

UNIVERSIDADE FEDERAL DE OURO PRETO

Caracterização e Análise de uma Rede de Ingredientes e Receitas

Willyan Michel Ferreira
Universidade Federal de Ouro Preto

Orientador: Fabrício Benevenuto

Dissertação submetida ao Instituto de Ciências
Exatas e Biológicas da Universidade Federal de
Ouro Preto para obtenção do título de Mestre
em Ciência da Computação

Ouro Preto, Março 2014

Caracterização e Análise de uma Rede de Ingredientes e Receitas

Willyan Michel Ferreira
Universidade Federal de Ouro Preto

Orientador: Fabrício Benevenuto



C132 Ferreira, Willyan Michel.
Caracterização e análise de uma rede de ingredientes e receitas [manuscrito] / Willyan Michel Ferreira – 2014.

69f.: il.; color.; graf.; tabs.

Orientador: Prof. Fabrício Benevenuto.

Dissertação (Mestrado) - Universidade Federal de Ouro Preto. Instituto de Ciências Exatas e Biológicas. Departamento de Computação. Programa de Pós-graduação em Ciência da Computação.
Área de concentração: Recuperação e Tratamento da Informação.

1. Culinária – Teses. 2. Grafo – (Sistema de computador) - Teses. 3. Sistemas de recuperação da informação. I. Benevenuto, Fabrício. II. Universidade Federal de Ouro Preto. III. Título.

Catálogo: sisbin@sisbin.ufop.br



MINISTÉRIO DA EDUCAÇÃO E DO DESPORTO
Universidade Federal de Ouro Preto
Instituto de Ciências Exatas e Biológicas – ICEB
Programa de Pós-Graduação em Ciência da Computação



Ata da Defesa Pública de Dissertação de Mestrado

Aos 14 dias do mês de março de 2014, às 10 horas na Sala Multimídia do Instituto de Ciências Exatas e Biológicas (ICEB), reuniram-se os membros da banca examinadora composta pelos professores: **Prof. Dr. Fabrício Benevenuto de Souza (presidente e orientador)**, **Prof. Dr. Luiz Henrique de Campos Merschmann** e **Profa. Dra. Ana Paula Couto da Silva**, aprovada pelo Colegiado do Programa de Pós-Graduação em Ciência da Computação, a fim de argüirem o mestrando **Willyan Michel Ferreira**, com o título “**Caracterização e Análise de uma Rede de Ingredientes e Receitas**”. Aberta a sessão pelo presidente, coube ao candidato, na forma regimental, expor o tema de sua dissertação, dentro do tempo regulamentar, sendo em seguida questionado pelos membros da banca examinadora, tendo dado as explicações que foram necessárias.

Recomendações da Banca:

Aprovada sem recomendações

Reprovada

Aprovada com recomendações: _____

Banca Examinadora:


Prof. Dr. Fabrício Benevenuto de Souza


Prof. Dr. Luiz Henrique de Campos Merschmann


Profa. Dra. Ana Paula Couto da Silva


Prof. Dr. Haroldo Gambini Santos
Coordenador do Programa de Pós-Graduação em Ciência da Computação
DECOM/ICEB/UFOP

Ouro Preto, 14 de março de 2014.

Dedico este trabalho a minha mãe Vitória, sem ela nada disso seria possível.

Caracterização e Análise de uma rede de Ingredientes e Receitas

Resumo

A troca de receitas é um hábito de muitas pessoas. Um meio online e colaborativo de compartilhar esse tipo de informação é através de websites especializados que permitem que usuários postem receitas, comentem e avaliem receitas existentes. Apesar de extremamente populares, pouco se sabe sobre esses sistemas e os padrões de interações que eles permitem. Visando preencher essa lacuna, esse trabalho apresenta uma extensa caracterização do site Tudo Gostoso, um importante site brasileiro de compartilhamento de receitas. Para isso, nós coletamos todas as receitas existentes no site juntamente com informações associadas aos comentários e avaliações. Além de explorar as interações existentes entre os usuários do site, nosso trabalho analisa uma rede formada por ingredientes que co-ocorrem em receitas e investiga a viabilidade de se extrair possíveis alterações nas receitas a partir de comentários dos usuários do site. Nossas análises revelam padrões de uso de ingredientes fundamentais da culinária brasileira e podem ser úteis para inspirar a construção de diversas novas aplicações, como ferramentas de recomendação de receitas.

Characterization and Analysis of a network of Ingredients and Recipes

Abstract

Exchanging recipes is an habit of many people. An online and collaborative way of sharing this type of information is through specialized websites that allow users to post recipes, comment and rate existing recipes. Although extremely popular, little is known about these systems and about the patterns of interactions that they promote. Aiming at filling this gap, this paper presents an extensive characterization of the site Tudo Gostoso, an important Brazilian site for sharing recipes. To do that, we crawled all recipes on the site along with information associated to comments and evaluations. In addition to explore the interactions between users of the website, our work analyze a network of co-occurring ingredients in recipes and investigates the feasibility of extracting existing changes on recipes from commentaries. The analysis of this graph reveals patterns of ingredients to use basic ingredients of Brazilian cuisine.

Declaração

Esta dissertação é resultado de meu próprio trabalho, exceto onde referência explícita é feita ao trabalho de outros, e não foi submetida para outra qualificação nesta nem em outra universidade.

Willyan Michel Ferreira

Agradecimentos

Agradecimentos deveriam ser a parte mais fácil de um trabalho, porém demonstrar gratidão vem sempre com o receio de esquecer alguém que de certa forma foi importante nesse processo.

Primeiramente agradeço a Deus, aquele que me abençoou durante todo esse processo e iluminou meu caminho colocando pessoas como as que serão citadas abaixo.

Ninguém é nada sem a sua base, sua família e por isso agradeço a minha mãe Vitória, por toda força e por ser meu exemplo de vida, como sempre digo, se um dia eu for um terço do ser humano que ela é, já serei o segundo melhor. Agradecer também minha tia Conceição minha segunda mãe, a minha madrinha Vera que sem dúvida a pessoa que mais acredita no meu potencial. Além do meu pai Geraldo, meu padrinho Olemar e demais familiares.

Seguindo agradecer a família que fazemos por opção, os amigos. Nessa conquista não poderia deixar destacar Felipe e Alexandre que me ajudaram muito nesse processo nas tarefas com trabalho manual. Aos amigos que nesse processo ficaram um pouco mais distantes mas sempre me acolhendo em BH, Wander, Rafael, Edwaldo e Marcos.

Ao meu orientador, Fabrício Benevenuto, meu maior incentivador e motivador, que com paciência buscou sempre deixar focado no trabalho mesmo em momentos que eu fraquejei.

Aos professores Ana Couto da Silva e Luiz Merschmann, que engradeceram muito o trabalho com suas sugestões e participaram de um artigo produzido com resultados parciais desse trabalho.

Sumário

Lista de Figuras	xii
Lista de Tabelas	xiii
1 Introdução	1
1.1 Problemas e Objetivos	2
1.2 Contribuições do Trabalho	3
1.3 Publicações	3
1.4 Organização dos Capítulos	4
2 Referencial Teórico	5
2.1 Trabalhos relacionados	5
2.1.1 Redes envolvendo receitas culinárias	5
2.1.2 Recomendação de receitas	6
2.1.3 Análise de culturas gastronômicas diversas	8
2.1.4 Identificação de nomes de ingredientes em receitas	8
2.2 Redes complexas	9
2.2.1 Grau	9
2.2.2 Diâmetro	10
2.2.3 Betweenness	11

2.2.4	Coeficiente de Clusterização	11
3	Coleção de Dados	13
3.1	Dados Coletados	14
3.2	Identificação de Ingredientes	14
4	Análise dos dados	16
4.1	Ingredientes	17
4.2	Avaliações	18
4.3	Comentários	20
4.4	Usuários	21
4.4.1	Postagem de receitas x Comentários em receitas	23
4.4.2	Análise de gênero dos usuários	24
4.5	Relações entre ingredientes	26
5	Rede de ingredientes	28
5.1	Análise	28
5.1.1	Ingredientes mais importantes	29
5.1.2	Arestas mais importantes	30
6	Personalização de receitas	32
6.1	Identificação de sugestões de adição e remoção de ingredientes	34
6.2	Identificação de sugestões de substituição de ingredientes	34
6.2.1	Adições e Remoções	35
6.2.2	Substituições	36
7	Conclusão e Trabalhos Futuros	39

Referências Bibliográficas	41
Identificação de ingredientes	44

Lista de Figuras

2.1	Exemplo de grafo direcionado.	10
2.2	Exemplo Diâmetro.	10
2.3	Exemplo de cálculo de Coeficiente de Clusterização.	11
4.1	Total de ingredientes por receita.	17
4.2	Número de ingredientes distintos e únicos por categoria	18
4.3	Distribuição do total de avaliações.	19
4.4	Receitas avaliadas e quantidade de avaliações realizadas	19
4.5	Número de comentários e quantidade de receitas comentadas	21
4.6	Número de usuários ativos que comentaram e criadores de receitas	23
4.7	Correlação entre usuários que criam receitas e usuários que postam comentários.	24
5.1	Representação gráfica da rede de ingredientes.	31
6.1	Nuvem de palavras mais utilizadas nos comentários.	33
6.2	Rede de substituições.	38

Lista de Tabelas

4.1	Categories e número total de receitas.	16
4.2	Conceito médio e receitas com conceito máximo	20
4.3	Conceito médio e receitas com conceito máximo	22
4.4	Usuários do sexo masculino e feminino	25
5.1	Ranking de ingredientes, com os respectivos valores para cada métrica considerada	29
5.2	Ranking das arestas	30
6.1	Tipos de comentários encontrados na amostra.	33
6.2	10 ingredientes mais frequentes nas sugestões de adição e remoção.	36
6.3	10 ingredientes mais frequentes nas sugestões de substituição.	37

“O sucesso é ir de fracasso em fracasso sem perder entusiasmo.”

— Winston Churchill

Capítulo 1

Introdução

O advento da internet e a popularização da mesma, além da grande aceitação das redes sociais estão mudando o comportamento das pessoas. A era da informação trouxe com ela várias mudanças. Atualmente as pessoas tendem a estar conectadas a internet o tempo todo, uma vez que diversos aparelhos móveis possuem esse acesso.

Outra grande mudança é que a web é um espaço aberto para que as pessoas possam se expressar de diversas formas e redes sociais são um grande indício disso Castells (2003). Essas tendências de oportunizar o usuário a criar conteúdo levou a criação de sistemas colaborativos, como por exemplo, a Wikipedia, uma enciclopédia baseada no conhecimento criado por seus usuários Bruns (2008).

Atualmente compartilhar informações vem se tornando cada vez mais comum através desse meio, como exemplo, o Youtube recebe aproximadamente 20 horas de vídeo a cada minuto. Com esses valores, em aproximadamente 15 dias, o conteúdo gerado superaria o da Rede Globo em toda sua história Times (2009). Outro fator além da geração de conteúdo é a quantidade de usuários dessas redes sociais. O Facebook em 2012 superou a marca de 1 bilhão de usuários, número que aumentou 100% em 2 anos de São Paulo (2010). Tudo isso demonstra a tendência do compartilhamento de informações sejam opiniões, vídeos, fotos ou até receitas culinárias.

Receitas em seus primórdios eram tratadas como um tesouro de família, um segredo passado entre gerações Lole et al. (2002) Alves et al. (2008). Com o tempo essas começaram a serem compartilhadas por meio de livros e atualmente em larga escala em sistemas colaborativos de receitas culinárias, como por exemplo, o site *Allrecipes.com*¹,

¹<http://allrecipes.com/>

considerado um dos mais populares desse seguimento, e que possui mais de 46 mil receitas voltadas apenas para seus usuários americanos. Teng et al. (2012)

A internet pode ter sido o grande fator para o crescimento do compartilhamento de receitas por proporcionar ao criador uma interação com outros usuários e o reconhecimento do conteúdo por ele criado, por meio de comentários e avaliações. Além do alcance, antes restrito à família e aos amigos, essa informação online possibilita que pessoas de qualquer lugar do mundo possam preparar uma dessas receitas e postar suas impressões Ihle et al. (2009).

1.1 Problemas e Objetivos

Sites de receitas culinárias são sistemas colaborativos, onde usuários trocam experiências, comentam e avaliam receitas uns dos outros. São sistemas que permitem que receitas possam ser visualizadas por milhões de usuários, que interagem com outros usuários interessados em receitas ou mesmo com outros criadores. Essa troca de experiências e o retorno em larga escala, pode ser o fator determinante para crescimento desse tipo de sistema. Abreu et al. (2008).

Apesar do grande interesse, pouco se sabe sobre os tipos de interações que acontecem nesses sistemas e sobre o comportamento dos usuários que navegam e usam esses sistemas.

Este trabalho tem como objetivo caracterizar os padrões de interação de um sistema colaborativo de receitas culinárias, visando o desenvolvimento futuro de aplicações a partir dessa caracterização. Para tal será realizada a coleta de dados de um portal de receita. Esses dados serão minerados e quantificados, buscando padrões nas relações dos ingredientes em receitas.

Outra grande fonte de informação a ser trabalhada são os comentários dos usuários em receitas, uma vez que esses podem conter algum tipo de sugestão de modificação para receita, o que pode ser de grande valor para o desenvolvimento de um sistema, por exemplo, de recomendação de ingredientes que se substituem.

