

**UNIVERSIDADE FEDERAL DE SANTA CATARINA  
DEPARTAMENTO DE ENGENHARIA DE PRODUÇÃO  
E SISTEMAS**

Elizangela Dias Pereira

**A OPERAÇÃO PORTUÁRIA DO PORTO DO RIO  
GRANDE MODELADA COMO UMA REDE DE FILAS  
COM ALOCAÇÃO DE *BUFFERS***

Florianópolis

2019



Elizangela Dias Pereira

**A OPERAÇÃO PORTUÁRIA DO PORTO DO RIO  
GRANDE MODELADA COMO UMA REDE DE FILAS  
COM ALOCAÇÃO DE *BUFFERS***

Tese submetida ao Programa de Pós-Graduação em Engenharia de Produção para a obtenção do Grau de Doutora em Engenharia de Produção.

Orientador

Universidade Federal de Santa Catarina: Prof. Dr. Antônio Sérgio Coelho

Florianópolis

2019

Ficha de identificação da obra elaborada pelo autor,  
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Pereira, Elizangela Dias

A operação portuária do porto do Rio Grande modelada como uma rede de filas com alocação de buffers / Elizangela Dias Pereira ; orientador, Antônio Sérgio Coelho, 2019.  
99 p.

Tese (doutorado) - Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós Graduação em Engenharia de Produção, Florianópolis, 2019.

Inclui referências.

1. Engenharia de Produção. 2. Chegadas em bloco. 3. Rede de filas. 4. Alocação de buffers. I. Coelho, Antônio Sérgio . II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Engenharia de Produção. III. Título.

Elizangela Dias Pereira

**A OPERAÇÃO PORTUÁRIA DO PORTO DO RIO  
GRANDE MODELADA COMO UMA REDE DE FILAS  
COM ALOCAÇÃO DE *BUFFERS***

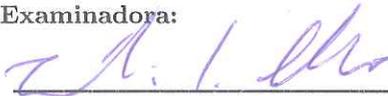
Esta Tese foi julgada aprovada para a obtenção do Título de “Doutora em Engenharia de Produção”, e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia de Produção.

Florianópolis, 01 de março 2019.

---

Profa. Dra. Lucila Maria de Souza Campos  
Coordenadora  
Universidade Federal de Santa Catarina

**Banca Examinadora:**



---

Prof. Dr. Antônio Sérgio Coelho  
Orientador  
Universidade Federal de Santa Catarina

---

Prof. Dr. Sérgio Fernando Mayerle  
Universidade Federal de Santa Catarina

---

Prof. Dra Vanina Macowski Durski Silva  
Universidade Federal de Santa Catarina  
(videoconferência)

---

Profa. Dra. Catia Maria dos Santos Machado  
Universidade Federal do Rio Grande - FURG

Dedico este trabalho ao meu irmão Luiz  
Ricardo, onde quer que ele esteja.



## AGRADECIMENTOS

Ao meu orientador, Prof. Dr. Antônio Sérgio Coelho, pelo aceite e acolhimento no decorrer do curso, por me compreender e permitir escolher o caminho a ser trilhado.

Aos meus pais Helen e Nilza pela incansável luta diária, por toda ajuda financeira e emocional, por todo amor dedicado dia após dia, sem eles eu nada seria nessa vida.

Aos meus irmãos Luiz Ricardo (*in memoriam*) e Mário César, e suas famílias, por acreditarem em mim e torcerem para dar certo.

À Catia, professora na graduação, orientadora no mestrado, colega de profissão e acima de tudo uma amiga presente em todos os momentos, responsável por parte das minhas escolhas e principalmente por me manter firme e forte no alcance dos meus objetivos.

Aos meus amigos Daniel, Rita e Carolina, meus fiéis escudeiros desde tempos remotos, por todas as vezes que desabei, que pensei em desistir e eles estavam ali, pra me dar a mão e me fazer continuar segurando forte no leme.

Aos amigos Emerson e Maria Inácia, colegas de curso e parceiros pra toda vida, pelos intervalos regados com muito café e boa conversa, pelo empréstimo de seus ombros para dividir o pesado fardo de uma pós-graduação.

Ao colega Alexandre pela parceria durante o curso, juntos na qualificação e na defesa, dividindo anseios e sonhos para nossa vida profissional.

Aos colegas de curso da Universidade Federal do Pampa por compreender meus afastamentos e de alguma forma aliviar meus encargos para que pudesse concluir essa etapa.

Ao aluno Welbert pela implementação dos algoritmos desenvolvidos para este trabalho, mostrando que nunca seremos seres humanos completos.

Aos meus vizinhos, Sônia e Antônio, que torceram por mim desde que cheguei em Bagé e me ajudam a vencer os dias longe da família.

E, finalmente, agradeço imensamente a Deus por ter-me concedido resiliência ao longo desses 6 anos para que todos os dias eu pudesse acreditar que é possível vencer, apesar de tudo.



*“Nãõ basta fazer o que se ama, é preciso amar o que se faz”. (Autor desconhecido)*



## RESUMO

Esta tese tem como escopo apresentar uma modelagem para a operação portuária de um terminal de contêineres, com intuito de subsidiar os tomadores de decisão quanto à melhor utilização dos recursos disponíveis. A operação portuária pode ser vista como uma rede de filas com alocação de *buffers*, formada por navios, caminhões e trens que efetuam a carga e/ou descarga de produtos, analogamente a um sistema de manufatura, que pode ser modelado por redes de filas, a rede portuária seria formada por sistemas de filas conectados entre si, em que os usuários se movem entre eles para receber um serviço. O principal desafio encontrado foi modelar a chegada de contêineres, uma vez que a entidade de interesse na rede de filas é o contêiner. Para tanto, após intensa pesquisa bibliográfica, foi desenvolvido um modelo de filas com chegadas em bloco. A utilização de modelagem e programação matemática nesta tese teve como objetivo verificar se o modelo proposto é exequível para o caso real estudado, considerando a natureza aleatória das operações portuárias como um todo. Foram utilizados algoritmos específicos para o problema de redes de filas, como o Método de Expansão Generalizada para o tratamento da rede e o algoritmo de otimização NSGA-II para solucionar o problema de otimização multiobjetivo. Foi obtido um modelo de redes de filas com alocação de *buffers* que representa uma operação portuária, e que fornece um conjunto de resultados passível de análise de custos em investimentos.

**Palavras-chave:** Chegadas em Bloco, Problema de Alocação de *Buffers*, Redes de Filas, Terminais de contêineres



## ABSTRACT

This thesis aims to present a model for the port operation of a container terminal, in order to subsidize the decision makers regarding the best use of available resources. The port operation can be seen as a net of queues with allocation of its buffers, consisting of ships, trucks and trains that load and / or unload products, similarly to a manufacturing system, which can be modelled by networks, the port network would be composed of connected queuing systems, which can be modelled for sequences of nets, the port net have been formed for a queues system connected to each other, wherein the users could move between them to receive an order. The main challenge was to model the arrival of containers, once the entity of interest in the queue net is the container, so, after an intense bibliographical research, a model of queue was developed with blocks arrivals. The use of mathematical modelling and programming in this thesis have an objective of verify if the proposed model is feasible for the actual case studied, considering the random nature of the port operations as a whole. Specific algorithms were used for the problem of queuing nets, such as the Generalized Expansion Method for the net treatment and the NSGA-II to solve the problem of optimization algorithm multiobjective. A queuing network model with its buffers allocation was obtained that represents a port operation, and provides a set of results that can be analyzed for cost analysis in investments.

**Keywords:** Batch Arrivals, Buffers Allocation Problem, Queue Networks, Container Terminals



## LISTA DE FIGURAS

Figura 1	Processo de tradução contextual.....	27
Figura 2	Fases de estudo.....	28
Figura 3	Portos organizados no Brasil.....	30
Figura 4	Localização do Porto do Rio Grande.....	32
Figura 5	Super Porto.....	33
Figura 6	TECON Rio Grande S/A.....	35
Figura 7	Esquema geral de uma operação portuária.....	37
Figura 8	Esquema lateral de um terminal de contêineres.....	39
Figura 9	Processo de descarregamento/carregamento de navios..	40
Figura 10	Movimentação do Gate in/out.....	41
Figura 11	Estrutura de um Sistema de Filas.....	50
Figura 12	Tipos de filas.....	55
Figura 13	Exemplo de rede fechada.....	56
Figura 14	Exemplo de rede aberta.....	56
Figura 15	Modelo Inicial de Rede de Filas.....	65
Figura 16	Modelo de Rede de Filas.....	66
Figura 17	Método de Expansão Generalizada.....	73
Figura 18	Fluxograma NSGA-II.....	77
Figura 19	Procedimentos de Seleção do NSGA-II.....	78
Figura 20	Comparação dos valores de $\mu$ e $\theta$ .....	89
Figura 21	Comparação dos valores de $K$ e $\theta$ .....	89



## LISTA DE TABELAS

Tabela 1	Parâmetros NSGA-II.....	81
Tabela 2	Distribuição de Probabilidade do Tamanho dos Blocos.	84
Tabela 3	Dados de entrada para o Algoritmo 1.....	85
Tabela 4	Resultados para fila chegada em blocos.....	85
Tabela 5	Dados fornecidos pelo terminal 1.....	87
Tabela 6	Soluções de Pareto selecionadas.....	88



## LISTA DE ABREVIATURAS E SIGLAS

ANTAQ	Agência Nacional de Transporte Aquaviário
FMI	Fundo Monetário Nacional
GNL	Gás Natural Liquefeito
TEU	twenty-feet-equivalent-unit
FEU	forty-feet-equivalent-unit
SUPRG	Superintendência do Porto do Rio Grande
AGV	Automated Guided Vehicles
FIFO	First-in-First-Out
LIFO	Last-in-First-Out
PAB	Problema de Alocação de <i>Buffers</i>
WIP	<i>work-in-process</i>
FCFS	<i>First Come - First Served</i>
NSGA	<i>Non-Dominated Sorting Genetic Algorithm</i>
GEM	<i>Generalized Expansion Method</i>
AG	Algoritmo Genético
AGMOs	Algoritmos Genéticos Multiobjetivos
SBX	<i>Simulated Binary Crossover</i>



## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	23
1.1	CONTEXTUALIZAÇÃO .....	23
1.2	OBJETIVOS .....	25
1.2.1	Objetivo geral.....	25
1.2.2	Objetivos específicos .....	25
1.3	LIMITAÇÕES E CONTRIBUIÇÃO CIENTÍFICA .....	25
1.4	MÉTODOS E TÉCNICAS DE PESQUISA .....	26
1.5	ESTRUTURA DA TESE.....	28
<b>2</b>	<b>CARACTERIZAÇÃO DO OBJETO DE ESTUDO</b>	29
2.1	INTRODUÇÃO.....	29
2.2	PORTO DO RIO GRANDE .....	31
2.2.1	<b>TECON Rio Grande</b> .....	34
2.3	TERMINAIS DE CONTÊINERES E OPERAÇÃO PORTUÁRIA .....	35
2.4	A OPERAÇÃO PORTUÁRIA DO TERMINAL ESTUDADO .....	40
<b>3</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b> .....	43
3.1	INTRODUÇÃO.....	43
3.1.1	Processos estocásticos.....	43
3.1.2	Cadeias de Markov.....	44
3.1.2.1	Equações de Chapman-Kolmogorov .....	45
3.1.2.2	Probabilidade de estado estável .....	46
3.1.3	Cadeias de Markov de tempo contínuo .....	47
3.1.3.1	Probabilidade de estado estável .....	48
3.2	TEORIA DAS FILAS.....	48
3.2.1	Introdução .....	48
3.2.2	Características de uma fila .....	49
3.2.3	Notação de Kendall .....	52
3.2.4	Medidas de desempenho .....	53
3.2.5	Relação entre $L$ , $W$ , $L_q$ e $W_q$ .....	54
3.3	REDES DE FILAS .....	54
3.3.1	Introdução .....	54
3.3.2	Definição e características .....	55
3.4	O PROBLEMA DE ALOCAÇÃO DE <i>BUFFERS</i> (PAB) .	58
3.4.1	Introdução .....	58
3.4.2	Definição.....	59
3.4.3	Abordagem multiobjetivo para o PAB .....	61

3.4.4	Aplicações do PAB .....	62
4	MODELO E MÉTODO DE SOLUÇÃO.....	65
4.1	DEFINIÇÃO DAS VARIÁVEIS .....	65
4.2	MODELAGEM DO PROBLEMA .....	66
4.2.1	Modelo de Chegadas em Bloco .....	67
4.3	MÉTODO APLICADO .....	72
4.3.1	Método de Expansão Generalizada .....	72
4.3.2	Algoritmo de Otimização Multiobjetivo .....	76
5	RESULTADOS E DISCUSSÃO .....	83
5.1	INTRODUÇÃO .....	83
5.2	MODELAGEM DA CHEGADA DE CONTÊINERES ...	83
5.3	RESULTADOS GEM E NSGA-II .....	86
6	CONSIDERAÇÕES FINAIS.....	91
6.1	PRINCIPAIS APONTAMENTOS .....	91
6.2	SUGESTÕES PARA PESQUISAS FUTURAS .....	92
	REFERÊNCIAS .....	93

# 1 INTRODUÇÃO

## 1.1 CONTEXTUALIZAÇÃO

O transporte marítimo é a modalidade do transporte aquaviário que utiliza os mares abertos, tendo sido esse modal responsável por 83,5% do volume de exportações brasileiras em 2018, totalizando 521 milhões de toneladas exportadas (ANTAQ, 2019). Esse modal é de suma importância para o país, impactando diretamente na balança comercial, que fechou em superávit no ano de 2018 com um aumento de 9,6% do volume de exportações e 19,7% de importação. Um dos principais papéis do sistema portuário diz respeito aos custos e à eficiência da logística de transportes do país, o que impacta diretamente na competitividade dos produtos nacionais no exterior. O Brasil apesar de constituir a oitava economia do mundo, ainda está aquém de oferecer excelência em infraestrutura de transporte.

O processo de modernização portuária, implementado no Brasil, a partir da Lei de Modernização dos Portos (Lei 8.630/93) (BRASIL, 26 fev. 1993), buscou solucionar os problemas gerados pelos altos custos, baixa produtividade, serviços não competitivos, excesso de pessoal, subsídios e burocracia governamental. Todas essas características configuram o modelo portuário latino-americano tradicional, e levam a um intenso processo de reestruturação e de reformas nos portos, objetivando a compatibilidade com o acelerado crescimento do comércio entre os países e os blocos econômicos e a demanda por uma produção mais eficiente (KITZMANN, 2010).

O setor portuário brasileiro é responsável por 95% do volume total referente ao comércio exterior. Segundo a Agência Nacional de Transporte Aquaviário (ANTAQ), a movimentação de contêineres nos portos brasileiros no ano de 2018 obteve um aumento de 7,2% em relação ao ano anterior, com movimentação de cerca de 10 milhões de TEU's (ANTAQ, 2019).

O Porto do Rio Grande, no Rio Grande do Sul, é um típico porto de 3ª geração, marcado por funções comerciais, industriais e de plataforma logística para contêineres do tipo *directed call hubport*<sup>1</sup>, isto é, um porto de chamada direta, como todos os demais portos que constituem o sistema portuário brasileiro. Segundo a ANTAQ (2019),

---

<sup>1</sup>Porto concentrador de cargas e de linhas de navegação com operação do tipo chamada direta entre portos.

o Porto do Rio Grande está entre os 5 maiores portos de contêineres do Brasil, que movimentam juntos 61% do total de TEU's no país.

Apesar do aumento na tonelagem transportada nos últimos anos, a ineficiência dos portos se faz presente prejudicando a participação do Brasil no mercado exterior. Esse fato deve-se principalmente à infraestrutura precária e estrutura de fiscalização considerada obsoleta e burocrática, e ao modelo administrativo inapropriado bem como congestionamentos nas rodovias de acesso aos portos (ANTAQ, 2019).

No entanto, a convergência de inúmeros projetos portuário-industriais, bem como de macrologística nacional e regional, apontam para uma perspectiva de rápida elevação nas tonelagens totais a serem operadas no Porto do Rio Grande no período de 2015-2020. Projetos avançados pela indústria naval e *offshore*, a estruturação de um grande polo energético na cidade e região a partir da utilização de GNL (Gás Natural Liquefeito) e parques eólicos, exploração de minerais pesados, indústrias químicas e de alimentos. Além disso, cargas hidroviárias vindas do Uruguai e a extensão de dois ramais ferroviários conectando o porto ao Cerrado Brasileiro, apontam para a triplicação das tonelagens atualmente movimentadas até 2020 (SUPRG, 2019).

Faz-se necessário, então, que haja uma modernização das estruturas físicas do porto, bem como de processos gerenciais dando suporte às operações portuárias. Nesse sentido pesquisadores buscam alternativas que possam colaborar com o melhor desempenho do setor portuário, no que diz respeito à eficiência do sistema, bem como à busca por investimentos em infraestrutura aliado a estudos de análise de capacidade e desempenho.

A presente tese trata da operação portuária como um sistema único, desde o momento em que o navio chega ao porto até o momento em que os contêineres deixam o terminal. Para tanto, será configurada como um rede de filas com alocação de *buffers* descrevendo as diferentes etapas de uma operação portuária buscando por possíveis gargalos no sistema e ainda, identificar ações que possam vir a serem tomadas por seus decisores.

## 1.2 OBJETIVOS

### 1.2.1 Objetivo geral

Esta tese tem como objetivo apresentar uma modelagem para a operação portuária de um terminal de contêineres como uma rede de filas com alocação de *buffers*, baseada na teoria de filas e redes de filas, com o intuito de subsidiar os tomadores de decisão quanto à melhor utilização dos recursos disponíveis.

### 1.2.2 Objetivos específicos

- a) Desenvolver um modelo de rede com alocação de *buffers* para a operação portuária;
- b) Formular matematicamente o modelo formulado;
- c) Modelar a chegada de contêineres no porto;
- d) Desenvolver um algoritmo específico para solucionar a modelagem da chegada de contêineres;
- e) Implementar um algoritmo para resolver o modelo proposto.

## 1.3 LIMITAÇÕES E CONTRIBUIÇÃO CIENTÍFICA

Por se tratar de um sistema complexo, algumas simplificações precisaram ser feitas a fim de que se tenha um modelo tangível. A obtenção de dados também foi um processo lento e dificultoso, pois o terminal não dispunha de todos os dados necessários, e os dados fornecidos de forma bruta tiveram de ser tratados e filtrados. Em algum momento foi necessário contar com ferramentas estatísticas para complementar tais informações, não representando um prejuízo ao estudo apresentado, uma vez que se tenha acesso a dados em sua totalidade, o modelo continua sendo exequível da mesma forma.

O modo escolhido para representar uma operação portuária exerceu forte impacto no estudo bibliográfico, uma vez que o assunto ainda é pouco difundido cientificamente, e mesmo tendo diferentes aplicações no campo da logística, em geral os estudos versam sobre problemas de manufatura e não sobre sistemas portuários. Nesse sentido tem-se

uma importante contribuição científica que é a difusão do modelo de redes de filas para sistemas portuários e, de forma adjacente, o tratamento especial dado as chegadas de navios no porto. Uma vez que o interesse do estudo está voltado à entidade contêiner, precisou-se tratar esse evento como uma fila com chegadas em blocos, assunto esse, pouco encontrado na literatura, principalmente de forma aplicada.

Dentro de um sistema de uma operação portuária são muitos os problemas enfrentados diariamente, ocasionando a impossibilidade de serem tratados ao mesmo tempo, ou seja, a partir de uma modelagem única. Sendo assim, a presente tese limita-se a construir um modelo com intuito de servir como ferramenta de análise do processo de operação, conforme configuração geral apresentada no decorrer do texto.

