

Quais ingredientes fazem o melhor sabor?

Adriano Zanei–11101613– Turma A2

Rafael Perez–11031513– Turma A2

Matheus Tognetti–11031013– Turma A2

Igor Bandim– 11064013– Turma A2

Henrique Morikawa–11057313– Turma A2

Giancarlo Maschietto–11016313– Turma A2

Jean Gabriel Giacon–11018013– Turma A2

Renan Degaspari– 11011113 – Turma A2

Fred Ventura–11104313–Turma A2

Universidade Federal do ABC-UFABC

Santo André, SP, Brasil

Abstract

Em nosso projeto final de Comunicação e Redes, o principal objetivo foi gerar uma rede real e, posteriormente, analisá-la segundo os conceitos aprendidos durante o curso. O tema escolhido pelo grupo foi a culinária, sendo que, nesse caso, a informação transmitida pelas arestas são as relações entre os ingredientes e os sabores formados por elas. Nesse trabalho, foram utilizados os programas Excel, Gephi e Anaconda, com a intenção de aplicar os conhecimentos de geração e manipulação de dados. Os resultados, análises e conclusões podem ser conferidos a seguir.

I-Introdução

O estudo de grafos teve sua primeira fundamentação teórica no ano de 1736, feita pelo matemático e físico Leonhard Euler. A teoria de redes complexas mostra a rede como um grafo onde temos os nós e as arestas que ligam esses nós, representando as relações entre os nós que ela conecta.

O surgimento dessa teoria deu grande importância para o estudo de redes em geral, importância essa que segue aumentando até os dias de hoje. Muitos estudiosos se dedicaram e ainda se dedicam à construção, constituição e propriedades dos mais diversos grafos. Os estudos realizados são fundamentais para o entendimento da relação do mundo que nos cerca nas mais diversas áreas desde a biologia até as redes de comunicações.

No caso da análise de uma rede de ingredientes de uma receita, que é o caso desse trabalho, observamos a relação de vários ingredientes em diversas receitas. Esse estudo procura nos indicar quais os ingredientes principais na receita para poder ver quais tem o sabor mais marcante ou o melhor sabor.

Do estudo dessa rede de ingredientes esperamos poder retirar informações como os ingredientes mais importantes, ingredientes com sabor mais marcante e ingredientes que transmitem seu sabor com maior facilidade.

Tendo todas essas informações poderemos pensar em melhores combinações e talvez até mesmo montar novas e melhores receitas culinárias.

II- Rede de Ingredientes de Receitas e suas Ligações

Os nós, que representam os ingredientes, também são classificados em grupos tanto em relação a sua origem (temperos, carnes, derivados animais e vegetais e produtos industrializados) quanto em relação ao seu sabor (doce, salgado, amargo e azedo).

A rede analisada foi montada seguindo a métrica de que para cada receita temos arestas indicando que o ingrediente X e Y estavam juntos nessa mesma receita. Cada aresta apresenta um peso referente às vezes que cada par de ingredientes apareceu em diferentes receitas.

É importante notar que vemos apenas um componente conexo, isso nos mostra que em todas as receitas analisadas não existe uma série de ingredientes isolados.