Estudos como esse são de grande valor para comunidade e impacta diretamente em duas grandes áreas de investigação. A primeira, do ponto de vista sociológico, sendo uma investigação de novas formas de interação na web, relativo a um elemento central existente em diversas culturas: receitas culinárias. O outro é do ponto de vista

tecnológico, visto que os padrões identificados no uso desses sistemas podem inspirar novas ferramentas, como sistemas de recomendações de receitas ou mesmo dar subsídios a mudanças na interface ou nos módulos de interação.

1.2 Contribuições do Trabalho

As principais contribuições deste trabalho são:

- Criação de uma base de ingredientes da culinária brasileira. Uma das grandes dificuldades do trabalho foi a identificação dos ingredientes e isso muito se deu pela falta de uma base como a desenvolvida neste trabalho, que poderá ser utilizada em trabalhos semelhantes.
- Identificação de padrões nas relações entre os ingredientes. Tais padrões podem caracterizar grupos de ingredientes que ocorrem junto nas receitas com frequência. Esses grupos ingredientes identificados podem ser a base para novas receitas, que seriam formados por um desses grupos com acréscimo de outros ingredientes.
- Caracterização dos usuários de um sistema colaborativo de receitas culinárias brasileiras. Nossa análise permite traçar o perfil do usuário brasileiro desse tipo de sistema e pode sugerir mudanças de interface desses sistemas mais voltadas às necessidades do público observado. Entender o comportamento dos usuários de um sistema permite ao mesmo evoluir de acordo com as expectativas desses.

1.3 Publicações

Resultados parciais desta pesquisa foram publicados e apresentados em evento científico através da seguinte publicação:

- FERREIRA, W. M.; COUTO DA SILVA, A. P.; MERSCHMANN, L. H. C.; BE-NEVENUTO, F.: *Comer, Comentar e Compartilhar: Análise de uma rede ingredientes e receitas*. In: *X Simpósio Brasileiro de Sistemas Colaborativos (SBSC)*, 2013, Manaus, Amazonas, Brasil.

No artigo descrevemos a base de dados utilizados, bem como a metodologia para que a coleta fosse realizada. As análises foram feitas de forma geral para toda a base, abordando apenas alguns pontos sobre as categorias de receitas da base.

1.4 Organização dos Capítulos

O restante do trabalho está organizado da seguinte forma. O **Capítulo 2** aborda outros esforços da área de pesquisa por meio de trabalhos relacionados. Em seguida, o **Capítulo 3** descreve a base dados utilizada neste trabalho, bem como se deu a coleta e processamento da mesma. Depois, no **Capítulo 4**, descrevemos de maneira geral as análises realizadas na base de dados, esse capítulo ainda é composto de seções onde são descritas análises por categoria. Seguindo no **Capítulo 5**, demonstramos a rede de ingredientes gerada no trabalho, além de explicitar resultados gerados a partir da análise dessa por meio de métricas de mineração de dados em grafos. Logo após o **Capítulo 6** demonstra quais alterações a base sofreria seguindo sugestões dos usuários do site utilizado como base de dados do trabalho. Foram observadas as sugestões de adição, remoção e substituição de ingredientes. Finalmente, no **Capítulo 7** concluímos o trabalho e discutimos direções para trabalhos futuros.

Capítulo 2

Referencial Teórico

Este capítulo se divide em duas seções, na primeira apresentaremos trabalhos que também envolvem receitas culinárias, recomendação de receitas, análise de culturas gastronômicas diversas e identificação de nomes de ingredientes de receitas. A segunda seção Redes complexas apresentamos métricas utilizadas em análises de redes complexas.

2.1 Trabalhos relacionados

Nesta seção serão apresentados trabalhos envolvendo redes com receitas culinárias.

2.1.1 Redes envolvendo receitas culinárias

Ahn, Ahnert, Bagrow e Barabási Ahn et al. (2011) criaram uma rede de ingredientes onde esses se conectam de acordo com seus componentes químicos.

O estudo foi realizado com receitas de dois repositórios diferentes, sendo um ocidental e um oriental, mas com diversas culturas de culinárias como: Norte Americana, Europa Ocidental, Sul da Europa, América Latina e Leste Asiático. Ao todo foram trabalhadas 56.498 receitas e foi observado que em média eles tem 8 ingredientes.

Foi constatado que as receitas da cultura Norte Americana e da Europa Ocidental tendem a ter ingredientes que compartilham componentes químicos. Em contraste a cozinha do Leste Asiático e do Sul da Europa tendem a receitas onde seus ingredientes não compartilham ingredientes químicos.

Em outro esforço Teng, Lin e Adamic Teng et al. (2012) desenvolveram duas redes, sendo uma envolvendo ingredientes que co-ocorrem e outro ingredientes que se substituem.

O trabalho foi realizado utilizando como base as receitas americanas do site *allrecipes.com* um dos mais populares sites de compartilhamento de receitas, ao todo foram coletadas e trabalhadas 46.337 receitas.

Para criação da rede de ingredientes co-ocorrentes, foi desenvolvido uma métrica que visa observar a tendência de dois ingredientes estarem juntos em receitas, a partir da probabilidade deles ocorrem juntos em receitas sobre a a probabilidade de ocorrem separados.

Essa rede foi criada no intuito de observar o impacto nas avaliações recebidas pelas receitas quando essas tem ingredientes que tendem a ocorrer juntos, como resultado foi detectado que isso ocorre de forma muita pequena.

Para a o desenvolvimento da rede de ingredientes que se substituem, o primeiro passo foi identificá-los e para isso optou-se por não utilizar de bases da *web* uma vez as pesquisas não proviam informações para todos os ingredientes, logo optou-se por indentificá-los por meio dos comentários realizados pelos usuários nas receitas.

A rede de substitutos serviu de base para algoritmo de recomendação de receitas proposto no trabalho citado. A rede desenvolvida nesse trabalho visa observar as relações dos ingredientes que compões as receitas, visando identificar os mais importantes.

2.1.2 Recomendação de receitas

Um dos primeiros estudos realizados envolvendo receitas conhecido e o primeiro envolvendo recomendação de receitas é o CHEF desenvolvido por Hammond Hammond (1986), um aplicativo com aprendizagem de máquina, que visa dar ao usuário a melhor receita baseando-se nas preferências, texturas, ingredientes e pesquisas passadas.

Vários esforços tem sido feitos no que tange a recomendação de receitas, Svensson, Höök e Cöster Svensson et al. (2005) fizeram um trabalho focando o usuário de um sistema colaborativo de receitas e algumas descobertas foram observadas por meio de questionários.

Eles conseguiram constatar que a maioria dos usuários, alvo do estudo, consideram

um receita realmente boa a partir dos comentários positivos que ela recebe, esse quesito tem peso maior do que uma métrica que deveria atestar a qualidade que são as avaliações da mesma.

Com isso eles observaram que avaliações talvez não seja a métrica ideal para recomendação de receitas, que no início do estudo as recomendações estavam bem aceitas usando esse parâmetro, mas no fim do estudo elas não foram bem aceitas e esse era o momento que a precisão delas já deveria ser maior, esse estudo se deu num período 6 meses e envolveu 302 usuários.

Ainda observando o usuário, Yajima e Kobayashi Yajima and Kobayashi (2009) desenvolveram um sistema de recomendação de receitas fáceis de se preparar voltadas a pessoas com pouca experiência na cozinha além dos ingredientes de preferências do usuário.

Em outro esforço Zhang e outros Zhang et al. (2008) focaram a recomendação a partir do ingrediente principal de uma receita acessada pelo usuário, por exemplo o usuário está visitando a receita "Bolo de cenoura", cujo o ingrediente principal é cenoura, logo ele recomenda outras receitas que possuem esse ingrediente levando em consideração as avaliações dessas receitas, esse tipo de abordagem tem sido utilizada em diversos trabalhos Kuo et al. (2012), como por exemplo Maruyama, Kawano e Yanai que desenvolveram um aplicativo para dispositivos móveis que baseado na imagem de ingrediente recomenda receitas com aquele ingrediente, ou seja, o usuário tira a foto de uma fruta e o aplicativo recomenda receitas com a fruta reconhecida Maruyama et al. (2012).

Ueda, Takahata e Nakajima Ueda et al. (2011) desenvolveram uma recomendação voltada as preferências do usuário, para isso eles se utilizaram do histórico de navegação de usuários a um site receitas, a partir disso eles identificaram os ingredientes e tipo(categoria) de comida preferida do usuário e com isso realizavam a recomendação. Bodlaender e outros Maarten Peter Bodlaender (2010) desenvolveram e patentearam um sistema de recomendação que desenvolveram, baseado em aprendizado de máquina, o sistema guarda as receitas a medida que o usuário vai utilizando o mesmo e recomenda a partir do histórico de utilização.

Geleijnse e outros Geleijnse et al. (2011) fizeram um trabalho segue o mesmo caminho dos já citados, porém suas recomendações tem o foco voltado para receitas saudáveis, logo previamente eles identificam ingredientes considerados mais saudáveis, identificam as preferências do usuário e assim recomendam receitas de forma que esse possa ter um ganho se considerar aspectos relacionados a sua saúde, outros trabalhos seguem essa

linha de pesquisa van Pinxteren et al. (2011) Freyne and Berkovsky (2010b) Freyne and Berkovsky (2010a) Mino and Kobayashi (2009).

Nosso trabalho não visa questões relacionadas a recomendação, mas sim uma caracterização de padrões de interação e uso de ingredientes em sistemas de compartilhamento de receitas. Acreditamos que resultados que a caracterização realizada em nosso estudo possa ser uma base consiste para um futuro desenvolvimento de bom sistema de recomendação de receitas voltado para o contexto específico das receitas brasileiras.

2.1.3 Análise de culturas gastronômicas diversas

Como já fora dito Ahn, Ahnert, Bagrow e Barabási Ahn et al. (2011) realizaram um estudo e observaram que as culturas norte americana e da europa ocidental os usuários tendem a preferir receitas com alguns ingredientes que compartilhem compostos químicos, ao contrário do leste asiático e sul da europa que preferem por ingredientes com compostos diferentes.

Observando a culinária chinesa Zhu e outros Zhu et al. (2013) realizaram um vasto estudo caracterizando diversos aspectos da mesma, eles observaram que a China possui 20 culinárias regionais e que a proximidade geográfica é fator determinante para isso e que o clima não influencia. Verificaram ainda que em médias as receitas possuem 10 ingredientes, variando entre 1 e 35.

2.1.4 Identificação de nomes de ingredientes em receitas

Hamon e Grabar Hamon and Grabar (2013) utilizaram técnicas LPN a partir de um sistema híbrido envolvendo também métodos de CRF e com aprendizagem de máquina, para o tratamento do texto e identificação dos termos foram criadas algumas categorias de possíveis entradas como stop words, utensílios de cozinha, ações, medidas, quantidades e ingredientes.

O estudo foi feito para uma base de dados contendo aproximadamente 20.000 receitas, os resultados obtidos foram considerados satisfatórios, porém ressaltam que alguns pontos não foram abordados, como correção ortográfica.

Teixeira desenvolveu um trabalho envolvendo a culinária brasileira e cita algumas especificidades da mesma, como também observamos, questões como regionalidade e adaptação

de nome de ingredientes de origem estrangeira Teixeira (2009).

Em nosso trabalho optamos pela identificação manual dos ingredientes para ter um controle maior da base e observar casos como agrupamento de ingredientes, regionalização de nomes o mesmo a correção ortográfica.

2.2 Redes complexas

Para caracterização e análise das redes de ingredientes, diferentes métricas de teoria de grafos são utilizadas, especialmente métricas topológicas. Métricas topológicas são definidas como medidas baseadas em atributos estruturais de um grafo. Estas métricas podem considerar cada vértice ou a visão global do grafo Newman (2003).

2.2.1 Grau

O grau de um vértice pode ser definido pela quantidade de arestas incidentes a ele, no caso dos laços (aresta com mesmo origem e destino) contando duas vezes.

AlvesBruno Leite Alves (2013) cita que o grau de um vértice é uma característica importante na estrutura de uma rede e normalmente segue a lei de potência, como a Internet Faloutsos et al. (1999) , a *Web* Barabási and Albert (1999) e redes neurais Braitenberg and Schüz (1998).

Para grafos direcionados o grau de entrada de um vértice, é o total de arestas que incidem no vértice. E o grau de saída de um vértice, é o total de arestas que tem origem no vértice. Nesse tipo de grafo o grau do vértice pode ser definido pela soma do seu grau de entrada com o seu grau de saída.

Na figura 2.1 temos um grafo com 3 vértices, todos de grau 2, porém com graus de entrada e saída diferentes. O vértice A possui grau de entrada igual a 1 e de saída igual a 1. O vértice B possui grau de entrada 0 e de saída 2. E por fim o vértice C possui grau de entrada 2 e de saída 0.

Para grafos não-direcionados o grau do vértice é exatamente a sua quantidade de arestas com outros vértices.

Uma métrica muito utilizada na análise de grafos é o cálculo do grau médio do mesmo. E esse é dado pela soma dos graus de todos os vértices do grafo dividido pelo número

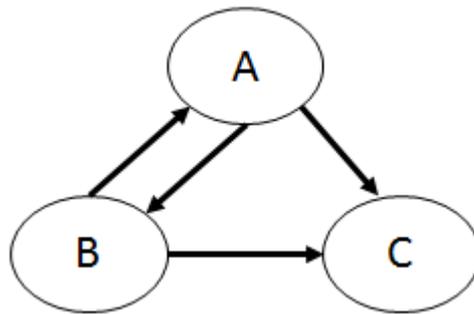


Figura 2.1: Exemplo de grafo direcionado.

de vértices, utilizando como exemplo a Figura 2.1 o grau médio seria 2.

