

Aplicação de modelos de redes de filas abertas no planejamento do sistema *job-shop* de uma planta metal-mecânica

Cláudio Rogério Negri da Silva

Reinaldo Morabito



Resumo

Este artigo ilustra a aplicação de modelos de redes de filas abertas para auxiliar no processo de tomada de decisões de projeto (ou reprojeto) e planejamento de sistemas discretos de manufatura. Utilizam-se métodos aproximados de decomposição paramétrica, para avaliar as medidas de desempenho da rede. Por meio de um estudo de caso de uma rede job-shop de uma fábrica metal-mecânica, mostra-se que estes métodos são adequados e efetivos para estimar, por exemplo, os estoques em processo (WIP) e os leadtimes de produção dos produtos. Mostra-se ainda que modelos de programação matemática de alocação de capacidade nas estações da rede, baseados nos métodos de decomposição, também são efetivos para otimizar o desempenho da rede em diferentes configurações. Curvas de trade-off entre investimento em capacidade e WIP ou leadtime são geradas e são úteis para auxiliar um gerente a decidir quanto alocar de capacidade na rede para atingir metas de desempenho. Estas curvas também são úteis para analisar o efeito em termos de capacidade se a variabilidade das chegadas externas de produtos, no mix de produtos ou na taxa de produção da rede muda.

Palavras-chave: Projeto de sistemas discretos de manufatura. Métodos de avaliação de desempenho. Redes de filas abertas. Modelos de otimização. Curvas de trade-off.

1 Introdução

A gestão de sistemas de manufatura tem se tornado mais complexa na medida em que: novos produtos estão surgindo, a demanda de produtos é incerta, os ciclos de vida dos produtos são mais curtos, e uma grande variedade deles compete pelos mesmos recursos. Modelos de redes de filas podem ser utilizados para representar sistemas de manufatura e auxiliar no processo de tomada de decisões de projeto (ou reprojeto) e planejamento destes sistemas. As redes são formadas de vários sistemas de filas conectados entre si, com usuários deslocando-se entre eles para receber serviço. Os nós da rede representam as estações de trabalho (*shops*); os produtos (*jobs*) correspondem aos usuários requerendo serviço das estações; e os arcos conectando os nós correspondem às rotas dos produtos.

Redes de filas são em geral classificadas como: redes de Jackson, aquelas cujos processos de chegadas externas de produtos são Poisson, e os tempos de serviço em cada estação têm distribuição exponencial, ou redes genéricas, aquelas cujos processos de chegadas externas não necessariamente são Poisson e/ou os de serviço não necessariamente são exponenciais, ou seja, são descritos por distribuições genéricas. As redes podem ser abertas

(o número de produtos circulando na rede pode variar a cada instante) ou fechadas (o número de produtos na rede é constante), sendo as abertas o foco deste estudo. Métodos exatos de avaliação de desempenho existem somente para redes de Jackson, mas a representação de sistemas de manufatura como redes de Jackson, em geral, superestima a variabilidade dos processos de serviço (e.g., BITRAN; TIRUPATI, 1988). Portanto, métodos aproximados são em geral utilizados para estimar o desempenho de redes mais genéricas, particularmente os sistemas discretos de manufatura do tipo *job-shop*. Estas aproximações configuram métodos descritivos, ou seja, que descrevem o desempenho da rede para uma dada configuração.

Neste artigo são utilizadas aproximações geradas pelo método de decomposição paramétrico, cujo objetivo é estender para redes mais gerais o resultado obtido para as redes de Jackson: cada nó da rede pode ser tratado como se fosse estocasticamente independente, e desta maneira, a análise de medidas de congestão da rede é simplificada. O método usa apenas dois parâmetros para caracterizar cada processo de chegada e cada processo de serviço em

cada estação, um para descrever a taxa média e o outro para descrever a variabilidade. Vários trabalhos em literatura abordaram aproximações e utilizaram o método de decomposição para avaliar o desempenho de redes de filas abertas, dentre os quais podem ser citados, por exemplo, Albin (1982, 1984), Bitran e Morabito (1995, 1996), Bitran e Sarkar (1994a, 1994b), Bitran e Tirupati (1988, 1989a, 1989b), Boxma et al. (1990), Buzacott e Shanthikumar (1993), Kouvelis e Tirupati (1991), Kuehn (1979), Ketzenberg et al. (2003), Morabito (1998), Segal e Whitt (1989), Shanthikumar e Buzacott (1981), Souza et al (2001, 2002), Suri et al. (1993), Van Vliet e Rinnooy Kan (1991) e Whitt (1983a, 1983b, 1994, 1995).

Aqui, os métodos descritivos de redes de filas abertas são aplicados para apoiar decisões tático-estratégicas no projeto e planejamento de sistemas *job-shops*. Uma questão central é a seleção entre as várias configurações para a rede, ou mais especificamente, como os recursos devem ser adequadamente distribuídos para alocar capacidade nas várias estações. Para isso, eles são utilizados em modelos prescritivos de programação matemática, ou seja, que prescrevem a melhor configuração do sistema, otimizando certo critério e satisfazendo determinadas restrições. A alocação insuficiente de capacidade pode causar altos níveis de inventário em processo (*work-in-process* ou WIP) e longos *leadtimes* de produção (tempo total que um produto leva para ser produzido, ou seja, a soma dos tempos de espera em fila mais os tempos de processamento em cada estação). Por outro lado, o excesso de capacidade pode resultar em desperdício de recursos onerosos, devido aos baixos níveis de utilização dos equipamentos. Alguns exemplos de trabalhos tratando modelos de otimização de redes de filas abertas em sistemas *job-shops* são: Bitran e Dasu (1992), Bitran e Morabito (1996, 1999), Bitran e Sarkar (1994b), Bitran e Tirupati (1989a, 1989b), Boxma et al. (1990), Bretthauer (1996), Calabrese (1992), Frenk et al. (1994), Morabito (1998), Seshadri e Pinedo (1999), Shanthikumar e Xu (2000), Souza e Ketzenberg (2002), Souza et al (2002), Sundarraj et al. (1994) e Van Vliet e Rinnooy Kan (1991).

Bitran e Dasu (1992), Bitran e Morabito (1999) e Govil e Fu (1999) classificam os problemas de projeto e planejamento de sistemas de manufatura em três classes: a) desempenho desejado do sistema (SP1, *Strategical Problem 1*); b) desempenho ótimo do sistema (SP2, *Strategical Problem 2*); e c) partição da instalação (SP3, *Strategical Problem 3*). Na classe SP1, o objetivo é minimizar o investimento no sistema sujeito às restrições de desempenho desejado para o sistema. Medidas de desempenho típicas são WIP, *leadtime* médio dos produtos e utilização média das estações. Na classe SP2 deseja-se otimizar o desempenho do sistema sujeito a limitações orçamentárias. Tanto o problema SP1 quanto o SP2

envolvem um *trade-off* (conflito) entre o capital de investimento e o capital de trabalho (ou o nível de serviço), se o WIP (ou o *leadtime*) for a medida de desempenho adotada. A classe SP3 busca particionar a planta em subplantas menores (plantas dentro da planta) para reduzir a complexidade do sistema e melhorar o desempenho global da rede. Esta medida pode requerer adição de capacidade e, portanto, há um *trade-off* entre o custo de adição de capacidade e a redução de complexidade gerencial do sistema. Os problemas SP1 e SP2 são objetos de estudo deste artigo; o problema SP3 foi recentemente estudado em Silva e Morabito (2006). No presente artigo objetiva-se:

- a) verificar a aplicabilidade e adequação de métodos de avaliação de desempenho (baseados em redes de filas genéricas, especificamente os aproximados por decomposição) para estimar o WIP, o *leadtime* dos produtos e o nível de utilização das estações de um sistema *job-shop* na prática. Para isso, realizou-se um estudo de caso em uma empresa do setor metal-mecânico e os resultados obtidos são aqui apresentados e analisados. Estes resultados são validados por modelos de simulação e análise do tomador de decisões da empresa; e
- b) verificar a aplicabilidade e adequação dos modelos de otimização SP1 e SP2 neste estudo de caso. Curvas de *trade-off* (ou de Pareto), obtidas através de avaliação paramétrica das medidas de desempenho dos problemas SP1 e SP2, são geradas para o estudo de caso com vistas a auxiliar no processo de tomada de decisões, particularmente aquelas que dizem respeito à alocação de capacidade na rede considerando alternativas discretas de capacidade. Ainda, observar a sensibilidade das soluções apresentadas a variações nos parâmetros de entrada da rede, tais como a variabilidade do processo de chegadas dos produtos, o *mix* de produtos e a taxa de produção da rede (*throughput*).

Estes objetivos vêm de encontro às perspectivas para pesquisas futuras propostas em Bitran e Morabito (1996, 1999); veja também Kouvelis et al. (2005). Poucos trabalhos analisaram e reportaram a aplicação prática de modelos de redes de filas em sistemas *job-shops* reais, tanto para avaliar o desempenho da rede, quanto para prescrever configurações para a rede que otimizam seu desempenho. Este artigo está organizado da seguinte maneira: a seção 2 discute brevemente redes de filas abertas e métodos de avaliação de desempenho, a seção 3 estuda resumidamente modelos de alocação/relocação de capacidade discretos sujeitos a restrições de limite de WIP e *leadtime* de produção. A seção 4 é importante para os objetivos deste trabalho porque apresenta e analisa os resultados computacionais obtidos ao aplicar as aproximações e os modelos de alocação de capacidade no estudo

de caso. Seis sistemas diferentes de aproximações foram testados, e as estimativas de medidas de desempenho obtidas foram comparadas com os resultados obtidos por simulação, para verificar qual dos sistemas era o mais apropriado. O sistema escolhido foi então utilizado num procedimento de otimização para alocação de capacidade discreta na rede. Também se verificou a sensibilidade das soluções a mudanças nos parâmetros de variabilidade de chegadas externas das classes de produtos. Por fim, a seção 5 discute as conclusões e perspectivas para pesquisas futuras.

2 Redes de filas abertas e métodos de avaliação de desempenho

Sistemas de manufatura consistem basicamente de máquinas e estações de trabalho, onde operações são realizadas sobre itens, partes, submontagens e montagens, para criar produtos que serão distribuídos para clientes. Os componentes de movimentação de materiais e de estocagem permitem que os itens se movam de estação para estação, que partes apropriadas estejam disponíveis para montagem, e que o trabalho seja mantido até poder entrar nas estações para processamento. Os sistemas discretos de manufatura são aqueles em que cada item processado é distinto. Por simplicidade, um item, parte, submontagem ou montagem processada por uma máquina ou estação é chamado simplesmente de item ou produto. Sistemas discretos podem ser classificados em função do volume e variedade dos itens/produtos. Tanto a variedade dos itens/produtos manufaturados pelo sistema (ou seja, o escopo), quanto o volume de cada item/produto produzido (ou seja, a escala), interferem diretamente no projeto e operação do sistema.

O presente estudo tem como foco o *job-shop*, que é uma forma organizacional do sistema de manufatura com capacidade de escopo (ou seja, capacidade de processamento de várias classes de produtos), mas com limitação em capacidade de escala devido aos problemas associados ao controle e movimentação dos produtos e ao tempo de preparação (*setup time*) das máquinas. Este sistema também é referido como layout por processo (ou layout funcional), dado que máquinas desempenhando funções semelhantes são agrupadas em um mesmo departamento ou estação. As máquinas podem ser intercambiáveis, ou seja, uma mesma máquina pode realizar operações em produtos pertencentes a classes diferentes, o que exige um certo tempo de preparação das máquinas. Daí, o sistema ser caracterizado em geral por longos *leadtimes* e altos níveis de WIP. Vantagens e desvantagens do emprego deste tipo de sistema de manufatura estão discutidas, por exemplo, em Askin e Standridge (1993), Buzacott e Shanthikumar (1993) e Krajewski e Ritzman (1996).

