



UFRPE

**UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA APLICADA**

Róger Alves Silva de Araújo

**Algoritmo Inteligente para Otimização do Plano
Diário em um Sistema de Recomendação
Alimentar**

Projeto de Dissertação

Recife, 2021

Róger Alves Silva de Araújo

**Algoritmo Inteligente para Otimização do Plano
Diário em um Sistema de Recomendação
Alimentar**

Orientador: Cicero Garrozi
Coorientador: Rodrigo Soares

Dissertação apresentada ao Curso de
Mestrado em Informática Aplicada da
Universidade Federal Rural de Pernambuco
como requisito parcial para obtenção do grau
de Mestre.

Recife, 2021

Dedico este trabalho aos meus amigos mais próximos, que entenderam a minha ausência nos meses em que minhas folgas eram comemoradas apenas com meus velhos amigos IDE, processador de texto e editor de planilha. Também a minha família, por todo o apoio e incentivo dados.

Agradecimentos

Ao professor Danilo Araújo, por todas as cobranças enquanto eu tentava me familiarizar com a computação evolucionária e análise de qualidade, me deu uma base crucial para o desenvolvimento de todo este trabalho. A Ananília, Paulo, Allan, Barbara e Patrícia, meus gerentes de projeto, que permitiram que eu tivesse tempo para me dedicar e não criaram barreiras nos momentos em que usei a expressão “é para o meu mestrado”. A minha amiga Ana Freitas, que me ajudou com ideias quando tudo que eu tinha era uma capa com título e uma grande interrogação. A Karla Silva, que me ajudou a vencer o monstro da formatação de texto. A Pedro, por compartilhar experiências que me ajudaram com os desafios que eu viria a enfrentar. Ao professor Thiago, responsável por muitos destes desafios, mas que sempre se mostrou solícito e confiante que eu iria conseguir vencê-los. Por fim, agradeço em especial aos meus orientadores Cicero e Rodrigo, por toda paciência (muita) e confiança depositada, mesmo com minhas limitações e falhas na agenda, sem esta insistência em acreditar este trabalho não teria enfim criado forma.

Resumo

Aplicações móveis para consulta e registro de informações nutricionais acerca dos alimentos ingeridos são populares nas lojas de aplicativos dos *smartphones*. O público destas aplicações geralmente está dividido entre dois grupos: os que desejam perder peso e os que necessitam do aumento de massa muscular, embora o acesso simplificado às informações nutricionais dos alimentos ingeridos diariamente possua relevância para todos os públicos. Este tipo de aplicativo possui limitações que podem desencorajar sua utilização por parte dos usuários em médio prazo, como por exemplo, o fluxo de uso, que impõe que cada alimento seja reportado manualmente, delegando muita carga de trabalho ao usuário. Também pode ser considerada uma limitação a ausência de funcionalidades que permitam o uso da ferramenta entre os profissionais de saúde e os pacientes para acompanhamento nutricional.

Algoritmos evolucionários vêm sendo estudados no contexto de otimização da geração de recomendações alimentares. Com esta otimização, o fluxo de uso de aplicativos nutricionais se torna menos dispendioso para os usuários. Estes estudos mostraram que, com o uso desta técnica, é possível montar recomendações com boa variabilidade de opções. Porém, os resultados apresentados se mostraram insatisfatórios em relação ao sabor e coerência da combinação dos alimentos. Ainda não foi apresentada uma modelagem de algoritmo que inclua estes conceitos subjetivos no processo de geração das recomendações.

Esta dissertação apresenta uma proposta de algoritmo evolucionário multiobjetivo que visa obter recomendações alimentares que contenham combinações coerentes dos alimentos, bem como satisfaçam os valores nutricionais adequados. O algoritmo evolucionário proposto foi modelado de forma a permitir que profissionais de saúde controlem variáveis importantes para a geração das recomendações, proporcionando a participação ativa no processo de otimização.

Palavras chaves: Algoritmo, Otimização, Multiobjetivo, Recomendação, Nutrição

Abstract

Mobile apps for querying and logging nutritional information about ingested foods are popular in smartphone app stores. The audience for these applications is often divided between two groups: those who want to lose weight and those who want to gain muscle mass, although simplified access to nutritional information on foods eaten daily has relevance to all audiences. This type of application has limitations that discourage users from using it in the medium term. Among them is the flow of use, which requires each food to be reported manually, delegating a lot of workload to the user. It can also be considered a limitation the absence of functionalities that allow the use as a tool for nutritional monitoring between health professionals and patients.

Evolutionary algorithms have been studied in the context of optimizing the generation of dietary recommendations. With this optimization, the flow of nutritional application usage becomes less costly for users. These studies showed that, using this technique, it is possible to make appropriate nutritional recommendations and with good variability of options. However, the results presented were inadequate to dish taste and logical combination of foods. An algorithm modeling that includes these subjective concepts in the recommendation generation process has not yet been presented.

This dissertation presents a Multiobjective evolutionary algorithm proposal that aims to obtain nutritional recommendations that are nutritionally adequate and assembled with attention to the logical combination of foods. It has also been modeled to allow healthcare professionals to control important variables for the generation of recommendations, enabling their active participation in the optimization process.

Keywords: Algorithm, Recommendation, Nutrition

Lista de Figuras

Figura 1: Fluxo básico do AG	33
Figura 2: Exemplos de funcionamento dos operadores de cruzamento, mutação e seleção	36
Figura 3: Distribuição dos alimentos no grupo de teste considerando as metas nutricionais do usuário com mais restrição (à esquerda) e com menos (à direita).....	52
Figura 4: Funcionamento do operador de cruzamento usado no Algoritmo proposto ..	55
Figura 5: Exemplo de ajuste dinâmico do desvio padrão usado no cálculo da aptidão das soluções no início e no fim da evolução.....	61
Figura 6: Exemplo de soluções com bom (à esquerda) e mau espalhamento (à direita).62	
Figura 7: Exemplo de soluções com bom (à esquerda) e mau espaçamento (à direita) .	63
Figura 8: Classificação das soluções para cálculo da taxa de acerto.....	65
Figura 9: Exemplo de gráficos de pareto montados com os resultados dos algoritmos mono-objetivo, da esquerda para direita: trabalho de Isowaka, único nutriente e combinação de nutrientes	72
Figura 10: Dispersão dos valores de taxa de acerto (eixo x) e repetição de subgrupos (eixo y) em 729 combinações de parametros numericos para os usuários 1, 2 e 3	84
Figura 11: Dispersão da taxa de repetição de subgrupos em relação as probabilidades de cruzamento (em cima) e reparação (em baixo)	86
Figura 12: Distribuição dos parâmetros numéricos selecionados com base nas taxas de repetição de subgrupos e acerto durante a primeira etapa do teste.....	88
Figura 13: Distribuição dos parâmetros numéricos selecionados com base nas taxas de repetição de subgrupos e acerto durante a segunda etapa de teste	89
Figura 14: Comparativo das métricas obtidas por meio das abordagens mono e multiobjetivo.....	92

Lista de Tabelas

Tabela 1: Exemplo de refeição com seus respectivos macronutrientes.....	28
Tabela 2: Fórmulas para cálculo do TMB segundo a OMS/FAO	29
Tabela 3: Faixas de valores de NAF de acordo com o estilo de vida do paciente	30
Tabela 4: Participação dos macronutrientes na recomendação energética de acordo com o Ministério da Saúde	49
Tabela 5: Perfis de usuário utilizados nos testes de performance	49
Tabela 6: Valores nutricionais diários recomendados pelo algoritmo para os usuários de teste.....	50
Tabela 7: Distribuição dos alimentos pertencentes ao grupo de teste nas faixas de proximidade das metas nutricionais para os usuários teste com maior e menor restrição nutricional.....	54
Tabela 8: Exemplo de recomendações geradas para o usuário de teste 1	58
Tabela 9: Conjunto inicial de parâmetros numéricos utilizados nos testes de performance	66
Tabela 10: Processo de análise dos resultados na terceira etapa	69
Tabela 11: Instâncias dos algoritmos mono-objetivo utilizadas para comparação	71
Tabela 12: Amostra dos resultados obtidos na primeira etapa de testes	74
Tabela 13: Configurações selecionadas na segunda etapa de teste	75
Tabela 14: Conjunto de parâmetros numéricos reajustados que foram utilizados no segundo ciclo de testes de performance	76
Tabela 15: Comparativo entre as 3 melhores soluções da primeira etapa de testes durante o primeiro e segundo ciclo.....	77
Tabela 16: Configurações selecionadas na segunda etapa de testes do segundo ciclo...	78
Tabela 17: Configurações vencedoras da etapa final de testes.....	79
Tabela 18: Comparativo de recomendações geradas com probabilidade de reparação diferentes	82
Tabela 19: Taxa de repetição de subgrupos nas soluções vencedoras para cada perfil de usuário	83
Tabela 20: Configurações finais para os usuários 1 e 3 considerando a coerência da refeição	90
Tabela 21: Amostra de recomendações geradas pela configuração vencedora do teste para os usuários 1 e 3	91

Tabela 22: Subgrupos alimentares criados	102
--	-----

Lista de Abreviaturas e Siglas

AG	Algoritmo Genético
ANVISA	Agência Nacional de Vigilância Sanitária
CE	Computação Evolucionária
CFN	Conselho Federal de Nutrição
ED	Evolução Diferencial
FAO	<i>Food and Agriculture Organization</i>
FDA	<i>Food and Drug Administration</i>
IA	Inteligência Artificial
IFC	<i>Intakes for Food Constituent</i>
IFG	<i>Intakes for Food Group</i>
IMC	Índices de Massa Corporal
IMDRF	<i>International Medical Device Regulators Forum</i>
MOEA	<i>Multi-Objective Evolutionary Algorithm</i>
MOGA	<i>Multi-Objective Genetic Algorithm</i>
NAF	Nível de Atividade Física
NE	Necessidade Energética
NEPA	Núcleo de Pesquisas em Alimentação
NHS	<i>National Health Service</i>
NSGA	<i>Non-Dominated Sorting Genetic Algorithm</i>
OMS	Organização Mundial da Saúde
SaMD	<i>Software as Medical Device</i>
SPEA	<i>Strength Pareto Evolutionary Algorithm</i>
TAA	<i>Two Archive Algorithm</i>
TACO	Tabela Brasileira de Composição de Alimentos
TC-SEA	<i>Taxi-Cab Surface Evolutionary Algorithm</i>
TMB	Taxa Metabólica Basal
UNICAMP	Universidade Estadual de Campinas

Sumário

1	Introdução.....	14
1.1	Uso de recomendação alimentar para reduzir a inserção manual de alimentos ...	14
1.2	Uma abordagem computacional para redefinir o fluxo padrão de uso dos aplicativos nutricionais	15
1.3	Objetivos.....	16
1.4	Justificativa.....	17
2	Revisão da Literatura.....	19
2.1	Soluções tecnológicas para apoio ao acompanhamento nutricional.....	19
2.1.1	Funcionalidades básicas presentes em aplicativos nutricionais	20
2.1.2	Inovações e limitações do mercado de aplicativos nutricionais	21
2.1.3	Demandas dos profissionais de saúde e usuários para o mercado de aplicativos nutricionais	23
2.2	Aplicações de saúde e segurança.....	25
2.3	Fundamentos teóricos da prescrição nutricional	27
2.3.1	Obtendo energia através dos alimentos	27
2.3.2	Definindo a recomendação energética para um indivíduo	28
2.3.2.1	Nível de atividade física	29
2.4	Fundamentos teóricos da computação evolucionária	30
2.4.1	Motivação e inspiração	31
2.4.2	Fluxo básico de execução de um Algoritmo Genético	32
2.4.3	Parâmetros numéricos e operadores	34
2.4.4	Objetivos e Algoritmos apropriados.....	37
2.5	Técnicas de otimização para geração de recomendações alimentares	38
2.5.1	Soluções mono-objetivo	38
2.5.2	Soluções multiobjetivo	43
3	Metodologia.....	46
3.1	Escolha da técnica computacional para gerar as recomendações.....	46
3.2	Escolha do algoritmo	47
3.3	Base de dados e aplicação	47
3.4	Conhecimentos do problema aplicados ao algoritmo.....	48
3.4.1	Definição das metas nutricionais.....	49
3.4.2	Representação e escopo da solução	51

3.4.3	Construção da população inicial.....	53
3.4.4	Operador de cruzamento.....	54
3.4.5	Operador de mutação.....	55
3.4.6	Função de aptidão.....	57
3.5	Indicadores de desempenho.....	61
3.5.1	Indicador de espalhamento.....	62
3.5.2	Indicador de espaçamento.....	63
3.5.3	Hipervolume.....	64
3.5.4	Taxa de acerto.....	64
3.6	Coleta de dados.....	65
3.7	Análise de Dados.....	67
3.7.1	Etapa 1: Eliminação das piores configurações.....	67
3.7.2	Etapa 2: Escolha das melhores configurações.....	68
3.7.3	Etapa 3: Escolha da configuração final do algoritmo.....	68
3.8	Comparação com a abordagem mono-objetivo.....	69
4.	Resultados.....	73
4.1	Constatações do primeiro ciclo de testes.....	73
4.1.1	Resultados da primeira etapa de testes.....	73
4.1.2	Resultados da segunda etapa de testes.....	74
4.2	Constatações do segundo ciclo de testes.....	76
4.2.1	Resultados da primeira etapa de testes.....	76
4.2.2	Resultados da segunda etapa de testes.....	78
4.2.3	Resultados da terceira etapa de testes.....	79
4.2.4	Taxa de repetição de subgrupos.....	81
4.3	Relação entre a coerência da refeição e os demais indicadores de qualidade.....	84
4.4	Teste com foco na coerência da refeição.....	87
4.4.1	Constatações da primeira etapa de testes.....	88
4.4.2	Constatações da segunda etapa de testes.....	89
4.4.3	Constatações da terceira etapa de testes.....	90
4.5	Algoritmo proposto x Abordagem mono-objetiva.....	92
5	Conclusões e trabalhos futuros.....	94
5.1	Análise geral dos resultados.....	94
5.2	Oportunidades de melhoria.....	96
5.3	Trabalhos futuros.....	97

5.4 Considerações finais	98
Apêndice A	102
Subgrupos definidos	102
Apêndice B	103
Trabalho Publicado	103

1 Introdução

O mercado de aplicativos móveis voltados para assistência à saúde mostra-se aquecido e apresenta soluções robustas ao longo do tempo. Dentre as diferentes áreas de atuação deste tipo de ferramenta, existem as que focam no monitoramento e/ou acompanhamento do consumo de alimentos no dia a dia dos usuários, com o objetivo de proporcionar uma alimentação mais saudável do ponto de vista nutricional. Franco (2016) realizou um levantamento dos aplicativos de acompanhamento nutricional mais utilizados e constatou que existe um conjunto de funcionalidades muito semelhantes entre eles. Ele destaca pontos principais que são comuns durante o fluxo de uso: em primeiro lugar, a necessidade que o usuário forneça algumas características fenotípicas relevantes para auxiliar o monitoramento, como idade, peso e altura. Além disto, na definição de metas nutricionais, geralmente usa-se como meta o consumo máximo diário de calorias ou macronutrientes (carboidratos, gorduras e proteínas). Franco também destaca a interface de acompanhamento do progresso destas metas que está ligada à interação do usuário através da inserção dos alimentos consumidos, pois é através dela que se verifica quão perto ou longe o usuário está de atingir sua meta.

A academia vem estudando este modelo e identificando suas limitações. Jensen (2015) e Franco (2016) apontam como exemplo de limitação a ausência de um canal de comunicação entre o usuário e os profissionais de saúde. Gabrielli (2017) mostra que o método manual de inserção dos alimentos é desestimulante e cansativo para o usuário. Neste sentido, Hongu (2015) e Franco (2016) apresentam métodos alternativos para registro dos alimentos; os resultados apontados mostram que há como reduzir a carga manual imposta ao usuário, porém não há uma solução de uso comercial, portanto as principais aplicações disponíveis do mercado ainda utilizam o método tradicional de adição manual das informações.

1.1 Uso de recomendação alimentar para reduzir a inserção manual de alimentos

Com o intuito de reduzir o impacto do problema de inserção manual de alimentos na experiência dos usuários dos aplicativos nutricionais, uma alternativa é a inserção

automática de alimentos. Para fomentar esta ideia, algumas questões podem ser levantadas:

- i) Esta automação se daria por meio da predição do que o usuário pretende comer em seus hábitos alimentares?
- ii) Sendo a intenção do aplicativo promover uma alimentação mais saudável, como sugerir uma dieta para um usuário com hábitos não saudáveis?
- iii) Qual é a fonte de informação que a aplicação deveria usar para fazer esta recomendação?

Trazer as recomendações de profissionais de saúde para o fluxo padrão de uso dos aplicativos nutricionais é um dos pontos mais importantes mencionados por Jensen (2015) para a garantia da segurança e confiabilidade dos resultados produzidos. Neste sentido, o algoritmo apresentado neste trabalho foi modelado para que o conjunto de alimentos possíveis em uma recomendação, bem como os limites nutricionais de cada refeição sugerida, sejam definidos por um profissional de saúde para cada caso, isto é, tipo de refeição, paciente ou tratamento.

Partindo do princípio de que um médico ou nutricionista pode fornecer as recomendações alimentares por meio da aplicação, surge o seguinte problema: embora o conjunto de alimentos disponíveis seja abrangente, a escolha otimizada dos que podem ser inseridos na dieta não é trivial. Pode não ser viável determinar manualmente todas as combinações possíveis, a depender das metas e preferências do usuário. Portanto, os nutricionistas trabalham normalmente com recomendações prontas, direcionadas para cada tipo de necessidade nutricional ou patologia (diabetes, hipertensão, dentre outros). Tal fato limita demasiadamente as opções do paciente. Recomendações pré-formatadas possuem o efeito indesejado de não se adaptar às características específicas de um usuário (gosto pessoal, hábitos, impossibilidades regionais e financeiras).

1.2 Uma abordagem computacional para redefinir o fluxo padrão de uso dos aplicativos nutricionais

Em Seljak (2006) pode ser observada uma modelagem realizada para o problema da escolha dos alimentos com o problema da mochila. Neste problema, tem-se uma mochila com uma determinada capacidade máxima e um conjunto de itens, cada um com um peso e um valor associado, e deve-se escolher quais itens levar na mochila de forma a maximizar o valor a ser carregado sem extrapolar sua capacidade [Dantas, 2016]. No

caso de uma recomendação alimentar, a capacidade da mochila pode ser definida, por exemplo, pelas metas nutricionais do paciente (quantidade limite de carboidrato, proteína, lipídio e sódio) e os itens seriam os alimentos, que possuem seus valores de cada um desses macronutrientes. Uma solução ótima seria a mochila ter sua capacidade totalmente atingida para todos os objetivos, sem excessos e nem faltas. É interessante notar que neste exemplo os objetivos podem ser conflitantes, por exemplo, um alimento com baixo teor de sódio pode ser rico em carboidratos, o que acrescenta mais complexidade ao algoritmo de busca por uma solução ótima. Para resolver este tipo de problema, a técnica utilizada no trabalho de Seljak, (2006) foi a otimização multiobjetivo.

Este trabalho, propõe a aplicação de um algoritmo evolucionário multiobjetivo, do inglês, *Multiobjective Evolutionary Algorithm* (MOEA) para a otimização da geração de recomendações alimentares. Com o auxílio deste algoritmo, um nutricionista não precisará montar cardápios manualmente, apenas definir os limites nutricionais que o algoritmo utilizará para produzi-los de forma automática, além de aumentar a variedade de opções aos pacientes e reduzir a necessidade de inserção manual de alimentos. Com o uso de técnicas de IA (Inteligência Artificial) será possível ainda sugerir um cardápio mais adequado ao gosto do usuário e adaptá-lo automaticamente, levando em consideração os alimentos consumidos com mais frequência, e combinando-os de maneira adequada para se adequar aos os limites nutricionais diários. A saída do algoritmo será as recomendações alimentares.

1.3 Objetivos

A presente pesquisa tem como objetivo geral desenvolver um algoritmo inteligente que seja capaz de produzir o plano diário de alimentação, baseado em metas nutricionais e restrições alimentares. Para atingir esse objetivo foram propostos os seguintes objetivos específicos:

- Desenvolver um mecanismo que permita a criação de grupos de alimentos adequados para perfis de usuário específicos;
- Recomendar automaticamente uma lista diária de alimentos a serem consumidos nas pequenas refeições, de acordo com o plano diário recomendado pelo algoritmo inteligente;
- Reduzir a carga relacionada com a inserção manual de alimentos, por meio da construção automática de recomendações alimentares.

1.4 Justificativa

De acordo com Pei (2009), está crescendo o número de pessoas preocupadas com a qualidade nutricional dos alimentos ingeridos bem como a preocupação com a utilização de conhecimento científico na montagem das dietas. Atento a isto, o trabalho de Franco (2016) mostra que o mercado de aplicativos voltados para a saúde tem procurado oferecer opções cada vez mais completas aos usuários de aplicativos nutricionais.

Em contrapartida, Shuren (2018) e Fernández (2015) mostram que existe uma crescente preocupação das organizações de saúde com esta descentralização do conhecimento e seus possíveis riscos. Sem regulamentação oficial, muitos aplicativos fornecem informações sem o embasamento de um profissional da área e, por falta de conhecimento, muitos usuários têm substituído o acompanhamento profissional por este tipo de tecnologia.

Existe também uma linha de pesquisa com computação evolucionária aplicada nesta área. Alguns pesquisadores chegaram em bons resultados, do ponto de vista técnico, com o uso de algoritmos evolucionários tradicionais, tais como algoritmos genéticos (AG) em Chen (2018) e Isowaka (2015), e de evolução diferencial (ED) em Pei (2009). Estes algoritmos conseguem obter resultados que preenchem razoavelmente as metas nutricionais pré-estabelecidas, mas que não conseguem atender à lógica comum de montagem de uma refeição, com uma combinação dos alimentos que faça sentido ao usuário.

O uso de operadores personalizados, com conhecimentos específicos do domínio do problema de recomendação alimentar em um algoritmo evolucionário para muitos objetivos, é uma abordagem capaz de atingir resultados adequados. Ao personalizar certos operadores heurísticamente, reduz-se o espaço de busca do algoritmo para que atue sobre uma região mais promissora para obtenção de melhores resultados, enquanto um algoritmo evolucionário instanciado com operadores padrão exploraria todo espaço de busca, o que na prática significa que o algoritmo dedicará tempo para lidar com soluções sem potencial de gerar bons resultados. De acordo com Wang (2012), uma refeição pode ser modelada como um problema para muitos objetivos. Neste contexto utilizou-se um algoritmo com suporte à representação de uma refeição considerando múltiplos nutrientes, fornecendo como saída um conjunto de opções com qualidade semelhante, ao invés de apenas a melhor solução no caso dos algoritmos para um único objetivo, o que é particularmente interessante no contexto de recomendação alimentar (CHEN, 2018).

Por fim, espera-se que as modificações propostas neste trabalho sejam capazes de contribuir para a melhoria da usabilidade dos aplicativos nutricionais, aumentando a adesão e a manutenção de uso de pessoas que buscam uma vida saudável, mas que frequentemente se desestimulam com o fluxo de uso tradicional dos aplicativos atuais devido à alta carga manual de trabalho, além de facilitar o acesso a um acompanhamento profissional especializado.

Esta dissertação está organizada da seguinte maneira: O Capítulo 2 detalha os principais conceitos que foram utilizados durante o trabalho; no Capítulo 3 é apresentada a metodologia usada no desenvolvimento e teste de desempenho do algoritmo proposto; o Capítulo 4 apresenta os resultados obtidos durante o teste do algoritmo e as análises sobre estes resultados; no Capítulo 5 são apresentadas as conclusões do trabalho e apontadas algumas possíveis direções para trabalhos futuros. O Apêndice A apresenta mais detalhes sobre a divisão dos alimentos em subgrupos alimentares (que são explicados na Seção 3.4.5) e o Apêndice B possui alguns detalhes sobre o artigo produzido sobre este trabalho e apresentado no XIV Congresso Brasileiro de Inteligência Computacional (CBIC 2019).

2 Revisão da Literatura

Neste capítulo, serão tratadas com mais detalhes as grandes áreas que compõem este estudo. A Seção 2.1 trata da relação entre a tecnologia e o acompanhamento nutricional. A Seção 2.2 disserta sobre movimentos regulatórios que devem nortear o desenvolvimento de soluções tecnológicas na área de saúde. As Seções 2.3 e 2.4 abordam conceitos relacionados com nutrição e otimização multiobjetivo que são relevantes para este trabalho. Por fim, a Seção 2.5 faz uma revisão dos trabalhos que relacionam computação evolucionária com a geração otimizada de recomendações alimentares.

2.1 Soluções tecnológicas para apoio ao acompanhamento nutricional

Gabrielli (2017) ressalta que uma das vantagens de um aplicativo de acompanhamento alimentar é a possibilidade de intervenções relacionadas a forma como o usuário se alimenta ao longo do tempo, mesmo após uma determinada dieta ter terminado. Assim, se torna mais rápido e eficiente o acompanhamento prestado pelo nutricionista ao paciente em longo prazo. Outras vantagens, do uso deste tipo de aplicação do ponto de vista do profissional de saúde, estão relacionadas com o aumento da fidelização dos pacientes e maior capacidade de atendimento.

Mesmo com as vantagens e o potencial dos apps nutricionais que foram ressaltados no parágrafo anterior, Jensen (2015) destaca que ainda há poucas opções no mercado de soluções que foram desenvolvidas em uma equipe multidisciplinar que continham ao menos um profissional da área de nutrição. Em seu trabalho foi desenvolvido um sistema móvel com foco em automonitoramento do paciente e utilizando uma equipe multidisciplinar, utilizando o chamado “Processo de Atenção Nutricional”, que define 4 etapas para um suporte nutricional de alta qualidade. Tais etapas foram utilizadas para nortear o desenvolvimento das funcionalidades do sistema e são: Assessoramento nutricional; Diagnóstico; Intervenção; Monitoração e avaliação. A atenção a estas etapas foi elogiada pelos profissionais de saúde que participaram do processo de avaliação do trabalho, bem como o uso das técnicas de mudança de comportamento que contemplavam, entre outras técnicas, a configuração de metas e o acompanhamento do progresso da dieta. O sistema, chamado de *NutriMetas*, não chegou a ser testado com pacientes, mas recebeu avaliações positivas superiores a 60% em todas as suas funcionalidades em um grupo de 185 nutricionistas. Destaca-se, contudo, que o

NutriMetas não possui um sistema de recomendação próprio, cabendo ao nutricionista o papel de definir a dieta para cada paciente e a este reportar seu progresso. Características que limitam o potencial do sistema como candidato a ser a resposta para os problemas já mencionados deste fluxo convencional de uso em aplicações deste tipo.

Apesar dos problemas apontados no aplicativo *NutriMetas*, a atenção à opinião dos profissionais de saúde ao longo do desenvolvimento dos apps nutricionais, bem como a escolha de técnicas apropriadas para definição do conjunto de funcionalidades destes é de grande importância. Nikolaou (2017) disserta sobre a mesma carência de aplicações que foram desenvolvidas com considerações de médicos ou nutricionistas, deixando claro que existe uma relação entre a participação de profissionais de saúde nos aplicativos com a taxa de retenção de usuário. Isso também é citado por Jensen (2015) que relaciona a efetividade da recomendação oferecida pelo aplicativo com a presença destes profissionais no processo. O trabalho de Nikolaou (2017) discutiu exatamente esta questão, catalogar e identificar características nos aplicativos voltados para saúde (não apenas nutrição). A pesquisa foi realizada nas lojas de apps de vários países e pôde quantificar o tamanho da carência. Foi constatado que apenas 0,05% dos 28905 aplicativos considerados relevantes para a pesquisa continham a participação de profissionais da área de saúde em sua equipe.

Franco (2016) realizou uma análise sistemática com as principais aplicações deste tipo disponíveis na Google Play Store e na Apple Store, relacionando as funcionalidades e fazendo uma comparação entre eles. Os resultados deste trabalho servem para identificar quais funções são comuns nos aplicativos populares e quais as diferenças entre eles, com as semelhanças pode ser extraído um conjunto mínimo de funcionalidades que devem estar presentes para posicionar um aplicativo nutricional entre os populares. Com as diferenças, pode ser identificado quais as inovações que os aplicativos em destaque estão trazendo como forma de se diferenciarem entre si e encontrar oportunidades ainda não trabalhadas no mercado.

