

Mapeando a Sinergia de Jogadores no Futebol Virtual

Uma Abordagem Analítica Através de Grafos e Exploração de Comunidades

Celso Vinícius Sudário Fernandes

Bacharelado em Engenharia da Computação

CEFET-MG, Campus V

Divinópolis, Brasil

celso.23@aluno.cefetmg.br

Resumo: Este estudo propõe uma metodologia inovadora para analisar as interações e sinergias entre jogadores de futebol, empregando uma abordagem fundamentada em redes complexas e grafos ponderados. Em contraste às limitações frequentemente encontradas nas abordagens convencionais, que negligenciam as dinâmicas intrínsecas entre jogadores, esta abordagem destaca a importância da contribuição individual dos jogadores para o sucesso coletivo. Por meio da análise de dados estatísticos de jogadores provenientes de um banco de dados de um jogo de futebol virtual, e através da aplicação de um algoritmo de análise de comunidades, este trabalho visa identificar combinações eficazes de jogadores, considerando a possibilidade de representação privilegiada de posições, alinhada com as demandas de entretenimento do jogo. Assim, esta abordagem explora como as interações entre posições diferentes, refletem e potencialmente preveem o sucesso tático e a eficiência da equipe, buscando oferecer percepções valiosas para a formação estratégica de equipes no futebol virtual.

Palavras-chave: Grafos, Comunidades, Futebol, Futebol Virtual, Conexão, Jogadores, Threshold, Equipes, Desempenho.

I. INTRODUÇÃO

A análise de grafos tem emergido como uma ferramenta robusta para a compreensão das complexas dinâmicas esportivas, particularmente no futebol, onde a interconectividade entre os jogadores é crucial para o sucesso coletivo, mesmo no futebol eletrônico. Este estudo se aprofunda na exploração das sinergias e estratégias táticas que moldam a dinâmica do jogo, empregando grafos ponderados para elucidar como as estatísticas de desempenho dos jogadores em um ambiente virtual refletem na performance da equipe. Investiga-se a premissa de que a estrutura e densidade dessas redes não apenas espelham, mas também podem ser indicativos do sucesso tático e do desempenho coletivo, com um olhar particular sobre a relevância das posições e a ênfase no ataque como vetor de entretenimento no contexto dos jogos virtuais de futebol.

II. CONTEXTUALIZAÇÃO

O estudo da análise de redes e grafos é fundamental para compreender as complexas interações em diversos sistemas, incluindo esportes como o futebol. Esta seção apresenta os fundamentos teóricos que embasam a pesquisa, abrangendo

conceitos de grafos, análise de redes complexas, e a aplicação do conceito de comunidades no contexto esportivo.

A. Teoria dos Grafos

A teoria de grafos é uma área da matemática e ciência da computação que estuda as relações e interconexões entre diferentes objetos. Em diversos campos, incluindo ciências sociais, biologia, informática e esportes, a teoria de grafos é aplicada para modelar e analisar sistemas complexos. No contexto do futebol, ela pode ser usada para representar jogadores e suas interações, ajudando a entender dinâmicas de equipe e estratégias de jogo.

B. Comunidades em Grafos

Na teoria dos grafos, uma comunidade ou *cluster* é um conjunto de vértices com conexões mais densas entre si do que com o restante do grafo. Este conceito é essencial para compreender redes complexas em diversos contextos. No futebol, por exemplo, um *cluster* pode representar jogadores que trocam passes frequentemente, revelando estratégias de jogo ou afinidades táticas. A análise desses *clusters* pode oferecer abstrações importantes sobre estilos de jogo e a eficiência de diferentes formações ou parcerias entre os jogadores.

C. Algoritmos Relevantes

1) *Kamada-Kawai*: Utilizado para a visualização do grafo em primeira análise do estudo, visa otimizar a disposição dos nós para proporcionar uma clareza visual ideal. Este algoritmo calcula posições dos nós de forma a minimizar a energia de um sistema de molas, promovendo uma distribuição uniforme e reduzindo sobreposições. No contexto deste estudo, o *Kamada-Kawai* foi aplicado para destacar as interações e a importância relativa dos jogadores dentro da rede.

2) *Greedy Modularity Communities*: Para a detecção de comunidades, o algoritmo de modularidade gananciosa, agrupa os nós em comunidades densamente conectadas internamente. Esse método é baseado na ideia de maximizar a "modularidade", uma métrica que quantifica a força da divisão de uma rede em módulos ou comunidades. No contexto do futebol, essas comunidades podem representar grupos de jogadores que frequentemente colaboram ou combinam em campo.