Redes de filas de manufatura são caracterizadas pelo número de estações (nós), roteiro dos produtos (seqüência de operações), e pelo tipo de rede (aberta, fechada ou mista). O número de nós na rede corresponde ao número de estações. O roteiro, determinístico ou probabilístico, pode ser seqüencial, seqüencial com realimentação, de montagem, cíclico e acíclico. Rotas probabilísticas podem ser empregadas para representar, por exemplo, a probabilidade de quebra de máquina ou a necessidade de retrabalho em um produto. Uma estação também pode ser visitada mais de uma vez por um mesmo produto para realizar operações diferentes. Neste trabalho, sistemas de manufatura são modelados por meio de redes de filas abertas, em que o número total de produtos circulando na rede é uma variável aleatória e pode variar pela chegada ou saída externa de produtos. Para trabalhos modelando sistemas logísticos de suprimento, distribuição e movimentação interna por meio de redes de filas, o leitor pode consultar, por exemplo, Kerbache e Smith (2004), Nakano e Ohno (1999) e Warsing et al. (2001).

Para especificar um nó na rede, devem ser identificados os processos que descrevem a chegada e o serviço de produtos, assim como a estrutura e a disciplina de serviço dos servidores. O processo de chegada é descrito pelo intervalo dos tempos entre chegadas de produtos no sistema, e pode ser determinístico ou probabilístico. A chegada na estação pode ser individual ou em lote (*batch*). Em determinadas situações é conveniente considerar um lote como sendo um produto, conforme é visto na seção 4. A segunda grandeza estatística que deve ser descrita é a quantidade de tempo demandada para processar o produto no servidor, usualmente referida como tempo de serviço, que pode ser determinístico ou probabilístico. Produtos podem ser processados individualmente ou em lotes. Em determinadas situações é conveniente representar um conjunto de recursos, como máquinas, operadores, ferramentas, etc, como uma única máquina, conforme é feito na seção 4. Com relação à estrutura e disciplina de serviço da estação, é preciso especificar: a) capacidade da fila; b) número de servidores; c) número de classes de produtos; e d) disciplina de atendimento, por exemplo, FCFS (*first-come-first-serve*).

Bitran e Tirupati (1988), Suri et al. (1993), Govil e Fu (1999) e outros autores discutiram classificações para os métodos de avaliação de desempenho, segundo os métodos de abordagem: métodos exatos; métodos aproximados; simulação; e técnicas relacionadas. Métodos exatos existem para os sistemas Markovianos (Redes de Jackson, de Kelly e BCMP). A falta de sucesso em obter soluções exatas para redes mais gerais motivou pesquisadores a desenvolverem aproximações para avaliar medidas de desempenho. Os métodos de aproximação podem ser classificados em: a) métodos de decomposição paramétricos; b) aproximação por difusão; c) análise

operacional; e d) análise do valor médio. Para os objetivos deste trabalho foram aplicadas as aproximações geradas pelos métodos de decomposição, por serem bastante flexíveis e de grande potencial para aplicações práticas. Estas aproximações inspiraram o desenvolvimento de vários pacotes comerciais de *software*, tais como: *Queueing Network Analyzer* (QNA), *ManuPlan*, *MPX*, e *Queueing Network Analysis Package* (QNAP). Outros *softwares* baseados em teoria de filas para análise de sistemas de manufatura estão descritos em Govil e Fu (1999).

A simulação discreta pode ser utilizada na ausência de resultados analíticos exatos ou aproximados. Ela tem sido empregada para avaliar sistemas complexos de redes de filas; é uma abordagem que permite o uso de hipóteses mais detalhadas e que estejam mais próximas da realidade operacional. Entretanto, o custo está relacionado aos requisitos computacionais, exigindo longos tempos de processamento e análise estatística, o que em geral implica em um pequeno número de alternativas de configuração sendo testadas (BITRAN; TIRUPATI, 1988). Além disso, há dificuldades para integrar simulação dentro de modelos de otimização, para prescrever a melhor configuração para o sistema. Neste trabalho, a simulação é empregada apenas como uma forma de validar os resultados gerados pelas aproximações analíticas. Leitores interessados em simulação podem consultar Law e Kelton (1991), Schriber (1991) e Banks et al. (1996).

Os métodos de decomposição paramétricos aproximam os processos de chegada de produtos por processos de renovação, decompõem a rede em estações individuais e as consideram como se fossem estocasticamente independentes. Tais métodos representam os processos de chegada e serviço por poucos parâmetros, daí o nome paramétrico. As medidas de congestão de cada estação são então descritas por fórmulas aproximadas que dependem somente destes parâmetros. Métodos de decomposição envolvem três passos essenciais (SHANTHIKUMAR; BUZACOTT, 1981; WHITT, 1983a, 1983b; GOVIL; FU, 1999):

- a) **passo 1:** Análise das interações entre as estações da rede de filas;
- b) **passo 2:** Decomposição da rede de filas em subsistemas com uma única estação, e avaliação das medidas de desempenho de cada estação; e
- c) **passo 3:** Recomposição dos resultados obtidos pelas estações decompostas com avaliação das medidas de desempenho da rede.

O Passo 1 é o mais importante e envolve três processos básicos para analisar a interação entre as estações: a) superposição das chegadas: representa a chegada agregada de produtos em uma estação, combinação dos fluxos individuais de chegadas de outras estações; b) partidas:

é o resultado da combinação deste processo de chegada agregado com o processo de serviço; e c) separação ou decomposição das partidas: decompõe o processo de partida agregada dos produtos nos fluxos individuais de partidas para outras estações. Por meio da solução de sistemas lineares derivados destes três processos básicos, o Passo 1 produz os seguintes parâmetros para aproximar os processos de chegada e de serviço em cada estação j da rede: a média $E(a_j)$ (ou taxa de chegadas de produtos $\lambda_j = 1/E(a_j)$) e o coeficiente quadrático de variação ca_j dos intervalos de tempo entre chegadas de produtos, e a média $E(s_j)$ (ou taxa de serviço de produtos $\mu_j = 1/E(s_j)$) e o coeficiente quadrático de variação cs_j dos tempos de serviço em cada estação j da rede, respectivamente (o coeficiente quadrático de variação de uma variável aleatória x é definido pela razão entre a variância de x e o quadrado do valor esperado de x , $cx = V(x)/E(x)^2$). Bitran e Tirupati (1988), Whitt (1983a) e outros autores afirmam que estes dois parâmetros são suficientes para descrever redes de filas genéricas com precisão usualmente aceitável para as decisões tomadas na prática.

Em Bitran e Morabito (1995, 1996) foram analisados seis diferentes sistemas lineares nas variáveis ca_j para o Passo 1, conhecidos da literatura, que foram aplicados neste trabalho. Estas aproximações referem-se a sistemas de redes de filas $GI/G/m$ e múltiplas classes de produtos com roteiros de fabricação determinísticos (i.e., a seqüência de nós visitados pelas classes de produtos é definida a priori). Os Passos 2 e 3 utilizam fórmulas em função dos parâmetros do Passo 1, para estimar medidas de desempenho para cada estação e a rede como um todo. Para mais detalhes de como estas aproximações são derivadas, o leitor pode consultar a referência acima e as referências nela citadas.

3 Modelos de alocação/realocação de capacidade

Este trabalho está preocupado em analisar questões tático-estratégicas de médio a longo prazo, relacionadas ao projeto/reprojeto e planejamento de sistemas de manufatura *job-shop* representados por redes de filas abertas. O projeto do sistema envolve, entre outros (BITRAN; DASU, 1992): escolha de tecnologia; seleção e alocação de máquinas nas plantas; alocação de produtos nas plantas; escolha da localização das plantas; e projeto do sistema de distribuição. Neste trabalho, o problema de projeto está restrito apenas aos dois primeiros fatores. Se várias estações de trabalho operam próximas a suas capacidades e há uma considerável diversidade nos requisitos de processamento dos diferentes produtos, então as filas nas estações tendem a ser longas; além disso, a variância dos tempos que um produto gasta no sistema tende a ser alta. As tarefas de um gerente operacional incluem atri-

buir e prever prazos de produção para produtos, controlar o fluxo de produtos na instalação tal que os prazos sejam respeitados, controlar os níveis de WIP, responder a perdas não antecipadas de capacidade (devido a fatores como quebra de máquina ou incertezas de rendimento), e empenhar-se para melhorar processos e produtos.

O projeto de sistemas envolve *trade-offs* entre: custos fixos e custos variáveis das máquinas; capacidade de processamento das máquinas; taxa média de processamento da planta; *leadtime* de produção; e complexidade para gerir a instalação. Para os propósitos deste trabalho, serão somente consideradas aquelas variáveis, objetivos e restrições que se relacionam mais diretamente à formação de filas nas estações. Conforme mencionado na seção 1, os problemas de projeto aqui tratados são agrupados em duas classes: desempenho desejado do sistema (SP1- *Strategical Problem 1*) e desempenho ótimo do sistema (SP2- *Strategical Problem 2*). A seguir utilizou-se o WIP como medida de desempenho para formular problemas das classes SP1.1 e SP2.1. Desde que WIP e *leadtime* são linearmente relacionados pela fórmula de Little, os algoritmos aqui utilizados também se aplicam para o *leadtime*. Sejam:

- μ_j e m_j : taxa de serviço de cada máquina e número de máquinas (idênticas) na estação j ;
- $F_j(\mu_j, m_j)$: custo de alocação da capacidade (μ_j, m_j) na estação j ;
- F_T : orçamento disponível para aquisição de capacidade na rede;
- $L_j(\mu_1, m_1; \mu_2, m_2; \dots; \mu_n, m_n)$: número médio de produtos na estação j como função da capacidade da rede. No texto que segue, em certos casos, por simplicidade o número médio de produtos na estação j é denotado simplesmente por L_j ;
- v_j : valor monetário médio de um produto na estação j (independente de sua classe); e
- L_T : limitante superior para o WIP na rede.

O WIP é um valor médio monetário do número de produtos na rede, aqui definido como $\sum_{j=1}^n v_j L_j(\mu_1, m_1; \dots; \mu_n, m_n)$. O valor v_j de cada estação j pode ser estimado usando experiência prática, ou pela média ponderada proporcional à taxa média de chegada dos produtos na estação j , conforme discutido na seção 4. Se $v_j=1$ para todo j , então o WIP corresponde simplesmente ao número de produtos na rede. O problema de WIP desejado *SP1.1* é o problema de determinar a capacidade (μ_j, m_j) para cada estação j , de maneira a minimizar o custo total de alocação de capacidade e satisfazer a restrição do WIP desejado para a rede. Este problema pode ser formulado como em Bitran e Dasu (1992) e Morabito (1998):

$$\begin{aligned} \min \sum_{j=1}^n F_j(\mu_j, m_j) \\ \sum_{j=1}^n v_j L_j(\mu_1, m_1; \dots; \mu_n, m_n) \leq L_T \\ (\mu_j, m_j) \in P_j, j = 1, \dots, n \end{aligned} \quad (1)$$

sendo P_j o domínio das variáveis. Similarmente, o problema de desempenho ótimo SP2.1 é o problema de determinar a capacidade (μ_j, m_j) para cada estação j , de maneira a minimizar o WIP da rede e satisfazer a limitação orçamentária para a rede. Este problema pode ser formulado de maneira análoga ao problema SP1.1. Para ambos os problemas SP1.1 e SP2.1, as variáveis de decisão são a capacidade de processamento de cada estação j , determinadas pelo número de servidores (m_j) e pela taxa de processamento de cada servidor (μ_j) .

