exemplo, a linha L1 produz em média 150 moldes por hora, enquanto a linha L2 produz 270 e a linha L3 produz 200. Naturalmente, a produtividade das linhas depende da peça a ser moldada e existem algumas peças que não podem ser produzidas por determinadas linhas (normalmente, por questão de qualidade).

3 Definição do Problema

A programação da produção é feita por turno de trabalho. A fundição estudada funciona 24 horas por dia, que é dividido em 3 turnos de trabalho. Os fornos funcionam ininterruptamente durante os 3 turnos, porém, uma fração de tempo do terceiro turno é utilizada para manutenção das linhas. Isto faz com que os fornos estejam com carga máxima no início do primeiro turno do dia seguinte. É também interessante observar que muitas vezes a programação feita pela empresa admite diminuir o tempo de manutenção. Veja os resultados na tabela 5, a qual mostra que a linha L3 foi utilizada além do tempo disponível, porém a modelagem matemática do problema não admite essa solução.

Para efeito de programação da produção, cuja ênfase está nas linhas de moldagem, é suficiente considerar apenas duas ligas, cinzenta e nodular e, portanto, as subdivisões de cada liga não são consideradas. Esta simplificação é perfeitamente aceitável, pois a transformação da liga (cinzenta ou nodular) para uma das sub-divisões é feita depois que a base estiver pronta, apenas adicionando-se diferentes elementos.

O setor 1 (fornos F1, ..., F5) e o setor 2 (fornos F6, ..., F9) têm capacidades de armazenamento de 10.000 kg e 8.000 kg, respectivamente. A energia contratada para os fornos atualmente é de 6.000 kw/h e, para a fabricação de 1.000 kg por hora de liga são necessários 750 kw de energia. Portanto, é possível fabricar, em uma hora, no máximo 8.000 kg de liga. Isto representa a capacidade total de reposição dos fornos por hora. A distribuição de energia, por hora, para o setor 1 é de, no máximo, 4.000 kw (ou seja, no máximo é possível produzir 5.330 kg de liga por hora no setor 1) e, para o setor 2 é de, no máximo, 2.500 kw por hora (capacidade em liga: 3.330 kg por hora).

Observe que a soma das capacidades máximas de reposição dos dois setores ultrapassa 8.000 kg por hora, que é a capacidade máxima de reposição total. Portanto, como deve ser feita a reposição nos dois setores é uma decisão a ser tomada.

Desta forma, na modelagem matemática (seção 4) existirão variáveis de decisão referentes ao armazenamento das ligas e variáveis de decisão referentes à reposição das ligas. Tais variáveis serão limitadas pelas capacidades de armazenamento e de reposição dos setores.

A carteira de pedidos de determinado mês foi fornecida pela indústria e é mostrada na Tabela 1. A indústria fabrica 8 diferentes famílias de produtos finais, os quais são decompostos em diferentes peças que devem ser fabricadas pela fundição. Para cada peça existe um molde específico que produz uma quantidade de peças de cada vez, dependendo do número de cavidades presentes no molde. O conjunto de peças feito em um só molde é chamado árvore.

A demanda interna de cada peça é constante em cada mês, ou seja, durante 1 mês, a cada dia tem-se a mesma quantidade demandada para cada peça. Para se fazer a programação

Tabela 1: Carteira de pedidos real.

Peça	Família	Demanda Peça	Cavidades	Demanda Árvore	Peso Peça	Peso Árvore	Peso Dia	Liga	Est. Inicial	Peso Est. Ini.
1	A	2580	6	430	2,73	16,4	7052	C	777	12745,53
2	B	4248	6	708	3,30	19,8	14018,4	C	397	7863,90
3	C	17496	12	1458	1,79	21,48	31317,84	C	1502	32264,75
4	D	11328	6	1888	2,78	16,7	31529,6	C	2706	45190,20
5	E	6576	6	1096	2,62	15,7	17207,2	C	1174	18434,42
6	F	5664	8	708	1,73	13,8	9770,4	C	700	9660,00
7	F	5652	12	471	1,03	12,3	5793,3	C	875	10762,50
8	G	404	4	101	4,13	16,5	1666,5	C	193	3176,25
9	B	4200	35	120	0,31	10,8	1296	C	280	3024,00
10	G	400	8	50	1,29	10,34	517	C	150	1551,00
11	D	11305	35	323	0,42	14,8	4780,4	C	189	2790,86
12	E	6545	35	187	0,35	12,2	2281,4	C	257	3137,14
13	G	408	12	34	0,65	7,8	265,2	C	350	2730,00
14	G	400	20	20	0,52	10,4	208	C	20	208,00
15	G	396	36	11	0,24	8,75	96,25	C	11	96,25
16	A	2600	40	65	0,22	8,9	578,5	C	185	1646,50
17	B	2562	42	61	0,17	7,3	445,3	C	452	3302,38
18	A	2592	48	54	0,14	6,85	369,9	C	125	856,25
19	G	384	24	16	0,30	7,3	116,8	C	171	1247,08
20	A	2583	21	123	0,75	15,85	1949,55	N	310	4905,95
21	B	4245	15	283	0,89	13,37	3783,71	N	360	4813,20
22	F	5655	15	377	0,69	10,3	3883,1	N	673	6935,33
23	C	5817	21	277	0,58	12,1	3351,7	N	732	8859,50
24	H	11616	48	242	0,35	17	4114	N	542	9208,33
25	D	11322	18	629	0,69	12,4	7799,6	N	581	7205,78
26	E	6570	18	365	0,70	12,62	4606,3	N	417	5258,33
27	G	400	10	40	1,23	12,3	492	N	110	1353,00
Total		133948		10137			159290		14239	209226,45

da produção para um dia de trabalho é preciso considerar o estoque inicial de cada dia. No estoque inicial, são consideradas as peças que estão prontas na fundição, bem como as peças que estão armazenadas no setor de usinagem.

Como é possível observar na Tabela 1, no mês em questão, a indústria tem 27 peças que devem ser programadas para a fabricação diária. A Tabela 1 apresenta ainda, a demanda diária de cada peça, o número de cavidades de cada molde, a demanda por número de árvores (que é calculada dividindo-se a demanda da peça pelo número de cavidades do molde), o peso bruto de cada peça, o peso bruto de cada árvore, a liga em que cada peça deve ser feita (C=Cinzenta ou N=Nodular) e o estoque inicial (em número de árvores e em peso) de cada árvore.

O problema da programação da produção na fundição estudada consiste em programar simultaneamente os fornos e as linhas de moldagem. Ou seja, deve-se determinar as quantidades adequadas das ligas produzidas pelos fornos em cada período do planejamento, respeitando-se as capacidades dos fornos (seja de reposição como de armazenagem), as quais serão usadas para a fabricação de peças nas máquinas mais adequadas (uma mesma peça pode ser produzida em máquinas diferentes, porém uma máquina pode ser considerada pelo programador como “mais adequada” que outra). Isto deve ser feito de modo a atender a demanda e que os atrasos sejam mínimos e os estoques em níveis desejados. Na seção 4 apresentamos um modelo matemático para este problema.

4 Modelagem Matemática

Antes de apresentar a modelagem matemática, é necessário fazer alguns comentários com respeito a algumas exigências impostas pelo gerente de produção da indústria, as quais devem fazer parte das decisões do planejamento e programação da produção e complementam as informações anteriores:

- a programação deve ser feita em número de árvores de cada tipo de peça, ou seja, deve-se determinar o número de vezes que o molde (que fabrica determinada peça) será utilizado, em cada linha de moldagem, em cada turno de trabalho.
- os lotes devem ser quebrados em 3, caso a demanda seja maior que 1000 árvores ou em 2, caso a demanda esteja entre 500 e 1000 árvores. Por outro lado, os tamanhos dos lotes não devem ser muito pequenos, pois, existe um tempo de preparação das linhas de moldagem, que envolve a troca dos moldes e um custo de preparação para a limpeza do molde que estava sendo utilizado (o tempo e o custo de preparação são independentes da seqüência). Embora o tempo e custo de preparação sejam pouco relevantes na produção, são considerados no modelo com o propósito de inibir a formação de lotes pequenos. Deste modo, não é importante considerar no modelo matemático o fato da linha de moldagem estar preparada para a produção de uma peça fabricada no período anterior (em problemas com tempo e custo de preparação altos esta consideração pode ser importante);
- é desejável que se tenha estoque de peças para um dia inteiro de trabalho. No entanto, o nível de estoque também deve ser limitado superiormente;

A seguir, um modelo matemático é apresentado:

Índices:

$m=1, \dots, M$	linhas de moldagem;
$k=1, \dots, K$	tipos de ligas;
$i=1, \dots, N$	tipos de peças;
$t=1, \dots, T$	períodos de tempo.
η_t	número máximo de sub-períodos no período t ($t=1, \dots, T$);
$F_t = 1 + \sum_{\tau=1}^{t-1} \eta_\tau$	denota o primeiro sub-período do período t ;
$L_t = F_t + \eta_t - 1$	denota o último sub-período do período t ;
$n = F_1, \dots, L_T$	sub-períodos.

Observação: Vale ressaltar que, cada turno de trabalho deve ser programado separadamente e, a cada dia, é preciso re-programar a produção para os 3 turnos. Como os custos são subjetivamente estimados e a carteira de pedidos é atualizada freqüentemente dentro do horizonte de planejamento (principalmente devido a paralisações na usinagem), o problema é resolvido considerando algumas variáveis e restrições somente para os períodos imediatos (para o problema prático em questão, serão considerados períodos imediatos apenas os períodos relativos ao primeiro dia de trabalho, ou seja, 3 turnos de trabalho). Posteriormente, o horizonte é rolado progressivamente para uma nova aplicação do modelo (Clark e Clark, 2000). Assim, a programação para o primeiro dia é feita de forma detalhada, ou seja, o primeiro período $t = 1$

(que equivale a um dia) será subdividido em 3 sub-períodos (que equivalem aos 3 turnos de trabalho) e será feita a programação para cada turno de trabalho separadamente. Além disso, será feita a programação de forma aproximada (sem subdividir em turnos de trabalho) para outros 4 dias de trabalho, o que resulta num total de 5 períodos (ou 5 dias), ou seja:

$$F_1 = 1 \text{ e } L_1 = 3; F_2 = 4 \text{ e } L_2 = 4; F_3 = 5 \text{ e } L_3 = 5; F_4 = 6 \text{ e } L_4 = 6; F_5 = 7 \text{ e } L_5 = 7.$$

Desta forma, tem-se $T=5$ períodos t e $n = F_1, \dots, L_5$, ou seja, $n=1, \dots, 7$ sub-períodos n (ver Figura 2).

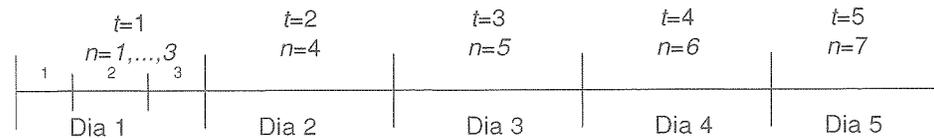


Figura 2: Períodos e sub-períodos do Método de Horizonte Rolante.

Dados do problema:

- Cap_k capacidade total de armazenamento da liga k (é determinada pela capacidade dos fornos que podem armazenar a liga k);
- CT capacidade total de armazenamento de todos os fornos;
- $CRep_k$ capacidade máxima de produção (reposição) da liga k numa hora de trabalho (é determinada pela capacidade de reposição dos fornos que podem produzir a liga k);
- $CTrep$ capacidade total de reposição de todos os fornos numa hora de trabalho (delimitada pela energia contratada);
- d_{it} demanda (em número de árvores) da peça i no período t ;
- ρ_i peso unitário bruto de cada árvore da peça i ;
- $S(k)$ conjunto de peças que utilizam a liga k ;
- a_{im} quantidade de árvores da peça i que a linha m consegue fazer por hora;
- h_{mn} número de horas disponíveis na linha de moldagem m no sub-período n ;
- q_{mn} número de horas de paralisações (quebras) na linha de moldagem m no sub-período n ;
- hf_n número de horas disponíveis nos fornos no sub-período n ;
- H_{it}^+ custo de estocagem de uma árvore da peça i no período t ;
- H_{it}^- custo pelo atraso na entrega de uma árvore da peça i no período t ;
- c_{im} custo para produzir uma árvore da peça i na linha m ($c_{im} = \infty$ significa que a árvore da peça i não pode ser produzida pela linha m);
- $Xmax_i$ tamanho máximo de um lote da árvore da peça i ;
- $Imax_{it}$ estoque máximo de árvores da peça i no período t ;
- s_m custo de preparação na linha m ;
- st_m tempo de preparação na linha m ;
- CF_{k0} quantidade de liga k no início do planejamento (também no início de cada dia).

Observação: As horas disponíveis (h_{mn}) para cada linha poderiam incluir as horas de paralisação (q_{mn}), pois para efeito de programação deve-se considerar: $h_{mn} - q_{mn}$. Porém, q_{mn} é um parâmetro aleatório que auxilia o programador em simulações.

Variáveis do problema:

- X_{imn} quantidade de árvores da peça i produzida na linha m no sub-período n ;
 I_{it}^+ quantidade estocada de árvores da peça i no período t ;
 I_{it}^- quantidade em atraso (ou, estoque negativo) de árvores da peça i no período t ;
 CF_{kn} quantidade disponível (em kg) de liga k no final do sub-período n ;
 CR_{kn} quantidade de reposição (em kg) da liga k no sub-período n ;
 Y_{imn} variável binária que indica a preparação para produzir a árvore da peça i na linha m no sub-período n .

Modelo:

$$\text{Minimizar } \sum_{i=1}^N \sum_{t=1}^T (H_{it}^- I_{it}^- + H_{it}^+ I_{it}^+) + \sum_{i=1}^N \sum_{m=1}^M \sum_{t=1}^T \sum_{n=F_t}^{L_t} c_{im} X_{imn} + \sum_{i=1}^N \sum_{m=1}^M \sum_{t=1}^T \sum_{n=F_t}^{L_t} s_m Y_{imn} \quad (1)$$

Sujeito a:

$$I_{i,t-1}^+ - I_{i,t-1}^- + \sum_{n=F_t}^{L_t} \sum_{m=1}^M X_{imn} - I_{it}^+ + I_{it}^- = d_{it} \quad i = 1, \dots, N \quad t = 1, \dots, T \quad (2)$$

$$\sum_{i=1}^N \left(\frac{X_{imn}}{a_{im}} + st_m Y_{imn} \right) \leq (h_{mn} - q_{mn}) \quad m = 1, \dots, M \quad n = F_1, \dots, L_T \quad (3)$$

$$0 \leq X_{imn} \leq X_{max_i} Y_{imn} \quad i = 1, \dots, N \quad m = 1, \dots, M \quad n = F_1, \dots, L_T \quad (4)$$

$$CF_{k,n-1} - \sum_{i \in S(k)} \sum_{m=1}^M \rho_i X_{imn} + CR_{kn} = CF_{kn} \quad k = 1, \dots, K \quad n = F_1, \dots, L_1 \quad (5)$$

$$0 \leq CF_{kn} \leq Cap_k \quad k = 1, \dots, K \quad n = F_1, \dots, L_1 \quad (6)$$

$$\sum_{k=1}^K CF_{kn} \leq CT \quad n = F_1, \dots, L_1 \quad (7)$$

$$0 \leq CR_{kn} \leq CRep_k(hf_n) \quad k = 1, \dots, K \quad n = F_1, \dots, L_T \quad (8)$$

$$\sum_{k=1}^K CR_{kn} \leq CTrep(hf_n) \quad n = F_1, \dots, L_T \quad (9)$$

$$\sum_{i \in S(k)} \sum_{m=1}^M \rho_i X_{imn} \leq CF_{k0} + CR_{kn} \quad k = 1, \dots, K \quad n = F_2, \dots, L_T \quad (10)$$

$$0 \leq I_{it}^+ \leq I_{max_{it}} \quad i = 1, \dots, N \quad t = 1, \dots, T \quad (11)$$

$$I_{it}^- \geq 0 \quad i = 1, \dots, N \quad t = 1, \dots, T \quad (12)$$

$$Y_{imn} \in \{0, 1\} \quad i = 1, \dots, N \quad m = 1, \dots, M \quad n = F_1, \dots, L_T \quad (13)$$

Observações: Note que certas restrições ocorrem a cada período (veja (2)), ou a cada sub-período (veja (3)), ou apenas para os primeiros sub-períodos a serem implementados (veja (5)), ou somente para os períodos futuros (veja (10)).

A função (1) é composta de três parcelas que objetivam a formação de estoques ou atrasos coerentes, a designação das peças para serem fabricadas nas máquinas mais adequadas e evitar a formação de lotes pequenos. Para isto, os coeficientes em (1) devem ser definidos pelo utilizador. Por exemplo, quando há interesse de se manter peças em estoques, de modo que, em caso de quebra de máquina, a demanda possa ser atendida pelo estoque, valores negativos são usados em H_{it}^+ para incentivar que as variáveis I_{it}^+ sejam positivas. As variáveis I_{it}^- correspondem a atrasos e o gerente de programação pode ajustar H_{it}^- de modo a proibir ou não atrasos de certos itens. A segunda parcela da função objetivo é também usado para levar em conta algumas proibições do processo produtivo, pois algumas peças não podem ser feitas por determinada linha, devido a alguma restrição da linha de moldagem. Assim, um valor alto para c_{im} faz com que o item i não seja produzido na máquina m . Além disso, o segundo termo ainda descreve algumas prioridades em relação a certas peças que, se possível, deveriam ser

feitas em linhas específicas, as quais são mais adequadas para produzir tais peças. O terceiro termo em (1) define os custos de preparação.

Além disso, os estoques negativos fornecem uma ferramenta para a avaliação dos prazos de entrega. Por exemplo, se $d_{i1}=30$ e uma solução tal que: $I_{i1}^- = 10$, $I_{i2}^- = 5$, $I_{i3}^- = 0$, significa que a demanda para o item i será totalmente atendida com 3 dias de atraso. Desta forma, é possível alertar o cliente e, se necessário, re-calibrar os parâmetros relativos a esta peça.

As restrições (2) são de balanceamento de estoque e consideram tanto estoques positivos, como negativos (atrasos). As restrições (3) são devido à limitação de tempo disponível das linhas de moldagem e (4) indicam que pode haver produção de determinada peça somente se houver preparação para esta peça e, a quantidade total produzida de árvores da peça i na linha m no sub-período n deve ser menor que o tamanho máximo do lote para a peça i . As restrições (5)-(7) se aplicam apenas para os L_1 primeiros sub-períodos (onde cada turno de trabalho é programado separadamente) e (5) são restrições de balanceamento da capacidade dos fornos, levando em consideração a quantidade de liga restante do turno anterior ($CF_{k,n-1}$), a capacidade de reposição de cada liga (CR_{kn}) e a quantidade (em peso) de liga utilizada em cada turno ($\sum_{i \in S(k)} \sum_{m=1}^M \rho_i X_{imn}$). Cabe observar que, CF_{k0} é um dado do problema e representa a quantidade de cada liga presente no início do planejamento, que normalmente é o primeiro turno de um dia de trabalho e os fornos estão carregados devido as máquinas de moldagem terem sido paradas para manutenção. Em (6) a quantidade de liga restante num turno (CF_{kn}) é limitada pela capacidade total de armazenamento da liga (Cap_k) e, em (7) a quantidade restante de todas as ligas ($\sum_{k=1}^K CF_{kn}$) é limitada pela capacidade total de armazenamento (CT). Em (8) a capacidade de reposição da liga k num sub-período n (CR_{kn}) é limitada pela capacidade máxima de reposição por hora da liga k , vezes o número de horas do sub-período ($Crep_k (hf_n)$) e, em (9) a capacidade de reposição de todas as ligas num sub-período ($\sum_{k=1}^K CR_{kn}$) é limitada pela capacidade total de reposição por hora (restrição devido à energia contratada), vezes a quantidade de horas do sub-período ($CTrep (hf_n)$). Em (10) as restrições de capacidade de produção de ligas para os outros dias de trabalho (quando os períodos não são divididos em sub-períodos) são consideradas, sendo a quantidade total de liga limitada pela quantidade inicial da liga no começo do dia, mais a capacidade de reposição da liga no dia de trabalho. As restrições (11)-(13) limitam as variáveis de estoque, de atraso e definem a variável de preparação como binária.

5 Resultados Computacionais

Para apresentar os resultados computacionais foram simulados 5 dias de programação utilizando os dados reais da indústria. O modelo foi resolvido 5 vezes (com o pacote de otimização inteira Ampl/Cplex 7.1, num Pentium III de 500 MHz com 512 MB de RAM) considerando apenas a programação detalhada para cada dia separadamente. Entre uma e outra resolução os dados foram atualizados, considerando o estoque inicial do dia, e todas as quebras das linhas ocorridas durante o dia. A demanda não precisa ser atualizada, pois é constante durante os 5 dias. A rigor, no primeiro período a demanda é $d_{i1} - I_{i0}^+$.

Para cada dia de programação, a resolução do modelo envolve 567 variáveis binárias Y_{imn} ,

567 variáveis lineares X_{imn} , 270 variáveis lineares I_{it}^+ e I_{it}^- , 28 variáveis lineares CF_{kn} e CR_{kn} e um total de 974 restrições. Na programação de cada dia o tempo do pacote foi limitado em 2 minutos e em nenhum dos problemas chegou-se a uma solução a mais de 3% do limitante inferior. Portanto, todas as soluções estão relativamente (em se tratando de um problema prático) próximas do ótimo. Isso mostra que o problema é relativamente pequeno permitindo chegar a bons resultados utilizando um pacote de otimização inteira num tempo computacional razoável.

A maioria dos dados utilizados é apresentada na Tabela 1. Observe que a demanda diária é de 10.137 árvores de peças e 159.290 kg e o estoque inicial e de 14.239 árvores de peças e 209.226 kg. Os custos (isto é, parâmetros da função objetivo) foram ajustados depois de várias reuniões com a gerência na empresa e de exaustivos testes computacionais. Vale ressaltar, a dificuldade em se estabelecer custos adequados e, mesmo depois de estabelecê-los é praticamente impossível saber se, realmente, são os que melhor representam os objetivos da empresa. Para isto é indispensável a participação de um programador experiente, que conheça bem a empresa. A seguir, são apresentadas algumas conclusões alcançadas após certa experiência com o modelo e com o problema prático.

Os custos de estocagem (H_{it}^+) devem ser negativos, pois é necessário manter um nível de estoque correspondente a um dia de trabalho. Os custos por atraso (H_{it}^-) devem ser altos (quando comparado com os outros custos) e positivos de forma a penalizar fortemente atrasos, os quais, são bastante indesejáveis, pois causam uma reação em cadeia no restante da empresa. Foram estabelecidos 4 custos diferentes relativos à produção dos itens nas linhas (c_{im}). Um deles é um valor muito alto que representa que a linha não pode produzir a peça, os outros 3 representam ótima, boa e inadequada associação de peças às linhas de moldagem. Finalmente, os custos de preparação são definidos pelo utilizador de modo a evitar lotes muito pequenos.

