

SIMULACIÓN Y ANÁLISIS DE INTERACCIONES ENTRE JUGADORES EN UN PARTIDO DE FÚTBOL.

Muñiz Morales Erick
Guzmán Morales Lázaro Abraham

22 de mayo de 2020

Resumen

En este trabajo se analizará el deporte de fútbol por medio de un sistema complejo, en función de dos torneos de relevancia mundial, uno varonil y otro femenino; partiendo de las formaciones más comunes usadas actualmente y simulando el movimiento del balón entre los jugadores para construir una red de pases, sobre la cual se calcularán diversas métricas con el objetivo de establecer criterios cuantificables para elegir que formaciones aprovechan de mejor manera a sus jugadores, haciendo una distinción en el modo de juego de ambos sexos.

Palabras clave: Pases, formaciones, *closeness*, *betweness*, posición, flujo.

INTRODUCCIÓN.

El hombre nació para los juegos. Un ejemplo claro está en los deportes que dan origen a eventos masivos de alcance global como el Mundial de Fútbol; Kurt Badenhausen, escritor de la revista *Forbes*, en 2018 estimó que la candidatura en conjunto de Canadá, México y Estados Unidos, anfitriones de la Copa Mundial de la Federación Internacional de Fútbol Asociación (*FIFA*) en 2026, generarán más de 14 mil MDD en ingresos hacia ellos y 11 mil MDD en ganancias para la *FIFA*. A raíz de esta relevancia internacional ha nacido un interés en aplicar el conocimiento científico para mejorar el desempeño de los equipos en torneo cruciales por medio de modelos matemáticos.

En un partido de fútbol participan dos equipos de 11 jugadores, estos deportistas compiten por la posesión de un balón con la finalidad de introducirlo en la portería rival. Para tener un mejor funcionamiento, cada equipo adopta formaciones que ayudan a explotar los atributos de sus jugadores que se encuentran distribuidos por todo el campo de juego.

En este trabajo analizaremos un partido de fútbol desde la perspectiva de un sistema comple-

jo: (i) para un correcto funcionamiento del equipo las posiciones actúan de forma interdependiente, si un defensa está mal posicionado puede que el equipo rival encuentre espacios para anotar gol, o un delantero mal colocado no aprovechará las jugadas creadas por el equipo; (ii) cada jugador de forma independiente debe de saber como actuar en su posición; (iii) la emergencia se encuentra cuando los jugadores se conocen muy bien entre si, son capaces de fabricar jugadas complejas ¹ gracias a que existe una sinergia entre ellos; (iv) a lo largo del partido, la formación se debe de mantener todo el tiempo posible, de forma autorganizada. En virtud de los cuatro puntos anteriores, el fútbol actúa como un sistema complejo que se puede modelar por medio de Teoría de Redes.

En este trabajo cada jugador se representa como un nodo y la relación que se establece entre ellos es: el jugador A se conecta con B si su posición en el campo está a lo más en 25 metros de distancia con B, notemos que no hay pesos ni dirección en los

¹Un ejemplo es el famoso *Tiki Taka*, estilo de juego fundamental en la obtención de 6 trofeos de relevancia mundial para el equipo F.C. Barcelona en la temporada 2009-2010 de la liga española

enlaces en esta red. Cada plantilla que analizaremos tiene una red que se construye de forma natural por las posiciones de cada jugador en el campo

Por medio de un modelo que hemos construido para simular el movimiento del balón en cada formación se obtendrá una red de pases, la cual sigue manteniendo a los jugadores como nodos pero ahora sus relaciones se dan gracias a las trayectorias que sigue el esférico. En esta segunda red se tienen pesos y aristas dirigidas, entonces se aplicarán diversas métricas con la finalidad de establecer criterios cuantificables para decidir la elección de mejores plantillas, importancia de posiciones y jugadores clave.

ANTECEDENTES.

Existen tres zonas fundamentales en un campo de fútbol, el área: de defensa, medio campo y delantera, por lo que es común representar las formaciones como $1 - 4 - 3 - 3$, lo significa que hay un portero, cuatro jugadores en la zona defensiva, tres jugadores en mediocampo y tres en la parte delantera, sin embargo, no siempre se ha tenido el mismo orden.

A través de la historia los jugadores se han distribuido en la cancha de forma distinta dependiendo del físico estándar entre los deportistas, en un principio era común encontrar equipos con dos defensas y hasta cinco delanteros, con un estilo de juego más compacto pues la complexión de los jugadores era más robusta sin tanta velocidad, pero con el tiempo los jugadores se volvieron más veloces, llegando al punto de tener mucho juego por las bandas, lugar en el que hay pocas personas y así recorrer grandes distancias en la cancha de forma rápida.

La inclusión de las mujeres en los torneos de relevancia mundial ha abierto las oportunidades a un modo de juego distinto al varonil, por ejemplo, en promedio las porterías son más bajas de altura que los porteros así que se ha observado una cantidad mayor de goles desde media cancha en el fútbol femenino a comparación del masculino, esto se ve reflejado en la cantidad de goles que anotan las mediocampistas.

Por ello, el modelo que se planteará más adelante busca resaltar estas diferencias.

