

Reconfiguração de kits cirúrgicos utilizando técnicas de redes complexas

Aluno: Gustavo Sasaki Roncaglia

Orientadora: Francisco Aparecido Rodrigues

Universidade de São Paulo, Instituto de Ciências Matemáticas e de Computação,

Departamento de Matemática Aplicada e Estatística

Dezembro de 2019

1 Introdução

A Santa Casa da Misericórdia de São Carlos, é uma instituição hospitalar da cidade de São Carlos, que atende tanto o SUS (60 %) por convênio e particulares (40 %), efetuando mais de 800 mil ligações anuais. Dentro destes, uma porção corresponde a várias cirurgias.

Em cada cirurgia são realizados vários procedimentos, por exemplo, para septoplastia (correção do desvio de septo), às vezes é necessário além da correção do septo, a remoção da carne esponjosa, totalizando 2 procedimentos. Para cada um desses procedimentos realizados, é utilizado um kit de produtos, composto por materiais, medicamentos e instrumentos específicos para cada tipo de procedimento a ser executado. Atualmente, existem quase 100 tipos de kits cirúrgicos, referentes a 17 especialidades. O que não é usado durante a cirurgia deve retornar à farmácia para reutilização. No entanto, sabe-se que alguns itens não utilizados são desperdiçados devido a contaminação ou desordem..

Considerando esse problema e sabendo que um funcionário gasta tempo para montar o kit, podendo levar até uma hora e meia. Este projeto tem como objetivo definir novos kits mais eficientes, cujos componentes são usados com mais frequência juntos, reduzindo a necessidade de remontagem dos kits.

Para tal, modelamos o banco de dados das cirurgias realizadas durante o período de um ano como um grafo, cada vértice representa um insumo médico e aplicamos

sobre ele técnicas de detecção da comunidade em redes complexas [4], de forma que cada comunidade corresponda a um novo kit.

2 Base de Dados

O banco de dados fornecido pela Santa Casa é composto por todas os insumos utilizadas em todos os procedimentos de 01/01/2017 a 31/12/2017. Uma primeira exploração dos dados fornece dois fatos importantes:

- Existe uma grande disparidade de ocorrências entre as especialidades. Nefrologia e radiologia, por exemplo, tem menos de 20 ocorrências cada, enquanto outras têm muitos mais casos, atingindo até um máximo de 31 % do total com ginecologia e obstetícia. O tipo de cirurgia cesariana em específico ocorreu aproximadamente 3560 vezes, correspondendo a 17 % do total. Com essas informações, foi decidido, no início do projeto, focar na redefinição de kits para a especialidade de ginecologia e obstrução (que inclui cesariana).
- Há casos onde há a falta ou excesso na quantidade de insumos usados num procedimento. Um dos motivos para essas instâncias se deve ao fato dos funcionários hospitalares as vezes registrarem o uso do insumo no procedimento errado da mesma cirurgia. Resultando numa base de dados bastante muito mais ruidosa.

Foi proposto para resolução deste problema a criação dos kits segundo a cirurgia em vez do procedimento. No entanto não foi adotado tal ideia pois cada ocorrência de um mesmo tipo de cirurgia pode apresentar procedimentos difer-

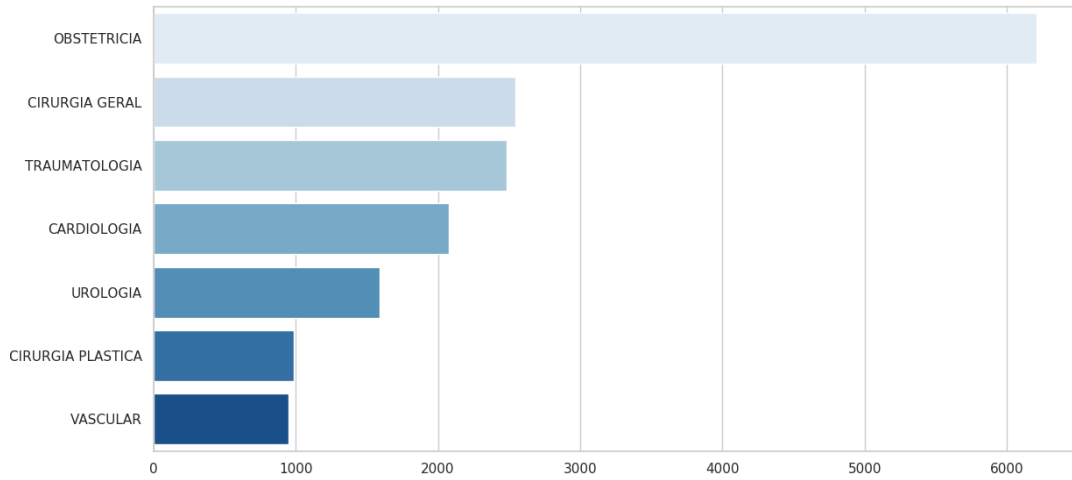


Figure 1: Tabela mostrando as especialidades com maior número de ocorrência. Juntas elas representam 84% de todas as cirurgias realizadas. Ginecologia e obstreticia compõe 31%, Cirurgia Geral e Ortopedia/Traumatologia cada uma compõe 12% do total, Cardiologia é 10%, Urologia 8%, cirurgia vascular 5% e cirurgia vascular 4% .

entes, como uma como infecção ou hemorragia inesperada, necessitando mais procedimentos que o normal.