Durante os 5 dias simulados a quantidade de paralisações por quebras das linhas foi bastante expressiva (ver Tabela 2), dificultando a programação da produção. Observe que, em geral, apenas 71,90% do tempo foi disponível para a produção.

Tabela 2: Tempo disponível nas 3 linhas a cada dia.

	1	2	3	4	5	Média
Tempo Total (h)	64,86	64,86	64,86	64,86	64,86	64,86
Tempo de Paralisações (h)	16,68	12,8	25,05	20,34	16,25	18,22
Porcentagem Disponível	74,28	80,27	61,38	68,64	74,95	71,90

Os critérios utilizados para comparar as duas programações (programação realizada na prática e a programação obtida pela resolução do modelo matemático) foram estabelecidos com o gerente de produção da indústria e estão nas Tabelas 3 e 4, onde nas colunas tem-se informação sobre os 5 dias que foram programados e a média para esses 5 dias e, nas linhas, são analisados o número de preparações obtidos em cada programação, o custo total de produção (c_{im}) (estabelecidos de acordo com o interesse em se utilizar determinada linha para fazer determinada peça), a quantidade (unidades) total em estoque e o peso (kg) total estocado a cada dia, a quantidade total em atraso e o peso total em atraso a cada dia, a quantidade total produzida e o peso total produzido a cada dia, a porcentagem de utilização das linhas sem preparações e a porcentagem de utilização das linhas com preparações.

Tabela 3: Resumo da programação da produção: prática.

	1	2	3	4	5	Média
Preparações	15	22	15	19	21	18,4
Custo de Prod.	9129	16677	14094	13332	17583	14163
Quant. Estoque	12098	10810	8694	7386	7226	9242,8
Peso Estoque	174971	152104	123140	106907	100650	131554
Quant. Atraso	0	0	247	580	1752	515,8
Peso Atraso	0	0	2821	6922	21909	6330
Quant. Produz.	7996	9272	7352	8496	8804	8384
Peso Produzido	125035	143346	120582	138956	138046	133193
Utilização	78,18	83,14	92,62	88,14	86,88	85,47
Utilização Prep.	81,92	88,21	97,14	93,27	92,06	90,20

Tabela 4: Resumo da programação da produção: solução modelo (1)-(13).

	1	2	3	4	5	Média
Preparações	22	27	24	26	27	25,2
Custo de Prod.	13298	14092	10222	11810	12482	12381
Quant. Estoque	13029	12721	9698	8129	7241	10164
Peso Estoque	204548	201938	154618	128936	117744	161557
Quant. Atraso	0	0	0	0	0	0
Peso Atraso	0	0	0	0	0	0
Quant. Produz.	8935	9829	7113	8568	9249	8738,8
Peso Produzido	154714	156681	111960	133606	148091	141010
Utilização	90,22	89,51	87,32	88,19	88,34	88,79
Utilização Prep.	95,7	95,74	94,56	95,19	95	95,27

Observe que, nas duas programações o nível de estoque foi reduzido dia a dia. Isso reflete as paralisações das linhas. Vale lembrar que, é desejável manter o nível de estoque para um dia de trabalho (10.137 moldes/dia e 159.290kg/dia). O nível de estoque da programação feita pela solução do modelo (Tabela 4) se mantém ligeiramente superior em todos os dias.

O peso total produzido na programação do modelo é 5,8% maior que o peso total produzido pela programação da indústria.

O número total de preparações obtido na solução do modelo é maior que as preparações feitas pela indústria. Tal fato pode ser explicado pela maior utilização das linhas, o que levou a uma maior produção e conseqüente redução para zero da quantidade atrasada (que é o principal objetivo da empresa, refletido nos altos valores para H_{it}^-).

Foi feita uma análise da utilização de cada linha de moldagem separadamente. Foram analisadas as porcentagens de utilização do tempo disponível. Como é possível observar nas Tabelas 5 e 6, a programação obtida pela solução do modelo (1)-(13) utilizou cerca de 20,49% a mais da linha L1, 6,71% a mais da linha L2 e 7,62% a menos a linha L3.

Tabela 5: Porcentagem de utilização das linhas (com preparações): prática.

	1	2	3	4	5	Média
L1	52,02	63,02	98,98	80,96	96,73	75,91
L2	93,75	96,92	82,82	88,90	76,27	88,03
L3	96,13	104,71	106,09	104,19	102,67	102,53

Tabela 6: Porcentagem de utilização das linhas (com preparações): solução do modelo.

	1	2	3	4	5	Média
L1	97,2	96,74	97,04	96,27	93,96	96,40
L2	94,05	95,27	92,79	95,21	95,58	94,74
L3	95,79	95,2	93,64	94,62	95,05	94,91

Após uma análise minuciosa das peças produzidas em cada linha foi possível constatar que houve uma troca entre as linhas L2 e L3, ou seja, algumas peças que estavam sendo produzidas na linha L3 na programação da indústria, passaram a ser produzidas pela linha L2 na programação obtida pelo modelo. Isso aumentou o rendimento da linha L3, pois esta não foi sobrecarregada (os 102,53% da programação da indústria representam que a linha foi utilizada em hora-extra, diminuindo o tempo de manutenção). Além disso, a solução do modelo utilizou bastante a linha L1, o que é refletido no aumento expressivo da produção e na redução dos atrasos.

6 Conclusões e Trabalhos Futuros

Neste trabalho foi feito um estudo de caso em uma fundição cativa de grande porte. Um modelo matemático de programação inteira mista foi proposto e resolvido com auxílio de um pacote de otimização inteira. O modelo matemático considera a programação das linhas de moldagem, bem como, a distribuição de capacidade nos fornos. Os resultados computacionais indicam que o modelo matemático proposto representa bem o problema de programação da produção na fundição estudada e permite um aumento do rendimento das linhas (devido às melhores associações de peças às linhas), aumento na utilização das linhas, o que leva também a um aumento na produção, de forma a evitar atrasos e a manter o nível de estoque. Além disso, o modelo também apresenta um controle da capacidade dos fornos, impedindo que se faça uma programação nas linhas de moldagem que seja inviável para os fornos, ou seja, é feito um balanceamento no peso das peças que estão sendo produzidas a cada turno de trabalho de modo que, os pesos totais das ligas sejam admissíveis para os fornos. Na programação feita pela indústria este fato ocorre muitas vezes, ou seja, a programação das linhas é feita sem considerar a capacidade dos fornos de modo que, muitas vezes, as linhas ficam paradas aguardando o metal que está sendo fundido nos fornos. Isso reflete numa baixa utilização das linhas de moldagem.

Nas próximas etapas da pesquisa, um método de solução específico para este problema,

baseado em metaheurísticas, deverá ser desenvolvido, bem como uma interface gráfica amigável de modo a viabilizar a implantação desta abordagem de solução na prática industrial.

Agradecimentos: Os autores agradecem o apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) e do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e ao revisor criterioso da revista *Investigação Operacional*.

7 Referências

- [1] Allahverdi, A., Gupta, J. N. D. e Aldowaisan, T. (1999). A Review of Scheduling Research Involving Setup Considerations. *Omega, International Journal of Management Science*, 27, 219-239.
- [2] Araujo, S. A. e Arenales, M. N. (2003). Dimensionamento de Lotes e Programação do Forno numa Fundição de Porte Médio. *Pesquisa Operacional*, v. 23, n. 3, p. 403-420
- [3] Araujo, S. A., Arenales, M. N. e Clark, A. R. (2003). Dimensionamento de Lotes e programação do Forno numa Fundição de pequeno porte. *Gestão & Produção*, v. 11, n. 2, p.165-176.
- [4] Araujo, S. A. (2003). *Modelos e Métodos para o Planejamento da Produção Aplicados no setor de Fundições*, Tese de Doutorado, ICMC-USP/SC.
- [5] Bahl, H. C., L. P Ritzman, e J. N. D Gupta (1987). Determining Lot Sizes and Resource Requirements: A Review. *Operations Research*, 35, 329-345.
- [6] Billington, P. J., McClain, J. O. e Thomas, L. J. (1983), Mathematical Programming Approaches to Capacity MRP Systems: Review, Formulation and Problem Reduction, *Management Science*, v. 29, n. 10, p. 1126-1141.
- [7] Bowers, M. R., Kaplan, L. A. e Hooker, T. L. (1995). A Two-Phase Model for Planning the Production of Aluminum Ingot. *European Journal of Operational Research*, 81, 105-114.
- [8] Cerny, V. (1985). A Thermodynamical Approach to The Traveling Salesman Problem: an Efficient Simulation Algorithm. *Journal of Optimization Theory and Applic.*, 45, 41-55.
- [9] Clark, A. R. e Clark, S. J. (2000), Rolling-Horizon Lot-Sizing When Setup Times are Sequence-Dependent, *International Journal of Production Research*, v. 38, n. 10, p. 2287-2307.
- [10] Drexel, A. e Kimms, A. (1997). Lot Sizing and Scheduling - Survey and Extensions. *European Journal of Operational Research*, 99, 221-235.
- [11] Fourer, R., Gay, D. M. e Kernighan B. W., (1993). *AMPL- A Modeling Language for Mathematical Programming*. Boyd and Frase, Danvers, Massachusetts. <http://www.ampl.com>.
- [12] Gonçalves-Vianna, A. C. e Arenales, M. (1995), Problema de Programação da Produção numa Fundição Automatizada, *Gestão & Produção*, v.2, p. 244-263.
- [13] Gravel, M., Price, W. L. e Gagné, C. (2000). Scheduling Jobs in an Alcan Aluminium Foundry Using a Genetic Algorithm. *International Journal of Production Research*, 38 (13), 3031-3041.
- [14] Goyal, S. K. e Gunasekaran, A. (1990), Multi-Stage Production-Inventory Systems, *European Journal of Operational Research*, v. 46, p. 1-20.
- [15] Hamada, K., Baba, T., Sato, K. e Yufu, M. (1995). Hybridizing a Genetic Algorithm with Rule-based Reasoning for production Planning. *IEEE Expert*, 10, 60-67.
- [16] Hendry, L. C., Fok, K. K. e Shek, K. W. (1996). A Cutting Stock and Scheduling Problem in the Copper Industry. *Journal of Operational Research Society*, 47 (1), 38-47.
- [17] ILOG CPLEX 7.1 User's Manual, (2001). ILOG S.A., BP 85, 9 Rue de Verdun, 94253 Gentilly Cedex, France. <http://www.ilog.com>.

- [18] Kuik, R., Salomom, M. e Van Wassenhove, L. N. (1994). Batching Decisions: Structure and Models. *European Journal of Operational Research*, 75, 243-263.
- [19] Lee, H. S., Murthy, S. S., Haider, S. W. e Morse, D. V. (1996). Primary Production Scheduling at Steelmaking Industries. *IBM Journal of Research and Development*, 40 (2), 231-252.
- [20] Lopes, L., Carter, M. W. e Gendreau, M. (1998). The Hot Strip Mill Production Scheduling Problem: a tabu Search Approach. *European Journal of Operational Research*, 106, 317-335.
- [21] Maes J. e van Wassenhove, L. N. (1988), Multi-Item Single-Level Capacitated Dynamic Lot-Sizing Heuristics: A General Review, *Journal of Operational Research Society*, v. 39, n. 11, p. 991-1004.
- [22] Petersen, C. M., Sorensen, K. L. e Vidal, R. V. V. (1992). Inter-process Synchronization in the Steel Production. *International Journal of production Research*, 30 (6), 1415-1425.
- [23] Potts, C. N. e Van Wassenhove, L. N. (1992). Integrating Scheduling with Batching and Lot-Sizing: a Review of Algorithms and Complexity. *Journal of Operational Research Society*, 43 (5), 395-406.
- [24] Potts, C. N. e Kovalyov, M. Y. (2000). Scheduling with Batching: a review. *European Journal of Operational Research*, 120, 228-249.
- [25] Santos-Meza, E., Santos, M. O. e Arenales, M. N. (2002). A Lot-Sizing Problem in an Automated Foundry. *European Journal of Operational Research*, 139 (2), 490-500.
- [26] Sounderpandian, J. e Balashanmugam, B., (1991). Multiproduct, Multifacility scheduling Using the Transportation Model: a Case Study. *Production and Inventory Management Journal*, 32 (4), 69-73.
- [27] Tang, L., Liu, J. Rong, A. e Yang, Z. (2000a). A Mathematical Programming Model for Scheduling Steelmaking-Continuous Casting Production. *European Journal of Operational Research*, 120, 423-435.
- [28] Tang, L., Liu, J. Rong, A. e Yang, Z. (2000b). A Multiple Traveling Salesman Problem Model for Hot Rolling Scheduling in Shanghai Baoshan Iron & Steel Complex. *European Journal of Operational Research*, 124, 267-282.
- [29] Tang, L., Liu, J. Rong, A. e Yang, Z. (2001). A Review of Planning and Scheduling Systems and Methods for Integrated Steel Production. *European Journal of Operational Research*, 133, 1-20.
- [30] ABIFA - Associação Brasileira de Fundição (2004). Relatório anual do setor de fundição, <http://www.abifa.org.br>.
- [31] Karimi B., Fatemi Ghomia S. M. T. e Wilson J. M. (2003). The capacitated lot sizing problem: A review of models and algorithms. *Omega*, v. 31, p. 365-378.

Programação de operações em sistemas com tempo de processamento variável e com buffers de preparação

José António Oliveira * M.^a Sameiro de Carvalho *
J. M. Valério de Carvalho *

* Departamento de Produção e Sistemas
Universidade do Minho
Braga - Portugal
{vc,sameiro,zan}@dps.uminho.pt

Abstract

In this work a scheduling operation problem with variable processing times is analysed. In particular this project is concerned with the scheduling of the truck load operations and the preparation of loads in buffers in an automatic warehouse. The truck load operations can be assumed as processing tasks with ready times (availability of the trucks) and due dates (trip cut off time). Additionally, it is possible to process a task before ready time by using a buffer where a truck load is assembled to be transferred to the truck in a later time period. Processing time is variable and depends upon the number of tasks simultaneously in progress. A minimum number of simultaneous tasks is assumed in order to guarantee a maximum, nominal, rate of expedition.

A procedure based on Dynamic Programming with truck expedition times ordered by EDD rule is presented. The algorithm efficiency is also discussed.

Resumo

Neste artigo apresenta-se um problema de planeamento de operações para a expedição de mercadorias em camiões num sistema automático de armazenamento. O carregamento de um camião pode ser entendido como o processamento de uma tarefa com data de disponibilidade e data de entrega. Os tempos de processamento das tarefas são variáveis e dependem do número de tarefas a serem processadas simultaneamente. Neste sistema de armazenamento existe a possibilidade de iniciar o processamento de uma tarefa antes da sua data de disponibilidade, utilizando para o efeito uma zona de preparação antecipada, que substitui temporariamente o camião. Diariamente, dado um conjunto de tarefas a processar, pretende-se elaborar o plano operacional para a sua execução de modo a minimizar a soma dos atrasos na expedição.

Apresenta-se um procedimento baseado em programação dinâmica para a definição do plano operacional relativo à expedição dos camiões com o objectivo de minimizar a soma dos atrasos, estudando-se a sua eficiência computacional.

Keywords: Logistics, Scheduling, Warehousing.

Title: Scheduling variable processing time operations in systems with buffers.

1 Introdução

Os modernos conceitos de armazenagem visam sincronizar os elementos da cadeia logística, proporcionando todo um conjunto acrescido de operações, tais como: a consolidação de encomendas (preparação de encomendas para clientes com produtos provenientes de vários produtores); “cross-docking” (operação de preparação de encomendas mas que não envolve armazenamento dos produtos); processamento (operações que acrescentam valor aos produtos; ex. embalagem), etc. Benefícios económicos resultam, em cada caso, de vários factores: diminuições nos custos de transporte, reduções nos custos de armazenagem e manuseamento dos produtos, menores níveis de stock, etc.

Os sistemas automáticos de armazenamento facilitam as acções de logística. Estes sistemas devem ser eficientes para justificar o investimento que envolvem e constituírem uma alternativa aos sistemas convencionais de armazenamento. A eficiência de um sistema automático de armazenamento depende, entre outros factores, do planeamento das operações de recepção e expedição de mercadorias.

O planeamento de operações deve ser efectuado de um modo cuidado de forma a garantir que os camiões partem do armazém antes das datas limites, e serem capazes de entregar as encomendas nos clientes sem atrasos, principalmente no caso dos armazéns automáticos que funcionam com (elevadas) taxas de expedição de mercadoria, próximas do valor limite para as quais foram dimensionados. Neste artigo, apresenta-se um algoritmo para a programação das operações associadas à expedição de mercadorias, com o objectivo de minimizar a soma dos atrasos. Em particular, estuda-se o planeamento das operações do sistema automático de colocação e recolha de paletes do armazém. O sistema automático de colocação e recolha de paletes é controlado pelo Warehouse Management System (WMS). Além do controlo do sistema automático, o WMS decide as posições de armazenamento das paletes, escolhe a sequência da recolha das paletes e determina também o movimento dos empilhadores.

Tipicamente, neste tipo de sistemas de armazenagem, é o supervisor que define o plano de operações, selecciona as cargas / viagens que são preparadas a cada momento. De acordo com o nosso melhor conhecimento, esta selecção de cargas é usualmente baseada em simples regras de despacho, sendo depois a eficiência das regras avaliada usando modelos de simulação Elsayed e Stern [2], Gibson e Sharp [4] e Randhawa e Shroff [6]. Neste trabalho desenvolve-se uma estratégia para o sequenciamento das cargas e um procedimento baseado em Programação Dinâmica, que gera soluções onde as viagens são agrupadas em blocos de processamento simultâneo. Estes modelos são inseridos no Warehouse Programming System (WPS), que determina os tempos de início das operações de carregamento de cada carga. O supervisor insere essa informação no WMS para o funcionamento do sistema. Obviamente, os planos obtidos devem ser tais que o WMS, seguindo as suas próprias regras, possa executar o plano dentro dos tempos previstos. O WMS e o WPS foram integrados num sistema desenvolvido pela empresa EFACEC Automação e Robótica, S. A. em 1998 para um centro de distribuição.

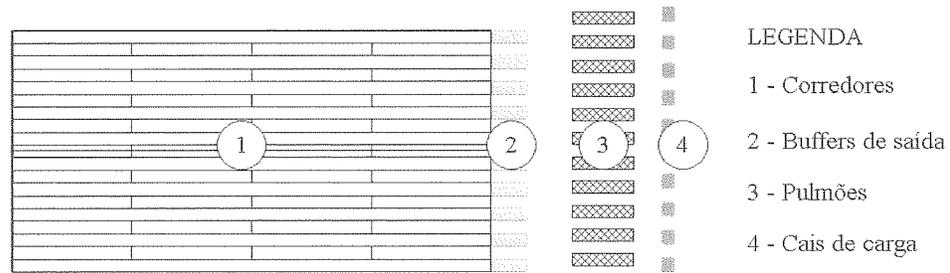


Figura 1: Layout do armazém

O artigo está organizado do seguinte modo: na secção seguinte, descreve-se o funcionamento de um armazém automático e identifica-se o problema de planeamento; a terceira secção é dedicada à caracterização do problema, na quarta secção procede-se à construção do modelo para este problema de planeamento; a quinta secção é dedicada ao algoritmo desenvolvido, baseado em programação dinâmica e à determinação da sua eficiência; por fim, discutem-se as conclusões sobre este trabalho.

2 Descrição do sistema de armazenamento

O sistema de armazenamento tem capacidade para armazenar cerca de 38 mil paletes em alvéolos. As paletes, com os artigos encomendados pelos clientes, são expedidas por camião. A viagem de um camião consiste no transporte de um conjunto de cerca de três dezenas de paletes. Os camiões disponíveis para efectuar o transporte das paletes têm idêntica capacidade. O conjunto de paletes que formam a carga de uma dada viagem é conhecido e fixo.

Na Figura 1 apresenta-se um esquema da planta do armazém. O corpo principal do armazém é constituído por onze corredores. Em cada corredor existe um “*stockador*” automático para movimentar as paletes desde os alvéolos até ao “*buffer de saída*”, colocado no topo do corredor. A jusante dos corredores, as paletes são movimentadas por empilhadores. Existe um conjunto de n_e empilhadores que recolhe as paletes nos buffers de saída dos corredores e as coloca nos camiões ou nos pulmões, conforme o modo de preparação da carga.

A preparação (da carga) de uma viagem significa a recolha, movimentação e formação do conjunto de paletes, que será transportado no mesmo camião. Uma viagem pode ser preparada de um de dois modos diferentes, relativamente à movimentação e formação do seu conjunto de paletes. No modo directo as paletes são agrupadas dentro do camião, através da movimentação das paletes, pelos empilhadores, desde o buffer de saída do corredor até ao camião. Abreviadamente diz-se que a viagem é preparada com uma operação (b-c). Alternativamente, no modo indirecto, as paletes são agrupadas numa área, que se designa por “pulmão”, através da movimentação das paletes, pelos empilhadores, desde o buffer de saída do corredor até ao pulmão. Abreviadamente diz-se que a viagem é preparada com uma operação (b-p). Posteriormente, o conjunto de paletes é transferido do pulmão para o camião numa operação que designamos por (p-c) e que é executada por um empilhador de transferência e_{pc} . Existe um conjunto de $n_{e_{pc}}$. A operação de transferência das paletes do pulmão para o camião só pode ser iniciada depois de a viagem ter sido completamente preparada no pulmão e, uma vez iniciada, não pode ser interrompida até estar terminada.

Os pulmões são áreas com capacidade equivalente à dos camiões, nas quais as viagens podem ser preparadas numa data anterior à chegada dos camiões ao armazém. Define-se data de disponibilidade da viagem como a data da chegada do camião ao armazém. A existência de pulmões permite relaxar as datas de disponibilidade das viagens que são preparadas nos pulmões.

Essencialmente a utilização de um pulmão permite a “armazenagem” de capacidade de processamento de alturas de menor solicitação para períodos onde o trabalho de expedição é mais intenso. A operação (p-c) é mais rápida do que a operação (b-c), o que permite aumentar a taxa de expedição de camiões, facilitando o cumprimento das datas de entrega. A expedição, sem atrasos, de um elevado número de viagens numa janela temporal pequena (definida pelas datas de chegada dos camiões e pelas datas de entrega) só poderá eventualmente ser possível se parte dessa procura for satisfeita através da transferência de cargas previamente preparadas em pulmões. Outra vantagem das preparações em pulmões é que o sistema pode funcionar de um modo mais equilibrado quanto ao número de processamentos de viagens por unidade de tempo.