A meta principal desta tese é formular um modelo de análise do processo de uma operação portuária, que servirá como ferramenta de apoio à decisão, utilizando um terminal específico como base de estudos. Sendo assim, o modelo aqui proposto deverá servir para terminais que operam de maneira semelhante, e com os ajustes pertinentes, esse mesmo modelo poderá ser aplicado em diferentes portos.

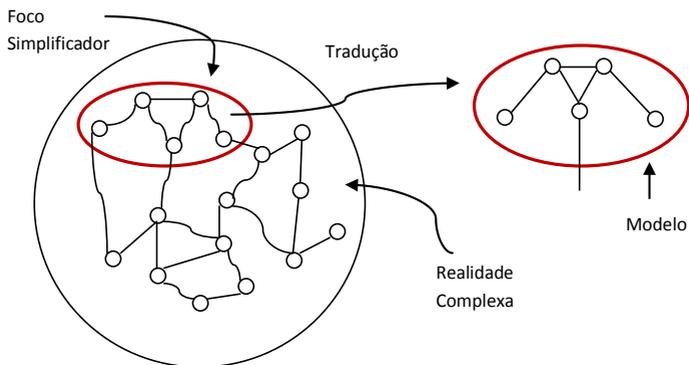
## 1.4 MÉTODOS E TÉCNICAS DE PESQUISA

Esta tese caracteriza-se como uma pesquisa empírica descritiva quantitativa, que, de acordo com Miguel (2012), tem por objetivo criar um modelo que descreva de forma adequada as relações causais que podem existir na realidade, o que leva a uma compreensão dos processos reais.

Cabe salientar que esta abordagem, baseada em modelagem quantitativa no âmbito de gestão de produção e operações, possui a premissa de que é possível construir modelos que expliquem pelo menos parte do comportamento de processos reais ou que é possível pelo menos capturar uma parte dos problemas de tomada de decisão encontrados em sistemas reais (MIGUEL, 2012).

Modelar consiste em abstrair da situação real um conjunto de variáveis dominantes que controlam o comportamento do sistema real, e este modelo deve ainda expressar de maneira tratável as funções matemáticas que representam o funcionamento do mundo real considerado (TAHA, 2008). Um bom modelo exige uma tradução contextual conveniente, a Figura (1) ilustra esse processo de tradução destacando os aspectos simplificador e estruturador necessários.

Figura 1: Processo de tradução contextual



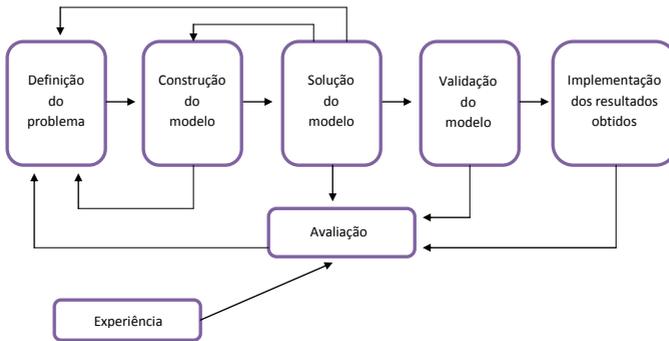
Fonte: Adaptada de (GOLDBARG; LUNA, 2000)

Um modelo que possui desempenho suficientemente próximo do desempenho da realidade e é de fácil experimentação, é dito um bom modelo. Dependendo do objetivo proposto inicialmente, a proximidade com a realidade torna-se variável, a fidelidade de um modelo aumenta à medida que ele incorpora características da realidade a partir da adição de novas variáveis. Quanto mais variáveis incorporadas, maior a complexidade e, por consequência imediata, a experimentação é dificultada, o que leva a considerar nesse momento o fator do custo-benefício quando se deseja melhorar o desempenho do modelo (TAHA, 2008).

As etapas da pesquisa seguem a proposta de (ANDRADE, 1989) conforme ilustradas na Figura (2). Uma vez definido o problema parte-se para a etapa de construção do modelo e a busca por métodos adequados de solução. Segundo Goldbarg e Luna (2000), um modelo é uma forma de visualização estruturada da realidade, ou ainda, uma representação simplificada da realidade que preserva, para determinadas condições e enfoques, uma equivalência adequada.

A etapa da modelagem ou construção do modelo pode ser considerada a fase de maior importância no desenvolvimento da pesquisa, sendo descrita em detalhes no Capítulo 4. Após a modelagem e a escolha pelo método de solução foi necessário a tomada de dados e o devido tratamento dos mesmos, o que foi feito a partir do software ARENA<sup>®</sup> e planilhas eletrônicas.

Figura 2: Fases de estudo



Fonte: Adaptada de (ANDRADE, 1989)

A utilização de modelagem e programação matemática nesta tese tem como objetivo verificar se o modelo proposto é exequível para o caso real estudado, considerando a natureza aleatória das operações portuárias como um todo.

## 1.5 ESTRUTURA DA TESE

Esta tese está organizada em seis capítulos, incluindo esta introdução. O Capítulo 2 traz os principais resultados que versam sobre operação portuária e a descrição do universo de pesquisa, no caso, um terminal de contêineres e suas peculiaridades. No Capítulo 3 são apresentados os estudos acerca do Problema de Alocação de *Buffers* encontrados na literatura. No mesmo capítulo são mostrados ainda detalhes a respeito de Teoria das Filas, Redes de Filas, Chegadas em bloco e estudos recentes que servem como suporte a toda a pesquisa realizada.

Os métodos de solução e o modelo proposto nesta tese são apresentados no Capítulo 4. Os principais resultados obtidos são relatados no Capítulo 5 e, finalizando, o último capítulo traz as considerações finais acerca do estudo.

## 2 CARACTERIZAÇÃO DO OBJETO DE ESTUDO

### 2.1 INTRODUÇÃO

Dentre os meios de transportes, o mais antigo é o transporte marítimo, sendo utilizado desde tempos remotos. No entanto, seu incremento aconteceu efetivamente após o término da Primeira Guerra Mundial, como resultado de grandes inovações no campo tecnológico levando a significativas evoluções no transporte de cargas. Dentre as principais mudanças pode-se destacar a melhoria em relação à capacidade de carga a ser transportada nos navios, além da criação de embarcações específicas, especializadas no transporte de determinados tipos de cargas.

Mais adiante, na década de 60 surge a evolução no transporte de carga por meio de contêineres, recipientes de metal com tamanhos padronizados. Os principais tamanhos são os contêineres de 20 pés (TEU - *twenty-feet-equivalent-unit*) e os de 40 pés (FEU - *forty-feet-equivalent-unit*), cujas dimensões são, respectivamente (6,1 m x 2,44 m x 2,59 m) e (12,2 m x 2,44 m x 2,59 m) (CARLO; VIS; ROODBERGEN, 2014b). A unidade de 20 pés (TEU) é utilizada para calcular o volume de um contêiner e serve ainda como unidade universal para medidas de produtividade por exemplo.

O sucesso na containerização de cargas é devido a diferentes fatores, como a queda no valor do frete, incentivado pela concorrência entre as companhias de navegação, bem como a possibilidade de redução no tempo de operação de um navio, que antes permanecia atracado por vários dias, e hoje realiza o carregamento e a descarga em poucas horas. Pode-se citar também a questão da armazenagem e o próprio transporte intermodal, em decorrência da padronização, os contêineres podem ser empilhados de forma mais eficiente em navios, trens e caminhões, e ainda podem ser facilmente manipuláveis por guindastes (AZEVEDO A. T.; CHAVES, 2014).

O papel mais importante do sistema portuário diz respeito aos custos e à eficiência da logística de transportes de um país, gerando assim um impacto direto na competitividade dos produtos nacionais no exterior. Dessa forma, os portos são o principal ponto intermodal, onde diferentes tipos de meios de transporte (navios, caminhões e trens) reúnem-se para troca de cargas e mercadorias.

No Brasil, ao todo existem 34 portos organizados (3) e nessa categoria estão os portos com administração exercida pela União, no

caso das Companhias das Docas, ou delegada a municípios, estados ou consórcios públicos. Um porto organizado é aquele construído e aparelhado para atender as necessidades da navegação e da movimentação e armazenagem de mercadorias, concedido ou explorado pela União, cujo tráfego e operações portuárias estejam sob a jurisdição de autoridade pública (ANTAQ, 2019).

Figura 3: Portos organizados no Brasil



Fonte: (ANTAQ, 2019)

Outra nomenclatura utilizada pela ANTAQ é de porto marítimo ou fluvial de acordo com a navegação de longo curso ou interior, que são portos aptos a receber linhas de navegação oceânicas, tanto em navegação internacional (longo curso) como em navegação de cabotagem (interior), independente de sua localização geográfica. Ademais, terminal portuário é uma instalação portuária, normalmente um cais ou pier especializado com equipamentos apropriados para movimentação de um determinado tipo de carga, sejam elas granel sólido, granel líquido, carga geral ou exclusivamente contêiner.

Segundo AZEVEDO A. T.; CHAVES (2014), terminais portuários podem ser classificados, de acordo com o tipo de carga movimentada, em quatro tipos principais:

- Contêiner: movimentação de produtos armazenados em recipientes que recebem o nome de contêiner.
- *Dry Bulk* ou Granel Sólido: produtos como soja, fertilizantes, carvão e minério de ferro.
- *Liquid Bulk* ou Granel Líquido: petróleo e seus derivados, gás na forma liquefeita.
- Carga Geral e terminais multi-propósito: são terminais que agregam diferentes tipos de cargas, como granel sólido e contêineres, e ainda podem movimentar produtos como carros.

Além dos principais tipos citados, ainda existem os terminais específicos para movimentação de frutas, pescado, passageiros ou veículos de passeio.

## 2.2 PORTO DO RIO GRANDE

O Porto do Rio Grande é um dos três portos organizados do Estado do Rio Grande do Sul, sendo o mais importante como único porto marítimo público, dotado de características naturais privilegiadas, com condições de atender à navegação de longo curso que exige boas profundidades. É localizado na cidade do Rio Grande, no litoral sul do Estado do Rio Grande do Sul, na margem oeste do Canal do Norte, que é o escoadouro natural da bacia hidrográfica da Lagoa dos Patos e está sob administração da Superintendência do Porto do Rio Grande (SUPRG).

Segundo dados da ANTAQ (2012), o Porto do Rio Grande é um típico porto de 3ª geração, marcado por funções comerciais, industriais e de plataforma logística para contêineres do tipo *directed call hubport*. Suas principais características são:

- acesso rodoviário, ferroviário e fluvial;
- canal de 230 m de largura e 16 m de profundidade;
- mais de 4.000 m de extensão de cais (Superporto, Porto Velho e Porto Novo);
- influência primária: RS, SC, Paraguai e Argentina;
- 20 armazéns com área total de 163.000 m<sup>2</sup>, área de pátio com 120.000 m<sup>2</sup> e 100.000 m<sup>2</sup> de pátio para automóveis;

- tipos de cargas movimentadas: soja em grão, farelo de soja, fertilizantes, cavaco de madeira, óleo de soja, trigo, celulose, fumo, frango congelado, calçados, maquinário agrícola e contêineres.

Sua área de atuação compreende os Estados do Rio Grande do Sul e Santa Catarina, assim como países vizinhos, como é o caso do Uruguai, o sul do Paraguai e o norte da Argentina. Devido à proximidade a esses países, o porto tornou-se conhecido como Porto do MERCOSUL. Na Figura 4 pode ser vista a localização do porto.

Figura 4: Localização do Porto do Rio Grande



Fonte: (SUPRG, 2019)

A área do Porto Organizado do Rio Grande é compreendida por três áreas distintas denominadas: Porto Velho, Porto Novo e Super-

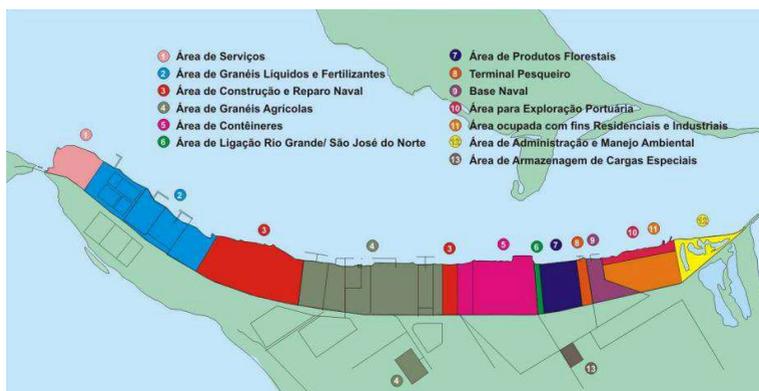
porto e mais uma área de expansão portuária que se localiza no município vizinho, São José do Norte, na margem esquerda do canal de acesso ao porto (SUPRG, 2019).

No Porto Velho concentram-se o cais pesqueiro e o terminal de passageiros para a travessia entre Rio Grande e São José do Norte e conta ainda, com uma área destinada às atividades de ensino e pesquisa vinculadas à Universidade Federal do Rio Grande - FURG.

O Porto Novo conta com sete áreas de atendimento portuário, sendo que as principais são o cais para movimentação de navios *roll-on-roll-off*<sup>1</sup> e o pátio automotivo, uma área para operação de contêineres e fertilizantes, e uma área destinada à construção e reparo naval.

O Superporto mostrado na Figura 5, dispõe ao todo de 1.552 metros de cais com profundidades variando de 5 m a 14,5 m, dividido em 13 áreas, onde estão instalados os seus principais terminais especializados, dentre eles, Pier Petrolero, Yara Brasil, Bunge Fertilizantes e Bunge Alimentos, Estaleiro Rio Grande, Termasa e Tergrasa, Bras-kem e o terminal especializado em movimentação e armazenagem de contêineres, TECON Rio Grande S/A.

Figura 5: Super Porto



Fonte: (SUPRG, 2019)

A área 5, na Figura 5, compreende o terminal especializado em contêineres, TECON Rio Grande S/A, arrendado dentro dos limites do porto organizado de acordo com a Lei 8.630/93 (BRASIL, 26 fev. 1993).

<sup>1</sup>Tipo de cargueiro que transporte automóveis e outros veículos

Na próxima seção será apresentado mais detalhes sobre este terminal, que é o foco do estudo dessa tese.

### 2.2.1 TECON Rio Grande

O TECON Rio Grande é a empresa que venceu a licitação e administra o terminal de contêineres do Porto do Rio Grande desde 1997, responsável pela maior movimentação de contêineres com cerca de 98% da carga containerizada que passa pelo porto. Dentre os diferenciais no atendimento, destacam-se as operações 24 horas por dia durante os 7 dias da semana e canais de atendimento *online* exclusivos para os clientes. (TECON, 2014).

O terminal conta com uma infra-estrutura física que compreende:

- Cais - 900 metros de acostagem + *Dolphin*<sup>2</sup> (37 m)
- Calado de 12,8 metros (42 pés)
- Área total - 735.000 m<sup>2</sup>
- Área pavimentada - 340.000 m<sup>2</sup>
- Capacidade estática - 25.000 TEUs
- 2800 tomadas para contêineres refrigerados
- Armazém para cargas especiais e gerais - área total de 18.000 m<sup>2</sup>
- 10 *in/out gates* para caminhões
- Acesso interno à malha ferroviária
- Prédio administrativo

Para o atendimento dos navios no cais, são disponibilizados nove guindastes STS Super *Post-Panamax*<sup>3</sup> e dois guindastes móveis com capacidade para 100 toneladas, quanto aos equipamentos para movimentação de pátio conta com 56 tratores de pátio e cerca de 50 empilhadeiras de diferentes tipos. A Figura 6 apresenta a vista aérea do terminal de contêineres.

---

<sup>2</sup>Estrutura portuária situada em local de maior profundidade com dimensões capazes de receber embarcações.

<sup>3</sup>Portêiner composto por uma lança em balanço equipada com um trilho que guia o volume, utilizada no embarque e desembarque de contêineres em um navio.

Figura 6: TECON Rio Grande S/A



Fonte: (SUPRG, 2019)

Os principais serviços oferecidos são a operação dos navios e a operação de pátio, mais especificamente armazenagem alfandegada de contêineres e mercadorias, armazenagem de contêineres vazios, fornecimento de energia e monitoramento em contêineres refrigerados, desova e estufagem de contêineres de importação ou exportação, inspeções e pesagem de contêineres ou mercadorias, além de limpeza e lavagem simples ou química em contêineres vazios.

Nas próximas seções serão apresentados o funcionamento geral de um terminal de contêineres, bem como o funcionamento geral de uma operação portuária nesses terminais.

## 2.3 TERMINAIS DE CONTÊINERES E OPERAÇÃO PORTUÁRIA

Um terminal de contêineres serve como uma ligação entre o oceano e o transporte rodo-ferroviário e possui duas principais funções: receber contêineres de exportação e descarregar contêineres de importação, podendo ainda oferecer serviços de armazenamento e afins.

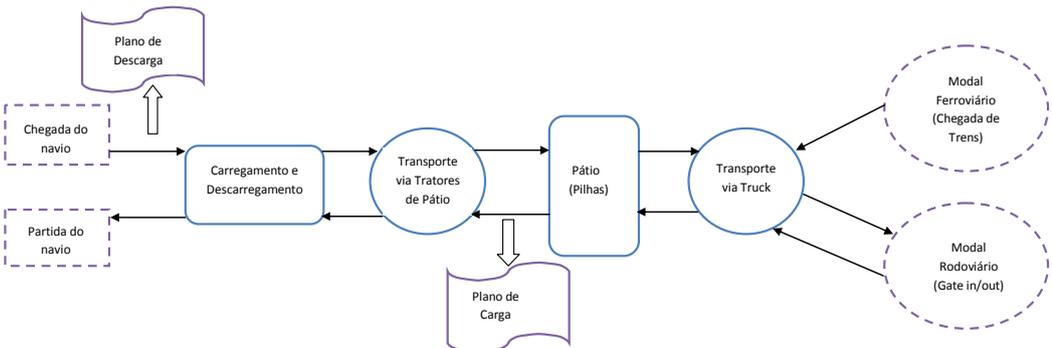
Dessa forma, sua eficiência é fundamental para possibilitar o incremento do fluxo de contêineres em uma cadeia global de suprimentos (MURTY et al., 2005).

Os terminais de contêineres mostram um alto nível de integração dos diferentes sistemas de informação e aplicações de engenharia de controle em um sistema global de informações. Esses terminais incluem serviços de tráfego marítimo, operações de pátio e planejamento de carga e descarga, gestão financeira e administrativa, gestão e controle das atividades de manuseio, equipamentos e manutenção (BIELLI; BOULMAKOUL; RIDA, 2006). Para Stahlbock e VoB (2008), os fatores chave para um terminal de contêineres são a eficiência no empilhamento e o transporte de um grande número de TEU's. A alta produtividade nesses quesitos é vital para um operador de terminal que deseja competir com outros terminais, nesse sentido, a gestão de um terminal portuário torna-se um processo complexo envolvendo um número elevado de decisões.

A operação portuária pode ser definida como o conjunto de todas as operações para realizar a passagem da mercadoria, desde o transporte marítimo até o transporte terrestre e vice-versa. O objetivo da operação portuária é sempre buscar a maior eficiência e eficácia. Em outras palavras, isso quer dizer minimizar os custos de transporte e armazenagem, e aumentar o fluxo (movimentação de cargas) dado um determinado período de tempo (OLIVEIRA, 2011).

O planejamento e a logística de operações estão intimamente ligados à competitividade de um terminal portuário. Basicamente uma operação portuária acontece a partir do plano de estiva, dividido em plano de descarga, para o caso de contêineres de importação e o plano de carga para os contêineres de exportação e/ou transbordo. As demais decisões, que envolvem os problemas mencionados na seção anterior são decorrentes do plano de estiva e estão intimamente interligadas. Na Figura (7) está representado o esquema geral de uma operação portuária que muito se assemelha a operação do terminal estudado.