III. TRABALHOS CORRELATOS

No cenário atual do futebol moderno, a análise do desempenho esportivo tem ganhado destaque, conforme evidenciado no trabalho de Francisco José Marques dos Santos, "Avaliação do desempenho de equipes de futebol profissional: Análise das interações entre jogadores" [6]. Este estudo resalta a relevância do trabalho em equipe e a influência da cooperação entre jogadores no desempenho coletivo, utilizando a metodologia *Design Science Research* para avaliar o impacto dessas interações.

Complementarmente, a pesquisa de Daniele Cristina Uchôa Maia Rodrigues em "Complex network measurements in graph-based spatio-temporal soccer match analysis" [7], explora a análise de redes complexas no futebol, com ênfase nos aspectos espaço-temporais das partidas. Utilizando grafos para mapear a dinâmica do jogo, o estudo destaca a importância das posições e movimentações dos jogadores, ampliando a aplicabilidade da análise de grafos no futebol.

Adicionalmente, o trabalho "Data Analytics in the Football Industry: The Portuguese Football Federation Case" [8], de Sofia Alexandra Grifo Cunha, foca na crescente relevância da análise de dados no futebol. Este estudo detalha a implementação de práticas de análise de dados e inteligência empresarial na Federação Portuguesa de Futebol, evidenciando como as decisões estratégicas no esporte estão se tornando cada vez mais orientadas por dados.

Desse modo, ao revisitar os trabalhos relacionados, este projeto busca demonstrar a diversidade e riqueza de abordagens na análise de dados no futebol, cada um contribuindo com perspectivas únicas que enriquecem o entendimento global da dinâmica do esporte. Assim, ao aplicar a teoria de grafos e o análise de comunidades, o presente estudo busca estabelecer uma ponte entre as análises qualitativas existentes das dinâmicas de equipe reais e das composições de jogadores no futebol virtual.

IV. METODOLOGIA

O presente estudo emprega o extenso conjunto de dados "EA Sports FC 24 Complete Player Dataset" [6], disponível na plataforma Kaggle, que inicialmente consistia em uma compilação abrangente de mais de 14.000 jogadores, incluindo estatísticas detalhadas de jogadores de futebol. Estes dados englobam atributos como nome, nacionalidade, clube, posição, idade, e avaliações de habilidades específicas: chute, passe, drible, defesa e finalização. A seleção dos dados visa representar um espectro amplo e diversificado de jogadores buscando uma análise abrangente.

A. Preparação dos Dados

A preparação dos dados envolveu um processo meticuloso de filtragem para refinar a lista para um subconjunto gerenciável que preservasse a integridade estatística e representatividade do conjunto original. Utiliza-se um algoritmo de filtragem personalizado para selecionar aproximadamente 400 jogadores, distribuídos equitativamente pelas posições definidas (atacantes, meias, zagueiros e laterais), buscando

assegurar a diversidade e relevância tática no contexto do futebol.

```
1 INICIO
2   IMPORTAR biblioteca pandas como pd
3   DEFINIR caminho do arquivo CSV
4   CARREGAR dados do arquivo CSV para 'players_data'
5   ORDENAR 'players_data' por 'Overall' decrescente
6   DEFINIR 'categories' com posicoes
   correspondentes
7   DEFINIR 'players_per_category' por categoria
8   INICIAR 'top_players_by_category' como DataFrame
   vazio
9   PARA CADA categoria EM 'categories'
10      SELECIONAR top jogadores
11      CONCATENAR ao 'top_players_by_category'
12      SALVAR 'top_players_by_category' em novo CSV
13 FIM
```

Pseudocódigo 1. Representação da Filtragem dos Dados.

B. Modelagem do Grafo

A modelagem do grafo foi realizada utilizando a linguagem de programação Python, com auxílio das bibliotecas NetworkX para a construção e manipulação do grafo e Matplotlib para a visualização.

1) *Vértices*: No grafo modelado, cada jogador é representado como um vértice. Os atributos dos vértices incluem informações como posição, clube e avaliação geral.

2) *Arestas*: As arestas são estabelecidas entre jogadores de posições que normalmente interagem em campo (por exemplo, atacante com meio-campista, meio-campista com lateral, etc.).

3) *Ponderação das Arestas*: As arestas são ponderadas com base no *threshold* (0 - 100) estabelecido entre as posições (por exemplo, a avaliação geral das habilidades de um atacante para conexões envolvendo meio-campistas é de 88).