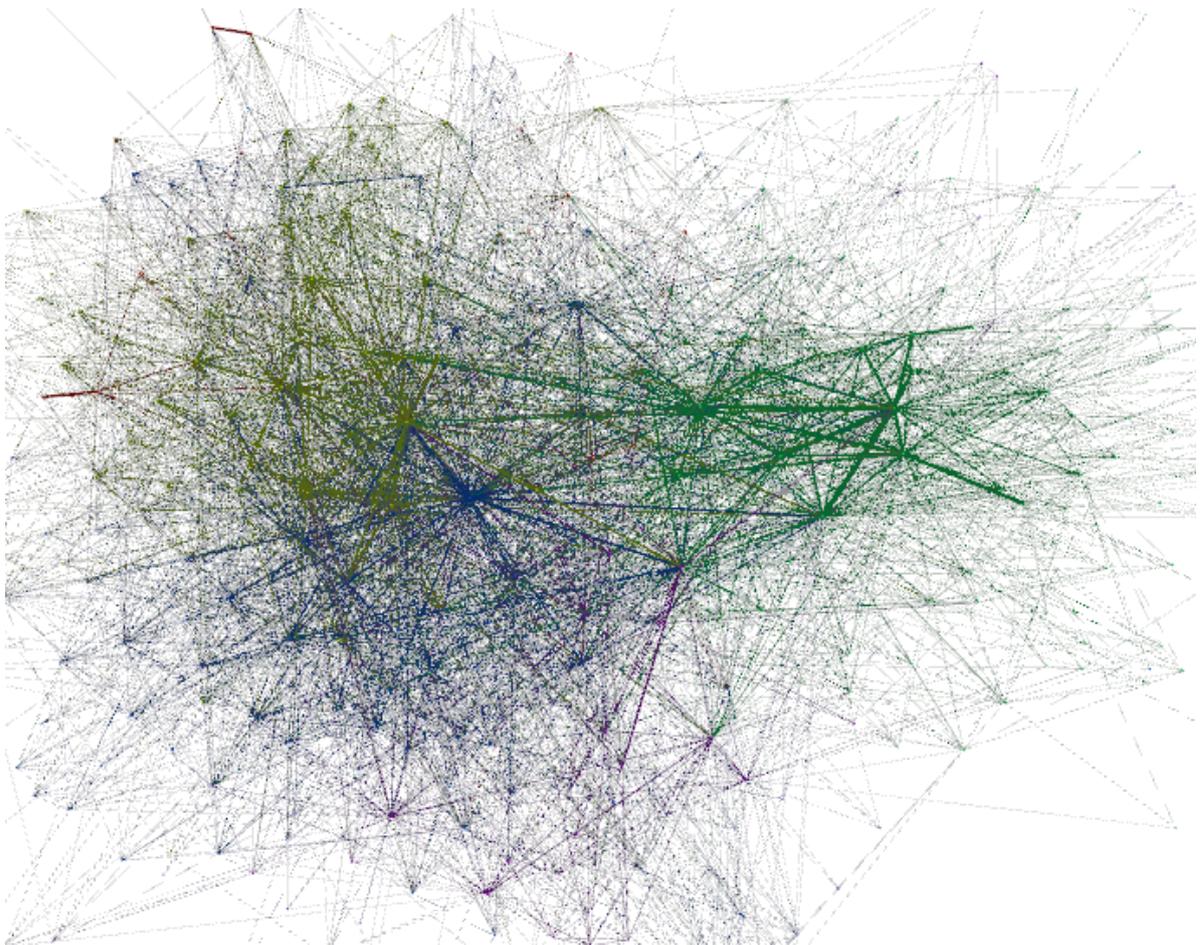


Imagem 1: Rede de ingredientes obtida.

III- Coleta de Dados

A coleta de dados para a construção da rede em questão começou com a manipulação e moldagem de scripts na linguagem python. Partindo dos scripts originais fornecidos pelo Professor Fabrício Olivetti, criamos 10 variações desses algoritmos que ao serem rodados no programa Anaconda retornaram cerca de 500 receitas aleatórias do site globo.

Uma vez que as receitas foram obtidas passamos para a parte de organização dos dados. A organização se deu na criação de três planilhas do Excel uma com todos os ingredientes (tabela de nós), uma com todas as ligações entre ingredientes formando as receitas e a repetição dessas ligações em diferentes receitas (tabela de arestas e respectivos pesos) e uma com a classificação dos ingredientes (origem e sabor).

Com as planilhas prontas partimos para a criação da rede em si. O programa usado foi o Gephi, bastou importarmos os dados das planilhas para o programa montar a rede e calcular suas características específicas.



Imagem 2: Layout do programa Anaconda.



Imagem 3: Layout do programa ghephi.

IV- Seção de Análise

Uma vez que a rede de ingredientes foi formada pudemos usar o programa gephi para calcular as suas particularidades. Todos os cálculos e dados obtidos são de extrema importância para atingirmos nosso objetivo de encontrar quais os ingredientes que fazem o melhor sabor.