O sistema de armazenamento deve ter uma taxa nominal de funcionamento μ_s que deve ser a mais elevada possível, para tornar eficiente a sua utilização e justificar o seu investimento. No sistema em análise, em termos de saídas, a capacidade prevista de manuseamento dos 11 stockadores é constante e igual a $\mu_m = 240$ paletes/hora.

Seja μ_e a capacidade de manuseamento de um empilhador expressa em paletes/hora. A capacidade de movimentar paletes dos n_e empilhadores é dada por $n_e\mu_e$, assumindo uma relação linear o número de empilhadores e a taxa de serviço. Os empilhadores devem ser em número suficiente de modo a garantir o funcionamento do sistema à taxa de expedição de paletes definida pelos stockadores. Se tal não acontecer, os empilhadores não são capazes de remover as paletes que os stockadores colocam nos buffers de saída, levando à imobilização dos stockadores quando os respectivos buffers de saída estão cheios. De facto, os stockadores e os empilhadores formam um sistema em série em que a capacidade do sistema é dada pela menor das duas capacidades

$$\mu_s = \min(\mu_m, n_e\mu_e) \quad (1)$$

No que se segue, vamos assumir que os n_e empilhadores existentes no sistema igualam a capacidade dos stockadores em termos de movimentação de paletes.

$$n_e\mu_e = \mu_m \quad (2)$$

Este sistema de armazenamento funciona continuamente. No entanto a expedição de viagens é feita entre as 6 horas e as 24 horas. O período entre as 0 horas e as 6 horas pode ser usado para preparar viagens em pulmão, reorganizar as paletes nos alvéolos e proceder a operações de entrada de mercadorias.

Existem 13 cais de carga, $N_c = 13$, onde os camiões são carregados, quer recebendo paletes directamente do corredor, quer do pulmão. Existem 21 pulmões para a preparação antecipada de viagens, $N_p = 21$.

O número de viagens expedidas por dia é da ordem de uma centena, tendo todas as viagens igual importância. Cada viagem tem um instante de saída do armazém, o qual deve ser cumprido planeando correctamente a hora de início de preparação da viagem, e ainda o

início de transferência de pulmão para camião, no caso das preparações feitas em pulmão. Como princípio, a programação não deve admitir atrasos. O plano operacional do armazém deve entrar em linha de conta com todos estes factores e minimizar a soma dos atrasos gerados pelas viagens que são concluídas após a hora limite de saída.

3 Características do problema

O estabelecimento de um plano de operações requer a definição do instante de início de preparação (da carga) de uma viagem, que pode ser uma preparação em camião ou num pulmão. Para as viagens preparadas em pulmão é necessário definir também o instante de início de transferência do pulmão para o camião.

Para auxiliar na construção do modelo para este problema de planeamento definiram-se as seguintes variáveis, que seguem de perto a notação usada em [1]:

v_j – designa a viagem j . Neste trabalho utiliza-se o termo “viagem” para denotar o mesmo que o termo “tarefa”, utilizado na literatura de planeamento da produção.

r_j – data de disponibilidade de v_j . Define o instante de chegada ao sistema do camião que transporta v_j . As operações (b-c) e (p-c) só podem ser iniciadas depois deste instante.

d_j – data de entrega de v_j . Define o instante em que o camião deve abandonar o sistema sem haver atraso.

t_{pc} – tempo de processamento de uma operação (p-c). Tem um valor fixo, porque é assumido que as viagens têm igual número de paletes.

s_j – instante em que se inicia a preparação de v_j , que pode ser com as operações (b-c) ou (b-p).

st_j – instante em que se inicia a operação (p-c), se v_j tiver sido preparada em pulmão.

B_j – instante em que termina a preparação de v_j , quer tenha sido preparada com operações (b-c) ou (b-p).

C_j – instante em que v_j sai do sistema de armazenamento.

Uma viagem diz-se “activa”, quando é iniciada a sua preparação em s_j , e permanece nesse estado até ao instante B_j em que recebe a última paleta. Por outro lado, uma viagem diz-se “aberta”, quando é iniciada a sua preparação em s_j , e permanece nesse estado até ao instante C_j , em que se conclui o carregamento do camião. O estado activo e o estado aberto são coincidentes no caso das preparações em modo directo, pois $B_j \equiv C_j$. Nas preparações em modo indirecto, o estado activo termina com a conclusão da recolha das paletes no pulmão, no instante B_j , mas a tarefa permanece aberta até ao fim da transferência do pulmão para o camião que ocorre no instante posterior C_j . Para qualquer tarefa verifica-se $B_j \leq C_j$.

A demora de uma viagem D_j é uma variável que toma valores não negativos e é medida relativamente ao instante de saída do sistema, $D_j = \max(C_j - d_j, 0)$.

A movimentação das paletes pelos stockadores e pelos empilhadores pode ser modelada como uma fonte que assegura um fluxo de paletes, que é constante ao longo do tempo. A capacidade de movimentação de paletes é medida em paletes por unidade de tempo. Iremos assumir que, em cada instante, esta capacidade é partilhada equitativamente pelas viagens que estão activas.

Um ponto crucial do modelo é que os tempos de processamento das viagens são dependentes

do número de viagens a serem processadas simultaneamente, isto é, o tempo necessário para completar uma viagem depende do número de viagens activas. Se o número de viagens activas for mais pequeno, mais rápida será a sua preparação.

De modo a facilitar o entendimento sobre os tempos de processamento vamos considerar um tempo nominal de processamento de uma viagem que teoricamente corresponderia à preparação da viagem com a totalidade do fluxo de paletes proveniente dos corredores, ou seja, corresponderia à preparação de uma única viagem à taxa $\mu_m = 240$ paletes/hora. Designa-se por tempo nominal de processamento de uma viagem p_j , o tempo teórico que demoraria a preparar a carga de uma viagem se fosse possível essa viagem receber todo o fluxo de paletes à taxa nominal de funcionamento do sistema, e independente de se tratar uma preparação num pulmão ou num camião. Dado que o número de paletes por viagem é aproximadamente constante, iremos considerar que o tempo nominal de processamento é igual para todas as viagens, verificando-se a relação $p_j = p, \forall j$.

Há, no entanto, que impor um limite mínimo ao número de viagens activas, N_{\min} . Se essa restrição não existir, pode não ser possível assegurar o fluxo nominal de funcionamento do sistema. Para ilustrar este ponto, vamos considerar a situação limite em que todo o sistema está apenas a executar uma viagem. Haverá um grande número de empilhadores a fornecer paletes para a mesma viagem, o que dá origem a congestionamentos, que se traduzem numa redução da capacidade do sistema, que passa efectivamente a funcionar abaixo da sua capacidade nominal. Na realidade é necessário haver um conjunto mínimo de viagens a serem processadas simultaneamente para escoarem o fluxo de paletes vindas dos corredores. Considera-se que $N_{\min} = 6$ é um valor razoável.

O número de pulmões e de cais disponíveis definem um limite superior ao número de viagens abertas, que se representa por N_{\max} .

Caso particular:

Um caso especial deste problema, em que não existem buffers de preparação $N_p = 0$, foi formulado matematicamente por Valério de Carvalho et al. [7], utilizando um conjunto de variáveis indexadas ao tempo (*time indexed variables*). A formulação apresentada “organiza” a distribuição das paletes disponíveis em cada intervalo de tempo t pelas diferentes viagens que estão activas nesse intervalo de tempo. A solução indica os instantes de activação das viagens, e os intervalos de tempo em que recebem as paletes, que minimizam a soma das penalidades das demoras para uma versão muito geral do problema, em que se consideram datas de disponibilidade e de entrega, e em que se permite que haja diferentes valores para as quantidades de paletes de cada viagem.

Valério de Carvalho et al. [7] mostraram que o problema de decidir se existe um plano válido obedecendo às datas de disponibilidade e de entrega, na versão do problema sem buffers de preparação, $N_p = 0$, pertence à classe de problemas NP-completos.

4 Uma metodologia de solução

Tratando-se de uma generalização de um problema NP-completo, e não se esperando que possa haver algoritmos eficientes para a sua solução, foi decidido usar uma estratégia de solução, baseada em programação dinâmica, com o objectivo de minimizar a soma das demoras das

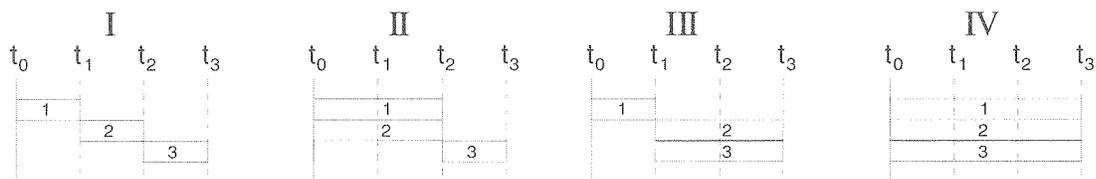


Figura 2: Diferentes modos de processar três viagens

viagens. Neste algoritmo, a ordenação da expedição das viagens é feita pela regra Earliest Due Date (EDD), ou seja, as viagens são sempre ordenadas por valores crescentes de data de entrega, $d_1 \leq d_2 \leq \dots \leq d_n$, e expedidas por essa ordem. A estratégia adoptada consiste em definir uma partição do conjunto de viagens em subconjuntos disjuntos de viagens que estão activas simultaneamente. Estes subconjuntos de viagens são designados por blocos. São também respeitados os limites para o número de viagens activas e para o número de viagens abertas, respectivamente N_{min} e N_{max} .

4.1 Datas de disponibilidade

As datas de chegada dos camiões são estabelecidas pelos responsáveis da gestão do armazém. Assim, e de acordo com o plano de preparação das viagens, é solicitado diariamente à empresa de transporte que aluga os camiões, um conjunto de camiões para efectuar o transportes das viagens desse dia, o que permite assumir que as datas de disponibilidade dos camiões são consistentes com as datas de entrega, verificando-se a relação $r_1 \leq r_2 \leq \dots \leq r_n$.

4.2 Tempos variáveis de processamento

Tal como foi referido na Secção anterior, os tempos de preparação de cada viagem são variáveis e dependem do número de viagens em processo simultaneamente. O esquema da Figura 2 apresenta 4 modos diferentes de planear a preparação de três viagens, com iguais tempos nominais de preparação. No caso I, ilustra-se o processamento individual de cada viagem. Nos casos II e III duas viagens são preparadas simultaneamente e no último caso as três viagens são processadas simultaneamente.

No entanto, salienta-se o facto de a viagem 3 ter sempre a sua conclusão no instante t_3 . Este facto resulta de a expedição acumulada de 3 viagens equivaler a uma quantidade (fixa) de paletes que são processadas pelo sistema a uma taxa que é igual nos 4 casos.

4.3 Definição dos blocos

Foi aceite que as viagens deveriam ser sequenciadas pela regra EDD. No modelo que vamos usar para justificar o uso desta regra, iremos considerar que todas as transferências pulmão-cais a executar no bloco podem ser feitas simultaneamente, mesmo que o número de empilhadores disponíveis para o efeito possa ser ultrapassado. Esta simplificação foi considerada aceitável.

Após o fim de um bloco, podemos caracterizar o estado do sistema pelo número de viagens

já saídas e pelo número de preparações disponíveis em pulmão. Vamos considerar que o bloco que a seguir se inicia é um bloco válido, em que o número de operações de cada um dos tipos, (p-c), (b-c) e (b-p), obedece às restrições impostas, em particular às datas de disponibilidade.

O número de operações de cada tipo permite identificar inequivocamente a quantidade de viagens activas e de viagens saídas no bloco. Todas as viagens activas têm os mesmos instantes de início e de fim, coincidindo com os instantes de início e de fim do bloco, e a duração do bloco depende apenas do número de viagens activas.

Em primeiro lugar, vamos definir a ordem relativa das operações (b-c) e (b-p). Dado que as datas de disponibilidade são consistentes, trocar a ordem EDD de duas viagens em operação (b-c) não traz melhoria ao valor da função objectivo. Portanto, todas as viagens em operação (b-c) devem seguir a regra EDD. Por outro lado, as viagens a que as operações (b-p) dizem respeito sairão em blocos subsequentes, pelo que corresponderão a viagens de maior índice segundo a ordenação EDD.

Em segundo lugar, vamos definir a ordem das operações (b-c) e (p-c) relativas às viagens que saem no bloco. O número de operações de cada tipo é bem definido. O estado do sistema no início do bloco permite identificar o número de preparações disponíveis em pulmão. Em blocos com operações (p-c) e (b-c), uma operação (p-c) ou deve dizer respeito a uma viagem de menor índice que todas as viagens em operação (b-c), ou a uma viagem de maior índice que todas as viagens em operação (b-c), não podendo essa preparação ser atribuída a uma viagem de índice intermédio, pelas razões que se expõem de seguida.

Vamos considerar os índices das viagens que resultam da sua ordenação EDD. Vamos supor que uma operação (p-c) era atribuída a uma viagem com índice EDD intermédio. Da troca dessa viagem por uma outra viagem processada em operação (b-c) resultariam sempre vantagens; a troca deveria ser feita com uma viagem de menor ou maior índice EDD, conforme o caso, como se analisa seguidamente.

O uso da operação (p-c) para uma viagem de menor índice EDD poderia diminuir a soma dos atrasos, por diminuir o atraso ou ajudar a respeitar a data de entrega da viagem de menor índice. Relembra-se que todas as viagens em operação (b-c) continuariam a ter o mesmo instante de início e de fim de processamento, em particular as que precedem e as que sucedem à viagem que correspondia à operação (p-c) antes da troca. Dado que as datas de disponibilidade são consistentes, se, antes da troca, todas as datas de disponibilidade eram respeitadas, após a troca, continuariam a sê-lo. A Figura 3 ilustra esta situação.

Por outro lado, pode ser necessário atribuir a operação (p-c) à viagem com maior índice segundo a ordem EDD para ser possível executar o bloco em análise. Isso acontece no caso em que há uma viagem que sai no bloco com uma data de disponibilidade posterior à data de início do bloco. A atribuição da operação (p-c) a uma viagem com índice intermédio poderia implicar a execução dessa viagem em operação (b-c) no final do bloco, o que não seria possível por não respeitar a data de disponibilidade. Neste caso, só mesmo a atribuição da operação (p-c) à viagem de maior índice EDD é que pode viabilizar a saída do número correcto de viagens. A Figura 4 ilustra esta a situação.

Esta opção de executar algumas transferências pulmão cais no final das viagens saídas no bloco só precisa de ser usada no caso de haver viagens no bloco com uma data de disponibilidade superior ao instante de início do bloco. Se todas as viagens do bloco puderem ser iniciadas no

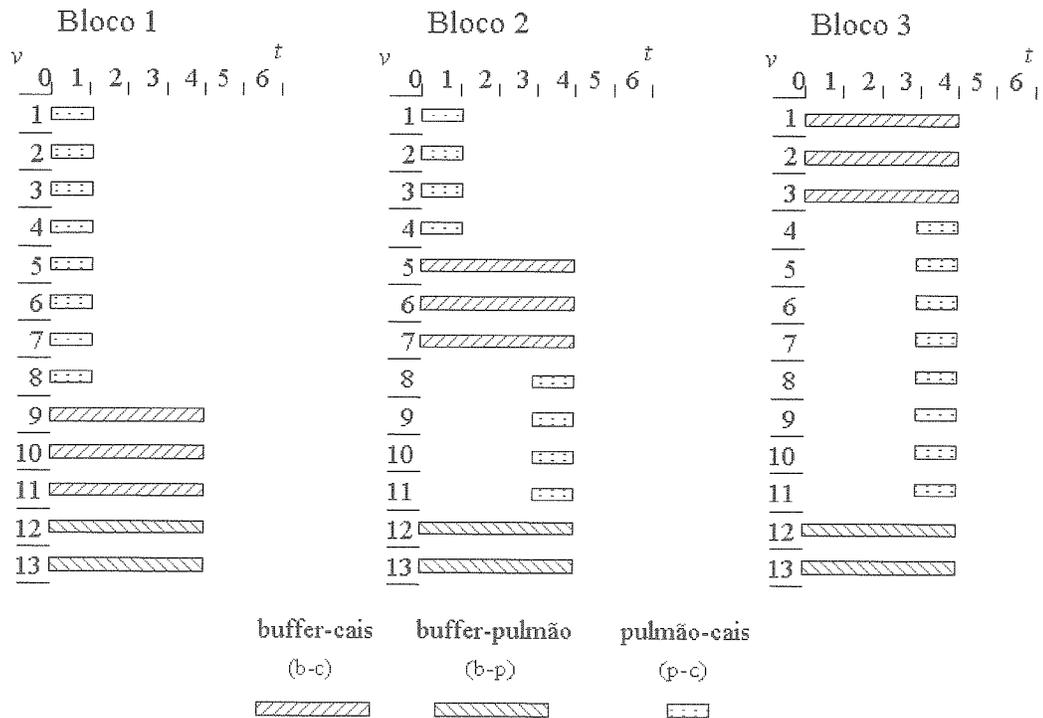


Figura 5: Modos de constituição de um bloco com operações (p-c), (b-c) e (b-p)

instante de início do bloco, a solução em que as operações (p-c) são atribuídas às viagens de menor índice EDD é uma solução que domina todas as alternativas, porque respeita as datas de disponibilidade e tem claramente menor soma de atrasos.

Esta estratégia de construção de blocos permite um melhor ajuste a diferentes configurações de linhas de datas de disponibilidade e de datas de entrega. Na Figura 5, são mostradas três situações em que a ordem das operações (p-c) e (b-c) permite responder a configurações algo diferenciadas.

4.4 Tempo de conclusão das viagens

Neste sistema, assumimos que os tempos reais de conclusão de todas as viagens do bloco coincidem com o instante de tempo em que o bloco termina, embora isso possa não acontecer na realidade para todas as viagens, estando garantido que terminam sempre antes do fim do bloco como se passa a explicar.

Um dos parâmetros que o sistema WMS utiliza para definir a sequência da recolha de paletes é a data de entrega das viagens que estão activas. As paletes pertencentes às viagens com menores datas de entrega têm maior prioridade e podem ser recolhidas em primeiro lugar. Face à sequência de recolha de paletes definida pelo WMS é possível que algumas viagens do bloco estejam terminadas antes do fim do bloco, quando lhes é entregue a última paleta da viagem. Tipicamente, o perfil para as viagens se completarem é ilustrado na Figura 6, em que as viagens são terminadas na sequência v_1, v_3, v_2, v_4, v_5 , coincidindo o fim do bloco com o fim da viagem 5. Dado não ser possível controlar a sequência da recolha das paletes, escolhemos

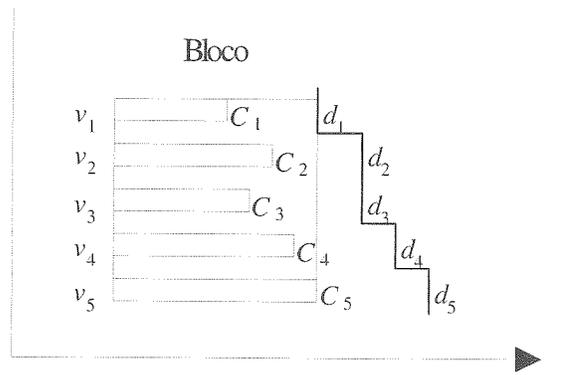


Figura 6: Perfil típico de conclusão das viagens de um bloco

uma opção conservadora, que garante que as datas de conclusão previstas no plano nunca serão ultrapassadas na execução real do plano: prevê-se que o instante de conclusão de todas as viagens do bloco seja o instante em que o bloco termina.

4.5 Duração de processamento dos blocos

A estratégia adoptada consiste em definir uma partição do conjunto de viagens em blocos de processamento, que podem ser constituídos pelos três tipos diferentes de operações de movimentos de paletes e em quantidades que são variáveis.

Considere-se a seguinte notação:

- α – quantidade de operações (b-c). Representa o número de preparações em camião;
- β – número de operações (p-c). Representa a quantidade de viagens preparadas em pulmão que são transferidas para o camião;
- δ – quantidade de operações (b-p). Representa o número de preparações em pulmão.

Vamos designar o conjunto de viagens tratadas no bloco por E . A duração de processamento do bloco, P_E , depende somente do número de preparações do bloco, isto é, depende do número $\alpha + \delta$ de viagens activas. A duração de processamento do bloco é uma função linear da dimensão do próprio bloco, $P_E(\alpha, \beta, \delta)$. Quanto maior for o número de preparações de um bloco, maior é o tempo de processamento das operações, porque o fluxo de paletes (com um valor constante por unidade de tempo) é dividido por um número maior de viagens.

Considere-se um bloco com α operações (b-c), β operações (p-c) e δ operações (b-p). A Figura 7 ilustra a duração de processamento do bloco formado pelas $\alpha + \delta$ de viagens activas.

Seja l o número de paletes por viagem. Num bloco com $\alpha + \delta$ operações, há um total de $l(\alpha + \delta)$ paletes a processar. A expressão que relaciona o número de paletes e o tempo P_E é a seguinte:

$$l(\alpha + \delta) = \int_0^{P_E} \mu_s dt = n_e \mu_e P_E \tag{3}$$

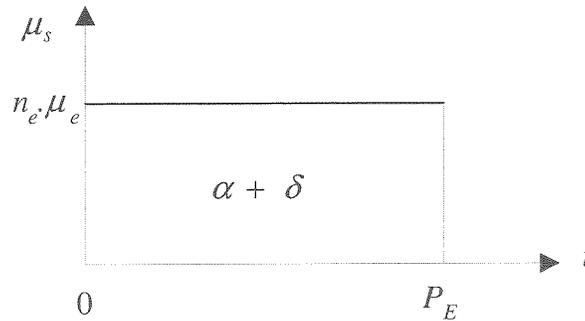


Figura 7: Duração do processamento de um bloco

A duração do bloco é dada por:

$$P_E = \frac{l}{n_e \mu_e} (\alpha + \delta) \tag{4}$$

Do ponto de vista prático, um plano pode não ser executável quando o número de operações (p-c) que é necessário fazer em simultâneo excede o número de empilhadores. No entanto, foi aceite que a execução das transferências pulmão-cais em grupos sucessivos constituía uma aproximação aceitável.

Seja s_{P_E} o instante de início de processamento do bloco. Dado que os tempos nominais de preparação são iguais, verifica-se que todas as viagens activas do bloco, ou sejam, as operações (b-c) e (b-p), têm o mesmo instante de início, $s_j = s_{P_E}$, e o mesmo instante de conclusão de preparação:

$$B_j = s_{P_E} + P_E \tag{5}$$

O número máximo de transferências possíveis em simultâneo é limitado pelo número de empilhadores “transferência” ne_{pc} . Caso se pretenda fazer um maior número de transferências, os tempos C_j do segundo e seguintes grupos de transferências serão maiores, porque as β operações (p-c) são divididas em $S + 1$ grupos, isto é:

$$\beta = S \cdot ne_{pc} + R, \quad S = 0, 1, 2, \dots \tag{6}$$

em que R representa o resto da divisão inteira de β por ne_{pc} .