2.1.1 Funcionalidades básicas presentes em aplicativos nutricionais

O fluxo comum dos aplicativos voltados para nutrição contempla a anamnese do paciente. O conjunto padrão das informações consultadas nesta etapa envolve idade, sexo,

peso e altura. Estas informações servem de entrada para o cálculo do Índice de Massa Corporal (IMC) do paciente, um dado informativo muito comum que é usado para categorização do paciente em níveis que variam entre abaixo do peso normal até o maior nível de obesidade. Além dessas informações, o estudo de Franco (2016) também constatou que os aplicativos populares permitem a inserção do nível de atividade física do paciente. Com esta informação, os aplicativos conseguem inferir a necessidade energética de cada paciente de forma personalizada.

O uso de dicas nutricionais baseadas no uso das aplicações também foi destacado por Franco (2016) que dividiu estas dicas em dois tipos: as que consideram o que o usuário está comendo no momento e sugere alimentos parecidos (advertindo quando representa algum problema no panorama geral da dieta) e dicas gerais relacionadas com hábitos saudáveis, como beber água, reduzir o consumo de gordura e praticar atividades físicas.

Todos os aplicativos populares possuem a funcionalidade de relatório de ingestão diária de alimentos, mecanismo por meio do qual o usuário pode acompanhar o progresso da dieta e atendimento as metas por ele estabelecidas. Esta pode ser considerada a funcionalidade principal dos aplicativos deste tipo, pois serve de entrada para a computação das informações fornecidas pelo usuário. Por fim, foi constatado que um gráfico de evolução do peso dos usuários está presente em todos os aplicativos, o que demonstra um direcionamento comum para o público que busca controle desse índice. Isto pode ser analisado como uma limitação destes aplicativos e será tratado na Seção 2.1.2.

2.1.2 Inovações e limitações do mercado de aplicativos nutricionais

Além das semelhanças citadas na Seção 2.1.1, alguns aplicativos tentam se diferenciar trazendo funcionalidades inovadoras que podem antecipar o caminho que o mercado irá seguir nos próximos anos. Também podem ser identificadas limitações que este grupo de aplicativos ainda enfrenta e encontrar oportunidades de melhoria ainda não exploradas. Uma das funcionalidades destacadas por Franco (2016) como inovadora entre todos as aplicações populares que participaram do seu estudo está a integração do sistema com um profissional da área de saúde. Como citado em Shuren (2018) e Fernández (2015)

esta é uma das deficiências nos aplicativos voltados para área de saúde. Franco (2016) constatou que entre os 13 aplicativos mais usados que participaram do estudo, apenas um deles fornecia esta funcionalidade. Pode-se constatar com isso que ainda existe uma barreira entre a atuação dos profissionais de saúde e estas tecnologias, o que foi destacado em Franco (2016) como uma situação potencialmente perigosa. Nesses aplicativos o peso ideal não é sugerido aos usuários, mas algumas vezes é solicitado como entrada. Uma data para se atingir um peso específico é também definida pelos usuários, entretanto, se isso for usado sem uma recomendação profissional pode levar os usuários a iniciar dietas não saudáveis e ocasionar um distúrbio alimentar (FRANCO, 2016).

Em Hongu (2015) foi desenvolvido o aplicativo *Recaller*, que apresentou uma funcionalidade considerada inovadora no mercado e que procurou amenizar uma das limitações mais citadas pelos usuários destes aplicativos que é a inserção dos alimentos. Nesta aplicação, embora ainda manual, o usuário pode fotografar o que está comendo para posterior consulta e inserção dos dados. Isso é útil pois os usuários nem sempre tem tempo para fazer o registro no momento da refeição e a ferramenta provê um meio para lembrar-se do que foi consumido quando for realizar o registro.

Por fim, avaliando a lista de funcionalidades apresentada em Franco (2016) podem ser destacadas aquelas que estavam presentes em no máximo dois aplicativos entre os estudados: Consideração do porte físico do usuário na usabilidade do aplicativo, utilização do GPS do aparelho para fins de monitoramento de atividades, pedômetro para contagem de passos, consideração de condições de saúde específicas do usuário, tais como diabetes, hipertensão e alergias; diário de anotações particular para o usuário registrar informações relevantes, registro de lista de compras de alimentos e gamificação baseada em hábitos saudáveis. A distribuição destas funcionalidades inovadoras foi uniforme entre os aplicativos participantes do trabalho, não havendo nenhum deles que tenha se destacado no quesito inovação.

Como já mencionado, um ponto em comum entre esses aplicativos nutricionais é o foco em perda de peso. Todos eles fornecem um gráfico de evolução do peso do paciente. Ao mesmo tempo que esta é uma funcionalidade útil, e que de fato pessoas que precisam perder peso fazem parte do público-alvo dos aplicativos nutricionais, é importante ressaltar que alimentação não saudável e falta de atividade física não estão relacionados necessariamente com o peso. Assim, precisa-se que os aplicativos nutricionais consigam clarificar este objetivo para os usuários. Em Hongu (2015) foi proposto um aplicativo para registrar o que os usuários estavam comendo por meio de

fotos tiradas pelo *smartphone* para posterior acompanhamento por um nutricionista. O objetivo era melhorar o atendimento e hábitos alimentares do paciente. Contudo, conforme relatado pelos usuários em Hongu (2015) “eles pensaram que isto só seria útil para pessoas que estão tentando perder peso ou iniciando um regime de exercícios”. Atingir o interesse de pessoas que não se enquadram na categoria atleta ou acima do peso pode ser considerado um desafio que é enfrentado pelos aplicativos nutricionais.

Outra limitação do atual conjunto de aplicativos nutricionais fornecidos está na área de automação, sendo o fluxo de uso ainda muito dependente das entradas dos usuários, assim, o papel destes aplicativos hoje é participar passivamente do processo de acompanhamento nutricional. Ou seja, ferramentas mais simples limitando-se a receber *inputs* (entradas) dos usuários (e raramente dos nutricionistas) e formatar esta informação de forma a oferecer conteúdo útil ao usuário. Como já mencionado anteriormente, essas limitações podem ter relação com o tempo de retenção dos usuários, que vão perdendo o interesse em manter a base de dados dos aplicativos alimentada devido a carga de trabalho manual. Soluções tecnológicas vem sendo testadas para melhorar este ponto, mas ainda existem limitações. Este assunto será tratado com mais detalhes na Seção 2.5. Pode-se ressaltar também que existem demandas recorrentes dos profissionais de saúde e usuários dos apps nutricionais que também podem ser traduzidas como limitações.

2.1.3 Demandas dos profissionais de saúde e usuários para o mercado de aplicativos nutricionais

Reconhecimento de imagem, processamento de linguagem natural e inteligência artificial não foram utilizadas nas aplicações citadas até este ponto. Existe uma oportunidade significativa de melhoria em termos de nutrição personalizada, que poderia incluir *feedback* individual, plano de dieta e educação nutricional (FRANCO, 2016). Essas tecnologias poderiam preencher uma área ainda não explorada pelos aplicativos nutricionais, que possuem muitas funcionalidades, mas pouca automação. O que pode explicar um comportamento observado no trabalho de Gabrielli (2017), que desenvolveu um aplicativo para acompanhamento nutricional contendo técnicas de mudanças de comportamento para aumentar a eficácia deste acompanhamento. Segundo relatado no trabalho, foi observada uma considerável queda no número de refeições inseridas diariamente no aplicativo durante o período de acompanhamento; o que confirma como

o registro de alimento ingerido pode impor um alto nível de esforço nos usuários e representa o principal desafio dos aplicativos.

As opiniões de médicos ou nutricionistas no processo de desenvolvimento dos aplicativos nutricionais já foi destacado como importante anteriormente. Em Gabrielli (2017) um grupo de 5 profissionais (2 nutricionistas e 3 pediatras) avaliou e forneceu sugestões de funcionalidades que deveriam estar presentes no aplicativo nutricional que foi desenvolvido no estudo. Este tipo de opinião é muito importante pois a academia busca formas de colaborar para que a tecnologia consiga melhor atender as necessidades destes profissionais. O objetivo principal do aplicativo desenvolvido no estudo era educação nutricional e foi testado por um grupo de 6 famílias, que usaram o aplicativo por um certo período para reportar como as crianças estavam se alimentando e se exercitando. Dentro deste escopo, as requisições de funcionalidades vindas dos profissionais foram:

- Interação simples, intuitiva e de rápido uso;
- Mecanismo digital para acompanhar o que os pacientes estão ingerindo;
- Possibilidade de verificar o grau de atendimento as diretrizes nutricionais prescritas pelo profissional;
- Mecanismo de acompanhamento automático das atividades físicas;
- Mecanismo de motivação para prática de atividades físicas;
- Possibilidade de gravar uma lista de compras de alimentos;

Da lista acima vale destacar que os profissionais veem como úteis as informações presentes nos diários de ingestão de alimentos e nos gráficos de progresso da dieta, funcionalidades apontadas pelo estudo de Franco (2016) como comuns a todos os aplicativos nutricionais populares. A integração com dispositivos de medição de atividades físicas, tais como pedômetro, GPS e dispositivos vestíveis também são requisitadas e aparecem em poucos aplicativos no mesmo estudo, assim como a lista de compras de alimentos que também estava presente como inovação no estudo de Franco (2016). Ou seja, o mercado de aplicativos de saúde oferece poucas opções que estão alinhadas com as necessidades dos profissionais da área, o que pode explicar o porquê da integração entre eles ainda ser pouco presente no mercado.

No que tange a experiência dos usuários, no aplicativo proposto em Gabrielli (2017) foi relatado na etapa de avaliação do estudo, um interesse por mecanismos de recomendação alimentar para melhor experiência com o aplicativo. Em Gabrielli (2017) “eles propuseram a adição de sugestões de receitas saudáveis para a preparação das

próximas refeições, com o objetivo de consolidar a adoção da dieta no seu dia a dia.”. O interesse dos usuários em sistemas de recomendação alimentar é também relatado no trabalho de Chen (2018), em alguns estudos onde foi questionado se as pessoas teriam interesse em seguir recomendações personalizadas baseadas em análise genética, os resultados foram favoráveis em até 50% dos entrevistados.

2.2 Aplicações de saúde e segurança

O número de aplicações com foco em saúde tem crescido de forma exponencial no mercado e, como já foi mencionado, são raros os casos que uma aplicação desta oferece suporte de um médico ou nutricionista ou é utilizado por eles. A qualidade da informação fornecida por estes apps tem sido discutida por organizações de saúde ao redor do mundo, e algumas sanções estão sendo configuradas. Até o momento, ainda não foi documentado algum dano ao usuário causado por uma aplicação mal construída ou que ofereça dados errados em comparação com os calculados utilizando a formulação acadêmica oficial, mas é habitual que algumas aplicações sejam retiradas das lojas de aplicativos por divergências em seus resultados (FERNÁNDEZ, 2015). Embora não forneçam diagnósticos de doenças e nem prescrevam tratamentos, a forma como os aplicativos nutricionais exibem as informações para os usuários podem levar a atitudes não saudáveis e causar doenças. Segundo a Agência Americana de Administração de Alimentos e Drogas, do inglês *US Food and Drug Administration* (FDA), agência federal do departamento de saúde e serviços humanos dos Estados Unidos, os aplicativos que devem estar sujeitos a regulamentação são os que pretendem tratar, diagnosticar, curar, mitigar ou prevenir doenças ou outras condições médicas (SHUREN, 2018). A FDA define os aplicativos que se enquadram na categoria de “sujeitos a regulamentação de software como dispositivo médico”, do inglês *software as medical device* (SaMD). Em 2016, durante o *21st Century Cures Act Congress* a FDA adicionou novas políticas para a supervisão de dispositivos de saúde que incluíram os SaMD e desde então vem liderando os trabalhos em conjunto com reguladores de outros países para definição de uma nova estrutura regulatória que atenda às necessidades de pacientes, médicos e desenvolvedores de aplicativos. O objetivo é colaborar com os desenvolvedores na promoção da tecnologia médica e promover a qualidade do suporte fornecido por este mercado de aplicações. Este grupo de reguladores, formam o Fórum Internacional de Reguladores de Dispositivos Médicos, do inglês *International Medical Device Regulators*

Forum (IMDRF), no qual o Brasil está representado pela Agência de Vigilância Sanitária (ANVISA). Mais recentemente, em 2017, a FDA lançou o Plano de Ação para Inovação em Saúde Digital, e a criação de um programa piloto de pré-certificação que propunha uma avaliação do design de aplicações médicas, teste e outras capacidades que julgarem necessárias a depender do risco dos produtos oferecidos antes de eles serem disponibilizados no mercado.

Além da FDA, na Europa existem diversos outros movimentos que buscam melhorar a qualidade das aplicações de saúde enquanto ainda não existe uma regulação oficial. Fontes de informação consideradas como referência em nível europeu são os sites do serviço nacional de saúde, do inglês *National Health Service* (NHS), e o diretório europeu de aplicativos de saúde, ambos serviços do Reino Unido. Nestes sites, os usuários têm acesso a uma série de informações de saúde e produtos voltadas para esta área, incluindo uma biblioteca de aplicações móveis. Além do Reino Unido, a agência de segurança sanitária de Andaluzia, na Espanha, criou um guia de recomendações para os desenvolvedores de aplicativos voltados para área de saúde que pretende servir de referência para o desenvolvimento de aplicações confiáveis. As recomendações têm foco em quatro principais aspectos: Design e Pertinência, Qualidade e Segurança da Informação, Prestação de Serviços e Confidencialidade e Privacidade. O desenvolvedor que construir o seu aplicativo seguindo as recomendações para todos estes aspectos pode submetê-lo a avaliação da agência e obter um selo de qualidade chamado *App Saudável*, do galego *App Saudable*. Além de ter seu app incluído no Catálogo de Aplicações de Saúde. O objetivo é incentivar os desenvolvedores a seguir o guia ao mesmo tempo que fornece uma base de dados de aplicações confiáveis aos pacientes.

Todas essas ações promovem o acesso a informações sobre o mercado de saúde e fomentam também o mercado de aplicativos voltados para esta área em direção a altos critérios de qualidade e confiabilidade dos serviços prestados. Em Fernández (2015) é possível ver a seguinte afirmação: “Todavia está para se definir um marco jurídico que define as regras de desenvolvimento destes produtos, a falta de uma normativa específica a nível mundial tem levado tanto os provedores como os desenvolvedores de aplicações a necessidade de validação prévia das apps, como único meio de assegurar e aumentar a segurança do paciente”. É interessante que os desenvolvedores estejam atualizados com as últimas movimentações do mercado e a estas regulamentações para evitar problemas futuros.

2.3 Fundamentos teóricos da prescrição nutricional

Esta seção explica conceitos que são utilizados no algoritmo proposto na definição de metas e objetivos. Em 2.3.1 é detalhado como é o processo para se obter o valor energético de uma refeição. Em 2.3.2, são apresentadas técnicas de definição da Necessidade Energética de um indivíduo.

2.3.1 Obtendo energia através dos alimentos

A alimentação tem dois papéis fundamentais para o organismo humano, prover substâncias essenciais e energia para manter os processos fisiológicos funcionando. O conjunto de atividades celulares que fica responsável pelo processo de obter energia por meio dos alimentos ingeridos é o metabolismo. São utilizadas a caloria (cal) e o joule como unidades de medida para representar a quantidade de energia produzida pelo metabolismo. A caloria é uma unidade de grandeza muito pequena, dessa forma é comum que o termo seja utilizado para representar a quilocaloria (kcal), onde $1 \text{ kcal} = 1000 \text{ cal}$. O metabolismo é um processo biologicamente sofisticado que está sendo realizado continuamente no nosso organismo para obtenção e liberação de energia. De forma resumida a energia é obtida degradando os nutrientes dos alimentos que são ingeridos. Um simples alimento pode ser composto por diversos nutrientes, contudo os que são convertíveis em energia no nosso corpo, os chamados nutrientes energéticos ou macronutrientes, são os carboidratos, as proteínas e os lipídios (também chamados de gorduras). Cada um desses nutrientes é capaz de ser convertido em energia com uma determinada taxa de conversão, de forma arredondada, são aceitos nas pesquisas os valores de 4 kcal/g de carboidrato ou proteína e 9 kcal/g de lipídio. Assim, a constituição dos alimentos é importante para determinar a sua contribuição energética (ou de calorias) em uma refeição; da mesma forma o cálculo da quantidade de calorias que determinada refeição tem é então a soma das calorias de cada alimento pertencente a ela. Como exemplo, considere os seguintes dados na Tabela 1 referentes aos itens de um café da manhã e seus macronutrientes:

Alimento	Porção (Gramas)	Carboidrato (Gramas)	Proteína (Gramas)	Lipídio (Gramas)
Pão francês	30	17,58	2,4	0,93
Manteiga	10	0,01	0,04	8,24

Café	100	1,5	0,7	0,1
------	-----	-----	-----	-----

Tabela 1: Exemplo de refeição com seus respectivos macronutrientes.

Para os dados da Tabela 1 a quantidade de calorias desta refeição seria:

$$Cal_{p\tilde{a}o} = (17,58 \times 4) + (2,4 \times 4) + (0,93 \times 9) = 88,29 \text{ kcal}$$

$$Cal_{manteiga} = (0,01 \times 4) + (0,04 \times 4) + (8,24 \times 9) = 74,36 \text{ kcal}$$

$$Cal_{café} = (1,5 \times 4) + (0,7 \times 4) + (0,1 \times 9) = 9,7 \text{ kcal}$$

Com base nessas informações, este café da manhã fictício possuiria um total de 172,35 kcal.

2.3.2 Definindo a recomendação energética para um indivíduo

Devido a particularidades do organismo de cada indivíduo, não há forma exata de determinar as necessidades energéticas de forma geral. Além disso, precisa-se distinguir os conceitos de Necessidade Energética e Recomendação Energética, onde o primeiro seria a quantidade de calorias que uma pessoa deve ingerir para manter seu corpo em funcionamento e executar as atividades do dia a dia, e o segundo teria o mesmo objetivo, mas considerando uma média da população saudável. Pode-se então dizer que as necessidades energéticas são individuais e as recomendações são mais gerais. Adicionalmente, pode-se estender as recomendações nutricionais acrescentando variáveis subjetivas do paciente, tais como objetivo de ganho ou perda de peso ou massa muscular. Vale ressaltar que um profissional de saúde, médico ou nutricionista, é a pessoa capacitada para fazer as recomendações energéticas, pois possuem o conhecimento necessário para considerar e interpretar as diversas variáveis relacionadas com o metabolismo de cada pessoa. Já que cada organismo possui diferenças na eficiência com que a energia proveniente dos alimentos é convertida e com isso existem pessoas que possuem mais ou menos facilidade de ganho ou perda de peso. Existem, contudo, alguns métodos matemáticos consolidados que visam obter um valor de recomendação nutricional que, embora não sejam precisos, podem fornecer um valor de recomendação energética genérica, considerando apenas algumas das variáveis anteriormente citadas.

O gasto energético de um indivíduo pode ser dividido em três componentes, primeiro a Taxa Metabólica Basal (TMB), que representa o valor de energia que o corpo consome para se manter funcionando, segundo o gasto energético do metabolismo que representa a degradação dos alimentos consumidos e obter energia por meio deles, e por

último o gasto energético com a realização de atividades físicas. Segundo Cuppari (2014) a TMB corresponde a até 70% do gasto total energético de uma pessoa, sendo então o componente mais importante a ser definido. Como já mencionado, a TMB pode ser obtida de várias formas, o método mais preciso é o laboratorial, porém apresenta um alto custo para ser executado, sendo usado apenas em estudos. Na prática, a TMB é calculada por um dos métodos consolidados existentes e aceitos pela academia, neste trabalho serão citados dois destes métodos: as fórmulas de Harris e Benedict e da Organização Mundial da Saúde (OMS) em conjunto com a Organização de Alimento e Agricultura (FAO).

A equação de Harris e Benedict leva em consideração o gênero, peso corporal, altura e a idade do paciente. Foi proposta em 1919 e revisada no trabalho de Roza (1984).

$$TMB_{Mulheres} = 447,593 + (9,247 \times P) + (3,098 \times A) - (4,330 \times I)$$

$$TMB_{Homens} = 88,362 + (13,397 \times P) + (4,799 \times A) - (5,677 \times I)$$

Onde:

- P = Peso em Kg
- A = Altura em cm
- I = Idade em anos

A equação proposta pela OMS desconsidera a altura do paciente, considerando gênero, peso corporal e faixa etária. Foi proposta em 2001 pela OMS e FAO mediante um estudo realizado com um grupo de pessoas saudáveis de diferentes etnias:

Faixa etária	Mulheres	Homens
18 a 30 anos	$14,818 \times P + 486,6$	$15,057 \times P + 692,2$
30 a 60 anos	$8,126 \times P + 845,6$	$11,472 \times P + 873,1$
Maior que 60 anos	$9,082 \times P + 658,5$	$11,711 \times P + 587,7$

Tabela 2: Fórmulas para cálculo do TMB segundo a OMS/FAO

Onde:

- P = Peso em Kg

2.3.2.1 Nível de atividade física

As necessidades energéticas de um organismo são compostas pela TMB, custo energético do metabolismo e custo energético com a realização de atividades físicas. O custo energético do metabolismo possui baixa participação no valor total das necessidades energéticas e, visto que as fórmulas do TMB não são precisas, é desconsiderado no cálculo final da necessidade energética. Assim resta definir o nível de

atividade física para calcularmos a quantidade de energia estimada para um organismo. Como existe mais de uma forma de obter o valor da TMB, segundo a OMS e FAO o Nível de Atividade Física (NAF) deve ser multiplicado pelo TMB para se obter o valor de necessidade energética diária. O NAF é definido mediante análise das atividades físicas e corresponde a razão entre o TMB e o gasto energético em um período de 24 horas podendo ser extraído da Tabela 3.

Estilo de Vida	NAF
Sedentário a atividades leves	1,40 a 1,69
Ativo a moderadamente ativo	1,70 a 1,99
Intenso a intensamente ativo	2,00 a 2,40

Tabela 3: Faixas de valores de NAF de acordo com o estilo de vida do paciente

Pode-se perceber que não há uma relação direta entre o estilo de vida do paciente e o valor do NAF, de fato a categorização do estilo de vida do paciente é algo subjetivo, e assim o valor do NAF deve ser definido segundo análise das atividades que o paciente informar que fazem parte do seu dia a dia. Existe um guia com exemplos de atividades físicas que são enquadrados dos estilos de vida apontados na Tabela 3 em Cuppari (2014).

Uma vez definido o NAF do paciente e com o TMB em mãos, é possível obter a sua necessidade energética. Como exemplo, considera-se uma paciente mulher, de 23 anos, altura de 1,62 metros e pesando 55 quilos e com estilo de vida ativo. Usando as fórmulas de Harris Benedict, da OMS e FAO tem-se como resultado os seguintes valores para a necessidade energética (NE) deste paciente hipotético (para este exemplo será usado NAF de 1,70):

- Usando a equação de Harris Benedict:

$$TMB = 447,593 + (9,247 \times 55) + (3,098 \times 162) - (4,330 \times 23) \cong 1358 \text{ kcal}$$

$$NE = 1358 \times 1,70 \cong 2309 \text{ kcal}$$

- Usando a equação da OMS e FAO:

$$TMB = 14,818 \times 55 + 486,6 \cong 1301 \text{ kcal}$$

$$NE = 1301 \times 1,70 \cong 2212 \text{ kcal}$$

Percebe-se que as equações de Harris Benedict e da OMS e FAO fornecem valores de TMB próximos, o que na prática leva a prescrições de dieta equivalentes.

2.4 Fundamentos teóricos da computação evolucionária

Nesta seção serão apresentados os conceitos relevantes que estão relacionados com a computação evolucionária (CE). Estes conceitos são necessários para entender

como o algoritmo proposto funciona. Em 2.4.1 é detalhado brevemente o que motivou a criação da computação evolucionária e sua inspiração, em 2.4.2 é apresentado o fluxo básico de execução de um Algoritmo Genético, algoritmo que serviu de inspiração para o algoritmo proposto. A Seção 2.4.3 define os conceitos de parâmetros numéricos e operadores e a Seção 2.4.4 explica o que são objetivos e as aplicações de algoritmos para um ou mais objetivos.

2.4.1 Motivação e inspiração

Para resolver um problema, em computação, é necessário modelá-lo em forma de uma função cuja entrada serão as variáveis do problema e a saída a solução pretendida. Esta função será então convertida em um algoritmo responsável por aplicá-la nas variadas combinações de valores para as suas variáveis de entrada. Nesta estratégia, os problemas são mapeados utilizando-se métodos diretos, cuja solução consegue ser obtida pelo computador com complexidade no máximo polinomial, isto é, o tempo para se encontrar uma solução aumenta de acordo com uma função polinomial, o que é caracterizado como um problema tratável do ponto de vista computacional. Porém a maioria dos problemas do mundo real é não-linear (Schwefel, 1994), sendo necessário o uso de métodos probabilísticos ao invés dos diretos para se encontrar uma solução em tempo adequado. Embora uma função direta consiga obter o resultado ótimo para a solução de um problema, na prática não se pode utilizar desta abordagem sempre. Em uma solução não-polinomial, a busca pela resposta ótima pode levar um tempo exponencial para ser encontrada, impossibilitando o uso desta estratégia, além do mais, o próprio processo de modelar o problema em uma equação pode ser difícil em alguns casos. Assim soluções não ótimas, mas que consigam ser obtidas em tempo aceitável passam a ser consideradas. Utilizando-se de heurísticas, os métodos probabilísticos costumam se inspirar em processos que podem ser observados na natureza para otimizar o processo de busca por uma solução viável para problemas do mundo real.

A computação evolucionária é um método probabilístico para busca de soluções em problemas de natureza não-polinomial, que se inspira na Teoria Sintética da Evolução ou Neodarwinismo (RIDLEY, 2004). Segundo esta teoria, os principais fatores relacionados ao processo evolutivo são a mutação, a recombinação genética e a seleção natural (AMABIS, MARTHO, 1985; RIDLEY, 2004). A computação evolucionária utiliza alguns dos conceitos do Neodarwinismo, transformando-os nos seguintes componentes básicos dos algoritmos evolucionários:

- População de indivíduos: Um conjunto de soluções candidatas. Uma solução recebe o nome de cromossomo, e cada uma possui características específicas que irão ser usadas para determinar o seu grau de capacidade para resolver o problema. Cada característica recebe o nome de gene.
- Recombinação e mutação genética: Como se propõem a resolver problemas cujo espaço de busca por uma solução é muito grande, dificilmente uma população inicial irá ter todas as características que atendem bem ao problema, sendo necessário que, assim como ocorre na natureza, características novas sejam obtidas por meio de cruzamento genético entre os indivíduos e/ou mutação genética.
- Seleção: Apenas indivíduos (cromossomos) com as melhores características serão selecionados para gerar a próxima geração (cruzamento) com a combinação entre os genes dos seus pais e/ou novos genes obtidos por mutação. Para a seleção, cada gene é avaliado e um valor que representa o grau de atendimento a solução é calculado. Chama-se este valor, que representa a qualidade do indivíduo, de aptidão. A função de avaliação, que calcula a aptidão, depende do problema a ser resolvido.

2.4.2 Fluxo básico de execução de um Algoritmo Genético

Em seu uso mais geral, os algoritmos genéticos se referem a uma família de modelos computacionais inspirados na evolução. Esses algoritmos codificam uma solução potencial para um problema específico em uma estrutura de dados simples do tipo cromossomo e aplicam operadores de recombinação a essas estruturas, a fim de preservar informações críticas (WHITLEY, 1994). O fluxo básico de execução deste algoritmo é composto dos seguintes passos, que estão demonstrados graficamente na Figura 1:

1. Inicialização da população: Um conjunto de soluções candidatas é criado. O critério para criação desta população pode ser aleatório ou direcionado a solução do problema com base em heurísticas.
2. Avaliação e seleção de pais: As soluções candidatas são avaliadas e então são selecionadas aquelas com melhor aptidão. O critério de seleção depende do operador utilizado, no AG clássico a escolha é feita de forma probabilística e

proporcional à aptidão. Assim, os melhores pais terão maior oportunidade de gerar descendentes e propagarem as suas características para as próximas gerações.