2.2.2 Diâmetro

O diâmetro pode ser definido com o maior caminho entre 2 vértices de um grafo. Para o cálculo dessa métrica é basicamente somar o número de arestas entre 2 vértices. Como no exemplo da Figura 2.2 o diâmetro é 3, sendo o maior caminho entre o vértice A e G.

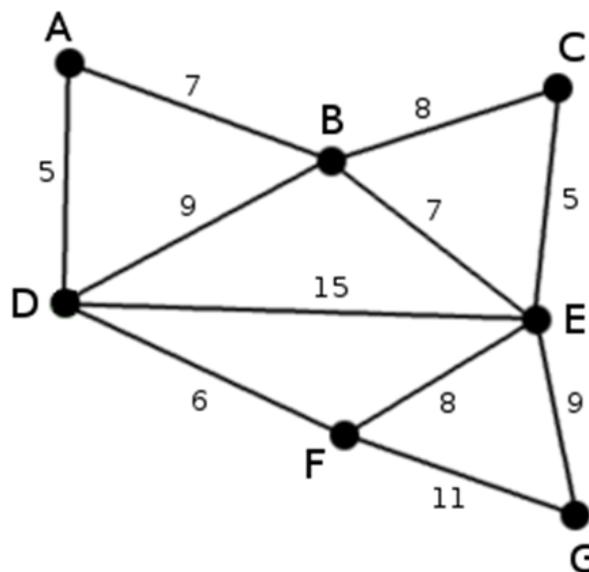


Figura 2.2: Exemplo Diâmetro.

2.2.3 Betweenness

O *Betweenness* de um vértice v é a fração dos caminhos mínimos, que ligam qualquer par de vértices e que passam pelo vértice v . Em outras palavras, seja $\sigma_{u,j}$ o total de caminhos mínimos entre u e j , e $\sigma_{u,j}(v)$ o número total de caminhos que passam por v . A métrica de *Betweenness* é definida como:

$$Betweenness(v) = \sum_{\forall v \neq u \neq j \in V} \frac{\sigma_{u,j}(v)}{\sigma_{u,j}}. \quad (2.1)$$

Também pode-se calcular o *betweenness* de uma aresta da rede, mas para esse caso é observado o número de caminhos mínimos que passam por determinada aresta.

Logo *Betweenness* reflete o grau de dispersão dos vértices do grafo \mathcal{G} , além de buscar quantificar a importância de vértices e arestas de uma rede. Benevenuto et al. (2011)

2.2.4 Coeficiente de Clusterização

O coeficiente de clusterização cc de um vértice v é definido como sendo a divisão da quantidade de arestas existentes entre os vizinhos de v e o número máximo de arestas possíveis entre os vizinhos do vértice v , como podemos observar na Figura 2.3.

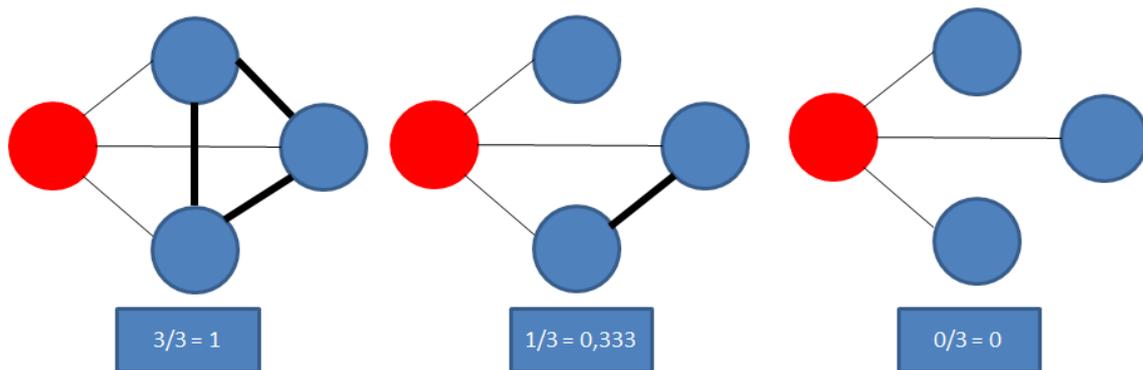


Figura 2.3: Exemplo de cálculo de Coeficiente de Clusterização.

O cálculo do coeficiente de clusterização de uma rede ou grafo é calculado a partir da média dos cc de todos os nós. Figueiredo (2011)

É uma métrica que visa indicar a conectividade entre os vértices de um grafo, observa-se que normalmente vértices tendem a criar grupos, logo a probabilidade de existirem

arestas entre vértices de forma a criarem esses grupos é maior do que a incidencia de ligações aleatórias Bruno Leite Alves (2013).

Capítulo 3

Coleção de Dados

Para entender as interações que ocorrem num sistema colaborativo de receitas culinárias brasileiras é necessária uma base de dados. Para tanto procuramos alguns portais nacionais de receitas e optamos pelo Tudo Gostoso, visto que esse era o que possuía o maior número de receitas.

O site citado, permite ao usuário algumas atividades como: acessar, postar, avaliar receitas, além de postar comentários nas receitas. Foram coletadas todas as receitas públicas, bem como os ingredientes que as compõem, além dos comentários, respostas e avaliações dos usuários do site. Para obtenção dos dados foi desenvolvido um crawler em *Python*.

Escolhida a linguagem, foi traçada uma estratégia de coleta que se dividiu em duas etapas. Na primeira etapa foram coletados os links das 11 categorias apresentadas no portal. Em seguida, o coletor desenvolvido caminhou por todas as receitas de cada categoria e armazenando na base todo o texto html das páginas, de onde futuramente seriam extraídas informações como:

- nome da receita;
- ingredientes;
- modo de preparo;
- número de avaliações;
- número de comentários;

- tempo de preparo;
- rendimento;
- nome do criador da receita;

A segunda etapa foi a de coleta dos comentários das receitas. Esses são dispostos de 10 em 10 em cada página, para coleta de todos os comentários foi desenvolvido um coletor para visitar todas as páginas de comentários de cada receita.

Essa coleta foi realizada entre os dias 2 e 5 de abril de 2013.

3.1 Dados Coletados

Nossos dados resultam de uma coleta exaustiva das receitas do site Tudo Gostoso e agrupam um total de 145.625 receitas. Foram observados também um total de 383.189 comentários e 320.504 avaliações nessas receitas.

Nossa coleta também observou um total de 236.189 usuários considerados ativos. Visto que só possuímos os dados públicos desse portal. Não podemos afirmar que esse é número total de usuários cadastrados no mesmo, logo consideramos usuários aqueles que realizaram alguma atividade, como postar ou comentar pelo menos uma receita.

A seguir, discutimos as estratégias utilizadas para identificação dos ingredientes presentes nas receitas.

3.2 Identificação de Ingredientes

Para formar a base de ingredientes, as receitas foram processadas visando essa identificação. Nesse processo foram retiradas as quantidades e medidas do texto, por exemplo, “1 colher de chá de sal” extraímos apenas “sal”, visto que o “1?” seria no caso a quantidade e “colher de chá?” uma medida.

A identificação desses ingredientes aconteceu de forma semiautomática. Num primeiro momento todos os ingredientes pré-identificados foram agrupados num arquivo e eliminados os duplicados, o que resultou em cerca de 5.000 termos, esses passaram por um processo de identificação manual, que contou a participação de 3 voluntários.

A opção por essa forma de identificação se deu pois não foi encontrada nenhuma fonte de dados para nomes de ingredientes. Outro fator determinante foram as várias especificidades encontradas no trabalho de identificação dos ingredientes, como a diferenciação regional da nomenclatura dos ingredientes, fato que se justifica sendo o Brasil um país continental e multicultural. Um exemplo desse caso é o ingrediente “mandioca”, também conhecido como “macaxeira” ou ainda “aipim”.

Outro caso a ser destacado é o agrupamento de ingredientes encontrados. Esses foram agrupados por se tratar do mesmo produto, porém como nomes diferentes como: “achocolatado” e “achocolato em pó” ou mesmo “canela em pau” que foi encontrada ainda como “canela de pau”, “canela em casca”, “canela em rama” e “pau de canela”.

Ainda ocorreu a adequação dos ingredientes identificados a língua portuguesa, por exemplo, “*whisky*” que se tornou “uísque”.

E por fim, foi feita uma exaustiva correção ortográfica nos ingredientes, uma vez que o conteúdo do site é criado por usuários e não passam por uma observação da administração do portal. Foram encontrados diversos erros, como por exemplo, para o ingrediente “óleo” que também apareceu como: “oio”, “olheo” e “olieo”.

Foram identificados 1.397 ingredientes distintos e um total de 1.079.835 ingredientes se observadas todas as receitas.

Todos os ingredientes que foram agrupados ou passaram por algum tipo de correção estão sumarizados no Apêndice e estão disponibilizados ¹.

¹<https://sites.google.com/site/comercomentarecompartilhar/>

Capítulo 4

Análise dos dados

A base de dados coletada possui um total de 145.625 receitas, divididas em 11 categorias descritas na Tabela 4.1, aonde se pode observar a categoria, o número de receitas e o número médio de ingredientes por receita.

Tabela 4.1: Categorias e número total de receitas.

Categoria	Número de Receitas	Número médio de Ingredientes por receita
<i>Sopas</i>	2.796	8,77
<i>Bebidas</i>	4.557	4,53
<i>Peixes e Frutos do Mar</i>	5.236	9,05
<i>Aves</i>	6.216	8,39
<i>Carnes</i>	6.839	8,06
<i>Prato Único</i>	7.057	9,61
<i>Massas</i>	16.777	8,37
<i>Lanches</i>	17.667	8,26
<i>Saladas, Molhos e Acompanhamentos</i>	19.315	7,52
<i>Bolos e Tortas</i>	26.948	7,64
<i>Doces e Sobremesas</i>	32.217	5,43
<i>Total</i>	145.625	7,4

Os números indicam a categoria “Sopas” como a de menor número de receitas. Isso pode ter relação com o clima do país, uma vez que estas são normalmente preparadas em períodos em que a temperatura é baixa e o Brasil possui um clima tropical. Mesmo no inverno as temperatura são amenas.

Analisando as categorias com maior número de receitas, sobressaem “Bolos e Tortas” e “Doces e Sobremesas”. Isto deve-se às receitas possuírem uma grande incidência de

ingredientes em comum e ao fato de que receitas desse tipo são as preferidas das mulheres, gênero que normalmente é maioria em sistemas de compartilhamento de receitas.

4.1 Ingredientes

Observando os ingredientes das receitas, aquela com maior número de ingredientes é a denominada “Cozido”¹, com um total de 33 ingredientes. E as receitas com menor número de ingredientes encontradas possuem apenas 1 ingrediente, como por exemplo, a “Bacon sem fumaça”², que utiliza apenas bacon.

A Figura 4.1 mostra a distribuição de probabilidades do total de ingredientes nas receitas do site.

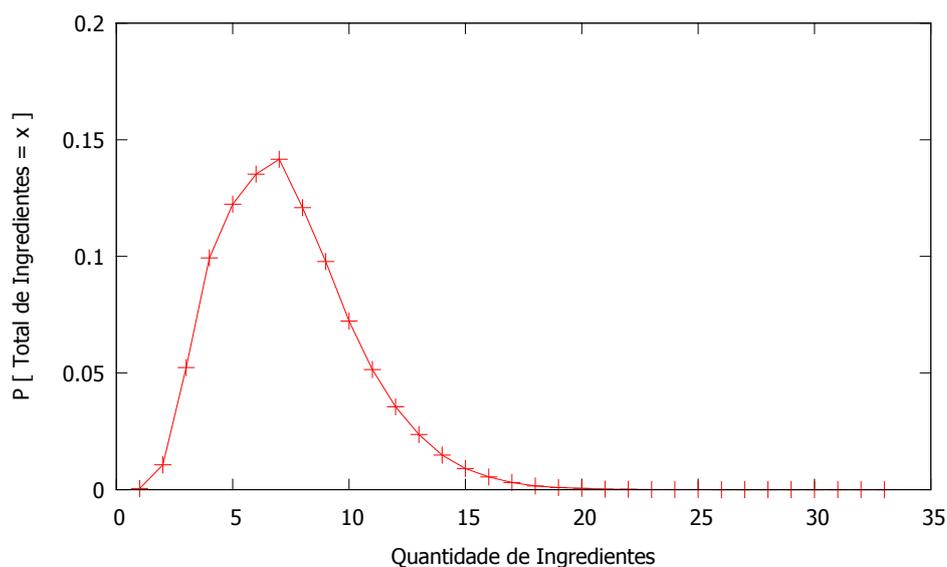


Figura 4.1: Total de ingredientes por receita.

A distribuição segue uma forma semelhante a uma distribuição normal com número médio de ingredientes igual a 7,4.