3.1 Modelos SP1.1 e SP1.2 com alternativas discretas

Este trabalho trata o problema de adicionar capacidade nas estações com o objetivo de alcançar um nível de WIP desejado, L_T , a custo mínimo. Supõe-se que adição de capacidade em cada estação possa ser modelada por um conjunto de alternativas discretas. Tratar capacidade como variável discreta pode fornecer considerável flexibilidade na modelagem de sistemas de manufatura. Adições de máquinas e/ou mão-de-obra e introdução de novos turnos de operação são exemplos de adições de capacidades que podem ser modeladas como variáveis discretas. Conforme mencionado na seção 2, neste trabalho foram aplicadas aproximações baseadas no método de decomposição paramétrico, o que resultou no menor desvio em relação aos resultados de simulação (no caso, o sistema linear 2 em BITRAN; MORABITO, 1995) foi escolhido. Esta aproximação é denotada a seguir simplesmente por:

$$\Phi(m, \lambda, ca, \mu, cs) \quad (2)$$

em que os vetores m, λ, ca, μ e cs correspondem, respectivamente, aos parâmetros $\{m_j, \lambda_j, ca_j, \mu_j, cs_j\}$ para toda estação $j, j = 1, \dots, n$. O número esperado de produtos na estação j pode ser aproximado por:

$$L_j(m_j, \lambda_j, ca_j, \mu_j, cs_j) = \frac{\lambda_j(ca_j + cs_j)}{2} Lq_j(M/M/m_j) + \frac{\lambda_j}{\mu_j} \quad (3)$$

em que $Lq_j(M/M/m_j)$ denota o número esperado de produtos em uma fila $M/M/m_j$ (definido em, e.g., KLEINROCK, 1975; TIJMS, 1986), para outras aproximações de L_j veja, e.g., (WHITT, 1993). Na seqüência, supõe-se que a tecnologia, o roteiro de fabricação dos produtos e a localização das estações estão especificados, e o problema de projeto resume-se a adicionar capacidade (máquinas) nas estações ao mínimo custo de modo a mini-

mizar o WIP. Ao invés de escolher m ou μ como variáveis de decisão do modelo, Bitran e Tirupati (1989b) consideraram um número finito de alternativas para mudanças de capacidade em cada estação. Seja n_j o número total de alternativas na estação j . Para cada alternativa t , $t = 1, \dots, n_j$, o seguinte dado de entrada é fornecido:

- μ_{jt} e m_{jt} = taxa de serviço de cada máquina e número de máquinas na estação j na alternativa t ; e
- f_{jt} = custo da estação j na alternativa t .

Seja u_{jt} a variável de decisão:

$$u_{jt} = \begin{cases} 1, & \text{se alternativa } t \text{ é escolhida para estação } j \\ 0, & \text{caso contrário} \end{cases} \quad (4)$$

em que $\sum_{t=1}^{n_j} u_{jt} = 1$, $j = 1, \dots, n$.

Observe que, deste modo, a variável binária u_{jt} carrega as informações de taxa de serviço (μ_{jt}) e número de máquinas (m_{jt}) para um conjunto discreto de alternativas t ($t = 1, \dots, n_j$) para cada estação j . A escolha de capacidade em cada estação j é representada pelo vetor $(u_{j1}, u_{j2}, \dots, u_{jn_j})$, no qual todos os elementos são nulos, exceto um. Desta maneira, tem-se $m_j = \sum_{t=1}^{n_j} m_{jt} u_{jt}$ e $\mu_j = \sum_{t=1}^{n_j} \mu_{jt} u_{jt}$.

Uma formulação para o problema de adicionar capacidade de modo a minimizar o WIP ao custo mínimo (modelo SP1.1 com alternativas discretas) é:

$$\min F(u) = \sum_{j=1}^n \sum_{t=1}^{n_j} f_{jt} u_{jt}$$

sujeito a (2) e (3)

$$\sum_{j=1}^n v_j L_j(m_j, \lambda_j, ca_j, \mu_j, cs_j) \leq L_T$$

$$\mu_j - \sum_{t=1}^{n_j} \mu_{jt} u_{jt} = 0, \quad j = 1, \dots, n \quad (5)$$

$$m_j - \sum_{t=1}^{n_j} m_{jt} u_{jt} = 0, \quad j = 1, \dots, n$$

$$\sum_{t=1}^{n_j} u_{jt} = 1, \quad j = 1, \dots, n$$

$$u_{jt} \in \{0, 1\}, \quad t = 1, \dots, n_j; j = 1, \dots, n$$

A função objetivo representa o custo total de adição de capacidade, e as restrições (2) e (3) representam as aproximações, discutidas anteriormente, para estimar o número médio de produtos em cada estação. A restrição seguinte do modelo limita o WIP da rede a L_T ; convém observar que esta restrição também pode ser substituída por uma restrição que representa o *leadtime* médio da rede $E(T)$, ou o *leadtime* médio de cada classe $E(T_k)$, conforme será discutido adiante. As duas restrições seguintes determinam a capacidade em cada estação. As duas últimas restrições, juntas, restringem a escolha de exatamente uma alternativa t para cada estação j . O modelo acima é um programa não-linear inteiro, difícil de ser resolvido otimamente. As restrições (2) e (3) são não-lineares e contêm um número de funções implícitas

que tornam o problema particularmente difícil. Segue daí a necessidade de hipóteses simplificadoras.

Em Bitran e Tirupati (1989b), Bitran e Morabito (1999) e outros trabalhos, experimentos realizados em redes de filas com número de classes de produtos relativamente grande mostraram que, admitir que ca e cs sejam invariantes a mudanças de capacidade, resulta numa aproximação razoável (e.g., os erros na estimativa de ca foram menores que 5% em redes com 10 classes de produtos). Uma consequência imediata desta hipótese é que o sistema de equações (2) pode ser resolvido inicialmente e os vetores ca e cs podem ser tratados como parâmetros para o modelo SP1.1 com alternativas discretas. Além disso, o número médio de produtos em cada estação é função de sua capacidade (μ_j, m_j). Assim, enquanto se escolhe uma alternativa t na estação j (ou seja, $u_{jt} = 1$ e $u_{jl} = 0$, $l \neq t$), pode-se referir a $L_j(m_j, \lambda_j, ca_j, \mu_j, cs_j)$ em (2) simplesmente como L_{jt} , em que $L_{jt} = L_j(m_{jt}, \lambda_j, ca_j, \mu_{jt}, cs_j)$. Desde que esta hipótese se aplique a SP1.1 com alternativas discretas, então este modelo pode ser reescrito pelo programa linear inteiro:

$$\min F(u) = \sum_{j=1}^n \sum_{t=1}^{n_j} f_{jt} u_{jt}$$

$$\sum_{t=1}^{n_j} u_{jt} = 1, \quad j = 1, \dots, n \quad (6)$$

$$\sum_{j=1}^n \sum_{t=1}^{n_j} (v_j L_{jt}) u_{jt} \leq L_T$$

$$u_{jt} \in \{0, 1\}, \quad t = 1, \dots, n_j; j = 1, \dots, n$$

Nesta formulação admite-se, sem perda de generalidade, que se $L_{jt} > L_{jl}$, então $f_{jt} < f_{jl}$, $t \neq l$, para $t, l = 1, \dots, n$. Um algoritmo iterativo baseado na hipótese anterior para tratar o problema SP1.1 com alternativas discretas de capacidade é descrito por (MORABITO, 1998):

Passo 1: Escolha ca^0 , os valores iniciais de a_j^0 para $j = 1, \dots, n$. Os valores iniciais de a_j^0 para redes genéricas processando múltiplas classes de produtos podem ser obtidos pela aplicação do método de decomposição, ou seja, resolvendo (2), por exemplo. Faça $iter = 1$.

Passo 2: Resolva o programa linear inteiro SP1.1 com alternativas discretas utilizando os valores atualizados de ca (a_j^{iter-1} , $j = 1, \dots, n$). Denote a solução por u_{jt}^{iter} (a_j^{iter-1}); observe que se $u_{jt} = 1$, então a capacidade (μ_{jt} , m_{jt}) é alocada na estação j . Seja $(\mu_j^{iter}, m_j^{iter})$ a capacidade alocada na estação j .

Passo 3: Recalcule ca (a_j^{iter} , $j = 1, \dots, n$) utilizando as novas capacidades (μ_j^{iter} , m_j^{iter}) no sistema (2). Pare se a_j^{iter-1} e a_j^{iter} forem suficientemente próximos. Senão, faça $iter = iter + 1$ e volte ao passo 2.

De maneira similar, este algoritmo também pode ser adaptado para tratar o problema SP2.1 com alternativas discretas de capacidade (MORABITO, 1998). Este algoritmo também pode ser aplicado reescrevendo o modelo

SP1.1 com alternativas discretas de forma que o *leadtime* seja a medida de desempenho a ser alcançada, ao invés do WIP. Esta modificação, denotada por SP1.2 com alternativas discretas, é realizada neste trabalho para permitir que o gerente obtenha o mínimo investimento em capacidade, sujeito a restrições de limitantes superiores para os *leadtimes* das classes de produtos. O modelo SP1.2 com alternativas discretas é definido por:

$$\begin{aligned} \min F(u) &= \sum_{j=1}^n \sum_{t=1}^{n_j} f_{jt} u_{jt} \\ \sum_{t=1}^{n_j} u_{jt} &= 1, \quad j = 1, \dots, n \\ \sum_{l=1}^{n_k} \sum_{j=1}^n \sum_{t=1}^{n_j} \{ (EWq_{jt} + \\ E(s_{kl})u_{jt}) I[n_{kl} = j] \} &\leq ET_k, \quad k = 1, \dots, r \\ u_{jt} &\in \{0, 1\}, \quad t = 1, \dots, n_j; \quad j = 1, \dots, n \end{aligned} \quad (7)$$

Na restrição de limite de *leadtime* de cada classe k , o lado esquerdo corresponde à soma dos tempos de fila e serviço em todas as operações das estações j que pertencem ao roteiro de fabricação da classe k . O termo $I[n_{kl} = j]$ denota uma função indicadora que assume valor 1, se o produto k realiza a operação l na estação j ; e valor 0, em caso contrário. Este modelo é útil para situações em que o gerente deve, por exemplo, reduzir o *leadtime* de alguns produtos para mantê-lo competitivo, e necessita saber quanto e em quais estações deve investir. Note que esta restrição permite que o gerente estabeleça um limitante para o *leadtime* médio de cada classe k , ET_k . Se esta restrição for manipulada para todas as classes $k = 1, \dots, r$, de modo a resultar na expressão:

$$\sum_{k=1}^r \lambda'_k \left(\sum_{l=1}^{n_k} \sum_{j=1}^n \sum_{t=1}^{n_j} \{ (EWq_{jt} + E(s_{kl})u_{jt}) I[n_{kl} = j] \} \right) / \sum_{k=1}^r \lambda'_k \leq \sum_{k=1}^r \lambda'_k E(T_k) / \sum_{k=1}^r \lambda'_k \quad (8)$$

o modelo resultante permite que o gerente especifique um limitante superior para o *leadtime* médio da rede ET, e desta forma guarda uma correspondência com o modelo com a restrição agregada para o WIP, discutido anteriormente (dado que, conforme já mencionado, o WIP e o *leadtime* da rede são linearmente relacionados pela fórmula de Little).

