



**UNIVERSIDADE DO ESTADO DO RIO GRANDE DO NORTE
UNIVERSIDADE FEDERAL RURAL DO SEMI-ÁRIDO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA
COMPUTAÇÃO**



KLEBER KROLL DE AZEVEDO SILVA

**GESTÃO EM TRANSPORTE MARÍTIMO: OTIMIZAÇÃO
DE REPOSICIONAMENTO DE CONTÊINERES VAZIOS
UTILIZANDO GRASP REATIVO COM APRENDIZAGEM
POR REFORÇO**

**MOSSORÓ - RN
2016**

KLEBER KROLL DE AZEVEDO SILVA

**GESTÃO EM TRANSPORTE MARÍTIMO: OTIMIZAÇÃO
DE REPOSICIONAMENTO DE CONTÊINERES VAZIOS
UTILIZANDO GRASP REATIVO COM APRENDIZAGEM
POR REFORÇO**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação - associação ampla entre a Universidade do Estado do Rio Grande do Norte e a Universidade Federal Rural do Semi-Árido, para a obtenção do título de Mestre em Ciência da Computação.

Orientador: Profº Francisco Chagas de Lima Júnior
Coorientador: Profº Carlos Heitor Pereira Liberalino

**MOSSORÓ - RN
2016**

**Catálogo da Publicação na Fonte.
Universidade do Estado do Rio Grande do Norte.**

Silva, Kleber Kroll de Azevedo.

Gestão Em Transporte Marítimo: Otimização De Reposicionamento De Contêineres Vazios Utilizando GRASP Reativo Com Aprendizagem Por Reforço – Mossoró, RN, 2016.

61 f.

Orientador (a): Prof. Francisco Chagas de Lima Júnior

Dissertação (Mestrado) Universidade do Estado do Rio Grande do Norte. Campus Central. Programa de Pós-Graduação em Ciência da Computação - associação ampla entre a Universidade do Estado do Rio Grande do Norte e a Universidade Federal Rural do Semi-Árido.

1. Reposicionamento de Contêiner. 2. Metaheurística. 3. GRASP. I. Lima Junior, Francisco Chagas de. II. Universidade do Estado do Rio Grande do Norte. III. Universidade Federal Rural do Semi-Árido. IV. Título

UERN/ BC

CDD 337

GESTÃO EM TRANSPORTE MARÍTIMO: OTIMIZAÇÃO DE REPOSICIONAMENTO DE CONTÊINERES VAZIOS UTILIZANDO GRASP REATIVO COM APRENDIZAGEM POR REFORÇO

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação para a obtenção do título de Mestre em Ciência da Computação.

APROVADA EM: 13/04/2016.

BANCA EXAMINADORA



Dr. Francisco Chagas de Lima Júnior-
UERN
Orientador



Dr. Carlos Heitor Pereira Liberalino-
UERN
Avaliador do Programa



Dr. Dario José Aloise - UERN
Avaliador do Programa



Dr. Gustavo Augusto Lima de Campos -
UECE
Avaliador Externo

Dedico este trabalho a minha família.

AGRADECIMENTOS

Agradeço a Deus, por tudo que tem proporcionado em minha vida.

A minha família, pelo incentivo e carinho incondicional.

Aos meus orientadores, Dr. Francisco Chagas de Lima Júnior e Dr. Carlos Heitor Pereira Liberalino, pela colaboração, paciência e pelos seus conhecimentos repassados durante todo o desenvolvimento do trabalho.

Aos colegas de trabalho do Departamento de Informática do Campus Central da Universidade do Estado do Rio Grande do Norte, pela excelente experiência profissional e bons momentos lá vividos.

Ao meu amigo sempre prestativo, Abílio Soares, e, em especial, ao Alex Trindade, pela sua atenção, presteza e colaboração para concretização deste trabalho.

À Universidade do Estado do Rio Grande do Norte e à Universidade Federal Rural do Semi-Árido por viabilizarem a minha qualificação acadêmica.

"O valor de um homem não se dá pelas roupas ou bens que possui...
E, sim, pelo caráter e beleza dos seus ideais."

Charles Chaplin

RESUMO

O reposicionamento de contêineres vazios entre os portos apresenta um custo elevado para os embarcadores e ocorre quando as quantidades de cargas exportadas e importadas são diferentes em uma determinada área de comércio marítimo, fazendo com que sobrem contêineres vazios em alguns portos e faltem em outros. Para atender à demanda dos clientes, as companhias podem desenvolver ações que causam prejuízo, como, por exemplo, repassar seus clientes para outras companhias, quando não houver disponibilidade de contêiner vazio ou alugá-lo, se disponível, no porto de demanda. Dessa forma, faz-se necessária a realização de estudos que apresentem soluções para o problema de reposicionamento de contêiner vazio, com o objetivo de minimizar os custos envolvidos nessa atividade. O presente estudo propõe uma solução algorítmica para o referido problema, utilizando método aproximativo (metaheurística), especificamente o algoritmo *Greedy Randomized Adaptive Search Procedure-GRASP Reativo*, conjugando a técnica de Aprendizado por Reforço (*Reinforcement Learning*) com o algoritmo *Q-Learning*. Este algoritmo é capaz de aprender diretamente a partir de experiências anteriores, o que minimiza a quantidade de cálculos a posteriori, otimizando, assim, o processamento computacional. Utilizou-se uma série de instâncias encontradas na literatura chamada *Carbin* na realização dos testes. Estes apontaram avanços (melhores soluções) em relação aos resultados obtidos por [Coelho \(2015\)](#) cujo trabalho também tratou o problema de reposicionamento de contêineres vazios, implementado com o algoritmo GRASP tradicional.

Palavras-chave: reposicionamento de contêiner vazio, metaheurística, GRASP, aprendizado por reforço.

ABSTRACT

Empty containers repositioning between ports has a high cost to shippers and occurs when the quantities of exported and imported cargoes are different in a certain maritime trade area, leaving empty containers at some ports and missing ones in others. To meet customer demand, companies can develop actions that cause losses, for example, pass on their customers to other companies when there is no availability of empty containers or rent these ones at demand port if they are available. Thus, making studies that present solutions for empty container repositioning problem is needed, in order to minimize the costs involved in this activity. This study proposes an algorithmic solution to this problem by using an approximation method (metaheuristics), specifically the *Reactive Greedy Randomized Adaptive Search Procedure*-GRASP algorithm with the Reinforcement Learning technique and *Q-Learning* algorithm. This algorithm can learn directly from the previous experience that minimizes amount of future calculations, optimizing computational processing. A series of instances found in the literature called Carbin was used on testing. These tests showed advances (best solutions) compared to the results obtained by Coelho (2015) whose work also dealt with the empty container repositioning problem, implemented with traditional GRASP.

Key-words: empty container repositioning, metaheuristic, GRASP, reinforcement learning.

LISTA DE FIGURAS

Figura 1 – Contêiner Padrão 20 pés	16
Figura 2 – Contêiner Padrão 40 pés	16
Figura 3 – Contêiner <i>Flat Rack</i>	17
Figura 4 – Contêiner Plataforma	17
Figura 5 – Contêiner <i>Open Top</i>	18
Figura 6 – Contêiner Ventilado	18
Figura 7 – Contêiner Refrigerado	19
Figura 8 – Contêiner Insulado	19
Figura 9 – Contêiner Tanque	20
Figura 10 – Eventos que compõem o ciclo de viagem de um contêiner. Fonte: (LAI; LAM; CHAN, 1995)	22
Figura 11 – Processo de interação entre um agente de aprendizagem por reforço e ambiente. Fonte: (SUTTON; BARTO, 1998)	35
Figura 12 – Exemplo do Problema de Fluxo Multiproduto com tempo discretizado	47
Figura 13 – Resultado para o contêiner tanque	50
Figura 14 – Resultado para o contêiner open-top	50
Figura 15 – Busca local 01	53
Figura 16 – Busca local 02	53

LISTA DE TABELAS

Tabela 1 – Tipos de contêineres em função da aquisição. Fonte:(LAI; LAM; CHAN, 1995)	20
Tabela 2 – Dados da instância exemplo	47
Tabela 3 – Característica das instâncias Carbin	54
Tabela 4 – Comparativo dos resultados GRASP Reativo com RL x GRASP x Cplex	56
Tabela 5 – Média, desvio-padrão e variância dos valores das soluções encontradas	56

LISTA DE ABREVIATURAS E SIGLAS

DF Dive-and-fix, página 28

GRASP Greedy Randomized Adaptive Search Procedure, página 29

PFM Problema de Fluxo Multiproduto, página 24

RF Relax-and-fix, página 27

TEU Twenty-foot Equivalent Unit, página 13

SUMÁRIO

	INTRODUÇÃO	9
1	ESTADO DA ARTE	12
1.1	Motivação	12
1.2	Problema	12
1.2.1	Transporte de contêineres	13
1.2.2	Problema de reposicionamento de contêineres vazios	21
1.3	Problema de Fluxo Multiproduto	24
1.4	Métodos de resolução	27
1.4.1	Métodos aproximativos	27
1.4.2	Greedy Randomized Adaptive Search Procedure - GRASP	29
1.4.3	GRASP Reativo	32
1.4.4	Aprendizagem por Reforço	34
1.5	Trabalhos relacionados	40
2	PROPOSTA	43
2.1	Metodologia	43
2.2	Modelo	43
2.2.1	Premissas do Modelo	44
2.2.2	Modelo Matemático	45
2.2.3	Exemplo Modelo Proposto	46
2.3	Políticas de ações do algoritmo Q-Learning	51
2.4	Algoritmo proposto	51
2.5	Fase de Busca Local	52
3	RESULTADOS	54
3.1	Instâncias utilizadas	54
3.2	Análise dos resultados numéricos	55
3.3	Dificuldades encontradas	57
4	CONCLUSÕES	58
	REFERÊNCIAS	59

INTRODUÇÃO

No ano de 1956, iniciou-se o uso de receptáculos para acondicionar mercadorias chamados *trailers*, que auxiliaram na unitização das cargas a serem transportadas, de forma que deram impulso ao conceito de multimodalidade em transporte, ou seja, este seria realizado em unidades padronizadas por meio das diferentes modalidades (rodoviária, ferroviária, marítima, fluvial e aérea) entre a origem e o seu destino. Ainda na década de 50, foram desenvolvidos os denominados contêineres de 20' e 40', como uma exigência de padronização para o transporte de mercadorias na Europa. O uso dos contêineres no transporte mundial de cargas foi iniciado, de fato, por volta dos anos 60, o que significou grande importância para o tráfego de mercadorias, devido ao aumento do volume de cargas transportadas. A utilização de contêineres facilita as operações de transporte de cargas em portos e outros pontos de transferência, permitindo maior agilidade e eficiência no manuseio, além de possibilitar a movimentação de múltiplas unidades de cargas simultaneamente (TEIXEIRA, 2011). Na década de 80, houve uma generalização da utilização dos contêineres para o transporte. Já na década de 90, houve um crescimento do volume de mercadorias movimentadas pelas várias modalidades, de maneira a suprir os diferentes padrões de exigência dos clientes no mundo (SAMPAIO, 2001).

O estudo de carregamento de contêiner aqui no Brasil começou ser realizado com mais ênfase há pouco tempo, por ter despertado interesses financeiros em empresas públicas e privadas, já que o transporte utilizando contêineres é oneroso e cobrado por container alugado, e não pela quantidade de itens que serão carregados. Por isso, a vantagem de aproveitar o volume do contêiner ao máximo (VENDRAMINI, 2007).

Segundo o Ministério do Desenvolvimento, Indústria e Comércio Exterior, a modalidade marítima é usada por mais de 90% do comércio exterior brasileiro, pois essa modalidade é a mais econômica em relação ao consumo de combustível, causando, assim, menor impacto ambiental, frente às alternativas aérea e rodoviária. O Ministério de Transportes de Québec afirma que, em 2007, para cada 1 litro de combustível, 1 tonelada de carga é transportada por mais de 241 km em navios; 95 km em trens e 28 km, em caminhões (GELAREH *et al.*, 2013).

Os embarcadores, clientes das companhias de navegação, solicitam o transporte de seus produtos por cabotagem¹, por exemplo, e deve ser realizado de modo unitizado (em módulos). Os contêineres, equipamentos responsáveis por essa modularização, podem pertencer ao embarcador ou à companhia de navegação. Normalmente, os contêineres pertencem à companhia de navegação e cabe aos embarcadores somente solicitar, uma ou mais vezes, o transporte de seus produtos dentro dos contêineres de

¹ Navegação entre portos de um mesmo país ou a distâncias pequenas, dentro das águas costeiras.

um porto de origem a um porto de destino. Consideram-se também solicitações de embarcadores cujos contêineres a eles pertencem. Nesse caso, cabe à companhia de navegação somente o transporte dos contêineres dos embarcadores em seus navios, sem a necessidade de disponibilização de contêineres para o atendimento (TEIXEIRA, 2011).

Considerando cenários em que a demanda no transporte de contêineres supera a oferta em certos portos dentro do horizonte de planejamento, as companhias de navegação precisam rejeitar uma ou mais solicitações de transporte. Além das limitações quanto à disponibilidade de contêineres e quanto ao embarque e desembarque dos contêineres nos portos de origem e de destino, há restrições dos navios para o atendimento dessas solicitações (TEIXEIRA, 2011).

Com relação ao transporte das cargas, este sofreu grandes mudanças nas últimas décadas, devido à globalização com maiores exigências e à competitividade entre as empresas. O conceito de *transporte* é apresentado como o traslado de uma mercadoria entre uma origem e um destino, o que demonstra a ligação direta com o comércio mundial. Existem seis tipos de modais de transporte: marítimo, fluvial, rodoviário, ferroviário, aéreo e por dutos (VIEIRA, 2013). Como cada modal possui características próprias, com suas vantagens e desvantagens, não há como afirmar, de forma abrangente, qual o melhor modal de transporte. Deve-se identificar todos os fatores que influenciam no transporte para a escolha do modal que melhor se apresenta para determinada situação. De fato, a criação dos contêineres trouxe revolução ao cenário do transporte mundial, todavia tem-se o problema de reposicionamento de contêineres vazios.

O presente trabalho visa a dar uma solução algorítmica do problema de reposicionamento de contêineres vazios em uma rede de transporte marítimo regular. Observa-se que ocorre este problema quando há uma diferença exorbitante no volume de cargas exportadas e importadas dentro de uma determinada área de tráfego comercial marítimo. Quando uma grande quantidade de contêineres cheios desembarca em um porto, caracterizando um processo de importação, após o seu descarregamento, é gerada a mesma quantidade de contêineres vazios em tal porto. Se este porto tivesse a mesma carga de exportação, poderia utilizar os mesmos contêineres, mas na prática sobram contêineres em determinados portos e faltam em outros.

A importância da abordagem desse problema está na prática constante do transporte de mercadorias por meio de contêineres. As exportações para outros continentes, na sua maioria, são feitas por meio marítimo, por ser mais cômodo, seguro e barato, fazendo ainda com que a carga tenha um maior aproveitamento. O transporte de mercadorias por contêiner proporciona maior segurança às empresas que o utilizam, além da sua rapidez, garantia de qualidade do produto transportado e principalmente formas de minimizar o custo de transporte. Se o carregamento do contêiner não for bem planejado, pode comprometer o valor final da mercadoria transportada ou até ser encarado como prejuízo para quem vendeu o produto que está sendo transportado

(VENDRAMINI, 2007).

Nesse contexto, faz-se necessário encontrar uma configuração ideal para suprir a demanda dos contêineres vazio entre os portos, levando-se em consideração restrição de capacidade de frota e de operação. Assim, esta pesquisa tem como objetivo apresentar uma solução algorítmica, utilizando metaheurística, que auxilie a resolução do problema do reposicionamento de contêineres vazios, considerando as variáveis e restrições envolvidas na minimização dos custos relacionados.

ORGANIZAÇÃO DA DISSERTAÇÃO

O trabalho está organizado da seguinte forma: no Capítulo 1, serão abordados a motivação, o problema e os métodos de resolução de problemas relacionados à temática, bem como as contribuições encontradas na literatura; no Capítulo 2, é apresentada a proposta para a resolução do problema, especificando a metodologia trabalhada, o modelo adotado e o algoritmo desenvolvido; e no Capítulo 3, são analisados os resultados obtidos na pesquisa e apresentadas as dificuldades encontradas durante o estudo. Em seguida, as conclusões deste trabalho são apresentadas e as referências que serviram como base para esta pesquisa.

1 ESTADO DA ARTE

1.1 MOTIVAÇÃO

Nas últimas décadas, observou-se um grande crescimento do comércio naval e o volume de cargas transportadas em contêineres ampliou-se de forma a se tornar a fração mais importante do tráfego de mercadorias.

As companhias navais competem no mundo inteiro pelo fornecimento de serviços de transporte em contêineres. Dessa forma, as empresas precisam melhorar o atendimento ao cliente e ao mesmo tempo reduzir seus custos para se tornarem mais competitivas e lucrativas. A margem de lucro das companhias pode ser incrementada controlando e reduzindo suas despesas ou respondendo de forma ágil às demandas dos clientes. Um componente importante do custo operacional total de uma companhia de navegação é decorrente do reposicionamento de contêineres vazios entre vários portos. Em função disso, é de vital importância haver uma política adequada para gerenciar a movimentação dos contêineres. Para ser mais eficiente, a companhia precisa também encontrar maneiras de evitar que negócios deixem de ser contratados devido à indisponibilidade de contêineres vazios quando requisitados pelos clientes (Bandeira, 2005 apud Lai, Lam e Chan, 1995).