A análise dos dados propriamente dita será separada em subseções para podermos observar cada detalhe da rede com mais atenção.

A-Panorama Básico da Rede

O grafo obtido é composto de 378 ingredientes com 4469 ligações entre eles nas cerca de 500 receitas analisadas. Como já foi mencionado antes foi obtido apenas um grande componente conexo que contém todos os ingredientes analisados.

Apesar da grande quantidade de nós a rede tem um diâmetro baixo, de cinco pra ser mais exato. O diâmetro pequeno nos mostra a facilidade para se percorrer a rede e encontrar diferentes sabores.

A densidade da rede indica a relação entre vértices e arestas. Por apresentar um número maior de arestas que de nós essa rede se caracteriza como uma rede densa. O valor calculado pelo programa para a densidade foi de 0,064.

Também é importante observarmos os graus médios da rede como um todo. Em média cada ingrediente tende a se ligar com outros 23,835 ingredientes. E se considerarmos os pesos das arestas (repetição de uma ligação em diferentes receitas) teremos o grau ponderado médio de 47,184. Isso reforça a ideia de uma rede muito conectada.

B-Características Individuais dos Ingredientes

Quando a análise de uma rede é feita um quesito de extrema importância a ser observado é a centralidade de cada nó. Existem diversos tipos de centralidades. Nesse trabalho analisaremos as centralidades de grau, proximidade, betweenness, autovetor.

O intuito da observação das centralidades está voltado para sabermos quais são os ingredientes com maior importância para diferentes funções em cada função.

A centralidade de grau nos indica quais são os ingredientes mais populares dentre as receitas. Ou seja, os ingredientes com maior centralidade de grau são os ingredientes que aparecem no maior número de receitas.

Os cinco ingredientes com as maiores centralidades de grau são:

1. Sal= 0.516
2. Cebola= 0.491
3. Alho= 0.473
4. Ovo= 0.406
5. Creme de Leite= 0.377

Já a centralidade de proximidade marca os ingredientes que conseguem transmitir o seu sabor de forma mais eficiente. Os ingredientes com grande centralidade de proximidade deixam rastros marcantes de sabor nas receitas das quais fazem parte.

Os cinco ingredientes que apresentaram maior centralidade de proximidade são:

1. Açúcar: 3,401
2. Salsicha: 3,377
3. Cebola: 3,340

4. Champignon: 3,270
5. Pimentão Verde: 3,086

A centralidade de betweenness por sua vez é responsável por nos mostra os ingredientes que poderiam ser chamados de “indispensáveis”, uma vez que esses ingredientes seriam os que causariam maior impacto caso fossem tirados da rede analisada, esses ingredientes marcam bem o sabor das receitas. São eles os nós que mais contribuem para o funcionamento da rede como um todo.

Os cinco ingredientes que mais marcam o sabor de suas receitas são:

1. Creme de Leite=7.505,225
2. Sal=6.517,017
3. Cebola=5.737,498
4. Açúcar=5.664,227
5. Ovo=5.221,936

A quarta centralidade é a centralidade de autovetor. Essa centralidade indica os ingredientes mais bem conectados na rede, são ingredientes que proporcionam melhores sabores através de suas conseguintes combinações.

As maiores centralidades de autovetor são dos seguintes ingredientes:

1. Sal= 1,000
2. Alho= 0,983
3. Cebola= 0,982
4. Tomate= 0,860
5. Ovo= 0,843

C-Parâmetros Gerais da Rede

O grafo gerado apresentou uma assortatividade de $-0,233$. Essa assortatividade negativa indica que na rede os ingredientes com grau alto tendem a se conectarem com aqueles que têm grau baixo. E na conjuntura das receitas isso nos indica uma recombinação de ingredientes que leva a formação de novos sabores.

O coeficiente de agrupamento médio da rede foi de $0,663$ que pode ser considerado alto que indica que a maioria dos ingredientes estão ligados a muitos outros ingredientes e eles formam muitos fechamentos triádicos nessa rede.