MARCO TEÓRICO

En 2010 se disputó la Copa Mundial de Fútbol en Sudáfrica, dejando una gran cantidad de datos

estadísticos como las asistencias y goles por cada jugador, la cantidad de pases que se hacen en un partido y la dirección que tienen, lo novedoso de todo esto es que ahora esas bases de datos se volvieron de dominio público por lo que, dos años después, los matemáticos Javier Lopez Peña y Hugo Touchette elaboraron una red de pases a partir de los datos recuperados del mundial anterior, el objetivo que tenían era hacer una representación cuantificable del estilo de juego de cada país participante.

Usaron la distribución de pases para construir una red con pesos y dirigida, en donde cada nodo corresponde a un jugador y el peso de las flechas es el número de pases exitosos entre los jugadores. Sobre ella se analizaron las medidas de centralidad como: *closeness*, *betweenness*, *Pagerank* y *clustering* con el objetivo de explicar el desempeño del equipo y jugadores de forma numérica. Las medidas antes mencionadas fueron modificadas para el uso exclusivo de su investigación.

A partir de la idea de Peña y Touchette hemos decidido aplicar un análisis parecido en dos torneos de relevancia mundial:

1. Liga de Inglaterra. También conocida como *Premier League*, es considerada la liga más competitiva del mundo. El hecho anterior se debe en gran medida a que tiene una regla especial, la cual dice que un jugador extranjero solo puede ser fichado por un equipo inglés si participa de forma activa en su selección nacional, esto provoca que los jugadores tengan un gran nivel de juego.

2. El Mundial Femenil de Fútbol, evento que ha cobrado más relevancia global en los últimos años. En este torneo juegan las mejores jugadoras de cada país, que a diferencia de los hombres, no tienen ligas locales de gran desempeño por la poca inversión que se tiene para esta área.

Con el fin de obtener criterios objetivos para el análisis de cada jugador, aplicaremos las siguientes medidas de centralidad:

1. *Closeness*, es la forma de cuantificar la manera en la que un nodo distribuye información de manera eficiente en una red, en nuestra representación sirve para identificar al jugador que más pases da.

2. *Betweenness*, sirve para cuantificar la relación de todos los caminos entre dos nodos y aquellos que pasan por un nodo i , esto ayuda a identificar el jugador por el que más pasa el balón.

3. *Pagerank*, ordena a los nodos en cuanto a relevancia que tienen en la red respecto a la incidencia

interna (por ser una gráfica dirigida) y el total de grados internos para cada nodo en la red. Aquí destacan los jugadores clave.

4. *Clustering*, es el coeficiente de agrupamiento; podemos asociarlo con los triángulos o las comunes ternas de compañeros con buena sinergia. Entre más grande sea el valor, el equipo tendrá un mejor desempeño.

Para hacer un análisis sobre el funcionamiento del equipo, hemos elegido usar las siguientes métricas:

1. *Conectividad de aristas*, es el número mínimo de aristas que hay que quitar para que la gráfica sea disconexa. Cuando queremos evaluar formaciones, nos gustaría una que logre unir a cada jugador de alguna manera.

Notemos que no es conveniente tomar *conectividad de nodo* (que es una idea análoga a la anterior, sustituyendo arista por nodo) ya que se espera que la gráfica sea conexa por los pases largos.

2. *Flujo máximo*; es la cantidad de flujo máximo que podemos llevar de un nodo A a B. La interpretación dada es ver la cantidad de pases máximos que puedo llevar desde el nodo "GK.^a" "GOAL", puesto que nos interesan los goles es válido hacer esta interpretación.

Hemos decidido incorporar una medida nueva que sirve para cuantificar que estilo de juego tuvieron los jugadores, lo definimos como un puntaje de efectividad definido como:

$$M_{df}(i) = \sum_{i \neq j} A_{i,j} - \sum_{i \neq j} A_{j,i}$$

Si un jugador A tiene un puntaje de efectividad negativo entonces fallo pases que dió, pues se espera que tenga un puntaje de cero, es decir, los pases que recibió son los mismos que devolvió. Si es positivo entonces el jugador movió el balón de una forma fluida y acertada.

Para el resto de medidas centrales usaremos las que vienen por defecto en la paquetería *NetworkX* para *Python*.

METODOLOGÍA

Primero hablemos acerca de la base de datos, fue recopilada por nosotros de la página web <https://fbref.com/en/> en donde hemos elegido la temporada 2018-2019 de la *Premier League* y el mundial de fútbol femenino disputado en 2018.

Para cada equipo que participa en los torneos mencionados usamos la formación que más usaron.

Por cada jugador tenemos dos tipos de estadísticas: Pases y Tiros. Por un lado, tenemos el número de intentos de pases a distancia media y larga, así como los acertados y errados, por otro, conocemos la cantidad de tiros, la cantidad de goles y el porcentaje de fallo del tiro.

A continuación explicaremos el modelo que construimos para simular las interacciones existentes entre los 22 jugadores de un partido de fútbol a partir de los datos conocidos, entre ellos, el intento de anotar goles y errar pases.

Los supuestos que hemos hecho es que no hay cambios de jugador, juegan los mismos por 90 minutos y el partido se simula por la cantidad de pases promedio que hay en un encuentro, que es de 600 en total. Ahora entremos de lleno en el modelo.

En cada formación de fútbol hay relaciones que se dan de forma natural por la estructura del campo de juego, la parte defensiva y de ataque se unen por el mediocampo, los defensas ubicados en la parte central juegan más con los defensas que se colocan cerca de las bandas y con los mediocampistas que tienen un rol defensivo, a su vez, los extremos que juegan por la bandas mandan pases para el delantero central con el objetivo de que anote goles. Analizaremos las siguientes formaciones: 1-4-2-2, 1-3-5-2, 1-4-5-1 y 1-4-3-3. Las formaciones se pueden ver en la Figura 1.