O problema de reposicionamento de contêineres vazios é frequente no comércio naval, sendo consequência do desequilíbrio entre importação e exportação, decorrente das diferentes necessidades econômicas de cada região. Por outro lado, o sistema de movimentação de contêineres cheios está diretamente vinculado ao sistema de controle de contêineres vazios, havendo uma relação direta entre as entradas e as saídas de cada sistema.

Desse modo, faz-se necessário estudar o problema de reposicionamento de contêineres vazios, a fim de se desenvolver novos métodos que otimizem a movimentação desses contêineres, de forma a encontrar soluções de menor custo para a empresa.

1.2 PROBLEMA

O processo de reposicionamento de contêineres no oceano não custa menos que 15 bilhões de dólares, o equivalente a 27% do custo total de funcionamento da frota mundial, e 20% de todos estes movimentos estão relacionados aos contêineres vazios. A gestão eficiente dos contêineres vazios é um problema importante no setor de transporte, possuindo efeito econômico e também de impacto ambiental e sustentável, uma vez que a redução da movimentação de contêineres vazios irá reduzir o consumo de combustível, congestionamentos e emissões de poluentes (SONG; CARTER, 2009).

No período de 1990 a 2010, o tráfego de contêiner vazio aumentou de 17,8 milhões para 116,9 milhões de TEUs¹, ao mesmo tempo que o tráfego de manipulação mundial aumentou de 87,9 milhões para 548,5 milhões de TEUs (TRAN; HAASIS, 2013).

Nesse contexto, observam-se a necessidade e a importância de implementação de procedimentos que visem à otimização das operações de reposicionamento de contêineres vazios, minimizando os custos envolvidos. Esta pesquisa aborda o problema do reposicionamento dos contêineres vazios, a fim de encontrar uma configuração ideal (disposição de contêineres vazios) para atender à demanda entre os portos, sob restrição de capacidade de frota e de operação, propondo uma estratégia de solução baseada na metaheurística GRASP Reativo, associado ao aprendizado por reforço, usando o algoritmo *Q-Learning*.

Nas próximas subseções, são descritos o problema do reposicionamento de contêineres vazios nos portos e alguns conceitos imprescindíveis para o desenvolvimento da pesquisa.

1.2.1 Transporte de contêineres

Segundo Preston e Kozan (2001), os contêineres apareceram pela primeira vez como método de transporte de cargas ao final dos anos 30. Eles foram introduzidos como vagões desmontáveis que podiam ser removidos de seus chassis e colocados em chassis de outros trens ou de algum transporte rodoviário. Esse sistema foi utilizado principalmente para eliminar o recarregamento que ocorria quando era necessária a troca da carga entre trens ou de caminhão para trem. Isso levou a um aumento da segurança, uma redução dos custos de transporte e facilitou a entrega de produtos de porta a porta. A indústria náutica rapidamente reconheceu o apelo econômico da utilização de contêineres e empregou-os como meio de carga.

O transporte marítimo em contêineres surgiu por volta dos anos 1950 e é um dos meios de transporte mais importantes no comércio internacional (TRAN; HAASIS, 2013). Seu início se deu quando Malcolm McLean, proprietário de uma empresa de caminhões em Nova York, começou um serviço regular utilizando um navio tanque readaptado, entre os Estados Unidos e Porto Rico. Algumas vantagens de utilizar o contêiner foram percebidas naquela época: rapidez de carga e descarga do navio e a facilidade para empregar a multimodalidade. Porém, os conservadores da época resistiram à utilização dos contêineres inicialmente, argumentando que haveria altos custos para substituir a infraestrutura e equipamentos existentes, o que a tornaria economicamente inviável, haja vista o crescimento da utilização no decorrer dos anos (VIEIRA, 2013).

O contêiner é um cofre de carga móvel, ou seja, provido de dispositivos que permitem sua manipulação, desenvolvido para o transporte multimodal, com volume

¹ Twenty-foot Equivalent Unit - É uma medida padrão, utilizada para medir o volume de um contêiner, equivalente a um contêiner de 20 pés.

interno de, no mínimo, 1 m³, e que pode ser usado várias vezes². Oferece também outras vantagens, entre elas a padronização do carregamento, aumento da segurança e facilidade de manipulação para o transporte intermodal. Por outro lado, apresentam problemas específicos, tais como: dimensionamento e gerenciamento de frota, decisão entre propriedade ou arrendamento dos contêineres, reposicionamento de contêineres que foram liberados e preparação anterior a um carregamento. O alto custo de aquisição, manutenção, manipulação e transporte de contêineres faz com que o seu gerenciamento seja um problema de grande relevância (Bandeira, 2005 apud Dejax e Crainic, 1987). O aço é o material de construção mais utilizado, devido à alta resistência e seu custo ser relativamente baixo. Existem também contêineres de alumínio e de madeira com fibra de vidro, para cargas leves, devido ao seu peso (VIEIRA, 2013).

O uso de contêineres facilitou muito as operações de manipulação em portos e em outros pontos de transferência, aumentando a eficiência e a velocidade do transporte, e permitindo que múltiplas unidades de carga fossem manipuladas simultaneamente. Os terminais portuários mudaram radicalmente após o surgimento dos contêineres, incluindo a introdução de equipamento especializado para sua manipulação, alterações na área de armazenamento e nos métodos de estocagem (técnicas de empilhamento). A padronização dos contêineres permitiu também a expansão de sistemas de transporte intermodal, ou seja, a carga acondicionada em contêineres pode ser transportada por múltiplas vias (rodoviária, ferroviária ou naval), com maior facilidade, devido a seu formato e dimensões padronizadas (BANDEIRA, 2005).

Pode-se ter uma noção da evolução do transporte de contêineres pela ampliação do porte dos navios. A primeira geração de navios porta-contêineres, no final da década de 60, tinha capacidade de cerca de 400 TEUs. No final dos anos 70, navios porta-contêineres tinham capacidade de 2000 TEUs, chegando na década de 90 a mais de 4000 (BANDEIRA, 2005 apud TALEB-IBRAHIMI, CASTILHO e DAGANZO, 1993). Atualmente, o MSC Oscar, fabricado pela Mediterranean Shipping Company, é o maior navio cargueiro, conseguindo comportar 19.224 TEUs.

Desde o princípio, o processo de "containerização" trouxe inúmeros benefícios: entrega segura, maior rapidez, frete oceânico mais barato e adequação dos cofres para o transporte multimodal. As vantagens no uso dos contêineres podem ser atribuídas a cada uma das partes envolvidas na cadeia de transporte (VIEIRA, 2013):

- Embarcadores: diminuição da manipulação das mercadorias, incidência de menos avarias e roubos.
- Transportadores terrestres: maior agilidade no transporte e menores reclamações em função da redução de roubos e avarias.

² conceito estabelecido pela ISO (International Organization for Standardization.)

- Armadores³: redução do tempo de estadia nos portos, em decorrência da maior agilidade em carga/descarga, implicando maior economia de escala.
- Portos: redução do número de escalas, devido à estratégia dos armadores aumentarem o tamanho dos navios.

Em contrapartida, verificam-se algumas restrições no uso de contêineres: espaços perdidos dentro da unidade de carga; exigência de equipamentos de alto investimento para a movimentação da unidade de carga; necessidade de reposicionamento dos contêineres vazios; pagamento de aluguel de contêiner; custos de reparo, reposição e retorno dos contêineres.

O transporte de contêineres é feito de forma unitizada e modular, permitindo um melhor aproveitamento dos espaços nas embarcações, o que torna uma das maiores vantagens no transporte de mercadorias. Além disso, facilita a padronização e a automação de todas as operações envolvidas em sua movimentação, fazendo com que aumente a eficiência nos pontos de transferência, terminais e armazéns de contêineres, além de proporcionar economia de tempo e redução de custos para toda a cadeia (TEIXEIRA, 2011).

No entanto, o aumento da demanda de cargas a serem transportadas por contêineres ocasiona, em certos pontos e períodos do horizonte de planejamento, incapacidades de atendimento em virtude das limitações de recursos logísticos disponíveis para o transporte de carga containerizadas. Com isso, as companhias de navegação responsáveis por administrar a alocação e transporte de contêineres precisam definir o atendimento específico da demanda. Conforme Lai, Lam e Chan (1995), para maior eficiência, a companhia de navegação precisa evitar que negócios deixem de ser contratados devido à indisponibilidade de contêineres vazios quando requisitados pelos embarcadores (TEIXEIRA, 2011).

Em virtude do desequilíbrio de demandas, que cria certos desafios logísticos no gerenciamento de contêineres vazios (LAI; LAM; CHAN, 1995), há necessidade de, além das decisões envolvendo seleção das cargas a serem transportadas, tomada de outras decisões de modo integrado, envolvendo reposicionamento de contêineres vazios que devem ser reutilizados inúmeras vezes por diversas cargas em diferentes períodos do planejamento. Na seleção das cargas a serem transportadas, busca-se a maximização da receita gerada com o atendimento do transporte das cargas. Em contrapartida, busca-se redução dos custos logísticos envolvidos, como no reposicionamento e outras situações, considerando que o contêiner é um equipamento dispendioso e não descartável (TEIXEIRA, 2011).

Os contêineres podem ser classificados de acordo com suas características físicas ou tipo de contrato de aquisição (SAMPAIO, 2001), sendo essas duas variações incluídas

³ Proprietários de embarcações de grande porte, específicas para o transporte containerizado.

no custo total da movimentação do contêiner disponível no modelo matemático deste estudo, ou seja, levou-se em consideração tanto o tipo de contêiner quanto a forma de aquisição do contêiner.

Com relação às características físicas, existem tipos diferentes que variam de acordo com sua utilização e tamanho, sendo que este último teve que passar por uma padronização de forma a se adequar à maioria dos navios existentes. As dimensões dos contêineres foram padronizadas pela ISO, no ano de 1968, em 10, 20, 30 e 40 pés. Os mais utilizados são os de 20 e 40 pés, associados à utilização da sigla TEU. Os tipos mais comuns são especificados abaixo:

- **Padrão 20 pés:** este tipo de contêiner é um dos mais utilizados para o transporte dos diversos tipos de carga: materiais secos, granéis, embalados, úmidas e substâncias líquidas. Recomenda-se o padrão de 20 pés também para cargas pesadas de menor volume.



Figura 1 – Contêiner Padrão 20 pés

- **Padrão 40 pés:** utiliza-se este contêiner para o transporte de cargas de grande volume, porém material leve.



Figura 2 – Contêiner Padrão 40 pés

- **Flat Rack:** indica-se este tipo de contêiner para cargas que possuem formas diversas, fora dos padrões convencionais, tais como máquinas para agricultura,

barcos, geradores, veículos etc. Com relação às dimensões, o tipo *Flat Rack* pode possuir tanto 20 quanto 40 pés.



Figura 3 – Contêiner *Flat Rack*

- **Plataforma:** semelhante ao *Flat Rack*, este tipo de contêiner é indicado para o transporte de cargas com formas não convencionais, sendo colocado no topo do navio, sobre os demais. As cargas compridas, largas, altas, sem formas regulares ou com problemas de acondicionamento são exemplos de material transportado no contêiner plataforma.



Figura 4 – Contêiner Plataforma

- **Open Top:** utiliza-se este tipo de contêiner quando a carga tem forma irregular e só pode ser inserida por cima de sua estrutura (abertura na parte superior). A proteção da carga é feita por meio de uma lona fixada ao topo do contêiner. Alguns exemplos de cargas que podem ser transportadas são tosas de madeira, empilhadeira, vidro, máquinas para a construção e agricultura, entre outros. O contêinter *Open Top* pode possuir tanto 20 quanto 40 pés.



Figura 5 – Contêiner *Open Top*

- **Ventilado de 20 pés:** as cargas que precisam de ar puro, ventilação natural, podem ser transportadas neste tipo de contêiner. Foram projetados para cargas perecíveis como café, cacau, feijão, grãos, fumo etc.



Figura 6 – Contêiner Ventilado

- **Refrigerado de 20 pés:** recomenda-se este contêiner para cargas perecíveis ou não, que não podem ficar expostas a temperaturas elevadas. É capaz de manter a temperatura interna em 20°C negativos, mesmo que a externa chegue a 40°C positivos.



Figura 7 – Contêiner Refrigerado

- **Insulado de 20 pés com *clip-on*:** os contêineres tipo insulado são separados e resfriados por ar circulante gelado no controle de temperatura. Assemelha-se ao contêiner refrigerado, diferenciando-se por a sua refrigeração não ser autônoma. O contêiner é ligado a um sistema de refrigeração do tipo *clip-on* existente no navio. Indicado para transporte de cargas altamente perecíveis como frutas, peixes, carnes etc.



Figura 8 – Contêiner Insulado

- **Tanque:** é indicado para transporte de líquidos de qualquer tipo, desde alimentícios até aditivos e diversos materiais químicos, à prova de corrosão e de vazamento e até com controle de temperatura.



Figura 9 – Contêiner Tanque

Os contêineres são modulares, sendo que se considera o contêiner de 20' (vinte pés) de comprimento como um módulo, ou seja, equivalente a 1 TEU. Todas as demais medidas de contêineres consideram a medida de equivalência em TEU. Por exemplo, o contêiner com medida de 40' (quarenta pés) pode ser definido como um contêiner de 2 TEUs, ou seja, equivalente a 2 módulos de 20' (vinte pés), sempre em relação ao seu comprimento.

Uma frota de contêineres, de acordo com a sua forma de aquisição, pode ser: própria, alugada ou mista, conforme detalhado na Tabela 1. De uma maneira geral, aqueles que adquirem tal frota são clientes (importadores/exportadores) ou transportadores. No caso dos clientes, estes escolhem a forma de aquisição da frota, de acordo com o tipo de mercadoria a ser movimentada. No caso dos transportadores, a escolha torna-se mais complexa, uma vez que estes são os intermediários dos clientes. Por exemplo, se sempre houver demanda por transporte de cargas densas como minério, os transportadores devem dispor de contêineres de meia altura. Por outro lado, se no movimento em questão houver incidência constante de blocos de granito, eles precisam adquirir uma boa quantidade de contêineres plataforma.

Tabela 1 – Tipos de contêineres em função da aquisição. Fonte:(LAI; LAM; CHAN, 1995)

Tipo de Contêiner	Caracterização
Próprio	O contêiner é armazenado nos depósitos do armador ou de terceiros, caso não seja utilizado, e estes cobram uma taxa diária de armazenagem.
Alugado a curto prazo (menos de 3 anos)	O contêiner é devolvido para a companhia de <i>leasing</i> em operação denominada <i>drop-off</i> quando não é utilizado. Esta devolução é sujeita a um prazo limite de aluguel que varia de porto para porto. Pode-se armazenar o contêiner em depósitos para uso posterior, caso não seja utilizado.
Alugado a longo prazo (mais de 3 anos)	A devolução do contêiner não é permitida dentro de um prazo estabelecido que varia normalmente de 3 a 5 anos, dependendo do contrato de aluguel.

1.2.2 Problema de reposicionamento de contêineres vazios

É cediço que a padronização das cargas em contêineres influenciou o comércio marítimo, dentro do escopo econômico e de segurança no transporte. Tal fato aumentou consideravelmente o volume de cargas transportadas em todas as modalidades, que levou à necessidade de se obter novas formas mais eficientes para integrar as inúmeras formas de transporte e otimizar o reposicionamento dos contêineres (cheios ou vazios) no sistema de transporte marítimo.

O problema do reposicionamento de contêineres vazios consiste em realocar os que estiverem disponíveis e determinar a quantidade que deve ser arrendada, a fim de atender à demanda de clientes, dentro de um determinado período de tempo. Apesar de contêineres cheios terem prioridade de alocação na embarcação, deve-se levar em consideração que contêineres vazios não podem ficar ociosos, pois são necessários para futuros carregamentos.

As quantidades de contêineres descarregados e carregados em um navio podem ser totalmente diferentes em determinado porto específico, os operadores das linhas de navegação frequentemente precisam redistribuir seus contêineres vazios ou arrendar contêineres de companhias de *leasing*, a fim de atender à demanda dos clientes. Todavia, a disponibilidade de contêineres é sujeita a vários parâmetros de incerteza: a demanda nos portos, o tempo de retorno de contêineres dos clientes consignadores e a capacidade do navio para contêineres vazios. Diante dessas incertezas, os operadores tendem a operar de modo conservador, o que acarreta fortes consequências econômicas para as regiões envolvidas no comércio internacional (BANDEIRA, 2005 apud CHEUNG e CHEN, 1998).

O problema para reposicionar contêineres vazios pode ser descrito da seguinte forma: supõe-se que um cliente solicita um contêiner vazio de um depósito, que normalmente está localizado próximo a um porto. Após carregá-lo, o cliente retorna o contêiner ao porto poucos dias antes da data prevista para autorizar a partida. O contêiner será então colocado no navio, que irá para o porto de destino, e, em seguida, será direcionado ao destinatário, que irá descarregar o contêiner e devolvê-lo vazio a um pátio de contêineres, a um depósito ou diretamente ao porto. Esse contêiner vazio poderá então ser utilizado para satisfazer a demanda deste porto, ou armazenado como estoque. Além disso, em um porto, quando os contêineres vazios disponíveis não são suficientes para atender à demanda, o operador da companhia naval pode necessitar arrendar contêineres de fornecedores ou realocar contêineres vazios de outros portos. Entretanto, o reposicionamento de contêineres vazios está sujeita à existência de espaço disponível no navio, onde geralmente contêineres cheios têm prioridade mais alta (BANDEIRA, 2005).