Figura 7: Esquema geral de uma operação portuária



Fonte: Elaborado pelo autor

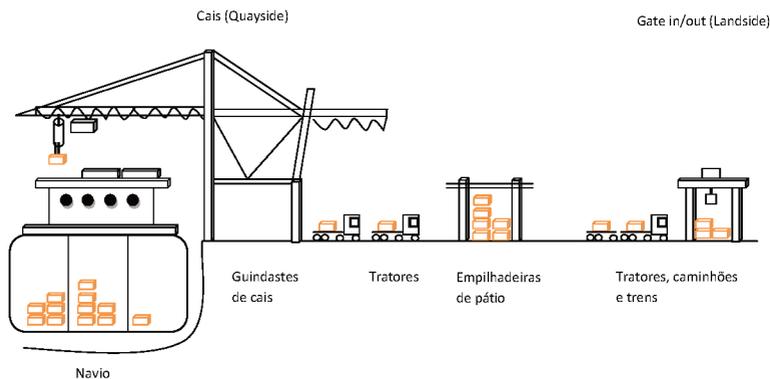
As operações diárias em um terminal de contêineres envolvem uma série de decisões inter-relacionadas, cujo objetivo final dessas decisões é minimizar o tempo de atracação de navios, os recursos necessários para o tratamento da carga de trabalho, o tempo de espera de caminhões e de clientes, bem como o congestionamento nas estradas, blocos de armazenamento e docas dentro do terminal e ainda, para fazer o melhor uso do espaço de armazenamento (STAHLBOCK; VOB, 2008), (CARLO; VIS; ROODBERGEN, 2014b) e (BIELLI; BOULMAKOUL; RIDA, 2006). Tais decisões podem ser divididas em cinco principais problemas:

1. Alocação de Berços: Programação da atribuição de berços no cais do porto para permitir o atendimento de navios;
2. Plano de Estiva: Programação de carregamento e descarregamento de contêineres entre um navio e um berço do porto, por meio de um ou mais guindastes portuários, observando um certo número de portos a serem percorridos;
3. Atribuição e *Scheduling* de Guindastes Portuários: Atribuição de guindastes portuários de modo a se elaborar uma programação de operação para cada seção do navio;
4. Transporte no Cais: Transporte dos contêineres de cada berço de atracação para o pátio do porto por meio de máquinas apropriadas como por exemplo *Automated Guided Vehicles* (AGV) e caminhões;
5. Transporte no Pátio: Carregamento e descarregamento de contêineres no pátio do porto (embarque nos navios ou o despacho para caminhões e trens) por meio de equipamentos como guindastes de pátio, pórticos, e alguns tipos específicos de empilhadeiras.

A Figura (8) mostra o esquema lateral de funcionamento de um terminal de contêineres, onde encontram-se alguns dos problemas citados.

Nesse contexto, cada vez mais pesquisadores despertam interesse na área portuária e alguns problemas pontuais já possuem métodos eficientes de solução, servindo como ferramentas de apoio à gestão. Estudos de manuseio e estocagem de contêineres que consideram os recursos existentes de operação, bem como operações de pátio de armazenamento ganham importância pois estão relacionados à redução de custos de transporte e tempo de operação, Zhang et al. (2003) faz

Figura 8: Esquema lateral de um terminal de contêineres



Fonte: Adaptada de (CARLO; VIS; ROODBERGEN, 2014b)

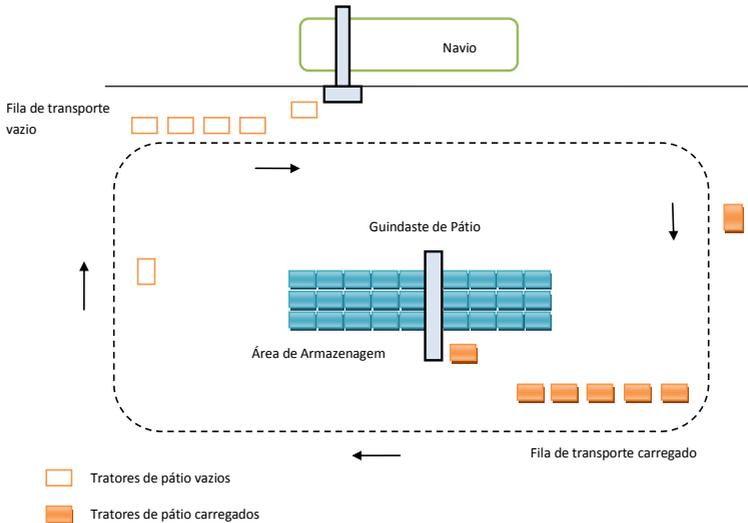
esse estudo através da simulação obtendo um equilíbrio nas operações, evitando possíveis gargalos nas operações de um terminal. Carlo, Vis e Roodbergen (2014a) faz uma revisão da literatura propondo uma classificação para as operações de pátios de armazenamento, identificando novos caminhos para a pesquisa na área portuária.

A Teoria de Filas e a Simulação são frequentemente os métodos empregados para análise de capacidade, eficiência ou desempenho de sistemas portuários, inclusive para determinação do número ótimo de berços em um terminal como pode ser visto no trabalho de El-Naggar (2010), cujos modelos matemático e econômico são aplicados no Porto de Alexandria no Egito. Em geral, análise de capacidade é de suma importância pois o desempenho operacional de um porto ou terminal portuário é essencial para assegurar a competitividade dos produtos nacionais no mercado externo, tendo forte impacto econômico em todo o sistema. Por essa razão grande parte dos estudos na área portuária são voltados à esse tipo de análise como pode ser visto nos trabalhos de Falcao e Correia (2012), Bichou (2013), Yun e Choi (1999), Lemke (2008) e Fernandes (2001).

## 2.4 A OPERAÇÃO PORTUÁRIA DO TERMINAL ESTUDADO

Baseado na entrevista realizada com o gerente de operações do TECON Rio Grande, na Figura (9) é apresentado o funcionamento unilateral de um processo de descarregamento de navios. Vale ressaltar que o processo para o carregamento funciona inversamente no mesmo esquema. Os contêineres são carregados por veículos próprios do terminal chamados “tratores de pátio” e são empilhados ou desempilhados dos lotes por meio das máquinas empilhadeiras. Para o embarque/desembarque do navio são utilizados o que chamam “guindastes portuários”.

Figura 9: Processo de descarregamento/carregamento de navios

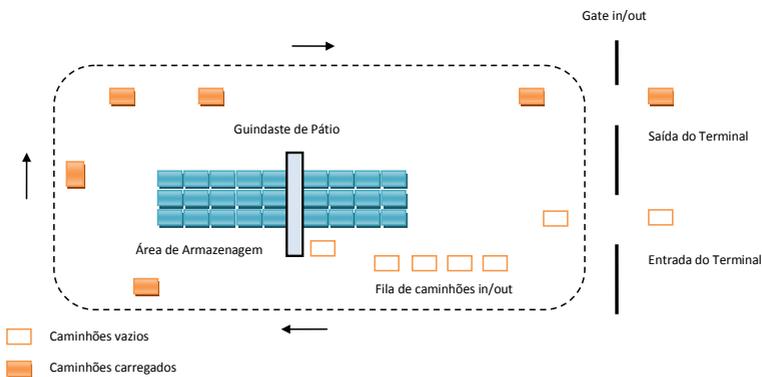


Fonte: Adaptada de (BIELLI; BOULMAKOUL; RIDA, 2006)

Na Figura (10) tem-se a movimentação a partir do *gate in/out*, ou seja, é apresentado o funcionamento do setor responsável pelo recebimento e o despacho de contêineres por via rodo-ferroviária. Nessa movimentação são utilizadas as mesmas máquinas empilhadeiras, disponíveis no pátio de contêineres, para manuseio dos contêineres. O transporte até o *gate in* ou a partir do *gate out* é feito por veículos particulares, normalmente transportadoras ou trabalhadores autônomos.

Uma vez que o contêiner encontra-se dentro do terminal, o seu manuseio é feito exclusivamente pelos tratores de pátio.

Figura 10: Movimentação do Gate in/out



Fonte: Adaptada de (BIELLI; BOULMAKOUL; RIDA, 2006)

A partir das operações descritas, criou-se um modelo de redes de filas contemplando a movimentação de contêineres no terminal, possibilitando uma análise detalhada das medidas de desempenho da rede. Essa movimentação será considerada desde o momento em que o navio que transporta os contêineres entra na fila até o momento em que o contêiner deixa o terminal. No capítulo seguinte serão apresentados os conceitos principais que orientaram a pesquisa e serviram de suporte à modelagem do problema.



## 3 FUNDAMENTAÇÃO TEÓRICA

### 3.1 INTRODUÇÃO

Segundo os objetivos desta tese, descritos no Capítulo 1 é necessário um levantamento bibliográfico a respeito do Problema de Alocação de *Buffers*, bem como de uma apresentação teórica dos conceitos de Teoria de Filas e Redes de Filas, importantes para o desenvolvimento da pesquisa.

#### 3.1.1 Processos estocásticos

Processos estocásticos representam sistemas em que seu estado muda ao longo do tempo. Essas mudanças não são completamente previsíveis, porém estão associadas a distribuições de probabilidade. Diversos fenômenos reais podem ser modelados através de processos estocásticos, por exemplo, um processo estocástico  $X_1, X_2, \dots$  pode representar as quantidades de carros que passam por um determinado ponto de uma rodovia, a evolução dos níveis de estoque semanais de uma empresa, evolução do número de desempregados num determinado país, dentre outros (HILLIER; LIEBERMAN, 2006).

Um processo estocástico é definido como uma coleção de variáveis aleatórias ( $X(t)$ ) indexadas por um parâmetro  $t$  pertencente ao conjunto  $T$ . Normalmente  $T$  é o conjunto dos inteiros não-negativos e  $X(t)$  representa uma característica mensurável de interesse no tempo  $t$  (HILLIER; LIEBERMAN, 2006).

Em outras palavras, um processo estocástico é a descrição de um fenômeno aleatório que varia com o tempo. Os valores assumidos por  $X(t)$  representam o *estado* do sistema e o conjunto de todos os estados possíveis é denominado *espaço de estados*.

Os processos estocásticos podem ser classificados em relação ao estado ou ao parâmetro (tempo) como discreto ou contínuo e as combinações de estado e tempo recebem a seguinte classificação (MARINS, 2011):

- Estado Discreto x Tempo Discreto = Cadeia estocástica de tempo discreto;
- Estado Discreto x Tempo Contínuo = Cadeia estocástica de tempo contínuo;

- Estado Contínuo x Tempo Discreto = Processo de estado contínuo e tempo discreto;
- Estado Contínuo x Tempo Contínuo = Processo de estado e tempo contínuo.

Para um valor fixado de  $t$ ,  $X(t)$  é uma variável aleatória que descreve o estado do processo no tempo  $t$ . Seja uma coleção finita  $t_1, t_2, \dots, t_n$  de tempos, então  $X_{t_1}, X_{t_2}, \dots, X_{t_n}$  é um conjunto de  $n$  variáveis aleatórias com distribuição conjunta definida como  $P[X_{t_1} \leq x_1, X_{t_2} \leq x_2, \dots, X_{t_n} \leq x_n]$  (CLARKE; DISNEY, 1979).

A probabilidade do processo estocástico estar no tempo  $t$  no estado  $e_t$  é denominada **probabilidade de estado** ( $p_t$ ) e é definida por:

$$p_t = P[X_t = e_t]. \quad (3.1)$$

O vetor  $p = [p_1, p_2, \dots, p_n, \dots]$  é chamado **vetor de probabilidade de estado**.

Basicamente, analisar um processo estocástico consiste em determinar a distribuição conjunta de cada conjunto de variáveis aleatórias e utilizá-las para prever o comportamento do processo no futuro, dado certo comportamento no passado (CLARKE; DISNEY, 1979).

Processos estocásticos descrevem o comportamento de um sistema operando ao longo de algum período e apresentam a seguinte estrutura, resumidamente:

O estado atual do sistema pode estar em qualquer uma das  $M + 1$  categorias mutuamente exclusivas denominadas estados. Esses estados são identificados como  $0, 1, \dots, M$ . A variável aleatória  $X$ , representa o estado do sistema no instante  $t$ , assim seus possíveis valores são unicamente  $0, 1, \dots, M$ . Observa-se o sistema pontos determinados do tempo,  $t = 0, 1, 2, \dots$  e dessa forma, o processo estocástico  $X_t = \{X_0, X_1, X_2, \dots\}$  fornece uma representação matemática de como o estado do sistema evolui ao longo do tempo (HILLIER; LIEBERMAN, 2006).

### 3.1.2 Cadeias de Markov

Para obter resultados analíticos é necessário levantar hipóteses referentes à distribuição conjunta de  $X_0, X_1, \dots$ , uma dessas hipóteses é de que o processo estocástico é uma cadeia de Markov, ou seja, o estado presente  $X_t = i$ , é independente dos eventos passados e depende

apenas do estado atual (HILLIER; LIEBERMAN, 2006).

Sendo assim, um processo estocástico  $\{X_t\}$  é dito ter a propriedade markoviana se  $P\{X_{t+1} = j \mid X_0 = k_0, X_1 = k_1, \dots, X_{t-1} = k_{t-1}, X_t = i\} = P\{X_{t+1} = j \mid X_t = i\}$ , para  $t = 0, 1, \dots$  e toda sequência  $i, j, k_0, k_1, \dots, k_{t-1}$ .

As probabilidades condicionais  $P\{X_{t+1} = j \mid X_t = i\}$  para uma cadeia de Markov são chamadas probabilidades de transição (uma etapa), e se, para cada  $i$  e  $j$ ,

$$P\{X_{t+1} = j \mid X_t = i\} = P\{X_1 = j \mid X_0 = i\}, \forall t = 1, 2, \dots \quad (3.2)$$

então as probabilidades de transição são ditas estacionárias, o que implica que essas não mudam ao longo do tempo.

As probabilidades de transição estacionária para  $n$  etapas  $P\{X_{t+n} = j \mid X_t = i\} = P\{X_n = j \mid X_0 = i\}$ , para todo  $t = 1, 2, \dots$  também implicam na estabilidade ao longo do tempo.

Para efeito de notação simplificada, as probabilidades de transição estacionárias serão definidas como:

$$\begin{aligned} p_{ij} &= P\{X_{t+1} = j \mid X_t = i\} \\ p_{ij}^{(n)} &= P\{X_{t+n} = j \mid X_t = i\} = P\{X_n = j \mid X_0 = i\} \end{aligned} \quad (3.3)$$

Como as  $p_{ij}^{(n)}$  são probabilidades condicionais, devem satisfazer as seguintes propriedades:

$$p_{ij}^{(n)} \geq 0, \forall i, j; n = 0, 1, 2, \dots \quad (3.4)$$

e

$$\sum_{j=0}^M p_{ij}^{(n)} = 1, \forall i; n = 0, 1, 2, \dots \quad (3.5)$$

### 3.1.2.1 Equações de Chapman-Kolmogorov

Para calcular as probabilidades de transição em  $n$  etapas, é fornecido um método por meio das equações de Chapman-Kolmogorov (HILLIER; LIEBERMAN, 2006):

$$\begin{aligned}
 p_{ij}^{(n)} &= \sum_{k=0}^M p_{ik}^{(m)} p_{kj}^{(n-m)}, & \text{para todo } k &= 0, 1, \dots, M, \\
 & & j &= 0, 1, \dots, M, \\
 & \text{e qualquer } m &= 1, 2, \dots, n-1, \\
 & & n &= m+1, m+2, \dots
 \end{aligned} \tag{3.6}$$

As equações resultantes de (3.6) dizem que ao ir do estado  $i$  para o estado  $j$  em  $n$  etapas, o processo se encontrará em algum estado  $k$  após exatamente  $m$  ( $m < n$ ) estados. Dessa forma,  $p_{ik}^{(m)} p_{kj}^{(n-m)}$  é a probabilidade condicional, partindo do estado  $i$  o processo vai para o estado  $k$  após  $m$  etapas e finalmente para o estado  $j$  em  $n - m$  etapas.

### 3.1.2.2 Probabilidade de estado estável

Existe uma probabilidade limitada de que o sistema se encontrará em cada estado  $j$  após um grande número de transições e essa probabilidade, por sua vez é independente do estado inicial. Segundo certas condições relativamente genéricas, essa propriedade de comportamento duradoura das cadeias de Markov de estados finitos tem validade e podem ser sintetizadas a seguir.

Para qualquer cadeia de Markov ergódica irredutível, o  $\lim_{n \rightarrow \infty} p_{ij}^{(n)}$  existe e é independente de  $i$  (SHAMBLIN; STEVENS, 1979). Além disso,

$$\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \pi_j > 0, \tag{3.7}$$

em que os  $\pi_j$  satisfazem única e exclusivamente as seguintes equações de estado estável

$$\begin{aligned}
 \pi_j &= \sum_{i=0}^M \pi_i p_{ij}, & \text{para } j &= 0, 1, \dots, M \\
 \sum_{j=0}^M \pi_j &= 1.
 \end{aligned} \tag{3.8}$$

Os  $\pi_j$  são ditas probabilidade de estado estável ou probabilidades estacionárias, o que significa que a probabilidade de encontrar o sistema em certo estado  $j$ , após um grande número de transições, tende a ser o valor  $\pi_j$ , independente da distribuição probabilística do estado inicial.

### 3.1.3 Cadeias de Markov de tempo contínuo

Sejam  $0, 1, \dots, M$  os possíveis estados do sistema, iniciando no instante 0 permitindo que o parâmetro  $t'$  execute continuamente para  $t' \geq 0$ , então a variável aleatória  $X(t')$  é o estado do sistema no instante  $t'$ . Ao longo de cada intervalo  $0 \leq t' \leq t_1$ ,  $X(t')$  assumirá um dos possíveis valores ( $M + 1$ ) e assim sucessivamente para os próximos intervalos  $t_1 \leq t' < t_2, \dots$ . Os pontos de transição  $t_1, t_2, \dots$  são pontos aleatórios e não necessariamente inteiros (CLARKE; DISNEY, 1979).

Consideremos três pontos no tempo, quais sejam:

- $t' = r, r \geq 0$ : tempo passado,
- $t' = s, s > r$ : tempo presente,
- $t' = r + t, t > 0$ :  $t$  unidades no futuro.

Observemos o sistema nos instantes  $t' = s$  e  $t' = r$ , a esses estados nomeamos  $X(s) = i$  e  $X(r) = x(r)$ . A distribuição probabilística do estado no instante  $t' = s + t$  é dada por

$$P\{X(s + t) = j \mid X(s) = i \text{ e } X(r) = x(r)\}, \text{ para } j = 0, 1, \dots, M \quad (3.9)$$

Obter essa probabilidade condicional é bastante difícil, no entanto, sua obtenção pode ser simplificada caso o processo estocástico possua a propriedade fundamental a seguir (HILLIER; LIEBERMAN, 2006).

Um processo estocástico de tempo contínuo  $\{X(t'); t' \geq 0\}$  possui a propriedade markoviana se  $P\{X(t+s) = j \mid X(s) = i \text{ e } X(r) = x(r)\} = P\{X(t+s) = j \mid X(s) = i\}$  para todo  $i, j = 0, 1, \dots, M$  e para todo  $r \geq 0, s > r$  e  $t > 0$  (WINSTON, 2003).

$P\{X(t+s) = j \mid X(s) = i\}$  é uma probabilidade de transição em que  $t$  não precisa ser um número inteiro. Se essas probabilidades forem independentes, de forma que  $P\{X(t+s) = j \mid X(s) = i\} = P\{X(t) = j \mid X(0) = i\}$  para todo  $s > 0$ , elas são denominadas probabilidades de transição estacionárias e denominamos como:

$$p_{ij}(t) = P\{X(t) = j \mid X(0) = i\} \quad (3.10)$$

onde  $p_{ij}(t)$  é a função de probabilidade de transição de tempo contínuo (WINSTON, 2003).