3. Cruzamento e mutação: As soluções selecionadas (os pais) passam pelo operador de cruzamento, que irá utilizar os genes de cada pai na criação de uma nova solução filha. As soluções filhas, por combinarem apenas genes das melhores soluções da atual geração, tendem a ter qualidade igual ou superior aos pais. Adicionalmente, para ampliar a área de busca e evitar que as soluções filhas fiquem limitadas as características da geração atual, os filhos passam por um operador de mutação que irá alterar aleatoriamente algum gene do indivíduo dentro de uma determinada probabilidade.
4. Avaliação e seleção: As novas soluções geradas serão avaliadas e então pais e filhos formam uma nova geração onde os sobreviventes serão escolhidos baseado na sua aptidão. O fluxo reinicia com a população sobrevivente e irá ter quantas iterações forem definidas para aquela execução do algoritmo, sempre evoluindo a qualidade da população ou até que o resultado ótimo seja encontrado.

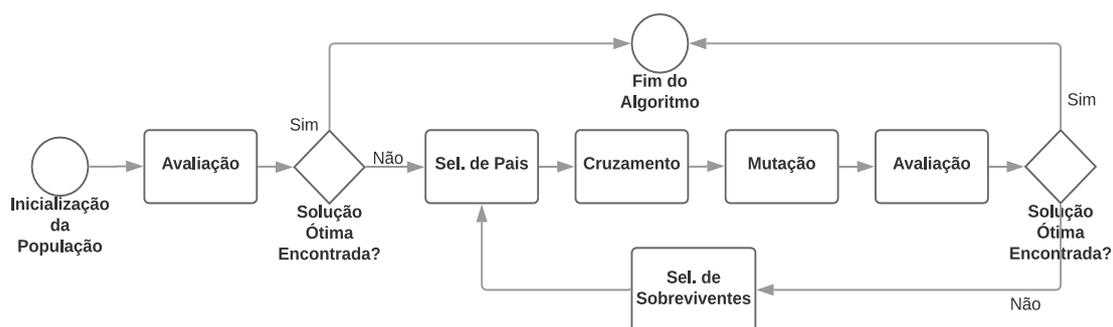


Figura 1: Fluxo básico do AG

Ao utilizar o AG é necessário configurá-lo de acordo com o problema que deseja resolver. A correta configuração do algoritmo é de extrema importância para que resultados satisfatórios sejam obtidos. Como algoritmos evolucionários são estocásticos, eles raramente conseguem obter a solução ótima, uma simples execução do algoritmo também não será suficiente para determinar se a configuração escolhida é a que obtém melhores resultados. A escolha dos operadores e parâmetros corretos para o algoritmo obter os melhores resultados já constitui um problema complexo, dadas as possibilidades de combinações entre eles, sendo assim, um método de avaliação dos resultados do algoritmo precisa ser definido e deve ser testado exhaustivamente com diferentes combinações de operadores e parâmetros de forma a garantir os melhores resultados

possíveis. Essa definição se dará de forma probabilística, utilizando-se de métodos estatísticos apropriados. Mais detalhes sobre este assunto na Seção , quando será explicado como mediu-se a qualidade de nosso algoritmo.

2.4.3 Parâmetros numéricos e operadores

O AG precisa ser configurado de acordo com a necessidade do problema a ser resolvido. Desta forma, o objetivo desta seção é apresentar o conceito de ajuste paramétrico do algoritmo. O ajuste paramétrico pode ser separado por parâmetros em dois grupos: Os **parâmetros numéricos** constituem um conjunto de valores que devem ser definidos e que serão relevantes na execução do algoritmo, os **parâmetros simbólicos** correspondem aos operadores. Na Figura 1 pode ser observado o fluxo básico de execução do AG, os operadores são responsáveis por realizar as etapas que fazem parte das iterações do algoritmo, ou seja, serão os responsáveis pelas variações que o algoritmo fará na população inicial e suas gerações. No AG clássico são utilizados operadores de Cruzamento, Mutação, Seleção de pais e sobreviventes.

Alguns dos parâmetros numéricos comuns são:

- Tamanho da população: O número de indivíduos da população inicial e que deve se manter após a seleção de sobreviventes. Populações muito pequenas podem levar o algoritmo a demorar muito tempo para explorar a área de busca da solução, em contrapartida, populações muito numerosas afetam a performance do algoritmo que precisa iterar sobre mais indivíduos a cada operador. O ideal é que o tamanho da população seja dimensionado de forma que cada geração tenha indivíduos que representem bem as possibilidades de solução, com soluções candidatas bem distribuídos ao longo do espaço de busca.
- Probabilidade de *crossover* (cruzamento): A chance que ocorra cruzamento entre dois pais. Normalmente é desejável valores altos neste parâmetro (podendo ser até de 100%). Valores muito pequenos podem afetar a velocidade de progressão da qualidade da população ao longo das gerações. Dependendo do problema, valores muito altos podem causar a convergência prematura do algoritmo, termo usado quando a qualidade das gerações para de crescer de maneira significativa em um espaço afastado das melhores soluções, ou ainda causar o efeito contrário, fazer o algoritmo nunca convergir.

- Probabilidade de mutação: A chance que ocorra mutação em um novo indivíduo gerado. A mutação é importante no fluxo do algoritmo evolucionário pois regula o tempo para convergência do algoritmo ao mesmo tempo que pode levar as soluções para regiões do espaço de busca ainda não exploradas pelas soluções da geração atual. Contudo ao mesmo tempo que pode haver uma grande melhora na qualidade do resultado, o oposto também pode ocorrer. O ajuste deste parâmetro deve ser feito com cuidado. Normalmente, valores pequenos são usados neste parâmetro.
- Número de iterações: Quantos ciclos, no máximo, o algoritmo deve executar caso a solução ótima não seja atingida. Assim como o tamanho da população, este parâmetro tem impacto direto na performance do algoritmo. Seu valor deve ser ajustado de maneira proporcional com a facilidade para o algoritmo convergir para bons resultados. Valores muito pequenos podem parar a execução do algoritmo antes que ele chegue nos melhores resultados. Valores muito altos podem consumir recursos de máquina desnecessariamente, fazendo cada iteração demorar além do necessário e sem acrescentar na qualidade final do resultado.

Como pode ser visto, os parâmetros numéricos participam de forma ativa na qualidade dos resultados produzidos pelo AG. Os parâmetros simbólicos (ou operadores) também influenciam bastante nestes resultados, porém as possibilidades para eles são menores. Mesmo assim, a depender do problema e do algoritmo evolucionário escolhido, podem existir diversas opções de operadores disponíveis, cada um com um comportamento distinto e que levarão a resultados diferentes. Alguns exemplos de como os operadores funcionam:

- Operadores de cruzamento: São os responsáveis por combinar os genes dos pais e criar um cromossomo a partir deles. Uma das formas de fazer isto é escolher um ponto a partir do qual os genes anteriores virão de um pai, e os posteriores do outro pai. A escolha do ponto é aleatória, assim cada pai contribuirá no máximo com o total dos seus genes menos um para a criação do filho e no mínimo com apenas um gene. Este operador é conhecido como *Crossover* de 1 Ponto. Existem outros como o de “N pontos” e o aritmético que combina o valor dos genes de cada pai na criação do filho usando uma fórmula aritmética.
- Operadores de mutação: Responsáveis por causar as chamadas perturbações genéticas nos cromossomos, ou seja, dado um cromossomo filho já criado, este

operador altera aleatoriamente um ou mais genes sem usar nenhuma informação dos pais. O resultado é um gene de valor totalmente novo e que pode pertencer a um espaço de busca ainda não alcançado pelas gerações anteriores. Vale ressaltar que esta operação pode de forma repentina melhorar ou piorar em muito a qualidade do indivíduo. Um exemplo de operação de mutação é escolher aleatoriamente um ou mais genes e trocar seu valor por outro dentro de um conjunto de valores possíveis para aquele gene.

- Operadores de seleção: Fazem a escolha de pais que serão cruzados e/ou a seleção das soluções que seguirão para a próxima geração (seleção de sobreviventes). Um algoritmo pode usar o mesmo operador para as duas etapas de seleção ou operadores diferentes. A aptidão de cada solução é a principal informação usada por estes operadores. O operador de seleção usado no AG clássico usa uma estratégia chamada de Roleta, na qual as soluções são dispostas em uma roleta com *slots* de tamanho fixo de forma que cada uma ocupe um número proporcional à aptidão. De forma aleatória uma das soluções é escolhida pelo número do *slot*. Note que pela forma como as soluções são distribuídas na roleta, as que possuem os melhores valores de aptidão terão mais chances de serem escolhidas.

Algumas ilustrações destes comportamentos podem ser vistas na Figura 2.

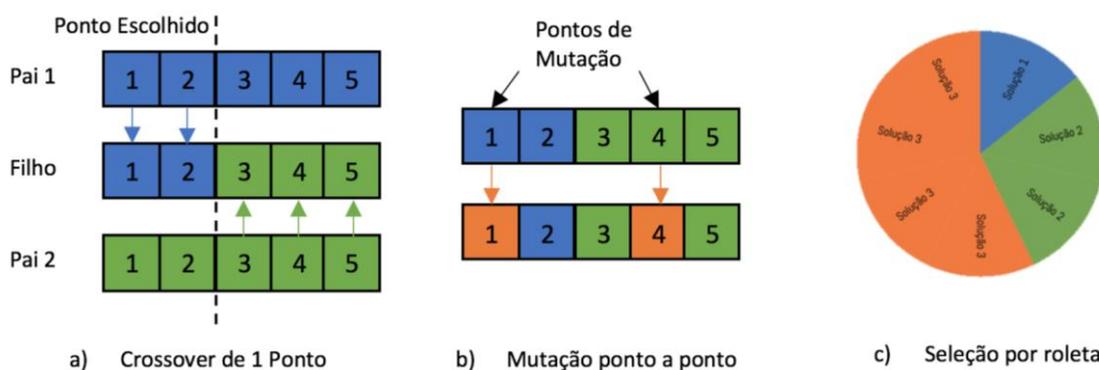


Figura 2: Exemplos de funcionamento dos operadores de cruzamento, mutação e seleção. Em “a” o ponto de corte foi definido entre os genes 2 e 3 do cromossomo. Em “b” os genes nas posições 1 e 4 sofreram mutação. Em “c” percebe-se que a solução 3 (laranja) possui mais posições na roleta, logo será maior a chance de ela ser escolhida no operador.

A escolha dos operadores adequados é decisiva para o sucesso do algoritmo evolucionário. Para definir os melhores operadores é necessário fazer testes com algumas configurações e medir o desempenho do algoritmo nessas condições. Quanto mais testes com diferentes combinações de operadores, menor a incerteza relacionada ao potencial da configuração escolhida para resolver o problema. Contudo, uma primeira filtragem de operadores pode ser feita antes mesmo dos testes. Alguns operadores são construídos para

operarem melhor com determinadas representações de gene, isto é, por definição alguns operadores são inviáveis para determinadas representações, as quais podem ser, por exemplo, binária, inteira ou ponto flutuante. Assim, um operador de mutação que atua invertendo valores binários (de 1 para 0 e vice-versa) em um cromossomo de representação binária não irá ser útil para um algoritmo cujo problema foi modelado com representação inteira. A representação que será utilizada no gene se torna uma primeira etapa antes do ajuste paramétrico, uma vez que a depender da que será escolhida para o problema o conjunto de operadores disponíveis pode ser bem diferente. A definição do tipo de representação é dependente do problema a ser resolvido e pode também ser variável para um mesmo problema a depender do que se pretende otimizar. Se é otimizado um problema de ligar ou desligar um conjunto de semáforos, a representação binária pode ser suficiente se desejarmos obter como solução as configurações de semáforo que devem estar ligadas e desligadas ao mesmo tempo, já se a otimização for sobre o tempo que cada semáforo deve permanecer em cada estado, uma representação de ponto flutuante é mais adequada. Por este motivo a escolha do que se quer otimizar precisa estar bem definida antes de iniciar a configuração do algoritmo, ou seja, é preciso que esteja claro quais são os objetivos do problema.

2.4.4 Objetivos e Algoritmos apropriados

A maioria dos problemas do mundo real envolvem solucionar mais de um objetivo simultaneamente [LAFETÁ, 2016]. A determinação da dimensão do problema é outra etapa importante para decidir qual algoritmo evolucionário usar para resolvê-lo. Divide-se os algoritmos em três grandes grupos sendo eles o mono-objetivo, multiobjetivo e muitos objetivos.

Nos algoritmos mono-objetivo o problema é modelado de forma a definir uma solução ótima com apenas um cromossomo, que será considerada como ponto de parada do algoritmo caso seja encontrada.

Quando se deseja utilizar mono-objetivo para resolver um problema com mais de um objetivo, é necessário um critério que normalize a importância de todos eles ou que atribua um peso a cada um de modo que uma única nota possa ser dada ao resultado. O problema desta abordagem é que não é trivial definir os pesos para cada preferência e elas por si só não são estáticas, mudando com o tempo e/ou com o usuário, além disso esta abordagem não é escalável, se tornando mais difícil normalizar os objetivos à medida que eles aumentam em quantidade e se tornam conflitantes entre si. Para evitar estes

problemas, os algoritmos multiobjetivo e para muitos objetivos passam a ser opções mais adequadas. A diferença entre essas abordagens está no número de objetivos com os quais cada uma se torna mais adequada, sendo os multiobjetivos recomendados para até 3 objetivos e os para muitos objetivos mais de 3.

Na otimização multiobjetivo existe uma função de aptidão para cada objetivo, e cada solução é classificada com mais de uma aptidão. A escolha pelas melhores soluções não pode ser dada de forma comparativa como na abordagem mono-objetivo, onde comparar o valor da aptidão de cada solução era suficiente para determinar qual era a melhor. No universo multiobjetivo, trabalha-se com o conceito de dominância. Como pode ser visto em Hashimoto (2004) dados dois vetores $x, y \in X^*$, é possível afirmar que o vetor x domina y se $f_i(x) \leq f_i(y), \forall i$. O conjunto de soluções não dominadas de um algoritmo multiobjetivo é chamado de **frente de pareto**. Uma frente de pareto ótima é composta pelas soluções não dominadas que resolvem o problema de forma ótima.

Diversas abordagens já foram propostas ao longo do tempo para resolver problemas multiobjetivo. Assim como nos algoritmos mono-objetivo, o algoritmo genético ganha destaque entre os multiobjetivo pioneiros, podendo mencionar por exemplo o *Multi-Objective Genetic Algorithm* (MOGA) e o *Non-Dominated Sorting Genetic Algorithm* (NSGA). Com o avanço dos estudos na área de algoritmo multiobjetivo houve evolução nestas abordagens e hoje os algoritmos multiobjetivos comuns são, por exemplo, o NSGAI e o *Strength Pareto Evolutionary Algorithm 2* (SPEA2). Mais recentemente, com aumento do interesse da sociedade científica pelo espaço com mais de 4 objetivos, vem ganhando destaque algoritmos como o NSGA-III.

2.5 Técnicas de otimização para geração de recomendações alimentares

Nesta seção são mostrados alguns exemplos de algoritmos propostos em outros trabalhos para a obtenção de recomendações alimentares. A Seção 2.5.1 aborda exemplos que utilizaram um objetivo, a Seção 2.5.2 aborda autores que utilizaram muitos objetivos.

2.5.1 Soluções mono-objetivo

Para melhorar a experiência de uso dos aplicativos nutricionais, muitos trabalhos acadêmicos têm sido desenvolvidos com o objetivo de tornar o processo de escolha de alimentos mais automatizado (CHEN, 2018; FRANCO, 2016; JENSEN, 2015). Técnicas de IA podem ser usadas para, dadas as necessidades nutricionais do paciente e suas

restrições, gerar recomendações que se adequem a estas necessidades. O desafio está em gerar recomendações que não apenas preencham corretamente os limites nutricionais dos usuários, como também consigam sugerir refeições que façam sentido do ponto de vista da coerência da refeição e gosto do usuário, sendo tudo isto em um tempo aceitável. Dentro desta proposta, Chen (2018) desenvolveu o sistema PERSON (do inglês *Personalized Expert Recommendation System for Optimized Nutrition*) que inclui o armazenamento e categorização automática dos alimentos em grupos, análise nutricional e recomendação automática para os usuários. Para cada uma das etapas, uma técnica diferente de IA pode ser usada, segundo o autor “na área de cuidados de saúde, acredita-se que a introdução de análise de *deep learning* em particular redes neurais é a razão do porquê os sistemas de recomendação têm melhorado nos últimos anos”. Nos módulos de recomendação a computação evolucionária é usada com diferentes tipos de algoritmos para a geração otimizada das sugestões. Diferentes estudos testaram a eficácia de diversos algoritmos e funções de aptidão, os resultados embora satisfatórios ainda possuem pontos de melhoria importantes a serem trabalhados antes que esta tecnologia seja aplicada no mercado.

Em Chen (2018) foi utilizado um algoritmo genético cujo cálculo da aptidão era feito conforme as seguintes etapas:

1. Assumiu-se que a probabilidade de um usuário seguir uma recomendação do sistema era máxima se metade dos alimentos disponíveis para uma determinada refeição eram oferecidos. A teoria é que mais de 50% dos alimentos torna-se um problema pois os usuários ficam em dúvida se o sistema está de fato filtrando os alimentos de acordo com suas necessidades. Em contrapartida, valores inferiores a 50% frustram os usuários devido ao baixo número de opções. Partindo deste pressuposto, é calculada a probabilidade de o usuário seguir a recomendação fornecida pelo sistema usando uma gaussiana.
2. Calcula-se a soma dos nutrientes ingeridos pelo usuário com a recomendação do sistema e o valor sem a recomendação.
3. Calcula-se a média dos nutrientes presentes no conjunto de alimentos possíveis para uma determinada refeição.
4. Calcula-se a redução no consumo de nutrientes pela subtração dos valores calculados em 2 e 3.
5. O valor calculado em 4 é dividido pelo número de alimentos pertencentes a recomendação e posteriormente normalizado.

6. Por fim, a aptidão era determinada pelo somatório do valor de redução de consumo normalizado calculado para cada um dos nutrientes considerados. Neste trabalho foram usados 5 nutrientes, sendo eles a energia, gordura, sal, açúcar e proteína.

O AG utilizado possuía o seguinte conjunto de configurações:

- Seleção de sobreviventes baseado no melhor valor de aptidão.
- Cruzamento de N pontos com *swapping*.
- Mutações ponto a ponto em alguns membros da população com baixa probabilidade.
- 1000 iterações.

Mediante medição da performance do algoritmo em 3 diferentes conjuntos de alimentos pertencentes a 4 grupos distintos que foram montados pelo próprio sistema no módulo de categorização, constatou-se que houve redução significativa no consumo de determinado nutriente em determinado grupo de alimento, como energia em arroz e massa e gordura em biscoitos e cereais. Tal algoritmo mostrou sugestões melhores para um determinado nutriente quando a maioria dos produtos pertencentes ao grupo de alimentos considerados não apresentava teores baixos ou até mesmo zero destes nutrientes.

Outra constatação importante do trabalho de Chen é referente ao processo de categorização dos alimentos. Como mencionado, o sistema PERSON incluía a categorização automática no sistema por meio de reconhecimento de palavras e redes neurais. Embora muito útil, Chen destaca que o sistema precisava ser manualmente ajustado para evitar problemas devido a erros de categorização quando na etapa de otimização com o AG. Esta dependência de grupos de alimentos bem definidos pode ser apontada como um ponto de melhoria no trabalho de Chen.

As porções de alimento utilizadas pelo algoritmo de Chen eram fixadas em 100 gramas (Gr). Não utilizar a porção como objetivo do algoritmo torna possível a utilização de representação inteira e diminui complexidade, mas reduz as opções que podem ser oferecidas. Sugerir uma porção de 100gr de peito de frango pode ser muito ou pouco dependendo da dieta. Neste contexto o tratamento das porções dos alimentos é outra oportunidade de melhoria observado no algoritmo de Chen.

Isowaka (2015) também aplicou a computação evolucionária com uma abordagem diferente. O objetivo do seu trabalho era desenvolver uma solução que fosse capaz de gerar recomendações de cardápio em um ambiente hospitalar para 5 dias, este ambiente

é particularmente interessante para este tipo de trabalho pois, ao mesmo tempo que é necessária a aplicação de uma dieta controlada, ainda existem várias restrições alimentares a depender do tipo de paciente e tratamento. Assim como uma preocupação em melhorar a qualidade de vida nos hospitais por meio de um cardápio mais adaptado ao gosto do paciente, o sistema de Isowaka precisava considerar uma combinação de requisitos e restrições no processo de confecção do cardápio. Em seu trabalho, ele ressalta que existem várias formas de controlar a ingestão de alimentos em uma dieta, neste estudo foram utilizados dois tipos: o controle de ingestão por nutrientes, do inglês *Intakes for food constituents* (IFCs), e ingestão por grupo de alimentos, do inglês *Intakes for food groups* (IFGs). Ambos os métodos de controle são considerados no cálculo da aptidão. A abordagem usada define a porcentagem de atendimento aos objetivos definidos dentro de cada método. No IFC, são definidos os limites nutricionais baseados nos valores de macro e micronutrientes, já no IFG é definido o número de alimentos em cada grupo alimentar, estes objetivos são definidos para cada refeição do paciente. O algoritmo genético usado faz a composição das refeições candidatas e realiza as seguintes operações para o cálculo da aptidão de um cromossomo:

1. É calculada a porcentagem que o valor de IFC e IFG tem em relação ao objetivo definido para cada nutriente/grupo alimentar;
2. São calculados a média e o desvio padrão entre os valores obtidos em “1”;
3. A diferença entre os valores obtidos em “2” é o valor da aptidão.

Para garantir que soluções inválidas não sejam fornecidas aos usuários a aptidão é ajustada para zero se uma solução apresentar dois ou mais alimentos iguais em uma mesma refeição e/ou contenha alimentos pertencentes a um subgrupo alimentar que o paciente seja alérgico. Ao mesmo tempo, para aumentar a chance de o paciente seguir a recomendação, o sistema permite a inserção de até 5 alimentos preferidos e os distribui uniformemente ao longo dos 5 almoços gerados.

O AG do trabalho de Isowaka apresentava o seguinte conjunto de configuração:

- Seleção de sobreviventes com roleta e elitismo.
- Crossover uniforme operando com 90% de probabilidade.
- Operador de mutação de abordagem não declarada com 1% de probabilidade.
- População de tamanho 60.
- Número de iterações não informado.

Isowaka realizou testes com o seu algoritmo usando usuários simulados, de diversos perfis físicos. Os resultados obtidos possuíam refeições com alto valor de aptidão e com boa variedade ao longo dos dias, o sistema era configurado para gerar refeições de almoço e jantar para um período de 5 dias. Para alguns perfis de usuário, algumas recomendações possuíam aptidão abaixo de 60%, isto foi identificado no trabalho como consequência do número limitado de alimentos presentes no banco de dados.

Algumas observações sobre o trabalho de Isowaka podem ser feitas. A aplicação de restrições no algoritmo pode ter o condicionamento a apresentar soluções adequadas, porém impedir que alimentos sejam repetidos em uma mesma refeição pode não ser uma boa prática para a escalabilidade do uso. Um número limitado de opções de alimentos foi usado nesse experimento, onde o banco de dados possuía 72 opções de alimentos, e neste contexto a performance de um algoritmo genético pode não se mostrar tão vantajosa quanto em problemas com maior volume de dados, onde seu uso é necessário devido à complexidade do problema de busca. Assim os resultados apresentados podem divergir bastante dos que seriam obtidos em um contexto com mais opções. Adicionalmente sobre a questão das restrições, encontra-se a mesma problemática que foi vista no trabalho de Chen a respeito das porções dos alimentos. Como não é possível ter duas porções de quase nenhum alimento em uma mesma refeição (é mencionado no texto que arroz branco é uma exceção a esta restrição), os resultados apresentados pelo algoritmo se tornam engessados e podem não agradar aqueles usuários que preferiam ter menos de um item e mais de outro no seu prato. Já a restrição de alergias foi implementada de modo que pode prejudicar a evolução. Ao eliminar soluções candidatas de forma brusca já no início da evolução, perde-se a possibilidade que essa solução seja cruzada com outra no futuro, eliminando o alimento alergênico de forma orgânica e aproveitando boa parte do cromossomo, que pode ser uma potencial solução final. Em contrapartida, o sistema de incentivo usado que incluía alimentos preferidos pelo paciente automaticamente no resultado também parece ser uma abordagem engessada, ao não dar nenhuma chance, mesmo que pequena, de algo novo vir a ser sugerido e assim, limitando a busca do AG. Por fim, Isowaka optou por usar um algoritmo genético que embora objetivasse encontrar uma solução que se adequasse a dois objetivos (IFC e IFG) combinava os resultados dos dois em um único objetivo, isso faz com que ao combinar os dois valores pode acontecer da deficiência em um corrigir um excesso do outro e isso pode vir a ser um problema. Também os algoritmos mono-objetivo apresentam apenas uma única solução candidata por vez, se apenas for usado o resultado de uma simples execução o usuário pode ficar

frustrado por ter poucas opções disponíveis, isso pode ser contornado executando o algoritmo mais de uma vez para um mesmo usuário nas mesmas condições de requisitos e restrições. Porém devido à natureza estocástica do AG soluções repetidas ainda podem acontecer, uma solução muito adequada as necessidades do paciente pode estar presente na população e ela vai impedir a evolução de soluções não tão boas ao combinar todos os objetivos, mas que pode ser melhor em um deles isoladamente e aceitável em outro, o que pode torná-la uma solução válida para apresentar ao usuário e aumentar seu poder de escolha.

2.5.2 Soluções multiobjetivo

Outros autores adotaram o uso de algoritmos multiobjetivos em sistemas de recomendação alimentar. Wang (2012) definiu o modelo de nutrição de alimentos como conceitualmente multiobjetivo, podendo ser modelado em um vetor. Pode-se considerar que qualquer alimento possui consigo um vetor de nutrientes, dos quais se deseja otimizar a busca de parte destes que combinados atendem a uma série de objetivos nutricionais. Em seu trabalho Wang propôs um novo algoritmo que pretendia ser mais performático no espaço multiobjetivo, se diferenciando de outros por ser uma solução mais escalável. Em seus testes ele confrontou algoritmos já consolidados no mercado como o SPEA e o NSGAI e obteve melhores resultados no espaço 4-D e 6-D. O algoritmo de Wang, intitulado no trabalho de FMOGA (do inglês *Fuzzy-Dominance-Driven Multiobjective Genetic Algorithm*) se baseava na procura por *fuzzy* para obter os melhores resultados em um espaço em que o tamanho da dimensão tornava difícil a busca por dominância tradicional dos algoritmos anteriores. Apesar dos resultados promissores do trabalho de Chen aplicados a área de recomendação alimentar, o FMOGA não se consolidou na comunidade científica enquanto ferramenta para otimização para muitos objetivos, hoje existem outras soluções mais consolidadas para trabalhar neste espaço como o NSGA-III, mas seu trabalho mostrou que o problema de recomendação alimentar é do tipo para muitos objetivos. De fato, um nutricionista pode desejar controlar ao mesmo tempo 4 ou mais nutrientes em determinadas dietas, sendo os algoritmos multiobjetivos tradicionais opções pouco viáveis nesse contexto. Para exemplificar, a Tabela Brasileira de Composição de Alimentos (TACO) usada como base de dados neste trabalho possui 23 macros/micronutrientes disponíveis para os alimentos, o nutricionista dificilmente vai

precisar regular todos em uma mesma dieta, mas limitar as escolhas a cerca de 3 é limitar demais o potencial da solução.

Pei (2009) usou Evolução Diferencial (ED) que propunha solucionar o problema de geração de recomendações de forma mais precisa, sendo indicado como uma melhoria em relação ao algoritmo genético. O algoritmo desenvolvido no trabalho de Pei, difere dos demais apresentados pelos seus objetivos. Enquanto os trabalhos anteriores focavam nos nutrientes e sua distribuição, Pei mostra que a computação evolucionária pode ajudar em outros tipos de problema. É citado entre outros exemplos, problemas de otimização do custo do alimento, objetivo que pode ajudar a trazer resultados mais adequados a realidade financeira de cada paciente. A estrutura do alimento, que pode ser necessária para algumas dietas, por exemplo as que restringem alimentos que possuem carboidratos não complexos, a distribuição dos alimentos ao longo das refeições, dentre outros. Assim, segundo Pei (2019) “O processo de solução deve encontrar todas essas variáveis para encontrar as condições de uma solução ótima. Claramente, este problema é uma combinação de problemas de otimização multiobjetivo”.

