

Teoría de Grafos y Sudáfrica 2010

Rodrigo Alfonsín, Ricardo Monter, Armando Rocha, Pablo Franqueblame.

Colegio Marymount Estrella del Norte #6, col. Rancho Tetela, Cuernavaca
62160, Morelos

E-mail: colegio@marymount.edu.mx

Palabras clave:

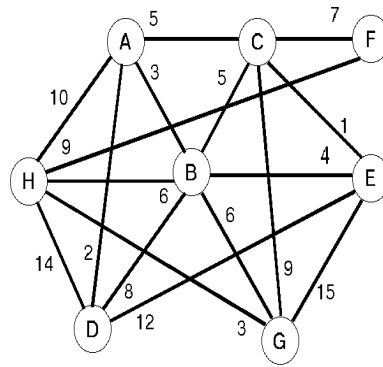
Teoría de grafos, Fútbol

Resumen

Se representó para cada partido de fútbol en la copa de Sudáfrica 2010 a los equipos participantes en ellos. Esto fue para poder calcular las eficiencias globales de ambos grafos (la definición de un *grafo* se encuentra más adelante en la parte de “Introducción”) y poder comparar si existe una eficiencia mayor en los equipos ganadores que en los perdedores. La eficiencia se define como el promedio de la suma de 1 entre el camino más corto para cada par de nodos. El camino más corto entre dos nodos es aquel que junta el nodo inicial con el final utilizando la menor suma de los valores de las aristas entre ellos (explicación más amplia de *eficiencia* en el capítulo de “Antecedentes”). Se utilizaron los datos que se encuentran en la página de la FIFA, los cuales muestran el número de pases entre cada par de jugadores dentro de un equipo para cada partido. En el grafo los nodos fueron representados con los jugadores de cada equipo y las aristas eran el número de pases entre cada par de jugadores. Estas fueron normalizadas con el número máximo de pases para ese equipo en ese partido. Se calcularon las eficiencias para cada partido y posteriormente se compararon las de los ganadores y las de los perdedores. Se obtuvo el promedio de las diferencias. Se simuló el Mundial de manera aleatoria (explicación en el capítulo de “Metodología y Materiales”) 10,000 veces con las eficiencias previamente calculadas y se obtuvo el promedio de las diferencias para cada una de las simulaciones. Posterior a esto se compararon el promedio de las diferencias del mundial de Sudáfrica con las de los mundiales simulados. Se llegó a la conclusión de que la eficiencia global no es un buen parámetro para medir el desempeño de un equipo ya que nuestro resultado no tuvo una diferencia significativa con respecto a los promedios de las diferencias de los mundiales simulados.

Introducción

Un grafo es un conjunto de dos tipos de objetos, los primeros llamados nodos, y los segundos, aristas (ver Figura 1₍₁₎). Los nodos son estados y las aristas muestran la relación entre éstos estados. Comúnmente un grafo es expresado mediante una serie de puntos (nodos) que están conectados por líneas (aristas). La teoría de grafos es una rama de las matemáticas especializada en el estudio de los mismos.



(Figura 1:
Ejemplo de un grafo con aristas ponderadas.)

Nosotros hicimos este proyecto porque el fútbol es un deporte que contiene muchos componentes aleatorios, por ejemplo, cuando un jugador se lesiona, hace una falta o el árbitro comete algún error de juicio por lo tanto encontrar alguna manera de predecir el resultado de un partido en tiempo real sería bastante complicado. Está es una buena opción ya que el método que explicamos más adelante puede ser actualizado en tiempo real conociendo el numero de pases para ese momento del partido.

Antecedentes

Encontramos un artículo⁽²⁾ en el que Javier López y Hugo Touchette (University of London) recopilaron y analizaron datos sobre los pases de balón en el mundial de Sudáfrica y ellos dicen que se puede predecir el ganador del mundial por medio de la teoría de grafos; sin embargo, no explican cómo llevaron a cabo el proyecto.

La eficiencia global de un grafo está definida como el promedio de la sumatoria de $1/C_{ij}$ donde C_{ij} es el camino más corto entre los nodos “i” y “j”. “i” y “j” van desde el nodo 1 hasta el nodo “n”. En este caso específico “n” es 10 que es el número de jugadores que estamos utilizando dejando fuera al portero ya que creemos que este no es un componente el cual llegue a modificar el flujo del balón dentro de un equipo.

Para calcular el camino más corto utilizamos el algoritmo de Floyd-Warshall lo que hace este algoritmo es encontrar el camino entre todos los pares de jugadores con sólo una ejecución. Por eso es que este algoritmo sirve para calcular el camino más corto entre cada par de nodos. El algoritmo utiliza 3 ciclos que viajan por todos los nodos desde el 1 hasta el n. El pseudocódigo es el siguiente.

```
Para k=1 hasta n
  para i=1 hasta n
    para j=1 hasta n
       $C_{ij} = \min(C_{ij}, C_{ik} + C_{kj})$ 
```

En el anexo1 se muestra el algoritmo de Floyd-Warshall aplicado a una serie de datos.