Como ilustra a Figura 10, as operações de exportação/importação originam um ciclo, cujo início e fim se dá em um depósito da empresa ou de uma companhia de *leasing*.

Quando um cliente demanda por um contêiner, é feita a solicitação para a empresa que manda um contêiner vazio de seu depósito (ou depósito *leasing*) para o cliente. Após o cliente fazer o carregamento, o contêiner cheio é levado para o depósito, onde espera alguns dias para ser levado ao terminal. Após chegar ao terminal o contêiner cheio é carregado no navio e segue viagem. O processo inverso é composto de duas possibilidades: se o contêiner for cheio, este segue do terminal para o depósito, em seguida para o cliente e finalmente segue vazio do cliente para o depósito; caso o contêiner descarregado seja vazio, segue do terminal para o depósito. Como se pode observar, as taxas de saída de unidades vazias do depósito da empresa para o cliente e de chegada dessas unidades dos clientes até o depósito, são aleatórias. Os contêineres cheios para exportação, partindo do cliente, devem ir diretamente ao terminal e vice-versa (LAI; LAM; CHAN, 1995). O resultado final desse ciclo de retorno vem a ser um montante de contêineres vazios em estoque, o que vai caracterizar cada terminal ou porto por seu tipo de *imbalance* ou seja, pelo excesso ou carência dessas unidades vazias.

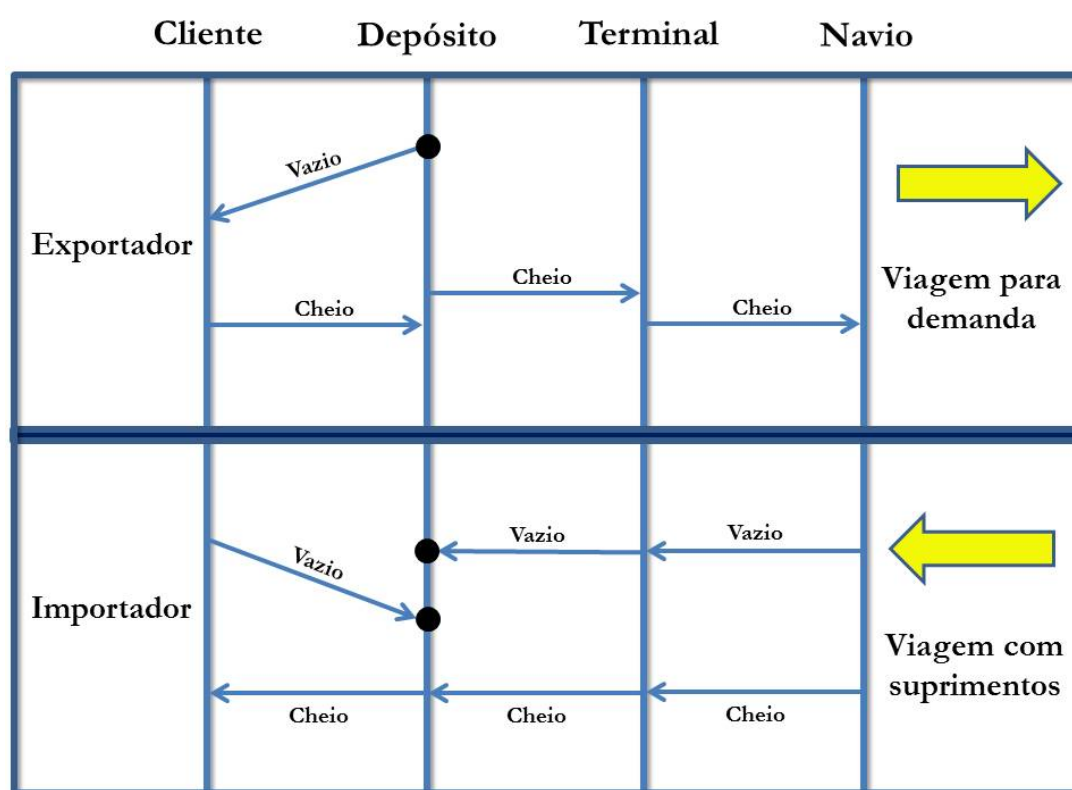


Figura 10 – Eventos que compõem o ciclo de viagem de um contêiner. Fonte: (LAI; LAM; CHAN, 1995)

Alguns portos são basicamente de importação e outros são de exportação, isto ocasiona um desequilíbrio (*imbalance*) no atendimento às demandas dos clientes de tal forma que faltam contêineres vazios em alguns deles e, em outros, há oferta maior que a demanda.

Nas situações em que há falta de contêineres vazios para o atendimento de determinada carga no porto de origem na data correspondente à disponibilização dos contêineres de propriedade da companhia de navegação, a decisão de rejeição da carga (parcial ou total) poderá se tornar necessária. Então, a companhia de navegação deve decidir, de maneira integrada, a seleção das cargas e o reposicionamento dos contêineres vazios, com objetivo de permitir o atendimento das demandas. Assume-se que os contêineres vazios somente podem ser transportados nas mesmas embarcações que transportam os contêineres cheios. Dessa maneira, o espaço utilizado nos navios deverá ser compartilhado entre contêineres cheios e vazios, não havendo viagens específicas dedicadas exclusivamente para o reposicionamento de contêineres vazios. Tendo em vista que o contêiner vazio ocupa espaço físico na embarcação, reduzem-se as oportunidades de utilização desse espaço para o transporte de contêineres cheios, responsáveis pela geração da receita (TEIXEIRA, 2011).

Nesse contexto, deve-se equilibrar entre duas situações: contrair custos com aluguel e estoque e o risco de não ter contêineres disponíveis suficientes quando necessários. Desta forma, o problema de reposicionamento de contêineres vazios pode ser representado pelo seguinte questionamento: "qual a quantidade de contêineres vazios que a empresa deve manter em cada porto e como reposicioná-los de modo a minimizar o custo total?".

Lai, Lam e Chan (1995) afirmam que as companhias têm como possíveis alternativas para atender à demanda dos clientes:

- Carregar contêineres vazios extras no navio;
- Alugar contêineres vazios no porto de demanda, caso estejam disponíveis;
- Encaminhar seus clientes a algum concorrente, quando não houver disponibilidade de contêineres vazios.

Estas possibilidades ocasionam perda de eficiência para a empresa, podendo resultar em prejuízo para as companhias navais e atendimento ineficiente ao cliente. Na realidade, a companhia tentará convencer o cliente a aguardar a disponibilidade de contêineres para que a carga seja transportada no próximo ciclo do navio.

Vários dias antes da data prevista para a partida de um navio, deve haver disponibilidade de contêineres vazios, para que os clientes possam buscá-los. Em geral, uma companhia atende um grande número de clientes em vários locais. Entretanto, o número exato de contêineres vazios necessários em um determinado período de tempo é altamente incerto. Por outro lado, o suprimento de contêineres vazios depende do tempo de retorno destes contêineres vindos dos consignadores, sobre os quais a companhia não tem controle. A situação é complicada, ainda mais pela prática comum de acumulação de contêineres em algum pátio (em terra), para então enviá-los em lotes

por trem, reduzindo custos (BANDEIRA, 2005 apud CHEUNG e CHEN, 1998). No entanto, armazenar contêineres vazios em portos envolve custo de armazenagem.

Caso a quantidade de contêineres vazios não seja suficiente para atender à demanda, faz-se necessário estabelecer critérios para definir a reserva de contêineres nos portos associados. Em alguns destes, poderá haver um número de contêineres insuficientes, o que deverá ser complementado com o aluguel de contêineres vazios por um curto prazo.

Diversos modelos foram propostos para resolver o problema de reposicionamento de contêineres, porém a característica de ser estocástico e o tamanho do problema tornam difícil o desenvolvimento e teste de métodos que encontrem soluções eficientes. Nos processos estocásticos, as decisões têm de ser tomadas antes que os parâmetros aleatórios sejam observados. Além disso, após os valores tornarem-se disponíveis, devem ser tomadas novas ações. O custo esperado depende da decisão atual e, geralmente, envolve problemas de otimização baseados em expectativas, tornando o problema muito difícil de solucionar.

Embora os movimentos no transporte de carga sejam levados em consideração no sistema de planejamento e no modelo de roteamento, apenas os movimentos de contêineres vazios aparecem como variáveis de decisão nos modelos de reposicionamento de contêineres. Tal fato ocorre porque os contêineres carregados são embarcados para fora da rede terrestre, desaparecendo para propósitos de planejamento. Os modelos devem ser projetados como parte de um sistema de suporte à decisão de planejamento de curto prazo para contêineres vazios e cheios. Para sua implementação, assume-se que deva existir um sistema de informação confiável. As previsões de suprimento e demanda de contêineres vazios devem ser precisas e têm de ser constantemente atualizadas, o que é crítico para o sucesso da implementação do modelo de reposicionamento.

1.3 PROBLEMA DE FLUXO MULTIPRODUTO

Apresentado inicialmente por Ford e Fulkerson (1962) e Hu (1963), o Problema de Fluxo Multiproduto (PFM) é caracterizado por vários produtos que compartilham os arcos em uma rede capacitada. A formulação típica destes problemas consiste em determinar o fluxo de custo mínimo, de forma a atender as restrições de capacidade e de conservação de fluxo da rede, para os vários produtos envolvidos. Portanto, busca-se satisfazer um único objetivo, sem, no entanto, proceder a qualquer análise relativa à estrutura da rede em questão. Há duas estruturas destes problemas: a primeira considera que as variáveis de fluxo são variáveis inteiras; a segunda assume que estas variáveis são binárias e, portanto, o fluxo de produtos é indivisível desde o nó de origem até o nó de demanda. Neste caso, cada produto possui um par de nós origem-destino,

correspondentes ao nó de oferta e ao nó de demanda deste produto.

Em geral, envolve um grande número de variáveis e restrições, além de possuir uma extensa variedade de aplicações, incluindo:

- Sequenciamento de carga;
- Roteamento de mensagens em redes de comunicação;
- Roteamento/sequenciamento de suprimentos militares;
- Localização de depósitos;
- Planejamento de sistemas de tráfego urbano e de produção de bens de consumo;
- Sequenciamento de operações em refinarias de petróleo.

A maior parte das aplicações está concentrada nas áreas de telecomunicação e de transporte.

Os Problemas de Fluxo Multiproduto, segundo [Albuquerque \(2011\)](#), podem ser classificados em três classes: lineares, inteiros e binários, sendo este último a categoria trabalhada neste estudo.

No caso dos problemas lineares, as variáveis de decisão, correspondentes ao fluxo de cada produto em cada arco, podem assumir valores não inteiros, ou seja, as variáveis envolvidas assumem valores pertencentes ao conjunto dos números reais. Na formulação inteira, cada variável de decisão assume um valor inteiro, o que significa que cada produto é indivisível, porém o fluxo de um determinado produto não precisa seguir uma única rota entre a origem e destino.

No problema de fluxo multiproduto binário, todas as unidades de um produto devem fazer uso de apenas uma rota, sendo um caso particular do problema inteiro, também conhecido na literatura como problema de fluxo multiproduto não-bifurcado.

Uma rede de fluxo multiproduto pode ser representada por um grafo $G = (N, A)$, onde N representa o conjunto de $n \in N$ nós da rede, enquanto A representa o conjunto de $a \in A$ arcos direcionados da rede. A rede também possui a variável K , onde representa os $k \in K$ produtos. Associado a cada arco $a \in A$, há um par de nós (i, j) , com uma capacidade e um custo para que cada unidade de cada produto trafegue pelo referido arco. Uma observação importante é que o custo para a utilização de um dado arco não é, necessariamente, o mesmo para todos os produtos, podendo variar pelo valor de cada um deles. Neste estudo, os produtos são representados pelos diferentes tipos de contêineres.

Para a formulação do problema, são estabelecidas as seguintes variáveis:

- c_{ij}^k representa o custo unitário do produto k no arco (i, j) ;

- x_{ij}^k representa a variável de decisão, que é o fluxo do produto k pelo arco (i, j) .

$$x_{ij}^k = \begin{cases} 1, & \text{se o fluxo do produto } k \text{ faz uso do arco } (i, j), \text{ no tempo } t; \\ 0, & \text{caso contrário.} \end{cases}$$

- $u_{i,j}$ representa a capacidade do arco (i, j) ;
- d^k representa a demanda do produto k ;
- b_i^k indica se o nó i é oferta ou demanda para o produto k . Considerando que cada produto k está associado a uma origem e destino, então b_i^k pode assumir os seguintes valores:

$$b_i^k = \begin{cases} d^k, & \text{se o nó } i \text{ é oferta para o produto } k; \\ -d^k, & \text{se o nó } i \text{ é demanda para produto } k; \\ 0, & \text{neste caso, } i \text{ é um nó de transbordo.} \end{cases}$$

Desta forma, a formulação binária para o problema de fluxo mínimo multiproduto é dada por:

$$\text{Min}(Z) = \sum_{k \in K} \sum_{(i,j) \in A} c_{ij}^k d_j^k x_{ij}^k \quad (1.1)$$

$$\text{Suj. a} : \sum_{(i,j) \in A} x_{ij}^k - \sum_{(j,i) \in A} x_{ji}^k = b_i^k, \forall i, j \in N, \forall k \in K \quad (1.2)$$

$$\sum_{k \in K} d_j^k x_{ij}^k \leq u_{ij}, \forall (i, j) \in A, \forall k \in K \quad (1.3)$$

$$x_{ij}^k \in \{0, 1\}, \forall k \in K, \forall (i, j) \in A \quad (1.4)$$

Sendo:

- A expressão (1.1) representa a função objetivo a ser minimizada, que indica os valores de custo de passagem dos k produtos pelos arcos (i, j) ;
- A expressão (1.2) representa as restrições de conservação de fluxo, que são responsáveis pelo gerenciamento do fluxo na rede;
- A expressão (1.3) representa as restrições de capacidade, sendo responsável por limitar a capacidade de cada arco;
- A expressão (1.4) representa a integridade das variáveis envolvidas.

1.4 MÉTODOS DE RESOLUÇÃO

Os problemas de otimização combinatória são computacionalmente resolvidos por algoritmos e podem ser definidos como uma "sequência bem definida de procedimentos computacionais que levam uma entrada a ser transformada em uma saída". Um dos objetivos de um método de otimização é solucionar o problema proposto de forma rápida e econômica. O termo rápido está relacionado à realidade do fim prático do algoritmo, enquanto econômico responde pela realidade da limitação dos recursos humanos e computacionais disponíveis. Nem sempre um algoritmo é executado eficientemente para algum problema particular e termina não retornando a resposta em tempo hábil.

Os problemas tratáveis, também chamados fáceis, podem ser resolvidos por algum algoritmo de tempo polinomial, caso contrário são intratáveis ou difíceis. Os problemas de otimização combinatória são tratados por, principalmente, dois tipos de métodos: os exatos e os aproximativos. Os primeiros realizam uma busca exaustiva no espaço de solução, testando todas as possibilidades até achar a ótima global. Já os aproximativos delimitam o espaço de solução, priorizando o tempo de resolução, sem necessariamente fornecer a solução ótima (CORMEN *et al.*, 2012).

1.4.1 Métodos aproximativos

Os métodos aproximativos possuem a vantagem de apresentar soluções boas para instâncias de grande porte em um tempo hábil, o que nem sempre pode ser alcançado por métodos exatos, porém não garantem que a solução ótima do problema seja encontrada. As heurísticas, as metaheurísticas, as matheurísticas e as hiper-heurísticas fazem parte da classe dos métodos aproximativos.

O termo "heurística" é usado porque não existe nenhuma prova matemática que garanta a qualidade da solução encontrada, ou seja, não possuem um limite matematicamente provado que garanta a qualidade de suas soluções, sendo estas avaliadas empiricamente, de maneira experimental. As "metaheurísticas" são métodos mais sofisticados, que utilizam heurísticas subordinadas a cada etapa, constituindo um modo mais inteligente de exploração do espaço de solução do problema tratado. Muitas delas são bem conhecidas na literatura e serviram de base para outras metaheurísticas mais sofisticadas ou direcionadas para um problema particular. Algumas metaheurísticas bem conhecidas são:

- Algoritmo Genético
- Simulated Annealing
- GRASP
- Busca Tabu

- Colônia de Formigas
- Colônia de Abelhas
- Otimização por Enxame de Partículas

As matheurísticas são métodos que utilizam programação matemática com o auxílio de heurísticas e metaheurísticas, caracterizadas pela exploração do modelo matemático. As matheurísticas de relaxação e fixação trabalham modificando diretamente a estrutura do modelo por meio da relaxação de algumas variáveis, objetivando assim melhorar o tempo de cálculo ou achar bons limites para a solução (LIBERALINO *et al.*, 2011). São elas:

- **Relax-and-fix (RF):** as heurísticas RF se baseiam na resolução de relaxações lineares parciais e sucessivas. A cada iteração, o conjunto de variáveis inteiras (normalmente de decisão) é repartido em vários conjuntos disjuntos. As variáveis do primeiro conjunto continuam sendo tratadas como inteiras, enquanto as outras são relaxadas. O problema modificado é então resolvido. Uma vez a solução é obtida, as variáveis do primeiro conjunto são fixadas aos valores encontrados. Elas são removidas do conjunto de variáveis e o processo continua com as demais, iniciando, assim, uma nova iteração similar à anterior. A operação termina quando todas as variáveis inteiras são fixadas. Esta heurística se baseia no fato de que os problemas modificados são menores e mais fáceis de resolver (FERREIRA; MORABITO; RANGEL, 2010).
- **Cut-and-fix:** procedimento similar ao anterior, diferenciando-se pela fixação de valores para as variáveis binárias de interesse. Esta heurística produz soluções melhores quando uma formulação mais compacta é usada, e quando a solução da programação linear correspondente tem menos restrições fracionadas. A variante *cut-and-fix* é a mesma heurística aplicada a uma formulação que tenha sido compactada com cortes ou uma reformulação estendida (POCHET; WOLSEY, 2006).
- **Dive-and-fix (DF):** segundo Liberalino *et al.* (2011), DF é uma estratégia que realiza uma busca em profundidade na árvore de decisão, utilizando critérios de limitação e fixação nas variáveis 0-1 do problema. Este processo permite simplificar o problema a cada iteração, uma vez que os limites impostos levam a uma solução viável. O método utiliza a relaxação linear a cada execução e, a partir do resultado obtido, aplica critérios para limitar e fixar os valores fracionários das variáveis em inteiros.