O número de ingredientes distintos e ingredientes únicos, aqueles utilizados apenas numa categoria específica, podem ser observados na Figura 4.2

A categoria “Sopas” aparece como a de menor número de ingredientes distintos e

¹<http://tudogostoso.uol.com.br/receita/77761-cozido.html>

²<http://www.tudogostoso.com.br/receita/119-bacon-sem-fumaca.html>

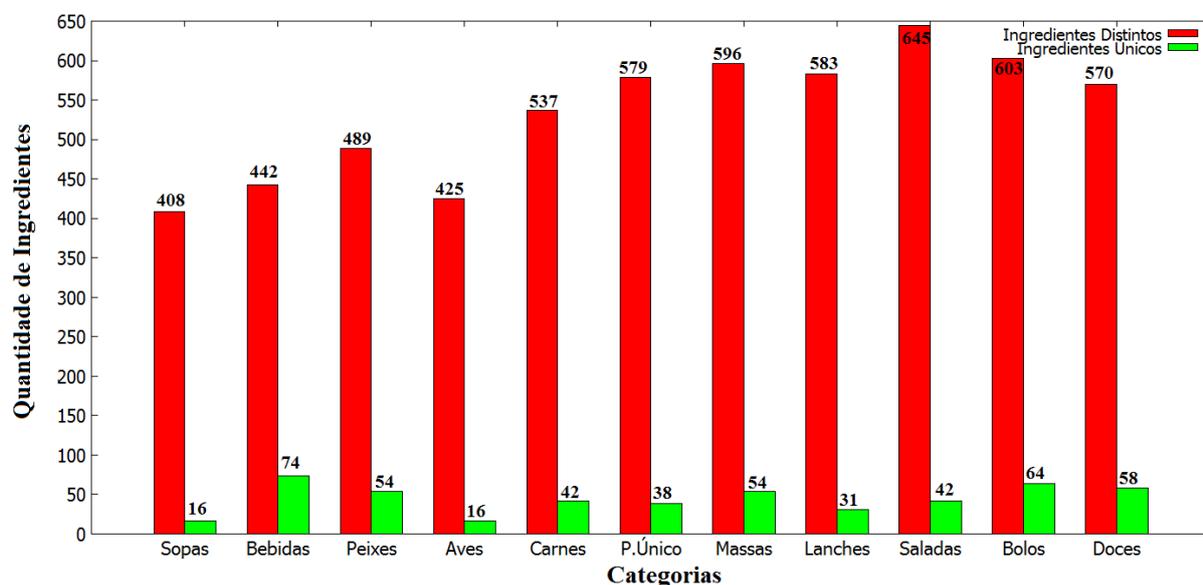


Figura 4.2: Número de ingredientes distintos e únicos por categoria

também com a de menor número de ingredientes únicos, o que pode estar ligado ao fato dessa ser a categoria com o menor número de receitas.

Ao se analisar a categoria com o maior número de ingredientes distintos, no caso “Bebidas”, imagina-se que isso ocorre, pois ingredientes como bebidas alcoólicas observadas na mesma, não são ingredientes comuns em receitas das outras categorias analisadas.

4.2 Avaliações

Como já foi citado, o site permite aos usuários avaliarem as receitas, atribuindo conceitos entre 0 e 5. Somente 31% das receitas foram avaliadas, sendo o conceito médio das avaliações igual a aproximadamente 4,4. A distribuição do total de avaliações para cada uma das receitas segue uma lei de potência, aonde poucas receitas recebem um número maior de avaliações e várias receitas recebem um número pequeno de avaliações, como pode-se observar na Figura 4.3.

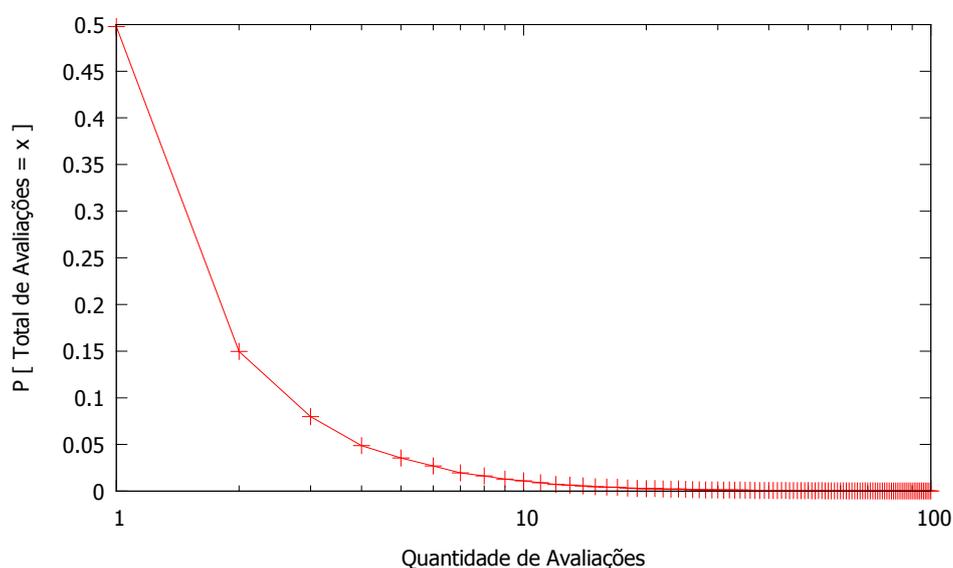


Figura 4.3: Distribuição do total de avaliações.

Também foi feito um estudo sobre a quantidade de avaliações realizadas e o número de receitas avaliadas por categoria, como mostra a Figura 4.4.

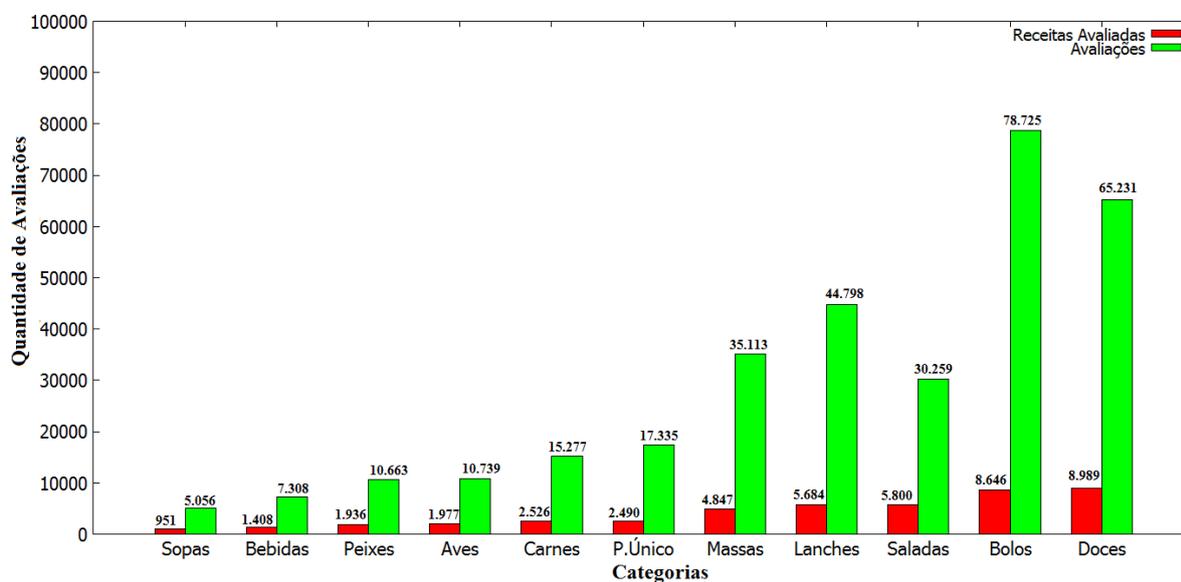


Figura 4.4: Receitas avaliadas e quantidade de avaliações realizadas

Como se observa, fica claro que a quantidade de receitas avaliadas está diretamente ligada ao número de receitas, visto que as categorias estão dispostas em ordem crescente do número de receitas. Mas isso não ocorre quando analisado o número de avaliações,

cuja categoria “Bolos e Tortas” se destacou como a que recebeu a maior quantidade.

As avaliações são feitas por meio de voto, variando entre os valores 0 e 5. O conceito geral das receitas do site foi de 4,4. Os valores por categoria e quantidade de receitas que receberam conceito máximo podem ser vistos na Tabela 4.2.

Tabela 4.2: Conceito médio e receitas com conceito máximo

Categoria	Conceito Médio	Receitas com conceito máximo(5)
<i>Sopas</i>	4,48	529
<i>Bebidas</i>	4,22	733
<i>Peixes e Frutos do Mar</i>	4,49	1.086
<i>Aves</i>	4,48	1.084
<i>Carnes</i>	4,44	1.300
<i>Prato Único</i>	4,48	1.372
<i>Massas</i>	4,41	2.547
<i>Lanches</i>	4,34	2.765
<i>Saladas, Molhos e Acompanhamentos</i>	4,41	2.987
<i>Bolos e Tortas</i>	4,37	4.269
<i>Doces e Sobremesas</i>	4,40	4.731
<i>Total</i>	4,40	23.403

Os conceitos são todos próximos à média geral, sendo o maior o da categoria “Peixes e Frutos do Mar” e o menor, o da categoria “Bebidas”. Como esperado o número de receitas que recebeu conceito máximo está diretamente ligado à quantidade de receitas da categoria.

4.3 Comentários

Em relação aos comentários nas receitas, a Figura 4.5 demonstra o número de comentários e a quantidade de receitas que receberam esses comentários.

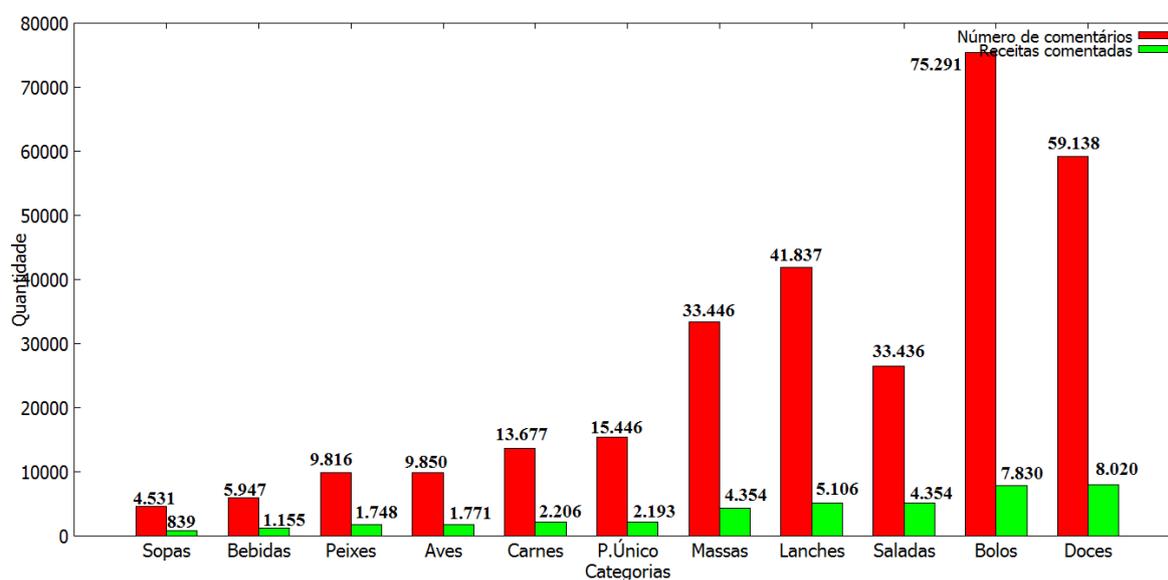


Figura 4.5: Número de comentários e quantidade de receitas comentadas

Fica notável que os números de comentários estão diretamente ligados às avaliações, sendo as categorias mais comentadas também as mais avaliadas. Sobre o número de receitas comentadas, estas estão relacionadas ao número de receitas da categoria.

4.4 Usuários

Os usuários ativos do site totalizam 236.189. Fora realizada uma análise por categoria desses usuários ativos, como se pode observar na Tabela 4.3.

Tabela 4.3: Conceito médio e receitas com conceito máximo

Categoria	Número de Usuários ativos
<i>Sopas</i>	5.417
<i>Bebidas</i>	7.135
<i>Peixes e Frutos do Mar</i>	10.719
<i>Aves</i>	11.137
<i>Carnes</i>	13.766
<i>Prato Único</i>	15.072
<i>Massas</i>	30.462
<i>Lanches</i>	30.702
<i>Saladas, Molhos e Acompanhamentos</i>	26.258
<i>Bolos e Tortas</i>	51.172
<i>Doces e Sobremesas</i>	46.889
<i>Total</i>	4,40

Assim, nota-se que o número de usuários ativos tende a seguir o número de receitas da categoria. Duas categorias se destacam: de maneira positiva, a categoria “Bolos e Tortas” é a categoria mais ativa, que possui mais receitas, usuários, comentários e avaliações; negativamente, a categoria “Saladas, Molhos e Acompanhamentos”, visto que os comentários, avaliações e usuários não seguem a tendência de acompanhar o número de receitas, sendo esta a terceira maior categoria quanto a esse quesito.

Outro estudo realizado foi quanto as atividades desses usuários e a Figura 4.6 demonstra o número de usuários que criaram e comentaram receitas.