Vamos considerar a transferência de β viagens, em que as transferências são divididas em grupos, da forma indicada na expressão (6). É possível acomodar num bloco qualquer número de transferências, menor que N_p , porque é possível determinar inequivocamente os tempos de conclusão. As transferências são divididas em grupos de dimensão do número de empilhadores transferência. No primeiro grupo as operações (p-c) são iniciadas no instante $st_j = s_{P_E}$, e concluídas no instante:

$$C_j = s_{P_E} + t_{pc} \tag{7}$$

Nos grupos seguintes, as operações (p-c) são iniciadas no instante da conclusão do grupo

anterior e concluídas nos instantes:

$$C_j = s_{PE} + kt_{pc}, \quad k = 2, 3, \dots \quad (8)$$

em que k representa a ordem em que o grupo é transferido. Nas operações (p-c) que são executadas no final do bloco, os tempos podem ser definidos de forma análoga.

As soluções do problema baseiam-se no dimensionamento dos vários blocos que devem respeitar os valores n_e , N_{\min} e N_{\max} . A definição dos conjuntos de viagens que devem ser transferidas e ou preparadas em paralelo, de modo a que os tempos de conclusão das respectivas preparações respeitem as horas limites de expedição das cargas, responde adequadamente ao problema operacional existente. Se tal não for possível, é necessário escolher a alternativa que minimiza a soma dos atrasos verificados nas horas de partidas.

5 Modelo de programação dinâmica

Em termos gerais, a solução deste problema consiste em determinar uma partição do conjunto das viagens em subconjuntos, tal que seja minimizada a soma dos atrasos.

5.1 Definição do modelo

Para efeitos de apresentação de um modelo de programação dinâmica, considere-se a seguinte notação:

- y – representa (num dado instante) o número de viagens já expedidas que podem ter sido através de operações (p-c) ou através de operações (b-c);
- z – representa (num dado instante) o número de viagens preparadas disponíveis em pulmão.

Os estados do modelo de programação dinâmica serão identificados pelo par ordenado $e_i = (y_i, z_i)$, que permite identificar todos os estados possíveis do sistema.

O estado inicial do modelo é definido para o instante em que se dá início ao período de expedição de viagens. Em princípio, o número de viagens preparadas antecipadamente deve esgotar os pulmões disponíveis, isto é $z_{ini} = N_p$, porque a transferência de pulmão para cais é uma operação mais rápida e que permite concluir as viagens em menor tempo, diminuindo os atrasos. Assim, o estado inicial corresponde a $e_0 = (0, z_{ini})$.

Para um dado estado e_i , é possível, com base na informação dada pelos elementos do par ordenado, calcular o número de viagens já processadas desde o estado inicial. Dado que as viagens estão ordenadas pela regra EDD, o estado e_i significa que as primeiras $y_i + z_i - z_{ini}$ viagens já foram processadas. Isto permite identificar o conjunto E_i de viagens já processadas, sendo $|E_i| = y_i + z_i - z_{ini}$.

As viagens do conjunto E_i representam uma quantidade bem definida de paletes. Atendendo à quantidade de paletes e à taxa de expedição de paletes, é possível estabelecer uma relação biunívoca entre o número de viagens processadas e o tempo de processamento. A expressão do tempo de preparação associado a um dado estado e_i pode ser derivada de uma forma semelhante à da Expressão (4):

$$t(e_i) = \frac{l}{n_e \mu_e} (y_i + z_i - z_{ini}) \quad (9)$$

Trata-se de uma função linear do número acumulado de viagens já preparadas. É de salientar que, para um dado estado e_i , este tempo é independente da forma como as viagens do conjunto E_i foram subdivididas e agrupadas nos blocos, conforme se ilustra no exemplo da Figura 2.

Havendo um tempo bem determinado associado a cada estado, é possível, ao enumerar as acções alternativas associadas ao estado, avaliar os atrasos que decorrem da expedição das viagens do bloco, e directamente calcular as contribuições de estágio, verificando uma condição fundamental para a aplicação da programação dinâmica [3].

Em cada estágio, para cada estado e_i , é necessário fazer o dimensionamento do próximo bloco de viagens. Seja Δ o conjunto de viagens do próximo bloco, isto é, o próximo bloco é constituído pelo processamento simultâneo de $|\Delta| = \alpha + \beta + \delta$ operações.

Seja Δ^{out} o subconjunto de Δ formado pelas viagens que são expeditas no bloco, $|\Delta^{out}| = \alpha + \beta$. Seja Δ^{act} o subconjunto de Δ formado pelas viagens activas neste bloco $|\Delta^{act}| = \alpha + \delta$.

A transição num estágio de um estado e_i para o estado e_j resulta do planeamento de um bloco formado por Δ viagens, $E_j = E_i \cup \Delta^{act}$ e tem associada uma contribuição de estágio $\gamma_{e_i}(\Delta)$. Esta acção conduz do estado $e_i = (y_i, z_i)$ ao estado $e_j = (y_i + \alpha + \beta, z_i - \beta + \delta)$.

Dado que, ao estado e_i está associado um tempo bem determinado e podemos calcular a duração do bloco, é possível determinar os instantes de conclusão das viagens de um bloco, para operações (b-c) e (p-c), que podem ser calculados de acordo com as Equações (5) e (7), respectivamente. Assim, a cada uma das acções alternativas, podemos associar um custo $\gamma_{e_i}(\Delta)$, dado pela soma dos atrasos respeitantes à expedição das Δ^{out} viagens que formam o bloco, ou seja, $\gamma_{e_i}(\Delta) = \sum_{j \in \Delta^{out}} D_j = \sum_{j \in \Delta^{out}} \max\{C_j - d_j, 0\}$.

A cada estado está associada uma função, Valor de Estado, $\Gamma(e_i)$, que representa o menor custo de planear as viagens representadas pelo estado. O Valor de Estado corresponde ao valor óptimo para resolver um problema reduzido, formado pelas viagens do conjunto E_i associado ao estado.

Dado que cada estado resulta de uma transição de um estado anterior, e se, para todas as possibilidades que conduzem ao estado e_j , escolhermos a que apresenta a menor soma, em termos de valor do estado anterior e da contribuição associado, obtém-se o menor custo para se planear as tarefas de E_j , isto é, $\Gamma(E_j) = \min_i \{\Gamma(E_i) + \gamma(\Delta_i)\}$, em que $E_j = E_i \cup \Delta_i$.

A validade da programação dinâmica requer a verificação das condições de separabilidade e optimalidade. A condição de separabilidade corresponde à validade de utilização de uma relação de recorrência. A relação de recorrência permite que o valor de cada estágio possa ser calculado em função do estágio antecedente e da contribuição de estágio.

$$\Gamma(\emptyset) = 0 \quad (10)$$

$$\Gamma(E_j) = \min_i \left\{ \sum_{j \in E_i} \max\{C_j - d_j, 0\} + \sum_{j \in \Delta_i^{out}} \max\{C_j - d_j, 0\} \right\}, \forall \Delta_i \quad (11)$$

$$\Gamma(E_j) = \min_i \left\{ \Gamma(E_i) + \sum_{j \in \Delta_i^{out}} \max\{C_j - d_j, 0\} \right\}, \forall \Delta_i \quad (12)$$

A condição de optimalidade implica a validade da hipótese de que a optimização global resulta da optimização em cada estágio. Na verdade, não pode haver outra forma de planear as tarefas de E_i com um custo inferior, senão seria possível diminuir o valor de $\Gamma(E_i)$ e dessa forma diminuir a soma, ou seja, diminuir o valor de $\Gamma(E_j)$.

A existência da relação de recorrência (12), que permite determinar a solução óptima, resulta de serem verificadas as condições de separabilidade e de optimalidade.

Exemplo:

Vamos considerar uma instância com 12 viagens, em que o tempo nominal de processamento de uma viagem é $p_j = 1$ viagem/u.t., $\forall j$. Os limites para a dimensão dos blocos são, respectivamente, $N_{min} = 3$ e $N_{max} = 4$. As datas de disponibilidade dos camiões e as datas de entrega das viagens são as seguintes:

v_j	1	2	3	4	5	6	7	8	9	10	11	12
r_j	0	0	0	0	0	2	2	4	4	4	4	4
d_j	2	3	4	5	7	7	9	9	13	13	13	13

Vamos considerar que o sistema tem 5 cais, que existe apenas 1 pulmão, e que esse pulmão tinha sido preparado antes do início do plano. O sistema possui apenas um empilhador “transferência” que executa uma transferência pulmão-cais em 0.4 u.t..

As regras impostas ao funcionamento do sistema podem ser expressas através das seguintes restrições:

a) A expressão que limita o número de viagens activas em cada bloco $N_{min} \leq |\Delta^{act}| \leq N_{max}$ condiciona o número conjunto de preparações em pulmão e movimentos buffer cais, ou seja, $3 \leq \alpha + \delta \leq 4$;

b) Foi assumido que não era possível descarregar um pulmão e prepará-lo no mesmo bloco. No caso em análise, em que existe apenas 1 pulmão, essa restrição traduz-se em $\beta + \delta \leq 1$;

c) O número de cais disponíveis limita o número de viagens que podem sair em cada bloco. Como $N_c = 5$, temos $|\Delta^{out}| \leq 5$, ou seja, $\alpha + \beta \leq 5$.

Estamos em condições de definir os estados possíveis com base na definição dos blocos possíveis. Cada bloco é definido por um triploto (α, β, δ) obedecendo às restrições enunciadas. Dado haver apenas um pulmão, os estados são definidos por $(y_i, 0)$ ou $(y_i, 1)$. Para melhor especificar as transições possíveis, vamos analisar separadamente as decisões alternativas a partir de estados em que o pulmão está cheio e em que está vazio.

Vamos considerar em primeiro lugar um estado em que o pulmão está cheio, o que implica

que não é permitido fazer preparações em pulmão, ou seja, $\delta = 0$. As transições possíveis são as seguintes:

Estado inicial	(α, β, δ)	Estado final	Número viagens saídas	Duração do bloco
$(y_i, 1)$	$(3, 0, 0)$	$(y_i + 3, 1)$	3	3
$(y_i, 1)$	$(3, 1, 0)$	$(y_i + 4, 0)$	4	3
$(y_i, 1)$	$(4, 0, 0)$	$(y_i + 4, 1)$	4	4
$(y_i, 1)$	$(4, 1, 0)$	$(y_i + 5, 0)$	5	4

O primeiro caso corresponde a um bloco em que são preparadas 3 viagens por movimentos buffer-cais, mantendo-se a preparação em pulmão para uso posterior. O número de viagens saídas é incrementado de 3 unidades.

No segundo caso, a preparação existente em pulmão é utilizada, sendo a sua transferência feita no início do bloco. No mesmo bloco, saem também 3 viagens por movimentos buffer-cais. Estes dois casos são os únicos possíveis para um número de viagens activas igual a 3. O terceiro e o quarto casos correspondem a blocos em que o número de viagens activas é igual a 4.

Vamos agora considerar os estados em que o pulmão está vazio, o que implica que não é permitido fazer transferências pulmão-cais, ou seja, $\beta = 0$.

As transições possíveis são as seguintes:

Estado inicial	(α, β, δ)	Estado final	Número viagens saídas	Duração do bloco
$(y_i, 0)$	$(2, 0, 1)$	$(y_i + 2, 1)$	2	3
$(y_i, 0)$	$(3, 0, 0)$	$(y_i + 3, 0)$	3	3
$(y_i, 0)$	$(3, 0, 1)$	$(y_i + 3, 1)$	3	4
$(y_i, 0)$	$(4, 0, 0)$	$(y_i + 4, 0)$	4	4

No primeiro caso, o bloco tem 3 viagens activas, havendo 2 movimentos buffer-cais e uma preparação em pulmão. Na definição do estado final, os valores do número de viagens saídas e o número de preparações em pulmão são incrementados de acordo com essas actividades. No segundo caso, saem 3 viagens por movimentos buffer-cais. O terceiro e o quarto casos correspondem a blocos semelhantes, mas com 4 viagens activas.

Se não for tido em conta a restrição imposta pelo cumprimento da data de disponibilidade dos camiões e de acordo com a definição apresentada, os estados possíveis para esta instância são os apresentados de seguida:

Estágio 1	
Estado inicial	Estados atingidos
$(0, 1)$	$(3, 1)$ $(4, 0)$ $(4, 1)$ $(5, 0)$

Estágio 2

Estado inicial	Estados atingidos
(3,1)	(6,1) (7,0) (7,1) (8,0)
(4,0)	(6,1) (7,0) (7,1) (8,0)
(4,1)	(7,1) (8,0) (8,1) (9,0)
(5,0)	(7,1) (8,0) (8,1) (9,0)

Estágio 3

Estado inicial	Estados atingidos
(6,1)	(9,1) (10,0) (10,1) (11,0)
(7,0)	(9,1) (10,0) (10,1) (11,0)
(7,1)	(10,1) (11,0) (11,1) (12,0)
(8,0)	(10,1) (11,0) (11,1) (12,0)
(8,1)	(11,1) (12,0) (12,1) (13,0)
(9,0)	(11,1) (12,0) (12,1) (13,0)

Só o estado (12,0) corresponde a um estado final válido, com um número final de viagens saídas igual a 12 e sem nenhuma preparação em pulmão. No Estágio 4, não há estágios de interesse para a definição do problema.

Se, por um lado, não considerarmos os estados que não conduzem a este estado final válido e, por outro lado, atendermos à restrição das datas de disponibilidade dos camiões obtém-se a rede apresentada na Figura 8. Os arcos representados correspondem às transições possíveis, e os valores que lhes estão associados traduzem a soma dos custos das demoras associadas à decisão. Os estados assinalados correspondem aos estados possíveis e têm valor finito; os seus valores óptimos são os apresentados na Figura 8. Aos restantes estados, está associado um valor $+\infty$.

Os tempos associados a cada estado, de acordo com a Equação (9), são os seguintes:

Estado	Tempo associado
(0,1)	0
(3,1)	3
(4,0)	3
(4,1)	4
(5,0)	4
(7,1)	7
(8,0)	7
(8,1)	8
(9,0)	8
(12,0)	11

O estado (4,0) tem o tempo associado 3 u.t.. A transição deste estado para o estado (8,0) implicaria a decisão de preparar 4 viagens, que inclui a preparação da viagem 8. Como esta viagem tem um instante de disponibilidade igual a 4, esta transição não é considerada.

A solução óptima corresponde a um plano em que é possível preparar todas as viagens sem atraso. Esta solução corresponde à seguinte transição de estados:

Estado origem	(α, β, δ)	Estado destino	Número viagens saídas	Duração do bloco
(0, 1)	(3,1,0)	(4, 0)	4	3
(4, 0)	(3,0,1)	(7, 1)	3	4
(7, 1)	(4,1,0)	(12, 0)	5	4

Os instantes de início de preparação, início de transferência, fim de preparação e fim de transferência são os seguintes:

v_j	1	2	3	4	5	6	7	8	9	10	11	12
r_j	0	0	0	0	0	2	2	4	4	4	4	4
s_j	-	0	0	0	3	3	3	3	7	7	7	7
st_j	0	-	-	-	-	-	-	7	-	-	-	-
Cp_j	-	3	3	3	7	7	7	7	11	11	11	11
C_j	0.4	3	3	3	7	7	7	7.4	11	11	11	11
d_j	2	3	4	5	7	7	9	9	13	13	13	13
D_j	0	0	0	0	0	0	0	0	0	0	0	0

□

5.2 Eficiência

A eficiência da metodologia proposta pode ser avaliada em termos da carga computacional. Na análise que se segue, é assumido que $n \gg N_p$ e $n \gg N_c$.

A carga computacional depende do número de estágios, do número de estados e do número de acções alternativas em cada estado no qual é avaliada a relação de recorrência, dada pela Equação (12). As acções alternativas em cada estado resultam das diferentes formas de dimensionamento dos blocos. A dimensão dos blocos deve ter um valor mínimo para permitir o funcionamento do sistema à taxa máxima. A dimensão máxima de um bloco é função do número de cais, do número de pulmões e do número de preparações disponíveis. Assim, num estado genérico $e_i = (y_i, z_i)$ há a considerar as acções alternativas $e_j = (y_i + \alpha + \beta, z_i - \beta + \delta)$, sendo necessário respeitar os seguintes limites:

- α – o número de operações (b-c) pode variar de 0 a N_c ;
- β – o número de transferências para cais pode variar de 0 até ao valor $\min(N_c, z_i)$;
- δ – o número de preparações em pulmão pode variar de 0 até ao valor $(N_p - z_i)$.

O número de acções alternativas possíveis é limitado superiormente por $N_c \times \min(N_c, z_i) \times (N_p - z_i)$. Tipicamente, nestes armazéns, o número de pulmões é superior ao número de cais, pelo que $\min(N_c, z_i) = N_c$. Assim, a ordem de grandeza do conjunto de acções alternativas, S , é $O(S) = O(N_c^2 \cdot N_p)$.

O número de estados possíveis para n viagens é da ordem $O(n \cdot N_p)$. Este valor resulta da análise aos valores máximos possíveis para cada elemento do par ordenado:

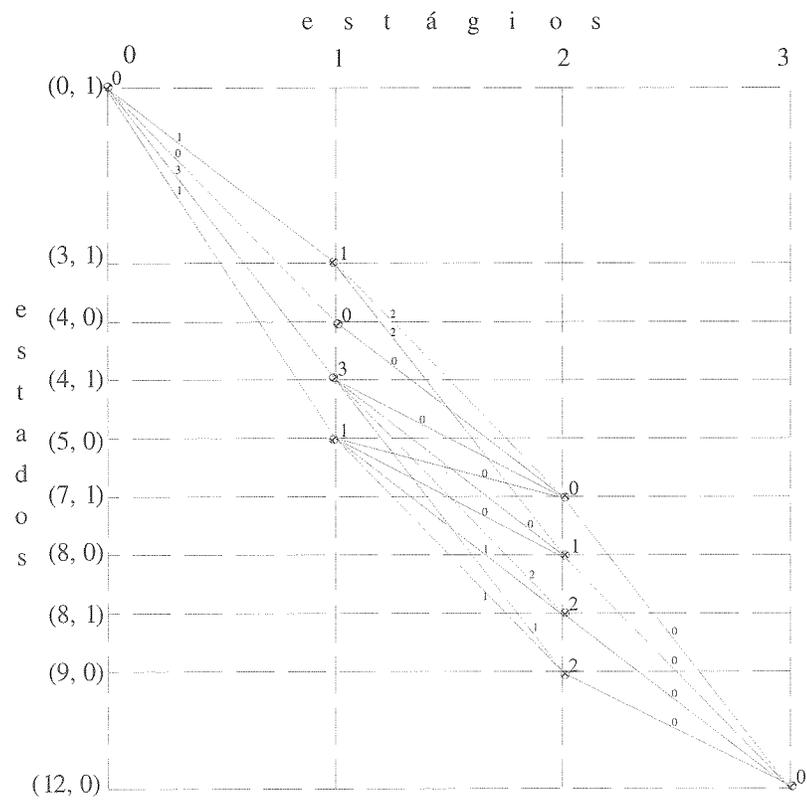


Figura 8: Enumeração das decisões alternativas

y_{\max} – o limite máximo é n , porque todas as viagens têm de ser expedidas quer por operações (b-c), quer por operações (p-c). No entanto a soma das duas parcelas não excede n ;

z_{\max} – o valor máximo para este parâmetro do par ordenado é dado pelo número de pulmões existentes, N_p .

O número máximo de estágios resulta da preparação das n viagens em blocos de dimensão mínima, $\frac{n}{N_{\min}}$, que é um valor de ordem $O(n)$.

A eficiência de um algoritmo de programação dinâmica é de ordem de grandeza do produto do número de estágios pelo número de estados e pelo número de alternativas [3]. Para o modelo em análise, a eficiência do algoritmo de programação dinâmica é $O(n) \cdot O(n \cdot N_p) \cdot O(N_c^2 \cdot N_p)$, ou simplificando, $O(n^2 \cdot N_c^2 \cdot N_p^2)$.

Tratando-se de um algoritmo polinomial, é possível estabelecer os valores máximos de tempo de computação para qualquer número de viagens. A resolução de uma instância de teste formada por 100 viagens demora 270 milissegundos numa máquina AMD K6 400 MHz.

6 Conclusão

Neste artigo, analisámos um sistema que assegura um fluxo constante, ao longo do tempo, de contribuições de processamento que são divididas equitativamente pelas tarefas activas. Isto dá origem a um problema com tarefas com um tempo de processamento variável, que depende do número de tarefas activas.

Quando se consideram datas de disponibilidade e datas de entrega, este problema pertence à classe de problemas NP-difíceis. Para obter soluções no caso em análise, foi adoptada uma heurística de solução baseada em programação dinâmica. Esta estratégia permite obter um procedimento de $O(n^2 \cdot N_c^2 \cdot N_p^2)$.

Agradecimentos

Agradecemos aos Drs. Jorge Vieira e Eduardo Silvério da Efacec, Sistemas de Informação, S. A., a sua colaboração na definição e modelação deste problema.

Este trabalho foi realizado no âmbito de um projecto da Linha de Engenharia de Sistemas e de Processos Industriais do Centro Algoritmi.

Bibliografia

- [1] Blazewicz, J., Ecker, K., Pesch, E., Schmidt, G. and Weglarz, J. *Scheduling in Computer and Manufacturing Processes*, Springer-Verlag, Berlin, 1996.
- [2] Elsayed, E. & Stern, R., Computerized algorithms for order processing in automated warehousing systems, *International Journal of Production Research* 21, 1983, 579-586.
- [3] French, S., *Sequencing and Scheduling: An Introduction to the Mathematics of the Job-Shop*, John Wiley, New York, 1982.

- [4] Gibson, D. e Sharp, G., Order batching procedures, *European Journal of Operational Research* 58, 1992, 57-67.
- [5] Rodrigues, A. *Investigação Operacional: Modelos Determinísticos*, Universidade do Minho, Braga, 1994.
- [6] Randhawa, S. e Shroff, R., Simulation-based design evaluation of unit automated storage/retrieval systems, *Computers and Industrial Engineering* 28, 1995, 71-79.
- [7] Valério de Carvalho, Carvalho e Oliveira, Programação de operações em sistemas com tempo de processamento variável, *Pesquisa Operativa*, 22, n.3, 2002.