4 Resultados da aplicação dos modelos no estudo de caso

Esta seção analisa a aplicação dos métodos de avaliação de desempenho (seção 2) e dos modelos de alocação ótima de capacidade discreta (seção 3) em um estudo de caso de uma empresa com sistema de manufatura *job-shop* e política de fabricação sob pedido (*make-to-order*).

A empresa localiza-se no interior de São Paulo, é especializada em usinagem de peças metálicas para aplicação industrial nos ramos eletrodoméstico (linha branca) e automotivo, e emprega cerca de 65 funcionários que operam 17,6 horas por dia, divididos em dois turnos de trabalho. O sistema de manufatura possui algumas máquinas desempenhando funções diferentes arranjadas para compor duas células de manufatura, e as demais máquinas arranjadas fisicamente por função (disposição funcional) para compor o sistema *job-shop*. As células de manufatura visam atender demandas mensais altas de poucas classes de produtos e as máquinas correspondentes são dedicadas a produzir somente aquelas poucas classes. Cada célula produz entre 50-65 mil peças/mês.

O sistema *job-shop* foi adotado pela empresa para satisfazer demandas mensais menores (pequenos lotes) para uma grande variedade de produtos. Para alguns produtos, lotes de 50-60 peças são solicitados somente uma vez ao ano, enquanto que para outros são pedidos lotes com 400-1600 peças/mês. Este sistema satisfaz principalmente um grande cliente, cujos pedidos são feitos por meio de um sistema de intercâmbio de dados eletrônicos (EDI) que são atualizados semanalmente. Este cliente disponibiliza os tamanhos dos lotes (que podem variar semanalmente) e o mês de entrega em um horizonte de tempo de aproximadamente um ano. Entretanto, as datas específicas (dia e mês) e o tamanho exato dos lotes são estabelecidos em um horizonte de tempo de dois meses, ou seja, os dados disponibilizados via EDI pelo cliente para produtos que devem ser produzidos e entregues em um horizonte de dois meses possuem datas específicas e tamanhos de lote fixos (geralmente não mudam semanalmente).

As principais dificuldades encontradas na aplicação dos modelos de redes de filas foram coletar e tratar estatisticamente os dados disponíveis na empresa, de modo a utilizá-los nos modelos. A primeira atividade foi representar a rede *job-shop* em estações de trabalho visitadas por classes de produtos (famílias) formadas por produtos que perfazem roteiros de fabricação similares. Em seguida, cada classe foi caracterizada por seu processo de chegadas externas e processo de serviço em cada estação do roteiro de fabricação. Essas atividades envolveram o diretor industrial e o engenheiro responsável pelo Planejamento e Controle da Produção (PCP) da empresa estudada.

O sistema *job-shop* foi modelado como uma rede com filas GI/G/1 de múltiplas classes e roteiro de fabricação determinístico. Estações com m_j máquinas operando em paralelo, por simplicidade, foram modeladas como contendo uma única "super máquina" com capacidade equivalente. A razão desta simplificação é que em certas estações as máquinas não são idênticas, ou são idênticas, mas não necessariamente trabalham em paralelo (por exemplo, duas máquinas e apenas um operador para

operá-las ao mesmo tempo). Deste modo, todas as estações da rede foram modeladas como possuindo uma única máquina (i.e., $m_j = 1, j = 1, \dots, n$). Convém mencionar que o *leadtime* de produção, assim como o WIP na rede, são superestimados ao se admitir que toda estação da rede é modelada como uma fila GI/G/1, ao invés de uma fila GI/G/m. Entretanto, esta hipótese foi adotada com vistas a simplificar o processo de coleta de dados e a modelagem do sistema, e parece razoável, considerando-se que as decisões de capacidade a serem suportadas pelos modelos são de médio a longo prazo (tático-estratégicas).

Juntamente com o diretor industrial e o engenheiro responsável pelo PCP, o sistema *job-shop* foi dividido em 15 estações de trabalho, a saber: estação 1 - máquina serra fita; estação 2 - torno automático; estação 3 - torno CNC (controlado numericamente); estação 4 - torno mecânico; estação 5 - torno revólver; estação 6 - máquinas furadeiras; estação 7 - estação de fresas; estação 8 - forno de indução; estação 9 - centro de usinagem; estação 10 - retíficas; estação 11 - rosqueadeiras de bancada; estação 12 - tamboreador vibratório; estação 13 - balancim; estação 14 - estação para inspeção intermediária ou final; e estação 15 - estação responsável pela embalagem. Operações realizadas por terceiros (e.g., para tratamento superficial das peças) não foram consideradas como estações de trabalho da rede, embora se tenha considerado 3 dias de operação efetiva para efeitos de estimativa de *leadtime* para aquelas classes de produtos que necessitam deste serviço terceirizado. De acordo com o gerente industrial, com pequenas variações, esse é o tempo médio gasto para que um lote deixe a empresa, realize o serviço e retorne à empresa. Além disso, há várias pequenas empresas disponíveis para realizar esse serviço, e quando uma empresa está congestionada, o lote é enviado para outra.

No total, este *job-shop* produz 20 classes de produtos diferentes. Entretanto, no período em que a amostra foi coletada (14 meses), algumas classes de produtos não tiveram pedidos de produtos, enquanto outras tiveram somente um ou dois pedidos. Assim, 6 classes de produtos (que juntas consumiram menos de 1,5% em hora-máquina do total consumido no período) foram desconsideradas na análise. Desta maneira, considerou-se que a rede processa 14 classes de produtos: classe 1 - núcleo da bobina; classe 2 - haste de inox I e haste de inox II (são produtos com roteiros de fabricação similares, mas com especificações técnicas diferentes); classe 3 - garfo; classe 4 - corpo da válvula; classe 5 - pino; classe 6 - pistão I; classe 7 - eixo; classe 8 - batente; classe 9 - adaptador I; classe 10 - adaptador II; classe 11 - pistão II; classe 12 - bucha de inox; classe 13 - luva; e classe 14 - parafuso de regulagem. Cada classe percorre um roteiro de fabricação diferente.

A Figura 1 ilustra o roteiro das classes no sistema *job-shop* estudado. Na Figura 1a, por exemplo, os arcos

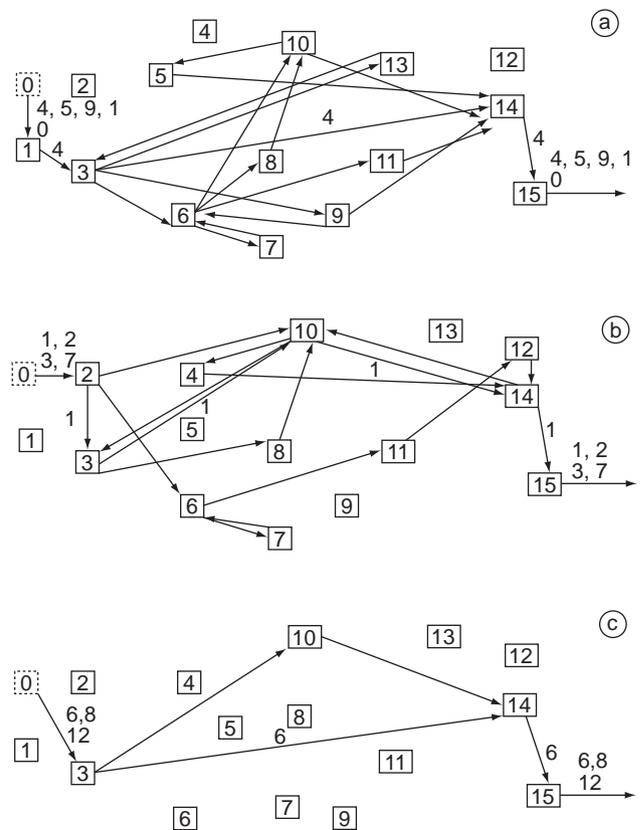


Figura 1. Sistema *job-shop* em estudo (0 é a estação fictícia representando a chegada externa). a) roteiros das classes 4, 5, 9, 10, 11, 13, 14; b) roteiros das classes 1, 2, 3 e 7; e c) roteiros das classes 6, 8, 12.

ilustram o roteiro de fabricação das classes 4, 5, 9, 10, 11, 13 e 14. Para que as figuras ficassem mais legíveis, identificou-se somente os arcos referentes aos roteiros das classes 4, 1 e 6 na Figura 1a,b,c, respectivamente. Em ambas as figuras, 0 é a estação fictícia representando as chegadas externas à rede. O roteiro de cada classe também é detalhado na Tabela 1 adiante. Para caracterizar os processos de chegada e de serviço de cada classe para as aproximações de rede de filas e a simulação, foi utilizada uma amostra compreendendo 14 meses de operação da fábrica. Os parâmetros de entrada necessários foram coletados através de relatórios emitidos de uma base de dados eletrônica disponível na empresa. Neste estudo, a inviabilidade de uma coleta direta no chão de fábrica deu-se pelo fato de boa parte das classes de produtos serem fabricadas quatro ou cinco vezes em um ano (o que as deixaria de fora em um processo de coleta de dias, semanas ou poucos meses).

4.1 Caracterização dos processos de chegadas externas e serviços

Para cada classe de produtos k ($k = 1, \dots, 14$), a taxa média de chegada externa λ'_k e o coeficiente quadrático

Tabela 1. Tamanho da amostra, tamanho médio e desvio-padrão do lote (entre colchetes), $E(a'_k)$, λ'_k , ca'_k , n_k (número de operações no roteiro da classe k), n_{kl} (estações visitadas para operação l no roteiro da classe k).

Classe	Tamanho da amostra	Tamanho do lote (produtos)	$E(a'_k)$ (h)	λ'_k (lote/h)	ca'_k	n_k	n_{kl}
1	13	493,6 [182,1]	280,03	0,003571	0,3818	5	2, 3, 10, 14, 15
2	36	896,0 [603,3]	131,29	0,007617	0,3377	8	2, 10, 3, 8, 10, 4, 14, 15
3	18	1583,7 [702,9]	257,00	0,003891	0,2569	9	2, 6, 7, 6, 11, 12, 14, 14, 15
4	9	621,0 [213,9]	473,26	0,002113	0,5092	4	1, 3, 14, 15
5	8	468,2 [236,1]	436,30	0,002292	0,2853	8	1, 3, 9, 6, 10, 5, 14, 15
6	5	1083,2 [104,1]	517,60	0,001932	0,4048	4	3, 14, 14, 15
7	8	900,1 [332,6]	538,50	0,001857	0,2939	7	2, 10, 14, 14, 10, 14, 15
8	5	4801,5 [3395,9]	671,14	0,001490	0,2934	5	3, 14, 14, 14, 15
9	13	640,6 [229,1]	357,14	0,002800	0,4131	8	1, 3, 9, 6, 8, 10, 14, 15
10	4	876,6 [238,6]	926,78	0,001079	0,2817	5	1, 3, 14, 14, 15
11	15	495,7 [327,0]	289,44	0,003455	0,4351	5	1, 3, 9, 14, 15
12	13	1069,2 [590,2]	349,41	0,002862	0,3607	4	3, 10, 14, 15
13	5	877,2 [499,4]	697,84	0,001433	0,3037	9	1, 3, 6, 7, 6, 11, 14, 14, 15
14	5	1108,5 [380,0]	738,01	0,001355	0,3528	7	1, 3, 13, 3, 9, 14, 15

de variação do intervalo de tempo entre chegadas externas a'_k , definidos como $\lambda'_k = 1/E(a'_k)$ e $ca'_k = V(a'_k)/(E(a'_k))^2$, em que $V(a'_k)$ e $E(a'_k)$ são a variância e a média do intervalo de tempo entre chegadas externas da classe k , respectivamente, foram estimados a partir da amostra. Por falta de informações mais precisas, o instante de chegada de um lote de fabricação de uma determinada classe de produto na primeira estação da sua rota foi considerado como o instante em que, segundo relatórios, foi aberta uma ordem de fabricação (OF) para aquela classe de produto. Obviamente, à abertura de uma OF antecedem regras de programação que são aplicadas pelo PCP da empresa, tais como a *Earliest Due Date* (EDD). Desta forma, o processo de chegada é o resultado desta programação, e não das chegadas dos pedidos na empresa ou no PCP. A consideração da chegada do pedido como o instante em que é aberta uma OF parece ser bastante razoável para caracterizar o processo de chegada e entrada no sistema *job-shop*, ao considerar que todas as classes possuem um atraso (*time lag*) aproximadamente constante entre o instante de abertura da OF e o início da fabricação.