### 3.1.3.1 Probabilidade de estado estável

A função de probabilidade de transição de tempo contínuo também satisfaz as equações de Chapman-Kolmogorov (HILLIER; LIEBERMAN, 2006). Para quaisquer estados  $i$  e  $j$  e números não-negativos  $t$  e  $s$  ( $0 \leq s \leq t$ ),

$$p_{ij}(t) = \sum_{k=1}^M p_{ik}(s)p_{kj}(t-s). \quad (3.11)$$

Um par de estados  $i$  e  $j$  comunicam-se entre si se houver tempos  $t_1$  e  $t_2$  tais que  $p_{ij}(t_1) > 0$  e  $p_{ji}(t_2) > 0$ . Além disso

$$\lim_{t \rightarrow \infty} p_{ij}(t) = \pi_j > 0, \quad (3.12)$$

sempre existe e é independente do estado inicial de uma cadeia de Markov, para  $j = 0, 1, \dots, M$ . Essas probabilidades são conhecidas como probabilidades do estado estável ou probabilidades estacionárias de uma cadeia de Markov (SHAMBLIN; STEVENS, 1979).

Os  $\pi_j$  satisfazem as equações

$$\begin{aligned} \pi_j &= \sum_{i=0}^M \pi_i p_{ij}(t), \quad \text{para } j = 0, 1, \dots, M \forall t \geq 0 \\ \sum_{j=0}^M \pi_j &= 1. \end{aligned} \quad (3.13)$$

Na próxima seção serão apresentados os principais conceitos acerca da Teoria de Filas, bem como as medidas de desempenho que podem ser calculadas em diferentes sistemas.

## 3.2 TEORIA DAS FILAS

### 3.2.1 Introdução

O congestionamento de clientes é o principal sintoma de que o funcionamento de um sistema é deficiente e que pode assim, estar dimensionado inadequadamente. A Teoria das Filas é um dos tópicos da Pesquisa Operacional que trata de congestionamento de sistemas

e sua característica fundamental é a presença de *clientes* solicitando *serviços* de alguma maneira (TAHA, 2008).

O objetivo central da Teoria das Filas é o desenvolvimento de modelos matemáticos que permitam prever o comportamento de sistemas de prestação de serviços. Essa abordagem matemática teve início no século XX com os estudos de A. K. Erlang em uma companhia telefônica. Porém somente após a Segunda Guerra Mundial foi que suas aplicações foram estendidas a diferentes problemas de fila, como por exemplo, planejamento e controle da produção, dimensionamento de sistemas de armazenamento, sistemas de transportes, sistemas de tráfego (rodo-porto-aéreo-ferroviário), manutenção de máquinas dentre outros (PRADO, 2009).

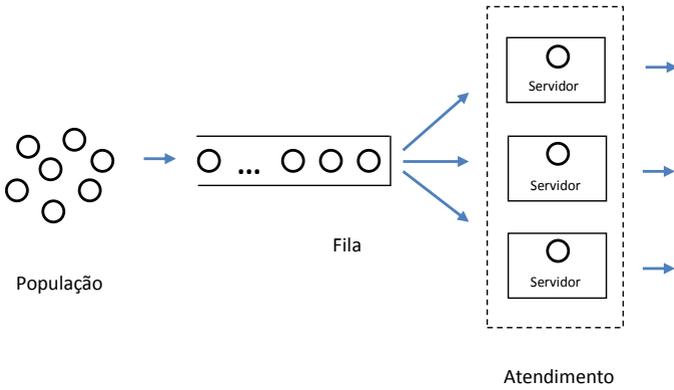
O maior interesse, do ponto de vista prático, em estudar este tipo de sistema se dá em virtude da possibilidade de se inserir modificações, que auxiliam de alguma maneira para aprimorar seu rendimento. Por exemplo, fornecer capacidade excessiva gera ociosidade, por outro lado, um atendimento mais enxuto gera insatisfação e consequente perda de clientes e de produção. Essas condições expressam uma relação muito forte e por isso é interessante ter disponível um modelo que possa prever o comportamento do sistema de acordo com as condições que lhe são impostas para, dessa forma fornecer subsídios na escolha da modificação mais adequada ao sistema. Outra implicação no estudo de modelos de filas é o problema econômico que está embutido, pois em qualquer fila existem dois custos envolvidos: o custo da fila e o custo do serviço (HILLIER; LIEBERMAN, 2006).

Essas irregularidades nas condições do sistema é que determinam seu funcionamento e que são expressas em termos probabilísticos no estudo de Teoria das Filas.

### 3.2.2 Características de uma fila

Um sistema de fila é formado basicamente por uma *população* de clientes ou produtos que necessitam de um determinado atendimento, chegam ao longo do tempo e entram na *fila*. Num dado momento, um membro da fila passa ao atendimento por meio de alguma regra pré-estabelecida conhecida por *disciplina da fila* (PRADO, 2009). Após receber o devido atendimento, o cliente ou produto deixa o sistema de filas. Esse processo está representado na Figura (11).

Figura 11: Estrutura de um Sistema de Filas



Fonte: Adaptada de (PRADO, 2009)

A seguir serão descritos os principais elementos de uma fila (MARINS, 2011), (HILLIER; LIEBERMAN, 2006), (KLEINROCK, 1975), (COOPER, 1981).

a) População

A população é composta pelos clientes que podem solicitar serviço em determinados períodos de tempo; pode ser de fonte finita ou infinita. Normalmente assume-se que a população é infinita ainda que seu tamanho real seja um número finito, pois os cálculos são mais simples para o caso infinito. Essa premissa deve valer como hipótese implícita para todo modelo de filas que não afirme o contrário. No caso finito o número de clientes no sistema de filas interfere no número de possíveis clientes fora do sistema a qualquer momento, por isso é mais difícil analiticamente.

b) Processo de chegada

A forma como os clientes ingressam no sistema determina o processo de chegada que pode ser determinístico ou estocástico. O modelo é dito determinístico quando o número de chegadas e o instante de tempo de ocorrência são conhecidos exatamente. É estocástico quando as chegadas no sistema ocorrem de forma aleatória e, dessa forma, obedecendo uma distribuição de probabilidade. Nesse contexto, o padrão estatístico pelos quais os

clientes chegam ao longo do tempo também deve ser especificado. É comum partir da hipótese de que eles chegam de acordo com um processo de Poisson. Esse caso é aquele na qual as chegadas ao sistema de filas ocorrem aleatoriamente, porém, com certa taxa média fixa, independentemente de quantos clientes já se encontrarem lá. A hipótese equivalente é que a distribuição probabilística do tempo entre as chegadas consecutivas é descrita por uma distribuição exponencial (HILLIER; LIEBERMAN, 2006).

c) Fila

A fila é o local onde os clientes aguardam antes de receberem atendimento e é caracterizada pelo número máximo de clientes que ela pode conter. As filas podem ser infinitas ou finitas. A hipótese de uma fila infinita é tida como padrão para a maioria dos modelos de filas, ainda que exista um limite superior finito (relativamente grande) para o número de clientes permitido, pois a análise se torna complicada com um limite superior destes. Entretanto, para sistemas de filas em que esse limite superior é suficientemente pequeno, ao ponto que possa efetivamente ser atingido com uma dada frequência, então é necessário supor uma fila finita.

d) Disciplina da Fila

A disciplina da fila diz respeito à ordem na qual os clientes são selecionados para atendimento. Normalmente obedecem o critério FIFO (*First-in-First-Out*), ou seja, o primeiro cliente que entra na fila é o primeiro a ser atendido. Outra disciplina de atendimento é o tipo LIFO (*Last-in-First-Out*) em que o último que entra é o primeiro a ser atendido, comumente utilizado em sistemas de estoques. A disciplina da fila ainda pode ser do tipo aleatória ou com estabelecimento prévio de determinadas prioridades.

e) Número de atendentes/servidores

Representa a quantidade de locais para a realização do atendimento. Podem ser instalações únicas com apenas um atendente, em paralelo com vários atendentes ou ainda em série. Um modelo de filas deve especificar a disposição das instalações e o número de atendentes (canais paralelos) em cada uma delas.

f) Processo de Atendimento

A principal característica do atendimento é a duração do mesmo, chamado também de tempo de permanência, que é o tempo decorrido entre o início do atendimento até o seu término para um

cliente em uma instalação. O modelo de um sistema de filas deve especificar a distribuição probabilística de tempos de atendimento para cada atendente, embora seja comum supor a mesma distribuição para todos os atendentes (HILLIER; LIEBERMAN, 2006).

g) Capacidade do sistema

O sistema pode ter capacidade finita ou infinita. Em alguns modelos de filas existe uma limitação física para o número de clientes que o sistema comporta, nesse caso o sistema funciona normalmente até que sua capacidade máxima seja atingida.

### 3.2.3 Notação de Kendall

Segundo a notação de Kendall, introduzida por David Kendall em 1953 (HILLIER; LIEBERMAN, 2006), um modelo de filas pode ser descrito por uma sucessão de símbolos da seguinte forma:  $A/B/c/K/m/Z$ , onde:

- $A$  e  $B$ : correspondem à distribuição dos intervalos de chegada e aos tempos de serviço, respectivamente, e de acordo com a distribuição de probabilidade adequada, assumem os seguintes símbolos:
  - $M$ : distribuição Exponencial (Markoviana);
  - $G$ : distribuição Geral (qualquer distribuição arbitrária permitida);
  - $D$ : distribuição Degenerada (tempos constantes);
  - $E_k$ : distribuição de Erlang com  $k$  fases; e
  - $H_k$ : distribuição Hiperexponencial com  $k$  fases.
- $c$ : quantidade de servidores (atendentes);
- $K$ : capacidade máxima do sistema;
- $m$ : tamanho da população; e
- $Z$ : disciplina da fila.

Para os casos em que a população é infinita, a capacidade do sistema é ilimitada e a disciplina adotada pela fila é FIFO, os três últimos símbolos podem ser omitidos.

### 3.2.4 Medidas de desempenho

A forma como um sistema se comporta pode ser definida como o seu desempenho. Para determinar o desempenho de um sistema é necessário definir suas características comportamentais de execução. Em Teoria das Filas é possível estimar as medidas de desempenho de um sistema e, dessa forma, dimensionar o mesmo para que cumpra a demanda de clientes, bem como evitar gargalos e desperdícios.

Existem algumas variáveis de decisão importantes para o desempenho do sistema, como o número de clientes na fila, o número de clientes no sistema, tempo de permanência na fila e no sistema, e ainda a ociosidade dos atendentes. Como essas variáveis, em geral, são aleatórias por depender do comportamento das chegadas e dos atendimentos, para medir o desempenho de um sistema de filas é necessário introduzir medidas de eficiências em função dos valores médios dessas variáveis aleatórias (HILLIER; LIEBERMAN, 2006).

As principais medidas de desempenho que podem ser determinadas pela Teoria das Filas, considerando o sistema em regime estacionário (estado de equilíbrio), ou seja, após um período longo de funcionamento, são (HILLIER; LIEBERMAN, 2006):

- $\lambda$ : taxa média de entrada de usuários no sistema;
- $\mu$ : taxa média de serviço;
- $\frac{1}{\lambda}$ : intervalo médio entre chegadas;
- $\frac{1}{\mu}$ : tempo médio de atendimento;
- $L_q$ : número médio de usuários na fila;
- $L$ : número médio de usuários no sistema;
- $W_q$ : tempo médio de espera na fila;
- $W$ : tempo médio de espera no sistema;
- $c$ : número de atendentes/servidores;
- $\rho = \frac{\lambda}{c \cdot \mu}$ : taxa de utilização do sistema;
- $P_n$ : probabilidade de que o número de clientes no sistema seja  $n$ .

### 3.2.5 Relação entre $L$ , $W$ , $L_q$ e $W_q$

John D. C. Little desenvolveu, no início da década de 60, uma das mais importantes relações em Teoria de Filas, ele relacionou o tamanho médio do sistema ao tempo médio de espera do cliente em estado de equilíbrio obtendo as seguintes relações (HILLIER; LIEBERMAN, 2006):

$$L = \lambda W \quad (3.14)$$

$$L_q = \lambda W_q \quad (3.15)$$

Considerando constante o tempo médio de atendimento, denotado por  $1/\mu$ , segue que:

$$W = W_q + \frac{1}{\mu}. \quad (3.16)$$

Essas relações são de extrema importância, pois permitem determinar imediatamente as quatro quantidades fundamentais ( $L$ ,  $L_q$ ,  $W$ ,  $W_q$ ) encontrando analiticamente apenas uma delas.

Outro resultado que pode ser extraído das fórmulas de Little e das relações entre  $W$  e  $W_q$  é o seguinte:

$$L - L_q = \lambda(W - W_q) = \lambda \left( \frac{1}{\mu} \right). \quad (3.17)$$

## 3.3 REDES DE FILAS

### 3.3.1 Introdução

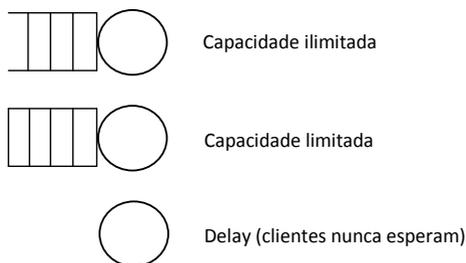
Frequentemente em Pesquisa Operacional, alguns dos sistemas são efetivamente *redes de filas*, ou seja, redes que contém instalações de atendimento onde os clientes recebem o serviço em uma ou mais instalações, como por exemplo, produtos em manufatura que devem ser processados por meio de uma sequência de grupos de máquinas (instalações de atendimento). Dessa forma, se faz necessário o estudo de toda a rede para obter informações, como o tempo de espera previsto total, o número de clientes esperados no sistema e as demais medidas de desempenho já vistas na Teoria das Filas (HILLIER; LIEBERMAN, 2006).

O estudo de redes de filas teve início com os trabalhos de Erlang em 1917 e Engset no ano seguinte, ambos na área de telefonia. No entanto, o formalismo de redes de filas foi introduzido nos anos 50 por Jackson, obtendo um enorme avanço em meados dos anos 70 quando surgiram propostas de soluções em forma de produto. Recentemente, as mesmas ideias aparecem em modelos de sistemas de computadores ou sistemas de transmissão de dados, sistemas de trânsito, sistemas de comando e controle dentre outros (LARSON; ODoni, 2007). Em virtude da importância de redes de filas, pesquisas nessa área vêm sendo cada vez mais pertinentes.

### 3.3.2 Definição e características

Um sistema que possui duas ou mais filas interligando caminhos que os clientes podem percorrer entre diferentes instalações, é denominado *redes de filas de espera* ou simplesmente *redes de filas*. É importante considerar os diferentes tipos de filas (Fig. 12) que um sistema pode assumir, pois eles determinam características importantes para uma rede de filas.

Figura 12: Tipos de filas

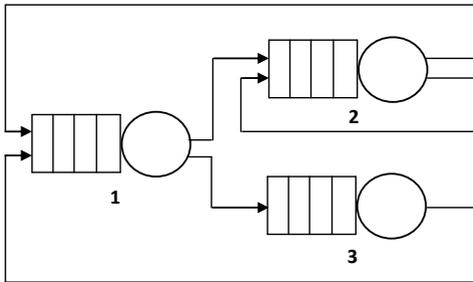


Fonte: Adaptada de (CHANIN, 2005)

De acordo com critérios ortogonais, as redes de filas podem ser classificadas da seguinte forma (HILLIER; LIEBERMAN, 2006):

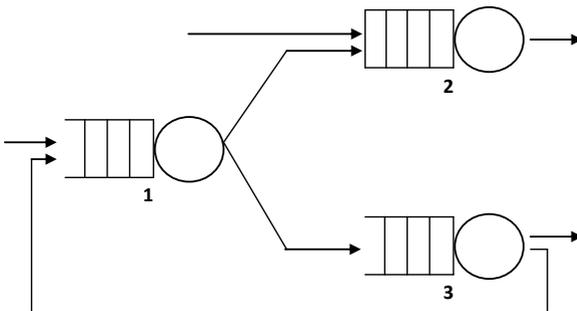
- a) permanência de clientes na rede: quando o número de clientes na rede é constante, ou seja, não há entrada ou saída de clientes, chama-se *rede de filas fechada*; caso contrário trata-se de uma *rede de filas aberta*, conforme Figuras (13) e (14), respectivamente;

Figura 13: Exemplo de rede fechada



Fonte: Adaptada de (CHANIN, 2005)

Figura 14: Exemplo de rede aberta



Adaptada de (CHANIN, 2005)

Fonte:

- b) números de servidores por instalação: quando possui apenas um servidor é dita *monoservidora* e quando possui mais de um servidor em cada instalação é chamada *multiservidora*;
- c) tipo de clientes: existe a rede *monoclasse* em que todos os clientes são iguais e recebem os mesmos serviços e a rede *multiclasse* onde os clientes recebem tratamento diferenciado em conformidade com a classe que pertencem.

Dentre os diferentes resultados já obtidos para redes de filas, um deles merece atenção especial por ser fundamental a toda discussão acerca do estudo de redes. Esse resultado é a propriedade de equivalência para o processo de entrada e processo de saída de clientes para certos sistemas de filas (HILLIER; LIEBERMAN, 2006).

**Propriedade de equivalência:** Suponha que uma instalação de atendimento com  $c$  atendentes e uma fila infinita tenha entrada de Poisson com parâmetro  $\lambda$  e a mesma distribuição exponencial de tempo de atendimento com parâmetro  $\mu$  para cada atendente (modelo  $M/M/c$ ), em que  $c\mu > \lambda$ . Então a saída de estado estável dessa instalação de atendimento também é um processo de Poisson com parâmetro  $\lambda$ .

A principal consequência desse resultado é que se os clientes precisam ir para a próxima instalação de atendimento, essa segunda instalação também terá uma entrada de Poisson. Esse processo se verifica a todas as passagens entre as instalações desde que os processos de atendimento sigam uma distribuição exponencial (HILLIER; LIEBERMAN, 2006).

Suponha um sistema de filas infinitas em série cujo processo de chegada acontece de acordo com o Processo de Poisson com parâmetro  $\lambda$ . Cada instalação  $i$  ( $i = 1, 2, \dots, m$ ) possui uma distribuição exponencial de tempos de atendimento com parâmetro  $\mu_i$  para seus  $c_i$  atendentes e  $c_i\mu_i > \lambda$ . Sob condições de estado estável, decorre da propriedade de equivalência que cada instalação de atendimento terá uma entrada de Poisson com parâmetro  $\lambda$ . Dessa forma, o modelo  $M/M/c$  pode ser utilizado para analisar cada instalação de atendimento independente dos demais.

Essa é uma simplificação para a obtenção das medidas de desempenho para esse sistema, por exemplo, a probabilidade de ter  $n$  clientes em determinada instalação é dada por  $P_n$  e a probabilidade conjunta de  $n_1$  clientes na instalação 1,  $n_2$  clientes na instalação 2,  $\dots$ , é portanto, o produto das probabilidades individuais, podendo ser expressa por

$$P[(N_1, N_2, \dots, N_m) = (n_1, n_2, \dots, n_m)] = P_{n_1} P_{n_2} \dots P_{n_m}. \quad (3.18)$$

A expressão (3.18) é chamada **solução em forma de produto** e de forma similar podem ser calculadas as demais medidas de desempenho de um sistema desse tipo.

Para o caso de filas finitas a propriedade de equivalência não é válida, ou seja, não existe uma solução em forma de produto simples. Nesse caso, as instalações devem ser analisadas em conjunto e os resultados obtidos até então são limitados. Esse tipo de fila é importante na prática, pois em muitos casos existe uma limitação definida no comprimento da fila que antecede as instalações em redes de filas. Por exemplo, um sistema de linha de produção oferece uma pequena quantidade de espaço de armazenamento em *buffer* em frente as instalações (estações) de serviço (HILLIER; LIEBERMAN, 2006).

