

Redes Complexas

Trabalho Prático

Alunos: Ana Clara Medina Fonseca
Lucas Borges Grandolpho

Professor: Vinicius da Fonseca Vieira

1 CONTEXTO

Os jogos olímpicos fazem parte de um evento centenário de imensa tradição no cenário mundial. Com a finalidade de estimular a competição saudável entre povos e nações, atletas e esportistas de todo o mundo reúnem-se a cada 2 anos, intercalando entre jogos de verão e inverno, para disputar medalhas de ouro, prata e bronze.

Desde sua criação, os jogos têm como principal função a união e a inclusão, mas é questionável se isto acontece realmente na realidade. Com a grande desigualdade existente entre as nações participantes, e o nível de investimento que cada uma delas recebe para sua participação, é no mínimo discutível se países menos desenvolvidos conseguem igualar ou ao menos se aproximar ao desempenho das grandes potências mundiais.

Observando estas questões, torna-se válido questionar se os jogos olímpicos efetivamente atenuam ou agravam ainda mais estas diferenças.

2 OBJETIVOS

O objetivo geral deste trabalho é descobrir se há relação entre o alto desenvolvimento de um país e o desempenho daquela nação nos jogos olímpicos. Isso será feito utilizando dados sobre medalhistas olímpicos, considerando os jogos desde Atenas 1896 até Rio 2016, e relacionando-os com métricas gerais sobre os países participantes. Com isso, este trabalho visa auxiliar no entendimento se há foco em investimento esportivo em países que enfrentam adversidades em outros aspectos, e se os jogos olímpicos ajudam neste processo.

3 METODOLOGIA

3.1 Extração dos dados

Os dados foram extraídos a partir de duas bases de dados. Uma chamada “120 Anos de História Olímpica: Atletas e Resultados”, e outra com nome de “Relatório de Felicidade Mundial”. Foram extraídos dados de ambas as bases, que por fim foram unidos para a realização dos experimentos. Isto é, ao fim da leitura e tradução de dados, cada país terá uma pontuação de medalhas e um índice de felicidade associados.

3.1.1 Base Medalhistas Olímpicos

Esta base inclui dados básicos sobre atletas e medalhas obtidas entre os jogos de Atenas 1896 e Rio 2016, divididos entre dois arquivos de dados. No arquivo “*athlete_events.csv*” estão inclusos mais de 270 mil linhas, cada uma delas representando um atleta olímpico competindo em um evento olímpico, e 15 colunas que representam:

1. ID (número identificador único para cada atleta);
2. Nome;
3. Gênero do atleta;
4. Idade;
5. Altura (em centímetros);
6. Peso (em quilogramas);
7. País que o atleta representou;
8. Comitê olímpico que o atleta representou (em siglas de 3 letras);
9. Jogos olímpicos que o atleta participou (ano e temporada);
10. Ano de participação;
11. Temporada (verão ou inverno);
12. Cidade sede dos jogos
13. Esporte;
14. Evento;
15. Medalha obtida (ouro, prata, bronze ou N/A).

O outro arquivo desta base, de nome “*noc_regions.csv*”, mostra exatamente qual país cada sigla de comitê olímpico encontrada no arquivo anterior representa.

É possível notar que existem muitos dados nesta base que não são interessantes para este trabalho. Por esse motivo, foi feito um trabalho de filtragem e tradução de dados.

Apesar da bases tratar de atletas individualmente, este trabalho deseja construir relações entre os países. Deste modo, foi preciso consolidar os dados de todos os atletas de um único país. Outro ponto importante é que foi usado os comitês olímpicos como país de origem para cada atleta, uma vez que, por questões políticas e culturais, por vezes um atleta representava mais de um país nos jogos, enquanto o comitê representante sempre é único. Para não ser necessário tratar com as siglas dos comitês, inicialmente foi feita uma tradução das siglas para o nome do país que ela representa.