La relación de equipos con su respectiva formación está en el Cuadro 1.

La notación es: *GK* Goalkeeper, *LB* Left Back, *RB* Right Back, *CBL* Central Back Left, *CBR* Central Back Right, *DM* Defensive Midfielder, *AM* Attack Midfielder, *CML* Camp Midfielder Left, *CMR* Camp Midfielder Right, *LW* Left Winger, *RW* Right Winger, *FW* Forward, *FWL* Forward Left, *FWR* Forward Right.

Ahora que tenemos la red, debemos simular el movimiento del balón. Cada jugador debe de pensar que va a hacer cuando tenga el balón, así que el proceso mental que lo lleva a a decidir se puede modelar de la siguiente forma.

1. El jugador A tiene que decidir si hará un tiro a portería para intentar anotar un gol o dar un pase a compañero. Las estadísticas que poseemos son los tiros a portería T y los pases P que efectuó a lo largo de cierto período de tiempo, el total de veces que movió el balón es $M = S + T$.

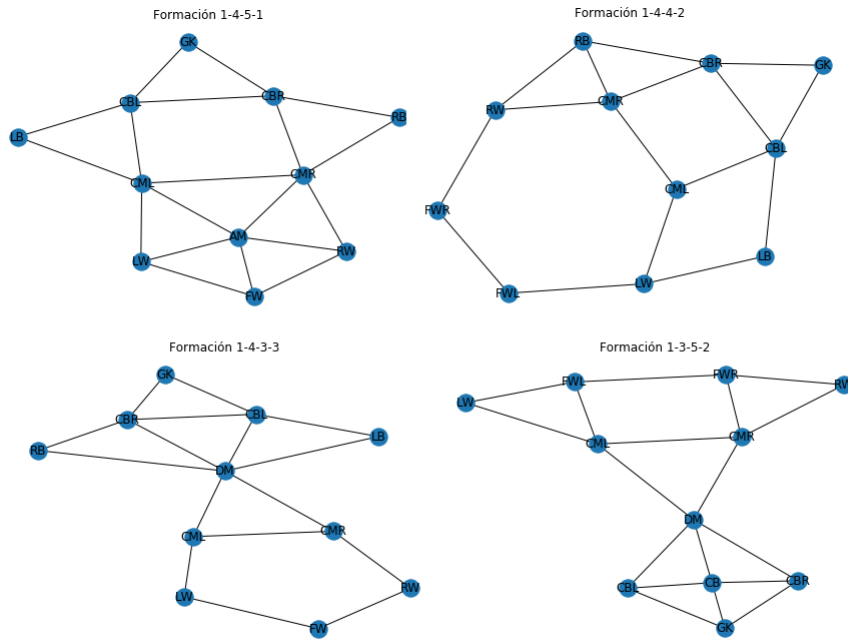


Figura 1: Redes que sirven como base para mover el balón y así construir la red de pases

Equipo	Formación	Equipo	Formación
ManchesterCity	1433	Argentina	1433
Liverpool	1433	Australia	1433
Chelsea	1433	Brasil	1442
Tottenham	1451	Camerun	1433
Arsenal	1451	Canada	1442
ManchesterUnited	1433	Chile	1433
Wolverhampton	1352	China	1442
Everton	1451	Inglaterra	1433
Leicester	1433	Francia	1442
WestHam	1451	Alemania	1442
Watford	1442	Italia	1433
CrystalPalace	1442	Jamaica	1433
NewCastle	1352	Japon	1442
Bournemouth	1442	Corea	1451
Burnley	1442	Holanda	1451
Southampton	1352	NuevaZelanda	1352
Brighton	1433	Nigeria	1433
Cardiff	1451	Noruega	1442
Fullham	1451	Escocia	1451
Hundersfield	1433	Sudafrica	1433
		España	1433
		Suecia	1451
		Tailandia	1433
		EUA	1433

Cuadro 1: Equipos y sus formaciones.

1.1. Añadimos un nuevo proceso (ver Figura 2). Cuando el jugador procede a tomar un pase, se cuenta la cantidad de rivales en su alrededor inmediato (nodos rojos). Puesto que solo hay 11 jugadores es válido dividirlo entre un número entero entre 1 y 11 para obtener una probabilidad. Haciendo una consideración de que la base nos da como un máximo de 5 enemigos, se procedió a dividirlo por este número para obtener la probabilidad deseada, así si alguien está completamente rodeado, pierde el balón y se pasa a uno éstos. Y se repite el caso con éste desde el paso 1. En caso contrario, podemos pasar al paso 2.

2. Se simula un número uniforme entre (0,1), digamos α .

2.1. El jugador da un tiro a gol si α es menor a la cantidad $\frac{T}{M}$. Dependiendo de la efectividad respecto a tiros y goles por cada jugador se decide si se anota o no. En caso de que sea gol se registra y el balón pasa a la posesión del equipo contrario repitiendo el proceso desde el paso 1, ahora en un jugador del equipo rival. Si el tiro falla, se anota el error y se cede la posesión al equipo rival, regresamos al paso 1.