Existe outro tipo de rede que admite solução em forma de produto, chamada *Redes de Jackson*. Nessa rede as características são as mesmas de um sistema de filas infinitas em série, a diferença é que os clientes visitam as instalações em ordens diferentes e não necessariamente visitam todas as instalações. Em cada instalação, os clientes podem ser procedentes de fora do sistema como de outra instalação, de acordo com um processo de Poisson.

Cabe salientar que existem outros tipos de redes de filas em que as instalações de serviço podem ser analisadas de forma individual, sendo mais complicado encontrar uma solução. Assim, encontrar redes de filas com uma solução em forma de produto simples tem sido uma meta de pesquisa sobre redes de filas (HILLIER; LIEBERMAN, 2006).

## 3.4 O PROBLEMA DE ALOCAÇÃO DE *BUFFERS* (PAB)

### 3.4.1 Introdução

Sistemas de manufatura e de telecomunicações são sistemas facilmente representados por redes de filas com *buffer* finito e, por isso, são exemplos de maior interesse prático. O Problema de Alocação de *Buffers* (PAB) em redes de filas é um problema de otimização combinatória NP-*hard* e é formulado como um problema de programação matemática inteira mista (CRUZ; DUARTE; WOENSEL, 2008).

Basicamente, o PAB consiste em encontrar o tamanho total ótimo de *buffer* (área de espera) para ser dividido ou alocado nas diferentes áreas de um sistema de produção a fim de atingir um objetivo específico. Segundo Demir, Tunali e Eliiyi (2014), a principal razão de ter uma área de armazenamento (*buffer*) é permitir que estações de trabalho sequenciais possam operar de forma independente. Uma consequência disso é a redução do tempo ocioso devido a indisponibilidade de novas entradas e os bloqueios causados pela falta de espaço entre as estações. Menos tempo ocioso tende a aumentar a taxa média de produção.

Por outro lado, a alocação de *buffers* requer investimento de capital adicional, além de espaço físico disponível. Assim, pode tornar-se um recurso oneroso. Um *buffer* também aumenta o estoque em processo que gera alto custo. Se a área de armazenamento é muito grande, o custo do capital investido pode superar o ganho do aumento de produtividade. Ao contrário, se esse espaço for pequeno, as estações de trabalho serão subutilizadas e a demanda poderá não estar sendo cumprida (DEMIR; TUNALI; ELIYI, 2014).

De acordo com Cruz, Duarte e Woensel (2008), devido aos custos críticos para alocação de *buffers*, é crucial determinar otimamente essa alocação, a fim de garantir o máximo desempenho do sistema com o menor custo possível.

### 3.4.2 Definição

O PAB consiste em atribuir um determinado número fixo de áreas de espera  $N$  entre  $K - 1$  localizações intermediárias de um sistema de produção. É de interesse particular em gestão de operações em que a alocação de espaço pode representar a principal flexibilidade para a organização (LARSON; ODoni, 2007).

Há pelo menos três tipos básicos para o PAB que dependem da função objetivo escolhida, ou seja, de acordo com o interesse da organização que pode ser a maximização das taxas de atendimento (*throughput*), a minimização do *work-in-process* (WIP) ou a minimização do número total de *buffers* (PAPADAPOULOS et al., 2009). Para cada um deles tem-se as seguintes formulações:

- **PAB-1** maximização da taxa de atendimento

$$\text{Max } X(n) = \text{Max } X(N_2, \dots, N_K)$$

s.a.:

$$\sum_{j=2}^K N_j = N$$

$$N_j > 0 \quad \forall j = 2, \dots, K$$

onde  $n = (N_2, N_3, \dots, N_K)$  denota o vetor de tamanho dos *buffers*,  $N_j$ , com  $j = 2, 3, \dots, K$ , são números inteiros e  $X(n) = X(N_2, \dots, N_K)$  representa o rendimento de produção da  $K$ -estação em função do vetor  $n$ .

- **PAB-2** minimização do WIP

$$\text{Min } L(n) = \text{Min } L(N_2, \dots, N_k)$$

s.a.:

$$X(n) = X(N_2, \dots, N_k) \geq X_0$$

$$\sum_{j=2}^K N_j \leq N$$

$$N_j \geq 0 \quad \forall j = 2, \dots, K$$

onde  $L(n) = L(N_2, \dots, N_k)$  denota a média WIP em função de  $N_j$ ,  $j = 2, \dots, K$  que são números inteiros e  $X_0$  é o nível de rendimento específico.

- **PAB-3** minimização do número total de *buffers*

$$\text{Min } N = \sum_{j=2}^K N_j$$

s.a.:

$$X(n) = X(N_2, \dots, N_k) \geq X_0$$

$$N_j \geq 0 \quad \forall j = 2, \dots, K$$

onde  $N_j$ ,  $j = 2, \dots, K$  que são números inteiros e  $X_0$  é o nível de rendimento específico.

Diferentes abordagens para a função objetivo são encontradas na literatura, como por exemplo, a maximização do nível de serviço ao cliente e a minimização do tempo médio de espera por serviço (DEMIR; TUNALI; ELIYI, 2014). Os resultados para o PAB-1 podem ser utilizados para resolver o problema PAB-3, pois segundo Papadapoulos

et al. (2009) o problema de maximização das taxas de atendimento é uma função crescente do espaço total de *buffers*, o que configura um *trade-off* crítico entre a área de espera total alocada e o *throughput*.

O PAB é um problema difícil por diferentes razões, sendo as principais: a ausência de uma relação algébrica ou expressão de forma fechada para a taxa de transferência e tamanho do *buffer*, e a natureza de otimização combinatória inerente ao problema (HILLIER; LIEBERMAN, 2006). Por exemplo, para uma linha de produção com  $K$  máquinas e  $N$  o número total de *buffers*, tem-se o número total de possíveis configurações para o problema calculado como segue:

$$\binom{N + K - 2}{K - 2} = \frac{(N + 1)(N + 2) \dots (N + k - 2)}{(K - 2)!} \quad (3.19)$$

Dessa forma, o número total de soluções viáveis cresce exponencialmente quando  $N$  e  $K$  são muito grandes, conforme equação (3.19). Devido à natureza discreta da alocação de *buffer* e a indisponibilidade de expressões em forma fechada, a solução desses problemas requer abordagens numéricas, ainda que em situações com número relativamente pequeno para  $N$  e  $K$  (STAHLBOCK; VOB, 2008).

### 3.4.3 Abordagem multiobjetivo para o PAB

Cruz et al. (2012) propõem uma abordagem multiobjetivo para solucionar o PAB, considerando o compromisso (*trade-off*) entre a área de espera total, as taxas de serviço e a taxa de saída do sistema. O alto custo que representam as áreas de espera e os serviços efetuados, ocasionam a restrição no tamanho da área de espera alocada e na capacidade total do serviço. Da mesma forma, alcançar uma taxa de saída maior, que é afetada diretamente pelas áreas de espera e taxas de serviço, gera conflito entre esses objetivos.

Uma rede de filas é definida como um grafo direcionado  $G(N, A)$ , onde  $N$  é o conjunto finito de nós e  $A$  é o conjunto finito de arcos (CRUZ et al., 2012).

A formulação multiobjetivo que maximiza a taxa de atendimento ( $\theta$ ), enquanto minimiza a área de espera global ( $K$ ) e as taxas de serviço ( $\mu$ ), proposta para o PAB, é dada pelas equações (3.21), (3.22) e (3.22)

$$\text{Min } F(K, \mu) \quad (3.20)$$

s.a.

$$K_i \in \{1, 2, \dots\}, \forall i \in N \quad (3.21)$$

$$\mu_i > 0, \forall i \in N, \quad (3.22)$$

em que  $K_i$  e  $\mu_i$  são, a área de espera e a taxa de serviço, respectivamente, para a  $i$ -ésima fila, e representam as variáveis de decisão do modelo. A função objetivo  $F(K, \mu) \equiv (f_1(K), f_2(\mu), -f_3(K, \mu))^T$  representam:

- $f_1(K) = \sum_{\forall i \in N} K_i$ , área de espera global alocada;
- $f_2(\mu) = \sum_{\forall i \in N} \mu_i$ , taxa de serviço total;
- $f_3(K, \mu) = \theta(K, \mu)$ , taxa de atendimento.

A proposta dos autores acrescenta a função  $f_2$  associada às taxas de serviço ( $\mu_i$ ) em cada uma das filas, o que difere de outras abordagens que a consideram como um valor fixo. Nesse ponto surgem questionamentos sobre a relevância da variação dos valores de ( $\mu_i$ ), fato que se justifica em situações em que um servidor pode tornar-se mais eficiente, por exemplo, a partir de uma substituição de equipamentos ou mesmo algum tipo de estratégia de capacitação dos servidores. Esse aumento de eficiência acaba por acarretar certo custo, que poderá ou não ser justificado por um ganho na taxa de atendimento do sistema (CRUZ et al., 2012).

### 3.4.4 Aplicações do PAB

Conforme mencionado anteriormente, alguns sistemas de filas com mais de uma instalação se configuram como redes de filas e, havendo limitação de espaço de espera, tem-se um problema de alocação de *buffers* nessas redes. Esse tipo de problema é comum em sistemas de manufatura, como pode ser visto no estudo de caso feito por Bitran e Morabito (1995b), onde os autores apresentam um modelo de redes de filas abertas para um sistema de manufatura discreto, revendo a avaliação de desempenho do sistema para estimar medidas de desempenho como estoque em processo, tempo de ciclo de produtos, taxa de

produção e o nível de utilização de equipamentos. São analisados os modelos de avaliação de desempenho para redes de Jackson e redes de Jackson generalizadas a partir de métodos de decomposição, modelos de otimização combinando técnicas de programação matemática e teoria das redes de filas são apresentados em outro trabalho dando sequência aos estudos realizados pelo autor (BITRAN; MORABITO, 1995a).

Estudo para alocação de capacidade em um sistema *job-shop* de uma fábrica metal-mecânica é apresentado em Silva e Morabito (2007) em que foram utilizados modelos de programação matemática e métodos aproximados de decomposição paramétrica para avaliar as medidas de desempenho da rede, com o objetivo de auxiliar no processo de tomada de decisão de projeto ou reprojeto da rede.

Manitz (2008) utiliza o método de decomposição e descrevem um procedimento de aproximação para avaliação do rendimento de uma linha de montagem com múltiplos estágios, que por ser um processo assíncrono com tempos de processamento estocásticos em cada estação, acaba formando filas. Expressões na forma produto são obtidas para atrasos médios em uma rede ferroviária modelada como uma rede de filas em Huisman, Boucherie e Dijk (2002). Os autores utilizam essas expressões para calcular as medidas de desempenho e avaliar novos projetos de rede, bem como simulações para expansão de capacidade.

O processo de filas de caminhões em um terminal de contêineres também pode ser modelado como uma rede de filas de forma que se possa estimar o comprimento da fila e otimizar o tempo de giro do caminhão dentro do terminal. Modelos de otimização baseados em algoritmos genéticos mostram-se eficientes para esse tipo de modelagem (ZHANG; ZENG; CHEN, 2013).

Em geral, estudos de redes de filas compreendem um problema de otimização multiobjetivo, onde se busca a melhor alocação de espaço de *buffer* bem como a minimização de taxas de serviço, por exemplo. Para esse tipo de problema, os algoritmos genéticos apresentam boa aplicabilidade por serem facilmente adaptáveis à função objetivo, bem como para problemas de análise de topologia da rede com o intuito de encontrar a melhor configuração atendendo aos múltiplos objetivos estabelecidos. Para o problema de análise de topologia de rede, Smith, Cruz e Woensel (2010) apresentam um método de aproximação para estimar o desempenho de redes configuradas em diferentes topologias. Em seu trabalho é feita uma abordagem abrangente apresentando as fórmulas de probabilidade de bloqueio e o método de otimização baseado em algoritmos genéticos.

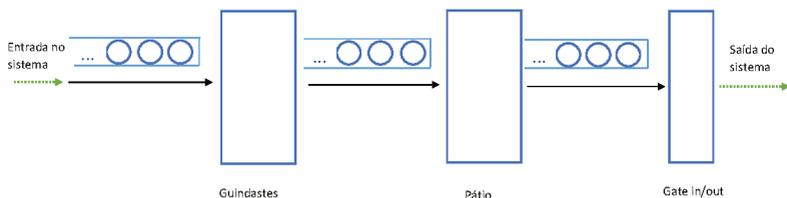
No próximo capítulo serão apresentados o modelo de redes de filas utilizado para representar a operação portuária do Terminal de Contêineres do Porto do Rio Grande e os algoritmos do método de solução escolhido.

## 4 MODELO E MÉTODO DE SOLUÇÃO

### 4.1 DEFINIÇÃO DAS VARIÁVEIS

Essa tese tem foco na modelagem de uma operação portuária como uma rede de filas com *buffers*, dessa forma, é necessário definir as variáveis que compõem o modelo. No caso da operação portuária, cada contêiner será tratado como um cliente que solicita serviço em uma sequência de estações de trabalho, quais sejam: guindastes, pátio e *gate in/out*, conforme representado na Figura 15. Estas estações compõem a rede de filas idealizada nesse trabalho. A escolha das mesmas foi baseada na configuração da operação portuária do terminal estudado.

Figura 15: Modelo Inicial de Rede de Filas



Fonte: Elaborada pela autora

Na frente de cada estação tem-se a formação de uma fila de clientes, que nesse caso, são os contêineres, e ainda, cada estação conta com uma taxa de serviço própria e independente. O cliente chega no sistema, é atendido na primeira estação (guindastes) e segue para o atendimento na estação de pátio formando uma nova fila, onde aguarda atendimento das máquinas especializadas para serem devidamente armazenados. Por último, uma nova fila é formada em frente à estação *gate in/out*, onde o cliente (contêiner) aguarda para deixar o sistema (terminal de contêineres).

De acordo com os conceitos vistos no Capítulo 3 e a partir da entrevista com o gerente de operações do terminal estudado, pode-se enumerar algumas características do modelo proposto. A população é considerada infinita e todas as filas formadas em frente às estações são finitas, ou seja, existe uma capacidade máxima  $K$  para o sistema considerado.

A disciplina de fila adotada para o modelo é FIFO, em que foram

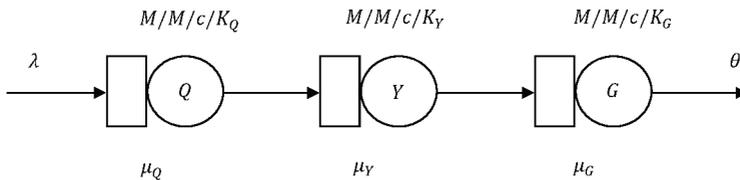
desconsideradas as diferentes movimentações de contêineres no empilhamento dos mesmos. Nesse caso, foi levado em consideração que o serviço executado na segunda estação (pátio) requer a disponibilidade de tratores de pátio e o tempo médio do serviço. O mesmo foi considerado para a estação *gate in/out*.

Os processos de chegada e de atendimento de cada estação foram admitidos como sendo processos markovianos seguindo a distribuição exponencial, considerando as taxas correspondentes como aquelas informadas pelo gerente de operações do terminal, uma vez que não foi possível obter dados brutos para verificação das mesmas.

## 4.2 MODELAGEM DO PROBLEMA

A partir da representação do modelo inicial na Figura 15 e da definição das distribuições dos tempos entre chegadas e tempos de serviço, foi possível modelar a rede de filas com alocação de *buffers* para a operação portuária do terminal de contêineres, representada na Figura 16.

Figura 16: Modelo de Rede de Filas



Fonte: Elaborada pela autora

A rede é modelada como um grafo  $G(N, A)$ , em que cada nó representa uma fila do sistema associada às estações correspondentes, respectivamente ao cais (*Quay*) ( $Q$ ), ao pátio (*Yard*) ( $Y$ ) e ao Gate in/out ( $G$ ). Cada fila segue o modelo  $A/B/c/K$  com taxa de chegada externa  $\lambda$  conhecida e taxas de serviço  $\mu_i$  conhecida em cada estação. As taxas de entrada nas estações  $Y$  e  $G$  são assumidas como as taxas de saída das estações anteriores.

As demais variáveis consideradas são:

- $\lambda$  = taxa de entrada/chegada externa no sistema;
- $\mu_Q$  = taxa de serviço da estação  $Q$ ;
- $\mu_Y$  = taxa de serviço da estação  $Y$ ;
- $\mu_G$  = taxa de serviço do servidor  $G$ ;
- $\theta$  = taxa de saída/atendimento (*throughput*) do sistema;

Para o modelo tem-se o problema de otimização multiobjetivo conforme proposto no Capítulo 3, equações (3.21), (3.22) e (3.22), cuja função objetivo  $F(K, \mu) \equiv (f_1(K), f_2(\mu), -f_3(K, \mu))^T$  busca simultaneamente, a minimização dos custos das áreas de espera alocadas ( $f_1(K) = \sum_{\forall i \in N} K_i$ ) e das taxas de serviço ( $f_2(\mu) = \sum_{\forall i \in N} \mu_i$ ) e a maximização da taxa de atendimento ( $f_3(K, \mu) = \theta(K, \mu)$ ).

Para a construção do modelo de redes de filas considerou-se os contêineres como sendo os clientes na rede, e dessa forma, a chegada desses clientes estando atrelada à chegada de navios no porto, necessita do tratamento como uma chegada em blocos, para que se tenha a chegada efetiva de clientes e não de navios. Nesse contexto, foi necessário modelar a chegada de navios para obter a taxa de chegada efetiva  $\lambda$  de clientes na rede.

#### 4.2.1 Modelo de Chegadas em Bloco

Tendo em vista a modelagem do problema como uma rede de filas onde os clientes são os contêineres, é necessário a obtenção da taxa efetiva de chegada de clientes na rede. Essa informação não é disponibilizada pelo operador portuário, dessa forma é preciso proceder com o tratamento da fila de navios como uma fila com chegadas em bloco, onde cada navio é representado como um bloco de tamanho  $X$ , em que  $X$  representa a quantidade de contêineres.

O modelo de chegadas em bloco foi baseado em Laxmi e Gupta (2000), em que os navios serão representados por blocos de tamanho  $X$  com probabilidade  $P(X = s) = g_s$   $s \geq 1$  e média  $E(X) = \bar{g}$ , sendo  $X$  a variável aleatória que representa o número de clientes na chegada, que pode ser em blocos.

A configuração do modelo é dada por  $M[X]/M/c/K$  onde  $M[X]$  representa a distribuição exponencial dos tempos entre chegadas dos blocos,  $M$  a distribuição, também exponencial, para os tempos de serviço,  $c$  é o número de servidores e  $K$  a capacidade do sistema, ou

seja, o número máximo de clientes permitidos. Os tempos entre chegadas são variáveis aleatórias independentes e identicamente distribuídas, com função distribuição acumulada  $T(z)$ , função densidade de probabilidade  $t(z)$ ,  $z \geq 0$  e média  $a$ .

A intensidade de tráfego é dada por  $\rho = \frac{\bar{g}}{ac\mu}$ , a disciplina da fila é FCFS. Os servidores  $c$  são independentes, com tempo de serviço exponencialmente distribuído e tempo médio igual a  $\frac{1}{\mu}$ . Ainda, os tempos entre chegadas, tamanho dos blocos e tempo de serviço são mutuamente independentes.

Como existe limitação de espaço de *buffer*, duas condições precisam ser observadas na chegada de um lote/bloco: são preenchidos os espaços vagos e os demais clientes do lote são dispensados (rejeição parcial) ou, todo o lote é rejeitado se não houver espaço suficiente (rejeição total) (GONTIJO et al., 2011).

Em uma operação portuária, na prática os dois casos podem ocorrer, ou seja, um navio (lote) pode ser totalmente rejeitado por falta de espaço para a operação ou pode ser aceito para atendimento parcial. Para efeito da modelagem nesse trabalho, optou-se pela rejeição total, uma vez que a decisão entre uma condição ou outra não pode ser prevista no modelo.

