



UNIVERSIDAD
DE GRANADA

PRÁCTICA 4

CASO PRÁCTICO DE ANÁLISIS
Y EVALUACIÓN DE REDES

MAURICIO LUQUE JIMÉNEZ

78004003D

MAULUJIM@CORREO.UGR.ES

17 DE DICIEMBRE DE 2023

REDES Y SISTEMAS COMPLEJOS

ESPECIALIDAD SISTEMAS DE INFORMACIÓN

4º GRADO EN INGENIERÍA INFORMÁTICA

PREGUNTA DE INVESTIGACIÓN

Durante los últimos años, el fútbol, así como el resto de ámbitos de nuestro día a día, ha visto cómo la tecnología y el uso de datos eran elementos más que comunes y a los que todos estamos habituados. En este caso, cada vez hay más estadísticas sobre todo lo que pasa en el campo, de manera que todo queda registrado y todo se puede analizar posteriormente. A este proceso de renovación se somete en esta investigación el Real Madrid, equipo que durante los últimos años ha impuesto una hegemonía a nivel europeo como no se veía desde hace más de cincuenta años. Ese equipo que dominó los años diez, aunque ya se ha reinventado y sienta su base sobre otros jugadores, todavía puede ser estudiado de muchas formas distintas; entre ellas, mediante el uso de redes. ¿Se puede intentar explicar el triunfo del Real Madrid de los Cristiano Ronaldo, Karim Benzema, Sergio Ramos y compañía? Esa es la pregunta que se va a intentar desarrollar aquí: mediante el estudio de redes, se va a intentar conocer qué patrones caracterizaban al equipo entrenado por aquel entonces por Zinedine Zidane (aunque también se estudiarán partidos previos a esa etapa) para conocer la importancia de algunos jugadores que actúan como elementos claves en la red.

DESCRIPCIÓN DEL CONJUNTO DE DATOS Y DE CÓMO SE HA OBTENIDO.

Para este estudio, se van a analizar las redes de pases de ciertos partidos concretos que supusieron puntos de inflexión (o simplemente momentos importantes) en los años en los que el Real Madrid aumentó su leyenda. De esta manera, las redes que nos vamos a encontrar son de la siguiente manera: son redes de tamaño pequeño, puesto que sólo cuentan con once nodos, que representan a los once jugadores con más minutos en los partidos en cuestión. Los enlaces entre los nodos serán los pases que se dieron los jugadores durante el transcurso del partido (estos enlaces serán dirigidos, puesto que es importante saber quién da un pase y quién lo recibe), y el tamaño de ambos (nodos y enlaces) representará el volumen de acciones de éstos durante el juego. Por ejemplo, si Cristiano Ronaldo dio muchos más pases que el resto, el nodo que le represente tendrá un tamaño mayor. De la misma manera, si Toni Kroos le dio muchos pases a Luka Modric durante un partido, su enlace entre ellos será mayor que el resto, aunque no tengan por qué ser los jugadores que más pases dieron.

Estas redes de pases se han construido a partir de datos proporcionados por la empresa StatsBomb, especializada en la ciencia de datos orientada al deporte. Estos datos se pueden encontrar gratuitamente en el repositorio de GitHub de la propia StatsBomb ([enlace en la bibliografía al final de esta práctica](#)). Siguiendo un tutorial que se adjunta en la bibliografía (pensado para generar redes de pases mediante imágenes, no archivos de grafo), se han leído los datos de pases (emisor y receptor, por decirlo de alguna manera) entre los jugadores y según el número de veces que aparezca ese mismo pase, se ha incrementado el peso del enlace en la red, a la vez que el tamaño de los nodos que formen ese enlace. Una vez se ha conseguido eso, se ha generado un grafo que se ha pasado a un fichero legible por Gephi, en el que se han transformado los datos para que acepte el peso como una medida externa ilegible desde el código fuente.