Para testar o seu algoritmo, Pei utilizou um conjunto de dados composto por 18 tipos de alimentos, distribuídos em 3 exemplos de refeição. Como medida de comparação foi usada a taxa de acerto dos resultados obtidos por cada abordagem, tendo sido testados dois algoritmos, um usando GA e outro usando DE, este último obteve resultados mais admissíveis. As conclusões tiradas pelo autor foram de que o DE é um algoritmo mais em linha com os hábitos alimentares dos humanos e superior ao GA em termos da precisão do resultado oferecido. Contudo o algoritmo também demonstra uma deficiência na composição da refeição, o autor apontou como ponto de melhoria para próximos estudos considerar o sabor final da refeição montada baseado na mistura entre os alimentos.

Embora pareçam promissores, nota-se que o conjunto de dados de teste utilizado por Pei foi bastante limitado, pois poucos alimentos foram usados e apenas um perfil de usuário de teste. Além disso, a diferença entre os resultados apresentados pelo DE e pelo GA foram pouco significantes, menor que 10%. Tendo o DE, apresentado os mesmos problemas de coerência da refeição apontados em trabalhos anteriores. Concluindo assim que o DE pode ser considerado como uma alternativa viável, mas não superior ao GA no problema de recomendação alimentar.

Na análise da literatura disponível referente a aplicação de computação evolucionária em problemas de recomendação alimentar, nota-se que as pesquisas estão progredindo ao longo dos anos, mas alguns problemas são antigos e recorrentes em todas

elas, como a coerência da refeição e as limitações de base de dados. Podem ser identificadas também duas vertentes de pesquisa para confecção das soluções, o uso de algoritmos mono-objetivo ou de multiobjetivo. Os trabalhos mais recentes desta área deram preferência pelos algoritmos mono-objetivo, com diferentes modelagens de função de aptidão e estrutura da solução. A vertente multiobjetivo apresenta resultados satisfatórios, mas com foco em construir algoritmos novos e compará-los com algoritmos já consolidados que apresentam resultados satisfatórios em um número limitado de objetivos. Embora o problema de confecção das recomendações tenha sido classificado como para muitos objetivos pelos autores, não foi encontrado o uso de um algoritmo consolidado pela academia, como o NSGA-III, aplicado a este problema.

O próximo capítulo irá apresentar como as informações coletadas nesta revisão da literatura foram utilizadas para nortear nosso trabalho bem como a metodologia adotada para o desenvolvimento do algoritmo proposto.

3 Metodologia

Este capítulo detalha os processos utilizados para a definição do algoritmo proposto. As Seções 3.1 e 3.2 dissertam sobre as motivações para escolha da abordagem multiobjetivo e o algoritmo utilizado como base para o desenvolvimento. A Seção 3.3 apresenta a base de dados utilizada pelo algoritmo. A Seção 3.4 detalha o algoritmo em termos de tecnologias utilizadas, representação e avaliação da solução bem como os operadores utilizados. Em 3.5 são apresentados os indicadores de desempenho definidos para avaliar os resultados do algoritmo. As Seções 3.6 e 3.7 apresentam o processo de coleta dos resultados do algoritmo e como eles são avaliados respectivamente e, por fim, a Seção 3.8 trata de como comparou-se os resultados do nosso algoritmo com algoritmos que utilizam otimização mono-objetivo para a geração de recomendações alimentares.

3.1 Escolha da técnica computacional para gerar as recomendações

O algoritmo desenvolvido neste trabalho fornecerá um plano alimentar diário baseado em metas nutricionais. Isto é, dado um usuário com determinadas metas de consumo de nutrientes e uma base de alimentos, o algoritmo deve escolher um conjunto de alimentos que satisfaçam, ou que estejam o mais próximo possível de satisfazer, estas metas. Nota-se que o sucesso de uma recomendação está fortemente atrelado ao cumprimento das metas alimentares definidas; isto é, o algoritmo deve ter como objetivo otimizar o valor nutricional de uma recomendação. Se um usuário tem como meta ingerir em uma refeição 660 kcal com cerca de 25 gramas de carboidrato, as refeições sugeridas devem conter alimentos cuja soma das calorias e de carboidratos esteja o mais próximo possível deste valor (a sugestão ótima seria aquela com valor exato). Note que as metas nutricionais podem ser generalizadas para qualquer macro ou micronutriente, tais como proteína, sódio, ferro, dentre outros. Em termos computacionais, tem-se um problema de combinação de alimentos em um conjunto de tamanho finito. Pode-se fazer uma analogia ao problema da mochila que, como já foi mencionado na Seção 1.2, pertence à classe de problemas NP-Difíceis, justificando assim a escolha de uma abordagem heurística, como algoritmos evolucionários. Além disto, vale ressaltar que segundo Chen (2018) o número de opções fornecidas em uma recomendação é relevante para a aceitação da mesma pelo usuário. Assim, utilizar uma técnica computacional estocástica como o AG se torna uma

opção superior às técnicas determinísticas, pois a cada iteração soluções diferentes e com nível similar de qualidade serão oferecidas.

3.2 Escolha do algoritmo

Neste trabalho, a aplicação da computação evolucionária utilizou algoritmos consolidados cuja eficiência em problemas de otimização já está testada e comprovada. Isto é, não é proposto um novo algoritmo evolucionário genérico para problemas de otimização, e sim uma versão aplicada de um algoritmo consolidado enriquecido com heurísticas específicas do problema de recomendação de alimentos.

O NSGA-III foi o algoritmo evolucionário utilizado como base para o algoritmo proposto neste trabalho. Analisando o problema da recomendação de alimentos da perspectiva de que cada nutriente (macro ou micro) pode ser um objetivo, tem-se que a quantidade de objetivos pode facilmente ser maior que três; na Seção 2.6.2 abordou-se o trabalho de Wang (2012) onde há cenários de otimização neste contexto para o universo 4D e 6D. Como mostrado em Deb (2014) para este número de objetivos o NSGA-III possui performance reconhecidamente superior ao NSGA-II por se enquadrar como um algoritmo para muitos objetivos. Além disso, os trabalhos anteriores que envolviam utilizar otimização multiobjetivo não tinham explorado o potencial de algoritmos para muitos objetivos, como o NSGA-III, aplicados a este tipo de problema.

A escolha das configurações do algoritmo ocorreu através da análise dos resultados apresentados em comparação com um conjunto de instâncias de configuração de teste, utilizando para isso critérios de qualidade definidos para o problema, que estão detalhados na Seção 3.5.

3.3 Base de dados e aplicação

Muitos dos trabalhos relacionados na Seção 2.5 apresentam uma prova de conceito de seus algoritmos sobre uma base de dados controlada, com poucos alimentos e poucas opções dentro dos grupos alimentares. Essa abordagem serviu para mostrar o potencial que os algoritmos evolucionários têm para gerar recomendações alimentares. Neste trabalho foi utilizada uma base de dados mais ampla, contendo mais opções de escolha para o algoritmo considerar ao longo da evolução. O desafio foi obter os dados com o maior número possível de informações nutricionais para os alimentos, assim o usuário poderia ter mais opções na hora de definir as metas. Vale ressaltar que, conforme pode ser visto na Seção 2.2, os dados fornecidos pelos apps nutricionais devem vir de fontes

confiáveis. Assim, um pré-requisito para a escolha dos dados é a procedência das informações prestadas.

Neste sentido, estes critérios foram atendidos pela Tabela Brasileira de Composição de Alimentos (TACO), uma base de informação com a composição dos alimentos consumidos no Brasil, gratuita e abrangente. Os dados foram obtidos por meio de análise laboratorial seguindo procedimentos validados e coordenados pelo Núcleo de Pesquisas em Alimentação da Universidade Estadual de Campinas (LIMA, 2011). Ela é composta por 597 alimentos, cada um com até 23 informações nutricionais distintas. A modelagem do banco de dados foi simplificada em duas tabelas: *Food* (alimento) e *FoodGroup* (grupo de alimentos), onde a primeira contém todos os 597 alimentos da TACO, e a segunda os 15 grupos alimentares definidos por ela. Essa forma de modelar possui flexibilidade para que sejam criados tantos grupos alimentares quanto forem necessários.

Este trabalho visa aplicar o algoritmo evolucionário em um contexto de uso onde um nutricionista pode definir múltiplos critérios nutricionais e, opcionalmente definir um conjunto de alimentos para determinada situação. Para fins de simulação, foi idealizado um cenário hipotético onde um nutricionista define metas nutricionais e um grupo de alimentos possíveis para compor o café da manhã do paciente, assim um décimo sexto grupo denominado “Café da Manhã” foi criado no banco de dados e alguns alimentos dentre os 597 foram selecionados para montar este grupo que foi utilizado ao longo das simulações de teste.

3.4 Conhecimentos do problema aplicados ao algoritmo

Esta seção mostra como os conceitos, relacionados ao problema de recomendação alimentar, foram utilizados na construção do algoritmo proposto nesta pesquisa. Em 3.4.1 é mostrado como as metas nutricionais foram definidas para os perfis de testes utilizados, uma vez que nesta etapa da pesquisa não houve a participação de usuários reais e nutricionistas. Em 3.4.2 é mostrado como uma solução candidata no algoritmo proposto é representada. A Seção 3.4.3 detalha como se deu o processo de filtragem da população inicial para cada perfil de usuário. As Seções 3.4.4 e 3.4.5 tratam dos operadores definidos bem como as modificações feitas para melhorar os resultados no domínio do problema. Por fim, a Seção 3.4.6 apresenta como é calculada a aptidão de uma solução candidata no algoritmo apresentado.

3.4.1 Definição das metas nutricionais

Conforme determinado pelo Conselho Federal de Nutrição (CFN) compete ao Nutricionista o papel de elaborar prescrição dietética (CFN, 2018). Na ausência deste profissional no grupo de pesquisa deste trabalho utilizamos, para fins de teste do algoritmo, as fórmulas para cálculo da Necessidade Energética apresentadas na Seção 2.3. Primeiramente, como visto na Seção 2.3.2, sabe-se que com os dados coletados na anamnese de uma pessoa é possível obter a recomendação energética por meio da equação de Harris e Benedict ou da OMS e FAO e do índice de atividade física. A forma de distribuir essa energia pode variar a depender da necessidade da dieta; é possível contudo, de forma genérica utilizar as recomendações estabelecidas segundo diretrizes nacionais, como as determinadas pelo Guia Alimentar para a População Brasileira, documento do Ministério da Saúde do Brasil e de distribuição gratuita que estabelece as seguintes diretrizes para a participação dos macronutrientes no valor da recomendação energética:

Macronutriente	Participação
Carboidrato	55% a 75%
Proteína	15% a 30%
Lipídio	10% a 15%

Tabela 4: Participação dos macronutrientes na recomendação energética de acordo com o Ministério da Saúde

O algoritmo proposto leva em consideração os dados da Tabela 4, as fórmulas para o cálculo da recomendação energética e as taxas de conversão apresentadas na Seção 2.4.1 para determinar a recomendação energética de forma autônoma. Assim é possível executar o algoritmo simulando o valor da recomendação energética utilizando valores próximos de uma situação real de uso.

Por fim, para a determinação das metas nutricionais, foram escolhidos três perfis de usuário distintos de modo que cada perfil tenha níveis diferentes de restrição em relação a quantidade de calorias ingeridas. Com o uso destes perfis o algoritmo pode ser testado para dietas mais e menos restritivas.

Perfil	Sexo	Idade	Altura (cm)	Peso (kg)	Índice de Atividade
1	Homem	21	172	70	Moderado
2	Mulher	25	160	55	Leve
3	Homem	25	180	80	Intenso

Tabela 5: Perfis de usuário utilizados nos testes de performance

Para cada perfil detalhado na Tabela 5 foi calculada a NE em kcal. Em seguida é calculada a proporção de calorias para cada nutriente utilizando a diretriz especificada em Brasil

(2014) e fixando as proporções em 68% para carboidrato, 12% para proteína e 20% para lipídios. Dado o valor em calorias para cada macronutriente o último passo é converter esse valor em gramas de nutriente utilizando as proporções mencionadas em 2.4.1 para então chegar no valor em gramas de carboidrato, proteína e lipídio que a recomendação do algoritmo deve ter como objetivo. Vale ressaltar que após essa última conversão os objetivos ficaram na mesma unidade de medida dos valores dos nutrientes da base de dados, o que permite que o algoritmo possa considerar o cálculo de aptidão das soluções candidatas de forma simplificada. Os valores das recomendações nutricionais diárias obtidos pelo algoritmo para os três usuários teste, bem como a distribuição em gramas dos macronutrientes utilizando as proporções mencionadas encontram-se na Tabela 6.

Usuário	Calorias (kcal)	Carboidratos (gramas)	Proteínas (gramas)	Lipídios (gramas)
1	2685,00	456,45	80,55	59,67
2	1845,39	313,72	55,36	41,01
3	3246,21	551,86	97,39	72,14

Tabela 6: Valores nutricionais diários recomendados pelo algoritmo para os usuários de teste

Caso a intenção do algoritmo seja recomendar todos os alimentos que o usuário deve consumir ao longo do dia, os valores informados na Tabela 6 seriam os objetivos da evolução. Na prática os testes do algoritmo foram feitos considerando 15% desses valores, correspondendo às metas nutricionais para uma refeição específica. Três objetivos foram definidos, cada um correspondendo a um dos três macronutrientes, pois como abordado na Seção 2.3.1, eles compõem o valor calórico do alimento. A vantagem de fazer a otimização sobre os macronutrientes ao invés das calorias é especificar melhor como deve estar composto o valor calórico do alimento e assim seguir as diretrizes do Ministério da Saúde mostradas na Tabela 4.

Com o número de objetivos definido em 3, pode parecer que o potencial do NSGA-III está sendo subutilizado neste trabalho. De fato, ao se trabalhar com até 3 objetivos, algoritmos multiobjetivo como o NSGA-II apresentam boa performance. A principal motivação para o uso do NSGA-III já foi explanada na Seção 3.2, e se embasa na flexibilidade para que sejam escolhidos quantos nutrientes, entre os 23 disponíveis, sejam necessários para cada dieta. A decisão de trabalhar no espaço 3D e não com mais dimensões foi tomada devido à inviabilidade de representar os resultados do algoritmo de

forma gráfica. Assim, para fins de estudo a proposta deste algoritmo utiliza três objetivos de modo a facilitar a visualização de alguns resultados, que serão apresentados em 3D.

3.4.2 Representação e escopo da solução

Nos exemplos de uso apresentados ao longo da Seção 2.2 os sistemas desenvolvidos elaboravam a recomendação sobre uma determinada refeição, isto é, um almoço ou jantar, sobre uma base de alimentos restrita. No algoritmo proposto neste trabalho, este comportamento é replicado, porém com um conjunto de dados mais rico. A intenção é mais uma vez simular um cenário de uso real, onde um nutricionista deseja montar, por exemplo, as sugestões para o café da manhã do paciente. Usando o algoritmo proposto neste trabalho, deve ser criado um conjunto que contenha todos os alimentos que o usuário espera consumir nesta refeição específica. Vale ressaltar que nesta etapa não é necessário se preocupar com a porção dos alimentos, que será definida pelo algoritmo, e sim com os alimentos que fazem sentido para o nutricionista/paciente naquela refeição. Dessa forma, tem-se que, para diferentes perfis de usuário, diferentes conjuntos de café da manhã podem ser definidos, sendo um contendo mais opções de frutas, outro contendo mais opções de raízes e assim por diante. Para o teste do algoritmo foi definido um grupo de café da manhã contendo mais de 100 itens dentre os presentes na tabela TACO e que foi utilizado nos testes de geração de recomendação para os três perfis de usuário. Vale ressaltar que, ao utilizar um grupo de alimentos por refeição, reduz-se as chances de uma combinação de alimentos incoerente ser oferecida, visto que alimentos esperados para uma determinada refeição e pertencentes a um grupo montado para um perfil específico de usuário tem maiores chances de combinarem entre si.

O balanceamento das calorias entre as refeições muda de acordo com a região e com o perfil de usuário, podendo o almoço ou o jantar ter ou não a presença de lanches entre eles. Foi adotado para fins de teste do algoritmo que as calorias esperadas para o café-da-manhã correspondem a 15% do valor diário recomendado. Com essa porcentagem, as metas nutricionais assumem valores que permitem que boa parte dos alimentos sejam elegíveis em uma recomendação como pode ser visto na Figura 3.

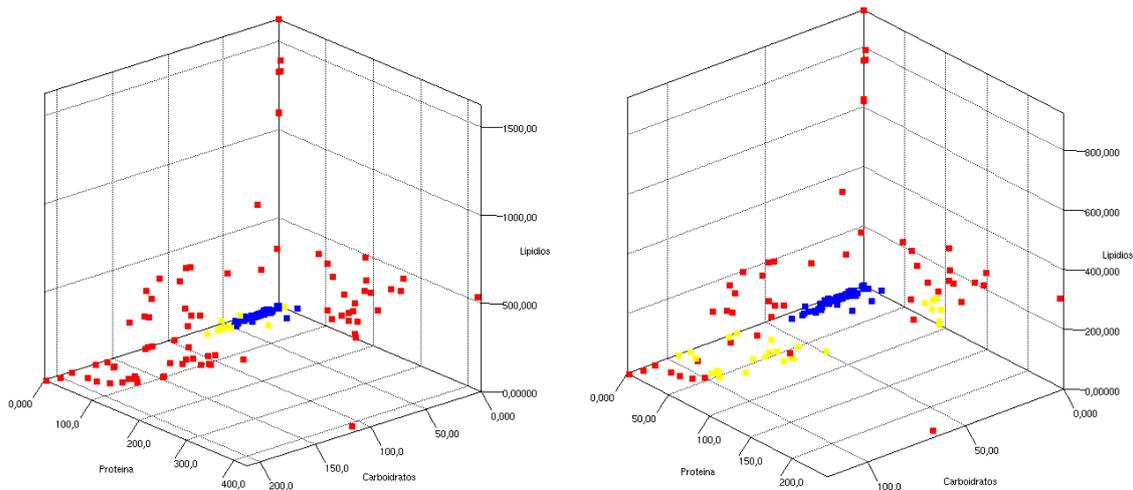


Figura 3: Distribuição dos alimentos no grupo de teste considerando as metas nutricionais do usuário com mais restrição (à esquerda) e com menos (à direita).

A representação de uma solução no algoritmo é feita por meio de uma sequência de inteiros de tamanho variável, onde cada posição da sequência representa o identificador de um alimento. Como visto nos exemplos da Seção 2.2, um dos problemas dos sistemas de recomendação é trabalhar com porções de tamanho fixo dos alimentos, e assim obter recomendações limitadas a estas porções. Neste trabalho foi adotada a estratégia de permitir a repetição de um alimento na solução (duas ou mais porções de um mesmo alimento) e utilizar tamanho variável na recomendação para aumentar o número possível de resultados. Assim, um alimento mais calórico que não poderia ser indicado em uma recomendação de tamanho fixo por deixar pouco espaço para outros alimentos, passa a ser recomendado para o usuário com o ônus de vir acompanhado de menos itens. Espera-se aumentar a possibilidade de o usuário gostar das recomendações fornecidas pelo algoritmo, baseando-se na premissa de Chen (2018) de que o número de opções fornecidas está relacionado com as chances de ele seguir a recomendação. Um ponto importante a se destacar é a granularidade das porções. Na tabela TACO, os dados nutricionais fornecidos para os alimentos estão relacionados a uma porção de 100 gramas. Assim, a estratégia adotada pelo Algoritmo proposto, usando apenas a base de dados da TACO, forneceria recomendações com porções variando em múltiplos de 100 gramas. Isto não é enxergado como um problema pois a granularidade não é limitada pelo algoritmo e sim pela base de dados, assim basta que novas entradas para outras porções sejam feitas e o algoritmo já está preparado para lidar com estas novas porções. Estratégia semelhante a esta é usada pelo aplicativo “Contador de Calorias”, ou no inglês *My Fitness*

Pal, que foi citado no trabalho de Franco (2016) como um dos mais populares aplicativos da área de nutrição. Ao decidir por não controlar as porções dos alimentos na forma de objetivo a representação da solução é simplificada (podendo ser inteira) e o algoritmo focado em otimizar os nutrientes.

3.4.3 Construção da população inicial

Dados os três perfis de meta nutricional exibidos na Tabela 5, tem-se que o usuário 2 possui metas nutricionais mais restritas (permite valores menores de nutrientes) e o usuário 3 possui metas menos restritas. Dessa forma, a Figura 3 mostra, através do esquema de cores usado, quais os alimentos dentre o universo pertencente ao grupo de teste escolhido são candidatos a fazer parte da recomendação final do algoritmo e como eles estão distribuídos. No gráfico, os pontos azuis representam os alimentos cujos macronutrientes são inferiores a 50% da meta, os amarelos superiores a 50% e os vermelhos superiores a 100%. Percebe-se que à medida que a dieta se torna mais restrita a variabilidade de alimentos disponíveis diminui (na Tabela 7 estão relacionadas a quantidade de alimentos dentro de cada faixa). Nota-se ainda que há uma maior concentração de alimentos azuis, assim as opções são reduzidas, mas ainda variadas, enquanto alguns alimentos destoaram bastante dos demais em um nutriente específico, não podendo ser recomendados para nenhum dos perfis de teste utilizados. Vale ressaltar que os alimentos com a última característica citada podem vir a ser elegíveis em uma recomendação desde que estejam presentes na base de dados em porção inferior a 100gr e/ou o paciente possua metas nutricionais mais permissivas.

Dada a forma como uma solução é representada neste trabalho, por meio dos identificadores dos alimentos no banco de dados e com porções múltiplas de 100 gramas, os alimentos que se enquadram na região vermelha do gráfico apontado na Figura 3 não precisam estar presentes em nenhuma etapa da evolução, visto que não seria possível que o gene que representa um alimento vermelho consiga espaço para fazer parte de uma solução viável em futuras gerações. Assim, foi inserido no algoritmo um método, de criação da população inicial, customizado que elimina os alimentos vermelhos da população. A busca por soluções viáveis pode evoluir sobre uma população com indivíduos realmente capazes de gerar um resultado satisfatório.

Faixa	Usuário 2	Usuário 3
-------	-----------	-----------

	Quantidade	Porcentagem	Quantidade	Porcentagem
< 50%	64	39,51%	75	46,30%
> 50%	11	6,79%	29	17,90%
> 100%	87	53,70%	58	35,80%
Total	162	100%	162	100%

Tabela 7: Distribuição dos alimentos pertencentes ao grupo de teste nas faixas de proximidade das metas nutricionais para os usuários teste com maior e menor restrição nutricional

Os alimentos azuis possuem maior probabilidade de estarem presentes em recomendações finais com maior variabilidade de opções, dada a distância entre os macronutrientes com a meta para a refeição. Já os alimentos amarelos, por estarem mais próximos do limite proposto, tendem a participar com menos opção de itens. Com a exclusão dos alimentos vermelhos, já na etapa da criação da população inicial, tem-se um conjunto (de alimentos possíveis) final de 75 (~46,30%) para o usuário 2 e 104 (~64,20) para o usuário 3. Vale ressaltar que mesmo no caso com menos alimentos disponíveis, há variabilidade das opções para os usuários em comparação com os trabalhos anteriores citados na Seção 2.2.

3.4.4 Operador de cruzamento

Ao realizar um cruzamento com pais de tamanhos diferentes, o principal cuidado que o operador deve ter é com o gerenciamento do tamanho das soluções filhas. Neste contexto, os pais contribuem com pesos diferentes em cada filho e soluções de tamanho maior terão mais chance de transmitir seus genes para futuras gerações. O *SmartNutriOnePointCrossover* (operador referenciado no código) elege dois pontos de corte, um em cada pai de acordo com o tamanho de cada solução, e realiza a troca de genes a partir do ponto escolhido até o fim da solução pai. A Figura 4 demonstra como o cruzamento é realizado. Neste operador, dois pais “A” e “B” irão gerar dois filhos “C” e “D”, de tamanhos variados, de acordo com as seguintes etapas:

1. São criados dois filhos “C” e “D” inicialmente iguais (cópias de “A” e “B”, respectivamente);
2. Um ponto de corte limitado ao tamanho de cada filho é definido;
3. Um limite de cruzamento é definido, baseado na solução de maior tamanho;
4. Para cada filho, a partir do ponto de corte definido em “2”, os genes do outro pai (o que não foi copiado) são transmitidos, sobrescrevendo os originais até o limite

de cruzamento definido em “3”. Se o outro pai não possui os genes (ele é menor do que o limite de cruzamento estabelecido) os genes originais são apagados.

Nesse passo a passo, nota-se que diferente dos operadores de cruzamento tradicionais, a contribuição de um pai pode não ser por meio da entrega de genes, e sim por meio da remoção dos genes do outro pai. Dependendo do ponto de corte escolhido, as soluções filhas podem ser mais ou menos parecidas com uma das soluções cruzadas; e como pode ser visto na Figura 4, pode-se dizer que soluções com tamanho maior possuem uma tendência a terem genes dominantes no processo de evolução.

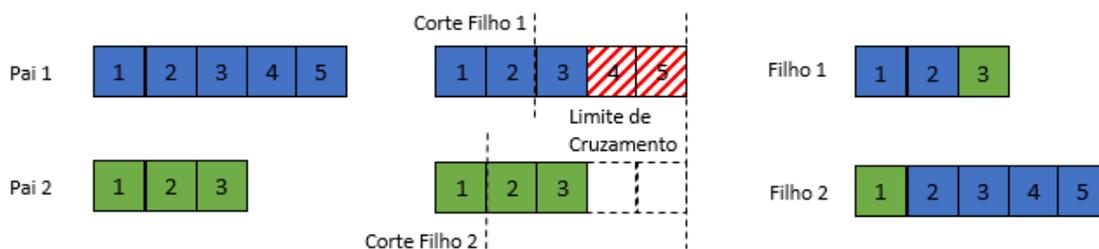


Figura 4: Funcionamento do operador de cruzamento usado no Algoritmo proposto

3.4.5 Operador de mutação

Assim como no operador de cruzamento, foi necessário definir o próprio operador de mutação para possibilitar soluções com tamanho variável. Além dessa motivação, houve a necessidade de tratamento do problema da “coerência da refeição”. Como citados nos trabalhos referenciados na Seção 2.5, o problema da coerência da refeição foi observado também nas soluções apresentadas nas primeiras versões do algoritmo. A modelagem da solução, utilizando grupo de alimento por refeição, já reduz a ocorrência deste tipo de problema, ao não misturar alimentos que comumente fazem parte de uma refeição com os de outra refeição. Contudo em uma base de dados variada como a tabela TACO ainda podem ocorrer situações em que itens que fazem parte de um mesmo grupo de alimentos não façam sentido em uma mesma configuração. Um exemplo de situação em que este efeito indesejado ocorre é ao recomendar o item açúcar. A tabela possui três tipos de açúcar: cristal, refinado e mascavo, mas não faz sentido em uma mesma refeição estes três itens estarem presentes. Foi para reduzir a ocorrência deste tipo de questão que foram criados subgrupos na entidade *Food*. A intenção é que os subgrupos sejam uma especialização dos grupos alimentares básicos, por exemplo, na tabela TACO tem-se o grupo de alimentos “Leite e Derivados”, dentro deste grupo pode-se extrair os subgrupos “Bebidas de Leite” e “Laticínios”. A definição dos subgrupos alimentares foi feita por

meio manual, visto que com as informações presentes na base de dados usada não foi possível reconhecer um padrão para extração desses subgrupos por meio de técnicas automáticas. Reconhecimento de palavras sobre o nome do alimento não foi cogitado pois já havia se mostrado deficiente em Chen (2018), mas há a hipótese de que seria possível identificar algum padrão mediante a análise comportamental das refeições reportadas pelos usuários. Esta última técnica pode ser considerada após a implantação do algoritmo em um sistema com usuários reais, de modo que se tenha os dados necessários para fazer esta inferência. A estratégia manual então foi preterida nesta etapa do trabalho. Ao todo, foram definidos 13 diferentes subgrupos alimentares, os quais podem ser consultados no Apêndice A.