O objetivo desse modelo é obter a taxa de entrada efetiva de clientes no sistema,  $\lambda$ , além de calcular as medidas de desempenho para uma fila com chegadas em bloco e espaço de *buffer* finito com rejeição total. A probabilidade de bloqueio ocorre quando não há servidores disponíveis e, se tratando de filas com chegadas em bloco é necessário considerar as probabilidades de bloqueio do primeiro e último cliente,  $P_{BF}$  e  $P_{BL}$ , respectivamente, bem como a de bloqueio arbitrário,  $P_{BA}$ . No caso de rejeição total  $P_{BF} = P_{BL}$ , já que todo o bloco é dispensado caso não possa ser atendido em sua totalidade.

As medidas de desempenho procuradas são calculadas pelas seguintes equações (LAXMI; GUPTA, 2000):

$$L_q = \sum_{s=1}^K (s - c) \pi_s \quad (4.1)$$

$$L = \sum_{s=1}^K s \pi_s \quad (4.2)$$

$$P_{BL} = P_{BF} = \sum_{s=0}^K \pi_s \sum_{t=K-s+1}^{\infty} g_t \quad (4.3)$$

$$P_{BA} = \sum_{s=0}^{\infty} \pi_s \sum_{t=K-s+1}^{\infty} \frac{tg_t}{\bar{g}} \quad (4.4)$$

$$\lambda' = \frac{\bar{g}(1 - P_{BA})}{a} \quad (4.5)$$

$$W_A = \frac{L}{\lambda'} \quad (4.6)$$

$$W_{qA} = \frac{L_q}{\lambda'} \quad (4.7)$$

Onde  $L_q$  e  $L$  são o tamanho médio da fila e o tamanho médio do sistema, respectivamente;  $\lambda'$  representa a taxa efetiva de chegada e  $W_A$  e  $W_{qA}$  são, respectivamente, o tempo médio de espera no sistema e o tempo médio de espera na fila, definidos a partir das Fórmulas de Little.

As probabilidades de estado no período pré-chegada  $\pi_t$ 's, são determinadas a partir do sistema linear de equações (LAXMI; GUPTA, 2000)

$$\left\{ \begin{array}{l} \pi_t = \sum_{s=0}^K p_{st} \pi_s, \quad \text{para } t = 0, 1, 2, \dots, K-1 \\ \sum_{s=0}^K \pi_s = 1 \end{array} \right. \quad (4.8)$$

resultando em um sistema de  $K+1$  equações a  $K+1$  incógnitas, que na forma matricial é dado por:

$$\begin{bmatrix} (p_{0,0} - 1) & p_{1,0} & \cdots & p_{K,0} \\ p_{0,1} & (p_{1,1} - 1) & \cdots & \vdots \\ \vdots & \cdots & \ddots & \vdots \\ p_{0,K-1} & \cdots & \cdots & (p_{K,K-1} - 1) \\ 1 & \cdots & \cdots & 1 \end{bmatrix} \begin{bmatrix} \pi_0 \\ \pi_1 \\ \vdots \\ \vdots \\ \pi_K \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \quad (4.9)$$

A variável  $p_{st}$  presente no sistema (4.8) representa as probabilidades de transição de um passo, e são calculados pelas equações:

$$p_{st} = \begin{cases} \sum_{r=t-s}^{K-s} \beta_{s+r-t} g_r, & t > s \geq 0, t \geq c \\ \sum_{r=1}^{K-s} \beta_{s+r-t} g_r + \beta_{s-t} \sum_{r=K-s+1}^{\infty} g_r, & c \leq t \leq s \\ \sum_{r=\max\{1, t-s\}}^{K-s} V_{s+r,t} g_r + V_{s,t} \sum_{r=K-s+1}^{\infty} g_r, & s \geq 0, 1 \leq t \leq c-1 \\ 1 - \sum_{t=1}^K p_{st}, & t = 0 \end{cases} \quad (4.10)$$

onde

$$V_{r,t} = \begin{cases} \frac{\lambda}{\mu} \frac{r! \Gamma\left(t + \frac{\lambda}{\mu}\right)}{t! \Gamma\left(r + \frac{\lambda}{\mu + 1}\right)}, & t \leq r \leq c \\ \frac{\lambda}{\mu} \left(\frac{c\mu}{\lambda + c\mu}\right)^{r-c} \frac{c! \Gamma\left(t + \frac{\lambda}{\mu}\right)}{t! \Gamma\left(r + \frac{\lambda}{\mu + 1}\right)}, & r > c, 0 \leq t \leq c-1 \end{cases} \quad (4.11)$$

em que a função  $\Gamma$  é uma extensão da função fatorial, definida como  $\Gamma(n+1) = n!$  ou  $\Gamma(n) = (n-1)!$ , e

$$\beta_m = \frac{\lambda}{\lambda + c\mu} \left(\frac{c\mu}{\lambda + c\mu}\right)^m, \quad m \geq 0 \quad (4.12)$$

com  $m = s+r-t$ , sendo  $s$  e  $t$  os estados do sistema representando a quantidade de clientes no período pré-chegada e  $r$  o tamanho do lote.

Para as equações (4.10), (4.11) e (4.12) são considerados lotes (navios) de clientes de tamanho aleatório  $X$  chegando em períodos  $\tau_n$ ,  $n = 0, 1, 2, \dots$ .  $\tau_n^*$  denota o período de tempo logo antes do instante da chegada  $\tau_n$ . Assim, em  $\tau_n^*$  um lote de tamanho  $r$  ( $r \geq 1$ ) chega e encontra o sistema no estado  $s$  com ( $0 \leq s \leq K$ ), então o total de clientes no instante  $\tau_n$  é  $s+r$ . No instante seguinte, ou seja, em  $\tau_{n+1}^*$ ,  $t$  ( $0 \leq t \leq K$ ) clientes precisam ser aceitos, dessa forma ( $s+r-t$ ) clientes devem deixar o sistema durante o período entre chegadas. Para

melhor entendimento, o alcance de  $t$  é dividido em duas partes, como segue:

(i)  $0 \leq t < c$ : sendo o tamanho do lote e os tempos de serviço independentes, a probabilidade de  $t$  clientes no término do tempo entre chegadas é dada por  $V_{s+r,t}g_r$ , onde  $V_{s+r,t}$  significa a probabilidade de  $(s+r-t)$  clientes completarem o serviço durante o tempo entre chegadas;

(ii)  $c \leq t \leq K$ : todos os servidores estão ocupados continuamente ao longo do período entre chegadas e a taxa de serviço é constante e igual a  $c\mu$ . A probabilidade a ser calculada é dada por  $\beta_{s+r-t}g_r$ , onde  $\beta_m (m \geq 0)$  significa  $m$  partidas durante o tempo entre chegadas.

O cálculo das medidas de desempenho foi feito a partir do Algoritmo 1, implementado em *Python*, versão 3.6, desenvolvido em ambiente Linux.

<b>Algoritmo 1: Algoritmo de Chegadas em Bloco</b>	
<b>Entrada:</b>	capSistema, taxaServico, taxaChegada, nServidores, intensTrafego, tamBlocos [ ]
<b>Saída:</b>	medidasDeDesempenho
<b>1 início</b>	
<b>2</b>	# preenche cada item da primeira linha com um vetor com os dados das probabilidades;
<b>3</b>	# criando assim a p [s][t];
<b>4</b>	<b>enquanto</b> $s < \text{capSistema}$ <b>faça</b>
<b>5</b>	p[s]=controeColuna( );
<b>6</b>	s + +;
<b>7</b>	<b>fim</b>
<b>8</b>	# cria um vetor auxiliar na criação de pi;
<b>9</b>	<b>enquanto</b> $s < \text{capSistema}$ <b>faça</b>
<b>10</b>	b[0][s]=0;
<b>11</b>	s + +;
<b>12</b>	<b>fim</b>
<b>13</b>	b[0][capSistema]=1;
<b>14</b>	# calcula vetor pi utilizado nas medidas de desempenho;
<b>15</b>	pi=b.matrizInversa(p);
<b>16</b>	escreva medidasDeDesempenho( );
<b>17 fim</b>	

Fonte: Elaborada pela autora

No próximo capítulo serão apresentados os resultados obtidos

pelo Algoritmo 1, bem como o tratamento dado aos tamanhos dos blocos.

Na próxima seção serão descritos os métodos de solução escolhidos para o problema de otimização multiobjetivo associado à modelagem da operação portuária como uma rede de filas com alocação de *buffers*.

### 4.3 MÉTODO APLICADO

O objetivo central dessa tese é modelar a operação portuária de um terminal de contêineres como uma rede de filas com alocação de *buffers*, gerando assim, um problema de otimização multiobjetivo, conforme apresentado no Capítulo 3. Para sustentar essa modelagem é necessário solucionar o problema associado, o que para tanto recorreu-se aos métodos de solução específicos para esse tipo de rede, encontrados na literatura.

Uma vez a rede modelada, o tratamento inicial dessa rede é feito pelo Método de Expansão Generalizada idealizado por Kerbache e Smith (1987), em que a mesma é definida como um grafo e devidamente expandida para comportar a possibilidade de bloqueio entre as estações e consequente atualização das taxas de entrada e de serviço em cada nó da rede.

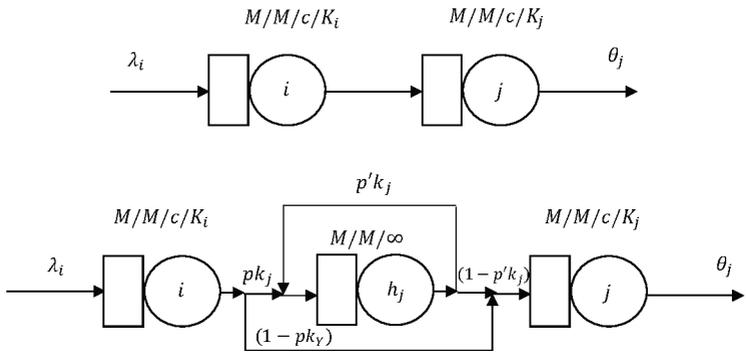
Por fim, para solucionar o problema de otimização multiobjetivo proposto nessa tese, é implementado um algoritmo genético com elitismo e ordenação não-dominada, conhecido como NSGA-II (do inglês *Non-Dominated Sorting Genetic Algorithm*) desenvolvido por Deb et al. (2002). A escolha pelo NSGA-II se justifica pela adaptação dos algoritmos genéticos a diferentes tipos de problemas de otimização, sendo este elaborado especificamente para um problema de otimização multiobjetivo.

#### 4.3.1 Método de Expansão Generalizada

O GEM (do inglês *Generalized Expansion Method*), é um método para obtenção de boas aproximações para as medidas de desempenho em redes de filas finitas, caso estudado na presente tese. O objetivo do algoritmo é proporcionar um esquema de aproximação das medidas de desempenho da rede de filas, levando em consideração o bloqueio

após serviço<sup>1</sup> (KERBACHE; SMITH, 1987). O método consiste em uma combinação de tentativas repetidas, decompondo a rede nó a nó para cada fila  $i$ , que é sucedida por uma fila finita  $j$ , conforme mostra a Figura 17.

Figura 17: Método de Expansão Generalizada



Fonte: Elaborada pela autora

O método é dividido em três estágios, o primeiro deles é a reconfiguração da rede que consiste em criar um nó artificial  $h_j$  para cada fila finita  $j$ , com taxa de serviço  $\mu_j$  e modelado como uma fila  $M/M/\infty$  com o objetivo de contabilizar os itens bloqueados, modelando o atraso adicional. Cada cliente que se dirige à  $j$  pode ser bloqueado com uma probabilidade  $p_{k_j}$ , caso contrário  $(1 - p_{k_j})$ . Quando ocorre o bloqueio, o cliente é encaminhado à fila  $h_j$ , onde espera até que seja liberado espaço na fila  $j$ . Dessa forma o nó  $h_j$  registra o tempo de espera do cliente antes de entrar na fila  $j$ .

O segundo estágio do método é quando são estimados os parâmetros  $p_{k_j}$ ,  $p'_{k_j}$  e a taxa  $\mu_{h_j}$ , de probabilidade de bloqueio, probabilidade de um segundo bloqueio e taxa de serviço do nó artificial, respectivamente. A probabilidade de bloqueio  $p_{k_j}$  é obtida pela Equação 4.13.

<sup>1</sup> Bloqueio que ocorre no caso de um cliente que conclui o serviço em uma estação, mas encontra a estação seguinte ocupada

$$p_{kj} = \frac{\left(\frac{\lambda}{\mu_j}\right)^{K_j} (c!)^{-1} (c^{K_j-c})^{-1}}{\left( e^{\frac{\lambda}{\mu_j}} \Gamma\left(c, \frac{\lambda}{\mu_j}\right) \Gamma^{-1}(c) + \frac{\left(\frac{\lambda}{\mu_j}\right)^c \left(1 - \left(\frac{\lambda}{\mu_j}\right)^{K_j-c+1}\right)}{(c!) \left(1 - \frac{\lambda}{\mu_j}\right)} \right)} \quad (4.13)$$

Por falta de uma solução de forma fechada para a probabilidade de um segundo bloqueio, recorre-se à aproximação via técnica de difusão, que foram empregadas com sucesso por Labetoulle e Pujolle (1980).

$$p'_{kj} = \left( \frac{\mu_j + \mu_{hj}}{\mu_{hj}} - \frac{\lambda((z^{K_j} - y^{K_j}) - (z^{K_j-1} - y^{K_j-1}))}{\mu_{hj}((z^{K_j+1} - y^{K_j+1}) - (z^{K_j} - y^{K_j}))} \right)^{-1} \quad (4.14)$$

onde  $y$  e  $z$  são raízes do polinômio  $\lambda - (\lambda + \mu_{hj} + \mu_j)x + \mu_{hj}x^2 = 0$  e  $\lambda = \lambda_j - \lambda_{hj}(1 - p'_{kj})$ , em que  $\lambda_{hj}$  é a taxa de chegada real ao nó artificial e  $h_j$  é a taxa de chegada real para o nó finito  $j$ , dadas pela equação 4.15.

$$\lambda_j = \lambda_i (1 - p_{kj}) = \lambda_i - \lambda_{hj} \quad (4.15)$$

sendo  $\lambda_i$  a taxa de atendimento na fila antecessora.

Por fim, o cálculo da taxa de serviço no nó artificial é dado pela equação 4.16

$$\mu_{hj} = \frac{2\mu_j}{1 + \sigma_j^2 \mu_j^2} \quad (4.16)$$

onde  $\sigma_j^2$  é a variância do tempo de serviço. Porém se a distribuição do tempo de serviço na fila finita fazendo o bloqueio é exponencial, pela teoria da renovação tem-se que  $\mu_{hj} = \mu_j$ , ou seja, a taxa de serviço no nó artificial é a mesma taxa de serviço do nó finito  $j$ .

O último estágio é a eliminação do laço de retroalimentação no nó  $h_j$ , pois este cria uma forte dependência no processo de chegada no nó  $j$ . Para eliminá-lo é feito um acréscimo adequado de tempo de serviço do nó  $i$  e assim tem-se uma taxa de serviço ajustada para o nó  $h_j$  dada por

$$\mu'_{hj} = (1 - p'_{kj})\mu_{hj} \quad (4.17)$$

O tempo médio de serviço em um nó  $i$  precedente a um nó finito é dado por  $\mu_i^{-1}$  na fase insaturada ( $1 - p_{kj}$ ) e  $\mu_i^{-1} + \mu'_{hj^{-1}}$  na fase saturada ( $p_{kj}$ ). Assim, o tempo médio de serviço no nó  $i$  precedente a um nó finito é dado pela equação 4.18

$$\widehat{\mu}_i^{-1} = \mu_i^{-1} + p_{kj}\mu'_{hj^{-1}} \quad (4.18)$$

O objetivo desse último estágio além da eliminação do laço de *feedback*, é também atualizar as taxas de serviço de cada uma das filas  $i$  que são precedidas por filas finitas e possibilitar a obtenção de uma aproximação acurada para a taxa de saída  $\theta_j$ .

Pela propriedade PASTA<sup>2</sup>, a taxa de chegada efetiva vista pelos servidores é dada por  $\lambda_{eff} = \lambda(1 - p_{kj})$  e assim, a taxa de saída/atendimento pode ser dada por  $\theta_j = \lambda_{eff}$ , (GROSS et al., 2009).

Para a expansão da rede e a devida atualização das taxas de serviço em cada um dos nós da rede, foi desenvolvido o Algoritmo 2 especializado, cujos resultados serão apresentados no próximo capítulo.

<b>Algoritmo 2:</b> Algoritmo de Expansão Generalizada	
<b>Entrada:</b>	redesNos
<b>Saída:</b>	novaRede
<b>1 início</b>	
<b>2</b>	# leia redeNos;
<b>3</b>	# $N$ recebe o numero de nos na rede original;
<b>4</b>	$N =  \text{redeNos} $ ;
<b>5</b>	# cria rede vazia que deverá ser a nova rede;
<b>6</b>	# novaRede=vazio;
<b>7</b>	# percorre cada item da rede $i = 1$ ;
<b>8</b>	<b>enquanto</b> $i < N - 1$ <b>faça</b>
<b>9</b>	faça novaRede=novaRede $\cup$ redeNos[ $i$ ];
<b>10</b>	novaRede=novaRede $\cup$ noArbitrario(redeNos[ $i$ ],redeNos[ $i + 1$ ];
<b>11</b>	<b>fim</b>
<b>12</b>	$i + +$ ;
<b>13</b>	# fim;
<b>14</b>	escreva novaRede( );
<b>15 fim</b>	

Fonte: Elaborado pela autora

<sup>2</sup>Poisson Arrivals See Time Averages: Propriedade dos fluxos de Poisson.

O algoritmo 2 foi implementado em *Python*, versão 3.6, desenvolvido em ambiente Linux. A rede expandida obtida por esse algoritmo serve como dado de entrada no próximo algoritmo de otimização para solução do problema proposto nessa tese.

### 4.3.2 Algoritmo de Otimização Multiobjetivo

Em problemas de otimização multiobjetivo contendo objetivos conflitantes, geralmente não existe uma única solução ótima que contemple a função objetivo. Sabendo que uma solução para problemas desse tipo é constituída por um conjunto de pontos, os Algoritmos Genéticos (AG) podem ser utilizados, já que a cada geração do algoritmo é considerada uma população de soluções (BARBOSA; RIBEIRO; ARANTES, 2010).

Nesse sentido, algoritmos evolutivos foram desenvolvidos para tratar desse tipo de problema de otimização e são conhecidos como Algoritmos Genéticos Multiobjetivos (AGMOs), sendo que alguns deles utilizam o conceito de Fronteira de Pareto, que consiste em uma curva no espaço de soluções em que estão as soluções não dominadas (KONAK; COIT; SMITH, 2006).

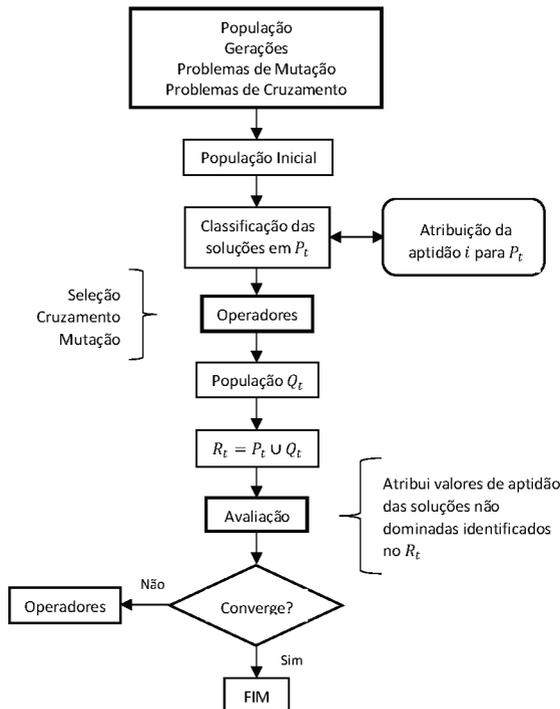
O objetivo é então aproximar essa curva classificando todas as possíveis soluções como dominadas e não dominadas. Por exemplo, dada uma solução  $x$ , se existe uma outra solução viável  $y$  que não seja pior que  $x$  para todos os objetivos, então  $x$  é dita solução dominada.