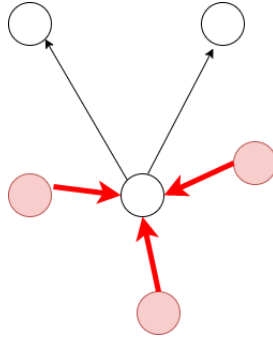


Figura 2: El nodo central tiene el balón. Ahora la proba de un pase efectivo disminuye conforme tenga más nodos contrarios rojos.

2.2 Cuando se da el pase a un compañero, el jugador debe decidir si será un pase medio o largo. Se considera un pase medio cuando se le da a un compañero a una distancia máxima de 25m, de está forma, todos los jugadores que se encuentran en la primera vecindad de un jugador A pueden ser objetivos de pases medios, al complemento de este conjunto los denominamos como objetivos de pases largos. Conocemos las estadística de los pases largos P_l y medios P_m para cada jugador, así que la cantidad $\frac{P_l}{P}$ se distribuye uniforme entre los jugadores que están a una distancia larga y el número $\frac{P_m}{P}$ hace lo mismo pero para la primera vecindad. De está forma hemos emulado la manera en la que un jugador elige a quién pasar el balón.

2.2.1 Una vez que se elige el destino del pase, el jugador tiene cierta efectividad de pases, dependiendo de ésta se decide si el pase fue completado o interceptado. En caso de ser interceptado la posesión se pasa al equipo contrario, en el otro caso, la posesión se mantiene y se repite el proceso desde el paso 1.

El proceso anterior simula la manera en la que un jugador mueve el balón a lo largo de la cancha. Al final de la simulación se tiene una gráfica dirigida con pesos que corresponden a la cantidad de pases entre dos jugadores.

Notemos que el modelo anterior es muy flexible para simular partidos de cualquier liga, torneo o época, solo se necesitan los datos de pases medios y largos, tiros, efectividad de pases y goles.

El proceso fue implementado en *Python 3.7*. El código se puede encontrar en la siguiente liga:

github.com/ErickMM98/SMA.LGME_project

Posteriormente, se calculan las medidas de centralidad antes mencionadas bajo la siguiente notación: *Clustering* ($Clus$), *Pagerank* (PR), *betweenness* (δ_{bet}), *closeness* ($Clos$), *eigenvector centrality* (E_c), grado (δ), grado central (δ_c) y nuestra medida definida en el marco teórico M_{df} . En cuanto a medidas globales: *pases*: pases realizados (p), *Conectividad de aristas dirigidas* (E_c), *Conectividad de aristas no dirigidas* (E_{end}), *Promedio de Clustering*: el promedio del valor de los jugadores en su Clustering ($prom_{clus}$), *Promedio de Eigencentrality*: el promedio del valor de los jugadores en su eigenvalor ($prom_{eigen}$), *Flujo máximo* ($flujo$) y el promedio de los pases que se realizan por equipo antes de anotar un gol ($P2G$).

Notemos que el proceso que hemos explicado es un algoritmo para simular un partido entre dos equipos. Como tenemos dos competencias, las cuales no se pueden combinar por ser de distinto sexo, cada equipo tiene asociado 19 juegos y 23 juegos para masculino y femenino respectivamente. Por lo cual, la red final que es el molde para realizar las estimaciones es la unión de cada uno de los juegos; respetando los pases realizados y sumándolos.

Resultados

Mejor formación en hombres.

El criterio para clasificar si una formación es buena o no, consiste en identificar quien tiene los mejores valores en cada medida que hemos calculado respecto a los demás. Los datos se encuentran en el cuadro 2, el cual la siguiente división: *clus* es el promedio de los promedios de los clustering de los jugadores con esa formación, *edge_c* es la conectividad de aristas promedio, *edge_{nd}* es la conectividad