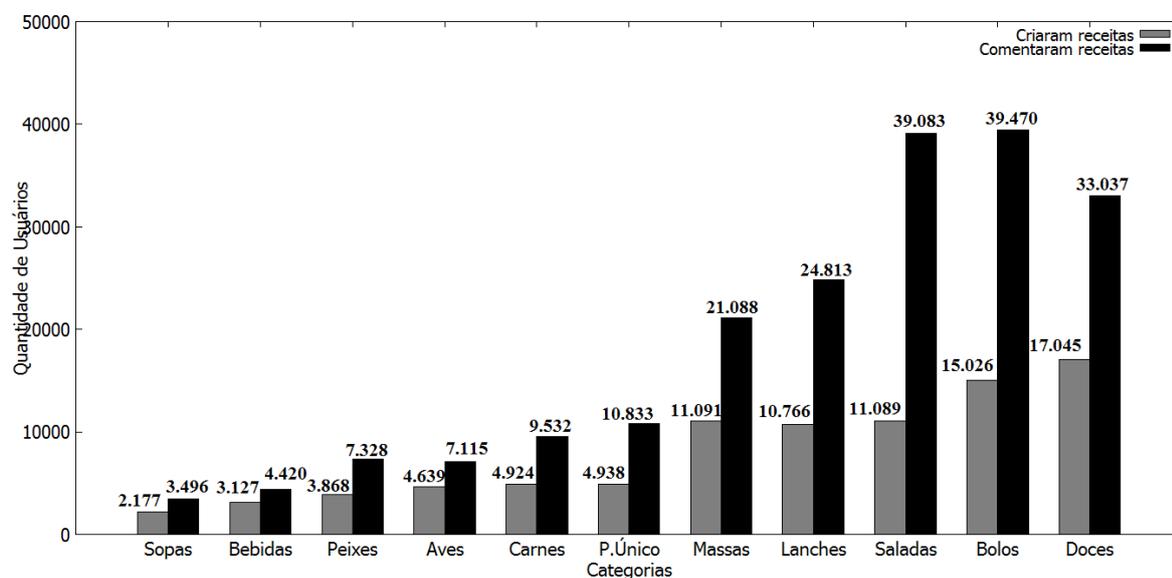


Figura 4.6: Número de usuários ativos que comentaram e criadores de receitas

4.4.1 Postagem de receitas x Comentários em receitas

Uma questão interessante investigada é foi a correlação entre usuários que mais criam receitas e aqueles que mais comentam em receitas. Para tal, foi calculado o coeficiente de correlação de Pearson.

O valor para esta métrica foi igual a 0,87, o que significa que usuários que participam ativamente na criação de receitas são também aqueles com o maior número de comentários postados.

A Figura 4.7 mostra o gráfico de dispersão entre criação de receitas e comentários.

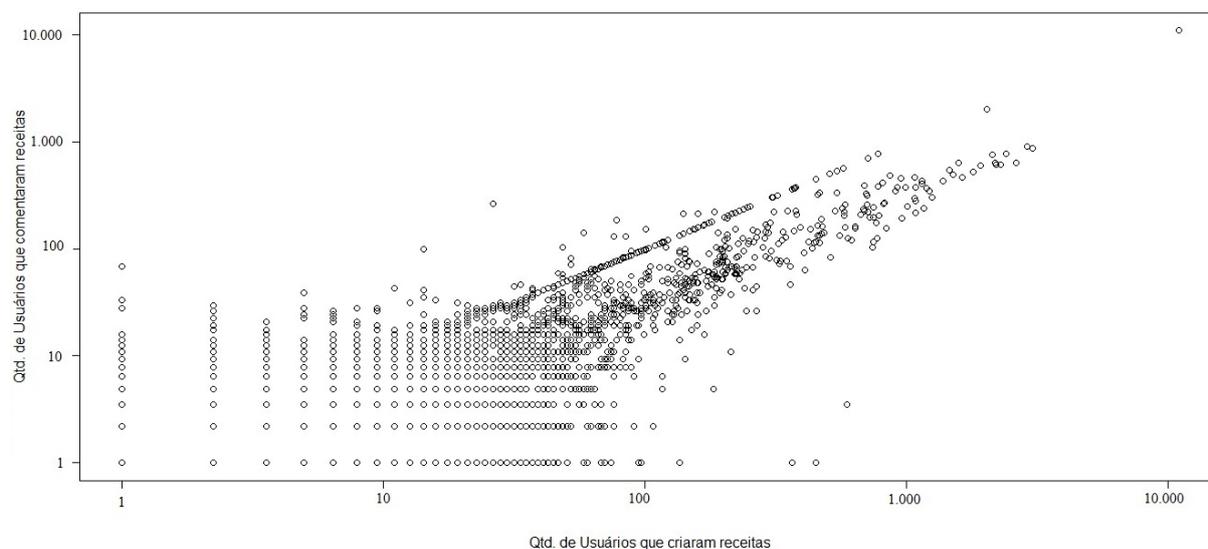


Figura 4.7: Correlação entre usuários que criam receitas e usuários que postam comentários.

Um outro ponto a ser verificado é se usuários simplesmente postam receitas ou se posteriormente também realizam comentários e/ou respondem a dúvidas relacionadas com suas receitas. Observou-se que 3,86% do total de usuários criadores de receitas (6.499 usuários) comentaram suas próprias receitas, sendo responsáveis por um de total 7.366 dos comentários. Em relação a respostas, 12,21% do total de usuários que criaram receitas (1.033 usuários) responderam a algum comentário feito em suas respectivas receitas.

4.4.2 Análise de gênero dos usuários

Nesta seção, será abordada a identificação de sexo dos usuários do site. Isso deu-se de forma semiautomática, uma vez que houve um casamento com listas de nomes e uma parte de identificação manual.

Nesse processo inicialmente fez-se um tratamento nos nomes dos usuários, aonde se aproveitou apenas o primeiro nome. Além disso, todos foram passados para minúsculo e seus acentos retirados. Por exemplo, o usuário ?João José da Silva?, passaria a ser tratado como “joao”.

Posteriormente, foram levantadas listas de nomes para identificação do gênero. A junção de algumas formou a base de nomes, divididas em dois grupos: sexo masculino e

Tabela 4.4: Usuários do sexo masculino e feminino

Categoria	Homens (%)	Mulheres (%)	Não Identificado (%)
<i>Bolos e Tortas</i>	3.000 (5,9)	47.768(93,3)	403 (0,8)
<i>Carnes</i>	1.281 (9,3)	12.370 (89,9)	115 (0,8)
<i>Aves</i>	818 (7,3)	10.226 (91,5)	93 (0,8)
<i>Peixes e Frutos do Mar</i>	954 (8,9)	9.693 (90,4)	72 (0,7)
<i>Saladas, Molhos e Acom.</i>	1.922 (7,3)	24.022 (91,5)	313 (1,2)
<i>Sopas</i>	381 (7,0)	5.004 (92,4)	32 (0,6)
<i>Massas</i>	2.184 (7,2)	28.038 (92,0)	241 (0,8)
<i>Bebidas</i>	746 (10,5)	6.311 (88,5)	78 (1,1)
<i>Doces e Sobremesas</i>	2.953 (6,3)	43.535 (92,8)	400 (0,9)
<i>Lanches</i>	2.161 (6,4)	31.296 (92,9)	245 (0,7)
<i>Prato Único</i>	1.014 (6,7)	13.944 (92,5)	114 (0,8)
Total	16.446 (7,0)	217.878 (92,2)	1.865 (0,8)

sexo feminino. Os nomes dessas listas também foram submetidos ao tratamento utilizado nos nomes dos usuários.

Em seguida, a lista de usuários já tratada foi ordenada alfabeticamente e comparada à lista de nomes do sexo masculino. Todos os nomes que estavam contidos nas duas listas foram incluídos numa terceira de “nomes masculinos identificados” e retirados da lista de usuários. Os nomes para o sexo feminino foram submetidos ao mesmo processo e gerou a lista “nomes femininos identificados”.

O arquivo com os nomes de usuários, mesmo após submetido à identificação dos sexos, continuou com alguns. Constatou-se eram apelidos utilizados pelos usuários. A identificação dos apelidos foi realizada também de forma manual.

Alguns nomes não foram identificados, uma vez que normalmente eram compostos por símbolos, o que não torna possível a identificação dos mesmos. Porém, foram apenas 0,8% do total. Analisando os gêneros por categoria foi notável a supremacia no número de usuários do sexo feminino (92,2%). como podemos ver na tabela 4.4.

Cabe ressaltar que proporcionalmente os homens se destacaram na categoria “Bebidas”, sendo a única na qual foram maior que 10%. As mulheres obtiveram seu maior percentual na categoria “Bolos e Tortas”.

4.5 Relações entre ingredientes

Analizamos quais os ingredientes mais co-ocorem em receitas, essa análise foi feita em todas as categorias. Foi analisado grupos de ingredientes de acordo com a média de ingredientes nas receitas da categoria e de acordo com a média geral 7. Foram obtidos 6 resultados, 3 de cada estudo e segue relatado o grupo que melhor representa a categoria.

Outra análise realizada foi em relação aos ingredientes que mais co-ocorreram em receitas. Sendo exploradas todas as categorias, foram observados grupos de ingredientes de acordo com a média destes nas receitas de uma categoria (obtendo 3 resultados) e de acordo com a média geral 7 (3 resultados). Segue relatado o grupo que melhor representa cada categoria.

- **Sopas:** Cebola, Sal, Alho, Batata, Água, Cenoura, Tomate;
- **Bebidas:** os resultados obtidos mostraram claramente dois tipos de grupos que representam a categoria e isso se dá pela divisão natural entre Bebidas Alcoólicas e Bebidas Não Alcoólicas.
 - **Bebidas Alcoólicas:** Açúcar, Água, Canela em pau, Gengibre, Licor\Vodca;
 - **Bebidas Não Alcoólicas:** Achocolatado, Açúcar, Canela em pau, Café, Leite em pó.
- **Peixes e Frutos do Mar:** Azeite, Cebola, Sal, Tomate, Alho, Leite de Côco, Pimentão, Camarão, Coentro;
- **Aves:** Cebola, Sal, Alho, Tomate, Leite, Milho verde, Óleo;
- **Carnes:** Azeitona, Carne bovina moída, Cebola, Muçarela, Ovo, Presunto, Sal;
- **Prato Único:** Alho, Cebola, Farinha de Trigo, Leite, Margarina, Óleo, Ovo, Queijo, Sal, Tomate;
- **Massas:** Açúcar, Farinha de Trigo, Fermento em pó, Leite, Margarina, Óleo, Ovo, Queijo, Sal;
- **Lanches:** Cebola, Farinha de Trigo, Fermento em pó, Leite, Óleo, Ovo, Sal, Tomate;
- **Saladas, Molhos e Acompanhamentos:** Azeite, Berinjela, Cebola, Orégano, Pimentão, Sal;

- **Bolos e Tortas:** Achocolatado, Açúcar, Farinha de Trigo, Fermento em pó, Leite, Leite Condensado, Margarina, Ovo;
- **Doces e Sobremesas:** Açúcar, Clara, Gema, Leite, Leite Condensado, Ovo.

Capítulo 5

Rede de ingredientes

Para análise dos ingredientes que compõem as receitas de nossa base, utilizamos uma rede. Rede essa que pode ser representada por um grafo $\mathcal{G}(\mathcal{V}, \mathcal{E})$, sendo \mathcal{V} o conjunto de vértices e \mathcal{E} o conjunto de arestas.

O grafo não direcionado \mathcal{G}_i modela a rede de ingredientes. Nesta rede os vértices são os ingredientes das receitas analisadas e as arestas se caracterizam quando dois ingredientes estão presentes numa mesma receita e o peso é o número de vezes que essa aresta ocorreu na base. Sem perda de generalidade, uma aresta $v \rightarrow u$ ocorre se, e somente se, os ingredientes v e u são utilizados em uma receita r_k . Assim, todos os ingredientes utilizados em uma receita r_k , estarão ligados entre si.

A rede pode evidenciar os ingredientes mais utilizados nas receitas da culinária brasileira, bem como os que mais combinam, analisando a sua co-ocorrência em receitas.

5.1 Análise

Esta seção explicita a análise da rede de ingredientes formada a partir da coleta das receitas do site Tudo Gosotoso.

A rede representada pelo grafo \mathcal{G}_i , é composta por 1.397 vértices, número total de ingredientes distintos utilizados nas receitas. Foram encontradas 88.837 arestas.

Utilizando algumas métricas para análises de grafo, podemos observar que nossa rede tem o diâmetro $\langle d \rangle = 3$, sendo então essa a maior distância de um ingrediente

a qualquer outro nessa rede. O valor do menor caminho médio obtido foi de 2,1. Tais métricas indicam que esta é rede bem conectada.

Quando observamos o grau do vértices k , podemos identificar como grau médio $\langle k \rangle$ de nossa rede é 95,54. O grau pode ser um fator importante para determinar afinidade entre os ingredientes, uma vez que a partir de k podemos dizer que em média um ingrediente harmoniza com ≈ 96 outros. De maneira mais ampla identifica-se também que cerca 80% dos ingredientes estão ligados a no máximo 150 outros ingredientes. O ingrediente de maior grau observado foi o “sal” com $\langle k \rangle = 1096$.

Outra métrica importante para análise de grafos calculada, foi o coeficiente de clusterização da rede, e o valor obtido foi 0.867, o que indica uma alta densidade da rede, fato também evidenciado pelo diâmetro e pelo valor do menor caminho médio.