3 Grafo

A base de dados fornecida pela Santa Casa foi modelado como um grafo ponderado não direcionado $G = \{V, A\}$ com 625 nós totalmente conectado, no qual cada vértice $n \in V$ representa um insumo e cada aresta $l \in A$ é um par ordenado $\langle x, z \rangle$ representando o uso dos dois insumo, x e z nas mesmas operações. O $peso_{xz}$ é um float calculado pela fração $peso_{xz} = F/E$, sendo F a quantidade cirurgias que utilizaram tanto x quanto z; e sendo E a quantidade de cirurgias que utilizaram x ou z. Quanto mais perto de um este valor, maior a frequência desses dois insumos serem utilizados juntos.

Com o intuito de facilitar uma análise inicial e permitir a criação de uma intuição do problema, foi utilizada nesses 6 meses um subgrafo de G com apenas instâncias que foram utilizadas em cesariana. O grafo resultante não apresenta distribuição power law, ele é altamente conectados, mas grande parte de suas conexões são muito fracas (menos de 0.2).

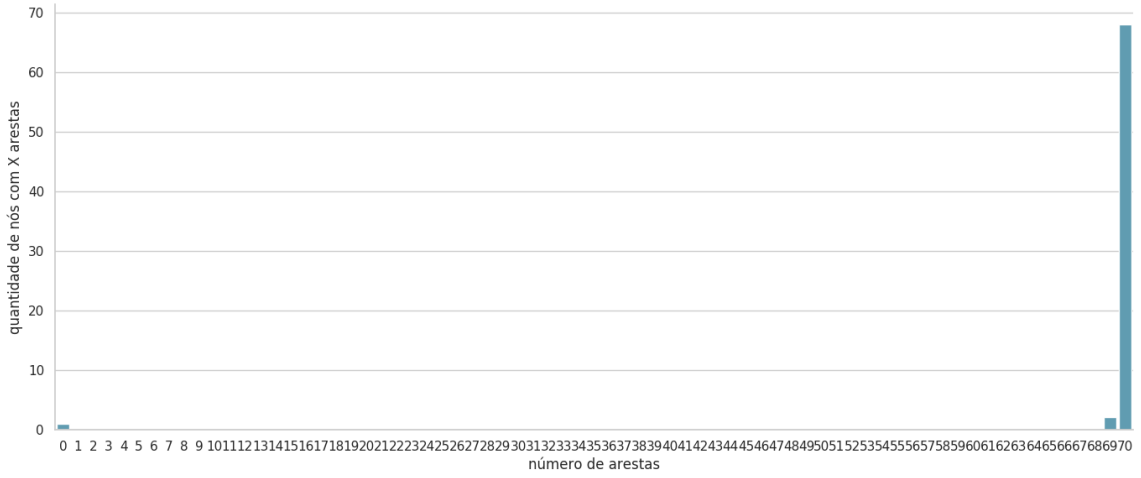


Figure 2: Histograma da quantidade de nós com X arestas. Quase todos os nós estão conectados com todos os outros nós.

4 Método

No projeto foi utilizado o algoritmo Louvain [2], este é um método de detecção de comunidades que se baseia na otimização da modularidade [5] do grafo completo. A modularidade é um valor entre -1 e 1 que representa o quão denso são as ligações internas das comunidades comparado com as ligações externas. Sua fórmula é:

$$M = \frac{1}{2m} \sum_{ij} [A_{ij} - \frac{k_i k_j}{2m}] \delta(c_i, c_j)$$

- M é o valor da modularidade

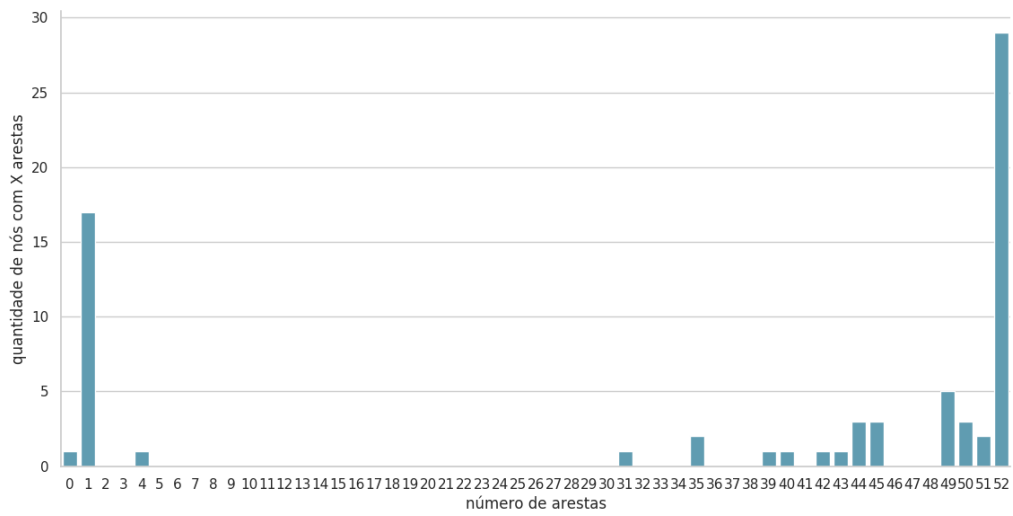


Figure 3: Cerca de 52% das arestas tem um valor abaixo de 0.2 , removendo elas, mostra a existência de nós isolados. Mesmo com mudança, percebe-se que grande parte dos nós continuam com quantidade parecidas de conexões, podendo eles indicarem a existência de uma comunidade.

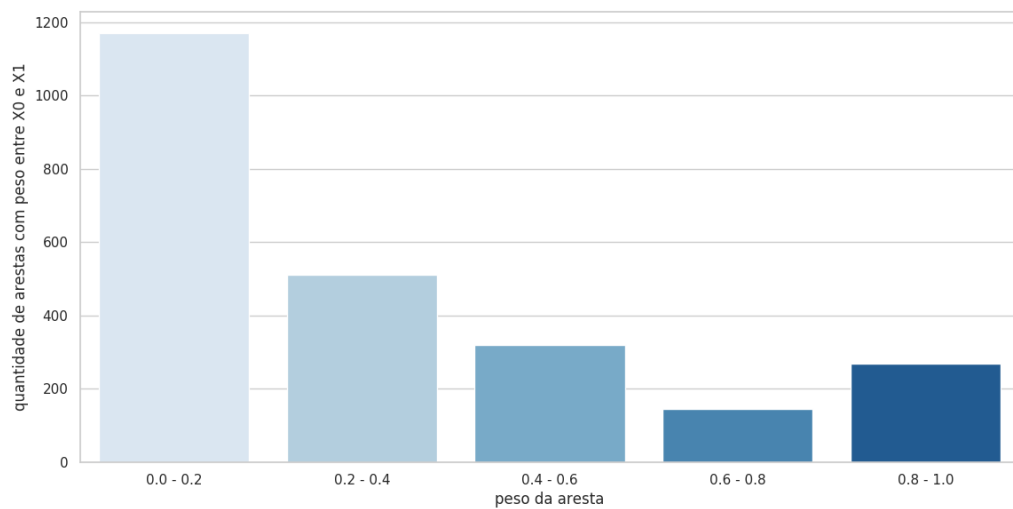


Figure 4: Este gráfico mostra a quantidade de arestas com valores entre 0 a 1 em intervalos de 0.2 .

- A_{ij} é o peso da aresta que liga i com j
- $2m$ é a soma dos pesos de todas as arestas
- δ é a função Kronecker delta. Caso c_i igual a c_j tem valor 1, caso contrário 0
- k_i, k_j são as somas dos pesos das arestas ligadas a i e a j respectivamente
- c_i, c_j são as comunidades de i e de j respectivamente

O seu funcionamento em linhas gerais:

Inicialização do grafo, cada nó começa em uma comunidade composta só por ele

for each nó in grafo **do**

grafo.modularidade \leftarrow *calcularModularidade(grafo)*

for each vizinho in nó **do**

grafoAux \leftarrow *grafo*

grafoAux.n.comunidade \leftarrow *grafoAux.vizinho.comunidade*

grafoAux.modularidade \leftarrow *calcularModularidade(grafoAux)*

if *grafoAux.modularidade* > *grafo.modularidade* **then**

grafo \leftarrow *grafoAux*

grafo.modularidade \leftarrow *grafoAux.modularidade*

end if

end for

end for

Na execução normal do Louvain, depois de realizado o código acima seria juntado todos os itens de uma comunidade num nó, combinando as arestas de todos os seus participantes. Com este novo grafo criado, você aplica novamente o Louvain e conseguir melhores valores de modularidade. No contexto do projeto, tal ação não traz benefícios pois a cada iteração, a comunidade engloba mais insumos, o que resulta em grandes comunidades com itens que são usados menos freqüentemente juntos.