O operador de mutação, chamado no algoritmo de *SmartNutriMutationWithRepair*, executa a alteração de uma solução mediante duas abordagens: perturbação genética ponto a ponto e reparação. Ambas as abordagens são baseadas nos subgrupos, mas possuem objetivos diferentes. A primeira é mais próxima de uma mutação convencional trocando o valor de um gene por outro pertencente ao mesmo subgrupo alimentar e, assim, evitando o problema de ter em uma mesma refeição os alimentos que não fazem sentido. A segunda abordagem realiza alterações mais profundas no cromossomo, removendo todas as repetições de subgrupo pertencentes em uma solução. Diferente da abordagem de perturbação genética, a reparação não executa sobre cada gene dentro de uma probabilidade, e sim sobre todo o cromossomo. Esta característica mais impositiva sobre a geração das soluções foi o motivo da reparação ter sido colocada no operador de mutação ao invés do cruzamento (onde é mais comum que ocorra a reparação). Quando há muita alteração no fluxo de execução de um algoritmo evolucionário, corre-se o risco de convergir os resultados de forma tendenciosa, reduzindo a variabilidade característica desta técnica. Como o operador de mutação trabalha com probabilidades pequenas em comparação com o de cruzamento, colocar a reparação neste operador garante que o algoritmo está executando com certa liberdade sobre a evolução, evitando a convergência tendenciosa.

Para funcionar corretamente, o operador conta com as probabilidades de mutação e de reparação, ao invés de apenas uma como é comum nos operadores tradicionais. No momento em que o operador é executado uma das abordagens é selecionada de acordo com as probabilidades ajustadas e então o cromossomo será alterado conforme os passos listados:

- Na mutação

1. Para cada gene do cromossomo é decidido se vai haver a mutação mediante a probabilidade de mutação;
 2. Em caso de mutação, o valor do gene é substituído por outro que represente um alimento do mesmo subgrupo do original. Assim, garante-se que a troca não irá causar, por exemplo, a situação dos açúcares.
- Na reparação:
 1. O subgrupo e o valor do primeiro gene do cromossomo são salvos em uma coleção;
 2. É verificado o subgrupo do próximo gene. Se estiver presente no mapa e o valor for diferente (outro alimento do mesmo subgrupo), um outro alimento pertencente a um subgrupo ainda não presente no mapa é escolhido e o valor do gene é substituído;
 3. O operador prossegue trocando os alimentos que são de grupos repetidos até que todo o cromossomo é percorrido.

Nota-se que, no pior caso, o cromossomo pode ter todos os seus genes (exceto o primeiro) alterados no caso de uma reparação e que nesse caso a performance do operador só é garantida se o número de subgrupos disponíveis seja maior do que o tamanho da solução. Nos testes do operador isso não foi um problema pois as soluções geradas tinham entre 2 e 6 opções de alimentos (há 13 opções de subgrupo). O algoritmo proposto foi configurado para garantir que soluções com número de itens maior do que o de subgrupos disponíveis não aconteçam.

3.4.6 Função de aptidão

Uma solução candidata no algoritmo é avaliada de acordo com os valores dos nutrientes pertencentes a cada alimento presente na solução. Durante o desenvolvimento do trabalho, duas formas diferentes de calcular a aptidão foram testadas. A primeira baseava-se na diferença entre a soma dos nutrientes de um determinado objetivo em uma solução candidata e o valor da meta nutricional, calculada usando a equação de Harris e Benedict apresentada na Seção 2.3. A segunda, utiliza um sistema de pontuação para cada objetivo que relaciona a distância entre o valor do macronutriente na refeição sugerida e a meta definida. No contexto da primeira forma de cálculo, ao observar os valores dos objetivos para cada usuário percebe-se que, como a proporção em gramas é diferente entre cada macronutriente, o valor da diferença em gramas por si só não é um bom indicador

da qualidade da solução como pode ser observado na Tabela 8, que apresenta exemplos de soluções do Algoritmo proposto para o café da manhã do usuário 1.

Macronutriente	Meta	Solução 1	Solução 2	Solução 3
Carboidratos	68,47	69 (0,53)	67 (-1,47)	73 (4,53)
Proteínas	12,08	14 (1,92)	14 (1,92)	12 (-0,08)
Lipídios	8,95	9 (0,05)	9 (0,05)	8 (-0,95)

Tabela 8: Exemplo de recomendações geradas para o usuário de teste 1. O valor em parêntesis representa a distância entre a aptidão da solução e o objetivo

Para fins de cálculo da aptidão foi considerado apenas a distância absoluta entre a meta nutricional estabelecida e o que foi apresentado na solução candidata, ou seja, avaliar com a mesma aptidão soluções que tenham 1gr de diferença para mais ou para menos do objetivo. Assim, os valores da aptidão de cada solução seriam os valores absolutos das diferenças exibidas entre parênteses da Tabela 8. Porém, como o valor de cada objetivo apresenta ordens de grandeza diferentes entre cada nutriente, o valor numérico da diferença pode ser mais ou menos significativo a depender do objetivo. Usando a solução 3 da Tabela 8 como exemplo, as diferenças apresentadas em termos percentuais seriam aproximadamente 6,21% para os carboidratos, 0,67% para as proteínas e 11,88% para os lipídios. Note que se fosse considerar apenas a ordem de grandeza numérica, esta recomendação seria mais deficiente no objetivo carboidrato. Já analisando em termos percentuais, o pior objetivo seria o valor dos lipídios. Assim o valor percentual se torna uma medida mais adequada para medir a aptidão de uma solução. Porém alguns pontos devem ser considerados antes de chegar à solução final adotada para o cálculo da aptidão. As primeiras execuções do algoritmo com a aptidão calculada considerando o percentual de diferença para as metas nutricionais apresentavam poucos resultados satisfatórios, por vezes até nenhum, a cada execução do algoritmo. A definição de satisfatório neste contexto é dada por uma convenção, onde dado um valor de meta nutricional é considerado aceitável que uma recomendação possua um desvio de $\pm 10\%$ do valor da meta. O valor de 10% foi obtido mediante uma conversa com um profissional da área de nutrição e foi inicialmente utilizado durante os testes, sendo colocado como parâmetro no algoritmo para possibilitar que seja alterado de acordo com as necessidades de cada aplicação. Na versão final do algoritmo, este conceito foi flexibilizado em prol da eficiência da evolução, o que levou a resultados melhores e mais variados de acordo com as métricas de qualidade estabelecidas.

Algumas hipóteses foram levantadas para tentar explicar o baixo número de soluções que conseguiam sobreviver ao longo da evolução e fazer parte do conjunto de soluções final do algoritmo:

- I. A base de dados usada não continha alimentos que possibilitariam resultados satisfatórios;
- II. Boas soluções estavam sendo perdidas ao longo da evolução;
- III. Alguns parâmetros precisavam ser ajustados de acordo com o momento da evolução de modo a tornar o algoritmo mais ou menos permissivo em relação à sobrevivência das soluções.

Referente a investigação da primeira hipótese, a Figura 3 mostra que existem alimentos suficientes na base de dados para a confecção de boas soluções do ponto de vista nutricional (ainda não se pode afirmar que as soluções façam sentido lógico para o gosto pessoal do usuário). A hipótese dois isoladamente leva a uma técnica simples e comum utilizada em algoritmos evolucionários, o elitismo. Em problemas em que boas soluções são muito difíceis de serem encontradas, faz sentido salvar cada boa solução para sobreviver transversalmente ao longo das gerações. Este conceito já está incorporado no NSGA-III, algoritmo usado como base para a construção do Algoritmo proposto. Da análise das hipóteses 2 e 3 em conjunto, criou-se uma estratégia que leva em conta o número da iteração do algoritmo para ajustar o critério de avaliação e assim alterar a permissividade para as novas gerações, aumentando a qualidade dos resultados gradualmente com a evolução. Para tal, abandonou-se o uso da diferença entre meta e aptidão da solução candidata e se passou a fazer o cálculo da aptidão baseado no valor da soma dos macronutrientes da solução, convertido em uma nota de 0 a 10 cujo objetivo passou a ser a maximização da nota. Para fazer esta conversão, foi usada a fórmula da gaussiana, a qual é apresentada na equação (1).

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (1)$$

A equação (1) distribui seus resultados na forma de uma curva normal, isto é $X \sim N(\mu, \sigma)$, para uma dada média μ e um desvio padrão σ . Este comportamento é utilizado no algoritmo proposto para a determinação das notas das soluções.

Em termos da função de aptidão do algoritmo, foi usada uma função gaussiana cuja média é a meta nutricional e o desvio padrão é a diferença aceitável para este valor, isso nos permite obter notas sempre positivas independente do sentido da diferença (para mais ou para menos) e diretamente proporcionais com a proximidade da meta nutricional.

O valor obtido pela função gaussiana é normalizado de modo que a nota seja máxima, quando o valor de x for igual à meta nutricional. Assim, tem-se a seguinte equação para o cálculo da aptidão de uma solução em determinado macronutriente M :

$$fitness_M = \left[\frac{10 \times \left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\sum_{i=1}^n N_i - \mu)^2}{2\sigma^2}} \right)}{\frac{1}{\sigma\sqrt{2\pi}}} \right] \times -1 \quad (2)$$

Onde:

- $\sum_{i=1}^x N_x$: Somatório de todos os nutrientes de um determinado tipo “ x ” para todos os “ n ” alimentos pertencentes à solução sob avaliação;
- μ : Meta nutricional para o nutriente em questão;
- σ : Desvio padrão aceitável, ajustado em cada iteração do algoritmo dinamicamente conforme segue.

O segundo termo do numerador da equação fornece o valor da gaussiana para o valor do somatório dos nutrientes da solução sob avaliação; o denominador é o valor da gaussiana no ponto máximo, isto é, o valor da meta nutricional. O resultado é multiplicado por 10 para normalização e posteriormente multiplicado por “-1” pois o algoritmo proposto é de minimização. Assim, a melhor nota possível para o nosso problema é -10.

O ajuste da permissividade do algoritmo ocorre ao alterar o desvio padrão. No início da evolução, o desvio padrão é maior, assim as soluções terão maiores chances de obter boas notas e não serem descartadas pelo algoritmo. À medida que a evolução avança, o desvio padrão vai sendo reduzido, tornando o algoritmo mais rigoroso no critério de seleção e melhorando a qualidade das soluções sobreviventes. Este processo segue desde a primeira até a última geração da população de indivíduos. A taxa com que o desvio padrão é decrementado depende do ajuste do desvio padrão máximo e mínimo configurado no algoritmo e do número de iterações da evolução. O primeiro é um parâmetro numérico adicional que o algoritmo proposto possui e foi testado com diversos valores juntamente com os demais parâmetros do algoritmo.

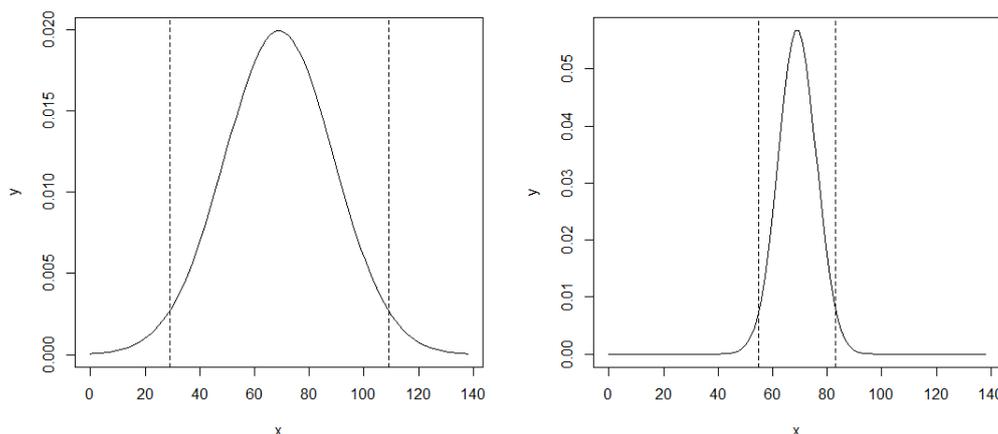


Figura 5: Exemplo de ajuste dinâmico do desvio padrão usado no cálculo da aptidão das soluções no início e no fim da evolução

Usando a meta apresentada na Tabela 8 para o objetivo carboidrato (68,47 gramas) e considerando que a evolução inicia em 30% da meta nutricional como valor admissível para o desvio padrão e termina em 10%, o comportamento do cálculo da aptidão será como mostrado na Figura 5. Nota-se que nas primeiras gerações, um número maior de soluções consegue obter boas notas, e terão mais chances de sobreviver para as próximas gerações, enquanto ao fim da evolução o critério para se obter a nota máxima se torna mais restrito. Uma solução que obteve nota satisfatória para seguir adiante na primeira geração, devido ao relaxamento dos critérios de avaliação, pode ser mal avaliada após algumas gerações e teria poucas chances de chegar ao fim da evolução. Contudo até este momento, a solução pode já ter compartilhado bons genes para as futuras gerações. Na prática, os resultados do algoritmo com esta abordagem de avaliação se mostraram promissores durante os testes, com aumento do número de recomendações oferecidas.

3.5 Indicadores de desempenho

Comumente os indicadores de desempenho para algoritmos multiobjetivos são ancorados na relação entre a curva de pareto do algoritmo com a curva de pareto ótima. Este conceito se torna por vezes complexo na prática, pois a solução ótima não está disponível para a maioria dos problemas. No caso específico do algoritmo proposto nesta pesquisa, em que diversas customizações com a finalidade de resolver um tipo específico de problema são feitas, faz sentido que os indicadores de qualidade sejam definidos de modo a garantir que o algoritmo está atingindo as premissas pelas quais ele foi preparado. Partindo deste raciocínio foram utilizados quatro indicadores de qualidade para medir o

desempenho do algoritmo proposto sob variadas combinações de parâmetros de configuração e para os cenários de teste explanados nas Seções anteriores. Os indicadores foram determinados de modo a avaliar principalmente a diversidade das soluções apresentadas, a convergência na região de soluções dentro de uma faixa aceitável de desvio dos objetivos e a qualidade do ponto de vista da montagem de uma refeição.

3.5.1 Indicador de espalhamento

Para indicar o quanto as soluções apresentadas estão distribuídas entre os extremos do espaço dos resultados é usado o indicador de espalhamento. Em um conjunto de refeições sugeridas, é desejável que existam opções que sejam adequadas do ponto de vista nutricional, mas que apresentem configurações distintas. Assim, aumenta-se a variabilidade de alimentos oferecidos no geral da recomendação e diminui-se as chances de o usuário não gostar de alguma das opções oferecidas, uma vez que ele terá uma faixa maior de alimentos para escolher. Este indicador promove a ideia de diversidade entre as soluções apresentadas, ao verificar a distância entre elas no conjunto de soluções. Para efeitos de comparação ao usar esse indicador, uma solução com maior valor de espalhamento é preferível.

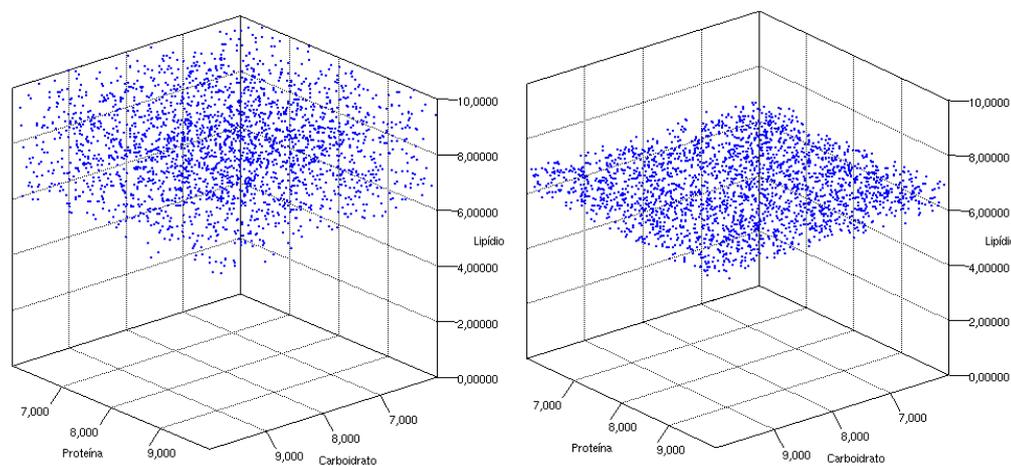


Figura 6: Exemplo de soluções com bom (à esquerda) e mau espalhamento (à direita).

Na Figura 6, os dois gráficos apresentam recomendações dentro de uma região de soluções aceitáveis. Porém as recomendações do gráfico à esquerda estão mais afastadas entre si, ou seja, estão mais espalhadas ao longo da região de boas soluções. Nestas condições, o indicador de espalhamento apresenta um valor maior para a solução a esquerda.

3.5.2 Indicador de espaçamento

Usado para verificar a convergência dos resultados, o indicador de espaçamento verifica a distância mínima entre as soluções. Enquanto o indicador de espalhamento tem relação com os extremos da área coberta pelas soluções, o de espaçamento verifica se dentro desta área as soluções estão bem distribuídas, isto é, preenchem uniformemente a área demarcada pelos extremos das soluções. Para o cálculo, foi utilizada a Equação (3) proposta em Van Veldhuizen (1999):

$$S = \sqrt{\frac{1}{n} \sum_{i=1}^n (\bar{d} - d_i)^2} \quad (3)$$

Na Equação (3):

- \bar{d} : Média das distâncias mínimas entre cada ponto da curva de pareto da solução.
- d_i : Distância euclidiana mínima de um ponto e os demais pontos da curva de pareto da solução.
- n : Número de pontos na curva de pareto.

Note que, se todos os pontos da curva estiverem equidistantes, o valor do espaçamento será zero. Ou seja, ao contrário do indicador de espalhamento, uma solução tem melhor espaçamento quando o valor do indicador é menor.

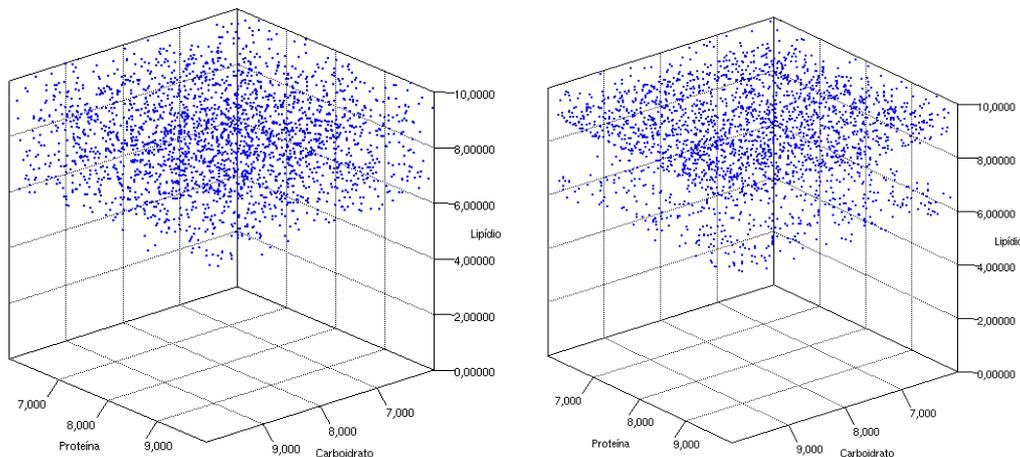


Figura 7: Exemplo de soluções com bom (à esquerda) e mau espaçamento (à direita)

Ambos os gráficos exibidos na Figura 7 apresentam uma área total coberta similar em termos de espalhamento. Porém, observa-se que o gráfico à direita contém mais espaços vazios, ou seja, as soluções estão mais afastadas entre si do que no gráfico à esquerda. O indicador de espaçamento terá um valor melhor para o gráfico da esquerda.

Em termos práticos, o que se pretende com este indicador é identificar soluções cujas opções estejam divergindo uniformemente. Isto é, os alimentos vão sendo

substituídos aos poucos, fornecendo recomendações com refeições semelhantes e com pequenas diferenças entre elas, como uma sugestão que troca apenas um alimento em relação a uma outra, mantendo o valor nutricional.

3.5.3 Hipervolume

Enquanto o espalhamento verifica a diversidade e o espaçamento averigua a convergência, o Hipervolume pode ser considerado um indicador híbrido, pois analisa a qualidade da solução nos dois aspectos. Seu funcionamento é observado quando, uma vez definido um ponto em uma região contrária à da convergência da evolução, calcula-se a soma da área compreendida entre este ponto e cada ponto pertencente à curva de pareto em análise. Valores altos neste operador indicam que as soluções estão bem espaçadas e com bom espalhamento, sendo portando um indicador ratificador dos apresentados anteriormente. Como foi mencionado na Seção 3.4.2.7, o valor máximo de aptidão que uma solução pode ter no Algoritmo proposto está no ponto (-10, -10, -10), logo o ponto de referência definido para o cálculo do hipervolume é o ponto (0, 0, 0). Este é um indicador cuja performance é bastante sensível ao número de objetivos, sendo considerado viável neste trabalho uma vez que a quantidade de objetivos foi fixada em 3 nos testes do algoritmo proposto.

3.5.4 Taxa de acerto

O indicador Taxa de Acerto relaciona a quantidade de recomendações pertencentes à curva de pareto da solução com as soluções que estão em uma região de 10% da meta nutricional estabelecida. É similar ao inverso do indicador Taxa de Erro, o qual verifica a porcentagem de soluções da curva de pareto sob análise que não estão presentes na curva de pareto ótima. No contexto de definição de aptidão usado no Algoritmo proposto, o conceito de curva de pareto ótima não se aplica, visto que o melhor dos cenários está concentrado em um único ponto do espaço de objetivos, que é o ponto (10, 10, 10). O uso da taxa de erro neste contexto não agregaria o valor desejado as métricas de qualidade, pois a maior parte das soluções teriam 100% de taxa de erro, o que não indica necessariamente que todas as soluções apresentam opções fora do limite de desvio aceitável definido. A Equação (4) foi usada no indicador de taxa de acerto é:

$$TA = \frac{\sum_{i=1}^n a_i}{n} \quad (4)$$

Na Equação (4) “n” é o número de pontos na curva de pareto sob análise e a_i será 1 se todos os objetivos do ponto “i” estiverem dentro da região aceitável e 0 caso contrário.

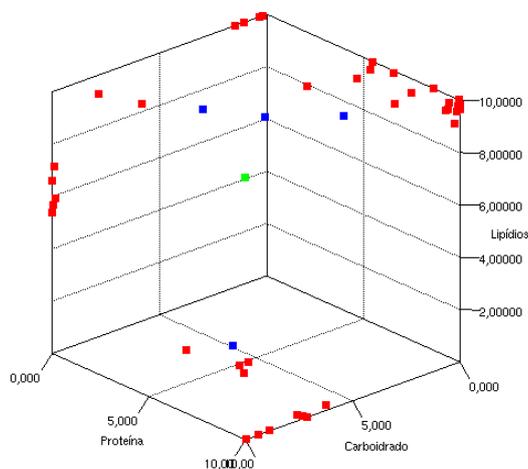


Figura 8: Classificação das soluções para cálculo da taxa de acerto

Na Figura 8, as recomendações de uma solução estão dispostas no espaço objetivo. O ponto marcado em verde representa o ponto ótimo (10, 10, 10) e não é considerado no cálculo do indicador. Os pontos em vermelho apresentam um ou mais objetivos fora da região aceitável (10% da meta nutricional) e os pontos em azul fazem parte desta região. No exemplo em questão, a curva de pareto tem 44 pontos, dos quais 4 estão marcados de azul. Logo, a taxa de acerto calculada é: $\frac{4}{44} \cong 9\%$. Note que este é um indicador onde se deseja maximização, ou seja, em teoria a melhor solução tem taxa de acerto igual a 100%.

O operador de taxa de acerto foi usado em conjunto com os outros três abordados nas Seções anteriores no processo de escolha da configuração final do algoritmo. De maneira resumida, a qualidade dos resultados é avaliada sob os aspectos de diversidade, convergência e atendimento às prescrições nutricionais. Os métodos utilizados para a manipulação destas informações estão detalhados na próxima seção.

3.6 Coleta de dados

Enquanto um algoritmo evolucionário pode ser definido por sua estratégia de evolução e seus operadores, uma instância de um AE corresponde ao conjunto de parâmetros numéricos que é aplicado a este algoritmo. A escolha de qual instância será utilizada depende da performance do algoritmo medida através das métricas de qualidade que foram estabelecidas na seção anterior. Para fazer esta avaliação, uma vez que o AE

tem natureza estocástica, é necessária a execução do algoritmo com determinada instância por um número definido de vezes para que sejam considerados os valores médios de cada métrica. O primeiro conjunto de valores para os parâmetros numéricos do algoritmo estão dispostos na tabela abaixo, onde foram utilizados valores oriundos de trabalhos anteriores e definidos de forma empírica. Durante o ajuste paramétrico, estes valores são modificados de acordo com os resultados do algoritmo em cada configuração.

Parâmetro	Valor 1	Valor 2	Valor 3
Número de Iterações	500	1000	1500
Probabilidade de Cruzamento	80%	90%	100%
Probabilidade de Mutação	5%	10%	20%
Probabilidade de reparação	50%	70%	90%
Desvio padrão inicial	10%	20%	50%
Desvio padrão final	5%	7,5%	10%

Tabela 9: Conjunto inicial de parâmetros numéricos utilizados nos testes de performance

Os parâmetros específicos do algoritmo foram definidos a partir das seguintes premissas:

- I. Para a reparação, o objetivo é verificar o nível de dependência desta abordagem na evolução e seu impacto nas demais métricas.
- II. Os valores de desvio padrão inicial e final irão demonstrar quão permissivo o algoritmo pode ser no início da evolução sem impactar a qualidade final dos resultados, bem como o quão restrito pode ser ao fim da evolução sem reduzir muito os pontos disponíveis na curva de pareto.

Para a etapa inicial de testes, foi utilizada a classe utilitária *MixedRangeCombinationIterable*, pertencente à biblioteca “javagl/Combinatorics” de Hutter (2013). Com o auxílio desta classe, foi possível obter 729 combinações distintas dos parâmetros detalhados na Tabela 9. Cada instância do algoritmo decorrente destas combinações foi executada repetidas vezes e os resultados dos indicadores de qualidade para cada execução foram salvos em um arquivo no formato CSV para posterior análise estatística. Os critérios e etapas de análise dos resultados estão detalhados na próxima seção. Por fim, além dos 4 indicadores de qualidade definidos no algoritmo, foi acrescentada ao arquivo a cardinalidade do conjunto de pareto das execuções, pois como mencionado, um dos focos do algoritmo é fornecer uma recomendação com diversidade de opções.

3.7 Análise de Dados

A definição da instância do algoritmo que melhor atende ao problema se deu em 3 etapas. Cada uma com o propósito de filtrar as melhores configurações sob circunstâncias variadas. A configuração considerada vencedora é a que apresentar o melhor resultado para cada tipo de usuário ao final das três etapas. Como os valores iniciais dos parâmetros numéricos foram definidos empiricamente a partir das hipóteses, ao final do primeiro ciclo de análise é verificado se a configuração vencedora está no valor extremo de algum dos parâmetros, o que pode indicar que há possibilidade de melhoria dos resultados com ajustes para além dos extremos pré-definidos. Neste último caso, é realizado um segundo ciclo de testes, com todas as 3 etapas, a fim de verificar se os resultados serão ainda melhores. A análise prossegue nesta lógica executando tantos ciclos quanto forem necessários.

Em cada etapa, são calculados os seguintes valores estatísticos para cada um dos cinco parâmetros coletados ao longo das execuções:

- Média Aritmética: Qual valor está equidistante dos extremos obtidos para o indicador ao longo das execuções.
- Mediana: Qual valor está no centro dos valores obtidos para o indicador ao longo das execuções. Uma mediana igual à média indica que os valores são distribuídos de forma simétrica.
- Desvio padrão amostral: O quanto os valores obtidos para o indicador oscilam em torno da média, expressando o nível de estabilidade dos resultados do algoritmo.

Os cálculos estatísticos são feitos em código Java, utilizando uma classe criada especificamente para o projeto e usando como base as equações consolidadas de cada métrica.