5.1.1 Ingredientes mais importantes

Buscamos identificar os 10 ingredientes mais importantes da rede 3 métricas, incidência nas receitas, grau do vértice e betweenness, descritos na tabela 5.1.

Tabela 5.1: Ranking de ingredientes, com os respectivos valores para cada métrica considerada

Ranking	Incidência	Grau	Betweenness
1	Leite (74.056)	Sal (1069)	Sal (82.749)
2	Sal (62.028)	Leite (956)	Leite (72.075)
3	Ovo (54.282)	Cebola (872)	Açúcar (50.047)
4	Açúcar (47.899)	Água (868)	Água (40.895)
5	Cebola (38.625)	Açúcar (844)	Cebola (38.197)
6	Far. Trigo (37.943)	Ovo (844)	Ovo (30.997)
7	Leite Cond. (31.889)	Alho (832)	Alho (30.927)
8	Ferm. Pó (30.915)	Far. Trigo (788)	Azeite (23.965)
9	Óleo (29.066)	Óleo (778)	Far. Trigo (23.033)
10	Tomate (25.948)	Azeite (766)	Tomate (22.826)

Conforme esperado, os ingredientes com maiores valores de centralidade obtidos a partir do cálculo do betweenness dos vértices, são os ingredientes básicos, como sal, leite,

água, ovo, que estão presentes tanto em receitas salgadas como em receitas de doces.

Outra observação a ser feita é que a importância desses ingredientes, se os mesmos fossem retirados da rede, que até então possuía um único componente, a rede passaria a ser formada por 5 componentes, o que deixa claro a função de ligação desses. Vale ainda destacar o “leite” como o ingrediente com maior poder de ligação se observarmos a geração de novos componentes a partir da retirada de um ingrediente, uma vez que apenas com a sua remoção da rede, essa passaria ter 3 componentes.

5.1.2 Arestas mais importantes

Uma análise também foi feita observando as principais arestas da rede de ingredientes. Essa aconteceu observando o peso das arestas, na identificação os pares de ingredientes que mais ocorrem juntos em receitas. A tabela 5.2 demonstra as arestas mais frequentes da rede.

Tabela 5.2: Ranking das arestas

Ranking	Aresta	Peso
1	Ovo ↔ Leite	31.150
2	Açúcar ↔ Leite	29.287
3	Açúcar ↔ Ovo	28.980
4	Sal ↔ Leite	25.725
5	Far. Trigo ↔ Ovo	24.043
6	Sal ↔ Cebola	23.594
7	Leite Condensado ↔ Leite	22.986
8	Fermento ↔ Ovo	22.870
9	Far. Trigo ↔ Leite	22.845
10	Sal ↔ Ovo	22.561

Como pode se observar as principais arestas são formadas pelos ingredientes mais importantes da rede, antes identificados.

A representação gráfica da rede foi desenvolvida pelo software Gephi¹, para facilitar a visualização do grafo e suas conexões como podemos observar na Figura 5.1. O tamanho

¹gephi.org/

Capítulo 6

Personalização de receitas

Em nossa base existem receitas muito similares, muitas vezes que se diferenciam por apenas um ingrediente. Um dos fatores que podem favorecer esse fato são preferências pessoais de quem executa a receita. Logo os ingredientes que se diferenciam nessas receitas são muitas vezes ligados diretamente ao sabor, por exemplo, o recheio de uma massa, onde a maioria dos ingredientes estão na composição da massa e o fator que torna uma diferente de outra é o recheio. Esses poucos ingredientes diferentes podem ter sido fruto de uma adição, substituição ou até mesmo da remoção de 1 ou mais ingredientes da receita original.

Como já fora observado em outros trabalhos, os usuários de um sistema colaborativo de receitas culinárias tendem a sugerir esse tipo de modificação por meio de comentários Teng et al. (2012). Logo resolvemos investigar se isso também é um costume de usuários brasileiros.

Nossa base possui um total de 319.219 comentários. Para que a análise fosse realizada num primeiro momento deveriam ser identificados termos que caracterizam sugestões de modificações. Buscando a identificação desses termos utilizamos uma amostra da base composta de 1.000 comentários que foram extraídos de 100 receitas aleatórias entre as receitas que possuem mais de 10 comentários, onde eram extraídos os últimos 10 comentários dessas.

A análise da amostra se deu de forma manual, para identificar todos os tipos de comentários realizados nas receitas. Foram encontrados 11 tipos diferentes de comentários segundo a classificação realizada, como podemos observar na Tabela 6.1.

Tabela 6.1: Tipos de comentários encontrados na amostra.

Tipo de Comentário	Quantidade
<i>Elogio</i>	721
<i>Crítica</i>	92
<i>Sug. Adição</i>	90
<i>Sug. Remoção</i>	29
<i>Sug. Substituição</i>	24
<i>Dúvida</i>	122
<i>Resposta</i>	71
<i>Dica de modo de preparo</i>	42
<i>Modificação de medidas</i>	57
<i>Sug. de data para execução</i>	23
<i>Ainda vai preparar</i>	38

Em alguns comentários era possível observar mais de um tipo comentário, ou seja, um usuário sugeria a adição de ingrediente e também elogiava a receita, exemplo, “Muito boa, adicionei leite de condensado e ficou ainda melhor”.

A Figura 6.1 apresenta uma nuvem de palavras com os mais frequente termos utilizados nos comentários.



Figura 6.1: Nuvem de palavras mais utilizadas nos comentários.

6.1 Identificação de sugestões de adição e remoção de ingredientes

Após esse processo manual e levantamento de termos o próximo passo foi automatizar o processo de identificação de conteúdo nos comentários. Para identificação dos comentários que sugerem remoções ou adições a estratégia adota é similar uma vez que esse tipo de sugestão seguem o padrão “termo+ingrediente”, exemplos:

- adicionei(termo) sal(ingrediente);
- tirei(termo) o tomate(ingrediente);
- retirei(termo) o sal, tomate e batata palha(ingredientes);

Outra observação importante é que logo após encontrar um ingrediente o algoritmo seguia a procura de outros, uma vez que como terceiro exemplo acima, num mesmo comentário foi sugerido a retirada de mais de um ingrediente.

A expressão regular abaixo foi elaborada para demonstrar a identificação das sugestões de adição de ingredientes, sendo X os ingredientes:

“acrescentei—acrescento—acrescente—acrescetar—acrescentamos—acrescentei —adicionei X1,X2,X3...”

Para exemplificar a identificação das sugestões de remoção de ingredientes a seguinte expressão regular foi desenvolvida:

“nao coloquei—nao acrescentei—nem colocar—nao uso—nao coloco—nao pode colocar—n coloquei—tirar—nao coloque—nao usei—n colokei X1,X2,X3...”

6.2 Identificação de sugestões de substituição de ingredientes

Para identificação dos comentários com sugestões de substituição de ingredientes observou-se duas formas diferentes. Uma que segue o padrão “termo + ingrediente + termo + ingrediente” e outra que segue “termo+ingrediente+ingrediente”, como podemos observar nos exemplos abaixo:

- coloquei(termo) tempero(ingrediente) no lugar(termo)do sal(ingrediente)
- troquei(termo) sal(ingrediente) por tempero(ingrediente);

Vale ressaltar que foram criadas para esse processo duas listas de termos uma vez que a primeira citada utiliza inicialmente um termo próximo aos que indicam a adição de ingredientes. Outro ponto que merece destaque é que a substituição é realizada de forma direcionada. como foi visto no exemplo acima uma mesma substituição foi indicada seguindo os dois padrões observados e fica claro a inversão da ordem dos ingredientes, no primeiro caso o ingrediente que vai entrar na receita aparece primeiro do que será retirado e o segundo ocorre o inverso.

Os casos citados podem ser exemplificados por meio das expressões regulares abaixo, sendo X o ingrediente que irá entrar na receita e Y o que deve ser retirado:

“substitui—em vez—troquei—substituida Y por X”

“coloquei—adicionei X no lugar—ao invés de Y”

Uma ressalva deve ser feita que quando num mesmo comentário foi detectado a sugestão de adição de um ingrediente e a sugestão de remoção de um ingrediente, esse cenário também caracteriza uma substituição, seguindo o padrão “termoAdc + ingrediente + termoRem + ingrediente”:

- adicionei(termoAdc) tempero(ingrediente) e retirei(termoRem) sal(ingrediente)

E que por meio de expressão regular pode ser demonstrado como:

“acrescentei—acrescento—acrescente—acrescetar—acrescentamos—acrecentei —adicionei Y e nao coloquei—nao acrescentei—nem colocar—nao uso—nao coloco—nao pode colocar—n coloquei—tirar—nao coloque—nao usei—n colokei X.

Esse último caso acontece quando num mesmo comentário ocorre a sugestão de adição e remoção de ingredientes.

6.2.1 Adições e Remoções

Após o processamento de todos os comentários de nossa base foram encontradas 22.800 sugestões de adição envolvendo 461 ingredientes distintos (33% dos ingredientes) e

4.885 sugestões de remoção contemplando 311 ingredientes diferentes (22% dos ingredientes). Pode-se observar que normalmente as sugestões envolvem ingredientes de grande incidência na base.

A Tabela 6.2 mostra os 10 ingredientes que apareceram com maior frequência nas sugestões de cada uma das categorias de adição e remoção. Nessa tabela, os valores entre parênteses correspondem às frequências de ocorrência dos ingredientes.

Tabela 6.2: 10 ingredientes mais frequentes nas sugestões de adição e remoção.

Ranking	Sugestão de Adição	Sugestão de Remoção
<i>1</i>	Leite (1.733)	Leite (396)
<i>2</i>	Creme de Leite (922)	Creme de Leite (216)
<i>3</i>	Queijo (884)	Açúcar (213)
<i>4</i>	Coco (781)	Água (153)
<i>5</i>	Molho de Tomate (549)	Queijo (137)
<i>6</i>	Açúcar (518)	Sal (129)
<i>7</i>	Far. Trigo (505)	Coco (123)
<i>8</i>	Milho Verde (480)	Cobertura de Sorvete (113)
<i>9</i>	Chocolate (453)	Ovo (111)
<i>10</i>	Cebola (429)	Chocolate (89)

6.2.2 Substituições

Além das adições e remoções, o processamento dos textos de comentários nos permitiu identificar 5.511 substituições, das quais 2.319 são distintas e que utilizaram 632 ingredientes diferentes (45%). A Tabela 6.3 apresenta os 10 pares de ingredientes que apareceram com maior frequência nas sugestões de substituição. Nessa tabela, os valores entre parênteses correspondem às frequências de ocorrência dos pares de ingredientes.

Buscando expandir a análise das sugestões dos usuários foi desenvolvido um grafo a partir dessas. Um grafo direcionado onde cada nó representa um ingrediente e uma aresta indica a sugestão de substituição do ingrediente origem pelo ingrediente destino, ou seja, uma aresta do ingrediente “açúcar” para o “adoçante” indica a sugere tirar o açúcar e incluir adoçante.

Tabela 6.3: 10 ingredientes mais frequentes nas sugestões de substituição.

Ranking	Sugestão de Substituição
1	Água - <i>ç</i> Leite (142)
2	Leite - <i>ç</i> Leite de Coco (69)
3	Óleo - <i>ç</i> Azeite (66)
4	Leite - <i>ç</i> Água (65)
5	Açúcar - <i>ç</i> Leite Condensado (55)
6	Margarina - <i>ç</i> Manteiga(52)
7	Óleo - <i>ç</i> Margarina (46)
8	Açúcar - <i>ç</i> Adoçante (37)
9	Margarina - <i>ç</i> Óleo (37)
10	Leite - <i>ç</i> Creme de Leite (31)

A Figura 6.2 demonstra a rede de substituições gerada. Alguns recursos visuais facilitam o entendimento da mesma como por exemplo, o tamanho de cada nó está associado com o seu grau, ou seja, os nós maiores indicam que aquele ingrediente está vinculado a uma maior quantidade de outros ingredientes nas sugestões de substituição. Desse modo, é possível observarmos, por exemplo, que o creme de leite é o ingrediente mais flexível com relação às substituições, dado que ele aparece vinculado a um grande número de outros ingredientes.

Essa caracterização dos comentários aliada às análises das sugestões de adição, remoção e substituição de ingredientes são etapas fundamentais para a proposição de sistemas de recomendação que tenham como meta tornar as receitas mais saudáveis e/ou adequadas ao gosto de um determinado usuário.