A construção inicial do grafo foi realizada com o intuito de mapear as relações funcionais entre os jogadores. Utiliza-se dados de desempenho dos jogadores para criar nós no grafo, representando cada atleta. As conexões, ou arestas, entre eles foram estabelecidas com base em uma política de conexões predefinida que reflete as interações táticas comuns no futebol. A visualização do grafo foi aprimorada através do algoritmo *Kamada-Kawai*, destacando as interações por posição, este passo é demonstrado a seguir e mostra-se fundamental para entender como as habilidades individuais podem influenciar a sinergia da equipe como um todo.

```
1 INICIO
2   IMPORTAR bibliotecas necessarias (pandas ,
   networkx , matplotlib)
3   CARREGAR dados dos jogadores para o DataFrame '
   players_data'
4   CRIAR um grafo vazio 'G' usando networkx
5   ADICIONAR cada jogador como um no no grafo 'G'
6   DEFINIR 'connection_policy' para estabelecer as
   regras de conexao
7   PARA CADA par de jogadores
8       SE 'connection_policy' for atendida
9           ADICIONAR uma aresta entre os jogadores
10      EXIBIR o grafo com nos e arestas
11 FIM
```

Pseudocódigo 2. Grafo de Interações entre Jogadores por Posição.

Na segunda etapa, o foco foi identificar agrupamentos ou comunidades de jogadores que mostraram um nível de interação e colaboração mais significativo. Através da aplicação do algoritmo de clusterização *greedy-modularity-communities* exemplificado à frente, conhecido por sua eficácia em redes complexas, foi possível discernir o "Cluster de Maior Conectividade", uma comunidade-chave com potencial impacto no desempenho coletivo. Além disso, determinou-se que os jogadores mais centrais dentro deste cluster, proporcionando uma compreensão mais profunda das influências individuais na rede.

```

1 INICIO
2   UTILIZAR o algoritmo '
   greedy_modularity_communities' para detectar
   comunidades
3   IDENTIFICAR o maior cluster dentre as
   comunidades encontradas
4   CALCULAR a centralidade de grau para cada
   jogador dentro do maior cluster
5   SELECIONAR os 'top k' jogadores com maior
   centralidade de grau
6   EXIBIR o grafo destacando o maior cluster e os '
   top k' jogadores
7   SALVAR a lista de clubes dos 'top k' jogadores
   para análise posterior
8 FIM

```

Pseudocódigo 3. Grafo de Maior Cluster e Top k Jogadores.

Os resultados gerados pela aplicação do algoritmo foram traduzidos em visualizações gráficas, que facilitaram a interpretação dos padrões de interação entre os jogadores. A discussão sobre a relevância dessas interações e como elas se correlacionam com estratégias de jogo reais é apresentada detalhadamente na seção "V. Resultados e Análise". A metodologia adotada busca a reprodutibilidade da análise, procurando assegurar que a aplicação dos procedimentos descritos conduza a resultados consistentes e verificáveis.

V. RESULTADOS E ANÁLISE

A presente análise concentra-se em uma rede estruturada de aproximadamente 400 jogadores, cuidadosamente filtrada para atestar uma representação equitativa das posições-chave dentro do futebol. A base de dados foi meticulosamente selecionada para incluir um número igual de jogadores para cada posição designada, com o objetivo de visar uma análise imparcial das interações e dinâmicas de jogo.

A. Reprodução de Modelo

A fim de atestar a transparência e a reprodutibilidade dos resultados apresentados neste estudo, todo o código-fonte e a documentação detalhada deste estudo estão disponíveis em um repositório GitHub público. O diretório "src/" contém os scripts Python usados para a análise, permitindo que outros pesquisadores validem, reproduzam ou expandam a pesquisa. A URL do repositório Git é fornecida na seção de Referências [5] deste artigo.

B. Funcionamento Geral do Grafo

O grafo foi construído com base em um modelo de rede que designa cada jogador como um nó e estabelece as arestas de

acordo com as interações táticas predefinidas e um *threshold* (limite) de valor geral de habilidade, ou *Overall*. Este limite foi ajustado para refletir o alto calibre dos jogadores incluídos na análise, com o pressuposto de que jogadores de nível superior estão mais propensos a interagir em contextos de jogo reais. A seleção estratégica de jogadores e o design criterioso do grafo visam proporcionar uma visão realista das estruturas de equipe e das sinergias em campo.