Uma hiper-heurística é uma heurística que pode ser utilizada para lidar com qualquer problema de otimização, desde que a ela sejam fornecidos alguns parâmetros,

como estruturas e abstrações, relacionados ao problema considerado. Pode se referir a qualquer processo heurístico que administre outras heurísticas na resolução de problemas de otimização. As hiper-heurísticas têm sido aplicadas a alguns problemas práticos e apresentadas como métodos de grande potencial, no que diz respeito à capacidade de possibilitar o desenvolvimento, em tempo bastante reduzido, de algoritmos capazes de lidar satisfatoriamente, do ponto de vista prático, com problemas de otimização complexos e pouco conhecidos. No entanto, é difícil situar as hiper-heurísticas em algum nível de qualidade e avaliar a robustez dessas abordagens, caso não as apliquemos a problemas para os quais existam diversas instâncias disponíveis publicamente e já experimentadas por algoritmos relevantes. Uma implementação natural de uma hiper-heurística é, por exemplo, um algoritmo que recebe um conjunto de heurísticas e, em cada ponto de decisão, seleciona uma delas para aplicar no passo seguinte. Porém, a principal motivação para o estudo das hiper-heurísticas é a potencial combinação entre facilidade de aplicação e robustez, o que nos leva a enfatizar os métodos cuja inteligência é quase completamente construída sem o emprego de informações sobre um problema específico de otimização (SUCUPIRA, 2007).

Nas próximas seções, serão tratadas a metaheurística GRASP Reativo e a aprendizagem por reforço, técnicas utilizadas no presente estudo para solucionar o problema de reposicionamento de contêineres vazios.

1.4.2 Greedy Randomized Adaptive Search Procedure - GRASP

A metaheurística Procedimento de Busca Adaptativa Aleatória Gulosa (do inglês *Greedy Randomized Adaptive Search Procedure-GRASP*) é basicamente um processo iterativo do tipo multipartida onde cada iteração é composta de duas fases: construção e busca local. A fase de construção consiste de um procedimento para determinar soluções viáveis. As vizinhanças destas soluções são investigadas na fase de busca local até que um ótimo local seja encontrado. Cada iteração GRASP trabalha de forma independente e o resultado final é então a melhor solução dentre todas as encontradas ao final da execução de todas as iterações (FEO; RESENDE, 1995).

A cada execução da etapa de construção, é definida uma lista de candidatos (LC) que representa os elementos que podem ser adicionados à solução gerada de forma incremental, onde em cada passo um novo elemento é inserido na solução parcial. Para seleção do próximo elemento a ser incorporado na solução, todos os candidatos são avaliados segundo uma função gulosa g e seleciona-se um subconjunto formado pelos melhores candidatos a serem inseridos em uma lista denominada lista restrita de candidatos (LRC). O elemento a ser introduzido na solução parcial é selecionado aleatoriamente entre os pertencentes à LC. Estas escolhas aleatórias permitem gerar soluções diferentes a cada execução do algoritmo. Uma vez determinado o elemento, este é inserido na solução parcial, os custos da função de avaliação são recalculados e a

LRC é atualizada. A fase de busca local tenta aprimorar a solução s gerada na fase de construção, efetuando uma varredura na vizinhança de s até que um ótimo local seja encontrado (GONÇALVES, 2005).

Feo e Resende (1995) apresenta a estrutura básica de um procedimento genérico GRASP para um problema de maximização, apresentado no Algoritmo 1.

Algoritmo 1 Procedimento GRASP padrão para um problema de maximização

```

1:  $s^* \leftarrow \emptyset$ 
2: enquanto Criterio de parada não for satisfeito faça
3:    $s \leftarrow \text{ConstroiSolucaoInicial}()$ 
4:    $s' \leftarrow \text{BuscaLocal}(N(s))$ 
5:   se  $c(s') > c(s^*)$  então
6:      $s^* \leftarrow s'$ 
7:   fim se
8: fim enquanto
9: devolve  $s^*$ 

```

De acordo com o Algoritmo 1, a solução global s^* é um conjunto vazio inicialmente e, em seguida, o método inicia suas duas fases. Na fase de construção (linha 3), uma solução viável s é construída para que seja submetida à fase de busca local (linha 4). Esta, por sua vez, realiza uma investigação na vizinhança $N(s)$ com o objetivo de obter a solução de melhor custo, armazenando-a em s' . Se o custo da solução obtida pela busca ($c(s')$) for melhor que o custo da solução global corrente ($c(s^*)$), (linha 5), s^* é atualizada (linha 6). Dessa forma, após a execução de todas as iterações do procedimento, diversas regiões do espaço de busca terão sido pesquisadas e a melhor solução encontrada é então retornada como resultado (linha 9) (RESENDE, 2003 apud OLIVEIRA; SANTOS; MARTINHON, 2008).

Fase de Construção

O GRASP é um método heurístico que combina as boas características dos algoritmos puramente gulosos e dos procedimentos aleatórios na sua fase de construção de soluções factíveis. A fase construtiva é iterativa, gulosa-aleatória e adaptativa. Ela constrói uma solução de forma iterativa adicionando um elemento por vez à solução em construção. Cada elemento é escolhido de forma aleatória, a partir da lista restrita de candidatos. O caráter adaptativo decorre do fato de a escolha do próximo elemento a compor a solução parcial ser influenciada pelas escolhas anteriores. Ou seja, quando o elemento é inserido na solução parcial, ele é removido da lista restrita de candidatos possíveis, sendo esta então atualizada e os custos incrementais recalculados.

Considere um problema de minimização e um conjunto $LC = \{e_1, e_2, \dots, e_n\}$, onde $N \leq |E|$ e E é o conjunto formado por todos os possíveis elementos candidatos a fazerem parte da solução em construção (uma lista de candidatos). Seja $c(e_i)$ uma função gulosa que designa o custo incremental associado à incorporação do elemento $e_i \in LC$ na solução

corrente parcial. Representa-se por c^{min} e c^{max} , o menor e o maior custo incremental, respectivamente.

Relacionado à construção da LRC existe o parâmetro $\alpha \in [0;1]$ que determina o tamanho da LRC. O processo construtivo da LRC utiliza os elementos $e_i \in LC$ com os menores custos incrementais $c(e_i)$, e pode ocorrer de duas formas:

1. Baseada na cardinalidade da LRC: a LRC é constituída de acordo com os p elementos com os melhores custos incrementais, onde p é um parâmetro. Ou seja, selecionam-se os primeiros p elementos a partir de uma lista de candidatos LC ordenada de forma decrescente segundo a função gulosa $c(e)$. O valor de p é calculado por:

$$p = 1 + \alpha(N - 1) \quad (1.5)$$

onde N é o número total de elementos candidatos. Este processo de construção da LRC foi usado no corrente estudo.

2. Baseada na qualidade dos elementos: a lista apresenta tamanho variável, já que depende da quantidade de elementos com bons valores de critério. A LRC é formada por todos os elementos $e_i \in LC$ cujos valores associados da função $c(e)$ respeitam a condição imposta pela equação:

$$LRC = \{e \in LC | c(e) \leq c^{min} + \alpha(c^{max} - c^{min})\} \quad (1.6)$$

O controle entre gula e escolha aleatória é feito pelo parâmetro α , de modo que no caso $\alpha = 0$, o algoritmo é puramente guloso, pois a LRC conterà apenas o elemento c^{min} , enquanto $\alpha = 1$ corresponde ao algoritmo totalmente aleatório, já que a lista conterà todos os candidatos possíveis. Para um problema de maximização, a LRC será constituída por todos os elementos cujo valor da função de avaliação está no intervalo $[c^{max} + (1-\alpha)(c^{min} - c^{max}) ; c^{max}]$, de forma a não ser, totalmente, gulosa nem aleatória.

A presença da aleatoriedade na geração da solução de partida garante que novos mínimos locais possam ser alcançados, e também representa o aspecto probabilístico da metaheurística. Dessa forma, a fase construtiva é responsável pela exploração da busca no espaço de soluções. Sem o elemento aleatório, fase construtiva mais busca local só poderia ser aplicada uma única vez (RESENDE, 2003).

Fase de Busca

A fase construtiva não garante a geração de ótimos locais, sendo, portanto, fortemente recomendada uma etapa de busca local. Esta fase representa a etapa de exploração da metaheurística, pois investiga a vizinhança da solução gerada até encontrar um mínimo local.

Costumeiramente, vizinhanças simples são utilizadas. Dado um conjunto discreto de soluções S , uma vizinhança (proximidade) de uma solução $s \in S$ é um subconjunto ou mapeamento de S , ou seja, $N(s) = \{s_1, s_2, \dots, s_k\}$, $k \leq |S|$, onde qualquer elemento de $N(s)$ é um vizinho de s . Um vizinho é alcançado a partir de uma solução, por meio de um movimento. Uma solução s^+ é dita um ótimo local se não houver qualquer solução vizinha melhor que ela. Um ótimo global (ou solução ótima) s^* é a melhor solução presente no espaço de busca.

Os algoritmos de busca local apresentam três etapas durante a exploração de uma determinada vizinhança no espaço de soluções: a) a partida é uma solução obtida por um método construtivo; b) na etapa de iteração, a solução corrente é substituída sucessivamente pela melhor encontrada por uma busca na sua vizinhança; c) a parada do algoritmo ocorre quando um ótimo local é encontrado.

Em problemas de sequenciamento, tais como o Problema do Caixeiro Viajante e o Problema de Alocação de Tarefas, normalmente são utilizados movimentos baseados em trocas de posição entre elementos da sequência que forma a solução, de modo a obter um novo ordenamento. Movimentos do tipo *2-opt* e *3-opt* são os mais comuns, e significam, respectivamente, que dois e três movimentos são necessários para transformar uma sequência em outra. O método *2-opt* parte de uma solução corrente, testa todas as combinações obtidas com a troca de posição entre dois elementos e escolhe a combinação que retorna o melhor valor da função de custo (JOHNSON; MCGEOCH, 2003).

A eficiência de um procedimento de busca local depende de vários aspectos, tais como: a estrutura da vizinhança, a estratégia de busca na vizinhança, a rápida avaliação da função de custo dos vizinhos, e da qualidade da solução de partida (RESENDE, 2003).

1.4.3 GRASP Reativo

A diferença do GRASP Reativo para o tradicional é somente o parâmetro α (que será utilizado na aprendizagem por reforço), cujo valor é autoajustado ao longo das iterações, levando em consideração a qualidade das soluções obtidas anteriormente. Este método foi proposto por Marcelo Prais e Celso Ribeiro (PRAIS; RIBEIRO, 2000), motivados pelas discussões feitas por Feo e Resende (1995) a respeito da influência do valor do α na qualidade e na diversidade das soluções que foram geradas: ao utilizar um valor fixo para este parâmetro, frequentemente se impede a obtenção de soluções de melhor qualidade que poderiam ser geradas caso outros valores de α fossem utilizados. O estudo mostra, ainda, que a estratégia com variação do parâmetro α frequentemente encontra melhores soluções para o problema proposto.

Segundo Lima Júnior (2009), no GRASP Reativo, o valor do α a ser utilizado na fase de construção da iteração corrente é selecionado de forma aleatória, a partir de um conjunto $\Psi = \{\alpha_1; \alpha_2; \dots; \alpha_m\}$ contendo m valores pré-determinados. A cada $\alpha_k \in \Psi$

existe uma probabilidade p_k associada, que será utilizada como base para escolha do α_k . Inicialmente, todas as probabilidades são iguais, ou seja, $p_i=1/m$; com $i = 1, \dots, m$. As probabilidades $\{p_1; p_2; \dots; p_m\}$ são recalculadas de maneira adaptativa, normalmente a cada y iterações, para favorecer valores de α que tenham produzido boas soluções. Dessa forma, estes valores terão maior probabilidade e, conseqüentemente, terão maiores chances de serem escolhidos nas próximas iterações.

O valor de α influencia na qualidade e diversidade da solução gerada na fase de construção. Valores baixos para α geram soluções gulosas de boa qualidade mas com pouca diversidade. Um valor alto para α , próximo a cardinalidade de LC, leva a uma grande diversidade com soluções de qualidade inferior. Isso também influencia o processo de busca local, pois soluções de qualidade inferior tornam o processo de busca local mais lento.

Para atualizar a distribuição de probabilidades é utilizada a seguinte expressão:

$$p_i = \frac{q_i}{\sum_{j=1}^m q_j}, \quad \forall i, j = 1, 2, \dots, m \quad (1.7)$$

onde

$$q_i = \left(\frac{f(s^*)}{m_i} \right)^\delta, \quad \forall i = 1, 2, \dots, m, \quad \delta \in \mathbb{R} \quad (1.8)$$

Na equação 1.8, considera-se $f(s^*)$ o valor da função objetivo da melhor solução encontrada até o momento e m_i o valor médio de todas as soluções encontradas usando $\alpha = \alpha_i$, com $i = 1, 2, \dots, m$. O parâmetro de amplificação δ tem o objetivo de penalizar as probabilidades cuja média m_i é muito maior que $f(s^*)$, amplificando a diferença entre os valores q_i .

Quanto menor for o valor de m_i , maior será o valor de q_i e, por conseqüência, a probabilidade recalculada p_i será maior. Dessa forma, nos blocos de iterações seguintes, o GRASP Reativo tende a usar os valores de α que encontraram, em média, as melhores soluções.

A estratégia genérica do GRASP Reativo é mostrada no Algoritmo 2. Na linha 9, um $\alpha_i \in \Psi$ é selecionado aleatoriamente, com probabilidade p_i , para ser usado na iteração corrente. As linhas de 15 a 19 correspondem à etapa de atualização das probabilidades p_i e são executadas a cada bloco de y iterações. Na linha 20, o algoritmo passa para a próxima iteração.

Pela observação do Algoritmo 2, nota-se que o GRASP Reativo adiciona apenas uma pequena sobrecarga sobre a versão tradicional do GRASP, já que a complexidade das fases construtiva e de busca local é preservada.

A eficácia do GRASP Reativo reside no fato de que a utilização de diferentes valores de α em iterações distintas permite a construção de LRC diferentes, possibilitando

Algoritmo 2 Pseudocódigo do algoritmo GRASP reativo

```

1: procedimento GRASP REATIVO
2:    $f(s^*) \leftarrow +\infty$ 
3:    $it \leftarrow 1$ 
4:   Definir  $\Psi = \{ \alpha_1, \alpha_2, \dots, \alpha_m \}$ 
5:   para  $i \leftarrow 1$  até  $m$  faça
6:      $p_i \leftarrow 1/m$  ;  $m_i \leftarrow 0$ 
7:   fim para
8:   enquanto não CritérioParada faça
9:     Selecione  $\alpha \leftarrow \alpha_i \in \Psi$  com probabilidade  $p_i$ 
10:     $s_1 \leftarrow FaseConstrutiva(\alpha)$ 
11:     $s_2 \leftarrow BuscaLocal(s_1)$ 
12:    se  $f(s_2) < f(s^*)$  então
13:       $s^* \leftarrow s_2$ 
14:    fim se
15:    se  $it \bmod y = 0$  então
16:      Recalcular  $m_i, \forall i \in \{1, 2, \dots, m\}$ 
17:      Recalcular  $q_i, \forall i \in \{1, 2, \dots, m\}$  pela equação 1.8
18:      Atualizar as probabilidades  $p_i, \forall i \in \{1, 2, \dots, m\}$  pela equação 1.7
19:    fim se
20:     $it \leftarrow it + 1$ 
21:  fim enquanto
22:  retorne  $s^*$ 
23: fim procedimento

```

a geração de soluções diferenciadas que não seriam obtidas por meio do uso de um único valor fixo de α . O GRASP Reativo pode ser visto como uma estratégia que usa uma memória de longo prazo. Resumidamente, o que ele faz é aprender durante o processo que valores de α resultaram em solução com melhores custos e favorecer a escolha de tais valores (ALMEIDA, 2014).