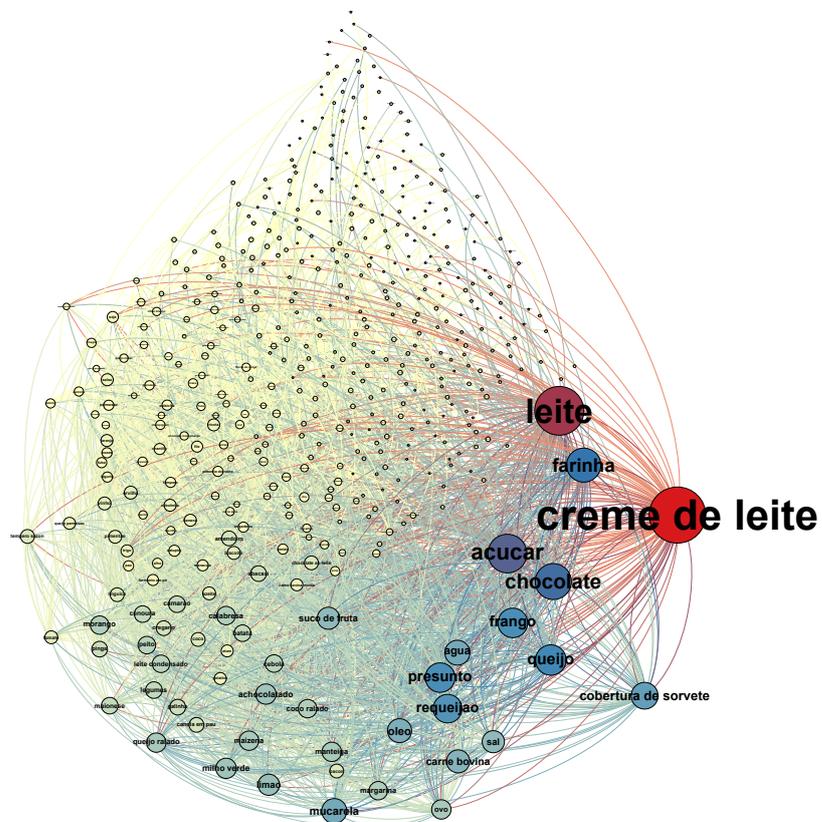


Figura 6.2: Rede de substituições.

Capítulo 7

Conclusão e Trabalhos Futuros

Apresentamos neste trabalho uma extensa caracterização de um importante site brasileiro de compartilhamento de receitas: o Tudo Gostoso. Em relação aos usuários, a análise revela que um número elevado de usuários compartilham suas opiniões (48,4%). Estes comentários englobam desde sugestões para a época do ano mais adequada para preparo da receita, sugestões que tornam, por exemplo, as receitas mais saudáveis e dúvidas referentes ao preparo das receitas. Adicionalmente, a maioria dos comentários realizados são feitos por mulheres, uma vez que também este é o sexo predominante entre os usuários.

Em relação as receitas, a categoria com maior número de receitas é a de Doces e Sobremesas, o que pode revelar uma tendência de gosto dos usuários do site. Interessante é que a categoria de Sopas possui somente 1,92% das receitas compartilhadas, fato este que pode estar ligado ao fato do Brasil como país tropical, ter invernos amenos. Receitas possuem, em média, 7,4 ingredientes distintos.

A rede de ingredientes baseadas nas receitas coletadas fornece indícios de padrões de uso de ingredientes encontrados na culinária brasileira. As receitas compartilhadas no site são compostas por 1.397 ingredientes distintos, sendo que estes ingredientes formam uma rede extremamente densa, sugerindo que, apesar do número grande de receitas no site, os ingredientes não variam muito de receita para receita. Além disso, cada ingrediente, em média, provavelmente harmoniza com outros ≈ 96 ingredientes.

Além da caracterização, foi desenvolvido um processo automático para identificação dos casos de adição, remoção e substituição de ingredientes em receitas. A análise realizada, juntamente com o processo automatizado proposto, são essenciais para o

desenvolvimento de sistemas de recomendação que poderão ser utilizados para tornar receitas mais saudáveis, ou mais próximas ao gosto de um determinado usuário.

Como trabalhos futuros, além do desenvolvimento de sistemas de recomendação para personalização de receitas, pretende-se formalizar e analisar a rede de sabores presentes nos ingredientes utilizados nas culinária brasileira, estendendo os resultados apresentados em Ahn et al. (2011) porém voltados a culinária local.

Outro trabalho a ser realizado é o clusterização dos ingredientes da base de dados, para que novas análises sejam feitas não observando categorias, mas sim grupos de ingredientes que são partes de um mesmo cluster.

Referências Bibliográficas

- Abreu, N. R. d., Baldanza, R. F. and Sette, R. d. S. (2008). Comunidades virtuais como ambiente potencializador de estratégias mercadológicas: lócus de informações e troca de experiências vivenciadas, *Perspectivas em Ciência da Informação* pp. 116–136.
- Ahn, Y.-Y., Ahnert, S. E., Bagrow, J. P. and Barabási, A.-L. (2011). Flavor network and the principles of food pairing, *Nature, Scientific Reports* (196).
- Alves, M. V. B., Dória, C. A. and de Souza, G. G. (2008). *Cozinheiro nacional, ou, Coleção das melhores receitas das cozinhas brasileira e européias; para a preparação de sopas, molhos, carnes, caça, peixes, crustáceos, ovos, leite, legumes, pudins, pastéis, doces de massa e conservas para sombra; acompanhado das Regras de servir à mesa e de trincar; ornado com numerosas estampas finas*, Senac.
- Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks, **286**(5439): 509–512.
- Benevenuto, F., Almeida, J. and Silva, A. (2011). Coleta e análise de grandes bases de dados de redes sociais online, *Jornadas de Atualização em Informática (JAI)*, pp. 11–57.
- Braitenberg, V. and Schüz, A. (1998). *Cortex: statistics and geometry of neuronal connectivity*, Vol. 249, Springer Berlin.
- Bruno Leite Alves, Alberto Henrique Frade Laender, F. B. (2013). *Um estudo sobre a evolução temporal de comunidades científicas*, Mestrado.
- Bruns, A. (2008). *Blogs, Wikipedia, Second Life, and beyond: From production to produsage*, Vol. 45, Peter Lang.
- Castells, M. (2003). The power of identity: The information age: Economy, society and culture, volume ii (the information age).
- de São Paulo, F. (2010). Facebook mostra o raio-x de 1 bilhão de usuários, <http://www1.folha.uol.com.br/tec/1163808-facebook-mostra-o-raio-x-de-1-bilhao-de-usuarios.shtml>.
- URL:** <http://www1.folha.uol.com.br/tec/1163808-facebook-mostra-o-raio-x-de-1-bilhao-de-usuarios.shtml>

- Faloutsos, M., Faloutsos, P. and Faloutsos, C. (1999). On power-law relationships of the internet topology, *ACM SIGCOMM Computer Communication Review*, Vol. 29, pp. 251–262.
- Figueiredo, D. R. (2011). Introdução a redes complexas, *Jornada de Atualizações em Informática* pp. 303–358.
- Freyne, J. and Berkovsky, S. (2010a). Intelligent food planning: Personalized recipe recommendation, *Proceedings of the 15th International Conference on Intelligent User Interfaces*, ACM, New York, NY, USA, pp. 321–324.
- Freyne, J. and Berkovsky, S. (2010b). Recommending food: Reasoning on recipes and ingredients, *User Modeling, Adaptation, and Personalization*, Lecture Notes in Computer Science, Springer Berlin Heidelberg, pp. 381–386.
- Geleijnse, G., Nachtigall, P., van Kaam, P. and Wijgergangs, L. (2011). A personalized recipe advice system to promote healthful choices, *Proceedings of the 16th international conference on Intelligent user interfaces*, pp. 437–438.
- Hammond, K. (1986). CHEF: A Model of Case-Based Planning, *Proceedings of AAAI-86*, pp. 267–271.
- Hamon, T. and Grabar, N. (2013). Extraction of ingredient names from recipes by combining linguistic annotations and crf selection, *Proceedings of the 5th international workshop on Multimedia for cooking & eating activities*, pp. 63–68.
- Ihle, N., Hanft, A. and Althoff, K.-D. (2009). Extraction of adaptation knowledge from internet communities, *ICCBR 2009 Workshop Proc., Workshop Reasoning from Experiences on the Web*, pp. 269–278.
- Kuo, F.-F., Li, C.-T., Shan, M.-K. and Lee, S.-Y. (2012). Intelligent menu planning: Recommending set of recipes by ingredients, *Proceedings of the ACM Multimedia 2012 Workshop on Multimedia for Cooking and Eating Activities*, ACM, New York, NY, USA, pp. 1–6.
- Lole, E., Oliveira, M. d. and Wiggers, S. (2002). Receitas culinárias, segredos de família.
- Maarten Peter Bodlaender, Arvid Randal Nicolaas, O. S. M. S.-N. A. G. (2010). System and method for recommending recipes.
- Maruyama, T., Kawano, Y. and Yanai, K. (2012). Real-time mobile recipe recommendation system using food ingredient recognition, *Proceedings of the 2Nd ACM International Workshop on Interactive Multimedia on Mobile and Portable Devices*, ACM, New York, NY, USA, pp. 27–34.
- Mino, Y. and Kobayashi, I. (2009). Recipe recommendation for a diet considering a user’s schedule and the balance of nourishment, *Intelligent Computing and Intelligent Systems, 2009. ICIS 2009. IEEE International Conference on*, pp. 383–387.

- Newman, M. E. J. (2003). The structure and function of complex networks, *SIAM Review* **45**(2): 167–256.
- Svensson, M., Höök, K. and Cöster, R. (2005). Designing and evaluating kalas: A social navigation system for food recipes, *ACM Trans. Comput.-Hum. Interact.* pp. 374–400.
- Teixeira, E. D. (2009). Especificidades da tradução técnica de receitas—alguns problemas e possíveis soluções.
- Teng, C.-Y., Lin, Y.-R. and Adamic, L. A. (2012). Recipe recommendation using ingredient networks, *Proceedings of the 3rd Annual ACM Web Science Conference, WebSci '12*, pp. 298–307.
- Times, T. N. W. (2009). Uploading the avant-garde, <http://www.nytimes.com/2009/09/06/magazine/06FOB-medium-t.html>.
URL: <http://www.nytimes.com/2009/09/06/magazine/06FOB-medium-t.html>
- Ueda, M., Takahata, M. and Nakajima, S. (2011). User’s food preference extraction for cooking recipe recommendation, *SPIM*, pp. 98–105.
- van Pinxteren, Y., Geleijnse, G. and Kamsteeg, P. (2011). Deriving a recipe similarity measure for recommending healthful meals, *Proceedings of the 16th International Conference on Intelligent User Interfaces*, ACM, New York, NY, USA, pp. 105–114.
- Yajima, A. and Kobayashi, I. (2009). ”easy” cooking recipe recommendation considering user’s conditions, *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology - Volume 03*, Washington, DC, USA, pp. 13–16.
- Zhang, Q., Hu, R., Namee, B. M. and Delany, S. J. (2008). Back to the future: Knowledge light case base cookery., *ECCBR Workshops*, pp. 239–248.
- Zhu, Y.-X., Huang, J., Zhang, Z.-K., Zhang, Q.-M., Zhou, T. and Ahn, Y.-Y. (2013). Geography and similarity of regional cuisines in China, *ArXiv e-prints* .

Identificação de ingredientes

Seguem abaixo os termos que foram modificados na identificação de ingredientes, foram corrigidos erros de ortografia, os ingrediente foram padronizados a lingua portuguesa, também se padronizou ingredientes de acordo com regionalidade. O termo final demonstra como o ingrediente foi armazenado na base de dados e os termos todas as variações de entrada conhecida.

Termo final: **cachaca 51**

Termos: — **51** —

Termo final: **acerola**

Termos: — **acelora** —

Termo final: **achocolatado**

Termos: — **achocolatado em po** — **achocolato em po** — **achocolotado em po**
— **acholatado em po** — **po de chocolate** —

Termo final: **suco ades goiaba**

Termos: — **addes de goiaba** —

Termo final: **suco ades**

Termos: — **ades** —

Termo final: **suco ades pera**

Termos: — **ades pera** —

Termo final: **vinho almadem**

Termos: — **almadem** —

Termo final: **licor amaretto**

Termos: — **amaretto** —

Termo final: **licor amarula**

Termos: — **amarula** —

Termo final: **ameixa preta**

Termos: — **ameixas pretas** —

Termo final: **ameixa seca**

Termos: — **ameixas secas** —

Termo final: **amendoim**

Termos: — **amendoin** —

Termo final: **ovo**

Termos: — **vos** —

Termo final: **acucar**

Termos: — **a;ucar** —

Termo final: **creme de amendoim**

Termos: — **amendocrem** —

Termo final: **bitter angostura**

Termos: — **angostura** — **angostura** —

Termo final: **erva doce**

Termos: — **anis** — **aniz** — **estrelas de anis** —

Termo final: **agua**

Termos: — **aqua** —

Termo final: **rum bacardi**

Termos: — **bacardi** —

Termo final: **uisque ballantines**

Termos: — **ballantine** — **ballantines** —

Termo final: **batata frita**

Termos: — **batatas fritas** —

Termo final: **bicarbonato de sodio**

Termos: — **bicabornato** — **bicabornato de sodio** — **bicarbonato** — **bircarbonato de sodio** — **bicarbonato de sodio** — **bicarbonato de sodio** — **bicarbonato de sodio** — **bicarbonate** — **bircarbonato** — **bircarbonato em po** — **bircarbonato de sodio** —

Termo final: **biscoito oreo**

Termos: — **bolachas de oreo** —

Termo final: **bolacha**

Termos: — **bolacha doce** —

Termo final: **chocolate**

Termos: — **bombons** — **bombom** —

Termo final: **chocolate meio amargo**

Termos: — **bombom meio amargo** —

Termo final: **uisque bourbon**

Termos: — **bourbon** —

Termo final: **chocolate branco**

Termos: — **bombons ouro branco** — **ouro branco** —

Termo final: **conhaque brandy**

Termos: — **brandy** —

Termo final: **brigadeiro**

Termos: — **brigadeiro mococa** —

Termo final: **cacau em po**

Termos: — **cacal em po** —

Termo final: **cachaca**

Termos: — **cahaca** —

Termo final: **canela em pau**

Termos: — **canela de pau** — **canela em casca** — **canela em rama** — **pau de canela** —