Em seguida, foi feita a contagem de medalhas. Não seria ideal apenas contar as medalhas obtidas por cada país, já que uma medalha de ouro é mais valiosa que uma medalha de prata, que por sua vez é mais valiosa que uma medalha de bronze. Para resolver isto, foi criado um sistema de pesos para cada tipo de medalha. Neste sistema, uma medalha de ouro tem valor 2, uma medalha de prata tem valor 1,5, e uma medalha de bronze tem valor 1. Estabelecidos estes valores, foi realizada uma leitura da base de atletas por inteira. Para cada atleta, se ele obteve alguma medalha em algum evento, é adicionado o respectivo peso daquela medalha para o seu comitê olímpico (e, por consequência, para o seu país de origem). Assim, ao final da computação da base toda, cada país possui uma pontuação relacionada a quantidade e qualidade de medalhas obtidas em todos os jogos olímpicos.

3.1.2 Base Felicidade Mundial

Esta base foi construída em cima de uma pesquisa realizada pela Organização das Nações Unidas sobre a felicidade global. Sua primeira versão foi publicada em 2012, e vem sendo repetida anualmente até 2019, que foi o ano utilizado como base para este trabalho. O relatório classifica 156 países através de seus “índices de felicidade”, calculado utilizando respostas da população de cada país avaliando

algumas questões sobre a vida naquele lugar. O arquivo dispõe de 9 colunas que representam:

1. Ranking geral do país;
2. Nome do país;
3. Índice de felicidade;
4. PIB per capita;
5. Suporte social;
6. Expectativa de vida;
7. Liberdade;
8. Generosidade;
9. Percepção de corrupção.

Os valores encontrados nas colunas de PIB per capita, suporte social, expectativa de vida, liberdade, generosidade e percepção de corrupção indicam o quanto estes fatores contribuem para o cálculo do índice de felicidade de um determinado país. Essencialmente, é avaliado a importância daquela questão segundo a própria população daquela região.

Para este trabalho, foi simplesmente lido o nome de cada país e o seu respectivo índice de felicidade, para serem unidos com a pontuação de medalhas obtida com a base anterior.

3.2 Modelagem da Rede

Com todos os dados extraídos das bases, o próximo passo a ser dado é a modelagem e construção da rede em si.

A rede foi modelada de forma a relacionar as ideias da pontuação de medalhas em jogos olímpicos e o índice de felicidade de cada país. Para isso, foi definido que cada vértice da rede seria um país. A definição da existência de arestas entre dois vértices quaisquer foi feita a partir da diferença entre os índices de felicidade dos dois países representantes. Se dois países têm uma diferença entre seus respectivos índices de felicidade menor do que 0,05, será criada uma aresta entre seus vértices.

Deste modo, ficou garantido que países com índices de felicidade com valores próximos sejam ligados entre si e formem possíveis comunidades de vértices.

Para a pontuação de medalhas de cada país, foi utilizado um sistema de escala de cores. Esta escala parte da cor azul (para os valores mais baixos) e vai até a cor amarela (para valores mais altos), como ilustrado da figura abaixo:



Figura 1: Escala de cores utilizada para pontuação de medalhas

Assim, países com melhor pontuação de medalhas obtidas terão seu vértice com uma cor mais próxima da extremidade amarela, enquanto países com pontuações mais baixas possuirão um vértice com uma cor mais próxima da extremidade azul.