1.4.4 Aprendizagem por Reforço

Aprendizagem por Reforço-AR (*Reinforcement Learning-RL*, em inglês) é um paradigma computacional de aprendizagem em que um agente aprendiz procura maximizar uma medida de desempenho baseada nos reforços que recebe ao interagir com um ambiente desconhecido. Sua utilização é recomendada quando não se dispõe de modelos a priori, ou quando não se consegue obter exemplos apropriados das situações as quais o agente aprendiz irá enfrentar (LIMA JÚNIOR, 2009). Um agente aprendiz é aquele que, a partir da interação com o ambiente que o cerca, aprende ativamente de maneira autônoma, por experimentação direta, uma política ótima de atuação, sem ser ensinado por meio de exemplos fornecidos por um supervisor, de forma que maximize a quantidade de recompensa recebida ao longo de sua execução, independentemente do estado inicial (BIANCHI, 2004).

Outra característica marcante da AR refere-se ao dilema da exploração versus exploração do aprendizado. O agente deve ser capaz de decidir de forma autônoma quando irá obter um novo aprendizado ou utilizar o conhecimento adquirido até o momento. A dificuldade está na tarefa de examinar uma série de possibilidades: favorecer progressivamente aquelas que fornecem melhores resultados ou agir de forma a explorar ações ainda não selecionadas (ALMEIDA, 2014).

A Figura 11 (traduzida a partir da original, obtida em Sutton e Barto (1998)) representa o modo como um agente de AR interage com o ambiente durante processo de aprendizagem.

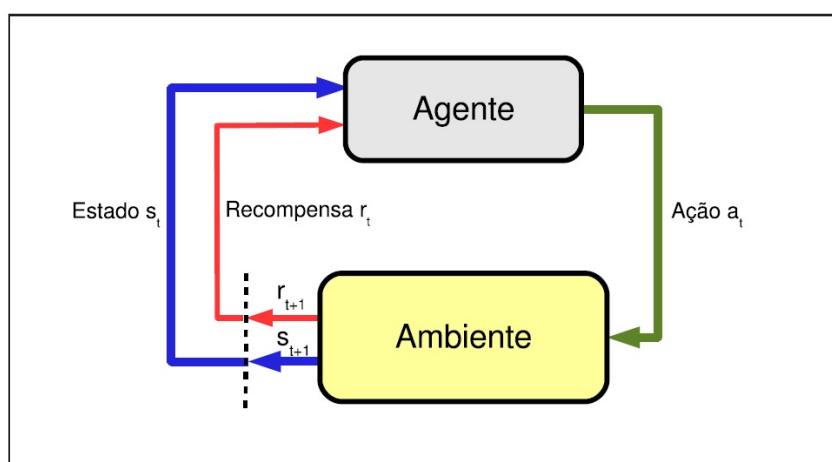


Figura 11 – Processo de interação entre um agente de aprendizagem por reforço e ambiente. Fonte: (SUTTON; BARTO, 1998)

De acordo com a Figura 11, a cada instante de tempo o agente percebe o ambiente e, com base nessa informação, executa uma ação. Esta ação produz dois resultados: o agente recebe uma recompensa imediata e o processo evolui para um novo estado, de acordo com uma determinada distribuição de probabilidades. No passo de tempo seguinte, o ambiente encontra-se alterado e o agente percebe um novo estado com um outro conjunto de ações disponíveis para realizar sua escolha.

CARACTERIZANDO UM PROBLEMA COM APRENDIZADO POR REFORÇO

Sutton e Barto (1998) considera, além do agente e ambiente, quatro elementos básicos do Aprendizado por Reforço abaixo relacionados:

- **Política:** uma política define o comportamento do agente aprendiz, isto é, como ele deve escolher suas ações em um dado instante de tempo. Representa o núcleo de um agente de AR no sentido que ela sozinha é suficiente para determinar o comportamento. Seja S o conjunto de estados do ambiente e $A(s)$ o conjunto

de ações disponíveis para um estado $s \in S$. Uma política π é uma função de mapeamento de estados para ações:

$$a = \pi(s), \forall a \in A, s \in S \quad (1.9)$$

ou seja, dado um estado s , a política determina qual é a ação a que o agente deverá tomar. Geralmente, as políticas podem ser estocásticas.

- **Função de recompensa:** é quem define o objetivo em um problema de AR. Ela associa cada estado (ou par estado-ação) a um sinal numérico (uma recompensa), indicando a atratividade do estado. Assim, o agente tem como objetivo a maximização da quantidade total de reforços recebidos ao longo da execução, denominado de retorno acumulado. Para que o valor do retorno não se torne muito grande um fator de desconto γ ($0 \leq \gamma \leq 1$) é introduzido para reduzir gradativamente os valores futuros.
- **Função valor:** é obtida utilizando os reforços atual e futuro. Enquanto a função de recompensa atribui um sinal imediato às boas escolhas, a função valor indica uma estimativa do ganho total que pode ser acumulado pelo agente durante a aprendizagem, partindo de um estado s , e considerando os estados que o sucedem. Por exemplo, um estado pode sempre produzir baixas recompensas imediatas, mas ainda assim ter um alto valor, pois ele é regularmente frequentado por outros estados que produzem altas recompensas. Quando a função valor considera apenas o estado s , ela é denominada função valor-estado, e denotada por $V(s)$:

$$V(s) = E\{r_{t+1} + r_{t+2} + \dots + r_{t+n} \mid s_t = s\} \quad (1.10)$$

onde E é o valor total esperado dos retornos acumulados pelo agente seguindo uma política π , a partir de um estado $s_t = s$, no instante t . No caso em que a função valor considera as ações possíveis para um dado estado, esta é denominada função valor estado-ação e é denotada por $Q(s,a)$, como segue:

$$Q(s,a) = E\{r_{t+1} + r_{t+2} + \dots + r_{t+n} \mid s_t = s, a_t = a\} \quad (1.11)$$

O que a difere da equação 1.10 é o fato de que o ganho esperado depende da ação escolhida no instante t .

- **Modelo do ambiente:** o modelo representa o comportamento do ambiente, ou seja, dados o estado e a ação, o modelo pode antecipar o próximo estado e a próxima recompensa de acordo com a probabilidade de transição.

O domínio de um problema (ou tarefa) de AR deve ser modelado como um processo de decisão de Markov. A próxima seção disserta sobre o formalismo em torno dos processos de decisão Markovianos.

PROCESSOS DE DECISÃO DE MARKOV

Um processo de decisão de Markov (MDP - Markov Decision Process) é uma forma de modelar processos onde as transições entre estados são probabilísticas, sendo possível observar em que estado o processo está e interferir neste periodicamente (em “épocas de decisão”), executando ações que têm uma recompensa (ou custo). Alternativamente, pode-se definir recompensas por estado apenas, sem que estas dependam da ação executada. São ditos “de Markov” (ou “Markovianos”) porque obedecem a propriedade de Markov: o efeito de uma ação em um estado depende apenas da ação e do estado atual do sistema (e não de como o processo chegou a tal estado); e são chamados “de decisão” porque modelam a possibilidade de um agente (ou “tomador de decisões”) interferir periodicamente no sistema executando ações, diferentemente de Cadeias de Markov, onde não se trata de como interferir no processo (PELLEGRINI; WAINER, 2007).

Segundo Puterman (2005), um processo de decisão sequencial pode ser denotado pela quintupla $M = \{T, S, A, R, P\}$, cujos componentes serão descritos a seguir:

- T : é um conjunto discreto de instantes de tempo em que as decisões são tomadas, denominados instantes de decisão.
- S : é um conjunto de estados do ambiente em que o processo pode estar. Os estados são as percepções do agente sobre o ambiente.
- A : corresponde a um conjunto de ações possíveis de serem executadas em diferentes instantes de decisão. $A(s)$ são as ações disponíveis quando o processo encontra-se no estado s , de forma que $A = \bigcup_{A(s)}, \forall s \in S$ é o resultado da união de todas as ações disponíveis ao agente em qualquer estado.
- $R: S \times A \Rightarrow \mathbb{R}$ representa uma função de retorno que atribui uma recompensa $r_t(s,a)$ como consequência pelo agente ter escolhido a ação $a \in A(s)$, quando no estado $s \in S$, no instante de decisão $t \in T$.
- $P: S \times A \times S \Rightarrow [0,1]$ é uma função de probabilidades de transição, de tal modo que a função $p_t(s'|s,a)$ corresponde à probabilidade de o processo passar para um estado $s' \in S$ no instante seguinte, quando se apresenta no estados no instante t e o agente decidiu executar a ação $a \in A(s)$.

Em um PDM, o conjunto de ações possíveis, os retornos e as probabilidades de transição dependem somente do estado e da ação atual do sistema, e não de como o

processo chegou a tal estado. Assim, a probabilidade de um estado s passar a um estado s' , dado um par estado-ação (s,a) , é dada por:

$$P_{s,s'}^a = Pr\{s_{t+1} = s' \mid s_t = s, a_t = a\} \quad (1.12)$$

onde Pr é o operador de probabilidade e representa a probabilidade do estado s_{t+1} ser s' , sempre que a ação a_t for a e o estado s_t for s no instante de decisão t . A equação 1.12 expressa a propriedade de Markov, a qual define se um processo pode ser considerado markoviano.

Analogamente, o valor do retorno esperado, dados o estado e a ação atuais e o estado seguinte s' , será:

$$R_{s,s'}^a = E\{r_{t+1} \mid s_t = s, a_t = a, s_{t+1} = s'\} \quad (1.13)$$

onde E corresponde ao valor esperado do retorno r_{t+1} , sempre que o estado s_t , no instante t , passa a ser s' no instante $t+1$, escolhendo a ação a_t .

ALGORITMO Q-LEARNING

Como já visto no início da seção 1.4.4, na aprendizagem por reforço, existe a presença de um agente aprendiz que interage com o ambiente, buscando estimar a ação de melhor resultado para cada estado, com o objetivo de estabelecer uma política ótima de atuação. Quando o agente não conhece o modelo de transição de estados do ambiente, torna-se impossível a utilização de programação dinâmica para a determinação desta política. Nem todos os algoritmos de aprendizagem por reforço necessitam de uma modelagem completa do ambiente, ou seja, não necessitam conhecer a matriz de probabilidades de transição e valores esperados do sinal de reforço para todos os possíveis estados e ações do ambiente. Este é o caso das técnicas de aprendizagem por reforço baseadas em diferenças temporais (LIMA JÚNIOR, 2009).

O *Q-learning* é um algoritmo baseado em Diferença Temporal (DT) que dispensa política (*offpolicy method*). Foi proposto por Watkins (1989) em sua tese de doutorado. O desenvolvimento desse algoritmo representa um dos maiores avanços na área de AR.

Um método de DT não exige um modelo exato do ambiente, permite ser incremental e busca estimar valores de utilidade para cada estado do ambiente. Apresenta a vantagem de atualizar as estimativas da função valor a partir de outras estimativas já aprendidas, sem a necessidade de alcançar o estado final de um episódio⁴ antes da atualização (ALMEIDA, 2014).

É capaz de aprender diretamente a partir da experiência, dispensando a necessidade de uma modelagem completa do ambiente para a determinação de uma política ótima de atuação. Assim, sua convergência para valores ótimos de Q (função valor) independe da política utilizada.

⁴ Intervalo de tempo em que o agente percorre uma sequência de estados até atingir um estado final.

A versão mais simples, *One-step Q-learning* (WATKINS; DAYAN, 1992), é definida pela seguinte expressão:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha_q [r_t + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)] \quad (1.14)$$

Seja s_t o estado corrente, a_t a ação realizada no estado s_t , r_t o reforço imediato obtido após executar a ação a_t em s_t , s_{t+1} o estado seguinte, $\max_a Q(s_{t+1}, a)$ a máxima função valor (que encontra e seleciona a ação do estado seguinte que a maximize), e $\gamma \in [0,1]$ e $\alpha_q \in [0,1]$, respectivamente, o fator de desconto e o coeficiente de aprendizagem⁵. A função valor estado-ação $Q(s_t, a_t)$ é atualizada a partir do seu valor atual. O fato de selecionar a ação que maximize a função valor no estado seguinte facilita a busca da função valor-ação ótima (ALMEIDA, 2014).

A política ε -gulosa é definida no algoritmo pela escolha da ação que possui o maior valor esperado, com a probabilidade definida por $(1-\varepsilon)$, e de ação aleatória, com probabilidade ε . Este processo permite que o algoritmo explore o espaço de estados e esta é uma das condições necessárias para garantir que algoritmos que utilizam aprendizado por reforço encontrem a ação ótima (SERRA, 2004).

A convergência do algoritmo *Q-learning* não depende do método de exploração usado. Um agente pode explorar suas ações a qualquer momento, não existem requisitos para a execução de ações estimadas como as melhores. No entanto, para melhorar o desempenho do sistema é necessária, durante o aprendizado, a busca das ações que maximizam o retorno (SERRA, 2004 apud Kaelbling, Littman e Moore, 1996).

O Algoritmo 3 apresenta o pseudocódigo do *Q-learning* de forma simplificada com base em Watkins e Dayan (1992):

Algoritmo 3 Pseudocódigo do algoritmo Q-Learning

- 1: procedimento Q-LEARNING ($\alpha_q, \varepsilon, \gamma, \text{NumEpis}$)
 - 2: Inicializar $R(s,a)$ {Matriz de Recompensas}
 - 3: Inicializar $Q(s,a)$ {Matriz dos Q-Valores inicialmente zerada}
 - 4: **Repita**
 - 5: Inicializar s {Estado inicial}
 - 6: **Repita**
 - 7: Selecionar uma ação a de acordo com a política apropriada
 - 8: Receber a recompensa $r(s,a)$ e observar o próximo estado s'
 - 9: Atualizar $Q(s,a)$ de acordo com a equação 1.14
 - 10: $s \leftarrow s'$
 - 11: **Até encontrar um estado final**
 - 12: **Até NumEpis**
 - 13: retorne $Q(s,a)$
 - 14: fim procedimento
-

⁵ Usa-se o índice q para diferenciar do parâmetro α utilizado no GRASP.

Serra (2004) enumera os mais importantes aspectos do algoritmo *Q-learning*:

- O objetivo do uso do algoritmo *Q-learning* é achar uma regra de controle que maximize cada ciclo de controle;
- O uso do reforço imediato é indicado sempre que possível e necessário, desde que ele contenha informação suficiente que ajude o algoritmo a achar a melhor solução;
- *Q-learning* é adotado quando o número de estados e ações a serem selecionados é finito e pequeno.

Neste trabalho, são usados o algoritmo *Q-learning* e o GRASP Reativo conjuntamente, que será abordado detalhadamente no capítulo 2.

1.5 TRABALHOS RELACIONADOS

O reposicionamento de contêineres vazios tem sido um problema em curso desde o início da containerização, mas tornou-se mais proeminente nas últimas décadas. Nessa subseção serão apresentados os principais trabalhos na área de algoritmos aproximativos que se propuseram a resolver o problema de reposicionamento de contêineres vazios.

Bandeira (2005) aponta que um dos artigos mais citados por vários dos autores pesquisados é o de Crainic, Gendreau e Dejax (1993). Eles trabalharam com o planejamento de parte do transporte terrestre do comércio marítimo internacional de contêineres. Para eles, esta é uma atividade extremamente complexa, especialmente se o modelo busca otimizar simultaneamente os aspectos de custo e serviço das operações de uma companhia, em um ambiente competitivo. Os componentes fundamentais deste problema são relativos ao reposicionamento de contêineres vazios aos clientes, à preparação de movimentações futuras da carga de exportação, e ao reposicionamento de outros contêineres vazios em depósitos ou portos de armazenagem, visando a utilizá-los para atender os próximos pedidos periódicos (esperados e previstos).

O modelo apresentado por Ang (2006) aborda uma solução heurística que trata da maximização da rentabilidade de um serviço de transporte de contêineres em um horizonte de planejamento de múltiplos períodos, sujeito a restrições de volume e peso. Neste caso, o modelo trabalha com a alocação de um espaço fixo nos navios para a movimentação dos contêineres vazios entre portos, independente da carga alocada ao navio.

Um exemplo de programação dinâmica aplicado ao problema é abordado por Lam, Lee e Tang (2007), onde os autores trataram do problema de reposicionamento de contêineres vazios por meio de uma heurística baseada em programação dinâmica aproximada.

O trabalho de [Shintani et al. \(2007\)](#) integra os problemas de roteamento de navios de contêineres e de reposicionamento de contêineres vazios, formulando o problema em dois estágios, tratado por heurística baseada em algoritmos genéticos.

O trabalho de [Chang \(2007\)](#) estuda a substituição entre contêineres vazios de diferentes tipos, a fim de reduzir o custo de vazio intercâmbio recipiente. A metodologia desenvolvida divide o problema em dependente e partes independentes, e aplica-se um procedimento de *branch-and-bound* para a parte dependente.

[Agarwal e Ergun \(2008\)](#) apresentam um modelo integrado, resolvendo o problema de escalonamento de navios e roteamento de cargas simultaneamente. Dada a complexidade da formulação em programação linear inteira mista, os autores empregam métodos heurísticos, geração de colunas e decomposição *Benders* para a solução do problema, apresentando experimentos para 20 portos e 100 navios.