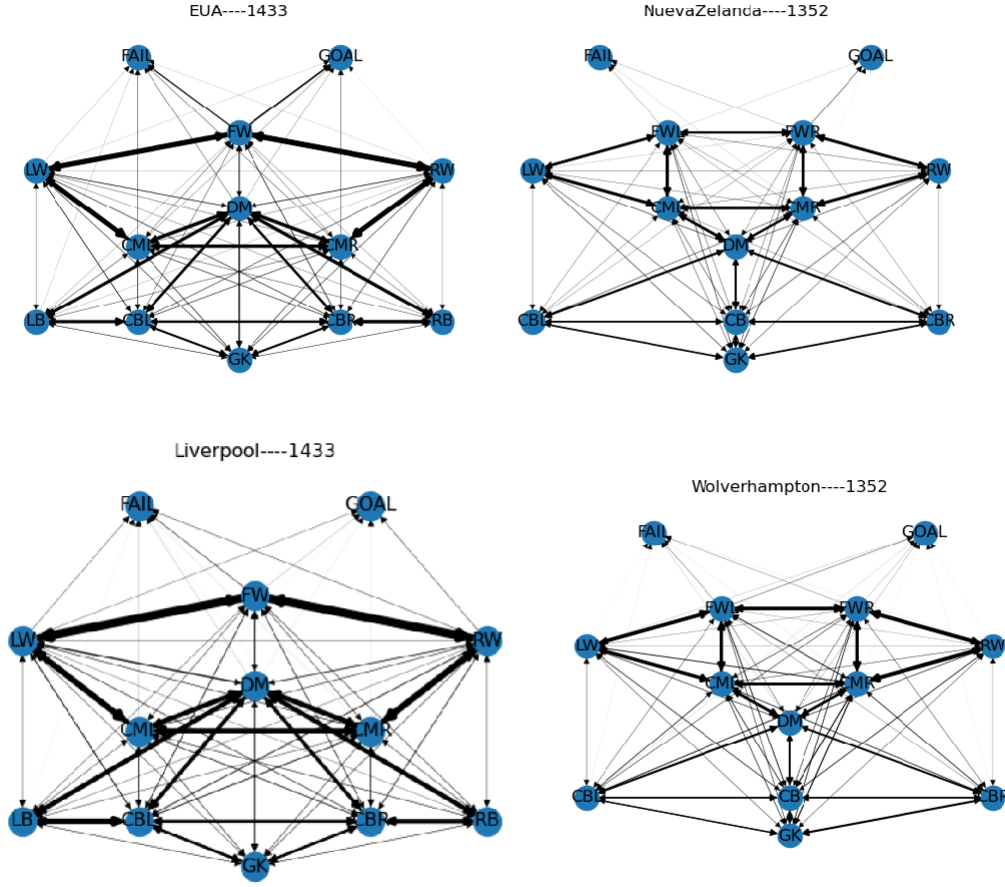


Figura 3: Ejemplos de redes de pases simuladas a partir del modelo explicado con anterioridad.

de aristas no dirigidas, *eigen* es el promedio de los promedios de los *eigencentralities* de los jugadores y la última columna (*P2G*) refleja el promedio de los pases realizados antes de anotar un gol.

Resulta que la formación 1–3–5–2 es la que tiene mejores valores en las medidas calculadas, pues de cinco categorías ocupa el primer lugar en dos y el segundo puesto en otros dos.

La formación 1-3-5-2 cumple con tener la menor cantidad de pases pero con un clustering alto en comparación con las otras (es la segunda mejor), lo cual significa que tiene muy buenas triagulaciones en sus jugadores. Además, el valor $edge_c$ es el mayor con lo cual es una red bastante conexas (se requieren al menos 10 aristas para hacerla disconexa) y es la formación que necesita, en promedio, menos pases para crear goles debido al flujo pequeño que tiene.

De los tres equipos varoniles que juegan esta formación (Wolverhampton, New Castle, Southampton), dos de ellos posiciones negativas (13ro y 14to), mientras que Wolverhampton se posiciona en 6to lugar, abajo de equipos en donde la gamma de jugadores es privilegiada (la plantilla de Wolverhampton vale menos de la mitad que las plantillas posicionadas por arriba); por lo que aprovecha al máximo a cada jugador que tiene.

Ahora hagamos un análisis más detallado en función de los valores del Cuadro 3 que representa los jugadores clave del equipo. Wolverhampton realizó juegos más ofensivos que defensivos; se nota en que el portero Rui Patricio obtuvo 0 en *betweenness centrality*, el mismo caso para sus defensas Conor Coady, y los medios Joao Moutinho y Rúben Neves con lo que deducimos que el juego se mantuvo

FORMACIÓN	EQUIPOS	H	M	Pases	clus	$edge_c$	$edge_{nd}$	$eigen$	flujo	P2G
1442	11	4	7	307.6363	0.0887	8.9090	5.6363	0.2522	59.2727	36.5259
1352	4	3	1	209.0	0.1270	9.75	5.25	0.2447	32.75	5.51
1451	10	6	4	302.4	0.1449	9.5	5.8	0.2490	51.4	12.8877
1433	19	7	12	278.94736	0.08405	8.3684	5.0	0.24323	63.7368	10.9280

Cuadro 2: Medida por formación .

en la parte media-alta de la cancha, además de que estos cuatro jugadores tienen de los mejores valores en *closeness*, lo cual significa que son muy buenos moviendo el balón; por la forma en la que construyó el modelo, la cantidad mencionada es equivalente a decir que fallan muy poco sus pases, pues sus enlaces son más pesados respecto a los demás.

Por estas justificaciones y considerando que el equipo Wolverhampton es un equipo mixto por el valor del equipo, concluimos que la formación 1-3-5-2 es eficiente para obtener un buen resultado en la rama varonil del fútbol sin necesidad de tener solamente jugadores de élite.

Mejor formación en mujeres.

Notemos que el caso femenino es distinto. El país Nueva Zelanda jugó con susodicha formación. Sin embargo, a pesar de tener esta plantilla ni siquiera logró clasificar en la fase de grupos quedando en último lugar. Esto ocurre porque el juego femenino tiende ser, en la mayoría de las veces, juegos en la parte media de la cancha.

Por lo que buscaríamos una formación ideal que se adapte al modo de juego femenino. Notemos que la mitad de los equipos femeniles optaron por jugar la formación 1-4-3-3, donde hay que recalcar que entre ellos se encuentran EUA e Inglaterra, campeón y 4to lugar respectivamente. Sin embargo, notemos que el segundo lugar es adquirido por la escuadra de Holanda con la formación 1-4-5-1.

Regresando al cuadro 2, vemos que la formación 1-4-3-3 obtiene una buena cantidad de pases aunque obtiene los demás valores bajos con respecto a las demás formaciones. Cabe señalar que la formación 1-4-5-1 es mejor en ese sentido, sin embargo sólo el equipo de Holanda obtiene un segundo lugar (lo demás ni siquiera clasificaron).

¿Por qué surge esta inconsistencia? Las diferencias radican en la eficiencia de las jugadoras; notemos que la jugadora coreana, Lee Min-ah, obtiene un valor de 0 en *betweenness* contrario a la holandesa Jackie Groenen con una cantidad no despreciable.