Graph-Based Structures for the Market Baskets Analysis

Luís Cavique *

* ESCS-IPL / IST-UTL
Portugal
lcavique@escs.ipl.pt

Abstract

The market basket is defined as an itemset bought together by a customer on a single visit to a store. The market basket analysis is a powerful tool for the implementation of cross-selling strategies. Although some algorithms can find the market basket, they can be inefficient in computational time. The aim of this paper is to present a faster algorithm for the market basket analysis using data-condensed structures. In this innovative approach, the condensed data is obtained by transforming the market basket problem in a maximum-weighted clique problem. Firstly, the input data set is transformed into a graph-based structure and then the maximum-weighted clique problem is solved using a meta-heuristic approach in order to find the most frequent itemsets. The computational results show accurate solutions with reduced computational times.

Keywords: data mining, market basket, similarity measures, maximum clique problem.

1 Introduction

This paper addresses the problem of finding market baskets in large databases. Current database capacities associated with bar code technology and growth of the Internet has led to a huge collection of customer transaction data.

Companies in different sectors such as banking, insurance, telecommunications and airlines have now become more customer oriented than never. To obtain the customer's profile there are two main data sources using the customer's personal data and the product-oriented data. In order to gather demographic, social, geographic, personality or lifestyle data of the customer, costly surveys are needed. On the other hand, product-oriented data, about the frequency and the quantity each customer buys of a certain item, already exists in the companies' database. In order to establish customer relationship strategy one needs to find out whom the best

customers are, how they respond to a campaign and be able to predict the next item each customer will buy, thus implementing a cross-selling strategy.

In section 2 we define the market basket problem and review the Apriori algorithm that solves this problem. Since this algorithm has a non-polynomial time complexity, we describe related work that tries to overcome this handicap.

In section 3 we present a swifter algorithm - Similis, which first of all transforms the input data set into a graph-based structure, and then the new problem, the weighted clique problem, is solved using a meta-heuristic approach. Each maximum-weighted clique corresponds to a quasi-most-frequent itemset.

In section 4 the Similis algorithm is validated with two real data sets and the computational results are presented.

Finally, in section 5 the conclusions point out the differences between the Apriori algorithm and the Similis algorithm.

In this paper we will refer to the market basket (or itemset) whenever it is related to physical data, on the other hand, if it refers to condensed data the terms graph-based structure (or clique) are used.

2 The Market Basket

2.1 The definition of the problem

The input for the market basket analysis is a file with a set of purchases. A market basket is composed of items bought together in a single trip to a store. The most significant fieldnames are the customer identification and item identification, ignoring the quantity bought and the price. Each transaction represents a purchase, which occurred in a specific time and place, and may be linked to an identified customer (usually carrying a card) or to a non-identified customer.

Definition 1: The file with multiple transactions can be shown in a relational database table $T(\text{customer}, \text{item})$. Corresponding to each attribute there is a non-empty set called domain. The $\text{domain}(\text{customer}) = \{1, 2, 3, \dots, n\}$ and the $\text{domain}(\text{item}) = \{a, b, c, \dots, z\}$. The table $T(\text{customer}, \text{item})$ can be seen as a set of all customer transactions $\text{Trans} = \{t_1, t_2, t_3, \dots, t_k\}$ where each transaction contains a subset of items $t_k = \{i_a, i_b, i_c, \dots\}$. The relational table $T(\text{customer}, \text{item})$ can also be seen as a set of item-clientele $\text{Itc} = \{i_1, i_2, i_3, \dots\}$ where each item-clientele contains a subset of customers $i_k = \{c_1, c_2, c_3, \dots\}$.

Based on the attributes (customer, item), the market basket will be defined as the N items that are bought together more frequently. Once the market basket with N items is known, we can move on to cross-selling. The next step is to identify all the customers having bought $N-m$ items of the basket and suggest the purchase of some m missing items. In order to make decisions in marketing applications, the market basket analysis is a powerful tool supporting the implementation of cross-selling strategies. For instance, if a specific customer's buying profile fits into a known market basket, the next item will be proposed.

One of the first approaches, in order to find the market basket consists in using the Collaborative Filtering software developed by Net Perception that finds a "soul mate" for each customer [Hughes 2000]. A customer's "soul mate" is a customer who has identical tastes and therefore the same market basket. The software is installed in hundreds of companies. However, its disadvantage is that it merely compares two individuals and this does not allow an overall view. For example: customer X bought four books about Data Mining and a book about Cooking, and customer Y bought the same four books about Data Mining. Therefore, the software will suggest the Cooking book as the next item for customer Y, leading to possible mistakes.

2.2 The Apriori Algorithm

In opposition to the previous algorithm, the Apriori Algorithm [Agrawal et al. 1996] takes all of the transactions in the database into account in order to define the market basket. The market basket can be represented with association rules, with a left and a right side $L \Rightarrow R$. For instance, given a itemset $\{A, B, C\}$ the rule $\{B, C\} \Rightarrow \{A\}$ should be read as follows: if a customer bought $\{B, C\}$ he probably would buy $\{A\}$. This approach was initially used in pattern recognition and it became popular with the discovery of the following rule: "on Thursdays, grocery store customers often purchase diapers and beer together" [Berry and Linoff 1997].

To evaluate the association rules two measures can be used - the support measure and the confidence measure. Let $\{A, B\}$ be an itemset and the let $A \Rightarrow B$ be the association rule. The support measure is equal to the relative frequency or $P(\{A, B\})$. The confidence measure is given by the conditional probability of B given A, $P(B|A)$, which is equal to $P(\{A, B\})/P(A)$.

The Apriori algorithm was implemented in commercial packages, such as Enterprise Miner from the SAS Institute [SAS 2000]. As input, this algorithm uses a table with purchase transactions. Each transaction contains the customer identification and the purchased items, as follows (customer, item). Input parameters are defined as the maximum number of acquired items (*max_k*) and the minimum support (*min_sup*) of a certain basket. The Enterprise Miner package from SAS Institute implemented this algorithm using *max_k=4* and *min_sup=5%* as default values. The *min_sup* is of extreme importance in the algorithm since it will prune the useless branches in the search.

The first step of the Apriori algorithm generates sets of market baskets. I_k is defined as the set of frequent items with *k* items bought together. Firstly, the algorithm filters the items with a frequency that is higher than the *min_sup*, generating I_1 . In the following stages, for each I_k it generates the I_{k+1} candidates, such as $I_k \subseteq I_{k+1}$. For each I_{k+1} candidate, the algorithm removes the baskets, which are lower than the *min_sup*. The cycle ends when it reaches I_{max_k} .

In the second step, the Apriori algorithm generates sets of market baskets and then generates association rules $L \Rightarrow R$. For each rule, the support measure and the confidence measure are calculated. In order to implement the cross-selling strategy the data analysts choose, firstly, the dimension of the basket, secondly, they choose the rules with the highest support measure. Finally, those with the highest confidence measure are chosen, among those with the highest support measure.

The outputs of the Apriori algorithm are easy to understand and many new patterns can be identified. However, the sheer number of association rules may make the interpretation of the results difficult. A second weakness is the computational times, due to the exponential complexity of the algorithm.

2.3 Related Work

Apriori algorithm has an exponential time complexity, and several passes over the input table are needed. To overcome these handicaps some proposals have been made.

The Apriori algorithm performs as many passes over the data as the size of the itemsets. The Dynamic Itemset Counting, the DIC Algorithm, reduces the number of passes made over the data, using itemsets forming a large lattice structure [Brin et al. 1997]. DIC starts counting the 1-itemset and then adds counters 2, 3, 4, ..., k itemsets. Thus, after a few passes over the data it finishes counting all the itemsets. Running DIC and Apriori, DIC outperformed Apriori in the majority of cases.

In [Aggarwal, Wolf and Yu 1999] a method for indexing market basket data for similarity search is discussed. The index structure requires the partition of the set of all k-itemsets into subsets. They create a graph so that each node corresponds to an item, and for each pair of items a weight is added, which is the inverse of the support measure. Finally, the set of items is divided into k-sets. This algorithm shows good scalability with an increase in the number of transactions.

Just like the DIC algorithm, the MARC algorithm [Liu, Lu and Lu 2001] avoids several passes over the databases. MARC algorithm will use the summarized cluster information. The algorithm analyzes the similarities between transactions and creates clusters of similar transactions.

In the GCTD algorithm [Chen et al. 2002], the concepts of similarity relationships and the clustering problem appear together in order to discover connected components in an undirected graph.

The condensed data representation is extremely useful [Jeudy and Boulicaut 2002], taking into account that the Apriori algorithm has a better performance using sparse data rather than using highly correlated data. The latter is considered difficult or even intractable.

In the developing of recommender systems, i.e., systems that personalize recommendations of items based on the customer's preferences, in [Lin, Alvarez and Ruiz 2002] the authors present an algorithm that does not require the min_sup measure. Having previously defined min_sup, a negative result can be expected, by cutting down either too many or too few itemsets.

In order to obtain frequent market baskets in reduced computational times, in the next section we present the Similis Algorithm [Cavique e Themido 2001] [Cavique 2002]. This algorithm reuses some of the mentioned strategies, such as the reduction of passes over the database, the reduction of the number of parameters (e.g. min_sup) and the aggregate measures (e.g. similarity measures).

3 The Similis Algorithm

To overcome the performance problems of the Apriori algorithm we have developed a new algorithm, Similis, meaning similar. This term was chosen just like Apriori, since they are both Latin names.

The Similis algorithm is developed in two steps - the transformation step and the search step. In this section, we firstly describe the problem transformation and the similarity measures, just like in [Chen et al. 2002]. Secondly, we justify the graph-based structure and present the search step and finally we present the algorithm in its totality.

3.1 Problem Transformation

The input for the Market Basket Analysis is a table with two attributes $T(\text{customer}, \text{item})$ that represents the item bought by a customer in a single transaction.

A possible way to condense the data is by transforming it into a graph structure. A graph is a pair $G=(V,E)$ of sets satisfying $E \subseteq [V]^2$ where elements of E are 2-element subsets of V . The elements of V are the vertices (or vertexes, or nodes, or points) of the graph $G(V,E)$ and the elements of E are its edges (or arcs, or lines). In market basket case each vertex corresponds to an item and each arc has a weight which represents the distance between the adjacent vertices. The distance between two items is given by the frequency that the two items are bought together.

To find the values for the weighted graph $G(V,E)$ some similarity measures can be used. The similarity value of the two items will be high if they are both included in frequent transactions.

Some authors study association measures between transactions [Liu, Lu and Lu 2001]. However, in this work the associations between items are analyzed. If two items are frequently bought in the same transactions, then they belong to a frequent market basket. In order to create sets of items, one association measure must be found, similarity or distance (dissimilarity) measures can be created.

Given two item-clientele A and B , the first two similarity functions that come up are the number of matches and the hamming distance. The number of matches is given by the cardinality of $(A \cap B)$, while on the other hand, the hamming distance is given by the sum of the cardinalities of the sets $(A - B)$ and $(B - A)$. A third measure is the euclidian distance, which is given by the square root of the sum of the squares of the differences.

Even though a variety of measures can be used, normalized measures are preferable. The most common ones are those based on vector comparisons.

In the generation of the weighted graph we will use the analogy with the Information Retrieval Systems. The Information Retrieval systems were developed in the middle of the 20th century to manage the vast amount of scientific information produced in universities, research centers and by the academic press. The commercial Information Retrieval systems have millions of recorded documents each containing many terms. Information Retrieval systems are very efficient in many fields that deal with documents and semi-structured data.

At this stage we will make an analogy between the Information Retrieval techniques and the market basket problem. Just as a document consists of many terms, a market basket contains several items. In the same way, the study of terms in a document is identical to the study of items in a basket. For each pair of items (A,B) a similarity measure $SIM(A,B)$ can be found, if the items are bought together many times they have a strong similarity, but they have a weak similarity if they are not usually bought together. For all items, an item similarity

matrix is generated, which can be represented by the adjacent matrix of the weighted graph $G(V,E)$.

To present the Information Retrieval System, in [Salton and McGill 1983] the authors describe the following similarity measures. All measures use binary matrices and return normalized values between 0 and 1. The Dice (sim1), Jaccard (sim2) and Cosine (sim3) coefficients are widely used given their simplicity.

$$sim1(A,B) = \frac{2 \cdot \sum_k (A_k \cdot B_k)}{\sum_k A_k + \sum_k B_k}$$

$$sim2(A,B) = \frac{\sum_k (A_k \cdot B_k)}{\sum_k A_k + \sum_k B_k - \sum_k (A_k \cdot B_k)}$$

$$sim3(A,B) = \frac{\sum_k (A_k \cdot B_k)}{\sqrt{\sum_k (A_k)^2 \cdot \sum_k (B_k)^2}}$$

Definition 2: A multiplicative model will be used to express the weight of an edge (A,B). The weight of the edge (A,B) takes into account the similarity and frequency of items, such as:

$$weight(A,B) = sim(A,B) \cdot frequency(A,B)$$

The similarity value of two items will be high if they are both included in the same transactions. The frequency of the item must be considered to guarantee a correspondence between high-weighted edges and items that appear in many transactions.

There are several ways to define item frequency. In this work we opt for the average of the relative frequency of the two items, given by:

$$frequency(A,B) = \frac{\sum_k A_k + \sum_k B_k}{2 \cdot n}$$

Using the Dice similarity measure, it is interesting to notice that the weight is equal to the support measure of two items.

$$weight1(A,B) = \frac{\sum_k (A_k \cdot B_k)}{n}$$

However, based on the computational results the Jaccard similarity measure was chosen.

Numeric Example 1:

To exemplify the first step in the Similis algorithm, a problem with 5 items and 7 customers is given, the domain(Item) = {a, b, c, d, e} and the domain(Customer) = {1, 2, 3, 4, 5, 6, 7}.

The item-clientele $I_{tc}=\{I_a, I_b, I_c, I_d, I_e\}$ where $I_a=\{2, 3, 5, 6\}$, $I_b=\{1, 2, 4, 7\}$, $I_c=\{1, 3, 4, 5, 7\}$, $I_d=\{1, 4, 5, 6\}$ and $I_e=\{3, 4\}$.

To obtain the adjacent matrix of graph $G(V,E)$, in the calculation of the weight of the edges the Jaccard similarity measure (sim2) is use. For instance, for items a and b , the weight (a,b) is given by:

$$\text{sim2} (I_a,I_b) = \frac{\sum_k (I_{a_k} \cdot I_{b_k})}{\sum_k I_{a_k} + \sum_k I_{b_k} - \sum_k (I_{a_k} \cdot I_{b_k})} = \frac{1}{4+4-1} = \frac{1}{7}$$

$$\text{frequency} (I_a,I_b) = \frac{\sum_k I_{a_k} + \sum_k I_{b_k}}{2 \cdot n} = \frac{(4+4)}{2 \cdot 7} = \frac{4}{7}$$

$$\text{weight2}(a,b) = \text{sim2} (I_a,I_b) \cdot \text{frequency} (I_a,I_b) = 1/7 \times 4/7 = 0.082$$

For items a and c , the weight (a,c) is given by:

$$\text{weight2}(a,c) = \text{sim2}(I_a,I_c) \cdot \text{frequency} (I_a,I_c) = 2/7 \times 4.5/7 = 0.184$$

The process is repeated for all the pairs of items, returning the adjacent matrix of the weighted graph G , presented in table 1.

Table 1: Adjacent matrix of the weighted graph $G(V,E)$

G(V,E)	b	c	d	e
a	0.082	0.184	0.190	0.086
b		0.321	0.190	0.086
c			0.321	0.200
d				0.086

In this way the data is condensed in a graph $G(V,E)$ using a procedure with time complexity $O(|V|^2)$, where $|V|$ is the number of vertexes i.e. the number of items.

3.2 Searching for the Maximum-weighted Clique

Given the undirected graph $G(V,E)$, then $G_1(V_1,E_1)$ is called a subgraph of G if $V_1 \subseteq V$ and $E_1 \subseteq E$, where each edge of E_1 is incident in the vertices of V_1 . A subgraph G_1 is said to be complete if there is an arc for each pair of vertices. A complete subgraph is also called a clique. A clique is maximal if it is not contained in any other clique. In the maximum clique problem the objective is to find the largest complete subgraph in a graph. The clique number is equal to the cardinality of the largest clique of the graph. If a weight is given to each edge, the subgraph is known as weighted subgraph, and the weight of the subgraph is given by the sum of the weights of the edges.

A clique can represent a common interest group. Given a graph representing the communication among a group of individuals in an organization, each vertex represents an individual, while edge (i,j) shows that individual *i* regularly communicates with individual *j*. Finding a common interest group where each individual communicates with all of the other group members, corresponds to finding a clique. In French "la clique" is defined as a closely knit group of individuals who share common interests. Finding the maximum clique means finding the largest common-interest group possible.

If a graph with weights in the edges is used, the most weighted clique corresponds to the common-interest group whose elements communicate the most among themselves. This structure allows the representation of sets of elements strongly connected.

We use the same principle in the market basket analysis, where the items bought together in a single transaction have a higher degree of connection than items bought in different transactions. In the last section we presented the data transformation of table T(customer, item) into a weighted graph G(V,E). In other words, the market basket problem can be transformed into the maximum-weighted clique problem, using a procedure with quadratic complexity.

Definition 3: As a way of comparing the weighted clique, given a size *k*, the clique weight must be calculated. Although it can be defined in several ways, the sum of the weighted edges was chosen, as follows:

$$\text{Clique_weight} = \sum \text{weight}(i,j), \forall \text{edge}(i,j) \in \text{Clique}$$

The Maximum Clique Problem is an important problem in combinatorial optimization with many applications, which include: market analysis, project selection and signal transmission [Berge 1991]. The Maximum Clique Problem is a hard combinatorial problem, classified as NP-Hard [Garey and Johnson 1979]. The interest in this problem led to the algorithm thread challenge on experimental analysis and algorithm performance promoted by Second DIMACS Implementation Challenge [Johnson and Trick 1996]. A recent state of the art survey is found in [Bomze et al. 1999].

In order to deal with large data volumes, a number of heuristic algorithms have been recently proposed: Genetic Algorithms [Balas and Niehaus 1998], Neural Networks [Jagota, Sanchis and Ganesan 1996], GRASP with Simulated Annealing [Feo, Resende and Smith 1994], Tabu Search [Soriano and Gendreau 1996] [Cavique, Rego e Themido 2002b] and Scatter Search [Cavique, Rego and Themido 2002a] [Cavique 2002].

To find the Maximum-weighted Clique with size *k*, the Primal-Tabu algorithm [Cavique, Rego and Themido 2002b] [Cavique 2002] was adapted by adding the restriction related to the clique size and the objective function was changed.

Three possible moves were identified: the related neighborhood structures are N^+ , N^- , and N^0 for addition, removal and swap of a vertex of the subgraph. To avoid cycling in the iterative search process, a short-term memory is incorporated. The short-term memory is usually implemented using a Tabu list, made up of a set of (reverse) moves, which are forbidden for a certain number of iterations.

The original version of Primal-Tabu combines the neighborhood structures N^+ , N^0 and N^- . At each step one new solution S' is chosen from the neighborhood $N(S)$ of the current

solution S . If it is possible to increase the clique value, one vertex is added to the clique using N^+ , or else try to swap vertices using the neighborhood structure N^0 , otherwise the heuristic moves backwards removing one vertex using N^- . At each iteration the current solution S and the best solution found S^* are updated whenever the clique value is increased.

The weighted version of the algorithm restricts the growth of $N^+(S)$ by using the condition $|S| < k$; the objective function was changed, instead of using the vertex number, the function for the clique weight is now used. The local search will run through the solution space searching for the maximum-weighted clique with a given dimension k . The Primal-Tabu Heuristic can be sketched as follows:

S is the current solution and S' the trial solution
 S^* is the best solution found
 Tabu is the tabu list
 $N^+(S)$, $N^-(S)$, $N^0(S)$ are the neighborhood structures

Primal-Tabu Meta-Heuristic for the Maximum-weighted Clique

```

input: weighted graph  $G(V,E)$ , size  $k$ ;
output: maximum-weighted clique  $S^*$  with size  $k$ ;
initialize  $S$ ,  $S^*$  e Tabu;
  while not end condition
    if  $N^+(S) \setminus \text{Tabu} \neq \emptyset$  and  $|S| < k$  choose the best  $S'$ ;
    else if  $N^0(S) \setminus \text{Tabu} \neq \emptyset$  choose the best  $S'$ ;
        else choose the best  $S'$  in  $N^-(S)$  and update Tabu;
    update  $S \leftarrow S'$ ;
    if  $\text{Clique\_weight}(S) > \text{Clique\_weight}(S^*)$  then  $S^* \leftarrow S$ ;
  end while;
return  $S^*$ ;

```

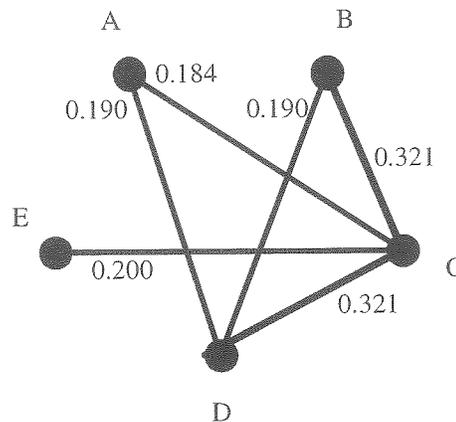


Figure 1: Weighted Graph

Numeric Example 2:

Following the previous numeric example, based on the found weighted graph, a chart presentation is given in figure 1, where only the edges which weigh more than 0.100 are shown, in order to clarify the figure. Two cliques with three vertices stand out from the graph: the clique (B,C,D) and (A,C,D) . For the clique (B,C,D) the sum of the edges is equal to 0.832

and for clique (A,C,D) the value is 0.695 is obtained. The most frequent market basket with 3 items occurs twice, is (B,C,D) and is also the most weighted clique, showing the aim was reached. The information given by the weighted graph seems to be richer than the frequency of a market basket with k items, since it also includes information from smaller markets baskets. The formal validation is given in section 4, where computational results show accurate solutions.

3.3 The Algorithm Description

To find the market basket patterns, i.e. the most frequent itemsets, the Similis algorithm is described in two steps - the data transformation step and the search step.

For the first step, the input is the table $T(\text{customer, item})$ and the output a weighted graph $G(V,E)$. For the second step, the input is the graph $G(V,E)$ and the size of the market basket k and the output is the market basket with k items. The search step can run more than once, depending on the market basket dimensions one is looking for.

Firstly a weighted graph $G(V,E)$ is generated based on the similarities of the items. In the graph $G(V,E)$ the vertex set V represents the itemset in the market basket. The weighted edge $(i,j) \in E$ represents the similarity between item i and item j . Two items are similar if they were bought together in many transaction.