[Cuoco \(2008\)](#) trata do problema integrado de movimentação de contêineres cheios e vazios. O autor implementou o modelo proposto como uma heurística gulosa, em que as cargas são selecionadas conforme uma ordenação de resultados de indicadores de rentabilidade comumente observados entre as companhias de navegação. Por se tratar de uma heurística gulosa, o trabalho de [Cuoco \(2008\)](#) não considera o sistema como um todo e somente um tipo de contêiner é permitido em sua modelagem. Quando a carga demandada possui mais de um tipo de contêiner, tal demanda é convertida em número equivalente de módulos de contêineres (TEU), sendo tratada, assim, por um único tipo de contêiner. Neste caso, não há restrições específicas por tipo de contêiner.

Em um estudo de caso com quatro armadores, [Filho \(2009\)](#) obteve resultados qualitativos que verificam a importância da gestão logística na reformulação de estratégias de geração de receitas e minimização de custos. O autor considerou que receitas não são vinculadas somente à venda e realização de transportes, mas também à produtividade dos contêineres. Os armadores buscam a maior rotatividade de carga por contêineres em cada período. Esse estudo permitiu ressaltar a importância do planejamento de utilização dos contêineres para redução de custos de movimentação e de oportunidade. Os custos de oportunidade, neste contexto, caracterizam-se pela falta de contêineres vazios em determinado porto para certo período do horizonte de planejamento, que pode ocasionar perda de negócios mais rentáveis e, conseqüentemente, oportunidades no transporte de cargas de clientes potencialmente mais rentáveis e a queda do nível de serviço acordado junto aos clientes. O trabalho realizado por [Filho \(2009\)](#) aborda o reposicionamento de contêineres cheios e vazios dinamicamente integrados. Aplicou uma metodologia dividida em três fases: i) formulação do sistema integrado; ii) desenvolvimento de heurística e; iii) comparações entre heurística e a formulação. A heurística proposta é inspirada no problema "Job-Shop", numa situação onde a quantidade de máquinas é igual ao número de tarefas.

[Zambuzi \(2010\)](#) apresenta um modelo de planejamento da distribuição de con-

têineres vazios. Tem por objetivo, principalmente, minimizar custos envolvendo a movimentação de contêineres vazios, em vista do desbalanceamento entre ofertas e demandas. Nesse modelo se consideram restrições de capacidade de transporte dos modais trabalhados. Baseou-se num modelo de fluxo de rede multiproduto ("multicommodity"), distribuída no tempo. Conhecem-se todas as demandas de cargas dos portos, somente considerando decisões de movimentação de contêineres vazios. Condiciona-se ao conhecimento prévio das rotas já programadas de navios. Permite que um navio passe mais de uma vez por um mesmo porto dentro do horizonte de tempo.

Coelho (2015) apresenta uma solução para o problema de reposicionamento de contêineres vazios, trazendo o modelo de fluxo multiproduto binário para caracterizar o problema. O modelo matemático levou em consideração o horizonte de planejamento, porém a distância entre os portos de origem e destino não foi implementada. O algoritmo utilizado neste trabalho foi o GRASP tradicional.

2 PROPOSTA

2.1 METODOLOGIA

A logística é uma área de grande importância no sentido prático, principalmente na otimização de processos e redução de custos. Dentre as atividades logísticas que uma empresa de transporte marítimo de contêineres realiza, este trabalho destaca o problema da logística do contêiner vazio, que está relacionado à necessidade de programar o posicionamento destes contêineres ao longo dos portos e das rotas dos navios. Este posicionamento torna-se necessário quando ocorrem desequilíbrios entre a oferta e demanda de determinados tipos de contêineres ao longo dos portos de uma ou mais rotas de navios que atendem uma dada região.

O modelo trabalhado neste estudo compõe o trabalho de [Coelho \(2015\)](#), cujo objetivo é minimizar o custo total de movimentação de contêineres vazios, composto pelas parcelas de custo de sua movimentação entre terminais, pelo custo de armazenagem e pelo custo de aquisição de contêineres fora do sistema (*leasing*, compra e empréstimo).

Motivados pela importância prática do problema de reposicionamento de contêineres vazios e buscando a resolução de instâncias reais, este trabalho propõe a construção de um algoritmo, utilizando o GRASP Reativo para a sua resolução. Aliado a este método, é utilizada a Aprendizagem por Reforço com o algoritmo *Q-Learning*.

Ao longo desse capítulo, será mostrado o modelo matemático usado para a implementação, além do algoritmo proposto e as técnicas de busca local utilizadas.

2.2 MODELO

Os contêineres vazios podem ser transportados desde qualquer origem inicial até qualquer destino final, dependendo apenas das demandas necessárias nos pontos de destino e do suprimento disponível nos pontos de origem. Antes de chegar ao seu destino final, os contêineres vazios circulam por um ou mais pontos intermediários (depósitos em terra e depósitos em portos), que operam como pontos de transbordo ([BANDEIRA, 2005](#)).

A modelagem matemática para o problema do reposicionamento destes contêineres vazios é apresentado ao longo desta seção e subseções. Este problema visa a minimizar os custos referentes ao processo de equilíbrio entre a quantidade de demanda e as ofertas dos contêineres vazios nos portos considerados, isto é, a necessidade dos portos onde faltam estes contêineres será suprida a partir dos portos onde existem em excesso.

O modelo de reposicionamento de contêineres vazios é representado por uma rede, cujos nós são compostos pelos pontos de origem e de destino, e pelos pontos de

transbordo (depósitos em terra e depósitos em portos). Os arcos da rede representam os movimentos possíveis entre esses pontos. Por definição, não são efetuadas ligações diretas entre clientes de suprimento e clientes de demanda. Também são excluídas as combinações que representam movimentos entre o mesmo depósito em terra ou entre o mesmo depósito em porto. Cada ponto de origem pode transportar contêineres vazios para depósitos em terra ou depósitos em portos que, ao atuar como pontos de transbordo, devem repassar esses contêineres para os pontos de destino, de acordo com sua demanda. O objetivo da distribuição de contêineres vazios é realocá-los de forma a atender a demanda de cada ponto pertencente ao modelo, visando a obter o menor custo possível de transporte global (BANDEIRA, 2005).

O objetivo da formulação matemática deste trabalho é determinar a quantidade de contêineres vazios dentro do sistema, levando em consideração as restrições de capacidade de transporte nos modais marítimo e terrestre.

As próximas subseções apresentam as premissas do modelo (seção 2.2.1) adotadas para tratar o modelo e a formulação matemática proposta para resolver o problema de reposicionamento de contêineres vazios (seção 2.2.2).

2.2.1 Premissas do Modelo

O modelo matemático proposto neste trabalho apresenta algumas regras que são adotadas para sua construção, de acordo com o trabalho de Coelho (2015). São elas:

1. Os portos considerados no sistema de movimentação representam, além do próprio porto em si, o agrupamento de todos os clientes que são atendidos por ele, ou seja, a demanda ou oferta de contêineres vazios associada aos portos representam a somatória das demandas ou ofertas de todos os clientes que aquele porto atende. Além disso, considera-se o porto também como local de armazenamento destes contêineres; não existem outros depósitos externos ou adjacentes ao porto, de modo que os eventuais deslocamentos terrestres entre portos e depósitos externos não serão modelados;
2. Os contêineres disponíveis em um determinado porto incluem aqueles descarregados de um navio, os que estão armazenados no próprio porto, os que devem retornar de clientes, além dos que podem ser alugados. Esses contêineres serão utilizados para satisfazer a demanda ou serão armazenados como estoque.
3. As operações de captação ou distribuição dos contêineres vazios para os clientes não serão modeladas; parte-se do princípio que, a partir do momento em que o contêiner vazio está disponível no porto de demanda, o cliente já foi atendido;

4. Alguns autores, como [Lai, Lam e Chan \(1995\)](#), consideram a possibilidade de substituição de um tipo de contêiner vazio por outro equivalente, como por exemplo: utilizar dois contêineres vazios de 20 pés para atender a uma demanda de um contêiner vazio de 40 pés. Neste trabalho, embora sejam considerados mais de um tipo de contêiner, o intercâmbio entre contêineres de diferentes tipos não será permitido pelo modelo proposto.
5. As restrições de capacidade a serem consideradas levam em conta as limitações nos espaços disponíveis nos modais de transporte. No caso do modal marítimo, o espaço disponível para o transporte de contêineres vazios depende da quantidade de contêineres cheios que serão transportados nesta mesma embarcação, uma vez que estes têm prioridade de transporte em relação aos vazios. No caso do transporte terrestre, esta limitação está relacionada à disponibilidade de veículos (caminhões ou vagões);
6. Por fim, o modelo adota como parcelas que compõem o custo total de movimentação, aquelas correspondentes às operações consideradas: carregamento e descarregamento do navio, transporte e aluguel de contêineres vazios.

2.2.2 Modelo Matemático

Este estudo adota o modelo do problema de fluxo multiproduto, trabalhado por [Coelho \(2015\)](#): inicialmente se exploram a função objetivo e as restrições do problema e, posteriormente, as definições das variáveis de decisão e dos índices e conjuntos do modelo.

O modelo adotado tem as seguintes variáveis, a saber:

- t representa o intervalo de tempo discretizado;
- c_{ij}^k representa o custo unitário de transporte do contêiner do tipo k no arco (i, j) ;
- x_{ijt}^k representa a variável de decisão, que é o fluxo do contêiner do tipo k pelo arco (i, j) no tempo t .

$$x_{ijt}^k = \begin{cases} 1, & \text{se o fluxo do contêiner do tipo } k \text{ faz uso do arco } (i, j), \text{ no tempo } t; \\ 0, & \text{caso contrário.} \end{cases}$$

- $u_{i,j}$ representa a capacidade do arco (i, j) ;
- d_j^k representa a demanda do contêiner do tipo k .
- b_i^k indica se o nó i é oferta ou demanda para o contêiner do tipo k . Considerando que cada contêiner do tipo k está associado a uma origem e destino, então b_i^k pode assumir os seguintes valores:

$$b_i^k = \begin{cases} d^k, & \text{se o nó } i \text{ é oferta para o contêiner do tipo } k; \\ -d^k, & \text{se o nó } i \text{ é demanda para o contêiner do tipo } k; \\ 0, & \text{neste caso, } i \text{ é um nó de transbordo.} \end{cases}$$

Assim, apresenta-se o modelo matemático para o problema do reposicionamento de contêineres vazios, que serviu de base para a implementação deste trabalho:

$$\text{Min}(Z) = \sum_{t \in T} \left[\sum_{k \in K} \sum_{(i,j) \in A} c_{ij}^k d_j^k x_{ij_t}^k \right] \quad (2.1)$$

$$\text{Suj. a} : \sum_{(i,j) \in A} x_{ij_t}^k - \sum_{(j,i) \in A} x_{ji_t}^k = b_i^k, \forall i, j \in N, \forall k \in K, \forall t \in T \quad (2.2)$$

$$\sum_{k \in K} d_j^k x_{ij_t}^k \leq u_{ij}, \forall (i, j) \in A, \forall t \in T, \forall k \in K \quad (2.3)$$

$$x_{ij_t}^k \in \{0, 1\}, \forall k \in K, \forall (i, j) \in A, \forall t \in T \quad (2.4)$$

Sendo:

- A expressão (2.1) representa a função objetivo a ser minimizada, que indica os valores de custo de passagem dos k tipos de contêineres pelos arcos (i, j) ;
- A expressão (2.2) representa as restrições de conservação de fluxo, que são responsáveis pelo gerenciamento do fluxo na rede;
- A expressão (2.3) representa as restrições de capacidade, sendo responsável por limitar a capacidade de cada arco;
- A expressão (2.4) representa a integridade das variáveis envolvidas.

2.2.3 Exemplo Modelo Proposto

A Figura 12 apresenta um exemplo de uma rede multiproduto com o tempo discretizado, onde $N = \{1, 2, 3, 4, 5\}$ representa a lista de portos envolvidos na rede, $A = \{a1, a2, a3, a4, a5, a6, a7, a8\}$, representa as rotas itinerantes entre os portos e $K = \{\text{tanque}, \text{open-top}\}$ representam os tipos de contêineres que serão reposicionados. Para efeito de simplificação o tipo *tanque* será identificado por P1 e o tipo *open-top* por P2. Será adotado o modelo proposto na equação 2.1, onde há inserção de uma variável $t \in T$, que possui a finalidade de tornar o modelo mais próximo da realidade do reposicionamento de contêineres vazios, discretizando o tempo com objetivo de melhor aproveitamento das capacidades dos arcos da rede.

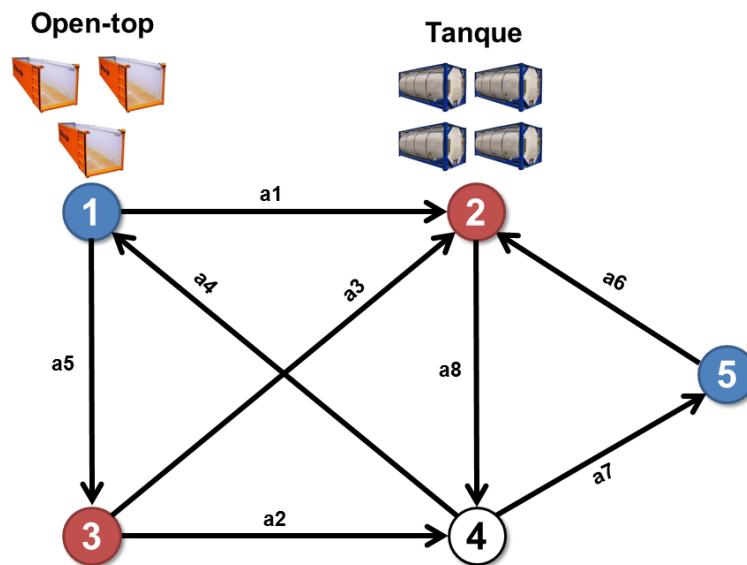


Figura 12 – Exemplo do Problema de Fluxo Multiproduto com tempo discretizado

Para a resolução desse problema, em particular, foi escolhido aleatoriamente o tempo $t=2$, ou seja, serão utilizados, caso necessário, 2 unidades de tempo distintas para o reposicionamento dos contêineres na rede, respeitando as restrições de conservação de fluxo e de capacidades das rotas. A Tabela 2 faz uma melhor apresentação do problema proposto:

Tabela 2 – Dados da instância exemplo

Tempo	Arco	Custo P1	Custo P2	Capacidade	Nós de ligação
1	a1	3	5	5	(1-2)
1	a2	5	4	5	(3-4)
1	a3	2	1	7	(3-2)
1	a4	3	3	6	(4-1)
1	a5	5	9	7	(1-3)
1	a6	9	2	7	(5-2)
1	a7	7	7	5	(4-5)
1	a8	1	7	7	(2-4)
2	a1	3	5	5	(1-2)
2	a2	5	4	5	(3-4)
2	a3	2	1	7	(3-2)
2	a4	3	3	6	(4-1)
2	a5	5	9	7	(1-3)
2	a6	9	2	7	(5-2)
2	a7	7	7	5	(4-5)
2	a8	1	7	7	(2-4)

Cada arco $a \in A$ compõe um par de nós (i,j) e está associado à uma determinada

capacidade, que pode ser representado pela matriz u :

$$u = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 \\ 5 & 5 & 7 & 6 & 7 & 7 & 5 & 6 \end{pmatrix}$$

A matriz c representa os custos para que cada unidade de cada contêiner trafegue pelo referido arco, onde a primeira linha refere-se ao *tanque* e a segunda linha refere-se ao *open-top*.

$$c = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 \\ 3 & 5 & 2 & 3 & 5 & 9 & 7 & 1 \\ 5 & 4 & 1 & 3 & 9 & 2 & 7 & 7 \end{pmatrix} \begin{matrix} P_1 \\ P_2 \end{matrix}$$

Cada contêiner possui quantidade de oferta igual à respectiva quantidade de demanda, isso pode ser observado na matriz b onde, na primeira coluna (*tanque*) tem-se uma oferta no nó 2 de 4 unidades e uma demanda no nó 3 de 4 unidades, da mesma forma na segunda coluna (*open-top*) tem-se uma oferta no nó 1 de 3 unidades e uma demanda no nó 5 de 3 unidades.

$$b = \begin{pmatrix} P_1 & P_2 \\ 0 & 3 \\ 4 & 0 \\ -4 & 0 \\ 0 & 0 \\ 0 & -3 \end{pmatrix} \begin{matrix} n_1 \\ n_2 \\ n_3 \\ n_4 \\ n_5 \end{matrix}$$

Como definido anteriormente, um nó de transbordo significa que servirá apenas de passagem para o contêiner, cujo valor é 0 na matriz b .