A la vez, que ésta misma formación radica la posición clave FW. La holandesa Lineth Beerens teyn obtiene un valor considerable de *betweenness* comparado con otra jugadoras de la misma formación. A su vez, notemos que no hay tantos ceros en la parte de *betweenness* con lo cual en la formación juegan todas. Además, superan en puntaje a esa posición. Por lo que, Holanda tenía una jugadora haciendo un papel de medio cuando en realidad era delantera.

Prestemos atención a la formación 1-4-3-3, los equipos que se estructuran de esta manera tienen una muy activa participación en media cancha y delantera, para EUA Alex Morgan obtuvo el valor máximo de *betweenness* en todas las jugadoras, la siguiente fue por parte de Kedeisha Buchanan de Camerún. De estas últimas dos escuadras, que juegan con la misma formación, EUA es campeón y Camerún ni logra clasificar. Sin embargo, comparten la peculiaridad que hay jugadoras que no juegan y sólo se vuelven 'paredes' para las jugadoras clave. Esto se ve reflejado en el Cuadro 5; 4 o 5 jugadoras no juegan y las que juegan necesitan valores estadísticos altos. Por lo que esta formación funciona cuando se tienen jugadoras estrellas como Alex Morgan y Megan Rapinoe.

Así, la formación 1-4-5-1 es funcional para cuando se obtiene un equipo mixto femenino. Mientras que la formación 1-4-3-3 es eficiente para un equipo con jugadoras estrella.

Diferencia marcada por sexo.

Prestemosle atención al Cuadro 6 y 7. Veamos que hay una diferencia marcada por el flujo. Recordemos que éste se puede interpretar como la cantidad máxima de pases que el equipo puede realizar desde el portero hasta el gol. Notemos que la parte masculina tiene un rango menor al femenino ya que el flujo mínimo lo reportó Everton y Fullham con 28 y el máximo con 72 en la escuadra de Leicester, mientras la rama femenino obtuvo como mínimo 17 por parte de Argentina y Corea, y el valor máximo

TEAM	POS	NAME	δ	δ_c	δ_{bet}	$Clos$	E_c	PR	Clus	M_{df}
Wolverhampton	CB	Conor Coady	20	1.66667	0.0	0.0646	0.17315	0.06788	0.15563	10
Wolverhampton	GK	Rui Patrício	20	1.66667	0.0	0.05519	0.14636	0.06181	0.13254	-76
Wolverhampton	CML	João Moutinho	20	1.66667	0.0	0.05208	0.41974	0.11945	0.20384	-27
Wolverhampton	CMR	Rúben Neves	22	1.83333	0.0	0.0499	0.42011	0.12324	0.18049	31

Cuadro 3: Jugadores clave en la formación 1-3-5-2.

TEAM	POS	NAME	δ	δ_c	δ_{bet}	$Clos$	E_c	PR	Clus	M_{df}
Holanda	CMR	Jackie Groenen	21	1.75	0.00758	0.03987	0.39913	0.11653	0.23716	-14
Holanda	FW	Lineth Beerensteyn	22	1.83333	0.16162	0.04006	0.32517	0.09404	0.15035	-41
Corea	CMR	Lee Min-ah	22	1.83333	0.0	0.04146	0.39635	0.11758	0.18359	8
Corea	FW	Jung Seol-bin	20	1.66667	0.0202	0.03805	0.3232	0.09236	0.16429	-26
Escocia	FW	Erin Cuthbert	22	1.83333	0.07576	0.04296	0.29568	0.0878	0.12969	5
Suecia	FW	Stina Blackstenius	22	1.83333	0.07323	0.04006	0.32066	0.09153	0.11958	1

Cuadro 4: Jugadores clave en la formación 1-4-5-1.

es 138 con Italia.

GEN	EQUIPO	FORMACION	flujo
H	ManchesterCity	1433	59
H	Liverpool	1433	48
H	Chelsea	1433	29
H	Tottenham	1451	56
H	Arsenal	1451	38
H	ManchesterUnited	1433	60
H	Wolverhampton	1352	46
H	Everton	1451	28
H	Leicester	1433	72
H	WestHam	1451	40
H	Watford	1442	47
H	CrystalPalace	1442	42
H	NewCastle	1352	33
H	Bournemouth	1442	55
H	Burnley	1442	42
H	Southampton	1352	24
H	Brighton	1433	33
H	Cardiff	1451	44
H	Fullham	1451	28
H	Hundersfield	1433	40

Cuadro 5: Flujo en hombres

M	Argentina	1433	17
M	Australia	1433	135
M	Brasil	1442	126
M	Camerun	1433	34
M	Canada	1442	47
M	Chile	1433	33
M	China	1442	34
M	Inglaterra	1433	102
M	Francia	1442	34
M	Alemania	1442	130
M	Italia	1433	138
M	Jamaica	1433	72
M	Japon	1442	36
M	Corea	1451	17
M	Holanda	1451	86
M	NuevaZelanda	1352	28
M	Nigeria	1433	21
M	Noruega	1442	59
M	Escocia	1451	83
M	Sudafrica	1433	51
M	España	1433	47
M	Suecia	1451	94
M	Tailandia	1433	97
M	EUA	1433	123

Cuadro 6: Flujo en mujeres.