Finally, to find the maximum-weighted clique that corresponds to the most frequent market basket, we adapted the Primal-Tabu Meta-heuristic. The main algorithm can be sketched as follows:

The Similis Algorithm

STEP 1 – Data Transformation

input: table $T(\text{customer, item})$

□ Generate graph $G(V,E)$ using the similarities between items

output: weighted graph $G(V,E)$

STEP 2 – Finding the maximum-weighted cliques

input: weighted graph $G(V,E)$ and size k

□ Find in $G(V,E)$ the clique S with k vertexes with the maximum weight, using the Primal-Tabu Meta-heuristic.

output: weighted clique S of size k that correspond to the most frequent market basket with k items.

In the first step, the condensed data is created, in such a way that the second step can run as many times as needed, altering the market basket size, thus showing one of the advantages of having condensed data.

The Apriori algorithm firstly finds the most frequent itemsets and secondly generates the association rules. We follow the same procedure by finding solutions and then creating association rules. To implement a cross-selling strategy, frequent itemsets with dimensions 2, 3, ..., k must be found, then association rules are generated and evaluated with the support and confidence measures.

4 Computational Results

In the validation of an algorithm some choices must be made such as the data files, the compiler, the computer and the performance measures. To validate the Similis algorithm, two data sets were chosen. The first includes the result of a survey carried out at inns named “Pousadas de Portugal” involving 43 inns (items) and 2500 customers [Cardoso, Themido e Pires 1999]. The second regards the distribution of frozen food items throughout Portugal by Nestle enterprise involving 158 items and 450 consumer centers. The computer program was written in C language and the Microsoft Visual C++ compiler was used. The computational results were obtained from a Pentium 200 MHz. The performance measures that we are going to use are the quality of the solutions and the computational times. The Apriori algorithm that has been implemented in the SAS Enterprise Miner will be used in the computational result comparisons.

The Apriori algorithm has two possible measures: the support and the confidence measure. The Similis algorithm retrieves only the clique weights as the performance measure, so the use of the market basket support seems pertinent for measuring the quality of both algorithms.

For each algorithm the top five solutions were reached. For the Apriori algorithm the solutions are ordered by decreasing support and in the Similis algorithm they are ordered by decreasing clique weight.

Definition 4: For each market basket with k items, found by the Similis algorithm, an accuracy measure is required. The Accuracy function is given by the average support of the top five solutions found by the Similis algorithm, divided by the average support of the top five solutions found by the Apriori algorithm, that is:

$$Accuracy = \frac{AverageSupport(Similis)}{AverageSupport(Apriori)} \cdot 100$$

The “Pousadas” data set includes the original file and two simulated files SX and SY with the same dimension. The computational time (in seconds) for the Apriori algorithm (TA), the Similis algorithm (TS) and the Accuracy obtained (Acc), are reported in table 2. The average accuracy is close to 97%. The Similis algorithm computational times are almost constant, while the computational times for the Apriori algorithm are exponential. What makes the real difference between the two algorithms is the computational time.

Table 2: Computational results of Pousadas data set

Basket Size	Acc Real	Acc SX	Acc SY	TA (sec.)	TS (sec.)
2	100	100	100	3	3
3	100	94	95	42	2
4	94	96	94	440	3

Acc- Accuracy; TA- Time Apriori; TS- Time Similis

The second data set was handed in by the Nestle enterprise and we generated some reduced files from the original one. The data set includes the Nestle20 with 20 items, the Nestle40 with 40 items and Nestle80 and Nestle120 with respectively 80 and 120 items.

The accuracy obtained by the Similis algorithm (Acc), the computational time in seconds for the Apriori algorithm (TA) and Similis algorithm (TS) are reported in table 3. The computational results are limited by the non-existence of cliques, in the Nestle20 case, but the great limitations are the Apriori computational times for Nestle40, Nestle80, Nestle120 and NestleReal.

Table 3: Computational results for Nestle data set for different basket sizes

basket size	Nestle20			Nestle40			Nestle80			Nestle120			NestleReal		
	Acc	TA	TS	Acc	TA	TS	Acc	TA	TS	Acc	TA	TS	Acc	TA	TS
2	100	0	18	100	1	56	100	5	204	100	12	384	100	23	573
3	97	1	17	92	8	43	100	89	159	100	347	293	100	1476	442
4	98	3	19	95	72	43	100	3771	159	99	25512	291	98	35235	455
5	86	10	17	88	486	44	93	21332	159						
6	93	26	16	92	2761	36									
7	95	51	16	97	13842	35									
8	90	80	16												
9	93	139	16												
10															

Acc- Accuracy; TA- Time Apriori; TS- Time Similis

The Similis algorithm has a good performance in the Acc with an average of 96%, and has constant computational times for baskets with different sizes for the Nestle data set.

In the top five solutions, the best solution found with the Apriori algorithm is also the one that is generally found by the Similis algorithm. The difficult problem is finding the most frequent baskets, since the frequency calculation corresponds to polynomial complexity procedures. After getting the maximum-weighted cliques (or baskets) using the Similis algorithm, it is possible to re-order the baskets based on the frequency in polynomial time.

When we need quasi-most-frequent market baskets involving limited computational effort, the Similis algorithm is very competitive as shown in the computational results of both data sets.

5 Conclusions

In this paper after having described the usefulness of the market basket problem, the Apriori algorithm was presented. The main disadvantage of the Apriori algorithm is the exponential time complexity, since it performs many passes over the data. Using few items or sparse data the algorithm is efficient, while when using correlated data the performance degrades significantly. In order to reduce the number of items in the input data, the `min_sup` parameter is used. However, since the `min_sup` is independent from the data table, it may lead to unpredictable data reductions. Finally, the Apriori algorithm generates a huge number of associative rules, of which only a few ones are used in cross-selling strategies.

Considering all of the handicaps of the Apriori algorithm, we developed the Similis algorithm because of its lower computational complexity, thus allowing the resolution of a greater number of real problems. In this innovative approach, the condensed data is obtained by transforming the market basket problem in a maximum-weighted clique problem. Firstly, the input data set is transformed into a graph-based structure and then the maximum-weighted clique

problem is solved using a meta-heuristic approach in order to find the most frequent itemsets. The computational results show accurate solutions with reduced computational times.

Both algorithms have very different characteristics. Whereas the Apriori algorithm uses a transaction file, the Similis algorithm organizes its condensed data in a weighted graph. They also differ in the procedure to obtain the solutions. The Apriori algorithm uses an exact counting method with exponential complexity, whereas the Similis algorithm uses a meta-heuristic approach. The output of the Apriori algorithm reports the exact market basket frequencies, whereas the Similis algorithm works with a transformed problem, using the “imperfect” information of the clique weights and reports quasi-most-frequent market baskets. Each algorithm can treat different data volumes. The Apriori algorithm can only work with limited data volumes, whereas the Similis algorithm can treat large data volumes.

The time complexity of the Apriori algorithm is dependent on the number of items, the number of transactions and the market basket size, whereas the time complexity of the Similis algorithm is only dependent on the number of items (that correspond to the number of vertexes of the graph) showing an excellent scalability with database size.

The Apriori algorithm performs efficiently with few items or with sparse data, as is already the case in banking and insurance implementations. Let us not forget that the default value for the maximum number of items in the Enterprise Miner of SAS is $\text{max_k} = 4$. The Similis algorithm expands the use of the market basket to hundreds of items in highly correlated data.

The Similis algorithm has some advantages when compared to data-condensed algorithms such as MARC and GCTP, since it doesn't require any additional parameters (like min_sup), making it easier to use.

To validate the Similis algorithm two real-case studies were presented. The first is from the “Pousadas de Portugal” with 43 items and the second is from Nestle enterprise with 158 items. The Similis algorithm performs well, the average of both data sets being equal to 96% for the accuracy function obtained with reduced computational times.

In this innovative approach, the ability to condense the data without using too many parameters and the capacity to find several market baskets with different sizes using reduced computational times, makes Similis algorithm very competitive. The market basket analysis can therefore be implemented in retail markets such, as supermarkets and on-line shopping. Its application can be extended to recommender systems due to its condensed data.

6 References

- [1] C.C. Aggarwal, J.L. Wolf and P.S. Yu, A New Method for Similarity Indexing of Market Basket Data, in *Proceedings of the 1999 ACM SIGMOD Conference*, Philadelphia PA, 1999, pp.407-418.
- [2] R. Agrawal, H. Mannila, R. Srikant, H. Toivonen and A. Verkamo, Fast Discovery of Association Rules, in *Advances in Knowledge and Data Mining*, U.M. Fayyad, G. Piatetsky-Shapiro, P. Smyth and R. Uthurusamy Eds, MIT Press, 1996.
- [3] E. Balas, W. Niehaus, Optimized Crossover-Based Genetic Algorithms will be the Maximum Cardinality and Maximum Weight Clique Problems, *Journal of Heuristics*, Kluwer Academic Publishers, 4, 1998, pp. 107-122.
- [4] C. Berge, *Graphs*, 3rd edition, North-Holland, 1991.

- [5] M. Berry and G. Linoff, *Data Mining Techniques for Marketing, Sales and Customer Support*, John Wiley and Sons, 1997.
- [6] I.M. Bomze, M. Budinich, P.M. Pardalos and M. Pelillo, Maximum Clique Problem, in *Handbook of Combinatorial Optimization*, D.-Z. Du and P.M. Pardalos Eds, 1999, pp.1-74.
- [7] S. Brin, R. Motwani, J.D. Ullman and S.Tsur, Dynamic Itemset Counting and Implication Rules for Market Basket Data, in *Proceedings of the 1997 ACM SIGMOD Conference*, Tucson, Arizona, 1997, pp.255-264.
- [8] M. Cardoso, I. Themido and F. Moura Pires, Evaluating a Clustering Solution: an Application in the Tourism Market, *Intelligent Data Analysis*, 3, 1999, pp. 491-510.
- [9] L. Cavique and I Themido, A New Algorithm for the Market Basket Analysis, Internal Report CESUR-IST, UTL, Portugal, 2001.
- [10] L. Cavique, C. Rego and I. Themido (a), A Scatter Search Algorithm for the Maximum Clique Problem, in *Essays and Surveys in Meta-heuristics*, C. Ribeiro and P. Hansen Eds, Kluwer Academic Publishers, 2002, pp. 227-244.
- [11] L. Cavique, C. Rego and I. Themido (b), Estruturas de Vizinhança e Procura Local para o Problema da Clique Máxima, *Revista de Investigação Operacional*, 2002, vol.22, pp. 1-18.
- [12] L. Cavique, Meta-heurísticas na Resolução do Problema da Clique Máxima e Aplicação na Determinação do Cabaz de Compras, dissertação de Doutoramento em Engenharia de Sistemas no Instituto Superior Técnico da Universidade Técnica de Lisboa, 2002.
- [13] N. Chen, A. Chen, L. Zou and L. Lu, A Graph-based Clustering Algorithm in Large Transaction Databases, *Intelligent Data Analysis*, 5, 2002, pp.327-338.
- [14] T.A. Feo, M.G.C. Resende and S.H. Smith, A Greedy Randomized Adaptive Search Procedure for Maximum Independent Set, *Operations Research*, 42, 1994, pp. 860-878.
- [15] M.R. Garey and D.S. Johnson, *Computers and Intractability: a Guide to the Theory of NP-Completeness*, W.H. Freeman and Company, New York, 1979.
- [16] A.M. Hughes, *Strategic Database Marketing*, McGraw-Hill, 2000.
- [17] A. Jagota, L. Sanchis, R. Ganesan, Approximately Solving Maximum Clique using Neural Network Related Heuristics, in *Clique, Coloring and Satisfiability, Second Implementation Challenge DIMACS*, D.S. Johnson and M.A. Trick Eds, AMS, 1996, pp. 169-203.
- [18] B. Jeudy and J.-F. Boulicaut, Optimization of Association Rule Mining Queries, *Intelligent Data Analysis*, 6, 2002, pp.341-357.
- [19] D.S. Johnson and M.A. Trick Eds, *Clique, Coloring and Satisfiability, Second Implementation Challenge DIMACS*, AMS, 1996.
- [20] W. Lin, S.A. Alvarez and C. Ruiz, Efficient Adaptive-Support Association Rule Mining for Recommender Systems, *Data Mining and Knowledge Discovery*, 6, Kluwer Academic Publishers, 2002, pp.83-105.
- [21] F. Liu, Z. Lu and S. Lu, Mining Association Rules Using Clustering, *Intelligent Data Analysis*, 5, 2001, pp.309-326.
- [22] G. Salton and M.J. McGill, *Introduction to Modern Information Retrieval*, McGraw-Hill, 1983.
- [23] SAS Software, *Enterprise Miner Documentation*, SAS Institute, 2000.
- [24] P. Soriano and M. Gendreau, Tabu Search Algorithms for the Maximum Clique, in *Clique, Coloring and Satisfiability, Second Implementation Challenge DIMACS*, D.S. Johnson and M.A. Trick Eds, AMS, 1996, pp. 221-242.

Heuristic Sensitivity Analysis for Baker's Yeast Model Parameters

Celina P. Leão *

Filomena O. Soares †

* Dept. of Production and Systems
School of Engineering - Univ. of Minho
Braga
cpl@dps.uminho.pt

† Dept. of Industrial Electronics
School of Engineering - Univ. of Minho
Guimarães
filomena.soares@dei.uminho.pt

Abstract

The baker's yeast, essentially composed by living cells of *Saccharomyces cerevisiae*, used in the bread making and beer industries as a microorganism, has an important industrial role. The simulation procedure represents then a necessary tool to understand clearly the baker's yeast fermentation process. The use of mathematical models based on mass balance equations requires the knowledge of the reaction kinetics, thermodynamics, and transport and physical properties. Models may be more or less complex, however they keep the basic feature of linking observations together into some pattern.

A FORTRAN90-based program was developed to simulate the baker's yeast fermentation process in order to predict the dynamic behaviour of a well-mixed reactor. Mass balances written for all the components define a system of ordinary differential equations of initial value problem type (IVP). Considering the kinetics and the gas transfer rates relations as part of the differential system, a differential-algebraic system (DAE) can be defined. The simulation results were compared with the experimental values obtained in a laboratorial five-litre fermenter, operated in fed-batch mode.

Prior to the parameter estimation procedure, an identification of the most significant model parameters was carried out. A heuristic sensitivity analysis was performed in order to adjust the model results with the experimental data. The Meyer and Roth method was used to minimise the objective function, defined as the sum of the relative square errors between the calculated and the experimental values (associated to the state variables: biomass, glucose and ethanol).

The yield coefficients and the maximum uptake rate for glucose and oxygen were found the most significant parameters.

Keywords: heuristic sensitivity analysis, numerical simulation, parameter estimation.

1 Introduction

Real processes in the chemical, biochemical and food industry are, in their vast majority, non-linear MIMO (Multiple Input Multiple Output) systems. Their dynamics and control are difficult to study both for theoretical and practical reasons. In many cases, experiments with real industrial processes are not carried out for economy and safety reasons, and frequently on-line measurements are not available or simply they are too expensive. Advanced control strategies rely on adaptive techniques based on the knowledge of the system state. When experimental observations are missing, 'software sensors' represent a major, and often, the only alternative to allow optimal process operation to be enforced. The robustness of such sensors needs however to be tested in flexible environments where difficulties such as measurement noises, time delays and loads to the process can be readily implemented on-line [1].

Living cells of *Saccharomyces cerevisiae*, which form baker's yeast, are predominantly used in bakery and beer industries. Apart from its industrial importance and economical significance, there is a scientific interest in baker's yeast fermentation [2].

Baker's yeast production is carried out in a fed-batch fermenter with inoculums of *Saccharomyces cerevisiae* culture and a glucose solution as substrate feed. We may distinguish three metabolic pathways: respirative growth on glucose, fermentative growth on glucose and respirative growth on ethanol. Respirative pathways occur in presence of oxygen and the fermentative one in its absence (with production of ethanol) [3].

The conventional approach for process modelling is based on mass, energy and/or population balance equations. This form of modelling requires knowledge about reaction kinetics, thermodynamic, transport and physical properties.

The simulation process requires the integration of a set of non-linear differential equations, for the state variables. A set of algebraic equations, concerning mass transfer relations and kinetics laws, can be considered as part of the system to be solved, defining a differential algebraic equation (DAE) system or, alternatively included in the model, transforming the system in an initial value problem (IVP) [4-5]. The simulation model needs then to be validated with experimental data.

For an identification of the more significant model parameters a heuristic sensitivity analysis was performed. The Meyer and Roth method [6] is used to minimise the objective function, defined as the sum of the relative square errors between the calculated and the experimental values. Two steps are considered. The first one identifies the most relevant kinetics parameters and the second one, with the previous identified parameters, estimates the most relevant yield coefficients. Depending on the research interests a multi-objective analysis can be also performed without previous identification of the most relevant parameters [7].

2 Baker's yeast fermentation - Modelling approach

A deterministic model for biological well-mixed fed-batch reactors, Figure 1, is obtained by writing macroscopic balance for all the components. The simulation model also includes a set of algebraic equations, related to the mass transfer relations and to the kinetics of baker's yeast growth [3].

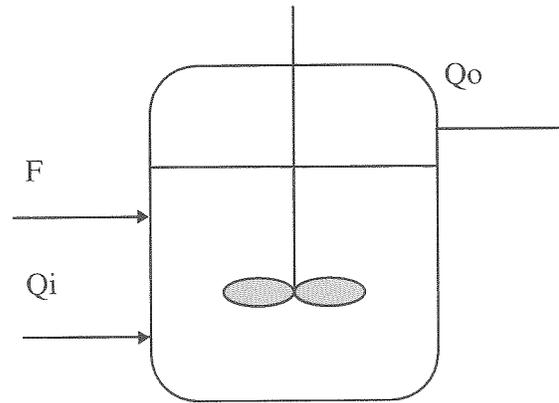


Figure 1: Fed-batch fermenter (F is liquid flow; Q_i and Q_o are gaseous flow at input and output, respectively)

2.1 Kinetic model

Depending on the environment conditions, for instance, availability of sugar and/or oxygen, yeast growth is characterized by three metabolic pathways, respiratory and/or fermentative, namely:

respiratory growth on glucose



fermentative growth on glucose



respiratory growth on ethanol



where μ_S^O , μ_S^r , μ_E^O : specific growth rates (1/h) for the three pathways, respectively; subscripts S and E mean glucose and ethanol.

The metabolic pathways of fermentative growth on glucose and oxidative growth on ethanol are competitive. This competition is governed by the respiratory capacity of the cells. If the instantaneous oxygen uptake capacity exceeds the oxygen need for total respiratory glucose uptake, then, all sugar uptakes follows the respiratory pathway (2.1) with the remaining oxygen being spent on ethanol (if present) respiratory uptake (2.3). Otherwise, if the instantaneous oxygen uptake capacity is not enough, then, part of glucose uptake follows the respiratory pathway (2.1) while the remaining follows the fermentative pathway (2.2). Figure 2 illustrates the mechanism.

Several kinetics models are proposed for baker's yeast growth [8]. In this work we considered Monod type equations, following Sonnleitner and Käppeli work [9], as described and explained in [3].

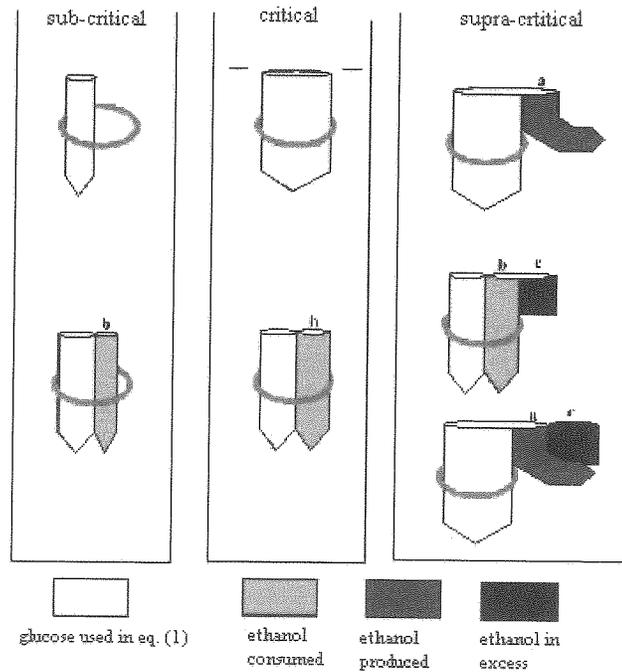


Figure 2: Metabolic mechanism of the baker yeast. The grey ring represents the respiratory bottleneck

The total specific growth rate, μ_t , is the sum of the growth rates for the three pathways

$$\mu_t = \mu_S^O + \mu_S^r + \mu_E^O. \quad (2.4)$$

The specific growth rates, μ_i , can be related to the corresponding substrate fluxes, q , and yield coefficients, Y , by

$$\mu_t = Y_{X/S}^O q_S^O + Y_{X/S}^r q_S^r + Y_{X/E}^{OE} q_E^O \quad (2.5)$$

where $Y_{X/S}^O$ and $Y_{X/S}^r$ represent the yield coefficients of biomass (X) in glucose (S) in the oxidative and fermentative phases, respectively; $Y_{X/E}^{OE}$ is the yield coefficient of biomass in ethanol (E) in the oxidative phase in ethanol.

Table 1: Baker's yeast kinetics equations for the respirative and respiro-fermentative regimes

$aq_S \leq q_O$ Respirative Regime	$\mu_S^O = Y_{X/S}^O \cdot q_S$ (2.11)
	$\mu_S^r = 0$ (2.12)
	$\mu_E^O = \min(\mu_{E_1}^O, \mu_{E_2}^O)$ (2.13)
$aq_S > q_O$ Respiro-Fermentative Regime	$\mu_S^O = Y_{X/S}^O \cdot \frac{q_O}{a}$ (2.14)
	$\mu_S^r = Y_{X/S}^r \cdot (q_S - \frac{q_O}{a})$ (2.15)
	$\mu_E^O = 0$ (2.16)

As ethanol uptake is influenced by the priority of glucose uptake, which functions as an inhibitor, the specific growth rate on ethanol can be described as

$$\mu_E^O = \mu_E^{\max} \frac{E}{E + K_E} \frac{K_i}{S + K_i} \tag{2.6}$$

where μ_E^{\max} is the maximal specific growth rate, K_i is the inhibition parameter and K_E is the saturation parameter.

However, this equation holds true only if there is an available respiratory capacity of the cells.

The glucose uptake, q_S , is slightly different because it follows two metabolic pathways: oxidative and fermentative

$$q_S = q_S^O + q_S^r \tag{2.7}$$

The glucose, q_S , and oxygen, q_O , uptake follows Monod kinetics, respectively

$$q_s = q_s^{\max} \frac{S}{S + K_s} \tag{2.8}$$

$$q_O = q_O^{\max} \frac{O}{O + K_O} \tag{2.9}$$

where q_s^{\max} is the maximal specific glucose uptake rate, K_S and K_O are saturation parameters and q_O^{\max} is the maximal specific oxygen (O) uptake rate.