A matriz M representa a matriz de incidência nó-arco para o problema apresentado, a qual foi utilizada nas restrições de conservação de fluxo no modelo matemático.

$$M = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 \\ 1 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ -1 & 0 & -1 & 0 & 0 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \end{pmatrix} \begin{matrix} n_1 \\ n_2 \\ n_3 \\ n_4 \\ n_5 \end{matrix}$$

A formulação matemática para a rede binária com tempo discretizado apresentada pela Figura 12 é dada por:

$$\text{Minimizar}(Z) = 12x_{12_1}^1 + 20x_{34_1}^1 + 8x_{32_1}^1 + 12x_{41_1}^1 + 20x_{13_1}^1 + 36x_{52_1}^1 + 28x_{45_1}^1 + 4x_{24_1}^1 + 15x_{12_1}^2 + 12x_{34_1}^2 + 3x_{32_1}^2 + 9x_{41_1}^2 + 27x_{13_1}^2 + 6x_{52_1}^2 + 21x_{45_1}^2 + 21x_{24_1}^2 + 12x_{12_2}^1 + 20x_{34_2}^1 + 8x_{32_2}^1 + 12x_{41_2}^1 + 20x_{13_2}^1 + 36x_{52_2}^1 + 28x_{45_2}^1 + 4x_{24_2}^1 + 15x_{12_2}^2 + 12x_{34_2}^2 + 3x_{32_2}^2 + 9x_{41_2}^2 + 27x_{13_2}^2 + 6x_{52_2}^2 + 21x_{45_2}^2 + 21x_{24_2}^2$$

Sujeito a:

$$\text{r1: } 4x_{12_1}^1 + 3x_{12_1}^2 \leq 5$$

$$\text{r2: } 4x_{34_1}^1 + 3x_{34_1}^2 \leq 5$$

$$\text{r3: } 4x_{32_1}^1 + 3x_{32_1}^2 \leq 7$$

$$\text{r4: } 4x_{41_1}^1 + 3x_{41_1}^2 \leq 6$$

$$\text{r5: } 4x_{13_1}^1 + 3x_{13_1}^2 \leq 7$$

$$\text{r6: } 4x_{52_1}^1 + 3x_{52_1}^2 \leq 7$$

$$\text{r7: } 4x_{45_1}^1 + 3x_{45_1}^2 \leq 5$$

$$\text{r8: } 4x_{24_1}^1 + 3x_{24_1}^2 \leq 6$$

$$\text{r9: } 4x_{12_2}^1 + 3x_{12_2}^2 \leq 5$$

$$\text{r10: } 4x_{34_2}^1 + 3x_{34_2}^2 \leq 5$$

$$\text{r11: } 4x_{32_2}^1 + 3x_{32_2}^2 \leq 7$$

$$\text{r12: } 4x_{41_2}^1 + 3x_{41_2}^2 \leq 6$$

$$\text{r13: } 4x_{13_2}^1 + 3x_{13_2}^2 \leq 7$$

$$\text{r14: } 4x_{52_2}^1 + 3x_{52_2}^2 \leq 7$$

$$\text{r15: } 4x_{45_2}^1 + 3x_{45_2}^2 \leq 5$$

$$\text{r16: } 4x_{24_2}^1 + 3x_{24_2}^2 \leq 6$$

$$\text{r17: } x_{12_1}^1 - x_{41_1}^1 + x_{13_1}^1 + x_{12_2}^1 - x_{41_2}^1 + x_{13_2}^1 = 0$$

$$\text{r18: } x_{12_1}^2 - x_{41_1}^2 + x_{13_1}^2 + x_{12_2}^2 - x_{41_2}^2 + x_{13_2}^2 = 1$$

$$\text{r19: } -x_{12_1}^1 - x_{32_1}^1 - x_{52_1}^1 + x_{24_1}^1 - x_{12_2}^1 - x_{32_2}^1 - x_{52_2}^1 + x_{24_2}^1 = 1$$

$$\text{r20: } -x_{12_1}^2 - x_{32_1}^2 - x_{52_1}^2 + x_{24_1}^2 - x_{12_2}^2 - x_{32_2}^2 - x_{52_2}^2 + x_{24_2}^2 = 0$$

$$\text{r21: } x_{34_1}^1 + x_{32_1}^1 - x_{13_1}^1 + x_{34_2}^1 + x_{32_2}^1 - x_{13_2}^1 = -1$$

$$\begin{aligned}
 r22: & x_{34_1}^2 + x_{32_1}^2 - x_{13_1}^2 + x_{34_2}^2 + x_{32_2}^2 - x_{13_2}^2 = 0 \\
 r23: & -x_{34_1}^1 + x_{41_1}^1 + x_{45_1}^1 - x_{24_1}^1 - x_{34_2}^1 + x_{41_2}^1 + x_{45_2}^1 - x_{24_2}^1 = 0 \\
 r24: & -x_{34_1}^2 + x_{41_1}^2 + x_{45_1}^2 - x_{24_1}^2 - x_{34_2}^2 + x_{41_2}^2 + x_{45_2}^2 - x_{24_2}^2 = 0 \\
 r25: & x_{52_1}^1 - x_{45_1}^1 + x_{52_2}^1 - x_{45_2}^1 = 0 \\
 r26: & x_{52_1}^2 - x_{45_1}^2 + x_{52_2}^2 - x_{45_2}^2 = -1 \\
 r27: & x_{12_1}^1, x_{34_1}^1, x_{32_1}^1, x_{41_1}^1, x_{13_1}^1, x_{52_1}^1, x_{45_1}^1, x_{24_1}^1, x_{12_1}^2, x_{34_1}^2, x_{32_1}^2, x_{41_1}^2, x_{13_1}^2, x_{52_1}^2, x_{45_1}^2, x_{24_1}^2, \\
 & x_{12_2}^1, x_{34_2}^1, x_{32_2}^1, x_{41_2}^1, x_{13_2}^1, x_{52_2}^1, x_{45_2}^1, x_{24_2}^1, x_{12_2}^2, x_{34_2}^2, x_{32_2}^2, x_{41_2}^2, x_{13_2}^2, x_{52_2}^2, x_{45_2}^2, x_{24_2}^2 \in \{0, 1\}
 \end{aligned}$$

As restrições estão identificadas pela presença da letra *r* no início da linha, onde as restrições de capacidade estão numeradas de *r1* até *r16*, e as restrições de conservação de fluxo estão numeradas de *r17* até *r26*. A restrição de integralidade do modelo está identificada por *r27*, onde diz que todas as variáveis envolvidas no modelo são binárias, ou seja, quando o arco for utilizado assume o valor 1, caso contrário o valor assumido será 0.

As rotas possíveis para o transporte dos contêineres da sua origem para o seu destino podem sofrer permutações, aumentando a sua quantidade significativamente. Uma alternativa de solução ótima para o transporte do contêiner tipo *tanque* utilizou o arco a_5 no tempo $t=1$ e os arcos a_4 e a_8 no tempo $t=2$, totalizando um custo ótimo para o primeiro tipo de contêiner com valor de 36, conforme apresentado na Figura 13. Para o contêiner tipo *open-top*, a alternativa ótima encontrada foi a utilização dos arcos a_1, a_7 e a_8 todos no tempo $t=1$, totalizando um custo ótimo de 57, conforme apresentado na Figura 14.

$$\begin{array}{cccccccc}
 & a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 \\
 \left(\begin{array}{cccccccc}
 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1
 \end{array} \right) & \begin{array}{l} t_1 \\ t_2 \end{array}
 \end{array}$$

Figura 13 – Resultado para o contêiner tanque

$$\begin{array}{cccccccc}
 & a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 \\
 \left(\begin{array}{cccccccc}
 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
 \end{array} \right) & \begin{array}{l} t_1 \\ t_2 \end{array}
 \end{array}$$

Figura 14 – Resultado para o contêiner open-top

Fazendo-se a aglutinação dos resultados apresentados nas Figuras 13 e 14, chega-se ao resultado de quais arcos os contêineres utilizaram para se deslocarem da origem

para seus destinos. Somando-se os custos referentes ao transporte dos dois tipos de contêineres, chega-se ao valor de 93, que é o valor ótimo para o problema apresentado.

2.3 POLÍTICAS DE AÇÕES DO ALGORITMO Q-LEARNING

O agente determina a ação entre diversificar ou intensificar em um problema de aprendizagem por reforço, de acordo com a política de seleção de ações. Ao logo das iterações, o algoritmo *Q-learning* deverá convergir para uma política ótima de ações e, para isto, faz-se necessário estabelecer uma política mista que sopesse as duas técnicas.

No algoritmo desenvolvido neste trabalho, em todos os estados, todas as ações estão disponíveis, porém não é permitida a ação de permanecer no estado atual. Uma função $f(x)$ retorna o custo da solução gerada, aplicando-se um valor α associado em um estado corrente s , cuja próxima ação será balizada da seguinte forma:

1. Aleatoriamente seleciona-se uma ação a dentro do conjunto de ações $A(s)$ e, em seguida, o α associado ao estado retornado por essa ação é aplicado à fase de construção do GRASP;
2. Se o custo da solução construída for menor que $f(x)$, o estado retornado pela ação selecionada será o próximo estado para o qual o processo seguirá;
3. Caso contrário, a ação a é retirada da lista de ações disponíveis, posteriormente retornando-se ao passo 1.

2.4 ALGORITMO PROPOSTO

No algoritmo proposto, a técnica reativa do GRASP baseado em probabilidades será substituída pelo o algoritmo *Q-Learning*. Ao longo do tempo, um valor α é selecionado com base nas informações contidas na matriz dos Q-valores retornada pelo *Q-Learning*, permitindo que experiências exitosas sejam repetidas em iterações futuras.

Almeida (2014) define a estratégia composta por quatro etapas utilizadas neste estudo:

1. A matriz dos Q-valores é inicializada com valor zero para todos os itens;
2. Em seguida, um índice da tabela Q é sorteado aleatoriamente para iniciar o processo de atualização desta. Este índice passa a ser o estado s_0 para o *Q-Learning*;
3. A partir do estado s_0 , utilizando a política descrita na Seção 2.3, o próximo estado s_1 é obtido;
4. Em posse dos estados s_0 e s_1 , o processo iterativo que atualiza a matriz Q é iniciado utilizando a equação (1.14). Após um determinado número de episódios, obtém-se a matriz Q da qual será escolhido o α utilizado na fase de construção. A escolha é feita da seguinte forma (LIMA JÚNIOR, 2009):
 - Seleciona-se aleatoriamente uma linha l da matriz Q ;

- A coluna que contiver o maior valor em l será o índice do α a ser utilizado na fase construtiva;

A matriz Q é atualizada pela execução do algoritmo Q -Learning após cada bloco de y iterações. Isto faz com que a qualidade das informações coletadas seja melhorada durante o processo de busca.

O pseudocódigo do método implementado é apresentado no Algoritmo 4, cujos parâmetros de entrada são: o coeficiente de aprendizagem α_q , o fator de desconto γ e o número de episódios $NumEpis$.

Algoritmo 4 Estrutura do algoritmo GRASP Reativo com Aprendizagem por Reforço

```

1: procedimento RL-RGRASP ( $\alpha_q, \gamma, NumEpis$ )
2:    $f(x^*) \leftarrow +\infty$ 
3:    $it \leftarrow 1$ 
4:   Definir  $\Psi = \{ \alpha_1, \alpha_2, \dots, \alpha_m \}$ 
5:   enquanto não CritérioParada faça
6:      $\alpha \leftarrow AprenderAlfa(Q)$ 
7:      $x_1 \leftarrow FaseConstrutiva(\alpha)$ 
8:      $x_2 \leftarrow BuscaLocal(x_1)$ 
9:     se  $f(x_2) < f(x^*)$  então
10:       $x^* \leftarrow x_2$ 
11:     fim se
12:     se  $it \bmod y = 0$  então
13:        $Q \leftarrow Qlearning(\Psi, f(x^*), \alpha_q, \gamma, NumEpis)$ 
14:     fim se
15:   fim do enquanto
16:   retorne  $x^*$ 
17: fim procedimento

```

Observa-se que nas linhas 6 e 13 ocorrem alterações significativas. Na linha 6, a função *AprenderAlfa* recebe a matriz dos Q-valores e seleciona um valor α a partir desta. Na primeira execução esta matriz está zerada, sendo a escolha do valor α totalmente aleatória. Na linha 13, o algoritmo Q -Learning é executado por meio da chamada da função *Qlearning*, que recebe o vetor de valores α pré-determinados Ψ , a melhor solução encontrada até o momento $f(x^*)$, necessária para o cálculo da função de recompensa, e os mesmos parâmetros do algoritmo proposto. O procedimento retorna a matriz dos Q-valores atualizada para ser utilizada no próximo bloco de iterações. Os procedimentos de construção e busca local não são afetados.

2.5 FASE DE BUSCA LOCAL

Neste estudo, a fase de busca local foi implementada em três ações sequenciais, a saber:

Estratégia da busca local 1: dada uma solução com n pedidos, será feita uma verificação se existe uma rota menor entre os portos de forma que não viole a capacidade

do arco utilizado, ou seja, procura-se uma rota que, preferencialmente, ligue os portos diretamente, sem intermediários.

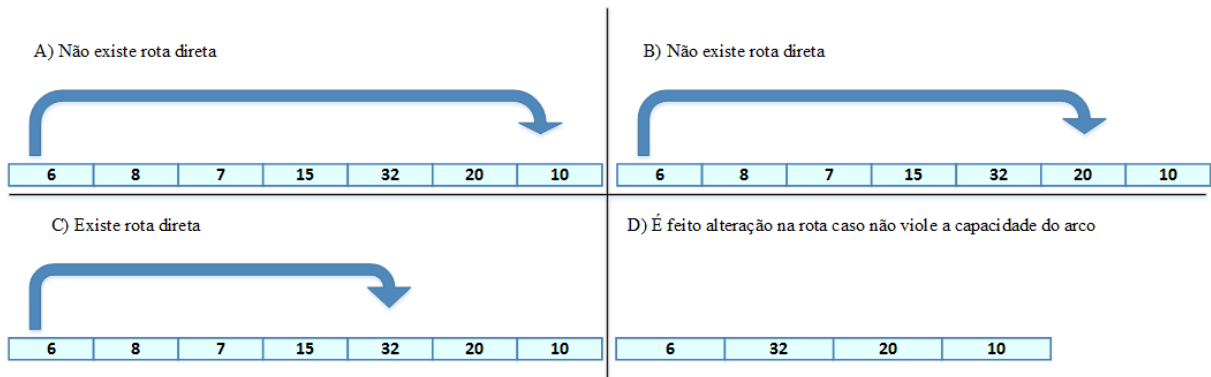


Figura 15 – Busca local 01

Estratégia da busca local 2: é realizada a busca por todos os pedidos e verificam-se quais destes possuem violação e se eles já trafegam pelo caminho mínimo desde o porto de origem até o destino. Caso contrário, a estratégia tenta montar o caminho mínimo, por meio do algoritmo Bellman-Ford, e verifica se este retorna uma rota sem violação.

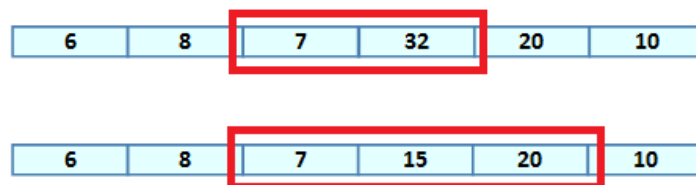


Figura 16 – Busca local 02

Estratégia da Busca local 3: uma busca é realizada por todos os pedidos que possuem valor da função objetivo maior que o valor de caminho mínimo a fim de reconstruir uma nova rota com caminho mínimo. Caso essa rota viole algum arco, tenta-se, a partir da rota de caminho mínimo, recriar uma nova rota, eliminando o arco violado.

3 RESULTADOS

Este capítulo aborda os resultados computacionais obtidos com a execução do algoritmo proposto no item 2.4 para o problema de reposicionamento de contêineres vazios descrito no Capítulo 1, cujo modelo matemático utilizado foi descrito no item 2.2.2. Outrossim, os resultados são comparados aos obtidos no trabalho de [Coelho \(2015\)](#) que abordou o algoritmo GRASP tradicional.

3.1 INSTÂNCIAS UTILIZADAS

O pacote de instâncias *Carbin* foi utilizado nos testes computacionais, criado e disponibilizado por [Alvelos \(2005\)](#), originado a partir de um gerador aleatório baseado na biblioteca *LEDA* (produz instância viável). É composto por 48 instâncias que podem ser classificadas de acordo com cinco características:

- número de produtos (contêineres);
- relação capacidade média dos arcos por demanda média dos produtos;
- tipo de custo (se o custo depende somente do arco ou se depende do arco e do produto que trafega por ele);
- variação dos custos;
- densidade da rede (número de arcos por número de produtos).

O pacote *Carbin* pode ser dividido em duas classes, cada uma delas contendo 24 instâncias: as que contêm um *s* em seu nome indicam que a relação capacidade/demanda é 1,5; as instâncias que apresentam um *l* em seu nome, essa proporção é 10. Com relação à natureza das instâncias, as numeradas com valores ímpares têm custo dependente somente dos arcos e as que contêm valor par em seu nome, o custo depende do arco e do produto que trafega por ele.

Tabela 3 – Característica das instâncias Carbin

Instância	Nós	Arcos	Produtos	Arco/Nós	Produtos/Nós	Penalidade
bx01-bx04	32	96	48	3	1,5	100.000
bx05-bx08	32	320	48	10	1,5	100.000
bx09-bx12	32	96	192	3	6	50.000
bx13-bx16	32	320	192	10	6	50.000
bx17-bx19	32	96	320	3	10	50.000
bx20-bx24	32	320	320	10	10	40.000

Na Tabela 3, o x contido nos nomes das instâncias na primeira coluna deve ser substituído por s ou l e, na segunda coluna, relaciona-se a quantidade de nós na rede multiproducto. Na terceira coluna, apresenta-se a quantidade de arcos e, na quarta coluna, a quantidade de produtos a serem transportados. Na quinta coluna, relacionam-se o número de arcos e o número de nós, na sexta coluna, tem-se a relação entre a quantidade de produtos e nós e, na última coluna, a penalidade aplicada à função objetivo, que será multiplicada pela quantidade de arcos violados da solução, caso haja.