A distribuição de graus da nossa rede segue uma lei de potência o que era esperado por termos ingredientes muito mais comum a todas as receitas (como temperos, por exemplo).

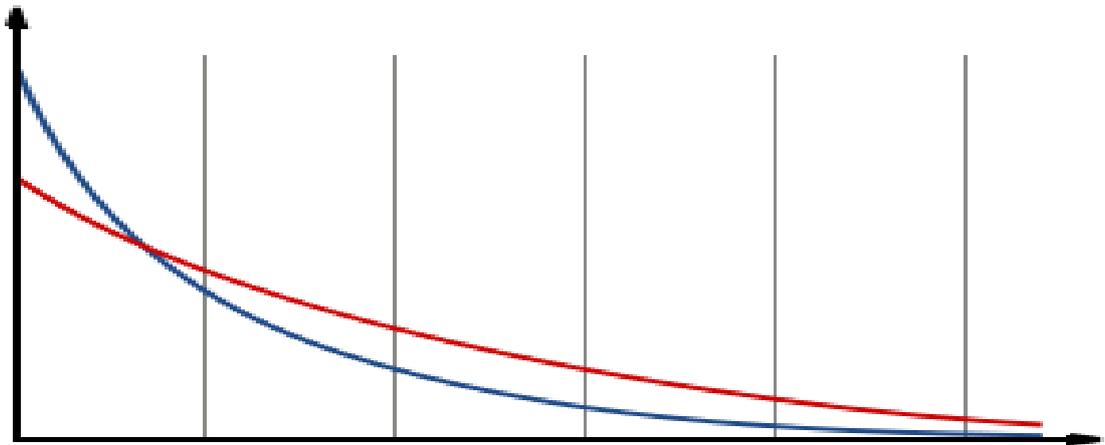


Imagem 4: Gráfico de distribuição de grau dos nós.

As arestas que mais aparecem, ou, os ingredientes que mais aparecem em conjunto em duas receitas indicam as combinações mais importantes ou as melhores combinações possíveis. Essas conexões são, em ordem:

1. Creme de leite-Leite Condensado= 31
2. Leite-Leite Condensado= 29
3. Sal-Cebola= 26

4. Açúcar-Rum= 24
5. Ovo-Leite Condensado= 23

Por fim achamos interessante dividir a nossa rede em comunidades baseadas na modularidade. Obtivemos 6 comunidades que podem ser divididas como:

1. Doces industrializados (Não Identificável na Rede, devido a ter poucos ingredientes)
2. Carnes e Temperos Suaves (Amarelo)
3. Temperos Fortes (Cor Rosa)
4. Vegetais, Derivados Animais (Cor Azul)
5. Temperos Doces (Cor Azul Claro)
6. Derivados Animais e Temperos Suaves (Cor Verde)