From (2.1), it can be seen that the oxidative glucose uptake depends on the availability of dissolved oxygen, and may be defined as

$$q_s^o = \frac{q_O^O}{a} \tag{2.10}$$

where a is the stoichiometric coefficient of the oxygen in the respiratory pathway of glucose and q_O^O is the oxygen uptake on glucose.

Two situations may occur: excess of oxygen that implies no fermentative growth of biomass or lack of oxygen and consequently excess of glucose that implies no respiratory growth on ethanol.

Table 1 resumes mathematically the two situations that coexist [10].

Two auxiliary equations, (2.17) and (2.18), must be added for the estimation of the specific growth rate on ethanol, defined as:

$$\mu_{E_1}^O = \mu_E^{max} \frac{E}{E + K_E} \frac{K_i}{S + K_i} \quad (2.17)$$

and

$$\mu_{E_2}^O = \frac{Y_{X/O}^{OE}}{Y_{X/E}^{OE}} (q_O - a_{qS}). \quad (2.18)$$

The relevant kinetic data were taken from Sonnleitner and Käppeli [9] and [11].

2.2 Mechanistic model

Considering that the yield coefficients, Y 's, are constant and the dynamics of the gas phase can be neglected, the following set of differential equations was obtained:

mass balance for the biomass

$$\frac{dX}{dt} = (\mu_S^O + \mu_S^r + \mu_E^O - D) X \quad (2.19)$$

mass balance for the sugar

$$\frac{dS}{dt} = \left(-\frac{\mu_S^O}{Y_{X/S}^O} - \frac{\mu_S^r}{Y_{X/S}^r} \right) X + (S_f - S) D \quad (2.20)$$

where S_f is the substrate concentration in the feed and D is dilution rate (ratio feed rate/volume),

mass balance for the ethanol

$$\frac{dE}{dt} = \left(\frac{\mu_S^r}{Y_{X/E}^r} - \frac{\mu_E^O}{Y_{X/E}^{OE}} \right) X - DE \quad (2.21)$$

mass balance for the oxygen

$$\frac{dO}{dt} = \left(-\frac{\mu_S^O}{Y_{X/O}^O} - \frac{\mu_E^O}{Y_{X/O}^{OE}} \right) X - DO + OTR \quad (2.22)$$

mass balance for the carbon dioxide

$$\frac{dC}{dt} = \left(\frac{\mu_S^O}{Y_{X/C}^O} + \frac{\mu_S^r}{Y_{X/C}^r} + \frac{\mu_E^O}{Y_{X/C}^{OE}} \right) X - DC - CTR \quad (2.23)$$

accumulation of the working volume during the fed-batch process

$$\frac{dV}{dt} = DV \quad (2.24)$$

The gas transfer rates are given by:

$$OTR = K_L^O a(O^* - O) \quad (2.25)$$

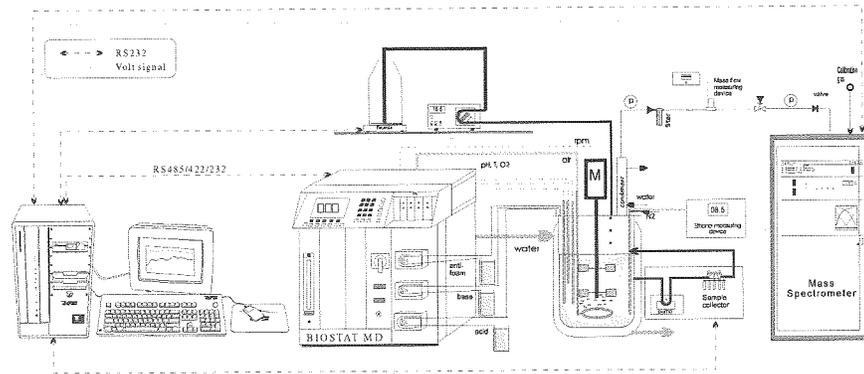


Figure 3: Experimental set-up for baker's yeast production

$$CTR = K_L^C a(C - C^*) \quad (2.26)$$

where $K_L^i a$ are overall mass transfer coefficients for oxygen and carbon dioxide and O^* and C^* are the corresponding equilibrium concentrations.

The set of equations (2.19–2.24) defines an IVP model and the set of equations (2.19–2.24, 2.25–2.26, 2.11–2.13 or 2.14–2.16) defines a DAE model [5]. These two different approaches were presented and discussed elsewhere [4].

3 Experimental set-up

In this section a brief description of the experimental set-up is presented. For more details see [12-13]. Figure 3 illustrates the macro experimental set-up used comprising a five-litre fermenter with temperature control, aeration and agitation. It is equipped with sensors for on-line measurement of environment variables, such as: temperature, pH and concentration of dissolved oxygen. These variables are monitored and controlled by a direct digital control unit (DCU, Biostat MD). Each variable has its own control loop, with appropriate parameters, which can be modified by the user or through the supervisor computer. The DCU actuates in each final control element, as for example, acid and base pumps for pH control. The pumps and electrodes calibration is done with help of this control unit.

The fed-batch fermentation starts with a two and a half litre volume substrate medium and an inoculum of 0.5 l, added through a peristaltic pump in a profile dictated by a control law. The substrate addition is monitored with the help of a balance, by means of mass variation of the glass that contains the glucose solution. The feeding ends when the maximum volume is attained (the fermentation time is about 15 to 20 hours).

The knowledge of liquid phase composition, in terms of state variables, is obtained by measuring the following variables: biomass, glucose, dissolved ethanol, oxygen and carbon dioxide. Each variable are measured and analysed by specific sensors and methodologies.

A computer and other hardware are linked to the DCU in order to implement data acquisition, monitoring and open and closed control loop strategies (in open loop to impose constant or variable feed flow and in closed loop to test PID laws and adaptive control algorithms). In

order to overcome the difficulties of off-line measures of biomass and glucose, the supervisor computer also determines their concentration by on-line estimation techniques.

4 Heuristic sensitivity analysis

As the baker's yeast fermentation is a highly non-linear system, differences in experimental and simulation results are significant. In order to overcome this problem, some of the model parameters are adjusted using an optimization technique based on Meyer and Roth method [6]. The objective function to be minimised, (4.27), is the sum of the errors associated to biomass, glucose and ethanol. Only three of the five state variables are used due to limitations on the optimization solver.

$$F_{obj} = \sum_{i=1}^{nexp} \left(\frac{X_i^{mod} - X_i^{exp}}{X_i^{exp}} \right)^2 + \sum_{i=1}^{nexp} \left(\frac{S_i^{mod} - S_i^{exp}}{S_i^{exp}} \right)^2 + \sum_{i=1}^{nexp} \left(\frac{E_i^{mod} - E_i^{exp}}{E_i^{exp}} \right)^2 \quad (4.27)$$

where subscripts *exp* and *mod* refer to experimental and simulation model results, respectively; *nexp* is the number of experimental points.

In order to quantify the performance of parameters estimation a new variable is defined, mean error, \overline{Er} :

$$\overline{Er} = \sqrt{\frac{F_{obj}}{3 \ nexp}} \quad (4.28)$$

being constant 3 the number of state variable considered in (4.27).

In the overall model (mechanistic and kinetics) there are 17 parameters to be studied (9 yield coefficients *Y* and 8 kinetic parameters). As previously mentioned only three state variables are considered in the optimization methodology. Biomass, glucose and ethanol are chosen due to their relevant role in the fermentation process. As a consequence, also only three parameters can be estimated simultaneously. As a first approach, an identification of the most relevant parameters in the model was performed. In this sense, relevant means the measure of their effect on the model. Heuristic sensitivity analyses were performed based on empiric evaluation of results. The first heuristic analysis calculates the most relevant kinetics parameters. The second heuristic analysis, with the previously identified kinetics parameters, estimates the most relevant yield coefficients. The methodology followed is described in the next section.

5 Modelling Simulation procedure

The simulation programs used for the sensitivity analysis of baker's yeast fed-batch production were developed in FORTRAN90 (adaptation of some subroutines described in [6]). All simulations were performed in a Pentium PC platform.

Five experiments were carried out with the initial conditions presented in Table 2. The experiments are named as EXPC0*i*, where *i* equal 1, 3, 4, 6 and 7. All the experiences were performed in the experimental set-up (Figure 3) previously described. EXPC07, considered for comparison purpose, was run in batch (*F*=0 l/h) and the others were performed as fed-batch

Table 2: Initial conditions in the five experiences

EXPC0i	S_f (g/l)	F (l/h)	Initial values					
			V (l)	X (g/l)	S (g/l)	E (g/l)	O (g/l)	C (g/l)
EXPC01	25	0.10	3.30	0.297	1.276	0.810	0.0066	0.0020
EXPC03	100	0.15	2.75	1.545	0.292	2.933	0.0066	0.0020
EXPC04	5	0.15	2.75	1.400	25.443	3.800	0.0066	0.0020
EXPC06	25	0.15	2.75	1.455	0.000	1.836	0.0066	0.0020
EXPC07	0	0	2.75	0.222	21	0.333	0.0066	0.0020

Table 3: Kinetics parameters in the sensitivity analysis run

<i>Simulation Run</i>	q_s^{max}	q_o^{max}	μ_e^{max}
0*	3.5	0.256	0.17
1	10.0	0.256	0.17
2	3.5	1.0	0.17
3	3.5	0.256	10.0
4	0.35	0.256	0.17
5	3.5	0.256	0.01
6	3.5	0.01	0.17

* literature values [9, 11].

fermentations, for different substrate flows (F). These values were chosen taking into account the set-up limitations.

6 Results and Discussion

In order to identify the most significant model parameters, several simulation runs were performed, changing one parameter at a time, and having as a reference the literature values [9, 11]. Due to their important role in the model definition, the kinetics parameters, q_s^{max} , q_o^{max} and μ_e^{max} , are the first to be chosen for sensitivity analysis. Table 3 resumes the values used (empirically assumed) in each simulation performed. Run 0 corresponds to the literature values.

Figure 4 shows the biomass (X), glucose (S) and ethanol (E) profiles obtained with the kinetics parameters presented in Table 3 (subscripts 0 to 6 correspond to the simulation runs). The dot points correspond to experimental data and the lines to simulation values.

Through Figure 4, it is clear that runs 4 and 6 change significantly the profiles (x4 and x6 profiles for biomass, s4 and s6 for glucose, and e4 and e6 for ethanol). These simulation runs correspond to the decrease on q_s^{max} and q_o^{max} kinetics parameters, respectively for glucose and oxygen (Table 3). For further study, these two parameters were identified to be the most

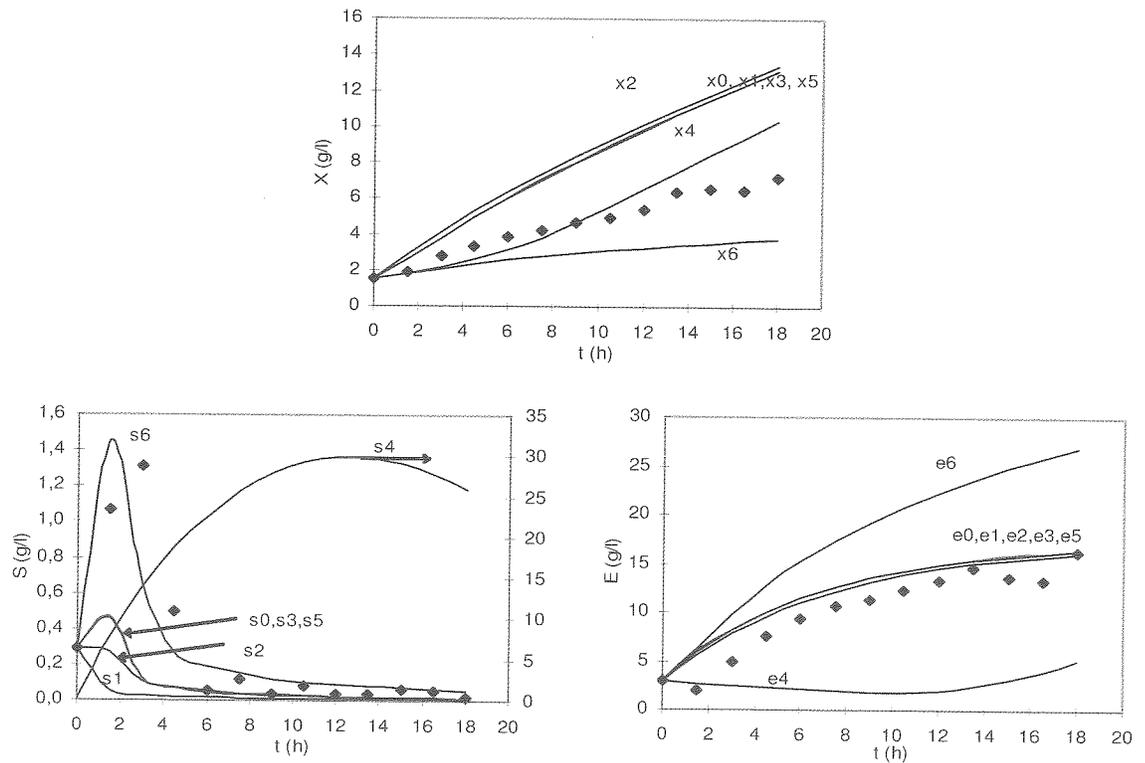


Figure 4: Biomass(X),glucose(S) and ethanol (E) profiles for the seven runs with the kinetics parameters presented in Table 3. The dot points correspond to experimental data and the lines to simulation values. Subscripts 0 to 6 correspond to the simulation runs.

relevant for baker's yeast production. The maximum specific growth rate on ethanol, μ_e^{max} , does not influence simulation state variables profiles (the x0, x3 and x5 profiles for biomass, the s0, s3 and s5 profiles for glucose and the e0, e3 and e5 profiles for ethanol are coincident).

Table 4 resumes the estimated values, for the two relevant parameters, q_s^{max} and q_o^{max} obtained in each run. The mean values of all the estimates and the literature values are also shown. These mean values enable an easily comparison with the literature ones. The objective function (4.27) and the mean error (4.28) obtained in each experiment are also presented.

There is in fact a large difference in the results obtained in the five experiences. The mean value for maximum uptake in glucose, q_s^{max} is similar to the literature but, q_o^{max} is approximately half of the value in the literature. Considering maximum glucose uptake parameter, it varies from 0.79 in EXPC04 to 8.31 in EXPC01. Taking into account the initial conditions in each run (Table 2) it is clear that as sugar concentration in feed stream decreases from 25 g/l in EXPC01 to 5 g/l in EXPC04, the glucose uptake also decreases sharply.

For a better understanding of the influence of these parameters in the X, S and E profiles, a graphical representation for EXPC03 was considered, Figure 5. The subscript (1) refers to literature parameters, (2) to the estimated values and (3) to the mean of estimated values of q_s^{max} and q_o^{max} . The dot points correspond to the experimental data.

Even with the estimated values for that experience, x2, s2 and e2 profiles in Figure 5,

Table 4: Estimated values for q_s^{max} and q_o^{max} , objective function and mean error in the five experiences under the initial conditions described in Table 2.

<i>Experiment</i>	q_s^{max}	q_o^{max}	<i>Objective Function</i> (4.27)	<i>Mean Error</i> (4.28)
EXPC01	8.31	0.154	9.35	0.490
EXPC03	3.16	0.114	11.54	0.544
EXPC04	0.79	0.071	3.33	0.304
EXPC06	4.65	0.184	6.90	0.421
EXPC07	1.78	0.1625	11.72	0.571
Mean	3.74	0.137		
Literature [9, 11]	3.5	0.256		

there are still a significant difference between the model and experimental results. Therefore, a second heuristic sensitivity analysis was carried out in order to improve the results for experiment EXPC03.

The same simulation program was employed, considering the initial conditions described in Table 2. The parameters referred below were selected and their values were changed arbitrary. This second analysis concludes that:

- q_s^{max} – only influences glucose profile;
- μ_e^{max} – not relevant;
- q_o^{max} – influences on the reverse order biomass and ethanol;
- K_{La} – influences on the reverse order biomass and ethanol;
- K_s – only influences glucose profile;
- K_o – influences on the reverse order biomass and ethanol;
- K_e and K_i – not relevant;
- Yield coefficients – $Y_{x/s}^r$, $Y_{x/e}^r$, $Y_{x/s}^o$ and $Y_{x/o}^o$ are relevant.

Considering EXPC03, with q_s^{max} and q_o^{max} values previously estimated an optimization run is performed in order to adjust $Y_{x/s}^r$ and $Y_{x/e}^r$ values. Figure 6 shows the new profiles obtained for biomass (X), glucose (S) and ethanol (E).

Comparing Figure 5, profiles (2), and Figure 6, there is a significant improvement; *i.e.* the experimental data is better adjusted by the model behavior.

In Table 5 the estimated and literature values for $Y_{x/s}^r$, $Y_{x/e}^r$ obtained are presented with the objective function and the mean error obtained.

With the new values for these four parameters, q_s^{max} , q_o^{max} , $Y_{x/s}^r$, $Y_{x/e}^r$, an optimization run is then performed in order to estimate $Y_{x/s}^o$ and $Y_{x/o}^o$. Figure 7 shows the best results attained

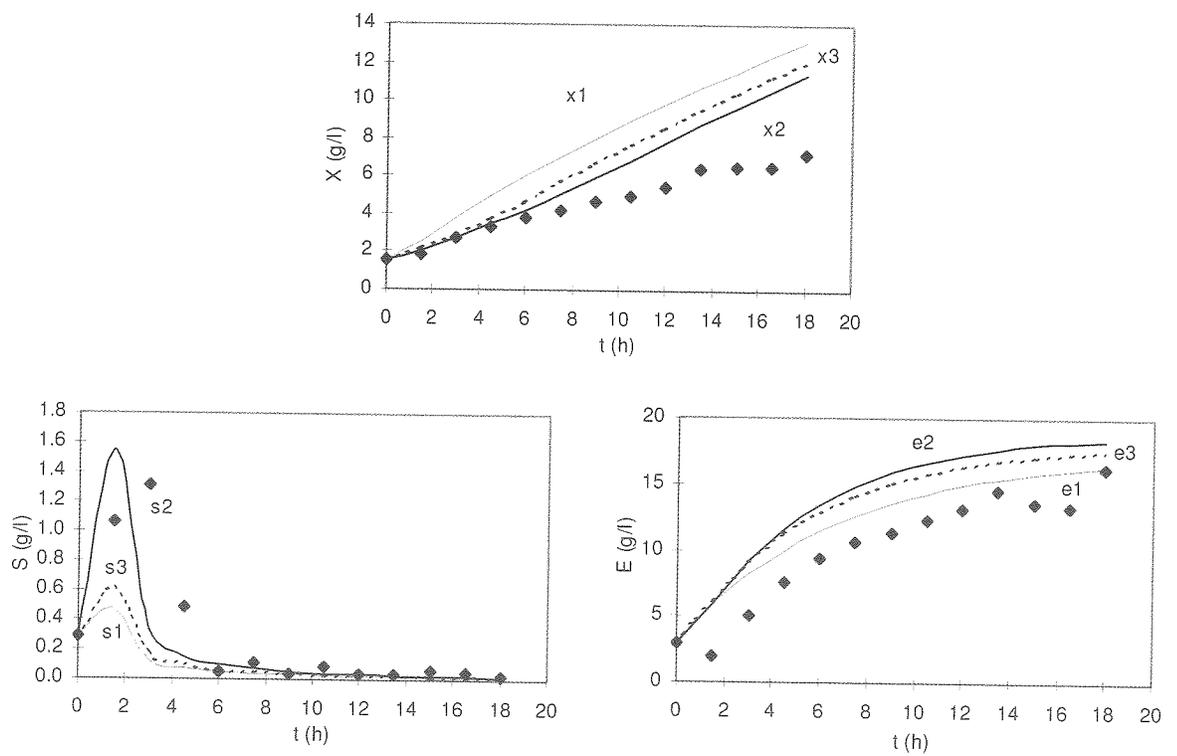


Figure 5: Biomass (X), glucose (S) and ethanol (E) profiles in experiment EXPC03. The subscript (1) refers to literature parameters, (2) to the estimated values and (3) to the mean of estimated values of $q_{x/s}^r$ and q_o^{max} . The dot points correspond to experimental data.

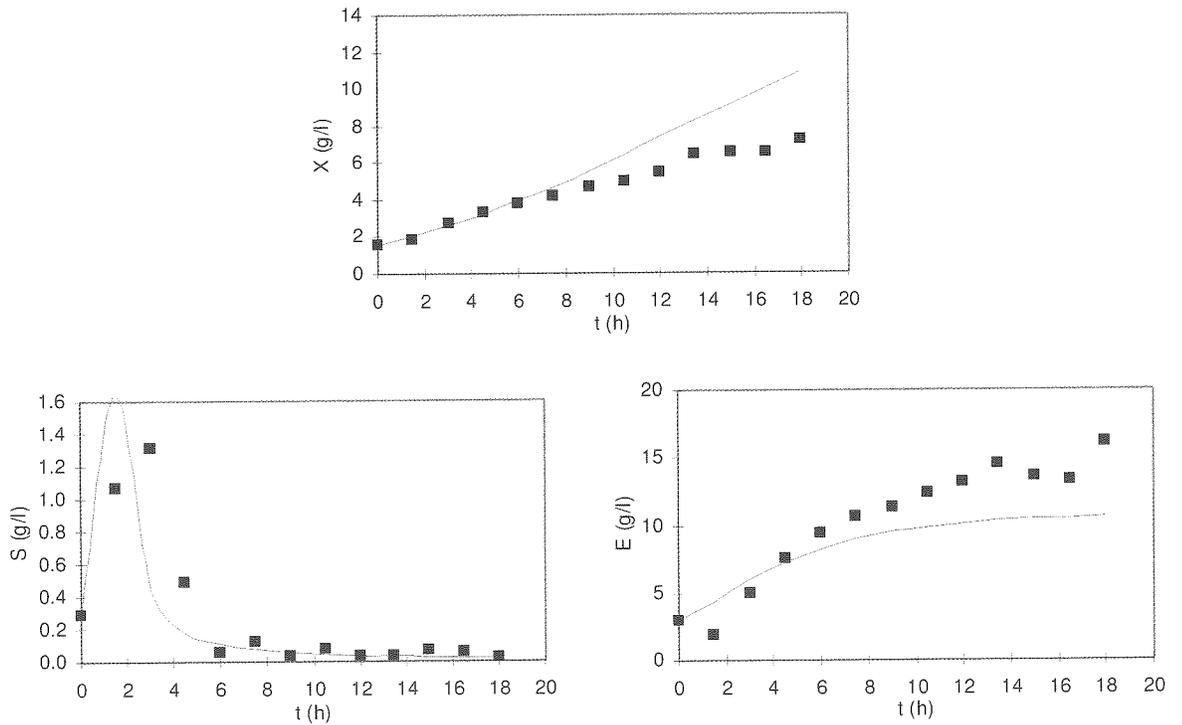


Figure 6: Biomass (X), glucose (S) and ethanol (E) in EXPC03 with q_s^{max} , q_o^{max} , $Y_{x/s}^r$, $Y_{x/e}^r$ estimated values (line) and experimental data (dot).