C. Análise Primária

A figura subsequente ilustra um grafo simplificado, evidenciando uma rede complexa de interações entre aproximadamente cem jogadores em cada posição, estruturada com base em um *threshold* diferenciado para cada categoria: "Atacantes" com *Overall* de 88, "Meias" com 86, "Zagueiros" com 84 e "Laterais" com 85. Esses limites foram estabelecidos para refletir as habilidades distintas exigidas para cada posição, garantindo que apenas os jogadores mais habilidosos e táticos fossem considerados na análise. Cada nó é colorido de acordo com a posição do jogador, promovendo uma visualização clara das contribuições táticas e técnicas individuais à dinâmica da equipe.

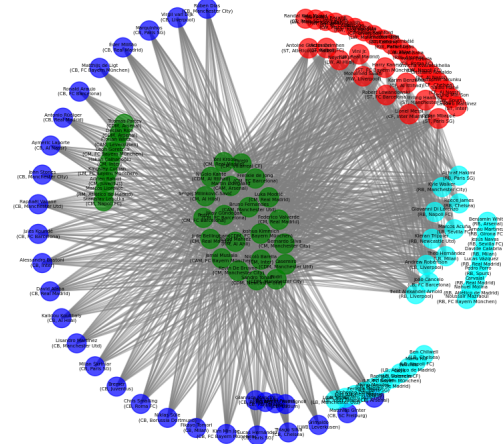


Fig. 1. Grafo de Interações entre Jogadores por Posição.

Em uma primeira análise, o grafo indica que jogadores classificados como "Meias" (verde) formam o núcleo central da rede, sugerindo que são fundamentais na construção de jogadas e na ligação entre defesa e ataque que os jogos de futebol virtual costumam agraciar. A concentração de conexões nessa categoria de posição pode ser interpretada como um indicativo de que os meias são essenciais para manter a posse de bola e distribuir o jogo, em alinhamento ao papel tradicionalmente atribuído a esses jogadores no futebol real e corrobora com a tendência de design de jogos eletrônicos em promover um estilo de jogo dinâmico e ofensivo.

Por outro lado, "Atacantes" (vermelho) e "Laterais" (ciano) aparecem mais periféricos na rede, mas ainda assim mantêm uma presença proeminente, alinhando-se com a preferência

dos usuários por uma experiência de jogo voltada para a ação e o espetáculo do gol. A posição mais reservada dos zagueiros reflete sua função defensiva, menos envolvida nas complexidades do meio de campo, mas não menos vital para uma estratégia equilibrada e uma defesa sólida no contexto virtual do futebol. A posição mais reservada dos "Zagueiros" (azul) reflete sua função defensiva, menos envolvida nas complexidades do meio de campo, mas não menos vital para uma estratégia equilibrada e uma defesa sólida no contexto virtual do futebol.

D. Análise de Comunidades

Após a implementação do algoritmo de detecção de comunidade *greedy-modularity-communities* em uma rede de jogadores de futebol, observa-se um padrão intrigante que realça a centralidade e a influência de certos jogadores no contexto geral do jogo. Os resultados obtidos revelam um maior cluster dominado por jogadores de clubes proeminentes, especificamente do Manchester City, Manchester United, FC Bayern München, Real Madrid, Chelsea, dentre poucos outros. Esta concentração de jogadores-chave em clubes de renome sugere uma tendência de aglomeração de talentos em equipes de alto calibre, potencialmente refletindo estratégias de imersão às performances reais desses clubes.

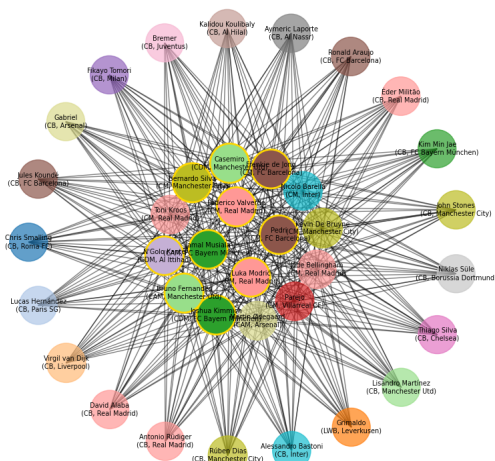


Fig. 2. Grafo de Maior Cluster e Top k Jogadores.