Considerou-se que as chegadas externas de classes de produtos ocorrem nos dias úteis em que a manufatura opera. Portanto, uma classe de produtos com ordem de fabricação aberta em uma segunda-feira, que possui outra ordem aberta na segunda-feira da semana seguinte, possui intervalo de tempo entre chegadas externas (em horas) igual a $17,6 \times 5 = 88$ horas. Isto ocorre se o *job-shop* opera de segunda à sexta-feira normalmente (e.g., sem feriados), com 17,6 horas de operação efetiva por dia, o que corresponde à quantidade de horas nominalmente trabalhadas pela empresa. Entretanto, ao considerar cada dia como sendo constituído de 17,6 horas, supõe-se que

todas as estações estejam disponíveis e operem efetivamente 17,6 horas por dia. Os modelos admitem que, quando uma estação está desocupada e ocorre a chegada de um lote de produtos nela, o serviço inicia-se instantaneamente. Esta situação nem sempre ocorre na prática, pois pode acontecer da estação estar desocupada e o lote ter de aguardar por falta de operador, material, quebra de máquina ou falta de energia. Em outras palavras, ao considerar que a estação opera efetivamente 17,6 horas por dia, superestima-se a capacidade da estação, pois são desconsiderados os fatores que ocasionam perda de capacidade.

Uma maneira de considerar a perda de capacidade é a redução das horas efetivamente trabalhadas em um dia de produção. De acordo com o gerente industrial da empresa, a capacidade efetiva das estações é em torno de 95%, ou seja, pode-se considerar que o dia efetivamente trabalhado possui $17,6 \times (1 - 0,05) = 16,7$ horas. Uma maneira mais precisa de se considerar tais perdas de capacidade é por meio da criação de classes de produtos fictícias (BUZACOTT; SHANTHIKUMAR, 1993; MORABITO, 1998) com roteiros probabilísticos ao longo das estações, com base em informações dos processos de falhas e manutenção de equipamentos. A empresa estudada não dispunha de dados dos instantes e durações de quebras de máquinas, faltas de operador, energia, etc., em cada estação, durante o período de coleta de dados, para que esta alternativa pudesse ser considerada nas aproximações analíticas e na simulação. Portanto, neste trabalho, a perda de capacidade foi considerada simplesmente pela redução das horas efetivamente trabalhadas para 16,7 horas/dia.

Os dados de entrada caracterizando os processos de chegadas externas e o roteiro de fabricação para cada classe são mostrados na Tabela 1. Para cada classe, a chegada de lotes de produtos de tamanhos variáveis (vide média e desvio-padrão do tamanho dos lotes, entre colchetes, na terceira coluna da Tabela 1) foi considerada como se fosse a chegada de um único item ou produto (*batch arrival*) nos modelos de redes de filas e simulação. Nota-se pelas classes 2 e 8 que o tamanho dos lotes é bastante variável. O roteiro de fabricação de cada classe é descrito pelas estações visitadas para cada operação desta classe (Tabela 1). Note que os coeficientes quadráticos de variação a'_k das classes são bem diferentes de 1, ou seja, os processos de chegadas externas são bem diferentes do processo de Poisson, caracterizando uma rede de filas genérica. Observe, por exemplo, que o roteiro da classe 6 é definido por 4 operações ($n_6 = 4$), em que $n_{61} = 3$, $n_{62} = 14$, $n_{63} = 14$ e $n_{64} = 15$ é a seqüência de estações visitadas, com duas operações consecutivas diferentes na estação 14.

Para caracterizar os processos de serviço, foram coletadas amostras no mesmo período de 14 meses, via relatórios da base eletrônica, dos tempos de serviço s_{kl} para cada classe k em cada operação l do seu roteiro. A partir dessas amostras, a taxa média de serviço e o coeficiente quadrático de variação dos tempos de serviço da operação l do roteiro da classe k , respectivamente $\mu_{kl} = 1/E(s_{kl})$ e $cs_{kl} = V(s_{kl})/(E(s_{kl}))^2$, foram determinados para os lotes processados, para servirem como dados de entrada para os modelos de redes de filas. Como se admitiu que cada OF aberta com lote de tamanho variável corresponde à chegada de um único produto, o tempo de serviço em cada estação corresponde ao tempo de processamento do lote (*batch service*). A Tabela 2 mostra os tempos médios

de serviço para cada classe e operação da rede. Os coeficientes quadráticos de variação dos tempos de serviço das classes, cs_{kl} , variaram de 0,0092 a 0,5002, ou seja, são bem diferentes de 1; portanto, não são exponencialmente distribuídos e reforçam a necessidade de se utilizar um modelo de rede de filas genéricas.

Algumas classes de produtos tais como as 3, 6, 7, 8, 10 e 13 realizam operações como tratamento superficial, que são terceirizadas. A Figura 2 ilustra em detalhes o roteiro de fabricação da classe 6, juntamente com o serviço terceirizado (ST). Conforme se pode perceber na Figura, a classe de produto 6 realiza sua primeira operação na estação 3, sendo em seguida encaminhada para realização de serviço terceirizado (ST), onde é feito o tratamento superficial. Após 3 dias (ou seja, $3 \times 16,7 = 50,1$ horas), o produto é devolvido para a fábrica e encaminhado para o fim da fila da estação 14, para ser inspecionado. Após inspeção na estação 14, o produto é enviado para outro serviço terceirizado (ST), que pode inclusive empregar outro ou o mesmo prestador de serviços, retornando à fábrica após 3 dias (ou 50,1 horas) no fim da fila da estação 14, para inspeção final. A partir daí, segue para embalagem na estação 15, para então aguardar a expedição. Conforme mencionado antes, todo prestador de serviços terceirizados possui um prazo de 3 dias para

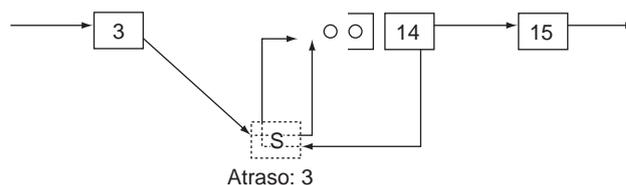


Figura 2. Roteiro de fabricação da classe de produtos 6: 3, ST, 14, ST, 14, 15.

Tabela 2. Tempos médios de serviço de um lote na operação l do roteiro da classe k , $E(s_{kl})$.

Classe	Operação (horas/lote)								
	1	2	3	4	5	6	7	8	9
1	5,761	27,831	4,773	5,507	2,346	-	-	-	-
2	19,197	8,664	27,989	12,488	31,368	14,848	6,971	12,639	-
3	26,935	33,828	35,662	47,808	9,139	4,744	4,519	3,247	2,557
4	0,718	17,596	1,770	2,197	-	-	-	-	-
5	1,080	22,038	18,353	14,022	7,604	6,580	8,919	2,700	-
6	42,332	3,968	3,028	2,558	-	-	-	-	-
7	14,606	18,966	3,293	8,052	32,346	4,937	6,638	-	-
8	38,146	3,891	2,033	6,490	3,103	-	-	-	-
9	23,900	30,973	28,480	25,128	6,399	9,095	6,281	3,620	-
10	54,625	41,868	4,530	7,158	4,320	-	-	-	-
11	1,261	11,395	21,002	5,216	1,820	-	-	-	-
12	24,265	8,565	3,731	2,676	-	-	-	-	-
13	3,207	19,688	10,119	17,693	10,978	4,761	2,623	3,399	1,600
14	0,566	24,156	13,566	19,469	48,550	6,274	6,881	-	-

retirar o produto, realizar a operação e devolvê-lo à fábrica, e ainda segundo o diretor industrial, raramente este prazo não é cumprido. Por este motivo, foi considerado – somente para efeito de estimativa de *leadtime* para aquelas classes que realizam operações terceirizadas – um atraso (*time lag*) determinístico de 50,1 horas para cada visita a terceiros que a classe de produtos realiza.

Convém observar que, para os modelos de rede de filas e simulação, o roteiro de fabricação considerado para a classe 6 é a seqüência de estações: 3, 14, 14, 15 (Tabela 1). Ou seja, não se considera no roteiro as visitas no serviço terceirizado, e a seqüência é tratada como retroalimentação imediata na estação 14 (*immediate feedback*, conforme WHITT, 1983a), visto que, após realização do serviço na estação 3, o produto segue para a estação 14, retorna novamente à fila da estação 14 e, posteriormente, segue para a estação 15. No entanto, o que acontece na prática é retroalimentação após atraso de 3 dias (ou 50,1 horas), que só é considerado na estimativa do *leadtime* das classes de produtos. Entretanto, tratar no roteiro como retroalimentação imediata, a retroalimentação com atraso de 3 dias, parece ser uma simplificação razoável, dado que a análise do sistema é feita em equilíbrio (*steady-state*), e espera-se que o estado da fila em um certo instante seja, em média, semelhante ao seu estado após 3 dias. Convém ressaltar que não se deve confundir a retroalimentação de um roteiro determinístico com a de um roteiro probabilístico, que deteriora a qualidade das aproximações dos métodos de decomposição e deve ser tratada de maneira diferente (WHITT, 1983a; MORABITO, 1998).

Com relação ao modelo de simulação para validar as aproximações analíticas, foram aplicados testes de aderência Kolmogorov-Smirnov (K-S) e Anderson-Darling (A-D), usando o *software* BestFit® com níveis de significância de até 15% para caracterizar os processos de chegadas externas e serviços. Em casos em que a amostra foi rejeitada nos testes de K-S e A-D para as 26 distribuições teóricas do BestFit®, utilizou-se a distribuição empírica sugerida em Law e Kelton (1991). As distribuições foram utilizadas pelo simulador GPSS/H® para gerar as chegadas externas e os serviços de cada classe em cada estação no modelo de simulação (para mais detalhes destas distribuições e do modelo de simulação, veja SILVA, 2005).