Interpretar los valores mínimos y máximos nos revela más. El fútbol masculino, por experiencia, ha sido más establecido, estudiado y patrocinado, por lo cual muchas estrategias están destinadas a

TEAM	POS	NAME	δ	δ_c	δ_{bet}	$Clos$	E_c	PR	Clus	M_{df}
EUA	GK	Alyssa Naehner	20	1.66667	0.0	0.0651	0.14626	0.05412	0.09973	6
EUA	CBR	Abby Dahlkemper	20	1.66667	0.0	0.07062	0.22547	0.07883	0.12829	27
EUA	CBL	Becky Sauerbrunn	20	1.66667	0.0	0.06313	0.23094	0.0812	0.13016	72
EUA	FW	Alex Morgan	22	1.83333	0.21843	0.05274	0.34768	0.10099	0.07316	90
EUA	DM	Julie Ertz	22	1.83333	0.0	0.0672	0.42102	0.1314	0.15031	-28
EUA	CMR	Rose Lavelle	22	1.83333	0.0	0.05669	0.3651	0.09844	0.10057	0
EUA	CML	Sam Mewis	22	1.83333	0.0	0.06313	0.38192	0.10097	0.10598	7
EUA	LW	Megan Rapinoe	22	1.83333	0.07071	0.05593	0.33688	0.0887	0.08209	-8
Camerun	DM	Raissa Feudjio	20	1.66667	0.0	0.12626	0.44881	0.14207	0.16839	15
Camerun	CBL	Estelle Johnson	20	1.66667	0.0	0.09058	0.22365	0.08567	0.10031	-55
Camerun	CMR	Jeannette Yango	22	1.83333	0.17424	0.08681	0.37374	0.09589	0.08573	41
Camerun	CML	Ajara Nchout	21	1.75	0.0	0.07375	0.38224	0.09695	0.0983	58
Camerun	CBR	Aurelle Awona	19	1.58333	0.0	0.08091	0.24708	0.09399	0.12086	-61
Camerun	RB	Augustine Ejangue	20	1.66667	0.20076	0.0817	0.17891	0.05566	0.07111	83

Cuadro 5: Jugadores clave y jugadoras 'pared' en la formación 1-4-3-3.

que los goles ocurran a una cierta cantidad de pases. Caso contrario, y desafortunado para las mujeres. La gran distancia entre estos valores se justifica con el hecho de los goles ocurren en media cancha. Ésto por la mala posición de la portera y el tamaño promedio de las mismas. Y, a su vez, el valor máximo ocurre por la poca la experiencia que se tiene en el área femenil creando mala comunicación generando muchos pases 'innecesarios' antes de marcar gol, ésto es provocado por la poca inversión y apoyo para las ligas femeniles. Ésta última idea se puede visualizar en la gráfica 5; existe cierta 'dominancia' de las mujeres a que los goles se empiecen a anotar o en muy poco pases o en mucho pases.

Recalamos que esto no es una condición de sexo o género, más bien una condición de negocios, el fútbol femenino no es tan rentable como el masculino. Cosa que esperamos cambie en un futuro con la promoción de eventos importantes para el fútbol femenino que aumenten la inversión en ésta área.

Análisis del puntaje de efectividad.

En la Figura 4 podemos tenemos los histogramas en los cuales hay una diferencia considerable. Los hombres fallan más pases que las mujeres, la diferencia se observa en la distribución que tienen pues los hombres tienen una configuración más equilibrada en valores menores a cero pero las mujeres solo tienen una columna en negativos situado entre -40 y -35 , además las mujeres llegan a tener valores de efectividad mucho mayores que los hombres. Una razón de esta diferencia que podemos dar es

que en el fútbol masculino se presenta un juego más agresivo e interceptivo mientras que algunas jugadoras pueden dar pases de forma más eficaz por su gran técnica de juego, recalando la diferencia que puede provocar un solo jugador elite respecto a los demás.

Conclusiones

Para concluir este trabajo por el momento, debemos comentar que los datos simulados presentan consistencia con los registros que se tienen. Además de que las medidas de centralidad nos ayudan a entender más acerca del comportamiento que tienen los jugadores a lo largo de un partido identificando los puntos fuertes que tiene cada posición.

Por lo mencionado en los resultados, podemos concluir que la mejor formación varonil es la 1-3-5-2 para equipos mixtos y la 1-4-3-3 para jugadores estrella. Análogamente para la femenil la formación 1-4-5-1 es funcional para una formación mixta mientras que la 1-4-3-3 es funcional para jugadores estrella.

La inclusión de los partidos de la rama femenil igual nos enseña que hay una manera distinta de jugar a comparación de los hombres.

Trabajo futuro

Notemos que nuestro modelo es estático tanto en la dinámica de las formaciones como en las habilidades de los jugadores. Una posible continuación de este modelo podría ser agregar éste dinamismo a las formaciones haciendo variar la cantidad de 'adversarios' conforme pase el tiempo y el momento del partido; se podría abrir un análisis de cómo va se va mutando una formación inicial, empezando el segundo tiempo y, en ciertas ocasiones, cuando un equipo va ganando o perdiendo (Ésto último se puede reflejar como la resistencia física).

Otra posible continuación de éste modelo es agregar un nuevo criterio al momento de 'dar el pase' considerando un *autoaprendizaje* del partido, ya sea mejorando o empeorando la efectividad de un pase; esto es, que la probabilidad de errar sea alguna función que dependa de las estadísticas propias del portador, los adversarios y el momento del partido.

