

3. *High Energy Physics*

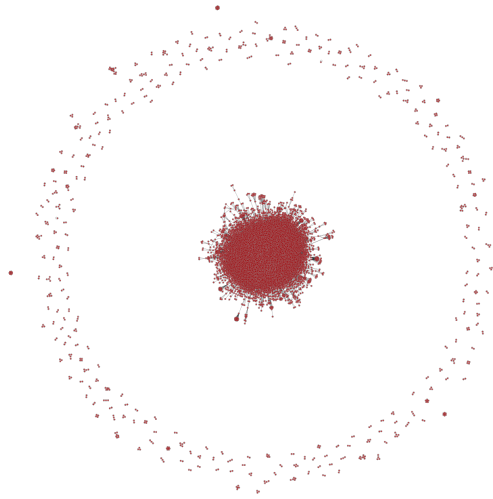


Figure 1: Grafo *High Energy Physics*

O grafo *High Energy Physics* é uma rede de colaboração entre autores de artigos científicos da categoria de *High Energy Physics*. Os autores são representados pelos vértices e se um autor i for co-autor de um artigo com o autor j , forma-se uma aresta unidirecional entre i e j . Se o artigo for de k autorias, forma-se um sub-grafo completo entre esses vértices. O grafo possui 12008 vértices e 118521 arestas.

Ph	degrees	distance	clustering	betweenness edges	betweenness vertex	components	slooseness
min	1	1	0	0	0		0,107145
max	491	13	1	0,004022	0,028025	11204	1
mean	19,740340	4,67	0,611166	3,43E-05	0,000266	43,194245	0,266357
std	46,637997		0,003540	0,000087	0,000871	670,58964	0,182584

- Clusterização Global : 0,65937213
- Quantidade de Componentes: 278

A rede de relacionamentos de *High Energy Physics* é bem parecida com a rede de emails Enron. Possui uma componente grande e diversas pequenas componentes que não são alcançadas pela maioria dos vértices, dessa forma, podemos dizer que o grafo não é fortemente conexo.

Analisando o gráfico (figura 2), podemos ver a distribuição do grau dos vértices em relação a sua probabilidade. O gráfico é resultado da estrutura desconexa,

tendo muitos vértices com grau muito baixo, entre 0 e 20, e poucos vértices com grau alto, justificando assim o grau máximo gerado.

Da mesma forma que o grafo anterior, esse grafo possui um diâmetro 13 e uma média de 4,67. Vale ressaltar que para alguns vértices, o valor da distância era infinita visto que existem diversas componentes desconexas.

Outra característica semelhante ao grafo de emails Enron, é o comportamento da distribuição dos índices de clusterização que pode ser visto no gráfico (figura 3). A maior parte encontra-se 0,2 e 0,8, seguido de um grupo de vértices possuem de 0,8 a 1,0. Uma parte dos vértices que está desconectada da componente maior acaba tendo um índice bem mais baixo. O índice relativamente alto da clusterização global pode ser justificado pela facilidade de surgimento de triângulo no grafo, visto que, a co-autoria de um artigo entre vários autores acaba criando um sub-grafo completo.

O gráfico (figura 4) traz as mesmas propriedades do grafo anterior. A maior parte dos vértices não fazem parte do caminho mínimo entre outros dois vértices, índice justificado pela quantidade de componentes desconexos e pela existência de vértices que interligam dois ou mais agrupamentos.

Se observarmos o tamanho da maior componente, perceberemos que 93% dos vértices encontram-se interligados, restando poucos vértices separados em 277 componentes menores, formados por menos de 20 vértices. Isso explica a média do tamanho das componentes baixa e o desvio padrão alto.

Por fim, o gráfico da distribuição do índice de *closeness* (figura 5). Comprovando os dados de clusterização e *betweeness*, constata-se que o índice está concentrado entre 0,2 e 0,4, que seriam os vértices que compõem a componente maior. Seguido deles, temos uma pequena parcela que possui índice próximos a 1, que são os vértices que conectam dois ou mais agrupamentos e outra pequena parcela que possui índice abaixo de 0,2, são os vértices que não fazem parte da componente conexa maior e possuem o grau baixo.

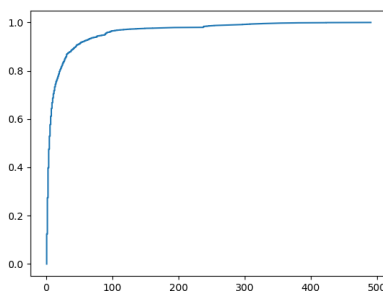


Figure 2: Grau x Probabilidade

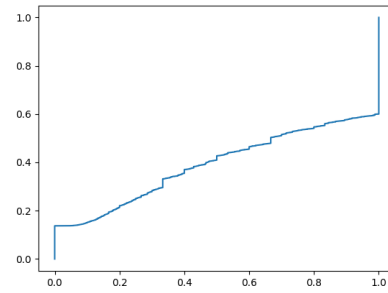


Figure 3: Clusterização x Probabilidade

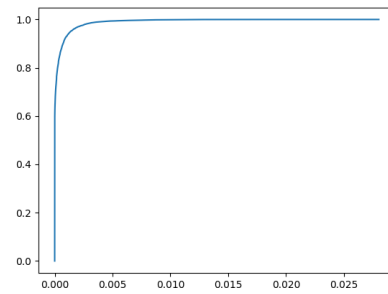


Figure 4: *Betweenness* x Probabilidade

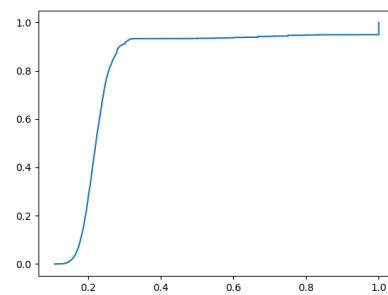


Figure 5: *Closeness* x Probabilidade