Termo final: **caninha**

Termos: — **caninha pitu** —

Termo final: **cappuccino**

Termos: — **capuccino** — **capucinno** — **puccino** —

Termo final: **cappuccino em po**

Termos: — **capuccino em po** —

Termo final: **cara**

Termos: — **cara.** —

Termo final: **cha**

Termos: — **cha.** — **sache de cha** —

Termo final: **cerveja preta**

Termos: — **caracu** — **cerveja escura** —

Termo final: **gengibre**

Termos: — **cardamomo** —

Termo final: **iogurte chambinho**

Termos: — **chambinho** —

Termo final: **licor chambord**

Termos: — **chambord** —

Termo final: **champanhe**

Termos: — **champagne** — **champagne** —

Termo final: **chantilly**

Termos: — **chantilly** — **chantilly** — **chantilly mococa** — **chantilly** — **chantineve** — **chanty** — **chantilly** —

Termo final: **licor chartreuse**

Termos: — **chartreuse** —

Termo final: **achocolatado choco milk**

Termos: — **choco milk** —

Termo final: **chocolate em po**

Termos: — **chocolate do padre** — **chocolate ralado** —

Termo final: **vermute cinzano**

Termos: — **cinzano** —

Termo final: **refrigerante coca cola**

Termos: — **coca** — **refrigerante de cola** — **coca cola** —

Termo final: **coco ralado**

Termos: — **coco picado** —

Termo final: **licor cointreau**

Termos: — **cointreau** —

Termo final: **leite condensado**

Termos: — **condensado** —

Termo final: **vermute contini**

Termos: — **contine** — **contini** —

Termo final: **licor contreau**

Termos: — **contreau** — **lic contreau** —

Termo final: **suco cramberry juice**

Termos: — **cramberry juice** —

Termo final: **licor contreau**

Termos: — **cravo-da-india** — **cravo** —

Termo final: **creme de cacau**

Termos: — **creme de cacau stock** —

Termo final: **creme de leite**

Termos: — **creme de leita** —

Termo final: **chocolate em bolas**

Termos: — **choco balls** — **crocante** — **m&m** —

Termo final: **licor curacau**

Termos: — **curacao** — **curacau** — **curacau azul** — **curacau blue** —

Termo final: **curau**

Termos: — **curau pronto** —

Termo final: **curau**

Termos: — **danone** — **danoninho** —

Termo final: **chocolate diamante negro**

Termos: — **diamante negro** —

Termo final: **goiabada**

Termos: — **doce de goiaba** —

Termo final: **conhaque domus**

Termos: — **domus** —

Termo final: **licor drambuie**

Termos: — **drambuie** —

Termo final: **suco artificial**

Termos: — **envelope de suco** — **suco art** —

Termo final: **erva doce**

Termos: — **erva - cidreira** — **erva-cidreira** —

Termo final: **erva doce**

Termos: — **erva-doce** —

Termo final: **vinho espumante**

Termos: — **espumante** —

Termo final: **essencia de pequi**

Termos: — **essencia da pequi** — **essencia de pequi** —

Termo final: **essencia de amarula**

Termos: — **essencia de marula** —

Termo final: **essencia**

Termos: — **essencia.** —

Termo final: **refrigerante fanta**

Termos: — **fanta** —

Termo final: **farinha lactea**

Termos: — **farinha lacta** —

Termo final: **feijao de soja**

Termos: — **feijao soja** —

Termo final: **fermento em po**

Termos: — **fermento biologico** — **fermento para pao** —

Termo final: **folha de laranjeira**

Termos: — **folhas de laranjeira** —

Termo final: **folha de laranjeira**

Termos: — **folhas de laranjeira** — **folhas laranjeiras** —

Termo final: **flocos de cereal**

Termos: — **flocos de cereais** —

Termo final: **fruta**

Termos: — **fruta.** —

Termo final: **gengibre**

Termos: — **gengibres** —

Termo final: **genipapo**

Termos: — **genipapos** —

Termo final: **germem de trigo**

Termos: — **germen de trigo** —

Termo final: **gim**

Termos: — **gyn** — **gin** —

Termo final: **glicose**

Termos: — **glucose** —

Termo final: **licor grenadine**

Termos: — **granadina** — **granadine** — **grenadine** —

Termo final: **licor grand marnier**

Termos: — **grand marnier** —

Termo final: **groselha**

Termos: — **grosselha** —

Termo final: **refrigerante**

Termos: — **guarana** — **gurarana** —

Termo final: **chocolate**

Termos: — **guarda-chuvinha** —

Termo final: **halls**

Termos: — **”halls”** —

Termo final: **hortela**

Termos: — **hotela** —

Termo final: **iogurte**

Termos: — **iogute** — **iorgut** — **iorgute desnatado** —

Termo final: **iogurte natural**

Termos: — **iogute natural** — **iorgurte natural** — **iorgute natural** —

Termo final: **hortela**

Termos: — **hotela** —

Termo final: **jabuticaba**

Termos: — **jaboticaba** —

Termo final: **uisque jack daniels**

Termos: — **jack** — **”jack daniels”** —

Termo final: **tequila jose cuervo**

Termos: — **jose cuervo** —

Termo final: **achocolatado joycolate**

Termos: — **joycolate** —

Termo final: **licor kahlua**

Termos: — **kahlua** —

Termo final: **cerveja kaiser bock**

Termos: — **kaiser bock** —

Termo final: **quiui**

Termos: — **kiwi** — **kiwwi** —

Termo final: **kirsch**

Termos: — **lata de kirsh** —

Termo final: **leite condensado**

Termos: — **leite moca** —

Termo final: **leite em po**

Termos: — **leite molico** — **leite ninho** — **leite, em po** —

Termo final: **leite**

Termos: — **letie** —

Termo final: **licor stock**

Termos: — **lic cremoso stock** —

Termo final: **licor de anis**

Termos: — **lic de aniz** —

Termo final: **licor de cacau**

Termos: — **lic de caucau** —

Termo final: **licor amarula**

Termos: — **lic marula** —

Termo final: **limao**

Termos: — **limoes** —

Termo final: **limao taiti**

Termos: — **limoes tahiti** — **limoes taiti** —

Termo final: **licor limoncello**

Termos: — **limoncello** —

Termo final: **maisena**

Termos: — **maizena** —

Termo final: **manjericao**

Termos: — **majericao** — **mangericao** —

Termo final: **licor malibu**

Termos: — **malibu** —

Termo final: **cerveja malzbier**

Termos: — **malzbier** —

Termo final: **mamao**

Termos: — **mamoes** —

Termo final: **licor marie brizard**

Termos: — **marie brizard sugar cane syrup** —

Termo final: **vermute martini**

Termos: — **martine rose** — **martini doce** — **vermouth martini bianco** —

Termo final: **cha mate**

Termos: — **mate** — **mate leao** —

Termo final: **mexerica**

Termos: — **mexirica** — **tangerina** — **ponkan** —

Termo final: **suco artificial mid**

Termos: — **mid refresco** — **suco mid** —

Termo final: **milho verde**

Termos: — **milho** —

Termo final: **miranta em po**

Termos: — **mirantam em po** —

Termo final: **vodka cranberry**

Termos: — **mix cramberry** — **monin cranberry** —

Termo final: **folha de morangueiro**

Termos: — **morangueiro** —

Termo final: **morango**

Termos: — **moranguinho** —

Termo final: **nata**

Termos: — **nata.** —

Termo final: **achocolatado nescau**

Termos: — **nescau** —

Termo final: **achocolatado nesquik**

Termos: — **nesquik** —

Termo final: **cereal mucilon**

Termos: — **mucilon** —

Termo final: **cereal neston**

Termos: — **neston** —

Termo final: **noz moscada**

Termos: — **nos - moscada** —

Termo final: **nutella**

Termos: — **nutela** —

Termo final: **ovomaltine**

Termos: — **ovo maltine** —

Termo final: **oz**

Termos: — **oz sweet and sour mix** —

Termo final: **pacoca**

Termos: — **pacoquinha** —

Termo final: **cha**

Termos: — **pacote de cha** — **pacotinhos de cha** —

Termo final: **suco artificial**

Termos: — **pacotinho de suco** — **pacotes de suco** — **pacote de suco** —

Termo final: **vermute paizano**

Termos: — **paizano** —

Termo final: **mamao papaia**

Termos: — **papaia** —

Termo final: **uva passa**

Termos: — **passas** —

Termo final: **creme de amendoim**

Termos: — **pasta de amendoim** —

Termo final: **refrigerante pepsi twist**

Termos: — **pepsi twist** — **pespi twist** —

Termo final: **licor pernod**

Termos: — **pernod** —

Termo final: **champanhe perrier**

Termos: — **perrier** —

Termo final: **petala de rosa**

Termos: — **petalas de rosas** — **rosas vermelhas** —

Termo final: **chantilly em po**

Termos: — **po para chantilly** —

Termo final: **polpa de futa**

Termos: — **polpas de futa** — **poupa** —

Termo final: **raspa de chocolate**

Termos: — **raspas de chocolate** —

Termo final: **uisque red label**

Termos: — **red label** —

Termo final: **suco artificial**

Termos: — **refresco art** —

Termo final: **suco artificial cligth**

Termos: — **refresco cligth** —

Termo final: **suco artificial tang**

Termos: — **refresco tang** — **tang** — **suco tang** —

Termo final: **suco artificial tang**

Termos: — **refresco tang** —

Termo final: **suco de fruta**

Termos: — **refresco da fruta** —

Termo final: **rosquinha mabel**

Termos: — **rosquinhas mabel** —

Termo final: **rum**

Termos: — **run** —

Termo final: **cachaca sagatiba**

Termos: — **sagatiba** —

Termo final: **licor san remy**

Termos: — **san remy** — **st. remy** —

Termo final: **cachaca sapucaia florida cristal**

Termos: — **sapucaia florida cristal** —

Termo final: **cachaca sapucaia florida ouro**

Termos: — **sapucaia florida ouro** —

Termo final: **refrigerante schweppes**

Termos: — **schweppes** — **schweppes citrus** —

Termo final: **semente de cardamon**

Termos: — **sementes de cardamon** —

Termo final: **coquetel sex on the beach**

Termos: — **sex on the beach** —

Termo final: **vodka smirnoff**

Termos: — **smirnoff** — **smirnoff ice** —

Termo final: **leite de coco**

Termos: — **sococo** —

Termo final: **refrigerante soda**

Termos: — **soda** — **soda limonada** —

Termo final: **suco sollys laranja**

Termos: — **sollys laranja** —

Termo final: **chocolate sonho de valsa**

Termos: — **sonho de valsa** — **sonhos de valsa** —

Termo final: **sorvete de creme**

Termos: — **sorverte de creme** —

Termo final: **sorvete de flocos**
Termos: — **soverte de flocos** —

Termo final: **refrigerante sprite**
Termos: — **sprit** — **spryte** — **sprite** —

Termo final: **steinhaeger**
Termos: — **stanheguer** — **stainhager** —

Termo final: **suco**
Termos: — **suco.** —

Termo final: **chocolate suflair**
Termos: — **souflair** —

Termo final: **molho tabasco**
Termos: — **tabasco** —

Termo final: **tequila**
Termos: — **tekila** —

Termo final: **achocolatado toddy**
Termos: — **toddy** —

Termo final: **licor triple sec**
Termos: — **triple-sec** —

Termo final: **refrigerante tubaina**
Termos: — **tubaina** —

Termo final: **bitter underberg**
Termos: — **underberg** —

Termo final: **adocante em po**
Termos: — **uniao diet sucralose** —

Termo final: **baunilha**

Termos: — **vanilina** —

Termo final: **cachaca velho barreiro**

Termos: — **velho barreiro** —

Termo final: **vermute**

Termos: — **vermouth** — **vermouth dry** — **vermut** —

Termo final: **vodka**

Termos: — **vodca** — **vodcka** — **vokca** — **vokda** — **voodka** — **wodka** —

Termo final: **waffer**

Termos: — **wafer** —

Termo final: **uisque**

Termos: — **whisk** — **whisky** — **wisk** — **wisky** —

Termo final: **uisque scotch**

Termos: — **whisky scotch** —

Termo final: **uisque teachers**

Termos: — **whisky teachers** —

Termo final: **yakult**

Termos: — **yakut** —

Termo final: **suco concentrado de fruta**

Termos: — **suco concentrado da fruta** —

Termo final: **suco de cranberry**

Termos: — **suco cramberry juice** — **suco de cramberry** — **suco de cranberry**

—

Termo final: **suco de cereja**

Termos: — **suco das cerejas** —

Termo final: **suco de goiaba**

Termos: — **suco de caixa de goiaba** —

Termo final: **suco de grape fruit**

Termos: — **suco de grapefruit** —

Termo final: **suco de laranja**

Termos: — **suco de laranja** —

Termo final: **suco artificial**

Termos: — **suco de saquinho** — **suco do pacote** — **suco em po** —

Termo final: **suco artificial fresh**

Termos: — **suco de saquinho fresh** — **suco fresh** —

Termo final: **suco**

Termos: — **suco natural** —

Termo final: **suco de acerola**

Termos: — **suco natural de acerola** —

Termo final: **vodka absolut vanilla**

Termos: — **absolut vanilla** —

Termo final: **danete**

Termos: — **danette** —