A visualização destaca um núcleo denso de jogadores altamente conectados, simbolizando a presença de um eixo central em torno do qual o jogo pode evoluir. Os "top k" jogadores, delineados em dourado no grafo, são em suma maiores meio-campistas e atacantes, e podem ser interpretados como os principais distribuidores ou receptores de jogadas, indicando sua relevância tática dentro de suas respectivas equipes.

Ademais, a predominância de jogadores altamente conectados, oriundos de clubes europeus de renome, evidencia uma modelagem do jogo que valoriza o controle do meio-campo e a criação de jogadas. Esta centralidade tática dos meios no grafo reflete uma tendência do jogo virtual em simular com

precisão as dinâmicas reais do futebol, onde a habilidade de distribuir o jogo e manter a posse de bola são vitais.

Assim como, a presença dominante de jogadores destes clubes dentro do maior cluster pode ser interpretada como um reflexo de suas estratégias bem-sucedidas de formação de elenco, onde a integração de jogadores de alta habilidade e a construção de relações táticas sólidas são priorizadas. Logo, a performance dos clubes no mundo real parece influenciar diretamente a configuração do jogo, com atributos dos jogadores ajustados para espelhar suas contribuições reais, promovendo uma experiência imersiva, porém amplificada no âmbito das jogadas de ataque e da emoção dos gols, que ressoa no entretenimento dos usuários e fãs do esporte.

VI. CONCLUSÃO

Este estudo revelou como a aplicação da teoria de grafos e análise de comunidades pode desvendar as complexas interações entre jogadores em um contexto de futebol virtual. A investigação detalhada das redes de jogadores destacou a centralidade dos meios e a influência tática dos atacantes, refletindo a realidade do futebol e as expectativas dos usuários em termos de jogabilidade e entretenimento. Os padrões de conexão entre jogadores de clubes renomados demonstraram a tendência dos desenvolvedores de jogos em espelhar a sinergia e a competência tática dos times reais. A análise sugere que os jogos de futebol virtual não apenas procuram recriar a autenticidade do desporto, mas também amplificam aspectos cruciais como a criação de jogadas e o espetáculo dos gols para enriquecer a experiência do usuário.

Embora os resultados obtidos tenham corroborado as expectativas iniciais e validado a eficácia da abordagem proposta, é crucial ressaltar a necessidade de estudos futuros que explorem uma comparação mais aprofundada dos dados obtidos a partir do banco de jogo de futebol virtual com a realidade do esporte. A validação dessas interpretações em cenários de jogo reais proporcionaria uma compreensão mais robusta e generalizável das dinâmicas identificadas, fortalecendo ainda mais a aplicabilidade prática das metodologias propostas.

REFERÊNCIAS

- [1] Newman, M. E. J. (2010). Networks: An Introduction. Oxford University Press.
- [2] Python Software Foundation. Python Language Reference, version 3.8. Disponível em: <https://www.python.org>
- [3] Kaggle. EA Sports FC 24 complete player dataset. Disponível em: <https://www.kaggle.com/datasets/stefanoleone992/ea-sports-fc-24-complete-player-dataset/>. Acesso em: 27 nov. 2023.
- [4] ROSEN, Kenneth H. Matemática discreta e suas aplicações [recurso eletrônico] / Kenneth H. Rosen; tradução técnica: Helena Castro, João Guilherme Giudice. – 6. ed. – Dados eletrônicos. – Porto Alegre : AMGH, 2010.
- [5] Celzin. Mapeando Rendimento no Futebol Virtual. Disponível em: <https://github.com/celzin/Mapeando-Rendimento-no-Futebol-Virtual>.
- [6] SANTOS, Francisco José Marques dos. Avaliação do desempenho de equipes de futebol profissional: Análise das interações entre jogadores. [S.l.]: [s.n.], 2022. Disponível em: <http://hdl.handle.net/10362/134203>. Acesso em: 27 nov. 2023.

- [7] RODRIGUES, Daniele Cristina Uchôa Maia. Complex network measurements in graph-based spatio-temporal soccer match analysis: Medidas de redes complexas na análise espaço-temporal baseada em grafos de jogos de futebol. 2017. - Universidade Estadual de Campinas, Instituto de Computação, Campinas, SP. Disponível em: <https://hdl.handle.net/20.500.12733/1632155>. Acesso em: 27 nov. 2023.
- [8] CUNHA, Sofia Alexandra Grifo. Data Analytics in the Football Industry: The Portuguese Football Federation Case. [S.l.]: [s.n.], 2022. Disponível em: <http://hdl.handle.net/10362/135781>. Acesso em: 27 nov. 2023.