3.7.1 Etapa 1: Eliminação das piores configurações

Nesta primeira etapa as 729 instâncias do algoritmo precisam ser executadas para cada um dos 3 perfis de usuário de teste detalhados na Tabela 5. Devido às limitações de hardware, e possivelmente melhorias no gerenciamento de memória do algoritmo, o tempo para obter os resultados nestas condições é elevado. Assim, para esta primeira etapa, cada instância do algoritmo é executada 10 vezes, totalizando 21870 execuções. O critério de seleção usado foi a média da taxa de acerto de cada instância, sendo as 15

melhores neste parâmetro para cada perfil de usuário escolhidas para a próxima etapa. Em caso de empate entre duas ou mais soluções, todas são selecionadas.

3.7.2 Etapa 2: Escolha das melhores configurações

Se não houve repetições entre as configurações vencedoras para cada usuário na etapa anterior, esta etapa conterà no máximo 45 instâncias do algoritmo para análise. Neste contexto o número de vezes em que cada instância é executada é maior, de modo que venha a melhorar a representatividade dos dados estatísticos calculados. Assim, nesta etapa cada instância é executada 50 vezes, totalizando um máximo de 6750 execuções (45 instâncias x 50 x 3 perfis de usuário). O critério de seleção é o mesmo usado na etapa anterior (média da taxa de acerto), a diferença é que apenas as 5 melhores configurações para cada perfil de usuário serão eleitas para a etapa final. Assim como na etapa 1, em caso de empate entre duas ou mais soluções, todas são selecionadas.

3.7.3 Etapa 3: Escolha da configuração final do algoritmo

Se não houve repetições entre as configurações vencedoras para cada usuário na etapa anterior, esta etapa conterà no máximo 15 instâncias do algoritmo para análise. Mais uma vez cada instância é executada 50 vezes para cada perfil de usuário teste, totalizando um máximo de 2250 execuções (15 instâncias x 50 x 3 perfis de usuário). Todos os indicadores são levados em consideração nesta etapa, bem como a média aritmética, a mediana e o desvio padrão amostral. A abordagem definida é baseada em um sistema de pontos. O melhor valor estatístico em determinado indicador vale um ponto para a configuração em questão. No final, a configuração que obtiver o maior número de pontos será a escolhida. Vale ressaltar que, a própria definição dos parâmetros do algoritmo evolucionário é um problema de otimização, pois enquanto determinada configuração melhora um indicador ele também pode piorar outro. A estratégia de pontos definida é uma forma de escolher a configuração que desempenha melhor no maior número de critérios, sem dar mais peso para um indicador específico.

De forma resumida, os resultados são avaliados de acordo com a natureza de cada indicador. Os indicadores de espalhamento, hipervolume, taxa de acerto e cardinalidade da curva de pareto devem ter a maior média, em contrapartida o indicador de espaçamento deve ter a menor média. Para todos os indicadores, deseja-se o menor desvio padrão, com a menor diferença entre a média e a mediana. A Tabela 10 exemplifica como acontece o processo de escolha por meio de pontos usado na etapa 3.

Configuração	Hipervolume			Taxa de Acerto		
	Média	Mediana	Desvio Padrão	Média	Mediana	Desvio Padrão
1	900	700	100	0,4	0,4	0,05
2	1000	950	30	0,4	0,3	0,1
3	700	700	10	0,6	0,1	0,3

Tabela 10: Processo de análise dos resultados na terceira etapa

No exemplo acima, cada célula em verde representa um ponto no par indicador/dado estatístico, a configuração escolhida nestas condições é a terceira, visto que ela obteve o maior número de pontos. Em caso de empate a configuração com melhor taxa de acerto é escolhida, se houver persistência do empate os critérios de escolha passam a ser menor desvio padrão e menor diferença entre média e mediana nesta ordem.

3.8 Comparação com a abordagem mono-objetivo

Como visto na Seção 2.5, a academia explorou a geração de recomendações alimentares por meio de algoritmos mono e multiobjetivo. Os resultados obtidos nos trabalhos anteriores revelaram que as duas estratégias eram adequadas para a solução deste problema. Portanto, após definida a instância do algoritmo multiobjetivo proposto neste trabalho com melhor performance na geração das recomendações alimentares, é realizada uma comparação entre os resultados obtidos por ele e algoritmos que usam a abordagem mono-objetivo, para validar que esta nova proposta é uma opção viável frente às demais abordagens. Três abordagens diferentes de algoritmo mono-objetivo foram desenvolvidas para fazer a comparação. A primeira, inspirada no trabalho de Isowaka (2015), realiza a otimização por meio dos conceitos de IFGs e IFCs. Para os IFCs foram utilizados os mesmos limites nutricionais detalhados na Tabela 6, para os IFGs foram definidos 4 subgrupos alimentares dentre os disponíveis no grupo de café da manhã escolhidos aleatoriamente. Apesar de todos os subgrupos alimentares estarem disponíveis para a evolução no Algoritmo proposto, a estratégia de Isowaka exigia que eles fossem definidos em cada refeição, assim uma adaptação foi necessária. As outras duas abordagens foram desenvolvidas especificamente para os testes, isto é, sem inspiração em trabalhos anteriores. Uma delas faz o cálculo da aptidão baseada na diferença em

gramas entre a meta nutricional e o que foi recomendado pelo algoritmo para um nutriente específico e a outra calcula a diferença considerando a combinação dos três macronutrientes utilizados como objetivos separados no Algoritmo. O trabalho de Chen (2018) não foi utilizado, pois o que é considerado como solução para o algoritmo mono-objetivo não é uma recomendação de refeição, e sim o equivalente a uma lista de alimentos para o próprio usuário montar a sua refeição. Esta diferença de proposta entre o algoritmo de Chen e o Algoritmo proposto impossibilita a comparação entre os resultados. A fórmula da aptidão para o mono-objetivo inspirado no trabalho de Isowaka está descrita na Seção 2.5, para as duas outras abordagens foram criadas as equações (5) e (6):

1. Otimização de um nutriente específico x :

$$aptidão = \left| T_x - \sum_{i=1}^n x_i \right| \quad (5)$$

Na equação, T_x é a meta nutricional para um nutriente “ x ”, “ n ” é o número de alimentos pertencentes à recomendação, nos testes todas as soluções tinham o tamanho fixo de cinco alimentos, e x_i corresponde ao valor que cada alimento da solução possui para o nutriente sendo otimizado.

2. Otimização da combinação de vários nutrientes:

$$aptidão = \left| T_x - \sum_{i=1}^n x_i \right| + \left| T_y - \sum_{i=1}^n y_i \right| + \dots + \left| T_\theta - \sum_{i=1}^n \theta_i \right| \quad (6)$$

De forma análoga à equação para nutrientes específicos, a combinação é feita mediante a soma das diferenças, em valor absoluto, entre cada nutriente do alimento e o seu respectivo valor objetivo.

Nas Equações, (5) e (6) o melhor valor para a aptidão é zero, ou seja, são problemas de minimização. Para o algoritmo inspirado no trabalho de Isowaka a melhor aptidão é 100% (totalmente adequado para todos os critérios definidos), sendo caracterizado como um problema de maximização. Os detalhes das instâncias de cada algoritmo estão relacionados na Tabela 11.

Parâmetro	Isowaka	1 ou N nutrientes
Operador de Seleção	Roleta	Torneio
Operador de Cruzamento	Crossover de 1 ponto	Simulação binária (SBX)
Operador de Mutação	Mutação Polinomial	Mutação Polinomial
Número de Iterações	1500	1500
Tamanho da população	60	100

Probabilidade de cruzamento	90%	90%
Probabilidade de mutação	1%	1%

Tabela 11: Instâncias dos algoritmos mono-objetivo utilizadas para comparação

Para o operador de seleção por torneio, o número de soluções participantes foi definido em 10. O operador de mutação utilizado por Isowaka não foi especificado no trabalho, portanto foi escolhido o mesmo operador definido para os demais algoritmos.

Com os três algoritmos implementados, a comparação entre os resultados obtidos é feita utilizando os mesmos indicadores de qualidade usados para avaliar os resultados do Algoritmo proposto, para tal, curvas de pareto são montadas utilizando um conjunto de resultados não-dominados obtidos por execuções repetidas do algoritmo. O método utiliza os passos a seguir:

1. O algoritmo é executado 100 vezes, as variáveis do resultado de cada execução são exportadas em um arquivo de texto separado por tabulações (.TSV);
2. 100 Soluções no formato definido pelo Algoritmo proposto são criadas a partir das variáveis exportadas em “1”;
3. As soluções criadas em “2” são avaliadas usando as fórmulas estabelecidas para o Algoritmo proposto, onde a aptidão das soluções em cada um dos três objetivos é definida;
4. Uma vez avaliadas, as soluções dominadas são removidas e as restantes formam uma curva de pareto;
5. As mesmas métricas definidas para avaliar as instâncias do Algoritmo proposto são calculadas para a curva de pareto construída em “4”;
6. Os passos de “1” a “5” são repetidos 100 vezes, isto é, 100 curvas de pareto são criadas;
7. São calculados os mesmos valores estatísticos definidos para as instâncias do Algoritmo proposto, ou seja, a média, mediana e desvio padrão entre os 100 resultados de cada indicador.

Especificamente para o algoritmo que realiza a otimização em um nutriente específico, o passo 1 é repetido para cada um dos três macronutrientes utilizados como objetivo pelo Algoritmo proposto, assim, o número de soluções consideradas para a construção de cada uma das 100 curvas de pareto é multiplicado por três. Por fim, apenas o usuário de teste 1 foi usado como referência para a otimização das recomendações geradas pelos algoritmos. Ele foi escolhido por apresentar restrições nutricionais

equilibradas em relação aos outros dois perfis, o que já é suficiente para fazer a validação proposta.

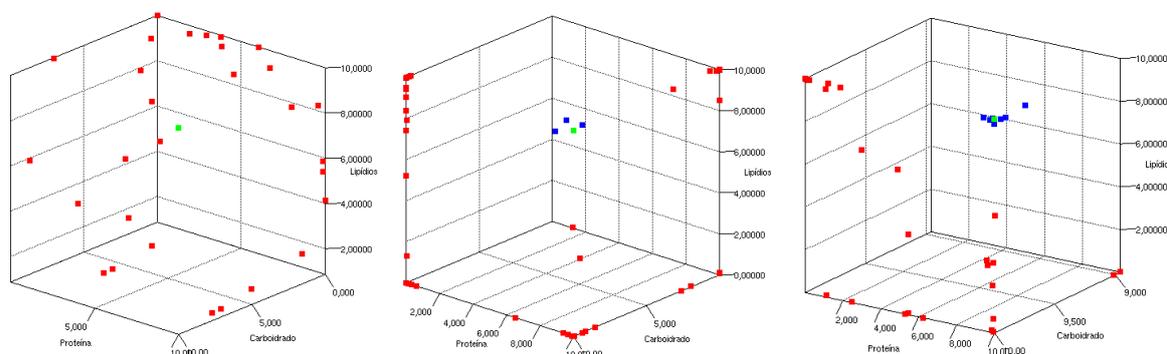


Figura 9: Exemplo de gráficos de pareto montados com os resultados dos algoritmos mono-objetivo, da esquerda para direita: trabalho de Isowaka, único nutriente e combinação de nutrientes

Os gráficos acima demonstram como os três algoritmos mono-objetivo usados para comparação distribuem seus resultados no espaço delimitado pelos três objetivos usados no Algoritmo (carboidrato, proteína e lipídios) quando executados conforme o método relatado. Os pontos em azul representam alimentos que estão dentro das metas nutricionais estabelecidas para o usuário de teste com desvio aceitável de $\pm 10\%$ e o ponto verde representa o ponto ótimo, ou seja, onde os três objetivos são iguais às respectivas metas nutricionais. Nota-se que o algoritmo de Isowaka espalha bem os resultados no espaço dos objetivos, mas não consegue obter nenhum resultado satisfatório do ponto de vista nutricional de acordo com os critérios de análise adotados no Algoritmo. Os demais algoritmos conseguem obter alguns resultados, mas a forma como os pontos são distribuídos com maior densidade nos extremos dos eixos do gráfico, mostram que na maioria das vezes esta técnica de otimização prejudica a qualidade de um nutriente em detrimento de outro. Vale ressaltar que apesar disso, existem situações em que os algoritmos conseguem encontrar pontos que todos os três objetivos estejam próximos das metas, mesmo com a otimização focando em apenas um nutriente (no gráfico do meio) ou em um valor único, resultado da combinação de todos os nutrientes (no gráfico a direita).

4. Resultados

Este capítulo apresenta os dados coletados durante os testes do algoritmo sob as diversas combinações de operadores numéricos e seguindo o processo de análise relatado no Capítulo 3. Foram realizados três ciclos de teste. A Seção 4.1 destaca os principais pontos do primeiro ciclo. A Seção 4.2 trata dos resultados do segundo ciclo e., na Seção 4.3 são feitas comparações entre os resultados anteriores e então é apresentada a relação entre a coerência da refeição e os demais indicadores de qualidade. A Seção 4.4 apresenta os resultados do terceiro ciclo de testes e como foi escolhida a instância final do algoritmo proposto. Por fim, na Seção 4.5 são apresentados os resultados obtidos pela comparação entre a instância do algoritmo proposto neste trabalho com outros três algoritmos que utilizam otimização mono-objetivo para gerar recomendações alimentares. Cada seção é dividida em subseções que representam cada etapa do ciclo de teste, conforme apresentado no capítulo 3.

4.1 Constatações do primeiro ciclo de testes

Nesta seção serão mostrados os resultados obtidos no primeiro ciclo de teste. Os resultados serão organizados divididos entre as duas primeiras etapas, cujo funcionamento foi explicado no Capítulo 3.

4.1.1 Resultados da primeira etapa de testes

Na primeira etapa de testes foram eliminadas configurações com desempenho desfavorável no indicador de taxa de acerto. De um total de 729 configurações distintas, apenas 37 (~5%) foram selecionadas para a segunda etapa. Apesar das divergências entre as melhores configurações para cada perfil de usuário observadas durante a primeira etapa, alguns padrões puderam ser observados. A Tabela 12 lista as melhores e piores configurações para cada perfil de usuário ordenadas por taxa de acerto.

Posição	Configuração						Usuário	$\bar{T}A$
	Iterações	Crossover	Mutação	Reparação	DP Inicial	DP Final		
1	1500	0,9	0,2	0,5	10%	5%	2	0,72
1	500	0,9	0,2	0,5	20%	7,5%	1	0,46
1	1500	1	0,2	0,5	10%	7.5%	3	0,34
729	1500	0,9	0,1	0,5	10%	10%	1	0,00

729	500	0,9	0,05	0,9	10%	10%	2	0,00
729	1500	0,8	0,05	0,7	10%	10%	3	0,00

Tabela 12: Amostra dos resultados obtidos na primeira etapa de testes

Os resultados dispostos na Tabela 12 mostram que a implementação do sistema de ajuste dinâmico do desvio padrão usado no cálculo da aptidão das soluções tem impacto direto na qualidade do resultado fornecido pelo algoritmo. Nota-se que nos dados da tabela, independentemente do valor dos outros parâmetros, quando o desvio padrão é mantido fixo, a taxa de acerto média é zero. Analisando todos os dados, das 81 configurações que contêm o desvio padrão fixo, a porcentagem destas que estão entre as 100 últimas posições do *ranking* por taxa de acerto é aproximadamente igual a 89%, 70% e 91% respectivamente para os usuários 1, 2 e 3.

Além do desvio padrão fixo, a etapa 1 também eliminou todas as configurações que continham o maior valor de reparação (90%). Nenhuma das 15 melhores soluções para cada perfil de usuário continha esse valor de reparação, o que demonstra que para a obtenção de bons resultados a interferência deste operador na evolução precisa ser reduzida. Também se observa que a taxa de acerto está relacionada com o perfil de usuário, o que tem dieta mais restrita (usuário 2) apresentou taxas de acerto superiores enquanto o menos restrito (usuário 3) apresentou taxas menores. A última configuração selecionada na primeira etapa para este usuário apresentou em média 26% de taxa de acerto, com cardinalidade da curva de pareto média igual a 52, ou seja, neste cenário o usuário receberia cerca de 13 opções de refeição.

4.1.2 Resultados da segunda etapa de testes

Das 37 configurações que foram testadas nesta etapa, 14 foram selecionadas para a etapa final de acordo com o *ranking* por taxa de acerto. Foram eliminadas todas as configurações com o menor valor de taxa de cruzamento e restaram apenas extremo superior de taxa de mutação e inferior de taxa de reparação. As taxas de acerto máximas foram reduzidas em comparação com os resultados da etapa anterior, isto pode ser explicado pelo maior número de dados na amostra usada para os cálculos (50 ao invés de 10), o que faz os valores encontrados nesta etapa serem mais confiáveis do ponto de vista estatístico. Diferente da rodada anterior, nenhuma configuração testada teve 0% de taxa de acerto, mas a diferença entre a última selecionada para cada perfil de usuário e a última colocada no *ranking* foi significativa, aproximadamente 77%, 97% e 43% para os

usuários 1, 2 e 3 respectivamente. Assim, foram removidas nesta etapa as configurações que limitavam a performance do algoritmo.

Posição	Configuração						Usuário	$\bar{T}A$
	Iterações	Crossover	Mutação	Reparação	DP Inicial	DP Final		
1	1000	0,9	0,2	0,5	50%	5%	1	0,39
2	1500	1	0,2	0,5	10%	5%	1	0,39
3	1000	0,9	0,2	0,5	50%	10%	1	0,39
4	1000	0,9	0,2	0,5	20%	10%	1	0,38
5	1000	1	0,2	0,5	10%	7,5%	1	0,37
1	1500	0,9	0,2	0,5	10%	5%	2	0,67
2	1000	0,9	0,2	0,5	20%	10%	2	0,64
3	1500	0,9	0,2	0,5	20%	7,5%	2	0,64
4	1500	0,9	0,2	0,5	10%	7,5%	2	0,64
5	1000	0,9	0,2	0,5	20%	7,5%	2	0,63
1	1500	0,9	0,2	0,5	20%	5%	3	0,26
2	1500	1	0,2	0,5	10%	7,5%	3	0,25
3	1000	0,9	0,2	0,5	50%	7,5%	3	0,25
4	1500	0,9	0,2	0,5	20%	10%	3	0,25
5	500	1	0,2	0,5	10%	7,5%	3	0,25

Tabela 13: Configurações selecionadas na segunda etapa de teste

A Tabela 13 contém as configurações vencedoras da segunda etapa de testes. As linhas marcadas em cinza representam configurações que foram comuns entre usuários de perfil diferente, por este motivo apenas 14 configurações foram selecionadas para a etapa 3. Com os dados da Tabela 13, pode-se constatar que:

- Em termos de taxa de acerto, todas as configurações apresentam resultados de qualidade semelhante. A decisão de qual é a melhor configuração deve ser feita considerando as demais métricas estabelecidas;
- Todas as configurações vencedoras tinham o maior valor de mutação dentre os disponíveis no conjunto inicial de valores, informados na Tabela 9. Isso pode indicar que resultados melhores podem ser obtidos com uma probabilidade de mutação maior;
- Todas as configurações vencedoras tinham o menor valor de reparação dentre os disponíveis no conjunto inicial de valores. Assim, reduzir a interferência deste operador na evolução pode melhorar os resultados ainda mais.

As constatações acima motivaram a criação do segundo ciclo de rodadas de teste, com reajuste dos valores de mutação e reparação para além dos limites atingidos. O segundo ciclo completo é compreendido pelas mesmas etapas e critérios de escolha, o conjunto de valores usados estão dispostos na Tabela 14.

No que tange à finalização do primeiro ciclo, a terceira etapa de testes foi realizada conforme processo estabelecido. Os resultados desta etapa não serão aprofundados nesta seção pois, com a execução do segundo ciclo de testes, os resultados desta etapa foram ignorados.

Parâmetro	Valor 1	Valor 2	Valor 3
Número de Iterações	500	1000	1500
Probabilidade de Cruzamento	80%	90%	100%
Probabilidade de Mutação	20%	30%	40%
Probabilidade de reparação	50%	40%	30%
Desvio padrão inicial	10%	20%	50%
Desvio padrão final	5%	7,5%	10%

Tabela 14: Conjunto de parâmetros numéricos reajustados que foram utilizados no segundo ciclo de testes de performance

4.2 Constatações do segundo ciclo de testes

O segundo ciclo de rodadas de teste comprovou a teoria que levou à sua criação e apresentou resultados melhores quando avaliado com os mesmos critérios do primeiro ciclo. São observados melhores resultados quando a probabilidade de reparação é menor, próxima da probabilidade de mutação. Nesta seção também serão apresentadas as motivações para a construção de um novo indicador de qualidade, a taxa de repetição de subgrupos, que será apresentada na Subseção 4.2.4.

4.2.1 Resultados da primeira etapa de testes

Uma característica observada durante o segundo ciclo de testes é que os perfis de usuário passaram a ser mais relevantes na configuração usada no algoritmo. Ocorreu um número maior de parâmetros distintos entre os perfis, sendo usadas para a segunda etapa 41 configurações das 729 possíveis (~6%). O comportamento dos resultados acerca da manutenção do desvio padrão fixo ao longo da evolução não foi alterado. Foi observado

que de fato uma taxa de reparação maior obteve resultados piores. A Tabela 15 compara as 3 melhores soluções de cada usuário entre os ciclos 1 e 2.

Posição/ Ciclo	Configuração						Usuário	T \bar{A}	Dif.
	Iter.	Cross.	Mut.	Rep.	DP I.	DP F.			
1/1	500	0,9	0,2	0,5	20%	7,5%	1	0,56	+26,8%
1/2	1000	0,9	0,3	0,4	50%	5%	1	0,71	
2/1	500	0,9	0,1	0,5	10%	5%	1	0,43	+62,8%
2/2	1000	0,9	0,4	0,5	10%	5%	1	0,70	
3/1	500	1	0,2	0,5	10%	7,5%	1	0,42	+66,7%
3/2	1000	0,9	0,3	0,4	20%	5%	1	0,70	
1/1	1500	0,9	0,2	0,5	10%	5%	2	0,72	+15,3%
1/2	1500	0,9	0,4	0,4	10%	7,5%	2	0,83	
2/1	1000	0,9	0,2	0,5	20%	5%	2	0,71	+15,5%
2/2	1500	0,8	0,4	0,3	50%	10%	2	0,82	
3/1	1500	0,9	0,2	0,5	50%	5%	2	0,71	+15,5%
3/2	1500	0,9	0,3	0,4	20%	10%	2	0,82	
1/1	1500	1	0,2	0,5	10%	7,5%	3	0,34	+17,6%
1/2	1500	0,9	0,2	0,3	50%	5%	3	0,40	
2/1	1500	1	0,2	0,5	20%	10%	3	0,31	+22,6%
2/2	1000	0,9	0,4	0,3	50%	5%	3	0,38	
3/1	1500	1	0,1	0,7	10%	7,5%	3	0,30	+23,3%
3/2	1500	0,9	0,2	0,3	10%	7,5%	3	0,37	

Tabela 15: Comparativo entre as 3 melhores soluções da primeira etapa de testes durante o primeiro e segundo ciclo

Analisando a amostra dos resultados dispostos na Tabela 15, pode-se perceber que em geral a média de execuções necessárias para obter os melhores resultados é superior com os novos valores dos parâmetros de mutação e reparação, e que a probabilidade de mutação se manteve oscilante entre as melhores soluções, indo do menor ao maior valor disponível neste ciclo. Já a probabilidade de reparação apresentou uma tendência de se manter próxima ao mínimo possível (30%), apesar de também oscilar. A conclusão desta primeira etapa mostrou que os resultados melhoraram significativamente, como mostrado na última coluna da Tabela 15, mas que são necessárias etapas adicionais de teste para

verificar, com dados estatísticos mais representativos, se os melhores resultados ainda convergem para algum extremo dentre as faixas de valores definidas neste ciclo para as probabilidades de mutação e reparação. Vale ressaltar ainda que devido à forma como o operador de mutação/reparação foi desenvolvido, o aumento na probabilidade de mutação ocasiona indiretamente em uma diminuição das chances de ocorrer a reparação.

4.2.2 Resultados da segunda etapa de testes

Resultados mais estáveis foram observados na segunda etapa de testes, isto é, ocorreu menos divergências entre as melhores soluções para os três perfis de usuário. O número de configurações que passaram para a terceira etapa de testes foi menor do que no primeiro ciclo (11 ao invés de 14), e diferente da segunda etapa anterior, não houve convergência nas probabilidades de mutação e reparação neste ciclo. Na Tabela 15 podem ser observadas as configurações selecionadas para a etapa final de testes.

Posição	Configuração						Usuário	\overline{TA}	Dif.
	Iter.	Cross.	Mut.	Rep.	DP I.	DP F.			
1	1000	0,9	0,2	0,3	50%	7,5%	1	0,65	+66,7%
2	1000	0,9	0,3	0,3	20%	7,5%	1	0,61	+56,4%
3	1500	0,9	0,2	0,3	10%	7,5%	1	0,57	+46,2%
4	1500	0,9	0,2	0,3	20%	10%	1	0,56	+47,4%
5	1500	0,9	0,3	0,3	20%	5%	1	0,55	+48,6%
1	1500	0,9	0,2	0,3	20%	10%	2	0,75	+13,6%
2	1000	0,9	0,3	0,3	20%	5%	2	0,75	+17,2%
3	1000	0,9	0,3	0,3	20%	7,5%	2	0,74	+15,6%
4	1500	0,9	0,3	0,4	20%	10%	2	0,73	+14,1%
5	1500	0,9	0,4	0,4	10%	7,5%	2	0,73	+15,9%
1	1000	0,9	0,2	0,3	10%	7,5%	3	0,33	+27,0%
2	1000	0,9	0,3	0,4	50%	5%	3	0,30	+20,0%
3	1500	0,9	0,2	0,3	50%	5%	3	0,29	+16,0%
4	1000	0,9	0,2	0,3	50%	7,5%	3	0,28	+12,0%
5	1500	0,9	0,2	0,3	20%	10%	3	0,28	+12,0%

Tabela 16: Configurações selecionadas na segunda etapa de testes do segundo ciclo

Como pode ser observado na Tabela 16, se manteve a tendência observada na primeira etapa de melhoria significativa entre os resultados obtidos entre as melhores

configurações dos primeiros e segundos ciclo de testes. O usuário de perfil 1, o que tem restrições nutricionais equilibradas em comparação com os demais, foi o principal beneficiado com taxa de acerto 66,7% superior em relação à melhor configuração do primeiro ciclo nesta etapa. Foram removidos os valores extremos de probabilidade de crossover (100% e 80%) bem como o maior valor de probabilidade de reparação, sendo observado que o desempenho do algoritmo do ponto de vista das métricas estabelecidas era prejudicado por uma taxa de reparação alta. As configurações marcadas em diferentes tons de cinza na Tabela 16 apareceram entre as selecionadas em mais de um perfil de usuário, onde cada tom de cor representa um par. Para a terceira etapa de testes, as 11 diferentes combinações de parâmetros numéricos da Tabela 16 serão utilizadas.

4.2.3 Resultados da terceira etapa de testes

Nesta etapa, todos os indicadores foram avaliados sob os resultados do algoritmo executado 50 vezes em cada uma das 11 configurações selecionadas na etapa anterior. De acordo com o sistema de pontos (Seção 3.7.3), foram escolhidas 3 diferentes configurações vencedoras para cada perfil de usuário. Nenhuma das configurações vencedoras, entretanto, foi unânime em todos os indicadores. Quando uma instância do algoritmo tem a menor taxa de erro, ela não possui o melhor hipervolume; quando tem um bom espalhamento entre os pontos da curva de pareto, o espaçamento entre os pontos não é o melhor. No final, as três configurações selecionadas para cada perfil de usuário foram as que obtiveram resultados mais equilibrados do ponto de vista geral dos indicadores utilizados. As configurações vencedoras de cada usuário estão relacionadas na Tabela 17.

Usuário	Iterações	Crossover	Mutação	Reparação	DP Inicial	DP Final	Pontos
1	1500	0.9	0.2	0.3	20%	10%	5
2	1000	0.9	0.3	0.3	20%	5%	5
3	1500	0.9	0.3	0.3	20%	5%	5

Tabela 17: Configurações vencedoras da etapa final de testes

Todas as configurações da Tabela 17 obtiveram o maior número de pontos dentre as demais para cada perfil de usuário, ressaltando que de um total de 15 pontos possíveis, todas as configurações vencedoras obtiveram 5 pontos (~33,3%). Na prática isso significa que as demais configurações que somam os 10 pontos restantes podem vir a ser opções viáveis a depender do perfil de usuário, o qual se mostrou importante na definição da

instância do algoritmo que vai ser utilizada. Uma configuração que obteve bons resultados para o perfil moderado e mais restrito pode vir a ser a pior escolha para o usuário com menor restrição nutricional. A dependência que a performance do algoritmo tem em relação ao perfil do usuário mostra que a estratégia para definir os parâmetros que serão utilizados em uma determinada execução do algoritmo precisa ser adaptativa com as metas nutricionais estabelecidas e a distribuição da base de dados disponível.