4.2 Custos e alternativas discretas para alocação de capacidade

O primeiro passo para implementação do modelo SP1.1 com alternativas discretas no algoritmo da seção 3 foi a estimativa do custo de um lote de produto qualquer na estação j , v_j . Os valores monetários aqui apresentados foram modificados por um certo fator para proteger interesses da empresa. Para cada classe de produto k e operação l , foram levantados os custos com matéria-prima e operação. Seja C_{kl} representando o custo acumulado operacional e com matéria-prima de um lote médio da classe k com a realização da l -ésima operação. Por exemplo, sabe-se que para cada produto da classe 14, gastam-se \$0,12 com matéria-prima. Como cada lote da classe 14 possui, em média, 1108,5 produtos (Tabela 1), então o custo com matéria-prima para produzir um lote de tamanho médio foi considerado $1108,5 \times 0,12 = \$133,73/\text{lote}$. Já o custo relacionado à operação 1 ($l=1$) para a classe 14 foi estimado da seguinte maneira: o custo operacional por hora (\$/hora) na estação em que ocorre a primeira operação da classe 14 ($n_{14,1} = 1$, vide Tabela 1) foi estimado e multiplicado pelo inverso da taxa de produção de um lote (hora/lote) da classe 14 na estação. Deste modo, o valor obtido para o custo operacional foi de \$13,13. Logo, o custo acumulado da classe 14 com a operação 1, $C_{14,1}$, foi de \$146,86/lote (\$13,13 + \$133,73). Da mesma forma, o custo operacional para a classe 14 na estação que realiza a operação 2 ($n_{14,2} = 3$, vide Tabela 1) foi estimado em \$972,04/lote, resultando em um custo acumulado da classe 14 com a segunda operação do roteiro de fabricação, $C_{14,2}$, de \$1118,90/lote (\$146,86 + \$972,04).

Este procedimento foi aplicado para cada classe k e operação l do roteiro de fabricação. Entretanto, para estimativa do capital de trabalho associado ao inventário em processo, deve-se estimar o custo médio de um lote de produto de uma classe qualquer na estação j , v_j . Para cada estação j ($j = 1, \dots, n$), foi computado o custo médio acumulado de todas as classes de produtos que utilizam a estação j , ponderado pela proporção de chegada dos produtos à estação, ou seja:

$$v_j = \frac{\sum_{k=1}^r \sum_{l=1}^{n_k} C_{kl} \lambda_k' 1[n_{kl} = j]}{\sum_{k=1}^r \sum_{l=1}^{n_k} \lambda_k'}, \quad j = 1, \dots, n \quad (9)$$

em que $1[n_{kl} = j]$ é uma função indicadora com valor 1 se a operação l da classe k é realizada na estação j ; e 0, em caso contrário. A Tabela 3 mostra os custos acumulados,

Tabela 3. Custo acumulado v_j para um lote de uma classe de produtos qualquer na estação j (em \$/lote).

Estação	1	2	3	4	5	6	7	8
Custo	799,0	2955,4	3369,9	10698,5	2825,0	5008,5	4361,2	6847,1
Estação	9	10	11	12	13	14	15	
Custo	2991,9	5888,5	7678,1	9664,2	1496,3	5526,9	5850,1	

v_j , para um lote de uma classe de produto qualquer na estação j da rede do estudo de caso.

Juntamente com o gerente industrial e o engenheiro encarregado do PCP, também foram consideradas as alternativas para adição (custos positivos) ou remoção (custos negativos) de capacidade nas estações, conforme apresentado na Tabela 4. A alternativa 1 corresponde à configuração atual e, por isso, o custo com adição ou remoção de capacidade é nulo para todas estações. A alternativa 2 considera a adição de uma máquina nas estações da rede, exceto na estações 14 e 15 porque correspondem às atividades de inspeção e embalagem, cujas tarefas são executadas por funcionários e não são de maior interesse da empresa estudada. A alternativa 3 considera a remoção de uma máquina naquelas estações em que a taxa de utilização da estação se mantiver menor que 1 (i.e., $\rho_j = \lambda_j / (m_j \mu_j) < 1$), que é a condição de equilíbrio da rede. Outras alternativas poderiam ter sido consideradas. Os dados de custos de máquinas foram pesquisados junto aos fornecedores dos equipamentos. Para o caso de remoção, foram considerados os custos de equipamentos obtidos dos fornecedores, levando-se em conta a depreciação do equipamento usado, uma vez que ao considerar a possibilidade da empresa remover a máquina, ela possa vendê-la a preço de mercado.

4.3 Avaliação do desempenho da configuração atual da rede

Após a representação do sistema de manufatura como uma rede aberta de filas composta de 15 estações de trabalho, 14 classes de produtos com roteiros de fabricação determinísticos e com os processos de chegadas externas e de serviços descritos por parâmetros de taxa e

coeficiente quadrático de variação, o próximo passo foi a implementação computacional dos métodos de avaliação de desempenho. As seis aproximações discutidas em Bitran e Morabito (1995, 1999) foram aplicadas para esta rede e os resultados obtidos foram comparados com os do modelo de simulação (usando o *software* GPSS/H®). Estes resultados estão detalhados em Silva (2005). A aproximação que resultou no menor desvio do número médio de lotes (-0,9%) e do *leadtime* médio (3,1%) em relação à simulação foi o sistema linear 2 em Bitran e Morabito (1995), composto neste caso de 156 variáveis e 156 equações. Ao apresentar os resultados obtidos para o engenheiro responsável pelo PCP da empresa, ele mencionou que os *leadtimes* obtidos estão próximos dos que ocorrem na prática da empresa, mas não pôde afirmar o mesmo a respeito do WIP estimado, já que o controle desta variável pela empresa é visual, sem registros que permitam uma comparação mais objetiva.

Ao supor a rede totalmente determinística, observou-se que o modelo subestima significativamente o número de lotes (desvio de -31,1% em relação à simulação). Por outro lado, ao considerá-la como uma rede de Jackson, o número de lotes é substancialmente superestimado (desvio de 69,0%). Também foram realizadas avaliações da sensibilidade das aproximações a alterações em parâmetros de entrada, tais como nas taxas de chegadas externas (lotes/hora) e no *mix* de produtos. As aproximações do sistema linear 2 apresentaram desvios pequenos no número de lotes e no *leadtime* na rede em relação à simulação. Para mais detalhes destes resultados o leitor pode consultar Silva (2005). Assim, o sistema linear 2 foi utilizado nos modelos de otimização para alocação

Tabela 4. Alternativas de adição/remoção de capacidade nas estações j : taxa de produção (μ_j) e custo (f_j) na estação j e alternativa t .

Estação	Alternativa 1		Alternativa 2		Alternativa 3	
	μ_{j1} (lote/h)	f_{j1} (\$1000)	μ_{j2} (lote/h)	f_{j2} (\$1000)	μ_{j3} (lote/h)	f_{j3} (\$1000)
1	0,104	0,0	0,208	9,0	0,104	0,0
2	0,057	0,0	0,066	22,0	0,047	-12,0
3	0,038	0,0	0,042	190,0	0,038	0,0
4	0,067	0,0	0,084	12,0	0,052	-7,0
5	0,152	0,0	0,228	6,0	0,076	-3,0
6	0,035	0,0	0,042	0,5	0,028	-0,3
7	0,032	0,0	0,037	40,0	0,027	-20,0
8	0,092	0,0	0,138	120,0	0,046	-60,0
9	0,038	0,0	0,057	130,0	0,019	-60,0
10	0,063	0,0	0,076	225,0	0,050	-110,0
11	0,126	0,0	0,252	2,0	0,126	0,0
12	0,211	0,0	0,422	2,0	0,211	0,0
13	0,074	0,0	0,148	0,5	0,074	0,0
14	0,199	0,0	0,199	0,0	0,199	0,0
15	0,201	0,0	0,201	0,0	0,201	0,0

ótima de capacidade a seguir, para avaliar o desempenho da rede em diferentes configurações.

4.4 Alocação de capacidade ótima na rede

A Tabela 5 mostra o $WIP (v_j E(L_j))$ em cada uma das estações j para a configuração atual da rede usando a aproximação 2. Observe que o WIP da rede é de \$26943,6 e o *leadtime* médio da rede corresponde a 187,6 horas. O algoritmo com modelo SP1.1 com alternativas discretas definido na seção 3 foi implementado no *software* GAMS/CPLEX® (SILVA, 2005) e resolvido para limitantes superiores de WIP iguais a \$19500, \$20000, \$21000, \$22000, \$23000, \$24000, \$25000, \$26000, \$27000 e \$28000. Os resultados, obtidos em poucos segundos num microcomputador Pentium III com 800 MHz, foram utilizados para gerar a curva de *trade-off* entre investimento em capacidade e WIP da rede (Figura 3), e investimento em capacidade e *leadtime* médio da rede (Figura 4). Percebe-se nas Figuras 3 e 4 que, na medida em que se aumenta o investimento em capacidade, o WIP e o *leadtime* médio da rede são reduzidos, como era esperado. Note nas figuras, no entanto, que a curva de *trade-off* pode ser não convexa. Estas figuras são bastante úteis na tomada de decisões tático-estratégicas, pois para cada custo de WIP ou valor de *leadtime*, a curva indica o ponto com mínimo custo de recursos. Nota-se na Figura 3, que a configuração atual da rede (26943,6; 0) é um ponto fora

da curva de Pareto e, portanto, pode ser melhorada. Por exemplo, o gerente pode migrar o estado do sistema de manufatura para o ponto (26965,7; -68,5). Ao migrar da configuração atual para esta configuração, o WIP e o *leadtime* médio da rede são mantidos aproximadamente constantes (WIP de 26943,6 para 26965,7 e *leadtime* de 187,6 para 187,7 horas, vide Figuras 3 e 4), mas o investimento necessário em capacidade é menor.

A Tabela 6 mostra as alternativas obtidas para a melhor configuração com WIP máximo igual a \$27000. Observe-se na tabela que foi adicionada capacidade nas estações 4, 6, 11 e 14, mas as estações 5, 7, 8, 12 e 13 necessitaram de menos capacidade, compensando a aquisição e resultando em uma economia de \$68500 em máquinas. Por outro lado, caso o gerente tenha interesse em reduzir o WIP da configuração atual de \$26943,6 para cerca de \$26000, o sistema deve migrar da configuração atual (26943,6; 0) para o ponto (25925,3; -6,5) (Figura 3). Neste caso, devem-se adicionar máquinas nas estações 4, 6, 11, 12 e 14 e disponibilizar capacidade das estações 3, 7 e 13 para venda, resultando em uma economia com capital de investimento de -\$6500 (SILVA, 2005). A redução do *leadtime* médio da rede é de 3,8 horas (de 187,6 para 183,8 horas, conforme a Figura 4).

Observe que se o gerente quiser reduzir o WIP da configuração atual para menos de \$22000, ele deve

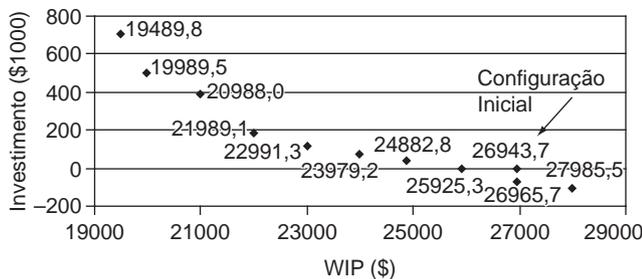


Figura 3. Curva de *trade-off* entre investimento em capacidade e WIP da rede.

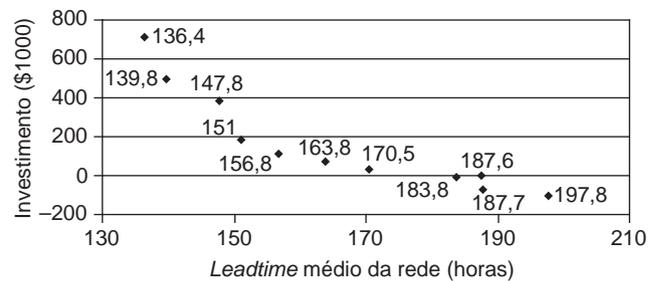


Figura 4. *Trade-off* entre investimento em capacidade e *leadtime* médio da rede.