VALORES DE LAS MEDIDAS GLOBALES DE LA RED

Nombre de la red	Número de nodos N	Número de enlaces L	Número máximo de enlaces L_{max}	Densidad $D=L/L_{max}$	Grado medio $\langle k \rangle$	Diámetro d_{max}	Distancia media d	Coefficiente medio de clustering $\langle C \rangle$	Nº componentes conexas	Nº nodos componente gigante N_g	% nodos componente gigante $\%N_g$	Nº aristas componente gigante L_g	% aristas componente gigante $\%L_g$
Real Madrid - Barcelona, Liga 2014/15	11	52	110	0,4727272727	4,727	2	1,055	0,473	1	11	100,00 %	52	100,00 %
Barcelona - Real Madrid, Liga 2014/15	11	52	110	0,4727272727	4,727	2	1,055	0,476	1	11	100,00 %	52	100,00 %
Real Madrid - Barcelona, Liga 2015/16	11	54	110	0,4909090909	4,909	2	1,018	0,491	1	11	100,00 %	54	100,00 %
Real Madrid - Atlético de Madrid, Liga 2015/16	11	52	110	0,4727272727	4,727	2	1,055	0,476	1	11	100,00 %	52	100,00 %
Barcelona - Real Madrid, Liga 2015/16	11	51	110	0,4636363636	4,636	2	1,073	0,467	1	11	100,00 %	51	100,00 %
Real Madrid - Atlético de Madrid, Champions League 2015/16	11	52	110	0,4727272727	4,727	2	1,037	0,473	1	11	100,00 %	52	100,00 %
Real Madrid - Juventus, Champions League 2016/17	11	53	110	0,4818181818	4,818	2	1,018	0,484	1	11	100,00 %	53	100,00 %
Real Madrid - Barcelona, Liga 2017/18	11	51	110	0,4636363636	4,636	2	1,073	0,466	1	11	100,00 %	51	100,00 %
Real Madrid - Liverpool, Champions League 2017/18	11	54	110	0,4909090909	4,909	2	1,018	0,491	1	11	100,00 %	54	100,00 %

Una vez tenemos las redes creadas, antes de ir caso por caso, podemos analizarlas generalmente, ya que vemos que sus medidas globales son prácticamente idénticas independientemente del partido. Esto es ciertamente lógico, puesto que el número de nodos (jugadores) nunca cambia y la manera de jugar del equipo, dado que la marcan sus futbolistas, no puede ser extremadamente diferente de un día para otro. Analizando algunas medidas concretas, vemos que el grado medio suele estar entre 4'5 y 5, lo que significa que, de media, un jugador pasa el balón a otros cuatro o cinco jugadores. Entendiendo el reparto de los jugadores sobre el campo y que se relacionan entre ellos con pases cortos, es algo lógico. El portero se relaciona con sus cuatro defensas (partimos de que el Real Madrid juega, de base, con cuatro defensas, tres o cuatro centrocampistas y tres o dos delanteros, respectivamente), los defensas se relacionan con los laterales y centrocampistas de su zona, y éstos últimos se relacionan con los tres delanteros, aparte de los pases que realicen hacia compañeros de su misma zona o más retrasados (siendo el Real Madrid un equipo ofensivo, es lo menos habitual). También vemos que el diámetro de la red es de 2 enlaces, lo que puede suponer que desde el portero al delantero más avanzado sólo hace falta pasar por un centrocampista, lo cual es factible. Al igual que antes se ha mencionado el grado medio, el coeficiente de clustering está en unas medidas similares, puesto que los vecinos de un nodo realizan la mitad de todos los posibles enlaces entre sí. Al final, en un equipo que tiene tanto el balón, es muy difícil que dos jugadores que no se junten en la misma zona del campo no se pasen el balón ni una vez. Si miramos al lateral izquierdo, resulta coherente que el centrocampista y central izquierdo conecten muy a menudo, porque es uno de los pases fáciles que evita perder el balón. Curiosamente, el coeficiente medio de clustering tiene valores prácticamente idénticos a la densidad media de la red, ya que, al ser una red tan pequeña, los vecinos de un nodo suelen ser aquellos con los que más interactúa, por lo que no intenta hacer enlaces con compañeros más lejanos.

BIBLIOGRAFÍA Y REFERENCIAS

- Fuente de datos: repositorio de GitHub de StatsBomb
 - <https://statsbomb.com/what-we-do/hub/free-data/https://github.com/statsbomb/open-data/tree/master>
- Consulta para generar red de pases a partir de un conjunto de datos de StatsBomb usando Python
 - <https://medium.com/@yogakrisanto1129/a-step-by-step-guide-to-using-python-to-create-football-passing-networks-e00e92fec99>
- Consulta para construir una red en formato Gephi a partir de un grafo creado en Python
 - <https://www.geeksforgeeks.org/saving-a-networkx-graph-in-gexf-format-and-visualize-using-gephi/>
 - <https://towardsdatascience.com/from-dataframe-to-network-graph-bbb35c8ab675>