As características de cada instância do algoritmo que obteve o maior número de pontos bem como sua performance para os demais perfis estão relacionadas abaixo:

- Configuração vencedora para o Usuário 1:
 - Apresentou os melhores resultados em todos os cálculos estatísticos para o indicador de hipervolume, isto é, média, desvio padrão e distância entre média e mediana. 3 pontos.
 - Maior média de espalhamento dentre as demais configurações, com a menor distância entre média e mediana neste indicador. 2 pontos.
 - Pontos obtidos nas análises dos outros usuários: 4 e 2 para os usuários 2 e 3 respectivamente.
- Configuração vencedora para o Usuário 2:
 - Apresentou os melhores resultados em todos os cálculos estatísticos para o indicador de espaçamento, isto é, média, desvio padrão e distância entre média e mediana. 3 pontos.
 - Maior média de tamanho da curva de pareto dentre as demais configurações, com a menor distância entre média e mediana neste indicador. 2 pontos.
 - Pontos obtidos nas análises dos outros usuários: 2 e 0 para os usuários 1 e 3 respectivamente.
- Configuração vencedora para o Usuário 3:
 - Menor diferença entre a média e a mediana no indicador de espalhamento. 1 ponto.
 - Apresentou os melhores resultados em todos os cálculos estatísticos para o indicador de espaçamento, isto é, média, desvio padrão e distância entre média e mediana. 3 pontos.
 - Maior média de tamanho da curva de pareto dentre as demais configurações. 1 ponto.

- Pontos obtidos nas análises dos outros usuários: 0 e 3 para os usuários 1 e 2 respectivamente.

Os dados acima mostram que as configurações não são, de maneira geral, intercambiáveis entre os perfis de usuário. Enquanto o usuário 3 é bem atendido por uma instância do algoritmo, esta mesma não pontuou bem para o usuário 1. Observa-se, que a instância usada para o usuário 1 também apresenta boa performance para o usuário 2 e, embora não seja a melhor, não está entre as piores para o usuário 3. Caso não seja usada uma abordagem de ajuste paramétrico adaptativo de acordo com o perfil de usuário, a configuração do usuário 1 pode então ser usada como uma estratégia mais genérica de atendimento aos três perfis.

Ao final do segundo ciclo de testes, todas as configurações vencedoras apresentavam mais uma vez o menor valor de probabilidade de reparação dentre as possíveis no ciclo. Esta constatação não indicou a necessidade de um terceiro ciclo de testes, pois outro valor de probabilidade de reparação esteve presente durante a etapa final do ciclo (o que não aconteceu no anterior) e apresentou resultados próximos das configurações vencedoras.

4.2.4 Taxa de repetição de subgrupos

Com a diminuição da probabilidade de reparação surgiu um cuidado com a repetição de subgrupos alimentares em uma mesma recomendação, que na maioria das vezes não é desejável. Para demonstrar o impacto que a diminuição da probabilidade de reparação teve sobre a geração das recomendações, o algoritmo foi executado nas configurações vencedoras para o usuário 1 (perfil com metas nutricionais equilibradas) nos dois ciclos de teste. A Tabela 18 exibe uma amostra das recomendações geradas para este usuário nas duas configurações, as três primeiras opções exibidas na recomendação foram usadas. Na tabela, é possível ter uma ideia de como a falta de reparação está relacionada com recomendações que podem não fazer sentido ao usuário. As opções 1 e 3 com taxa de reparação menor, por exemplo, oferecem três tipos diferentes de leite. As opções do ciclo 1, fazem mais sentido do ponto de vista do usuário, ainda que apresentando repetição de subgrupos como é o caso das frutas. Vale ressaltar que, do ponto de vista dos indicadores de qualidade definidos, as opções do ciclo 2 apresentam resultados com taxa de acerto maior, isto é, mais próximas das metas nutricionais do usuário. Se o foco for apenas otimizar o ponto de vista nutricional, as recomendações do segundo ciclo podem ser consideradas as melhores. Contudo, se não houver coerência nas

opções oferecidas dificilmente o usuário vai adotá-las. Assim, para avaliar a coerência da refeição baseado no número de opções ofertadas que apresentam repetição de subgrupos alimentares foi criado um indicador de qualidade adicional. Este indicador tem como objetivo verificar a relação entre a melhora dos demais indicadores e a taxa de repetição dos subgrupos. Para fazer esta análise, as 729 combinações de parâmetros numéricos do primeiro ciclo de testes foram executadas novamente, e a relação entre os resultados foi feita utilizando o aplicativo WEKA (EIBE FRANK, 2016).

Opção	Ciclo 1 (Reparação: 50%)	Ciclo 2 (Reparação: 30%)
1	Iogurte, natural, desnatado Mexerica, Murcote, crua Leite, de cabra Laranja, baía, crua Cuscuz, paulista	Leite, de vaca, integral Laranja, valência, suco Leite, fermentado (2 porções) Leite, de cabra Pamonha, barra para cozimento, pré-cozida
2	Iogurte, natural, desnatado Mexerica, Murcote, crua Leite, de cabra Laranja, baía, crua Cuscuz, paulista Pêssego, Aurora, cru	Leite, condensado, Laranja, Pêra, crua Bebida láctea, pêssego Café, infusão 10% (2 porções)
3	Laranja, lima, crua Iogurte, natural, desnatado Leite, de cabra Uva, Itália, crua Cuscuz, paulista Pêssego, Aurora, cru	Iogurte, natural, desnatado Leite, de vaca, integral Leite, de cabra Leite, fermentado Pamonha, barra para cozimento, pré-cozida Limão, tahiti, cru

Tabela 18: Comparativo de recomendações geradas com probabilidade de reparação diferentes

O indicador de taxa de repetição de subgrupos, chamado de “*MealSpreading*” no algoritmo, funciona de maneira semelhante ao de taxa de acerto. A diferença é que neste caso objetiva-se minimizar este valor. A Equação (7) foi usada para medir a taxa de repetição de subgrupos é apresentada abaixo:

$$MS = \frac{\sum_{i=1}^n a_i}{n} \quad (7)$$

Na Equação (7), “n” é o número de pontos na curva de pareto sob análise e a_i será 1 se houver um ou mais alimentos pertencentes ao mesmo subgrupo alimentar no ponto “i” e 0 caso contrário. Observe que no melhor dos cenários, nenhuma refeição contendo repetição de subgrupos foi sugerida, então o valor do indicador é zero.

Com o indicador de taxa de repetição de subgrupos definido, verificou-se qual o valor deste novo indicador nas três configurações vencedoras para cada perfil de usuário nos dois ciclos de teste. Para tal, cada configuração foi executada 50 vezes e então foram calculados a média, o desvio padrão amostral e a mediana deste indicador. Os resultados obtidos estão dispostos na Tabela 19.

Usuário	Ciclo	Reparação	Taxa de Repetição de Subgrupos
1	1	0,5	67,1%
1	2	0,3	62,4%
2	1	0,5	90,4%
2	2	0,3	88,4%
3	1	0,5	49,6%
3	2	0,3	31,4%

Tabela 19: Taxa de repetição de subgrupos nas soluções vencedoras para cada perfil de usuário

Como pode ser observado na Tabela 19, mesmo com taxa de reparação menor, a taxa de repetição de subgrupos não foi aumentada. Podem ser destacados dois pontos a partir destes dados:

- I. Embora com bons resultados do ponto de vista nutricional, as soluções apresentam taxa de repetição de subgrupos elevada, principalmente para os usuários 1 e 2.
- II. A probabilidade de reparação em todas as opções não é suficiente para reduzir o valor deste indicador.

Para entender melhor os motivos que ocasionam as taxas de repetição de subgrupos elevadas da Tabela avaliaram-se como os valores dos parâmetros numéricos influenciam os resultados dos indicadores, além da análise da relação entre os indicadores utilizados nas rodadas de teste anteriores e o novo indicador. A próxima seção aborda as observações realizadas a partir desta investigação.

4.3 Relação entre a coerência da refeição e os demais indicadores de qualidade

O conceito de coerência da refeição é abrangente e pode ser compreendido como uma combinação de fatores objetivos, subjetivos e culturais. Neste trabalho, este conceito foi generalizado de modo que a coerência da refeição pode ser definida por meio do indicador de taxa de repetição de subgrupos. Nessas condições, para uma refeição ser classificada como coerente ela deve ter uma combinação de alimentos que façam parte dos subgrupos definidos pelo nutricionista e sem repetição de alimentos de um mesmo subgrupo. Uma vez definido o indicador de taxa de repetição de subgrupos, será observado se há algum padrão que justifique a repetição de subgrupos nas soluções apresentadas na Tabela 19.

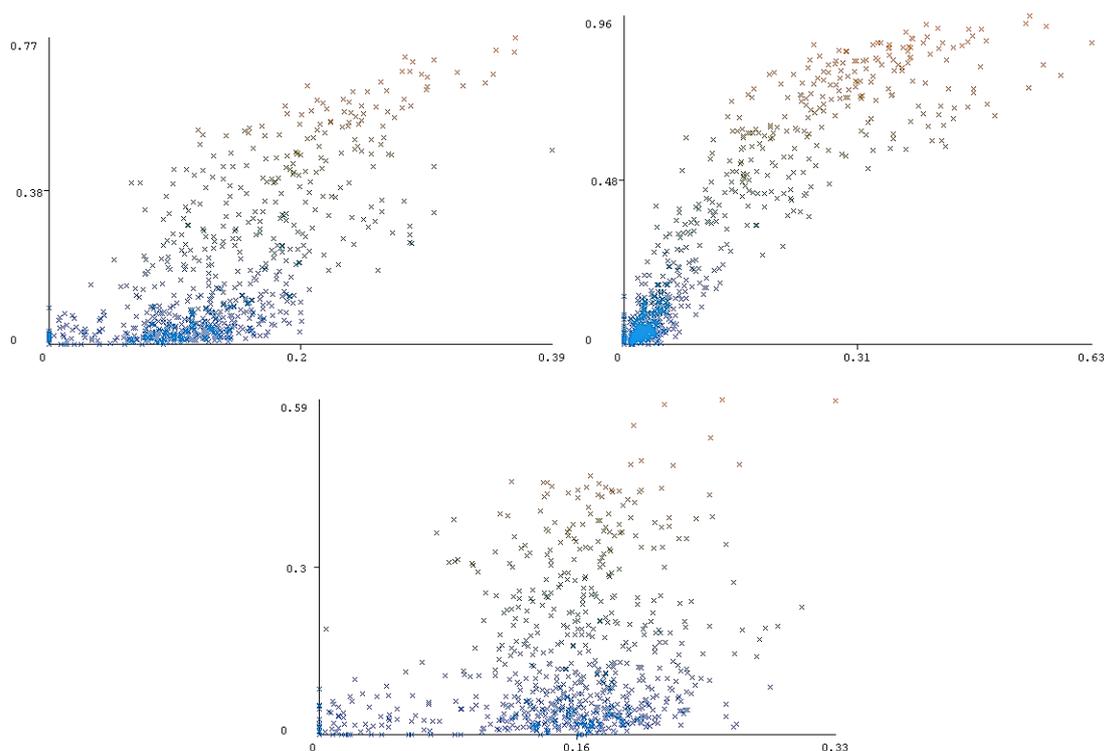


Figura 10: Dispersão dos valores de taxa de acerto (eixo x) e repetição de subgrupos (eixo y) em 729 combinações de parâmetros numéricos para os usuários 1, 2 e 3

A Figura 10 demonstra a existência de uma relação diretamente proporcional entre as taxas de acerto e repetição de subgrupos. As 729 combinações de parâmetros numéricos participantes do primeiro ciclo de teste foram executadas novamente e nas mesmas condições da primeira etapa (10 execuções em cada perfil de usuário). Os resultados foram importados no aplicativo de mineração de dados WEKA, que gerou os

gráficos dispostos na imagem. É possível observar que de fato quando a taxa de acerto é alta a taxa de repetição de subgrupos também é alta. A forma como um indicador varia em relação com o outro contudo não é linear, observe nos gráficos para os usuários 1 e 3 que é possível reduzir em mais da metade a taxa de repetição de subgrupos reduzindo a taxa de acerto em menos que a metade. Especificamente para o usuário 2, que possui metas nutricionais mais restritas, a escolha ocorre entre muitas soluções aceitáveis do ponto de vista nutricional e com problemas de coerência na composição da refeição ou entre poucas soluções que tenham mais chances de serem seguidas pelos usuários. A escolha entre uma abordagem e outra é especialmente problemática visto que o conceito de coerência da refeição em termos práticos é subjetivo e no modelo de algoritmo proposto a preferência do usuário não é considerada como entrada. Este trabalho limitou-se a identificar a relação de perda-ganho entre os dois indicadores, e as influências que os parâmetros numéricos possuem sobre a taxa de repetição de subgrupos. Cabe aos trabalhos futuros a aplicação do algoritmo com a configuração que melhor atenderá aos objetivos determinados.

Apesar da limitação vista no parágrafo anterior sobre a dependência existente entre o aumento da taxa de acerto e a taxa de repetição de subgrupos, vale ressaltar que o usuário cujo perfil apresentou este problema de maneira mais evidente possui metas nutricionais mais restritas, em outras palavras, pertence a uma classe específica de usuários. Isso permite que o algoritmo seja dimensionado de modo a seguir uma abordagem específica para este cenário, isolando a versão atual para os demais casos. Seguindo por esta linha de raciocínio, os gráficos da Figura 10 relacionados com os usuários 1 e 3 mostram que, uma vez identificada a causa do aumento da taxa de repetição de subgrupos, é possível melhorar significativamente a coerência da refeição com baixa degradação da taxa de acerto. Mais uma vez, foi usado o WEKA para auxiliar na determinação das configurações que mais influenciam na taxa de repetição de subgrupos. Na Figura 11 é mostrado o gráfico de dispersão da taxa de repetição de subgrupos, em relação aos dois parâmetros numéricos, que foram mais significativos: as probabilidades de cruzamento e reparação. Para os demais parâmetros do algoritmo, a dispersão da taxa de repetição de subgrupos se apresentou mais uniforme, não sendo, portanto, relevante para esta análise. Para reduzir o número de comparações, o usuário 2 foi excluído das análises a partir deste ponto, devido aos motivos supracitados.

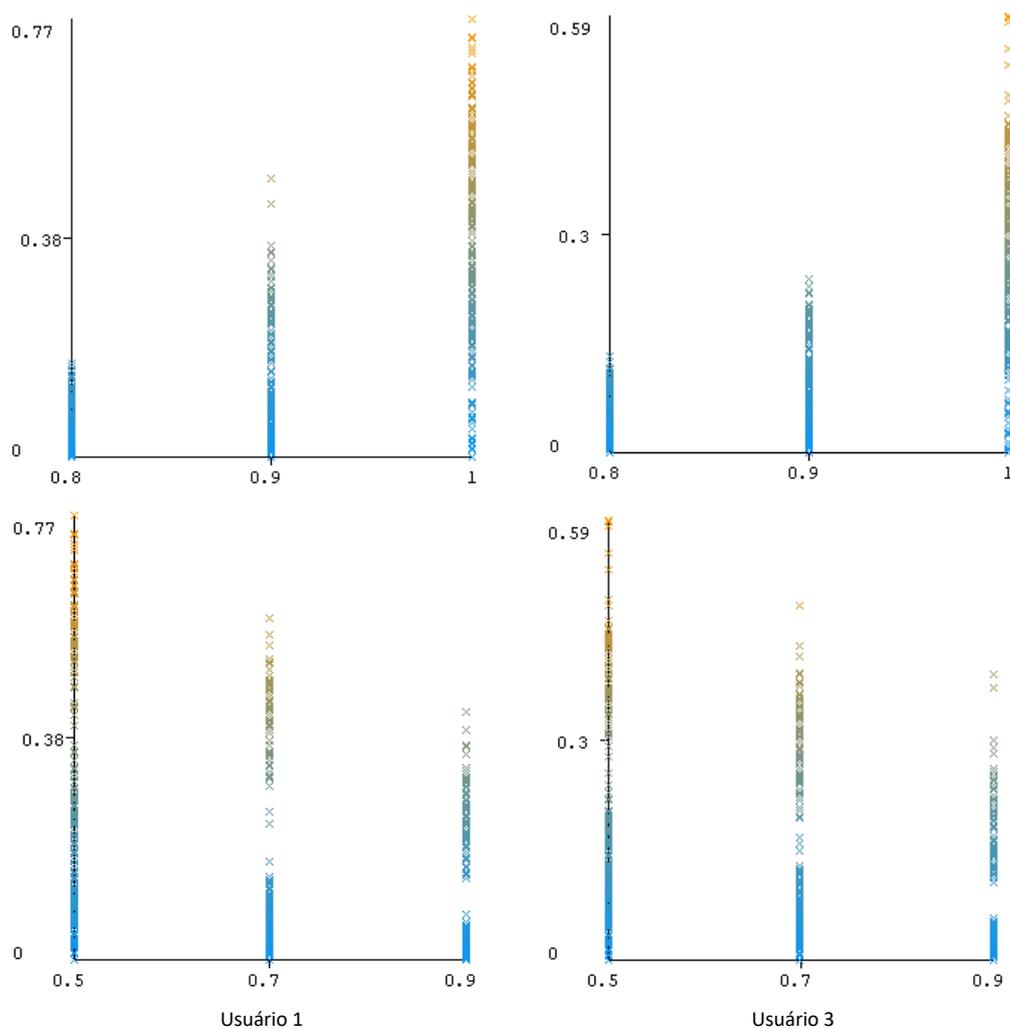


Figura 11: Dispersão da taxa de repetição de subgrupos em relação as probabilidades de cruzamento (em cima) e reparação (em baixo)

Na Figura 11 a primeira linha de gráficos apresenta a dispersão da taxa de repetição de subgrupos em relação à probabilidade de crossover. Observe que quando a probabilidade de crossover é alta, a taxa de repetição de subgrupos pode atingir o seu valor máximo. Em contrapartida, independente dos demais parâmetros, quando a probabilidade de crossover é menor, a taxa de repetição de subgrupos cai mais de 50%. Já com relação a probabilidade de reparação, embora tenha sido criada para este fim, observou-se que o seu impacto foi menor, embora relevante. Acredita-se que a explicação para o impacto do cruzamento ser superior ao da reparação está na forma como o segundo foi desenvolvido, dividindo espaço com o operador de mutação. O fato de não ocorrer reparação quando há mutação reduz a eficiência deste operador em reparar o que aparentemente é feito pelo operador de cruzamento.

Portanto, para gerar recomendações com aumento no número de opções que façam sentido ao usuário é necessário reconsiderar os valores dos parâmetros utilizados, em especial as probabilidades de crossover e reparação, e flexibilizar a taxa de acerto. A busca pela configuração que atenda a essas novas diretrizes precisa seguir um caminho diferente do adotado no primeiro e segundo ciclos de teste, uma vez que não podem apenas ordenar os resultados baseado em um único indicador devido à relação de perda-ganho entre eles. Além disso, como existem estudos prévios, a escolha dos valores dos parâmetros numéricos que serão usados nos testes pode se ancorar nos resultados anteriores. As decisões tomadas e seus resultados estão detalhados na Seção 4.4.

4.4 Teste com foco na coerência da refeição

O objetivo de repetir os testes com ênfase em melhorar a coerência da refeição é validar a performance do algoritmo em relação a este critério, uma vez que os ciclos de teste anteriores já mostraram a performance do algoritmo na obtenção de soluções com atendimento satisfatório às metas nutricionais estabelecidas pelo nutricionista. O processo de escolha das configurações vencedoras foi alterado devido à relação entre os operadores de taxa de acerto e repetição de subgrupos. Ordenar as soluções por taxa de acerto eliminaria as soluções com repetição de subgrupos baixa e vice-versa. Assim, o processo de escolha das soluções vencedoras ocorreu da seguinte forma:

- Para a etapa 1:
 - É definido um valor máximo de taxa de repetição de subgrupos igual a 1/3 do maior valor pertencente ao agrupamento dos menores valores definido no aplicativo WEKA para os resultados da etapa 1.
 - O subconjunto de configurações definido é ordenado de forma decrescente pela taxa de acerto. A metade superior das configurações é selecionada para a etapa 2.
- Para a etapa 2:
 - As melhores configurações em taxa de repetição de subgrupos são escolhidas usando o mesmo critério da etapa 1.
 - O subconjunto de configurações definido é ordenado pela taxa de acerto. As 5 primeiras configurações são selecionadas para a etapa final.

A etapa 3 utiliza o mesmo sistema de pontos dos demais ciclos de teste. Adicionalmente aos 15 pontos possíveis nos ciclos anteriores, os valores estatísticos

relacionados com a taxa de repetição de subgrupos são incluídos no sistema de pontos, aumentando o número máximo para 18.

4.4.1 Constatações da primeira etapa de testes

A primeira análise de dados foi feita sobre os resultados das 729 combinações de parâmetros numéricos, usados na construção dos gráficos da Figura 11, com o auxílio dos agrupamentos realizados na etapa de pré-processamento de dados realizada pelo aplicativo WEKA. Constatou-se que 347 combinações, aproximadamente 47,6% de todas as participantes da primeira etapa, apresentou taxa de repetição de subgrupos entre 0 e 0,077 para o usuário 1. A primeira filtragem de resultados eliminava todas as configurações que obtiveram repetição de subgrupos superior a 1/3 deste valor ($\sim 0,026$), para o usuário 1 isso significou uma redução das 729 combinações para 115. As combinações filtradas foram ordenadas pela taxa de acerto no sentido decrescente e a metade superior foi selecionada para a etapa 2. De forma análoga foi feita a análise baseada nos resultados para o usuário 3. No final, 76 combinações distintas de parâmetros numéricos foram selecionadas para a segunda etapa de testes, estas configurações obtiveram na primeira etapa taxas de acerto entre 9,7% e 26,4% e taxas de repetição de subgrupo entre 0% e 2,5%. Os valores dos parâmetros foram distribuídos como mostrado na Figura 12.

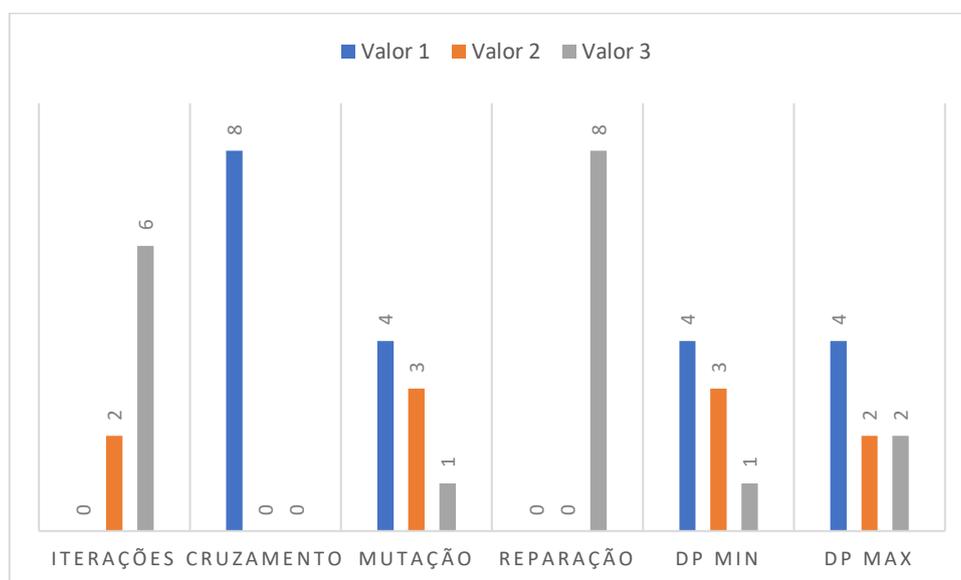


Figura 12: Distribuição dos parâmetros numéricos selecionados com base nas taxas de repetição de subgrupos e acerto durante a primeira etapa do teste

A Figura 12 mostra que os parâmetros mais significativos na criação de soluções com taxa de repetição de subgrupos baixa são o cruzamento e a reparação, com os

respectivos extremos superior e inferior praticamente eliminados já na primeira etapa de testes. As próximas duas etapas visam refinar os resultados por meio do aumento do número de execuções e por fim chegar à configuração que leva o algoritmo aos resultados que façam mais sentido ao usuário.

4.4.2 Constatações da segunda etapa de testes

Os resultados desta etapa consolidaram os resultados da etapa anterior e evidenciaram como os operadores precisam estar dimensionados para levar o algoritmo a resultados mais coerentes. Como pode ser visto na Figura 13, há um padrão para que os melhores resultados sejam obtidos com maior número de iterações e principalmente menor cruzamento e maior reparação. Portanto, para obter resultados mais lógicos, é preciso seguir o caminho oposto ao que aumenta o número de soluções aceitáveis. As taxas de acerto e repetição de subgrupos podem ser modeladas em trabalhos futuros como objetivos conflitantes, a fim de buscar uma alternativa mais equilibrada. Quanto aos resultados desta etapa, das 76 combinações apenas 8 foram selecionadas para a terceira etapa de testes utilizando os critérios abordados anteriormente, com taxas de acerto entre 9,8% e 18,0% e taxas de repetição de subgrupo entre 0,7% e 1,1%.

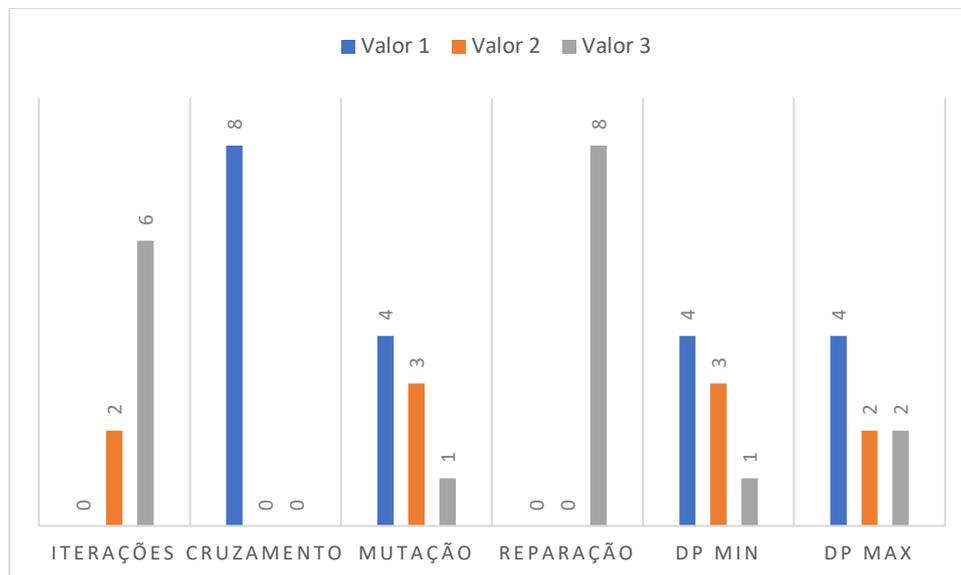


Figura 13: Distribuição dos parâmetros numéricos selecionados com base nas taxas de repetição de subgrupos e acerto durante a segunda etapa de teste

A terceira etapa dos testes levou em consideração os dados estatísticos de repetição de subgrupos em conjunto com os demais indicadores definidos e utilizou o mesmo sistema de pontos dos ciclos de teste anteriores. Os resultados estão detalhados na Seção 4.4.3.

4.4.3 Constatações da terceira etapa de testes

Assim como na terceira etapa do segundo ciclo de testes, os pontos não convergiram para uma configuração específica de algum dos dois perfis de usuário utilizados. Contudo a amplitude dos valores para os indicadores de taxa de acerto e repetição de subgrupos foi pequena. Na prática a performance entre as configurações vencedoras para o usuário 1 e 3, apresentadas na Tabela 20, é equivalente, podendo ser escolhida apenas uma delas para os dois perfis.