3.3 Experimentos

Uma vez modelada a rede, foi construído e gerado o grafo que a representa. O grafo, então, ficou da seguinte forma:

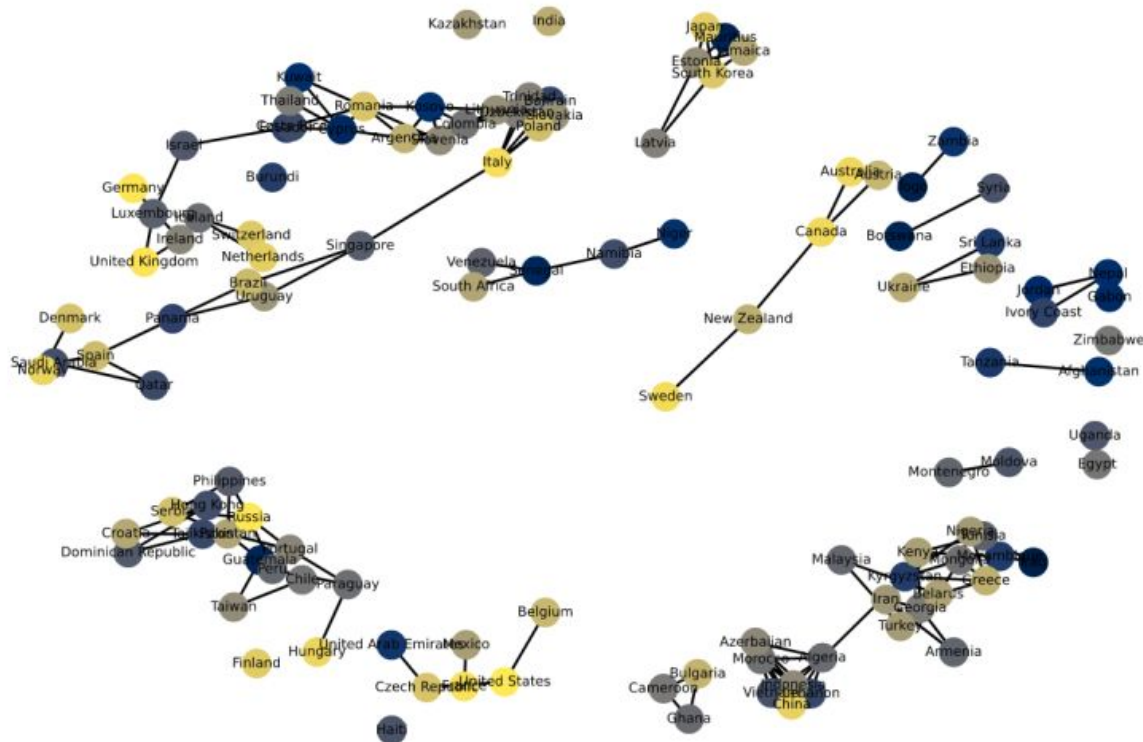


Figura 2: Grafo modelado a partir da rede

3.3.1 Análise Descritiva Básica

Serão apresentadas agora algumas características básicas do grafo criado¹:

- Número de vértices: 118
- Número de arestas: 167
- Densidade: 0,024
- Grau médio: 2,831
- Coeficiente de clusterização médio: 0,756
- Distribuição de graus:

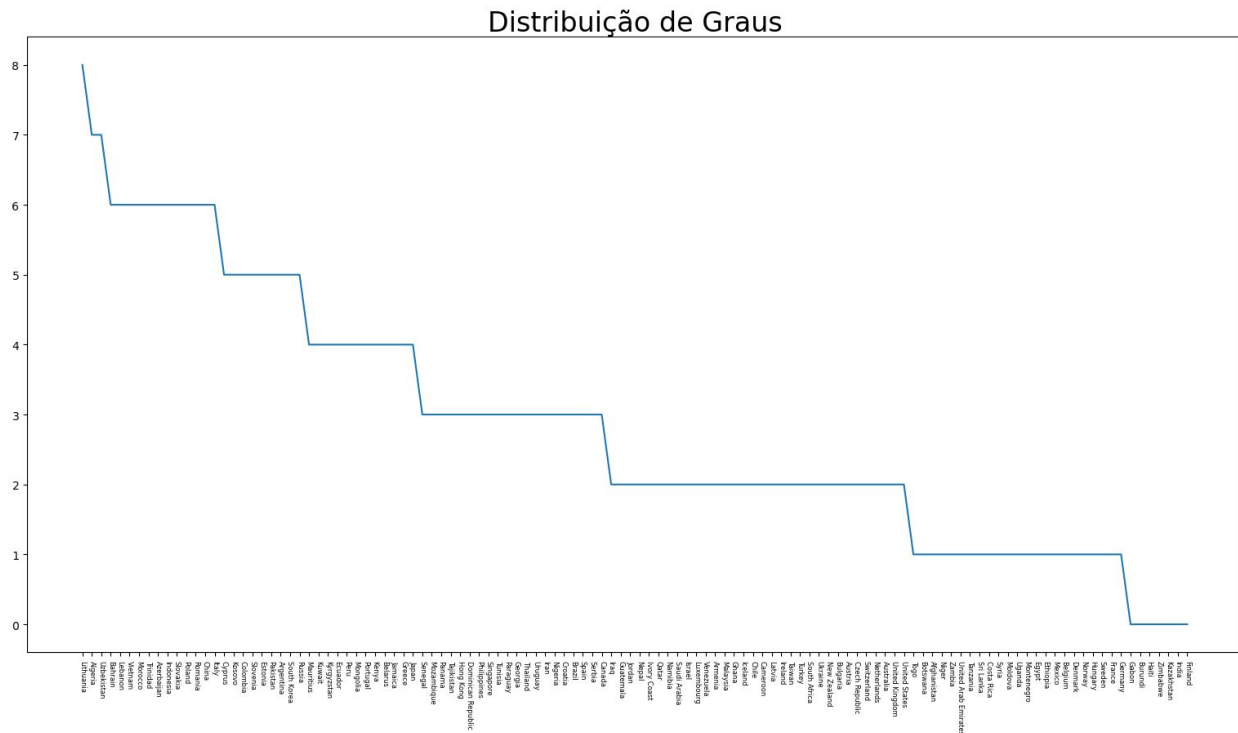


Figura 3: Gráfico de distribuição de graus

¹ Para encontrar essas características no código, foi utilizado a biblioteca *igraph*, apesar do grafo ter sido feito com a biblioteca *networkX*. Assim, foi utilizada uma função que converte um grafo feito em *networkX* para *igraph*. Esta função só existe na versão mais recente do *igraph*, então, caso ocorra um erro nesta parte da execução, favor atualizar o *python-igraph* instalado na máquina, e o problema deve ser resolvido.

3.3.2 Análise Topológica

É possível notar que foram formadas algumas comunidades no grafo gerado, como pode-se observar a seguir:



Figura 4: Exemplo de comunidade formada



Figura 5: Exemplo de comunidade formada

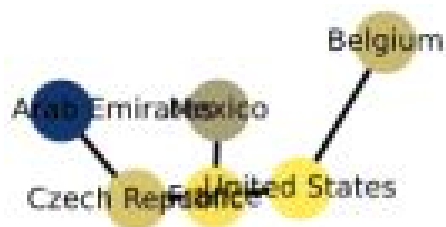


Figura 6: Exemplo de comunidade formada

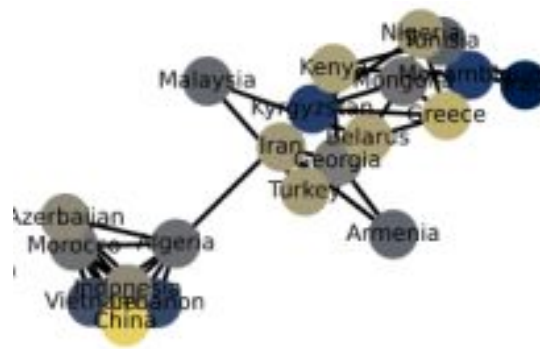


Figura 7: Exemplo de comunidade formada

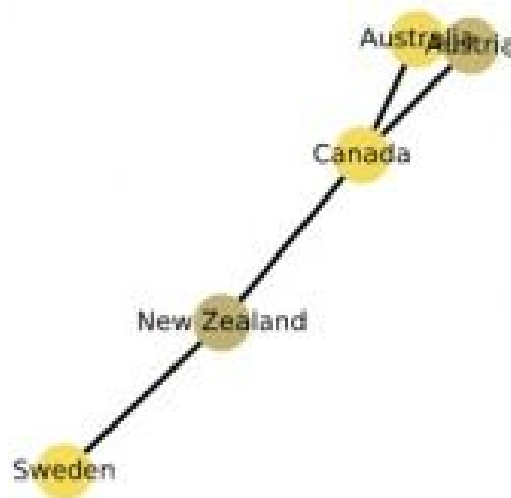


Figura 8: Exemplo de comunidade formada

4 DISCUSSÃO

Os experimentos são capazes de mostrar algumas coisas. Na análise básica, nota-se que o grafo possui 118 vértices, mesmo que os dados de felicidade classificavam 156 países. Isso ocorre pois provavelmente nem todos os países classificados conseguiram conquistar medalhas ou sequer participam dos jogos olímpicos. O número de arestas mostra que a forma como a criação de arestas foi definida foi adequada para formar um grande número de ligações sem que houvesse conexão entre a maioria dos países, o que tornaria difícil qualquer análise. Ainda assim a densidade resultou em um valor baixo, o que mostra que mais países não tem

conexão do que o contrário, o que é evidenciado também pelo resultado do grau médio. Já o coeficiente de clusterização médio teve um resultado relativamente próximo de 1, o que mostra que a probabilidade média de que dois vértices que são vizinhos de um terceiro vértice sejam também vizinhos entre si é relativamente alta.

As comunidades formadas vistas na análise topológica mostram que os vértices com cores próximas têm uma certa tendência em formar comunidades entre si, já que foi visto comunidades com cores predominantemente amarela, cinza ou azul. Isso significa que países com índices de felicidade próximos tendem a ter desempenho similar nos jogos olímpicos.

5 CONCLUSÃO

Analisando os resultados apresentados e discutidos, é possível concluir que os países que têm melhor desempenho em todos os jogos olímpicos até hoje são, de fato, aqueles que são melhores desenvolvidos. A causa mais provável para isso ocorrer é que países mais desenvolvidos são capazes de realizar grandes investimentos na área esportiva de sua nação, com estruturas de treinamento melhores e uma maior profissionalização de seus atletas.

Estes resultados mostram que, ainda que existam muitos atletas com um potencial imensurável no mundo, o lugar onde estes atletas nasceram acaba sendo um fator tão importante quanto, e pode ainda tornar-se um obstáculo para o seu sucesso, uma vez que o talento e o esforço por si só não conseguem garantir bons resultados à eles, resultados estes que são diretamente dependentes do investimento disponível para auxiliá-los.

6 REFERÊNCIAS BIBLIOGRÁFICAS

1. GRIFFIN, Randi H. *120 years of Olympic History: Athletes and Results*. 2018. Disponível em:
<<https://www.kaggle.com/heesoo37/120-years-of-olympic-history-athletes-and-results>>
2. SUSTAINABLE DEVELOPMENT SOLUTIONS NETWORK. *World Happiness Report*. 2019. Disponível em:
<<https://www.kaggle.com/unsdsn/world-happiness>>
3. DAHIS, Rafael. *Redes Complexas: Propriedades das Redes*. Rio de Janeiro. Universidade Federal do Rio de Janeiro. Disponível em:
<https://www.gta.ufrj.br/ensino/eel879/trabalhos_vf_2009_2/dahis/propriedades.html>
4. TODA MATÉRIA. *Olimpíadas (Jogos Olímpicos)*. 2020. Disponível em
<<https://www.todamateria.com.br/olimpiadas/>>