Conforme apresentado no item 1.2.1, cada tipo de contêiner é tratado de forma distinta em relação ao transporte, e, desta forma, assumem custos distintos para o manejo. Assim, os testes foram realizados somente com uma subclasse das instâncias (as numeradas com valores pares), pois os custos dependem do arco e do produto que trafega por ele.

As instâncias utilizadas foram usadas apenas para validar o método implementado, pois, na realidade, a quantidade de portos (nós) no mundo supera as 2.000 unidades, sendo 37 portos públicos no Brasil, segundo a Secretaria de Portos da Presidência da República.

3.2 ANÁLISE DOS RESULTADOS NUMÉRICOS

Os testes foram realizados em um computador com sistema operacional Mac OS X versão 10.9.5, com processador Intel Core I5 1,7 GHz e 4 GB de memória RAM.

As instâncias usadas para a realização dos testes foram BL02, BL04, BL06, BL08, BL10, BL12, BL14, BL16, BL18, BL20, BL22, BL24 e cada uma delas foi testada 30 vezes, devido à natureza aleatória do GRASP Reativo, com valores do alfa entre $\{0; 0,1; 0,2; 0,3; 0,4; 0,5; 0,6; 0,7; 0,8; 0,9; 1,0\}$. Desta forma, cada instância foi testada 330 vezes, com 1000 interações em cada teste.

A Tabela 4 apresenta os resultados obtidos pela função objetivo (F.O.) do algoritmo GRASP com Aprendizagem por Reforço (Q-Learning) e os obtidos por Coelho (2015), bem como o tempo de execução, expresso em segundos, das melhores soluções encontradas. Apresenta ainda uma análise de acordo com a proximidade da solução obtida (x) em relação a solução ótima (x^*), denominada de *gap*, dado por: $gap = 100 \times \frac{(x-x^*)}{x^*}$.

Tabela 4 – Comparativo dos resultados GRASP Reativo com RL x GRASP x Cplex

RGRASP-RL			GRASP		Cplex		GAP
Instância	F.O.	Tempo	F.O.	Tempo	F.O.	Tempo	%
BL02.FMC	1.512.992	1,16	1.506.680	657	1.487.224	0,55	1,73
BL04.FMC	16.633	1,16	17.404	752	15.912	0,45	4,53
BL06.FMC	455.665	1,26	483.498	210	401.021	1,42	13,62
BL08.FMC	5.916	1,27	6.107	1.002	1.336	1,23	342,81
BL10.FMC	5.965.291	3,38	6.026.450	2.602	5.884.016	2,03	1,38
BL12.FMC	62.662	3,26	65.483	1.996	62.012	2,20	1,04
BL14.FMC	2.657.639	4,09	2.919.560	756	2.202.905	9,81	20,64
BL16.FMC	28.770	4,06	36.114	1.597	7.865	7,25	265,79
BL18.FMC	10.169.270	5,64	10.258.300	4.824	10.146.000	2,40	0,22
BL20.FMC	108.032	5,62	109.649	547	107.050	2,54	0,91
BL22.FMC	4.517.140	6,71	5.147.090	1.409	3.750.055	20,90	20,45
BL24.FMC	49.806	6,81	76.238	1.446	43.329	19,30	14,94

De acordo com a Tabela 4, pode-se observar um ganho considerável nos resultados do algoritmo GRASP Reativo com RL em praticamente todas as instâncias (exceto a instância *BL02.fmc*), tanto no valor da função objetivo quanto no tempo de execução para encontrar este valor, em relação aos obtidos por [Coelho \(2015\)](#).

Com relação ao tempo de execução, percebe-se uma linearidade em seus resultados, com aumento gradativo através das iterações. Percebe-se que, a partir da execução da instância *BL22.fmc*, há uma diferença considerável entre os tempos de resposta do GRASP Reativo e o tradicional.

A qualidade do algoritmo GRASP Reativo com Aprendizagem por Reforço pode ser verificada também pela média, pelo desvio-padrão e pela variância obtidos a partir dos valores das soluções encontradas, os quais se apresentam na Tabela 5.

Tabela 5 – Média, desvio-padrão e variância dos valores das soluções encontradas

Instância	F.O	Média	Desvio-Padrão	Variância
BL02.FMC	1.512.992	1.518.513	4376,96	19.157.800,20
BL04.FMC	16.633	16.875,8	107,04	11.458,92
BL06.FMC	455.665	465.462,33	3350,66	11.226.925,60
BL08.FMC	5.916	6.043,93	47,15	2.223,71
BL10.FMC	5.965.291	5.997.987,36	15821,64	250.324.588,79
BL12.FMC	62.662	63.095,63	156,59	24.521,27
BL14.FMC	2.657.639	2.726.967,66	21.061,69	443.594.818,29
BL16.FMC	28.770	29.026,76	141,61	20.056,04
BL18.FMC	10.169.270	10.181.694,06	6.182,46	38.222.850,34
BL20.FMC	108.032	108.389,46	175,58	30.830,74
BL22.FMC	4.517.140	4.598.425,46	29.416,98	865.359.107,84
BL24.FMC	49.806	50.396,13	250,53	62.767,29

De acordo com a Tabela 5, nota-se que a diferença entre a média e o melhor valor

encontrado é pequeno, o que significa que há baixa instabilidade na mudança de valores entre as soluções através das iterações do algoritmo.

Vale salientar que o método foi validado com instância relativamente pequena e que, para instâncias robustas, ele deverá se comportar de forma satisfatória, quando o método exato não poderá dar uma resposta em tempo hábil ou simplesmente não retornar uma solução viável.

3.3 DIFICULDADES ENCONTRADAS

Algumas dificuldades foram encontradas durante o desenvolvimento do estudo, descritos a seguir:

- A implementação do problema abordado requer uma estrutura de dados que minimize o tempo de resposta, visto que o algoritmo utiliza uma quantidade considerável de dados que são manipulados durante a sua execução;
- A verificação das melhores soluções requer mudanças de vários atributos que influenciam no algoritmo, o que eleva o número de iterações do algoritmo para cada instância, ocasionando um menor tempo de resposta;
- Definir uma busca local que melhor se adeque ao problema.
- O problema de reposicionamento de contêineres vazios tem uma estrutura de dados não trivial, difícil de ser adequada outras metaheurísticas, como, por exemplo, o *Particle Swarm Optimization*.

4 CONCLUSÕES

Este trabalho apresentou uma alternativa para o problema de reposicionamento de contêineres vazios, baseado no problema de fluxo multiproduto binário, que consiste em conjugar a metaheurística *Reactive Greedy Reactive Adaptive Search Procedure-RGRASPe* a técnica Aprendizagem por Reforço utilizando o algoritmo *Q-Learning*.

A proposta deste trabalho se mostrou eficiente quando comparado à proposta de Coelho (2015) que trouxe a implementação do algoritmo tradicional GRASP para o problema. Algumas estratégias que trouxeram melhoria na implementação do problema:

- A variação do valor α do algoritmo GRASP Reativo, em contraponto ao valor fixo do α no GRASP tradicional;
- A aprendizagem por reforço do valor do α , a fim de que fosse escolhido o melhor valor para determinada iteração;
- Fase de busca do GRASP não descartar soluções não-factíveis, a fim de tratá-las e torná-las factíveis por meio de outras buscas locais.

Em praticamente todas as instâncias, foram obtidos valores melhores e tempo reduzido de resposta. Outrossim, os resultados obtidos a cada iteração se aproximaram do ótimo, como mostrado no desvio-padrão de cada instância.

O tempo de resposta do algoritmo implementado se apresentou de forma linear com o aumento da robustez da instância, inferindo-se que permanecerá com tal comportamento para outras instâncias.

Para a continuação deste trabalho, sugerem-se como trabalhos futuros:

- Verificar a influência de outras variáveis envolvidas no algoritmo implementado;
- Usar outras estratégias para a fase de busca local;
- Verificar a estratégia de parada do algoritmo de acordo com a estagnação de valores obtidos nas iterações.

REFERÊNCIAS

- AGARWAL, R.; ERGUN, O. Ship scheduling and network design for cargo routing in liner shipping. *Transportation Science*, v. 42, p. 175–196, 2008. Citado na página 41.
- ALBUQUERQUE, M. C. *Programação matemática e heurísticas para o problema de fluxo multiproduto binário*. Dissertação (Mestrado) — Centro Federal de Educação Tecnológica de Minas Gerais, 2011. Citado na página 25.
- ALMEIDA, I. I. de. *Metaheurística híbrida utilizando GRASP reativo e aprendizagem por reforço: uma aplicação na segurança pública*. Dissertação (Mestrado) — Universidade do Estado do Rio Grande do Norte e Universidade Federal Rural do Semi-árido, 2014. Citado 5 vezes nas páginas 34, 35, 38, 39 e 51.
- ALVELOS, F. P. P. da Cunha e. *Branch-and-price and multicommodity flows*. Tese (Doutorado) — Universidade do Minho, 2005. Citado na página 54.
- ANG, J. S. K. Model and algorithms for multi-period sea cargo mix problem. *European Journal of Operational Research*, 2006. Citado na página 40.
- BANDEIRA, D. L. *Alocação e movimentação de contêineres vazios e cheios: um modelo integrado e sua aplicação*. Tese (Doutorado) — Universidade Federal do Rio Grande do Sul, 2005. Citado 5 vezes nas páginas 14, 21, 40, 43 e 44.
- BIANCHI, R. A. da C. *Uso de heurísticas para aceleração do aprendizado por reforço*. Tese (Doutorado) — Escola Politécnica da Universidade de São Paulo, 2004. Citado na página 34.
- CHANG, H. A heuristic solution for the empty container substitution problem. 2007. Citado na página 41.
- COELHO, A. S. *Metaheurística aplicada ao problema de reposicionamento de contêineres vazios no transporte marítimo*. Dissertação (Mestrado) — Universidade do Estado do Rio Grande do Norte e Universidade Federal Rural do Semi-árido, 2015. Citado 9 vezes nas páginas 3, 42, 43, 44, 45, 54, 55, 56 e 58.
- CORMEN, T. H.; LEISERSON, C. E.; RIVEST, R. L.; STEIN, C. *Algoritmo: Teoria e Prática*. 3. ed. [S.l.]: Campus, 2012. Citado na página 27.
- CUOCO, M. *Otimização da seleção e alocação de cargas em navios de contêineres*. Dissertação (Mestrado) — Universidade de São Paulo, 2008. Citado na página 41.
- FEO, T. A.; RESENDE, M. G. C. Greedy randomized adaptive search procedures. *Journal of Global Optimization*, v. 6, p. 109–134, 1995. Citado 3 vezes nas páginas 29, 30 e 32.
- FERREIRA, D.; MORABITO, R.; RANGEL, S. Relax and fix heuristics to solve one-stage one-machine lot-scheduling models for small-scale soft drink plants. *Computers & Operations Research*, p. 76–88, 2010. Citado na página 28.
- FILHO, E. d. A. M. *Alocação e movimentação dinâmica de contêineres: Um modelo integrado de escalonamento*. Dissertação (Mestrado) — Universidade Federal do Rio Grande do Sul, 2009. Citado na página 41.

FORD, L. R.; FULKERSON, D. R. *Flows in network*. [S.l.]: Princetown University Press, 1962. Citado na página 24.

GELAREH, S.; MONEMI, R. N.; MAHEY, P.; MACULAN, N.; PISINGER, D. Single string planning problem arising in liner shipping industries: A heuristic approach. *Computers & Operations Research*, v. 40, n. 10, p. 2357 – 2373, 2013. ISSN 0305-0548. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0305054813001007>>. Citado na página 9.

GONÇALVES, L. B. *Heurísticas GRASP para um problema de roteamento periódico de veículos*. Dissertação (Mestrado) — Universidade Federal Fluminense, 2005. Citado na página 30.

HU, T. *Multicommodity network flows*. [S.l.]: Operations Research, 1963. Citado na página 24.

JOHNSON, D. S.; MCGEOCH, L. A. The traveling salesman problem: a case study in local optimization. 2003. Citado na página 32.

LAI, K. K.; LAM, K.; CHAN, W. K. Shipping container logistics and allocation. *Journal of the Operational Research Society*, v. 46, p. 687–697, 1995. Citado 7 vezes nas páginas 5, 6, 15, 20, 22, 23 e 45.

LAM, S. W.; LEE, L. H.; TANG, L. C. An approximate dynamic programming approach for the empty container allocation problem. *Transportation Research Part C: Emerging Technologies*, v. 15, p. 265–277, 2007. Citado na página 40.

LIBERALINO, C. H. P.; DUHAMEL, C.; QUILLIOT, A.; ALOISE, D. J. Modelo exato e metaheurísticas para o problema de dimensionamento de lotes integrado ao transporte. *SBPO*, v. 43, p. 1434–1445, 2011. Citado na página 28.

LIMA JÚNIOR, F. C. de. *Algoritmo Q-learning como Estratégia de Exploração e/ou Exploração para as Metaheurísticas GRASP e Algoritmo Genético*. Tese (Doutorado) — Universidade Federal do Rio Grande do Norte, 2009. Citado 4 vezes nas páginas 32, 34, 38 e 51.

OLIVEIRA, R. T. de; SANTOS, I. M.; MARTINHON, C. A. Aplicação da metaheurística grasp ao problema do maior conjunto controlado generalizado. *SBPO*, p. 1186–1196, 2008. Citado na página 30.

PELLEGRINI, J.; WAINER, J. Processos de decisão de markov: um tutorial. *RITA*, XIV, n. 2, 2007. Citado na página 37.

POCHET, Y.; WOLSEY, L. A. *Production Planning by Mixed Integer Programming*. [S.l.]: Springer US, 2006. Citado na página 28.

PRAIS, M.; RIBEIRO, C. C. An application to a matrix decomposition problem in tdma traffic assignment. 2000. Citado na página 32.

PRESTON, P.; KOZAN, E. *A tabu search technique applied to scheduling container transfers*. [S.l.: s.n.], 2001. Citado na página 13.

PUTERMAN, M. L. *Markov Decision Processes Discrete Stochastic Dynamic Programming*. New York: [s.n.], 2005. Citado na página 37.

- RESENDE, M. G. C. Handbook of metaheuristics. In: _____. [S.l.]: Kluwer Academic Publishers, 2003. cap. Greedy Randomized Adaptive Search Procedures, p. 219–249. Citado 3 vezes nas páginas 30, 31 e 32.
- SAMPAIO, L. M. D. *Alocação de contêneres vazios para a logística reversa do transporte de carga*. Dissertação (Mestrado) — Universidade Federal do Rio de Janeiro, 2001. Citado 2 vezes nas páginas 9 e 15.
- SERRA, M. R. G. *Aplicações de aprendizagem por reforço em controle de tráfego veicular urbano*. Dissertação (Mestrado) — Universidade Federal de Santa Catarina, 2004. Citado 2 vezes nas páginas 39 e 40.
- SHINTANI, K.; IMAI, A.; NISHIMURA, E.; PAPADIMITRIOU, S. The container shipping network design problem with empty container repositioning. *Transportation Research Part E: Logistics and Transportation Review*, v. 43, p. 39–59, 2007. Citado na página 41.
- SONG, D.-P.; CARTER, J. Empty container repositioning in liner shipping. *Maritime Policy & Management*, v. 36, p. 291–307, 2009. Citado na página 12.
- SUCUPIRA, I. R. *Um estudo empírico de hiper-heurísticas*. Dissertação (Mestrado) — Universidade de São Paulo, 2007. Citado na página 29.
- SUTTON, R. S.; BARTO, A. G. *Reinforcement Learning: an introduction*. [S.l.: s.n.], 1998. Citado 2 vezes nas páginas 5 e 35.
- TEIXEIRA, R. B. *Modelo Integrado para Seleção de Cargas e Reposicionamento de Contêineres Vazios no Transporte Marítimo*. Dissertação (Mestrado) — Escola Politécnica da Universidade de São Paulo, 2011. Citado 4 vezes nas páginas 9, 10, 15 e 23.
- TRAN, N.; HAASIS, H.-D. Literature survey of network optimization in container liner shipping. *Flexible Services and Manufacturing Journal*, Springer US, p. 1–41, 2013. ISSN 1936-6582. Disponível em: <<http://dx.doi.org/10.1007/s10696-013-9179-2>>. Citado na página 13.
- VENDRAMINI, E. *Otimização do problema de carregamento de container usando uma Metaheurística eficiente*. Dissertação (Mestrado) — Universidade Estadual Paulista, 2007. Citado 2 vezes nas páginas 9 e 11.
- VIEIRA, G. B. B. *Transporte Internacional de Cargas*. 2. ed. [S.l.]: Aduaneiras, 2013. Citado 3 vezes nas páginas 10, 13 e 14.
- WATKINS, C.; DAYAN, P. *Q-Learning. Machine Learning*. [S.l.: s.n.], 1992. Citado na página 39.
- WATKINS, C. J. C. H. *Learning from delayed rewards*. Tese (Doutorado) — University of Cambridge, 1989. Citado na página 38.
- ZAMBUZI, N. de castro. *Modelo de decisão para o planejamento da movimentação de contêineres vazios*. Dissertação (Mestrado) — Escola Politécnica da Universidade de São Paulo, 2010. Citado na página 41.