Diferentes algoritmos evolutivos multiobjetivo foram desenvolvidos nas últimas décadas. Para essa tese será utilizado o algoritmo NSGA-II desenvolvido por Deb et al. (2002).

O algoritmo NSGA-II baseia-se no conceito de ordenação elitista por dominância, tendo como objetivo classificar a população por fronteiras conforme o grau de dominância. Embora o NSGA-II seja baseado em algoritmos genéticos, existe uma diferença na forma como o operador de seleção é abordado, este é dividido em dois processos, o *Fast Non-Dominated Sorting* (Classificação Rápida por Não Dominância) e o *Crowding Distance* (Distância de Agrupamento). Para os demais operadores, o processo é similar aos AGs (DEB, 2001).

Na Figura 18 é apresentado um fluxograma geral para o algoritmo NSGA-II.

Figura 18: Fluxograma NSGA-II

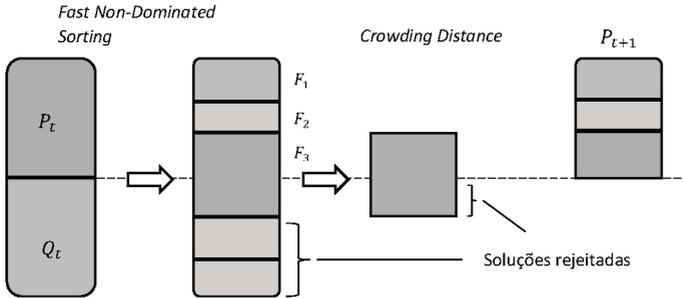


Fonte: Adaptado de Barbosa, Ribeiro e Arantes (2010)

A Figura 19 esquematiza o processo de seleção feito pelo NSGA-II, em que são aplicados os conceitos de ordenação elitista por dominância. O procedimento *Fast Non-Dominated Sorting* compara todos os indivíduos da população entre si para determinar o grau de dominância e classificá-los (DEB, 2001).

No segundo momento, os indivíduos cujo grau de dominância for igual a zero são ditos não dominados e inseridos na primeira fronteira de Pareto. Para os indivíduos restantes, aqueles cujo grau de dominância é diferente de zero, é recalculada a dominância entre eles, formando uma próxima fronteira. O processo continua até que a população esteja vazia (DEB, 2001).

Figura 19: Procedimentos de Seleção do NSGA-II



Fonte: Adaptado de Deb et al. (2002)

Encontrar soluções de Pareto-ótimas é um processo que tende a convergir para uma mesma região do espaço de busca. Para que as soluções encontradas estejam dispersas neste espaço, aplica-se o segundo processo do operador de seleção chamado *Crowding Distance* para ordenar as soluções dentro de uma mesma fronteira. Esse cálculo é feito de forma a selecionar indivíduos que estão mais distantes entre si na fronteira. O processo é executado de acordo com os critérios de otimização impostos inicialmente ao algoritmo, ou seja, para um problema que possui  $N$  critérios de otimização, cada um dos indivíduos terá  $N$  resultados  $Crowding_i$  e o valor final será o somatório de todos esses resultados (DEB et al., 2002).

Após a finalização do processo de classificação são selecionados os indivíduos pertencentes à primeira fronteira, aquela que possui os indivíduos não dominados, para compor a próxima geração até completar uma nova população.

Deb e Agrawal (1995) utilizaram como operador de cruzamento o *Simulated Binary Crossover* (SBX) que é conveniente quando pode-se codificar as variáveis por números reais. Os filhos  $x_{i,(*,t+1)}$  são obtidos dos pais  $x_{i,(*,t)}$ , conforme as equações 4.19.

$$\begin{aligned} x_{i,(1,t+1)} &= 0,5 \left[ (1 + \beta)x_{i,(1,t)} + (1 - \beta)x_{i,(2,t)} \right] \\ x_{i,(2,t+1)} &= 0,5 \left[ (1 - \beta)x_{i,(1,t)} + (1 + \beta)x_{i,(2,t)} \right] \end{aligned} \quad (4.19)$$

onde  $\beta$  é uma variável aleatória obtida da função de distribuição de probabilidade dada por:

$$f(\beta) = \begin{cases} 0,5(\eta + 1)\beta^\eta, & \text{se } \beta \leq 1 \\ 0,5(\eta + 1)\beta^{\frac{1}{\beta\eta+2}}, & \text{caso contrário} \end{cases} \quad (4.20)$$

Para cada uma das variáveis de decisão do problema em questão,  $K_i$  e  $\mu_i$ , a operação de mutação deverá ocorrer com uma probabilidade específica. De acordo com Deb e Agrawal (1995), foram adicionadas perturbações gaussianas à essas variáveis, fazendo  $K_i + \varepsilon_i$  e  $\mu_i + \varepsilon_{N+i}$ , com  $\varepsilon_i \sim N(0, 1)$ ,  $i \in \{1, 2, \dots, 2N\}$  para todo  $i \in N$ .

Para garantir a viabilidade das restrições após as operações de cruzamento e mutação, são aplicados operadores de reflexão (DEB, 2001) dados pelas equações

$$K_{\text{refl}_i} = 1 + |K_i - 1| \quad (4.21)$$

$$\mu_{\text{refl}_i} = \mu_{\text{inf}_i} + |\mu_i - \mu_{\text{inf}_i}| \quad (4.22)$$

onde 1 é o limite inferior para  $K_i$ ,  $\mu_{\text{inf}_i}$  é o limite inferior de  $\mu_i$  para assegurar que  $\rho < 1$ .

O critério de parada é aquele sugerido por Rudenko e Schoenauer (2004) em que é calculada uma medida de estabilidade das soluções não-dominantes após cada iteração.

O critério baseia-se na estabilização da máxima distância de aglomeração, definida como  $d_L$ , que é medida sobre  $L$  gerações e calculada a partir do desvio padrão dado na equação 4.23.

$$\sigma_L = \sqrt{\frac{1}{L} \sum_{l=1}^L (d_l - \bar{d}_L)^2} \quad (4.23)$$

onde  $\bar{d}_L$  é a média sobre  $L$  gerações. O algoritmo pode parar assim que  $\sigma_L < \delta_{lim}$ .

Rudenko e Schoenauer (2004) sugerem que os valores de  $L$  e  $\delta_{lim}$  sejam ajustados respectivamente, para 40 e 0,02. Assim tem-se como critério de parada a expressão  $\sigma_{40} \leq 0,02$ .

O Algoritmo 3 apresenta a síntese da implementação do algoritmo evolutivo multiobjetivo NSGA-II e assim como os demais, foi implementado em *Python*, versão 3.6, desenvolvido em ambiente Linux.

**Algoritmo 3:** NSGA-II

```

1 início
2   Recebe os parâmetros de inicialização → o tamanho da
   população  $N$ , o número máximo de geração  $G$ , a
   probabilidade de cruzamento rateCross, a probabilidade
   de mutação rateMut;
3   Recebe a rede já expandida → fila;
4   Recebe os parâmetros de parada → critério de parada  $\delta_P$ ;
5   GeraPopulaçãoInicial( $G_0$ );
6    $i = 0$ ;
7   enquanto  $\delta_i > \delta_P$  e  $i! = G$  faça
8     faça  $i + +$ ;
9      $P \rightarrow$  população de  $G_{i-1}$ ;
10     $Q \rightarrow$  população de  $G_i$ ;
11     $n \rightarrow$  número de indivíduos em  $G_i$ ;
12    enquanto  $n < N$  faça
13      faça  $P_1 \rightarrow$  indivíduo aleatório de  $G_i$ ;
14       $P_2 \rightarrow$  indivíduo aleatório de  $G_i$  diferente de  $P_1$ ;
15       $F_1 \rightarrow$  indivíduo gerado pelo cruzamento de  $P_1$  e  $P_2$ ;
16       $F_2 \rightarrow$  indivíduo gerado pelo cruzamento de  $P_1$  e  $P_2$ ;
17      Perturb → vetor com  $2^*$ número de nodos gerados da
      distribuição Normal  $[0,1]$  para todo  $j$  em  $F$ ;
18       $m_j \rightarrow$  probabilidade de  $F_j$  ser mutado;
19      se  $m_j < \mathbf{rateMut}$ ;
20       $F_j \text{Taxas} += \text{Perturb}_j$ ;
21       $F_j \rightarrow$  sofre alteração de reflexão (ajuste para os
      limites);
22       $Q \cup F_1$  e  $Q \cup F_2$ ;
23    fim
24     $T \rightarrow P \cup Q$ ;
25    ordenamentoPorFronteiras( $T$ );
26    ordenamentoPorDistância( $T$ ) em cada fronteira de  $T$ ;
27     $Q \rightarrow N$  melhores indivíduos de  $T$ ;
28     $\delta_i \rightarrow$  calculaCritérioParada();
29  fim
30 fim

```

Fonte: Elaborado pela autora

Os parâmetros utilizados do NSGA-II são aqueles sugeridos na literatura de referência e são apresentados a seguir na Tabela 1.

Tabela 1: Parâmetros NSGA-II

Descrição do parâmetro	Valor do parâmetro
Máximo de gerações	4000
Indivíduos da população	400
Probabilidade de mutação	0,2
Probabilidade de cruzamento	0,8
Parâmetro de dispersão	0,2
Gerações $L$	40
Desvio Padrão $\sigma_L$	0,02

Fonte: Elaborada pela autora

O objetivo central desta tese é a modelagem de uma operação portuária como uma rede de filas com alocação de *buffers*, o que para tanto buscou-se na literatura por trabalhos correlatos onde as principais aplicações encontradas estão voltadas à manufatura e estudo de filas de caminhões em um terminal de contêineres. Nesse sentido, optou-se por utilizar os métodos de solução de diferentes autores, conforme elucidado anteriormente, que são métodos comumente empregados para o tipo de problema que é colocado nessa tese.

Os algoritmos implementados para cada um dos métodos aplicados encontram-se à disposição com a autora para fins de estudos acadêmicos.

No próximo capítulo serão apresentados os principais resultados obtidos pelos algoritmos 1, 2 e 3 para o Problema de Rede de Filas com Alocação de Buffers aplicado à operação portuária do Terminal de Contêineres do Porto do Rio Grande.



## 5 RESULTADOS E DISCUSSÃO

### 5.1 INTRODUÇÃO

Neste capítulo são apresentados os principais resultados obtidos para a modelagem de uma operação portuária como uma rede de filas com alocação de *buffers*. Inicialmente é apresentada a modelagem da chegada de contêineres no terminal como uma fila com chegadas em bloco, passo necessário para a implementação dos Algoritmos 2 e 3. Em seguida são apresentados os resultados obtidos pelo Método de Expansão Generalizada (GEM) e pelo Algoritmo Multiobjetivo NSGA-II, finalizando com a discussão dos mesmos.

### 5.2 MODELAGEM DA CHEGADA DE CONTÊINERES

O primeiro passo, a partir da coleta de dados, foi o tratamento dos mesmos com a utilização de planilhas eletrônicas e o software ARENA<sup>®</sup>, com o objetivo de configurar as distribuições de probabilidade necessárias para a modelagem da chegada de contêineres no porto como uma fila com chegadas em bloco.

De acordo as premissas do Capítulo 3 é necessário determinar o modelo de fila que representará a chegada de contêineres no porto e, utilizando a ferramenta *Input Analyser* do software ARENA<sup>®</sup>, optou-se pela distribuição exponencial para os tempos entre chegadas, bem como para a taxa de serviço, essa última informada pelo gerente do terminal.

Para determinar a distribuição de probabilidade dos tamanhos do blocos foram utilizados os dados informados na visita *in loco*. Essa distribuição informa a probabilidade  $g_s$  de se ter um bloco de tamanho  $X$ , dado necessário para o modelo de fila com chegadas em bloco conforme descrito no capítulo anterior.

Com a ferramenta *Input Analyser* do software ARENA<sup>®</sup> foi possível gerar a distribuição de probabilidade para os tamanhos dos blocos conforme apresentada na Tabela 2.

Tabela 2: Distribuição de Probabilidade do Tamanho dos Blocos

Tamanho dos blocos $X$	Probabilidade $g_s$	Média $E(X) = \bar{g}$
82	0,00514	0,42148
137	0,00742	1,01654
191	0,01040	1,98640
245	0,01420	3,47900
299	0,01890	5,65110
354	0,02430	8,60220
408	0,03050	12,44400
462	0,03710	17,14020
517	0,04400	22,74800
571	0,05060	28,89260
625	0,05670	35,43750
679	0,06160	41,82640
734	0,06520	47,85680
788	0,06700	52,79600
842	0,06690	56,32980
897	0,06500	58,30500
951	0,06130	58,29630
1010	0,05630	56,86300
1060	0,06120	64,87200
1110	0,04350	48,28500
1170	0,03670	42,93900
1220	0,03020	36,84400
1280	0,02390	30,59200
1330	0,01850	24,60500
1390	0,01390	19,32100
1440	0,01020	14,68800
1490	0,00740	11,02600
1550	0,00500	7,75000
1600	0,00336	5,37600
1660	0,00219	3,63540
1710	0,00139	2,37690

Fonte: Elaborada pela autora

Além da distribuição de probabilidade para os tamanhos dos blocos, expostos na Tabela 2, e da taxa de chegada de navios, ambos obtidos pela ferramenta *Input Analyser* do software ARENA<sup>®</sup>, os demais dados de entrada do Algoritmo 1 foram apanhados a partir das informações tomadas na entrevista *in loco* com o gerente de operações do terminal. Os dados são apresentados na Tabela 3.

Tabela 3: Dados de entrada para o Algoritmo 1

Variáveis	Descrição	Valor/unidade
$\mu$	taxa de serviço	23,45 contêiner/hora
$\lambda$	taxa de chegada de navios	0,07 navio/hora
$c$	quantidade de servidores	3 berços
$K$	capacidade do sistema	3.000 contêiner

Fonte: Elaborada pela autora

Conforme o exposto no Capítulo 4 e os dados apresentados nas Tabelas 2 e 3, o modelo para a chegada de contêineres é dado por  $M[X]/M/c/K$ , onde  $M[X]$  representa a distribuição exponencial dos tempos entre chegadas dos blocos,  $M$  a distribuição, também exponencial, para os tempos de serviço,  $c$  é o número de servidores e  $K$  a capacidade do sistema.

Os resultados obtidos a partir do Algoritmo 1 estão dispostos na Tabela 4, onde a unidade de tempo está representada em horas e os tamanhos representados em contêineres.

Tabela 4: Resultados para fila chegada em blocos

Variáveis	Descrição	Valor
$L_q$	tamanho médio da fila	834
$L$	tamanho médio do sistema	837
$P_{BL} = P_{BF}$	probabilidade de bloqueio	0,0803
$P_{BA}$	probabilidade de bloqueio arbitrária	0,1005
$W_A$	tempo médio de espera no sistema	16,1804
$W_q A$	tempo médio de espera na fila	16,1377
$\lambda'$	taxa efetiva de chegada	51,7297
$\rho$	fator de utilização	0,82

Fonte: Elaborada pela autora

O dado de interesse é a taxa efetiva de chegada  $\lambda'$  que representa a taxa de chegada de contêineres e não mais de navios, uma vez que a modelagem da operação portuária como uma rede de filas leva em consideração o cliente como sendo a unidade contêiner. Essa taxa será utilizada como um dos dados de entrada para o Algoritmo 2.

As demais medidas de desempenho calculadas pelo algoritmo serviram para comparação com a situação real, por exemplo, o tempo

médio de permanência de um navio no terminal, obtido pelos dados informados, é em torno de 16,27 horas, valor aproximado ao calculado pelo algoritmo.

Outra medida analisada é o fator de utilização, que representa a taxa de ocupação do cais que, de acordo com dados informados gira em torno de 75%, uma vez que, segundo o gerente de operações, não é interessante para o terminal ocupar todos os berços simultaneamente. O valor calculado pelo algoritmo se aproxima do valor informado para o caso real.

Os resultados obtidos na modelagem da chegada de contêineres asseguram a continuidade do estudo, haja vista que o parâmetro  $\rho = 0,82 < 1$  garante o estado de equilíbrio no estudo de filas.

Na próxima seção são apresentados os resultados para a modelagem da operação portuária como uma rede de filas com alocação de *buffers*.

### 5.3 RESULTADOS GEM E NSGA-II

A modelagem da operação portuária como uma rede de filas finitas com espaço de *buffer* requer a expansão do grafo que representa a rede a fim de contabilizar os possíveis bloqueios entre os nós finitos e os recorrentes atrasos nos tempos de serviço conforme esquematizado na Figura 17, pág. 73.

Esses bloqueios podem ocorrer ao passo que um cliente liberado em uma estação de trabalho não pode seguir para a próxima estação, se esta ainda estiver ocupada, logo, esse cliente é encaminhado ao nó artificial que deverá contabilizar os atrasos ocorridos na rede.

Nesse sentido, o Algoritmo 2 basicamente calcula os bloqueios em cada um dos nós de serviço a montante, estima os parâmetros dos nós de espera e finalmente corrige as taxas de serviço eliminando os arcos de realimentação e os nós artificiais. Completados os três estágios do algoritmo, conforme descritos no Capítulo 4, tem-se uma rede devidamente expandida que será utilizada para o cálculo das medidas de desempenho para a rede inicialmente modelada.

Para a execução do algoritmo de expansão da rede são necessários os dados da Tabela 5 e a taxa efetiva de chegada calculada no algoritmo anterior,  $\lambda' = 51,7297$  contêiner/hora.

A unidade utilizada para as taxas de serviço na tabela abaixo é contêiner/hora, cujas médias foram informadas pelo gerente de operações do terminal de contêineres.

Tabela 5: Dados fornecidos pelo terminal 1

Descrição dos dados	Estação 1	Estação 2	Estação 3
Capacidade da fila	3000	300	100
Quantidade de servidores	3	30	15
Taxa de serviço	23,45	6	4

Fonte: Elaborada pela autora

O algoritmo gera um arquivo contendo os dados da rede expandida e suas medidas atualizadas considerando os possíveis bloqueios e atrasos em cada nó, para que possam ser lidos pelo algoritmo de otimização NSGA-II que procederá com o cálculo do problema de otimização multiobjetivo conforme as Equações 3.21, 3.22 e 3.22 (pág. 62).

Conforme descrito no Capítulo 4, o Algoritmo 2 de otimização apresenta um família de soluções (soluções Pareto-ótimas) oferecendo ao decisor um apanhado de alternativas a serem escolhidas de acordo com a necessidade específica do problema que se quer resolver. Dependendo do tamanho do conjunto de Pareto fornecido, podem ser utilizados outros mecanismos para filtrar as melhores soluções que estiverem relacionadas a um dos objetivos iniciais.

Essa filtragem final pode levar em consideração, por exemplo, as questões do alto custo atrelado ao espaço de *buffer*. Cabe ao decisor decidir o que se deve priorizar em relação aos multiobjetivos avaliados no problema posto, para que esse aparato de soluções filtradas seja de fato um mecanismo de apoio às devidas decisões referentes à utilização dos recursos disponíveis, bem como ao que se dispõe para possíveis investimentos.

Nesta tese, não foi possível obter o direcionamento para a filtragem dos resultados obtidos pelo Algoritmo 3, sendo assim, para efeito de análise dos resultados obtidos pelos algoritmos, dentre o conjunto de 100 soluções de Pareto geradas, foram escolhidas as 10 melhores soluções em relação a variável capacidade do sistema, ou seja, o somatório dos tamanhos das áreas de espera  $K$ .

