

PassFlow: Pass Route Visualizer

TEAM 096

Enoch Anim-Koranteng | Serdar Aydinoglu | Jonathan Eaton
Wayne Fong | Vigneshwar Perumal | Jackson Schieber

OVERVIEW

- Soccer (a.k.a. football) is the most popular sport in the world
- Teams have huge pressure and incentive to refine every aspect of their performance
- The most prevalent and important action in the sport is passing
- Current methods focus on simplistic analytics like number of assists or pass completion percentage
- More complex analytics content is often abstract and bereft of visualization or application
- A sophisticated, accessible analysis can have huge impact on coaching and training preparation
- PassFlow allows a user to easily, interactively, and effectively visualize goal-to-goal passing routes for the entire field with weighted probabilities

DATA

- PassFlow is built around StatsBomb's open source soccer data
- Python's statsbombpy package enables us to pull the data for 17 of Spain's La Liga seasons for all FC Barcelona matches
- Events data includes all events in a match, including dribbles, passes, shots, tackles, and more
- Each event includes the involved player's location at the time
- We analyze the most common starting 11 players for the season, making sure critical positions are represented
- We create a graph based on the average location of each player, with weighted directed edges representing passes made and received



LaLiga

APPROACH

- We aim to find the most likely passing route from the goalkeeper to the opponent's goal
- The 11 players on the field are nodes with the opponent's goal as the 12th node and the goalkeeper as node 1
- Pass probabilities are calculated using a Markov chain model, with a state transition matrix that is a 12 by 12 probability matrix representing the probability of a pass between 2 nodes
- All shots are considered a "pass" to the goal

- Markov chains are a well established method to compute probabilities in a graph
- We use the power iteration method to determine the state transition probabilities for each step of the passing sequence
- Each iteration step identifies the next node with the highest probability of having the ball, which becomes the initial node for the subsequent step
- Results are visualized using Python, HTML5 and D3.js



RESULTS



- The visualization allows the user to choose which of the 17 seasons to view
- User can choose to block a node, and the app calculates a new most probable route excluding that node
- Visualization adjusted and refined from peer usability tests
- No defined measure for model accuracy, but the model falls within reasonable expectation of people familiar with soccer