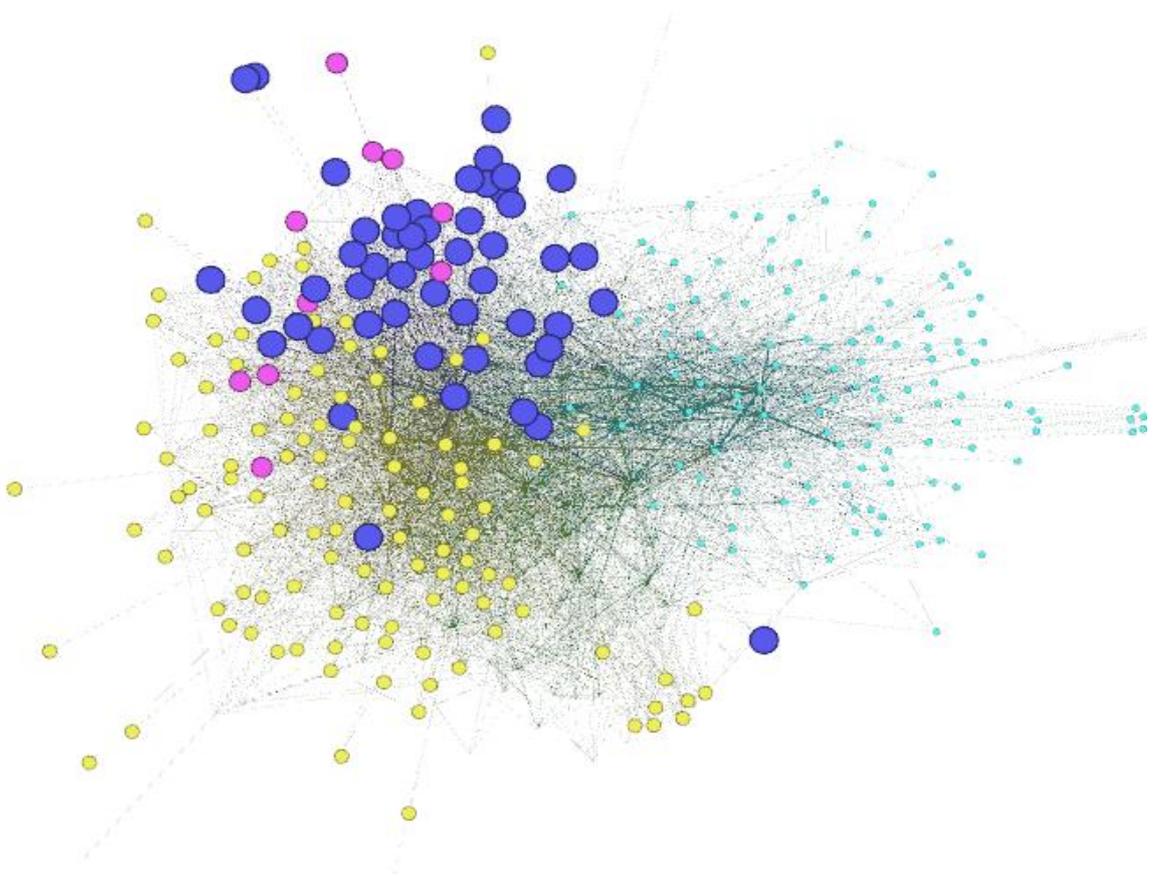


Imagem 5: Rede dividida em comunidades de acordo com a modularidade.

V-Análise e Observações

Uma peculiaridade que pudemos notar foi que ao separarmos a rede em comunidades três das seis comunidades são representadas por duas classes diferentes de ingredientes. É importante notarmos que, as comunidades que são identificadas como apenas uma única classe de ingredientes são compostas dos ingredientes de sua classificação que se conectam com diversos outros ingredientes, de diversas outras classes que não nos permitem escolher ou definir outra classe devidamente representativa, ou então se conectam diretamente com outras comunidades.

VI- Conclusão

Após a análise de tantos dados referentes à rede de ingredientes pudemos concluir que as melhores receitas podem ser feitas a partir de ingredientes que possuem maior afinidade entre seus sabores.

Também é importante que esses ingredientes tenham tendência a se ligarem com outros ingredientes semelhantes. Sendo que devemos nos lembrar que ingredientes com maior número de conexões se conectam com aqueles que menos se conectam.

Como já mencionado anteriormente no trabalho também concluimos que as centralidades são de extrema importância na análise dos ingredientes e de sua colaboração com as receitas e os sabores obtidos. Portanto ingredientes como cebola, sal e ovo (que aparecem entre os mais importantes de muitas centralidades) são definitivamente essenciais para a rede como um todo.

Tudo isso permite que nos aproximemos de um modelo próximo ao ideal para a criação de novas receitas que possam conseguir aperfeiçoar e maximizar o sabor de seus ingredientes.

Referências

Obtenção de receitas com os scripts usando receitas bases do site:
<http://tv.globo.com/receitas/>.

ALBERT, R.; BARABÁSI, A. L. Statistical Mechanics of Complex Networks. 2002.

GOLDBARG, ELIZABETH, GRAFOS: conceitos, algoritmos e aplicações. Editora ELSEVIER – CAMPUS, 2012.

SZWARFITER, J. L. Grafos e Algoritmos Computacionais. Rio de Janeiro, Editora Campus, 1984.