Por tanto foi realizado algo inverso a isso, no final de cada iteração do código, foi criado um subgrafo composto pelas comunidades maiores e aplicado sobre esse novo subgrafo o Louvain, resultando em kits mais concisos e eficientes.

5 Atividades Realizadas

Primeiro foi realizado a leitura de diversos artigos científicos relacionado a redefinição de kits médicos [3][6], resultando na pesquisa de diversos algoritmos, procurando um que se encaixa na base de dados, inclusive foi cogitado o uso de clusterização, mas essa ideia foi abandonada. Depois dessa primeira fase foi decidido definitivamente o uso de detecção de comunidades, logo outra seção de leitura foi realizado com livros e artigos acadêmicos para aquisição dos conceitos básicos da área [1], e avaliação de qual algoritmo específico utilizar.

Definido nosso método, foi recebida a base de dados e sobre ela foi realizada uma análise exploratória para um melhor entendimento dos dados, necessitando algumas semanas devido a falta de familiaridade com a área médica.

Em seguida foi melhorado a base de dados, adicionando novas features (colu-

nas), e limpando a base, retirando instâncias estranhas, outliers , devido a diversos motivos.

Com os dados preparados, foi possível estrutura-lo como uma rede complexa, aplicar algoritmos de decção (Louvain) e analisar os resultados.

Todo esses processos depois revisão bibliográfica foi implementada em Python3 , utilizando várias bibliotecas externas como Pandas, Numpy , NetworkX , SeaBorn e Matplotlib .

6 Resultados Parciais

Terminado os 6 primeiros meses do projeto, já está modelado toda a base de dados como uma rede complexa com dados muito mais limpas comparado inicialmente. Além disso, foi implementado o código que realiza a aplicação do Louvain, permitindo a definição de novos kits.

Como foi comentado antes, para os testes iniciais só foi utilizado instâncias de cesária. O número de insumos diferentes usados nessas cirurgias é 70, e o número de comunidades detectadas é 20; dessas, a maioria (mais que 10) correspondem a comunidade composta por 1 ou 2 itens utilizados muito pouco, dos restantes há 3 comunidade de tamanho médio (composta por 5 a 10 insumos), esses com aproveitamento do uso dos insumos de 40 a 60% e uma comunidade composta por 21 itens com aproveitamento dos insumos de 35%. Ainda é necessário aplicar outros passos juntos aos atuais para melhorar o aproveitamento.

7 Próximas Atividades

Como foi utilizada apenas um subgrafo com instâncias da cesária, o próximo passo é definir os kits para outros tipos de cirurgia. Para tal foi proposto duas formas:

- aplicar o Louvain para a rede inteira
- aplicar o louvain para vários sub-grafos, cada um composto por instâncias de uma especialidade com maior ocorrência.

O primeiro jeito permite o uso do mesmo kit para cirurgias de especialidades diferentes, apesar disso apresenta um consumo computacional muito maior e uma análise de resultados muito mais complicada. O segundo apesar de teoricamente menos poderoso, tem uma facilidade de manuseio muito maior, tornando não trivial a escolha do método.

Outro ponto a ser levando é o método utilizado. Há a disponibilidade online de outros algoritmos o qual detectam comunidades em grafos, aplica-los no projeto não deverá gastar tanto tempo e talvez se enquadre melhor nos dados fornecidos pela base de dados. Dentro esses destaca-se o Label Propagation[7].

Caso sobre tempo, seria interessante abordar detecção de insumos similares. Há itens que são utilizados com mesma finalidade, sendo dependente do médico na decisão de qual usar. Apesar disso, o custo desses não são iguais, logo caso sege possível achar essa relação de similaridade, é possível substituir insumos mais caros por mais baratos sem perda brusca de segurança.

Tais itens similares podem ser teoricamente detectados visto que são rotineiramente usados com outros insumos parecidos, mas nunca juntos; ou seja, eles são

vértices com mesmos vizinhos, mas com conexão fraca entre si.

References

- [1] Albert-László Barabási et al. *Network science*. Cambridge university press, 2016.
- [2] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008, 2008.
- [3] Brecht Cardoen, Jeroen Beliën, and Mario Vanhoucke. On the design of custom packs: grouping of medical disposable items for surgeries. *International Journal of Production Research*, 53(24):7343–7359, 2015.
- [4] Mark EJ Newman. The structure and function of complex networks. *SIAM review*, 45(2):167–256, 2003.
- [5] Mark EJ Newman. Modularity and community structure in networks. *Proceedings of the national academy of sciences*, 103(23):8577–8582, 2006.
- [6] Caetano Oliveira de Oliveira. Gestão de estoques a partir da lista de materiais (bill of materials): o caso de um hospital universitário. 2017.
- [7] Xiaojin Zhu and Zoubin Ghahramani. Learning from labeled and unlabeled data with label propagation. 2002.