Table 5: Estimated values for $Y_{x/s}^r$ and $Y_{x/e}^r$ and objective function and mean error obtain for EXPC03.

<i>Experiment</i>	$Y_{x/s}^r$	$Y_{x/e}^r$	<i>Objective Function</i> (4.27)	<i>Mean Error</i> (4.28)
EXPC03	0.04	0.16	7.30	0.433
Literature [9, 11]	0.05	0.10		

for this new experiment. Biomass and ethanol are significantly improved. However, glucose behavior in the first four hours is slightly worst than the one obtained without $Y_{x/s}^o$ and $Y_{x/o}^o$ estimates (Figure 6). We must point out that the six parameters were adjusted in pairs.

The estimated and literature values for $Y_{x/s}^r$, $Y_{x/e}^r$ are presented in Table 6 in addition with the respective objective function and the mean error.

Globally, it is apparent that a variability of the microorganisms metabolic behavior has, in a mathematically point of view, a relation to the variability of the characteristic parameters, namely yield coefficients and kinetics values [14].

Table 6: Estimated values for $Y_{x/s}^o$ and $Y_{x/e}^o$ and objective function and mean error obtain for EXPCO3.

<i>Experiment</i>	$Y_{x/s}^r$	$Y_{x/e}^r$	<i>Objective Function</i> (4.27)	<i>Mean Error</i> (4.28)
EXPCO3	0.86	0.63	5.68	0.385
Literature [9, 11]	0.49	1.20		

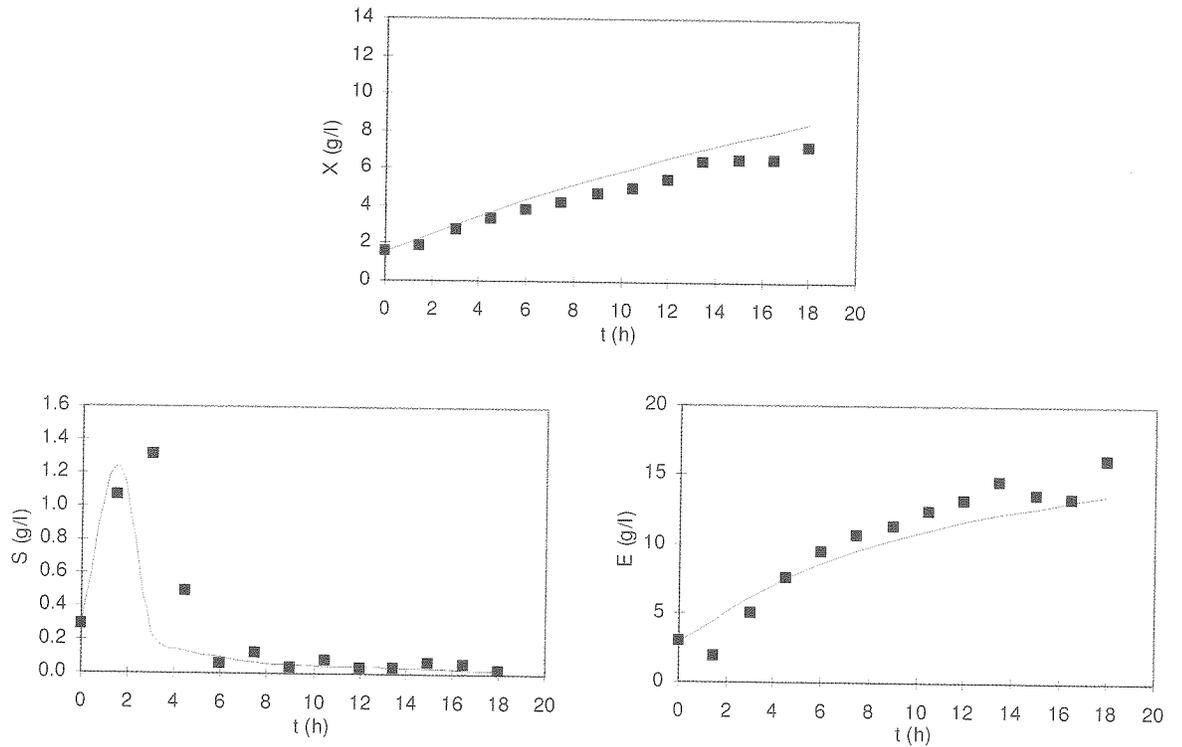


Figure 7: Biomass (X), glucose (S) and ethanol (E) in EXPCO3 with q_s^{max} , q_o^{max} , $Y_{x/s}^r$, $Y_{x/e}^r$, $Y_{x/s}^o$ and $Y_{x/o}^o$ estimated values (line) and experimental data (dots).

7 Conclusions

A deterministic model for fed-batch baker's yeast production was presented. The mathematical model, differential mass balance equations together with mass transfer and kinetics equations, considered as an IVP problem, was simulated in a FORTRAN90 based system.

The theoretical model was validated with a data set corresponding to five experimental open loop runs: four in fed-batch operation mode and one in batch mode.

As a significant discrepancy between simulated and experimental results was verified, there was a need for parameter estimation. In order to decide which parameters were the most relevant in model performance, a heuristic sensitivity analysis was taken. The maximum uptake rate for glucose and oxygen and the yield coefficients were seen to be the most significant model parameters. The number of parameter to be estimated at the same time was a software limitation (maximum three).

The apparent variability in microorganism behaviour motivates the thirst for process knowledge and understanding.

8 Suggestions for further research

As the optimization routine has a limitation in the number of parameters that can be estimated together (three) we are now starting a different method for estimating a larger number of parameters (depending only upon the experimental data available on each run), using a Minimum Square technique [15]. So, with the experimental fed-batch baker's yeast results, related to initial conditions (inputs) and final concentrations (outputs), yield coefficients and kinetic parameters are to be estimated together.

Nomenclature

C	dissolved carbon dioxide concentration	(g/l)
C*	carbon dioxide saturation concentration	(g/l)
E	ethanol concentration	(g/l)
\overline{Er}	mean error	
Fobj	objective function	
K_i	inhibition parameter	
$K_{L_a}^i$	a overall mass transfer coefficient for component i	(1/h)
O	dissolved oxygen concentration	(g/l)
O*	oxygen saturation concentration	(g/l)
S	glucose concentration	(g/l)
S_e	glucose concentration in the feed	(g/l)
V	liquid volume of reactor	(l)
X	biomass concentration	(g/l)
$Y_{x/s}^o$	yield coefficient biomass/glucose in oxidative pathway in glucose	
$Y_{x/s}^r$	yield coefficient biomass/glucose in fermentative pathway in glucose	
$Y_{x/e}^r$	yield coefficient biomass/ethanol in fermentative pathway in glucose	
$Y_{x/e}^{oe}$	yield coefficient biomass/ethanol in oxidative pathway in ethanol	
$Y_{x/o}^o$	yield coefficient biomass/oxygen in oxidative pathway in glucose	
$Y_{x/o}^{oe}$	yield coefficient biomass/oxygen in oxidative pathway in ethanol	
$Y_{x/c}^o$	yield coefficient biomass/carbon dioxide in oxidative pathway in glucose	
$Y_{x/c}^r$	yield coefficient biomass/carbon dioxide in fermentative pathway in glucose	
$Y_{x/c}^{oe}$	yield coefficient biomass/carbon dioxide in oxidative pathway in ethanol	

Small letters

a	a is the stoichiometric coefficient of the oxygen in the respiratory pathway of glucose	
nexp	number of experimental points	
q_o	total oxygen uptake	(1/h)
q_O^o	oxygen uptake on glucose	(1/h)
q_s	total glucose uptake	(1/h)
q_s^{max}	maximum glucose uptake	(1/h)
q_s^o	glucose uptake in the oxidative pathway	(1/h)
q_s^r	glucose uptake in the fermentative pathway	(1/h)
q_e^o	ethanol uptake in the oxidative pathway	(1/h)
q_o^{max}	maximum oxygen uptake	(1/h)
t	time	(h)

Greek symbols

μ	specific growth rate	(1/h)
μ_e^{max}	maximum oxidative specific growth rate on ethanol	(1/h)
μ_e^o	oxidative specific growth rate on ethanol	(1/h)
μ_{max}	maximum specific growth rate	(1/h)
μ_s^r	fermentative specific growth rate on glucose	(1/h)
μ_s^o	oxidative specific growth rate on glucose	(1/h)

References

- [1] Feyo de Azevedo, S., Pimenta, P., Soares, F. O., Ferreira, E. C., *Studies on on-line State and Parameter Estimation through a Real-time Process Simulator*, Modeling and Control of Biotechnical Processes (2nd IFAC Symp. and 5th Int. Conf. Computer Applications in Fermentation Technology, Keystone USA, 1992, M. Nazmul Karim, G. Stephanopoulos, Eds.), Pergamon Press, Oxford, 1992, pp.453-458.
- [2] Kristiansen, B. (Ed.), *Integrated Design of a Fermentation Plant-The Production of Baker's Yeast*, VCH, Weinheim, 1994.
- [3] Soares, F. O., *Monitorização e controlo de fermentadores – Aplicação ao fermento de padeiro*, Ph.D. Thesis, Universidade do Porto, 1997 (*in portuguese*).
- [4] Leão, C. P., Soares, F. O., *Two different strategies for baker's yeast fermentation process simulation*, in Recent Advances in Simulation, Computational Methods and Soft Computing, WSEAS Press, 2002, pp.11-16.
- [5] Brenan, K. E., Campbell, S. L., Petzold, L. D., *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*, Elsevier, New York, 1989.
- [6] Fayo de Azevedo, S., *Subrotinas em Fortran para utilização no sistema WangVS 80 da FEUP*, Internal Publication DEQ, FEUP, Porto, 1987 (*in Portuguese*).
- [7] Wang, F.-S., Sheu, J.-W., *Multiobjective parameter estimation problems of fermentation processes using a high ethanol tolerance yeast*, Chemical Engineering Science, 55 (18), 2000, pp. 3685–3695.
- [8] Nielsen, J., Villadsen, J., *Modelling of Microbial Kinetics*, Chem. Eng. Sci., 47:17/18, 1992, pp. 4225-4270.
- [9] Sonnleitner, B., Käppeli, O., *Growth of Saccharomyces cerevisiae is controlled by its Limited Respiratory Capacity: Formulation and Verification of a Hypothesis*, Biotech. Bioeng., 28, 1986, pp. 927-937.
- [10] Oliveira, R., Ferreira, E. C., Soares, F. O., Fayo de Azevedo, S., *A study on the convergence of observer-based kinetics estimators in stirred tank bioreactors*, Journal of Process Control, 6 (6), 1996, pp. 367-371.
- [11] Pomerleau, Y., *Modélisation et Contrôle d'un Procédé Fed-Batch de Culture des Levures à Pain (Saccharomyces cerevisiae)*. PhD Thesis, Université de Montreal, 1990 (*in French*).
- [12] Soares, F. O., Ferreira, E. C., Fayo de Azevedo, S., *Requisitos e Realização de uma Instalação Experimental para Estudos de Modelização e Controlo em Processos de Fermentação*, 1^o Encontro Português de Controlo Automático - CONTROLO 94, vol. II, Lisboa, Portugal, September, 1994, pp. 71-76 (*in portuguese*).
- [13] Soares, F. O., Correia, J.H., *Bioreactor-on-a chip: Application to Baker's Yeast Fermentation*, 1st Annual International IEEE-EMBS Special Topic Conference on Microtechnologies in Medicine & Biology, Lyon, FRANCE, 12-14 October, 2000, ISBN 0-7803-6603-4, pp. 45-48.
- [14] Fayo de Azevedo, S., Dahm, B., Soares, F. O., *Hybrid Modelling of Biochemical Processes: A Comparison with the Conventional Approach*, Computers and Chemical Engineering., 21, suppl, 1997, pp. S751-S756.
- [15] Leão, C. P., Soares, F. O., Fernandes, E. M. G. P., *Multiple Nonlinear Regression Analysis for the Baker's Yeast Fermentation Parameters Estimation*, Proceedings of the Workshop on Modeling and Simulation in Chemical Engineering, Coimbra, Portugal, 30 June-4 July, 2003, pp. 1-7.

REVISTA INVESTIGAÇÃO OPERACIONAL

Política Editorial

Investigação Operacional (IO) é a revista científica da APDIO - Associação Portuguesa de *Investigação Operacional*. A política editorial da IO é publicar artigos originais e de elevada qualidade que contribuam para a teoria, metodologia, técnicas e software de *Investigação Operacional* e a sua aplicação a diferentes campos. A Revista também publica artigos com revisões relevantes de temas de IO. Casos de sucesso na aplicação a problemas práticos são especialmente bem vindos.

Processo de Aceitação

Todos os manuscritos submetidos para publicação são revistos e aceites apenas com base na avaliação da sua qualidade, importância e adequação à política editorial. Será responsabilidade do Editor interpretar a avaliação dos revisores. A contribuição de cada artigo deve estar claramente evidenciada na Introdução. Critérios como a relação com literatura existente, comprimento e estilo do artigo são tidos em consideração. Uma indicação clara da viabilidade de aceitação do artigo é habitualmente dada na primeira fase de revisão do artigo.

Será requerido aos autores de um artigo aceite que transfiram os direitos de autoria para a APDIO, que assegurará a mais ampla disseminação possível de informação. Os volumes da Revista são publicados em papel, e distribuídos a todos os associados da APDIO, e em formato electrónico na rede SciELO - Scientific Electronic Library Online.

Resumos dos Artigos indexados em

IAOR - International Abstracts in Operations Research

Instruções aos Autores

1. *Submeter artigos* para publicação ao editor principal, de preferência por e-mail em Microsoft Word ou "Portable Document Format" (PDF) para jfo@fe.up.pt, ou por correio normal (quatro cópias) para o seguinte endereço: Prof. José Fernando Oliveira, Departamento de Engenharia Electrotécnica e de Computadores, Faculdade de Engenharia da Universidade do Porto, Rua Dr. Roberto Frias, 4200-465 PORTO, Portugal.
2. *Língua*. Os artigos devem ser escritos em Português, Inglês ou Espanhol.
3. *Os Manuscritos* devem ser impressos. Numerar as páginas consecutivamente.
4. *A primeira página* do manuscrito escrito em português ou em espanhol deve ter a seguinte informação: (a) Título; (b) nome, e-mail e afiliação institucional dos autores; (c) um resumo; (d) palavras-chave; (e) título em inglês (f) um resumo em inglês; (g) palavras-chave em inglês; (h) identificação do autor correspondente. Se o manuscrito for escrito em inglês, a primeira página deve ter a seguinte informação: (a) Título em inglês; (b) nome, e-mail e afiliação institucional dos autores; (c) um resumo em inglês; (d) palavras-chave em inglês; (e) identificação do autor correspondente.
5. *Agradecimentos*, incluindo informação sobre apoios, dever ser colocados imediatamente antes da secção de referências.

6. *Notas de rodapé* devem ser evitadas.
7. *Formulas* que são referenciadas devem ser numeradas consecutivamente ao longo do manuscrito como (1), (2), etc. do lado direito.
8. *Figuras*, incluindo grafos e diagramas, devem ser numerados consecutivamente em numeração árabe.
9. *Tabelas* devem ser numeradas consecutivamente em numeração árabe.
10. *Referências*. Citar apenas as mais relevantes e listar só as que são citadas no texto. Indicar as citações no texto através de parênteses rectos, e.g., [4]. No final do artigo listar as referências alfabeticamente por apelido do primeiro autor e numerá-las consecutivamente, de acordo com o seguinte formato: Artigos: autore(s), título, nome e volume da revista (ou livro, mas neste caso incluir o nome dos editores), ano e páginas. Livros: Autor(es), título, editor, ano.
11. *Artigos aceites* devem ser enviados pelo autor ao editor, de preferência na forma de um ficheiro fonte em LaTeX com ficheiros EPS para as figuras, juntamente com um ficheiro PDF ou Postscript. Em alternativa, ficheiros fonte em Word são também aceites. Para garantir uma boa qualidade gráfica, as figuras devem ser em formato vectorial; formatos raster como JPG, BMP, GIF, etc. devem ser evitados.
12. *Provas dos artigos* serão enviadas por e-mail como ficheiros PDF para o autor correspondente. Corrigir as provas cuidadosamente, e restringir as correcções apenas aos pontos em que as provas diferem do manuscrito. Desvios à versão aceite pelo editor são apenas possíveis com a autorização prévia e explícita do editor. Trinta separatas de cada artigo são enviados gratuitamente ao autor correspondente.

Informação sobre a Publicação

Investigação Operacional (ISSN 0874-5161) está registada na Secretaria de Estado da Comunicação Social sob o número 108335. Os volumes da Revista são publicados em papel, e distribuídos a todos os associados da APDIO, e em formato electrónico na rede SciELO - Scientific Electronic Library Online. O preço da assinatura anual é de 25 euros. Os volumes são enviados por correio normal. Informação adicional sobre assinaturas pode ser solicitada ao Secretariado da APDIO- CESUR, Instituto Superior Técnico, Av. Rovisco Pais, 1049-001 LISBOA, Portugal. Tel. +351 218 407 455 - www.apdio.pt - apdio@cesur.civil.ist.utl.pt

JOURNAL INVESTIGAÇÃO OPERACIONAL

Editorial Policy

Investigação Operacional (IO) is the scientific journal of APDIO - Associação Portuguesa de *Investigação Operacional* (the Portuguese Operational Research Association). The editorial policy of IO is to publish high quality and original articles that contribute to theory, methodology, techniques and software of Operational Research (OR) and its application to different fields. It also publishes articles with relevant reviews of OR subjects. Cases of successful application of OR to practical problems are specially welcome.

Acceptance Process

All manuscripts submitted for publication are refereed and accepted only on the basis of its quality, importance and adequacy to the editorial policy. It will be the responsibility of the Editor to interpret the referee's assessment. The contribution of each paper should be clearly stated in the introduction. Criteria such as relationship with existing literature, length and style are taken into account. A clear indication on the suitability of a manuscript is usually provided after the first round of refereeing. The authors of an accepted paper will be asked to transfer its copyright to the publisher, which will ensure the widest possible dissemination of information. The volumes of the journal are published in hardcopies, which are distributed to all APDIO associates, and in electronic format in SciELO - Scientific Electronic Library Online.

Articles are abstracted/indexed in

IAOR - International Abstracts in Operations Research

Instructions to Authors

1. *Submit papers* for publication to the main editor, preferably by e-mail in Microsoft Word or "Portable Document Format" (PDF) to jfo@fe.up.pt, or by ordinary mail (four copies) to the following address: Prof. José Fernando Oliveira, Departamento de Engenharia Electrotécnica e de Computadores, Faculdade de Engenharia da Universidade do Porto, Rua Dr. Roberto Frias, 4200-465 PORTO, Portugal.
2. *Language.* Papers must be written in Portuguese, English or Spanish.
3. *Manuscripts* should be typewritten or typeset. Number the pages consecutively.
4. *The first page* of the manuscript written in English should contain the following information: (a) Title; (b) names, e-mails and institutional affiliations of the authors; (c) an abstract; (d) keywords (f) identification of the corresponding author.
5. *Acknowledgements*, including support information, should be placed prior to the references section.
6. *Footnotes* should be avoided.
7. *Formulas* that are referred to should be numbered consecutively throughout the manuscript as (1), (2), etc. on the right.

8. *Figures*, including graphs and diagrams, should be numbered consecutively in Arabic numbers.
9. *Tables* should be numbered consecutively in Arabic numbers.
10. *References*. Cite only the most relevant references and list only those cited in the text. Indicate citations in the text by bracketed numbers, e.g., [4]. At the end of the paper list the references alphabetically by the surname of the first author and number them consecutively, according to the following formats: Articles: author(s), title, name and number of the journal (or book, but in this case include the editors names), year, pages. Books: Author(s), title, publisher, year.
11. *Accepted papers* are to be sent by the author to the editor, preferably in the form of a source file in LaTeX and EPS files for the figures together with a PDF or postscript file. Alternatively, source files in Word are also accepted. To ensure good publishing quality the figures should be in vector formats; raster formats like JPG, BMP, GIF, etc. should be avoided.
12. *Page proofs* will be e-mailed as a PDF file to the corresponding author. Correct proofs carefully, and restrict corrections to points at which the proof is at variance with the manuscript. Deviations from the version accepted by the editor are only possible with the prior and explicit approval of the editor. Thirty offprints of each paper are supplied free of charge to the corresponding author.

Publication information

Investigação Operacional (ISSN 0874-5161) is registered in the *Secretaria de Estado da Comunicação Social* under number 108335. The volumes of the journal are published in hardcopies, which are distributed free of charge to all APDIO associates, and in electronic format in ScieLO - Scientific Electronic Library Online. Subscription price is 25 euros. Issues are sent by standard mail. Additional subscription information is available upon request from APDIO Secretariat - CESUR, Instituto Superior Técnico, Av. Rovisco Pais, 1049-001 LISBOA, Portugal. Tel. +351 218 407 455 - www.apdio.pt - apdio@cesur.civil.ist.utl.pt

ÍNDICE

D.F. Pires, A.G. Martins, C.H. Antunes <i>Modelo multiobjectivo para apoio à localização de condensadores em redes radiais de distribuição</i>	139
E. Nunes, J. Faria, M. Matos <i>Abordagem hierárquica para avaliação da fiabilidade de sistemas de produção complexos com comportamento não-markoviano</i>	159
M.H. Mello, H.L. Quintella, J.C. Mello <i>Avaliação do Desempenho de Alunos Considerando Classificações Obtidas e Opiniões dos Docentes</i>	187
S.A. Araujo, M.N. Arenales <i>Planejamento e Programação da Produção numa Fundição Cativa Automatizada de Grande Porte</i>	197
J. Oliveira, M. Carvalho, J. Carvalho <i>Programação de operações em sistemas com tempo de processamento variável e com buffers de preparação</i>	211
L. Cavique <i>Graph-Based Structures for the Market Baskets Analysis</i>	233
C.P. Leão, F.O. Soares <i>Heuristic Sensitivity Analysis for Baker's Yeast Model Parameters</i>	247