Tabela 5. $WIP (v_j E(L_j))$ em \$/lote em cada estação e total para a configuração atual da rede.

Estação	1	2	3	4	5	6	7	8
WIP (\$)	130,5	951,7	10799,4	1282,8	42,6	3035,4	737,9	807,1
Estação	9	10	11	12	13	14	15	Total
WIP (\$)	849,9	4799,6	326,3	178,5	27,5	1671,4	1303,0	26943,6

Tabela 6. Alternativa t escolhida pela solução do modelo SP1.1 com alternativas discretas para limitante superior do WIP igual a \$27000 e investimento de -\$68500.

Estação	1	2	3	4	5	6	7	8
Alternativa	1	1	1	2	3	2	3	3
Estação	9	10	11	12	13	14	15	
Alternativa	1	1	2	3	3	2	1	

investir \$180500 em capacidade (Figura 3). Ou seja, para reduzir o WIP em \$4954,6 (de \$26943,7 para \$21989,1), há a necessidade de investimento da ordem de grandeza bem maior, que talvez não seja justificado. Entretanto, a redução do *leadtime* médio da rede é de 36,6 horas (de 187,6 para 151,0 horas), que corresponde a 2,19 dias de operação da empresa (lembre-se que a empresa opera 16,7 horas por dia). Esta redução pode ser significativa do ponto de vista de atendimento dos pedidos do cliente e, portanto, a análise do ponto de vista do *leadtime* pode ser de maior interesse do que do ponto de vista do WIP. Por este motivo, doravante o *leadtime* médio da rede ou dos produtos também é considerado nas observações e análise por meio do algoritmo com o modelo SP1.2 com alternativas discretas definido na seção 3. A Figura 5 mostra o *leadtime* das classes de produtos para a configuração atual da rede e para outras duas configurações (WIP = \$21679,0, F = \$180500 e WIP = \$24596,0, F=\$35500, com F representando o custo com adição/remoção de recursos), obtidas ao estabelecer diferentes limitantes superiores para o *leadtime* de cada classe na rede, $E(T_k)$, no modelo SP1.2 com alternativas discretas.

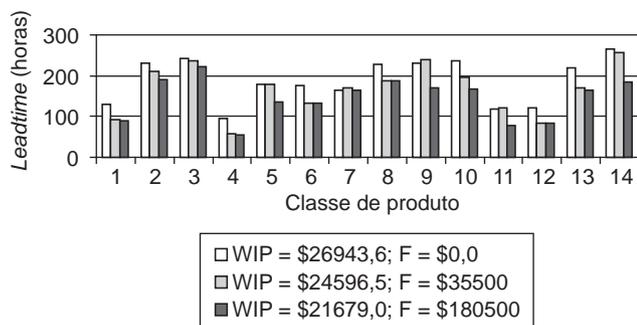


Figura 5. *Leadtime* das classes de produtos para três configurações de rede diferentes, obtidos ao resolver o modelo SP1.2 com alternativas discretas

Os limitantes superiores impostos para cada uma das classes nas configurações com $F = \$180500$ e $F = \$35500$ são apresentados na Tabela 7.

Note na Figura 5 que, em geral, quanto menor o investimento em recursos, maior é o *leadtime* das classes de produtos, como era de se esperar. Entretanto, podem resultar algumas classes com *leadtime* na configuração atual menor que o *leadtime* quando se investe $F = \$35500$, como é o caso das classes 7, 9 e 11. Isto ocorre porque as estações visitadas por estas classes apresentam alto nível de ocupação e, portanto, o tempo de espera nas respectivas filas das estações é grande, apesar do investimento na rede. Conforme a Tabela 1, a classe de produtos 7 realiza as sete operações do roteiro de fabricação nas estações 2, 10, 14, 14, 10, 14 e 15, respectivamente. De acordo com a Tabela 8, a estação 2 apresenta maior tempo de espera, apesar do investimento na rede de \$35500. Já as classes 9 e 11 também passam pelas estações 8 e 9 que, apesar do investimento na rede, apresentam maior tempo de espera em fila (Tabela 8). Isto acontece porque a solução de mínimo investimento em capacidade na rede utiliza a alternativa 3 para estas estações e, conforme a Tabela 4, esta alternativa corresponde à remoção de capacidade.

4.5 Sensibilidade do modelo a variações nos parâmetros de variabilidade

Uma maneira de reduzir o investimento em capacidade é a redução da variabilidade do intervalo de tempo entre chegadas externas das classes de produtos, ca'_k . Neste trabalho, conforme mencionado, a chegada externa foi considerada como o instante em que o PCP da empresa abre uma ordem de fabricação e o produto começa a ser produzido na primeira estação de seu roteiro de fabricação. Por isso, reduzir a variabilidade do processo de chegadas é uma estratégia possível de ser implementada, uma vez que depende de ajustes na política de programação da produção e, obviamente, de negociação

Tabela 7. Limitantes superiores (em horas) para o *leadtime* de cada classe de produtos, $E(T_k)$, nas configurações com necessidade de $F = \$180500$ e $F = \$35500$ de investimento em capacidade.

Classe	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$F = \$35500$	96	214	236	60	179	137	170	190	243	199	123	87	175	261
$F = \$180500$	96	195	224	59	141	138	171	191	173	171	80	87	169	192

Tabela 8. Tempos médios de espera nas filas das estações (em horas) para a configuração atual ($F = \$0$) e com investimento de $F = \$35500$.

Estação	1	2	3	4	5	6	7	8
$F = \$0$	1,63	1,38	70,63	0,89	<0,01	9,90	0,96	0,47
$F = \$35500$	1,63	2,65	33,32	0,89	0,01	5,67	1,74	2,93
Estação	9	10	11	12	13	14	15	
$F = \$0$	2,72	10,85	0,02	<0,01	<0,01	0,72	0,93	
$F = \$35500$	23,15	10,86	0,02	<0,01	<0,01	0,72	0,93	

de prazos e entregas com os clientes. Dois cenários são propostos e comparados com o cenário original da Tabela 1 (cenário 0):

- cenário 1: a taxa e o *mix* de produção são mantidos constantes, mas o coeficiente quadrático de variação do intervalo de tempo entre chegadas externas para a classe de produto 2 (ca'_k) é reduzido de 0,3377 (Tabela 1) para 0 (determinístico); e
- cenário 2: a taxa e o *mix* de produção são mantidos constantes, mas o coeficiente quadrático de variação do intervalo de tempo entre chegadas externas para todas as classes é reduzido à metade.

O investimento mínimo necessário para que o WIP não ultrapasse, por exemplo, o limitante superior de \$22000, foi obtido pelo modelo SP1.1 com alternativas discretas em \$180500 para o cenário 0, em \$159500 para o cenário 1, e em \$97500 para o cenário 2. Ou seja, houve economia de investimento em capacidade de \$21000 para o cenário 1 e \$83000 para o cenário 2, comparados ao cenário 0, que não considerou redução na variabilidade no processo de chegada. A Tabela 9 reporta os *leadtimes* médios obtidos para as classes de produtos e para a rede ($E(T)$) nos cenários 0, 1 e 2. Nota-se na Tabela que, apesar do cenário 2 exigir menor investimento em capacidade, o *leadtime* médio da rede também é, em geral, menor que para o cenário 0 devido à redução à metade dos parâmetros de variabilidade. A Figura 6 ilustra o *trade-off* entre investimento em recursos e WIP para os cenários 0 e 2 a limitantes superiores de WIP iguais a \$20000, \$20500,

\$21000, \$21500, \$22000, \$22500, \$23000, \$23500, \$24000, \$24500, \$25000, \$25500 e \$26000. Percebe-se na figura que para níveis desejados de WIP bem inferiores a \$22000, a diferença de investimento necessário entre os cenários 0 e 2 é bem significativa.

Os resultados aqui alcançados enfatizam a sensibilidade dos modelos de investimento em capacidade estudados a parâmetros de variabilidade, particularmente o coeficiente quadrático de variação dos intervalos de tempo entre chegadas externas das classes de produtos. Também foram realizados experimentos para analisar a sensibilidade do modelo a variações nas taxas de chegadas externas e no *mix* de produtos, que não foram aqui reportados por motivo de economia de espaço. Para mais detalhes destes experimentos, o leitor pode consultar Silva (2005).

5 Conclusões e perspectivas futuras

Neste trabalho verificou-se a aplicabilidade e adequação de modelos de rede de filas abertas para avaliação de desempenho e alocação/relocação de capacidade no sistema (SP1 e SP2), utilizando-se um estudo de caso de um sistema *job-shop* de uma empresa fabricante de peças metálicas. Uma das aproximações derivadas do método de decomposição paramétrico apresentou resultados razoavelmente próximos do WIP de produtos na rede (desvio médio de -0,9% com relação à simulação) e de *leadtime* médio de produção na rede (desvio de 3,1%). Os modelos de alocação/relocação de capacidade discreta também se mostraram adequados para a aplicação prática, permitindo a geração de curvas precisas de *trade-off* entre o capital de investimento e o capital de trabalho - WIP (modelo SP1.1 com alternativas discretas), e entre o capital de investimento e o nível de serviço - *leadtime* (modelo SP1.2 com alternativas discretas). Estas curvas podem ser úteis para auxiliar gerentes a reduzirem WIP ou *leadtimes* de produtos para níveis desejados. Por exemplo, no estudo de caso as curvas mostraram que um investimento de cerca de \$180 mil em capacidade permite reduzir o WIP da rede de cerca de \$27 mil para cerca de \$22 mil, o que pode não ser tão significativo em termos financeiros; por outro lado, este mesmo investimento permite reduzir o *leadtime* médio da

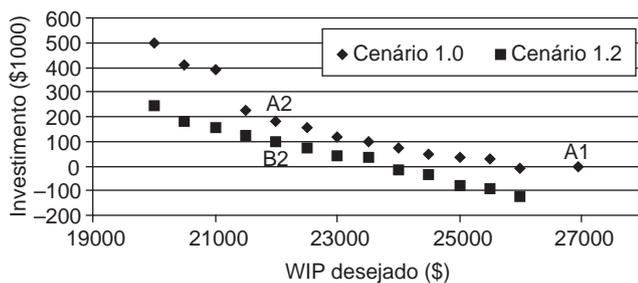


Figura 6. *Trade-off* entre investimento em recursos e WIP desejado para os cenários 0 e 2.

Tabela 9. *Leadtime* (em horas/lotte) do modelo SP1.1 com alternativas discretas para cada classe de produto para limitante superior do WIP igual a \$22000 para os cenários 0, 1 e 2.

Classe	1	2	3	4	5	6	7	8
Cenário 0	90,5	190,6	221,9	55,2	137,0	133,3	165,1	186,3
Cenário 1	92,4	194,4	226,7	58,5	139,4	134,7	166,5	187,7
Cenário 2	85,1	201,0	225,7	51,2	129,3	127,9	164,7	180,7
Classe	9	10	11	12	13	14	$E(T)$ da Rede	
Cenário 0	169,3	166,8	76,6	82,5	165,7	185,7	151,0	
Cenário 1	183,0	197,0	79,9	83,3	169,9	190,2	151,7	
Cenário 2	181,4	189,6	71,5	76,1	159,5	175,3	147,9	

rede em mais de 2 dias de operação da empresa (redução de 187,6 para 151,0 horas), o que pode ser significativo em termos de nível de serviço. Estes resultados são obtidos em poucos segundos utilizando o *software* GAMS/CPLEX® e um microcomputador Pentium III, o que reforça a aplicabilidade da metodologia aqui empregada.

