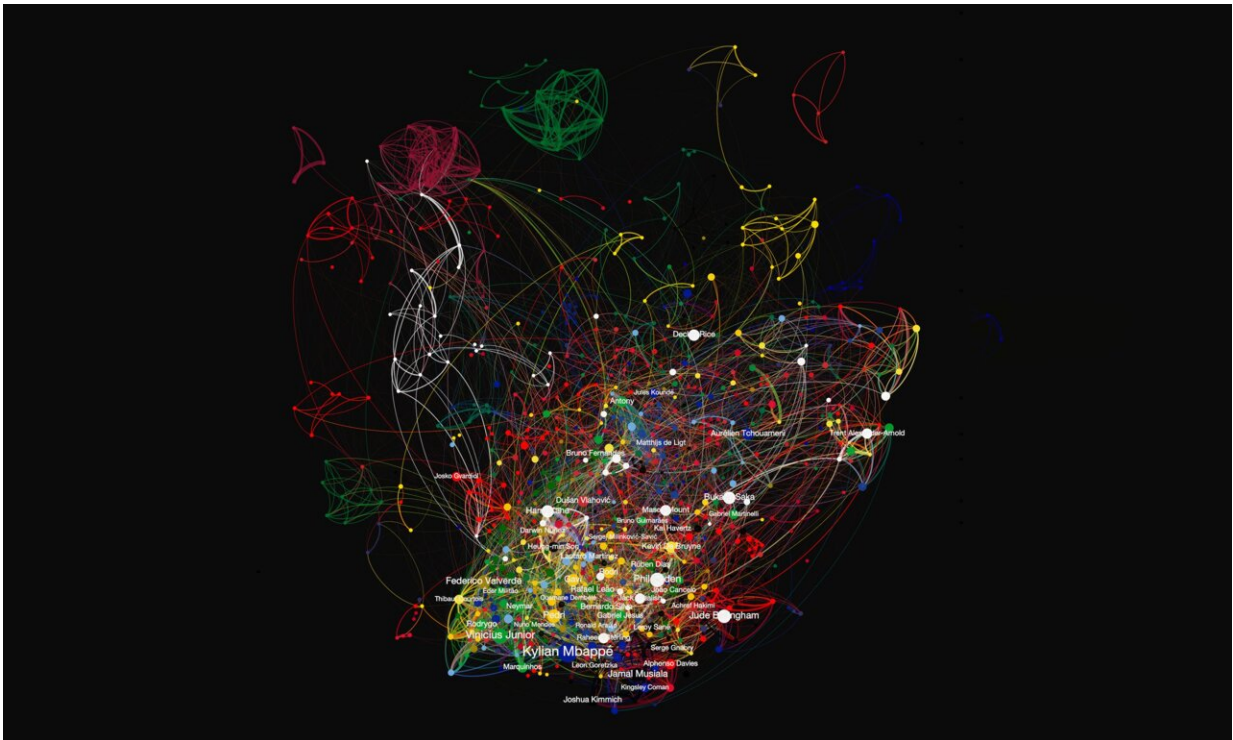# Exploring the FIFA World Cup 2022 using network science

February 2 2023, by Ingrid Fadelli

Credit: Milan Janosov and Patrik Szigeti.

Network science is the study of physical, biological, social and other phenomena through the creation of network representations. These representations can sometimes offer very valuable insight, unveiling interesting patterns in data and relationships between connected entities.

Milán Janosov and Patrik Szigeti, two [data scientists](#) working at Central European University, Baoba Inc. and Revolut recently used [network science](#) to examine the FIFA World Cup 2022. The network representations they created, outlined in a paper published on Research Gate, allowed them to shed some new light on the fascinating interconnected world of [soccer](#) stars and clubs.

"I am not a big soccer fan, so I haven't been closely following the recent FIFA championship either," Janosov told Tech Xplore.

"However, in my experience, network science and network visualizations are superb in summarizing and explaining complex systems in one image in a quick and objective way. So, I started wondering how much soccer knowledge I could pick up from one network, by answering questions like, who are the key players, and how does the whole ecosystem of soccer stars look like?"

"I thus reached out to my data scientist friend and co-author of this paper, Patrik Szigeti, who is also a self-educated soccer expert, to brainstorm about how we could build this network."

A network is essentially an object that consists of several nodes and links that connect these nodes. Network scientists like Janosov build these networks using data that relates to specific phenomena involving different interconnected parties or entities.

"To build a network, we need a [data source](#) that shows relationships between the entities we are studying," Janosov explained. "In the example of soccer, this could be a team just as much as individual players. So, first things first—we needed data. This is where expert knowledge is required, which led us to the transfermarkt.com website."

Janosov and Szigeti collected the data necessary to build their FIFA

World Cup 2022 Networks from transfermarkt.com, a soccer-related website owned by Axel Springer SE. This website contains a vast amount of information about soccer players and clubs, including players' team memberships and transfer histories, as well as both ongoing and past championship results.

"This kind of data is very much relational, so it can serve as a perfect input for network science—not surprisingly, building the team membership network of players was pretty straightforward," Janosov said. "Basically, if two players were on the same team during the same year, we considered them linked."

The player network representation created by the researchers contained 830 players, who were found to be linked to each other through approximately 6,400 past or present teammate relationships (i.e., they had been or were currently playing for the same team). The so-called average path length was 3, which essentially means that if two players were randomly selected from all of those playing in the FIFA World Cup, they would most likely both have had teammates who played for the same club at some point.

"Building a network of soccer clubs proved to be trickier," Janosov explained. "In this case, we wanted to capture which clubs are the main centers of gravity based on the typical directions in which players sign on and off."

To visually represent soccer clubs, Janosov and Szigeti extracted the club history of individual players and then organized links in their network to follow each player's unique professional path (i.e., from which teams and to which teams they transferred). This allowed them to uncover some very interesting patterns.

"One of the most exciting observations we gathered looking at the club

network is that, apparently, there are two main groups of clubs: spenders and mentors," Janosov said. "Mentors usually acquire players at an earlier stage of their career and for less cash, then sell them later for the big bucks. Spenders, on the other hand, reverse the flow of money and talent, and instead of training their own, they simply spend a tremendous amount on hiring stars."

Ultimately, Janosov and Szigeti also tried to use machine learning to predict some of the final World Cup rankings. Their model's predictions were not particularly accurate, as it yielded a modest accuracy of 60%. Nonetheless, their analyses identified some of the most relevant features when trying to predict a [club](#) ranking, such as the current market value of its players.

"While we are certainly not bookies, we gave a shot at trying to explain some parts of the final ranking of the World Cup using a simple machine learning model, fed with the network characteristics and market value of the players," Janosov said. "This analysis didn't yield particularly high accuracy, yet it was certainly cool to see that network features had significant predictive power as compared to the financials."

This recent study is merely the latest of a series of works by Janosov that looked at popular culture phenomena, including TV-series and movies, through the lens of network science. His efforts show just how valuable this emerging field of research can be when trying to better understand the connections underpinning [social phenomena](#).

"In my next works, I plan to use network science to dive deeper into sustainability," Janosov added.

© 2023 Science X Network

Citation: Exploring the FIFA World Cup 2022 using network science (2023, February 2) retrieved 17 April 2024 from https://techxplore.com/news/2023-02-exploring-fifa-world-cup-network.html