Referencias

- [1] LOPEZ PEÑA, JAVIER y TOUCHETTE, HUGO, *A network theory analysis of football strategies*, In C. Clanet (ed.), Sports Physics: Proc. 2012 Euromech Physics of Sports Conference, p. 517-528, Editions de l'Ecole Polytechnique, Palaiseau, 2013. (ISBN 978-2-7302-1615-9), 2012.
- [2] BULDU, J.M., BUSQUETS, J., ECHEGOYEN, I. y SEIRUL, F., *Defining a historic football team: Using Network Science to analyze Guardiola's F.C. Barcelona*, Scientific reports, 9(1), 1-14.
- [3] MIKE HUGHES, IAN FRANKS (2005): *Analysis of passing sequences, shots and goals in soccer*, Journal of Sports Sciences, 23:5, 509-514.
- [4] OLIVEIRA, A. P., & TYLER, H. R., (2015). *Measurement and comparison of passing networks in collegiate soccer*. Minnesota Journal of Undergraduate Mathematics, 1(1).
- [5] M. E. J. NEWMAN. *Networks : an introduction*. Oxford University Press, Oxford New York, 2010.
- [6] FBREF. (2019). 2019 Women's World Cup Stats. 29/03/2020, de Data Sports Group Sitio web: https://fbref.com/en/comps/106/Womens-World-Cup-Stats#all_stats_passing_squads
- [7] FBREF. (2019). 2018-2019 Premier League Stats. 29/03/2020, de Data Sports Group Sitio web: <https://fbref.com/en/comps/9/1889/2018-2019-Premier-League-Stats>

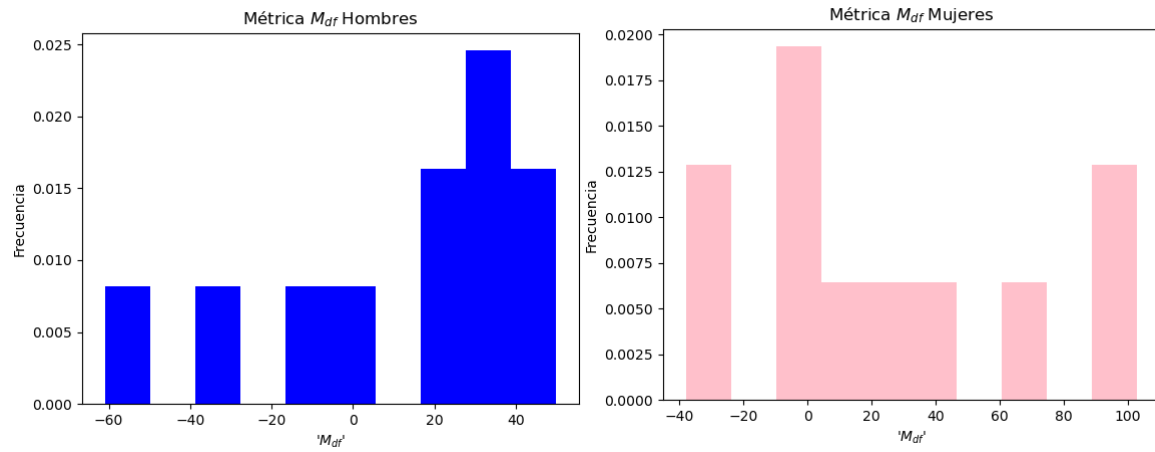


Figura 4: Comparación de histogramas de puntaje de efectividad en ambos sexos

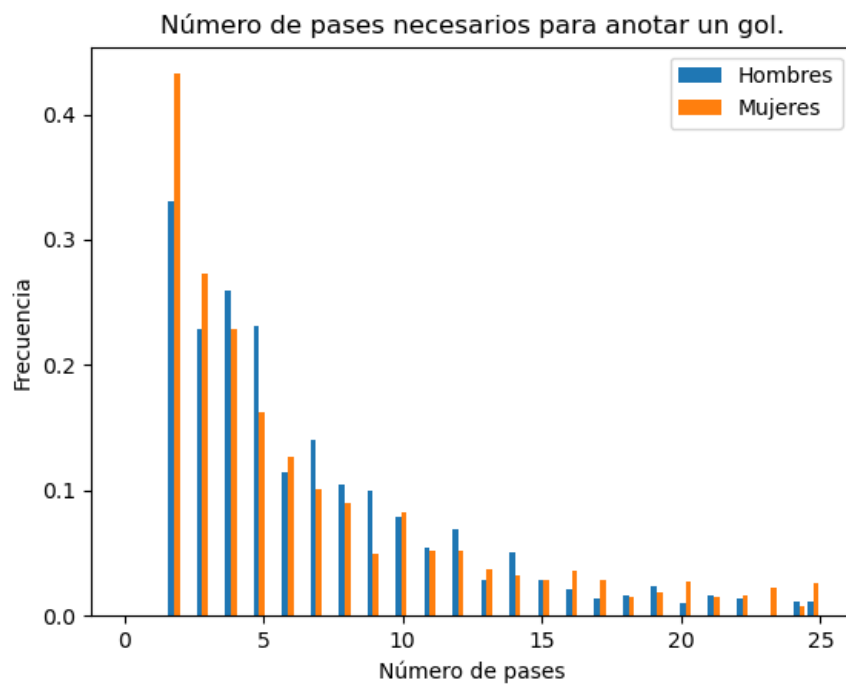


Figura 5: Número de pases antes de anotar un gol.