Usuário	Iterações	Crossover	Mutação	Reparação	DP Inicial	DP Final	Pontos
1	1500	0.8	0.1	0.9	50%	10%	6
3	1500	0.8	0.1	0.9	50%	5%	4

Tabela 20: Configurações finais para os usuários 1 e 3 considerando a coerência da refeição

Se a configuração usada para executar o algoritmo não for adaptativa ao perfil de usuário (isto é, se apenas uma das duas puder ser escolhida para instanciar o algoritmo) a opção para o usuário 1 possui características que a torna mais interessante do que a alternativa para o usuário 3. As justificativas para isso estão relacionadas abaixo:

- A configuração obteve 4 pontos para o usuário 3, sendo a segunda melhor configuração em número de pontos. Em contrapartida, a configuração do usuário 3 só recebeu 1 ponto quando avaliada com o usuário 1, sendo a sexta melhor nele. A opção do usuário 1 então consegue atender a ambos de forma mais genérica.
- A diferença da taxa de acerto para a melhor configuração do usuário 3 é de -2,1%, com uma curva de pareto em média de tamanho 35 para ambos. Na prática as duas configurações geram um número equivalente de soluções aceitáveis para os dois perfis de usuário.
- A taxa de repetição de subgrupos é inferior à melhor configuração do usuário 3.
- Como consequência do menor valor de desvio padrão final, a variabilidade das soluções apresentadas é reduzida, o que é inconveniente pois há uma diferença considerável no indicador de hipervolume entre a melhor solução do usuário 3 e a solução do usuário 1. A média do hipervolume calculada para a opção com desvio padrão final menor foi aproximadamente 60,3% inferior.

Dadas as justificativas acima, a Tabela 21, lista três opções recomendadas pelo algoritmo instanciado com a configuração de parâmetros vencedora para os usuários de teste.

Opção	Usuário 1	Usuário 3
1	Iogurte, natural, desnatado Queijo, petit suisse, morango Café, infusão 10% Laranja, valência, crua Pamonha, barra para cozimento, pré-cozida	Leite, de vaca, achocolatado Chá, mate, infusão 5% Capuccino, pó
2	Iogurte, natural, desnatado Queijo, petit suisse, morango Laranja, baía, suco Pamonha, barra para cozimento, pré-cozida Água de coco	Leite, de vaca, achocolatado Chá, mate, infusão 5% Aveia, flocos, crua
3	Graviola, polpa, congelada Queijo, petit suisse, morango Bebida láctea, pêssego Café, infusão 10% Pamonha, barra para cozimento, pré-cozida	Laranja, pera, crua Bebida láctea, pêssego Aveia, flocos, crua

Tabela 21: Amostra de recomendações geradas pela configuração vencedora do teste para os usuários 1 e 3

As recomendações listadas na Tabela 21 foram extraídas das primeiras três opções de uma recomendação alimentar gerada para cada usuário. Em comparação com as recomendações da Tabela 18, a redução de alimentos pertencentes a um mesmo subgrupo alimentar é relevante. Vale ressaltar que apesar da taxa de acerto ser inferior para esta instância do algoritmo em comparação com as vencedoras dos ciclos de teste anteriores, ela precisa ser interpretada como limitadora do número de opções oferecidas ao invés da qualidade das opções em termos nutricionais. Todas as recomendações das Tabelas 18 e 21 estão dentro da faixa aceitável de desvio dos limites nutricionais estabelecidos para os usuários. A diferença entre as instâncias que geraram as recomendações das duas tabelas é que as da primeira fornecem mais opções enquanto as da última fornecem opções que fazem mais sentido ao usuário.

4.5 Algoritmo proposto x Abordagem mono-objetivo

Os resultados obtidos pelas três instâncias de algoritmo usando a abordagem mono-objetivo revelaram que é possível ter recomendação de refeições semelhantes às obtidas pelo Algoritmo proposto em termos de número de opções e taxa de acerto. A exceção ocorreu com o algoritmo de Isowaka (2015), que apresentou resultados inferiores em comparação com os outros dois algoritmos mono-objetivo e o Algoritmo proposto. Os resultados do algoritmo de Isowaka revelam que a proposta apresentada por ele não é ortogonal em relação à base de dados e/ou à escalabilidade do problema de otimização, isto é, a alteração dos alimentos disponíveis e o número de critérios de avaliação nos indicadores IFG e IFC prejudicou a performance do algoritmo de Isowaka.

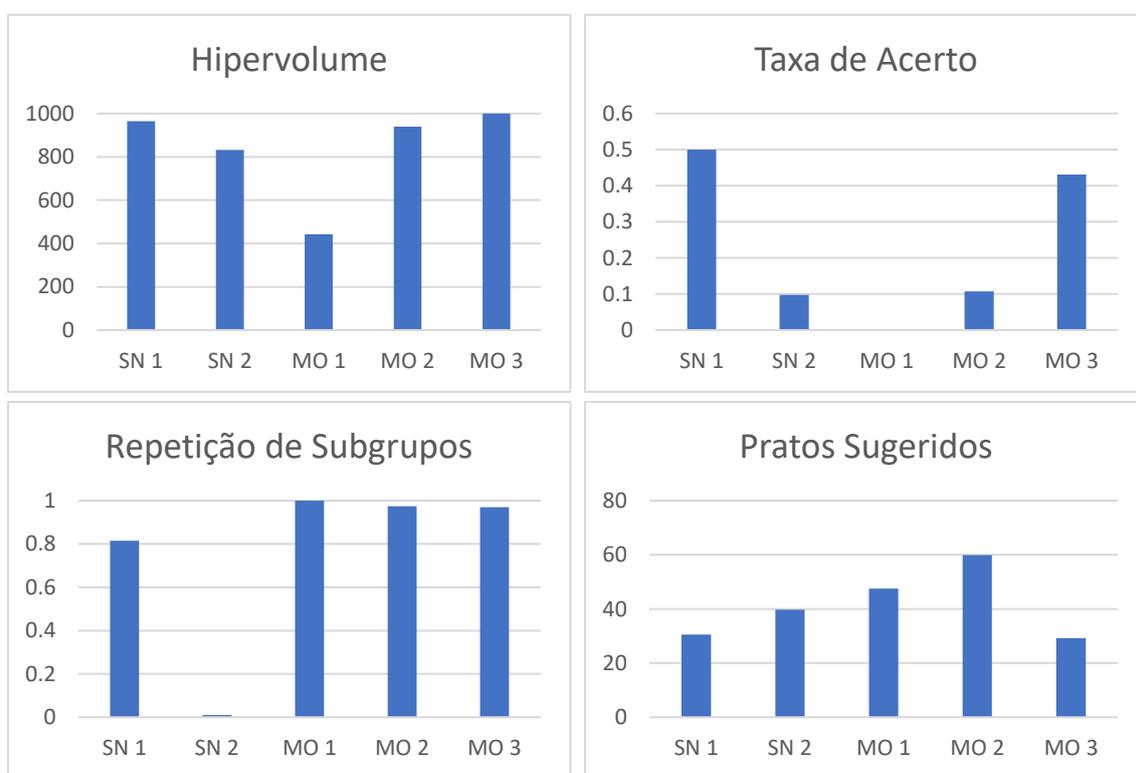


Figura 14: Comparativo das métricas obtidas por meio das abordagens mono e multiobjetivo.

Na Figura 14, SN 1 e SN 2 representam as instâncias do Algoritmo proposto com melhores resultados em relação a taxa de acerto e de repetição de subgrupos, MO 1 até MO 3 são respectivamente os algoritmos de Isowaka, com otimização de nutriente simples e combinação de nutrientes. Os valores mostrados representam a média do indicador em questão considerando todas as execuções realizadas durante os ciclos de teste detalhados anteriormente.

Comparando as abordagens mono e multiobjetivo, o destaque é a diferença no indicador de taxa de repetição de subgrupos. Neste indicador, os algoritmos mono-

objetivo obtiveram resultados piores até mesmo em comparação com a instância do algoritmo vencedora no ciclo de testes que priorizava a taxa de acerto sobre a coerência da refeição. Vale ressaltar, porém, que nenhum dos algoritmos mono-objetivo utilizados possuía um operador de reparação que auxiliasse na construção de soluções melhores do ponto de vista de coerência da refeição. Em resumo, o Algoritmo proposto mostrou ser uma opção para otimização da geração de recomendações alimentares com bom desempenho nas métricas estabelecidas neste trabalho. Além disso, é notável o impacto que a técnica de reparação das soluções candidatas acarreta na melhoria da qualidade da refeição oferecida ao usuário, o que confirma a hipótese inicial que incluir conhecimentos do problema em outras etapas da computação evolucionária (no caso deste trabalho o conceito de coerência da refeição no operador de mutação) pode levar a resultados melhores em comparação ao uso de um algoritmo clássico.

5 Conclusões e trabalhos futuros

A recomendação alimentar apoiada pela computação evolucionária é uma linha de pesquisa que demonstrou potencial para uso em aplicações destinadas ao acompanhamento e educação nutricional, perda de peso e tratamento de patologias. Neste Capítulo é realizada uma análise crítica dos resultados obtidos neste trabalho e apontadas algumas oportunidades de melhoria para aperfeiçoamento destes resultados. Também serão apontadas algumas direções que se acredita serem os próximos passos para que esse algoritmo seja aplicado em um contexto prático e real. As próximas seções dissertam sobre cada um destes itens.

5.1 Análise geral dos resultados

Esta pesquisa iniciou com uma análise do mercado de aplicações voltadas para saúde na atualidade e qual a participação da tecnologia no processo de confecção de recomendações alimentares. A análise dos trabalhos relacionados obtidos neste processo revelou que existe uma tendência a associar o acompanhamento nutricional com a necessidade de perda de peso ou atingir determinada forma física; este pensamento se mostra nas funcionalidades dos aplicativos voltados para esta área que foram apontadas no Capítulo 2. Além da visão limitadora sobre a utilidade de um bom acompanhamento nutricional, também foi constatado que existe um desalinhamento de expectativas entre os perfis de usuário relacionados neste processo. Os usuários não se estimulam a usar aplicações voltadas para a saúde possivelmente por não compreenderem a sua importância, e pelo fluxo de uso comum deste tipo de aplicação em forma de diário alimentar, que atribui ao usuário a obrigação de relatar tudo o que ele decidiu comer ao longo do dia. Do ponto de vista dos profissionais de saúde, as pesquisas mostraram que o uso das aplicações se limita basicamente a servir como um meio de comunicação entre o profissional e o paciente, e que mesmo para este uso o número de aplicações que oferecem esta funcionalidade é pouco expressivo. Este trabalho propôs atuar no estreitamento entre os objetivos dos perfis profissional e do paciente, aplicando a tecnologia de geração automática de recomendações alimentares com uso de otimização multiobjetivo. Com o algoritmo proposto torna-se possível uma alteração no fluxo de uso de aplicativos nutricionais, pois ao fornecer recomendações prontas e adequadas nutricionalmente ao paciente, reduz-se a carga relacionada à inserção manual dos

alimentos e torna a participação da tecnologia mais ativa neste processo. Ao mesmo tempo que atende aos usuários, esta tecnologia pode dar uma nova perspectiva aos profissionais de saúde com um aplicativo que pode ajudar na otimização do tempo e trabalho do dia a dia, como uma nova ferramenta de automação para montar fichas de dieta com maior variabilidade de opções.

As recomendações montadas pelo algoritmo se mostraram capazes de fornecer uma quantidade considerável de alimentos diferentes e adequadas nutricionalmente às especificações definidas, possibilitando a montagem do plano diário de alimentação referenciados como o objetivo geral deste trabalho. Os testes mostraram que dado um grupo alimentar personalizado e um conjunto de metas nutricionais, o algoritmo é capaz de gerar recomendações que se adequem às condições fornecidas. Mesmo com a variação das metas, por meio do uso de perfis diferentes de usuário, o algoritmo se manteve capaz de gerar recomendações.

A modelagem do algoritmo foi feita de modo a possibilitar que o grupo e os subgrupos alimentares utilizados durante a evolução sejam apenas variáveis do sistema. Isto possibilita que o nutricionista possa definir grupos de alimentos adequados a cada perfil de usuário livremente, cabendo ao algoritmo o trabalho árduo de combinar os itens de forma adequada. Em adição, acrescentar conhecimentos do universo do problema nas diversas etapas da evolução se mostrou positivo para os resultados, tal como o uso de soluções de tamanho variável, que aumenta o número de opções possíveis de refeição e torna possível a participação de alimentos que, por possuírem grandes quantidades de algum nutriente específico, poderiam nunca ser oferecidos em uma abordagem convencional de tamanho fixo. Vale destacar também a participação ativa do operador de reparação na melhora da coerência das refeições sugeridas conforme observado nas etapas de testes descritas no Capítulo 4.

O principal diferencial do algoritmo proposto em relação aos demais algoritmos testados ocorreu em dois quesitos. Por um lado, a coerência da refeição, que é em parte atribuída ao operador de reparação; e por outro lado, a dispersão dos resultados no espaço objetivo. Como pode-se observar na Figura 9, a distribuição dos resultados provenientes da abordagem mono-objetivo é em geral irregular, com os pontos condensados em pequenas regiões próximas dos extremos de cada objetivo. Com uso da abordagem multiobjetivo no algoritmo, as soluções oferecidas eram mais bem espalhadas ao redor do ponto ótimo; isso se traduz, em termos práticos, em soluções com maior diversidade de itens, o que aumenta a chance de alguma recomendação agradar ao usuário.

É importante considerar que o algoritmo deve ser enxergado como parte de um sistema maior de recomendação alimentar que envolve o desenvolvimento de aplicações clientes (web e móvel), para tornar possível o uso deste algoritmo. Por fim, a comparação com os algoritmos mono-objetivo foi importante para verificar se o algoritmo proposto neste trabalho apresenta alguma melhoria em relação as estratégias adotadas por outros pesquisadores e assim validar o problema.

5.2 Oportunidades de melhoria

Ao longo da pesquisa, limitações da abordagem adotada foram constatadas, algumas foram mitigadas como a exemplo da necessidade de um novo ciclo de testes envolvendo um novo indicador de taxa de repetição de subgrupos, e outras foram apenas assumidas como ações que fazem parte do escopo deste trabalho. Contudo há outros problemas encontrados que estão no escopo dos trabalhos futuros.

Os testes do algoritmo proposto foram baseados em repetições exaustivas da execução de diversas instâncias do algoritmo com variações nos parâmetros numéricos. Estes testes poderiam ser ainda mais exaustivos com a variação dos parâmetros simbólicos, isto é, dos operadores utilizados. A decisão de usar soluções com tamanho variável tornou o processo de escolha dos operadores complexo, pois o *framework* utilizado no desenvolvimento do algoritmo não fornece suporte a este tipo de solução, precisando cada operador ser desenvolvido especificamente para este trabalho. O desenvolvimento e teste de novos operadores com suporte a multiobjetivo e soluções de tamanho variável não é trivial e rápido. Neste trabalho a variabilidade de cenários para os testes do algoritmo considerou somente os perfis nutricionais e operadores numéricos.

Os algoritmos multiobjetivo utilizados em trabalhos anteriores obtiveram resultados satisfatórios, mas não foi encontrada referência a trabalhos nesta área que utilizavam o NSGA-III. O interesse por usar este algoritmo foi ainda maior por sua performance consolidada pela academia no espaço com muito objetivos (maior que 3), o que aumenta as opções possíveis de meta nutricional dada a quantidade de opções disponíveis de micronutrientes em adição aos três macronutrientes utilizados neste trabalho. Neste ponto, tem-se duas oportunidades de melhoria, a primeira é testar o uso de outros algoritmos *many-objective*, como o TAA e o TC-SEA; a segunda é o teste com mais de três objetivos. Apesar da possibilidade de o NSGA-III lidar com mais de três

objetivos, este valor foi adotado para facilitar a visualização dos resultados por meio de gráficos neste trabalho, o que seria inviável a partir do universo 4D.

5.3 Trabalhos futuros

Este trabalho apresentou um algoritmo multiobjetivo capaz de gerar recomendações para construção de um plano diário de alimentação. Ele pode ser encarado como o primeiro componente de um sistema maior que envolverá testes diretamente com o público-alvo desta pesquisa (profissionais de saúde envolvidos com nutrição e pessoas com interesse em receber um acompanhamento nutricional). Acredita-se que para atingir este estágio, o atual algoritmo deverá trabalhar em conjunto com técnicas de IA que possam incluir um componente ainda não explorado, que são as preferências do usuário. As variáveis de entrada limitaram-se às metas nutricionais definidas, podendo ser os três macronutrientes, o valor calórico final e/ou micronutrientes. Trabalhos futuros podem incluir uma lista de alimentos para o processo de geração das recomendações, isto é, o usuário poderia definir preferências e isso seria adicionado no fluxo de otimização do algoritmo.

Para além de algoritmos evolucionários, a mineração de dados também poderia auxiliar no processo de montagem das refeições, principalmente no que diz respeito à sua coerência. Enquanto o evolucionário trabalha reduzindo a taxa de repetição de subgrupos, com técnicas de mineração de dados seria possível extrair informações relacionadas às mais comuns baseado no histórico de alimentação dos usuários. Note que os dois processos são complementares; a abordagem atual elimina a possibilidade que ocorra “feijão preto” e “feijão carioca” em uma mesma recomendação, e com a mineração de dados seria possível atribuir mais pontos para a combinação “feijão” e “arroz” em uma região em que estes dois itens são comumente combinados em determinada refeição, para determinado perfil de usuário. Como sugestão de trabalho futuro, esta técnica pode ser usada para gerar recomendações iniciais quando ainda não há conhecimento suficiente a respeito do usuário. Também seria possível auxiliar o nutricionista na geração de grupos alimentares através da análise dos dados históricos relevantes.

Para permitir que os resultados do algoritmo sejam usados pelos usuários finais, ainda são necessários alguns ajustes: os usuários, suas preferências e as recomendações geradas podem ser armazenadas em banco de dados, além de usar alguma tecnologia que permita a construção de uma API que forneça a recomendação alimentar em forma de um

serviço web. Para o nutricionista, acredita-se que uma interface web seria mais adequada, pois possibilitaria o uso de um dispositivo com teclado e mouse para as atividades administrativas do fluxo de uso, como cadastro de pacientes, criação de grupos alimentares e análise de relatórios. Já para os pacientes, uma aplicação móvel seria mais adequada, pois facilitaria a consulta das recomendações alimentares para o dia/semana em qualquer lugar. Com essas interfaces de uso para os usuários e profissionais, seria possível testar o sistema de recomendação com usuários reais, obtendo assim avaliações muito importantes para o aperfeiçoamento da solução.

5.4 Considerações finais

Esta pesquisa apresentou uma nova aplicação da computação evolucionária para atender às deficiências detectadas em aplicativos nutricionais. Somando às outras pesquisas nesta mesma linha, acrescentou-se o uso de um algoritmo para muitos objetivos de performance consolidada (DEB, 2014) para ampliar o poder de escolha sobre os critérios nutricionais da recomendação. Também houve contribuição ao deixar a evolução ocorrer sobre um conjunto de dados maior e menos direcionado. Os resultados obtidos mostram que um passo a mais foi dado em direção à geração autônoma de recomendações alimentares com boa distribuição de opções e coesão entre os itens, e que a associação entre otimização multiobjetivo e conhecimentos do domínio do problema produz bons resultados.

Para finalizar, ressalta-se que o conhecimento relacionado à qualidade das refeições ingeridas é de interesse geral, como pode ser evidenciado pelas estatísticas de uso das aplicações nutricionais populares. O problema abordado neste trabalho já foi identificado tanto pelos usuários deste tipo de aplicação como pela academia. Espera-se que as experiências que foram relatadas nesta dissertação contribuam para o prosseguimento destes estudos e que novos trabalhos sejam desenvolvidos, de modo que esta tecnologia esteja cada vez mais próxima do dia a dia da sociedade, contribuindo para a conscientização acerca da qualidade nutricional das nossas refeições.

Referências

CHEN, C., MEMBER, S., KARVELA, M., SOHBATI, M., SHINAWATRA, T., TOUMAZOU, C. PERSON – Personalized Expert Recommendation System for Optimized Nutrition, **IEEE Transactions on Biomedical Circuits and Systems**, Vol. 12, NO. 1, 2018.

SHUREN J, PATEL B, GOTTLIEB S. FDA Regulation of Mobile Medical Apps. **JAMA**. 320(4):337–338. 2018

MARTINS, GILHERME P. V., SOUZA, EDSON P., VARGAS, ICLÉIA A., Parque Estadual do Prosa em Campo Grande/MS: uma análise ambiental, a mobilidade urbana atual e o ciclismo como alternativa de transporte sustentável e promotor da educação ambiental, **Braz. J. of Develop.**, Curitiba, v. 4, n. 7, Edição Especial, p. 3876-3892, nov. 2018

CFN, **RESOLUÇÃO CFN Nº 600**, Disponível em: https://www.cfn.org.br/wp-content/uploads/resolucoes/Res_600_2018.htm, Acesso em: 15 de Mar. De 2020

GABRIELLI, S., DIANTI, M., MAIMONE, R., BETTA, M., FILIPPI, L., GHEZZI, M., & FORTI, S. Design of a Mobile App for Nutrition Education (TreC-LifeStyle) and Formative Evaluation with Families of Overweight Children. **JMIR mHealth and uHealth**, 2017.

NIKOLAOU, CHAROULA K., LEAN, MICHAEL E J., Mobile applications (‘apps’) for obesity and weight management: current market characteristics. **International Journal of Obesity**, 2017

FRANCO, R. Z., FALLAIZE, R., LOVEGROVE, J. A., & HWANG, F. Popular Nutrition-Related Mobile Apps: A Feature Assessment. **JMIR mHealth and uHealth**, 2016.

DANTAS, B. D. A. Metaheurísticas para o Problema da Mochila Multidimensional. Tese, Faculdade de Computação, **Universidade Federal de Mato Grosso do Sul**, 2016.

EIBE FRANK, MARK A. HALL, The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Fourth Edition, 2016.

Disponível em:
https://www.cs.waikato.ac.nz/ml/weka/Witten_et_al_2016_appendix.pdf. Acesso em 22 jul. 2019

LEMOS EC, GOUVEIA GC, LUNA CF, SILVA GB. Programa academia da cidade: descrição de fatores de adesão e não adesão. **R. bras. Ci. e Mov** 2016, 24(4):75-84

LAFETÁ, Thiago Fialho de Queiroz. Algoritmos evolutivos many objectives aplicados ao problema de roteamento Multicast com qualidade de serviço. Dissertação de Mestrado - Universidade Federal de Uberlândia, Uberlândia, 2016.

JENSEN, M., QUESADA-LÓPEZ, C., ZUÑIGA, G., CHINNOCK, A., & JENKINS, M. Design, Development and Evaluation of a Mobile Application for Goal Setting and Self-

Monitoring of Dietary Behaviors. **Journal of the Academy of Nutrition and Dietetics**, 2015.

HONGU, N., POPE, B. T., BILGIÇ, P., ORR, B. J., SUZUKI, A., KIM, A. S., ROE, D. J. Usability of a smartphone food picture app for assisting 24-hour dietary recall: A pilot study. **Nutrition Research and Practice**, 2015.

FERNÁNDEZ ARANDA, M. I. Apps nutricionales para gestantes en atención primaria, aspectos claves para su uso. **Nutrición Clínica y Dietética Hospitalaria**, Vol. 35, Pág. 75–79, 2015

ISOWAKA, T., MATSUI, N. Performances in GA-based menu production for hospital meals. **IEEE Congress on Evolutionary Computation, CEC 2015** - Pág., 2498–2501, 2015

CUPPARI, L. *Nutrição Clínica no Adulto*. 3ª Edição. Editora Manole, 2014.

BRASIL, MINISTERIO DA SAÚDE. Guia Alimentar para a População Brasileira: promovendo a alimentação saudável, 2ª edição, **Ministério da Saúde**, Editora MS, 2014

K. Deb and H. Jain, “An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part i: Solving problems with box constraints,” *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 4, pp. 577–601, Aug. 2014.

HUTTER, MARCO. **Combinatorics**: Utility classes for combinatorics. Versão 4b72fac037bc3022f2b981bb2e0fd8ba367999c1
Disponível em: <https://github.com/javagl/Combinatorics>. Acesso em: 16 jul. 2019

WANG, G., SUN, J. A fuzzy-dominance-driven MOGA and its application in nutrition meal prescription. **International Symposium on Information Technologies in Medicine and Education, ITME 2012** – Pag, 851–855, 2012

LIMA, M.L, PADOVANI, R.M., RODRIGUEZ-AMAYA, D.B., FARFÁN, J.A, NONATO, C.T., LIMA, M.T. **Tabela Brasileira de Composição de Alimentos – TACO**, 4ª Ed. Campinas: NEPA/Unicamp, 2011

PEI, Z., LIU, Z. (2009). Nutritional diet decision using multi-objective difference evolutionary algorithm. **International Conference on Computational Intelligence and Natural Computing, CINC 2009**, Pag 77–80, 2009

SELJAK, B., Dietary menu planning using an evolutionary method, **International Conference on Intelligent Engineering Systems**. IEEE, 2006

HASHIMOTO, K. Técnicas de otimização combinatória multiobjetivo aplicadas na estimação do desempenho elétrico de redes de distribuição. Tese de Doutorado, Universidade de São Paulo - SIBUSP, 2004

RIDLEY, MARK. *Evolution*, 3ª edição, Oxford: Editora Blackwell Science Ltd, 2004

VAN VELDHUIZEN, D. A. Multiobjective Evolutionary Algorithms: Classifications, Analyses, and New Innovations. **Air Force Institute of Technology**, Wright Patterson AFB, OH, 1999.

SCHWEFEL, H. P. On the evolution of evolutionary computation. In: Zurada, J. M.; Marks, R.; Robinson, C. J., **Computation Intelligence - Imitating Life**, Pag 116-124. IEEE Press, 1994.

WHITLEY, D. A genetic algorithm tutorial. *Statistics and Computing*, v. 4, n. 2, jun. 1994.

AMABIS, J. M.; MARTHO, G. R. Curso básico de biologia, 3ª edição. São Paulo: Editora Moderna Ltda., 1985.

ROZA, A M, SHIZGAL, H M. The Harris Benedict equation reevaluated: resting energy requirements and the body cell mass, **The American Journal of Clinical Nutrition**, Vol 40, Pag 168–182, 1984.

Apêndice A

Subgrupos definidos

A Tabela 22 apresenta os subgrupos alimentares criados e a quantidade de alimentos pertencentes a cada um destes subgrupos.

Subgrupo	Quantidade de Alimentos
Açúcares	3
Bebidas de leite	17
Bebidas mistas	8
Cereais	9
Doces prontos	11
Frutas	59
Gorduras	12
Laticínios	9
Pães	9
Preparado de milho	7
Preparados de leite	2
Proteínas	14
Tubérculos	1

Tabela 22: Subgrupos alimentares criados

Apêndice B

Trabalho Publicado

Título: Algoritmo Inteligente para Otimizar a Geração de Recomendação Alimentar com Muitos Objetivos.

Autores: Róger Alves Silva de Araújo, Cicero Garrozi, Rodrigo Gabriel Ferreira Soares.

Evento: XIV Congresso Brasileiro de Inteligência Computacional (CBIC 2019).

DOI: 10.21528/CBIC2019-42

Resumo: Algoritmos evolucionários são utilizados para a obtenção de resultados satisfatórios em uma série de problemas em que algoritmos de busca tradicionais não são capazes de entregar soluções viáveis em um tempo aceitável. Existem exemplos na literatura de trabalhos que utilizaram o potencial destes algoritmos para otimizar a geração de recomendações alimentares para dados critérios nutricionais. Entre as propostas, estão exemplos com o uso de algoritmos mono-objetivo e multiobjetivo, em geral os estudos mostram que resultados nutricionalmente satisfatórios podem ser obtidos mediante o uso de ambas as técnicas, mas problemas relacionados com critérios subjetivos tais como o sabor do prato e o gosto do usuário foram recorrentes nos estudos. No presente trabalho é proposto um algoritmo que modela o problema da geração de recomendação alimentar para muitos objetivos, através do uso de um MaOEA (*Many-Objective Evolutionary Algorithm*) para a construção de recomendações alimentares, com o acréscimo de conceitos específicos do domínio do problema incorporados em seus operadores, objetivando melhorar os resultados do ponto de vista da combinação lógica dos alimentos para um usuário.