Uma perspectiva interessante para pesquisas futuras é estudar melhor em projeto de experimentos a relação entre as características de redes de filas abertas e a aplicabilidade das diversas aproximações encontradas em literatura para avaliação de desempenho das redes, utilizando o método

de decomposição paramétrico. O objetivo seria identificar melhor quais sistemas de aproximação são mais adequados para quais configurações de rede. Em outras palavras, para uma dada configuração da rede (e.g., com poucos ou muitos produtos, com intensidade de tráfego leve ou pesada, com produtos pouco ou muito similares em termos de roteiros de fabricação e tempos de processamento, com produtos com pequena ou grande variabilidade no processo de chegadas externas à rede, etc.), qual seria o sistema de aproximação mais adequado para descrever o desempenho da rede, dentre os sistemas disponíveis na literatura.

Application of open queueing network models for planning a job-shop system of an actual metallurgical plant

Abstract

This paper shows the application of open queueing network models to support the design (or redesign) decision making and planning of manufacturing systems. Approximate decomposition methods are used in order to evaluate the queueing network performance. From an actual metallurgical industry queueing network, it is shown that these methods are suitable and effective to estimate, for instance, the work-in-process (WIP) and the production leadtime. It is also shown that optimization models to allocate capacity at network stations, based on decomposition methods, are effective in optimizing the performance of the network under different configurations. Trade-off curves between capacity investment and WIP or leadtime are generated and are useful in supporting a manager to estimate how much capacity should be allocated to the network to achieve the performance goal. These curves are also useful in analysing the effects in terms of capacity if the variability of external arrivals of products, the product mix or the rate of production of network changes.

Keywords: Manufacturing system design. Performance evaluation methods. Open queueing network. Optimization models. Trade-off curves.

Referências bibliográficas

- ALBIN, S. L. Poisson Approximations for Superposition Arrival Processes in Queues. **Management Science**, EUA. v. 28, n. 2, p. 126-137, 1982.
- _____. Approximating a Point Process by a Renewal Process, II: Superposition Arrival Process of Queues. **Operations Research**, USA, v. 32, n. 5, p. 1133-1162, 1984.
- ASKIN, R. G.; STANDRIDGE, C. R. **Modeling and analysis of manufacturing systems**. New York: John Wiley & Sons, 1993.
- BANKS, J.; CARSON, J. S.; NELSON, B. L. **Discrete-Event System Simulation**. 2. ed. New Jersey: Prentice Hall, 1996. 548p.
- BITRAN, G. R.; DASU, S. A review of open queueing network models of manufacturing systems. **Queueing systems**, v. 12, p. 95-134, 1992.
- BITRAN, G. R.; MORABITO, R. Um exame dos modelos de redes de filas abertas aplicados a sistemas de manufatura discretos – Parte I. **Gestão & Produção**, São Carlos, v. 2, n. 2, p. 192-219, 1995.
- _____. Open Queueing networks: Optimization and performance evaluation models for discrete manufacturing systems. **Production and Operations Management**, USA v. 5, n. 2, p. 163-193, 1996.
- _____. An overview of trade-off curves in manufacturing systems design. **Production and Operations Management**, USA, v. 8, n. 1, p. 56-75, 1999.
- BITRAN, G. R.; SARKAR, D. Targeting Problems in Manufacturing Queueing Networks – An Iterative Scheme and Convergence. **European Journal of Operational Research**, North-Holland, v. 76, p. 501-510, 1994.
- _____. Throughput Analysis in Manufacturing Networks. **European Journal of Operational Research**, North-Holland, v. 74, p. 501-510, 1994.

- BITRAN, G. R.; TIRUPATI, D. Multiproduct queueing networks with deterministic routing: Decomposition approach and the notion of interference. **Management Science**, USA, v. 34, n. 1, p. 75-100, 1988.
- _____. Capacity Planning in Manufacturing Networks with Discrete Options. **Annals of Operations Research**, USA, v. 17, p. 119-136, 1989.
- _____. Tradeoff Curves, Targeting and Balancing in Manufacturing Queueing Networks. **Operations Research**, USA, v. 37, n. 4, p. 547-564, 1989.
- BOXMA, O. J.; RINNOOY KAN, A.; VAN VLIET, M. "Machine Allocation Problems in Manufacturing Networks". **European Journal of Operational Research**, North-Holland, v. 45, p. 47-54, 1990.
- BRETHAUER, K. M. Capacity Planning in manufacturing and computer networks. **European Journal of Operational Research**, North-Holland, v. 91, p. 396-394, 1996.
- BUZACOTT, J. A. ; SHANTHIKUMAR, J. G. **Stochastic models of manufacturing system**. New Jersey: Prentice-Hall, 1993.
- CALABRESE, J. M. Optimal workload allocation in open networks of multiserver queues. **Management Science**, USA, v. 38, n. 12, p. 1792-1802, 1992.
- FRENK, H. et al. Improved algorithms for machine allocation in manufacturing systems. **Operations Research**, USA v. 42, n. 3, p. 523-530, 1994.
- GOVIL, M. K.; FU, M. C. Queueing theory in manufacturing: A survey. **Journal of Manufacturing Systems**, USA, v. 18, n. 3, p. 214-240, 1999.
- KETZENBERG, M. E.; SOUZA, G. C.; GUIDE, V. D. R. Mixed assembly and disassembly operations for remanufacturing. **Production and Operations Management**, USA, v. 42, n. 3, p. 523-530, 2003.
- KERBACHE, L.; SMITH, J. M. Queueing networks and the topological design of supply chain systems. **International Journal of Production Economics**, North-Holland, v. 91, p. 251-272, 2004.
- KLEINROCK, L. **Queueing Systems**, vol. 1. John Wiley & Sons, New York, 1975. 417p.
- KRAJEWSKI, L. J. ; RITZMAN, L. P. **Operations management: strategy and analysis**. 4 ed. Massachusetts: Addison-Wesley Publishing Company, 1996.
- KOUVELIS, P.; TIRUPATI, D. Approximate performance modeling and decision making for manufacturing systems: A queueing network optimization framework. **Journal of Intelligent Manufacturing**, v. 2, p. 107-134, 1991.
- KOUVELIS, P.; CHAMBERS, C.; YU, D. Z. Manufacturing operations manuscripts published in the first 52 issues of POM: Review, trends and opportunities. **Production and Operations Management**, USA, v. 14, n. 4, p. 450-467, 2005.
- KUEHN, P. J. Approximate Analysis of General Networks by Decomposition. **IEEE Transactions Commun**, v. 27, n. 1, p. 113-126, 1979.
- LAW, A. M.; KELTON, W. D. **Simulation Modeling and Analysis**. 2. ed. New York: McGraw-Hill, 1991. 759p.
- MORABITO, R. **Análise de Curvas de Trade-off Baseada em Teoria de Redes de Filas para o Projeto e Planejamento de Sistemas Discretos de Manufatura**. 1998. 136 f.. Tese (Livredocência em Engenharia Produção) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, SP.
- NAKANO, M.; OHNO, K. Decomposition algorithm for performance evaluation of AGV systems. **Production and Operations Management**, USA, v. 8, n. 2, p. 193-205, 1999.
- SCHRIBER, T. J. **An Introduction to Simulation Using GPSS/H**. New York: John Wiley & Sons, 1991. 437p.
- SEGAL, M.; WHITT, W. A Queueing Network Analyzer for Manufacturing. In: BONATTI, M. **Teletraffic Science for New-Effective System, Networks and Services**. North-Holland: Elsevier, 1989. p. 1146-1152.
- SESHADRI, S.; PINEDO, M. Optimal allocation of resources in a job shop environment. **IIE Transactions**, v. 31, p. 195-206, 1999.
- SHANTHIKUMAR, J. G.; BUZACOTT, J. A. Open queueing network models of dynamic job shops. **International Journal of Production Research**, USA, v. 19, n. 3, p. 255-266, 1981.
- SHANTHIKUMAR, J. G.; XU, S. H. Strongly asymptotically optimal design and control of production and service systems. **IIE Transactions**, v. 32, p. 881-890, 2000.
- SILVA, C. R. N. **Aplicação de Modelos de Redes de Filas Abertas no Projeto e Planejamento de Sistemas Discretos de Manufatura**. 2005. 272 f.. Tese (Doutorado em Engenharia de Produção) – Programa de Pós-graduação em Engenharia de Produção, Universidade Federal de São Carlos, São Carlos, SP.
- SILVA, C. R. N.; MORABITO, R. Análise de problemas de partição de instalações em sistemas *job-shops* por meio de modelos de redes de filas. **Pesquisa Operacional**, v. 27, n. 2, p. 333-356, 2007.
- SOUZA, G. S.; WAGNER, H. M.; WHYBARK, D. C. Evaluating focused factory benefits with queueing theory. **European Journal of Operational Research**, North-Holland, v. 128, p. 597-610, 2001.
- SOUZA, G. S.; KETZENBERG, M. E.; GUIDE, V. D. R. Capacitated remanufacturing with service level constraints. **Production and Operations Management**, USA, v. 11, n. 2, p. 231-248, 2002.
- SOUZA, G. S.; KETZENBERG, M. E. Two-stage make-to-order remanufacturing with service-level constraints. **International Journal of Production Research**, UK, v. 40, n. 2, p. 477-493, 2002.
- SUNDARRAJ, R. P.; SUNDARARAGHAVAN, P. S.; FOX, D. R. Optimal server acquisition in open queueing networks. **Journal of the Operational Research Society**, v. 45, n. 5, p. 549-558, 1994.
- SURI, R.; SANDERS, J. L.; KAMATH, M. Performance evaluation of production networks. In: GRAVES, S. C. **Handbooks in OR/MS**. North Holland: Elsevier, v. 4, p. 199-286, 1993.
- TIJMS, H. C. **Stochastic modeling and analysis: A computational approach**. New York: John Wiley & Sons, 1986.
- VAN VLIET, M.; RINNOOY KAN, A. Machine Allocation Algorithms for Job-Shop Manufacturing. **Journal of Intelligent Manufacturing**, v. 2, p. 83-94, 1991.
- WARSING, D. P.; SOUZA, G. C.; GREIS, N. P. Determining the value of dedicated multimodal cargo facilities in a multi-region distribution network. **European Journal of Operational Research**, North-Holland, v. 133, p. 81-93, 2001.
- WHITT, W. The queueing network analyzer. **The Bell System Technical Journal**, USA, v. 62, n. 9, p. 2779-2815, 1983.
- _____. The queueing network analyzer. **The Bell System Technical Journal**, USA, v. 63, n. 9, p. 1911-1979, 1983.
- _____. Approximations for the GI/G/m Queue. **Production and Operations Management**, USA, v. 2, n. 2, p. 114-161, 1993.
- _____. Towards Better Multi-Class Parametric-Decomposition Approximations for Open Queueing Networks. **Annals of Operations Research**, USA, v. 48, p. 221-248, 1994.
- _____. Variability functions for parametric decomposition approximations of queueing networks. **Management Science**, USA, v. 41, n. 10, p. 1704-1715, 1995.

Sobre os autores

Cláudio Rogério Negri da Silva**Reinaldo Morabito**

Departamento de Engenharia de Produção, Universidade Federal de São Carlos – UFSCar,
Rodovia Washington Luís, SP-310, Km 235, São Carlos, SP, Brasil,
e-mails: claudio.negri@embraer.com.br; morabito@power.ufscar.br

Agradecimentos: Os autores agradecem aos revisores anônimos pelos úteis comentários e sugestões e pelo apoio da FAPESP e do CNPq.

Recebido em 22/8/2006
Aceito em 28/5/2007