No problema de otimização, a variável de decisão  $K$  deve ser minimizada, assim, as soluções escolhidas para análise representam as de menor somatório das áreas de espera. Os dados extraídos dos resultados estão organizados na Tabela 6.

Tabela 6: Soluções de Pareto selecionadas

$\sum_{i \in N} K_i$ (contêiner)	$\sum_{i \in N} \mu_i$ (hora/contêiner)	$\theta$ (contêiner/hora)
1870	0,062423562	62,473626
1880	0,053193498	66,516892
1980	0,062423562	68,410801
1990	0,053703184	64,722228
2110	0,044270315	62,473626
2120	0,046204956	61,731032
2340	0,046200042	66,516892
2350	0,053193498	69,382558
2410	0,053257154	66,516892
2460	0,053703184	68,583354

Fonte: Elaborada pela autora

Os dados foram ordenados de acordo com o somatório das áreas de espera na rede, ou seja, a capacidade de cada nó que representa as três estações de trabalho, quais sejam, o cais, o pátio e o *gate in/out*.

Em relação aos dados informados inicialmente pode-se dizer que o terminal poderia reduzir a capacidade de *buffer* atual que soma 3400 contêineres, assumindo o cenário que considera apenas o movimento de importação.

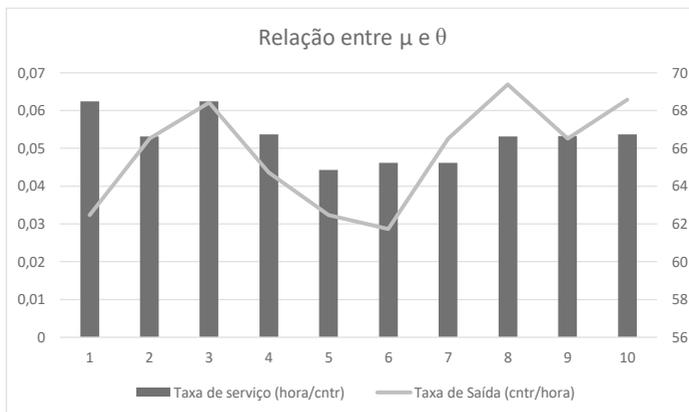
Comparando a melhor solução na primeira linha, cuja capacidade de *buffer* de 1870 contêineres e taxa de atendimento  $\theta = 62,473626$ , um aumento de 0,53% na área de espera (1880 contêineres) gera um aumento na taxa de atendimento de 6%, porém requer investimento em equipamentos ou pessoal qualificado para a obtenção de uma redução de 14,79% na taxa de serviço. Cabe nesse ponto uma análise dos custos em possíveis investimentos em contrapartida ao ganho na taxa de produtividade.

Diferentes análises podem ser feitas, como por exemplo, uma mesma taxa de produtividade  $\theta = 62,473626$  foi obtida para diferentes capacidades de *buffer*: 1870 e 2110 contêineres. Se a decisão a ser tomada estivesse pautada em quanto vale o investimento em área de espera para obtenção de um certo limiar para a taxa de produtividade, essa comparação mostraria que um aumento de cerca de 11% em área de espera gera a mesma taxa de atendimento.

Os gráficos 20 e 21 ilustram, respectivamente uma comparação entre os valores das taxas de serviço e das áreas de espera com os valores

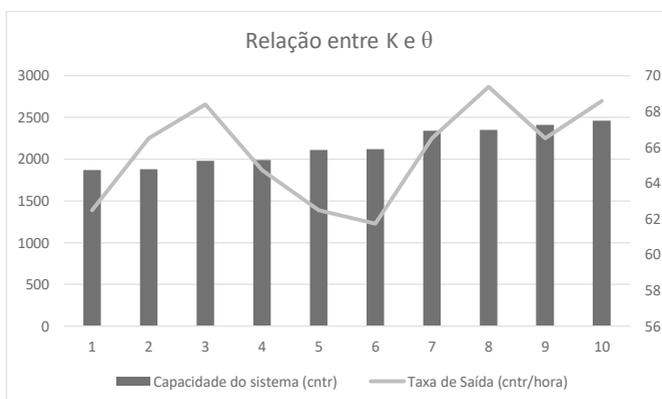
para a taxa de saída.

Figura 20: Comparação dos valores de  $\mu$  e  $\theta$



Fonte: Elaborada pela autora

Figura 21: Comparação dos valores de  $K$  e  $\theta$



Fonte: Elaborada pela autora

Como pode ser visto poucos são os casos em que um aumento na taxa  $\mu$  resulta em um aumento na taxa  $\theta$ , o mesmo pode ser observado em relação às áreas de espera  $K$ . A formulação multiobjetivo proposta evidencia a importância da análise do *trade-off* entre a taxa de saída/atendimento e a área de espera, fato que não fica claro numa formulação mono-objetivo.

A questão que se coloca nesse momento é o quanto investir em área de espera compensa um ganho ínfimo na taxa de saída. Ou ainda, um aumento na taxa de serviço, que pode significar um alto investimento em equipamentos ou contratação de pessoal qualificado, justifica um determinado aumento na taxa de saída. Questões como essas são devidamente respondidas pelo tomador de decisão com posse desses resultados.

Os resultados obtidos pelo algoritmo convergem para um conjunto de soluções aceitável, visto que a modelagem feita nesta tese levou em conta apenas o movimento de importação realizado pelo porto.

No próximo capítulo serão feitas as considerações finais do estudo e elencadas as possibilidades de pesquisas futuras sobre o tema abordado nesta tese.

## 6 CONSIDERAÇÕES FINAIS

### 6.1 PRINCIPAIS APONTAMENTOS

A proposta desta tese foi modelar uma operação portuária como uma rede de filas com alocação de *buffers*, sendo assim, procedeu-se com o desenvolvimento do modelo o qual necessitou da prévia modelagem das chegadas de contêineres ao porto, uma vez que a variável de interesse na rede deveria ser representada pela entidade contêiner e não pelo navio que os transporta.

Essa primeira modelagem das chegadas de contêineres significou uma intensa busca por referencial teórico, sendo este escasso, principalmente em termos de aplicação do modelo a problemas reais. Dessa forma, acredita-se que esta seja uma importante contribuição científica do trabalho realizado, tendo em vista os resultados que foram obtidos por essa modelagem.

Para a rede de filas, optou-se pelo método conhecido como Método de Expansão Generalizada, desenvolvido especificamente para esse tipo de rede com área de espera finita, em que devem ser considerados possíveis bloqueios entre as estações de trabalho e atrasos no tempo de serviço. O método não possui implementação disponível na literatura, dessa maneira foi necessário implementar o algoritmo a partir do referencial pesquisado.

Como consequência da proposta de modelagem de uma operação portuária, esta tese apresentou um problema de otimização multiobjetivo que consiste em maximizar a taxa de atendimento enquanto minimiza simultaneamente a área de espera e as taxas de serviço. Novamente, na literatura foi encontrado um algoritmo evolucionário multiobjetivo para esse tipo de problema, conhecido como NSGA-II. Baseado no algoritmo genético, após as operações padrão incorpora ainda o elitismo e a ordenação não-dominada, pertinentes ao tipo de problema, para gerar um conjunto de soluções de Pareto-ótimas.

Algumas soluções do NSGA-II foram filtradas para que se pudesse fazer uma análise dos dados obtidos, o que parecem estar de acordo com a modelagem que foi proposta, visto que nesta tese foi considerado apenas o movimento de importação e algumas aproximações precisaram ser feitas. Ainda assim, os resultados obtidos foram ao encontro do caso real estudado, dada a dimensão escolhida para o estudo.

Essa tese ainda representa uma importante contribuição acadêmica na área de Teoria de Filas, em que estudos aplicados com as carac-

terísticas aqui apresentadas ainda são pouco conhecidos.

## 6.2 SUGESTÕES PARA PESQUISAS FUTURAS

Algumas questões ficam pendentes e não esgotam o estudo proposto nesta tese, sugerindo assim investigações futuras. A primeira possibilidade de continuidade se dá no que tange à própria tomada de dados, repetir esse estudo com um universo maior de dados quantitativos reais permitirá a ampliação da veracidade do modelo, além de possibilitar uma efetiva calibração desses mesmos dados.

Outra oportunidade verificada diz respeito à modelagem de filas com chegadas em bloco dada a dificuldade encontrada na busca por referencial teórico. Esse estudo é de extrema aplicabilidade e pesquisas mais aprofundadas nessa área específica da Teoria de Filas representam uma importante contribuição científica.

Nesta tese apenas o movimento de importação foi levado em consideração, a ideia é que seja incorporado na mesma rede de filas também os movimentos de exportação e transbordo, que implicam, respectivamente, em fluxo inverso com múltiplas entradas e saídas, e laços nos nós da rede.

Quanto à possibilidade de obtenção de um elevado número de soluções geradas pelo algoritmo evolucionário multiobjetivo NSGA-II, é interessante aplicar uma busca pós-processamento, já que o algoritmo fornece um conjunto de Pareto-ótimo, mas não garante vasculhar todas as combinações possíveis para solução do problema.

Essas complementações no estudo visam a obtenção de um modelo completo e acurado para uma operação portuária vista como uma rede de filas, abrindo espaço para a pesquisa e colaborando com processos de tomada de decisão.

## REFERÊNCIAS

ANDRADE, E. L. **Introdução à pesquisa operacional**. [S.l.]: Rio de Janeiro, 1989.

ANTAQ. **Anuário Estatístico Anual**. [S.l.], 2012.

ANTAQ. **Anuário Estatístico Anual**. [S.l.], 2019.

AZEVEDO A. T.; CHAVES, A. A. N. L. S. Mini-curso: Introdução à otimização de operações portuárias. In: . [s.n.], 2014. Disponível em: <<http://www.din.uem.br/sbpo/sbpo2014/pdf/arq0001.pdf>>.

BARBOSA, A. M.; RIBEIRO, L. d. C.; ARANTES, J. M. d. O. Algoritmo genético multiobjetivo: Sistema adaptativo com eletismo. In: **9th Brazilian Conference on Dynamics Control and their Applications**. [S.l.: s.n.], 2010.

BICHOU, K. An empirical study of the impacts of operating and market conditions on container-port efficiency and benchmarking. **Research in Transportation Economics**, n. 42, p. 28–37, 2013.

BIELLI, M.; BOULMAKOUL, A.; RIDA, M. Object oriented model for container terminal distributed simulation. **European Journal of Operational Research**, v. 175, n. 3, p. 1731–1751, 2006. ISSN 0377-2217.

BITRAN, G.; MORABITO, R. Um exame de modelos de redes de filas abertas aplicados a sistemas de manufatura discretos: parte ii. **Gestão & Produção**, v. 2, n. 3, p. 297–321, 1995.

BITRAN, G.; MORABITO, R. Um exame dos modelos de redes de filas abertas aplicados a sistemas de manufatura discretos: parte i. **Gestão & Produção**, v. 2, n. 2, p. 192–219, 1995.

BRASIL. **Lei nº 8.630, Dispõe sobre o regime jurídico da exploração de portos organizados e das instalações portuárias e dá outras providências**. 26 fev. 1993.

CARLO, H. J.; VIS, I. F. A.; ROODBERGEN, K. J. Storage yard operations in container terminals: Literature overview, trends, and research directions. **European Journal of Operational Research**, n. 235, p. 412–430, 2014.

CARLO, H. J.; VIS, I. F. A.; ROODBERGEN, K. J. Transport operations in container terminals: Literature overview, trends research directions and classification scheme. **European Journal of Operational Research**, n. 236, p. 1–13, 2014.

CHANIN, R. **Avaliação Quantitativa de Sistemas**. Porto Alegre, 2005.

CLARKE, A. B.; DISNEY, R. L. **Probabilidade e processos estocásticos**. Rio de Janeiro: Livros Técnicos e Científicos, 1979.

COOPER, R. B. **Introduction to queueing theory**. 2. ed. New York: Elsevier North Holland, 1981.

CRUZ, F.; DUARTE, A.; WOENSEL, T. van. Buffer allocation in general single-server queueing networks. **Computers & Operations Research**, v. 35, n. 11, p. 3581 – 3598, 2008. ISSN 0305-0548. Part Special Issue: Topics in Real-time Supply Chain Management. Disponível em:

<<http://www.sciencedirect.com/science/article/pii/S0305054807000718>>.

CRUZ, F. R. et al. Throughput maximization of queueing networks with simultaneous minimization of service rates and buffers.

**Mathematical Problems in Engineering**, 2012. Disponível em: <<http://dx.doi.org/10.1155/2012/692593>>.

DEB, K. **Multi-objective using evolutionary algorithms**. England: John Wiley & Sons Ltd, 2001.

DEB, K.; AGRAWAL, R. B. Simulated binary crossover for continuous search space. **Complex Systems**, v. 5, p. 115–148, 1995.

DEB, K. et al. A fast and elitist multiobjective genetic algorithm. **IEEE Transactions on Evolutionary Computation**, v. 6, n. 2, p. 182–197, 2002.

DEMIR, L.; TUNALI, S.; ELIYI, D. The state of the art on buffer allocation problem: a comprehensive survey. **Journal of Intelligent Manufacturing**, Springer US, v. 25, n. 3, p. 371–392, 2014. ISSN 0956-5515. Disponível em: <<http://dx.doi.org/10.1007/s10845-012-0687-9>>.

EL-NAGGAR, M. E. Application of queueing theory to the container terminal at alexandria seaport. **Journal of Soil Science and Environmental Management**, v. 1, n. 4, p. 77–85, 2010.

FALCAO, V. A.; CORREIA, A. R. Eficiência portuária: análise das principais metodologias para o caso dos portos brasileiros. **Journal of Transport Literature**, v. 6, n. 4, p. 133–146, 2012.

FERNANDES, M. G. **Modelo econômico-operacional para análise e dimensionamento de terminais de contêineres e veículos**. Dissertação (Mestrado em Engenharia), 2001.

GOLDBARG, M. C.; LUNA, H. P. L. **Otimização combinatória e programação linear: modelos e algoritmos**. Rio de Janeiro: Campos, 2000. ISBN 85-352-0541-1.

GONTIJO, G. M. et al. Performance evaluation and dimensioning of  $g_i[x]/m/c/n$  systems through kernel estimation. **Mathematical Problems in Engineering**, 2011.

GROSS, D. et al. **Fundamentals of Queueing Theory**. 4. ed. New York, NY, USA.: Wiley-Interscience, 2009.

HILLIER, F. S.; LIEBERMAN, G. J. **Introdução à pesquisa operacional**. 8. ed. São Paulo: McGraw-Hill, 2006. ISBN 85-86804-68-1.

HUISMAN, T.; BOUCHERIE, R. J.; DIJK, N. M. v. A solvable queueing network model for railway networks and its validation and applications for the netherlands. **European Journal of Operational Research**, n. 142, p. 30–51, 2002.

KERBACHE, L.; SMITH, J. M. The generalized expansion method for open finite queueing networks. **European Journal of Operational Research**, n. 32, p. 488–461, 1987.

KITZMANN, D. **Ambiente portuário**. Rio Grande: Editora da FURG, 2010.

KLEINROCK, L. **Queueing Systems**. New York: John Wiley & Sons, 1975.

KONAK, A.; COIT, D. W.; SMITH, A. E. Multi-objective optimization using algorithms: A tutorial. **Reliability Engineering and System Safety**, v. 91, p. 992–1007, 2006.

LABETOULLE, J.; PUJOLLE, G. Isolation method in a network of queues. **IEEE Transactions on Software Engineering**, v. 6, n. 4, p. 373–381, 1980.

LARSON, R. C.; ODONI, A. R. **Urban operations research**. Belmont, U.S.A.: Dynamic Ideas, 2007. ISBN 0-9759146-3-4.

LAXMI, P. V.; GUPTA, U. Analysis of finite-buffer multi-server queues with group arrivals:  $G_i[x]/m/c/n$ . **Queueing Systems**, n. 36, p. 125–140, 2000.

LEMKE, N. **Análise das condições operacionais do Porto do Rio Grande utilizando a técnica de simulação**. Dissertação (Mestrado em Engenharia Oceânica), 2008.

MANITZ, M. Queueing-model based analysis of assembly lines with finite buffers and general service times. **Computers & Operations Research**, n. 35, p. 2520–2536, 2008.

MARINS, F. A. S. **Introdução à pesquisa operacional**. São Paulo: Cultura Acadêmica : Universidade Estadual Paulista, Pró-Reitoria de Graduação, 2011.

MIGUEL, P. A. o. C. **Metodologia de pesquisa em engenharia de produção e gestão das operações**. 2. ed. Rio de Janeiro: Elsevier: ABEPRO, 2012. ISBN 978-85-352-4891-3.

MURTY, K. G. et al. A decision support system for operations in a container terminal. **Decision Support Systems**, v. 39, n. 3, p. 309 – 332, 2005. ISSN 0167-9236. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0167923603001465>.

OLIVEIRA, C. T. d. **Modernização dos Portos**. 5. ed. São Paulo: Aduaneiras, 2011. ISBN 978-85-7129-592-6.

PAPADAPOULOS, C. T. et al. **Analysis and design of discrete part production lines**. New York: Springer, 2009. (Springer Optimization and its Applications, v. 31).

PRADO, D. S. d. **Teoria das filas e da simulação**. 4. ed. Nova Lima, MG: INDG Tecnologia e Serviços Ltda, 2009. ISBN 978-85-98254-40-1.

RUDENKO, O.; SCHOENAUER, M. A steady performance stopping criterion for pareto-based evolutionary algorithms. In: **6th International Multi-Objective Programming and Goal Programming Conference**. Hammamet, Tunisia: [s.n.], 2004.

SHAMBLIN, J. E.; STEVENS, G. T. **Pesquisa operacional: uma abordagem básica**. São Paulo, SP: Atlas, 1979.

SILVA, C. R. N. d.; MORABITO, R. Aplicação de modelos de redes de filas abertas no planejamento do sistema job-shop de uma planta metal-mecânica. **Gestão & Produção**, v. 14, n. 2, p. 393–410, 2007.

SMITH, J. M.; CRUZ, F. R. B.; WOENSEL, T. v. Topological network design of general, finite, multi-server queueing networks. **European Journal of Operational Research**, n. 2, p. 427–441, 2010.

STAHLBOCK, R.; VOB, S. Operations research at container terminals: a literature update. **OR Spectrum**, Springer-Verlag, v. 30, n. 1, p. 1–52, 2008. ISSN 0171-6468. Disponível em: <<http://dx.doi.org/10.1007/s00291-007-0100-9>>.

SUPRG. **Superintendência do Porto do Rio Grande**. Rio Grande - RS: [s.n.], 2019. Acesso em novembro de 2014. Disponível em: <<http://www.portoriogrande.com.br/site/index.php>>.

TAHA, H. A. **Pesquisa operacional: uma visão geral**. 8. ed. São Paulo, SP: Pearson Prentice Hall, 2008.

TECON. **TECON Rio Grande S/A**. Rio Grande - RS: [s.n.], 2014. Acesso em agosto de 2014. Disponível em: <<http://www.tecon.com.br/site/content/home/>>.

WINSTON, W. L. **Operations research applications and algorithms**. 4. ed. USA: Thomson, 2003.

YUN, W. Y.; CHOI, S. Y. A simulation model for container-terminal operation analysis using an object-oriented approach. **International Journal of Production Economics**, n. 59, p. 221–230, 1999.

ZHANG, C. et al. Storage space allocation in container terminals. **Transportation Research Part B**, n. 37, p. 883–903, 2003.

ZHANG, X.; ZENG, Q.; CHEN, W. Optimization model for truck appointment in container terminals. **Procedia - Social and Behavioral Sciences**, n. 96, p. 1938–1947, 2013.