Objetivo

Determinar si la eficiencia global es un parámetro que sirva para medir el desempeño de los equipos de fútbol.

Metodología y Materiales

Nuestro proyecto fue realizado en la Facultad de Ciencias de la UAEM.

Utilizamos una computadora HP w1907. El lenguaje operativo que utilizamos fue Fortran y el compilador fue F77.

Usamos los datos de la cantidad de pases entre cada par de jugadores para cada partido la Figura2₍₃₎ muestra los datos para el partido de la final entre España y Holanda.

Final

Netherlands - Spain 0:1 a.e.t.

Johannesburg / Soccer City Stadium / RSA

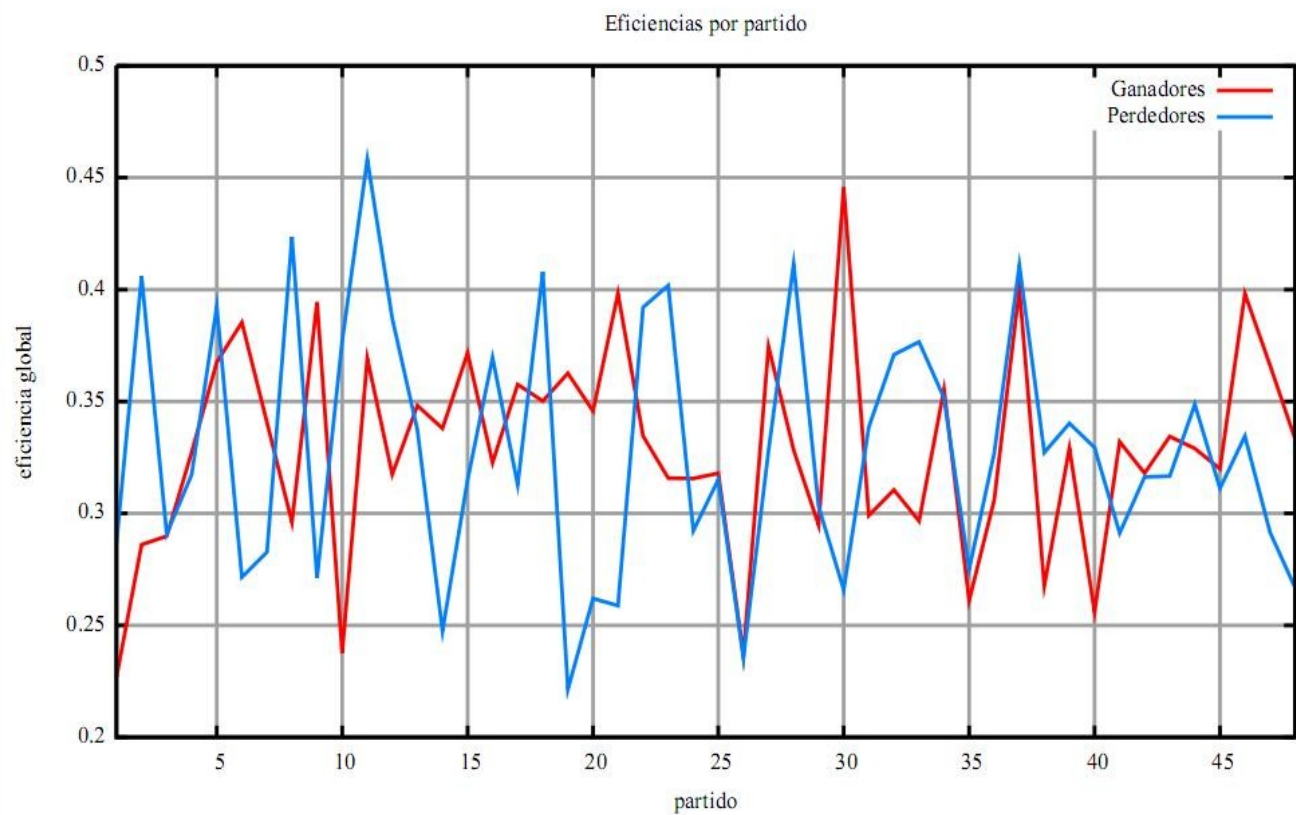


		To																										
		Maarten STEKELENBURG Gregory VAN DER WIEL John HEITINGA Joris MATHJUSEN Giovanni VAN BRONCKHORST Mark VAN BOMMEL Dirk KUYT Nigel DE JONG Robin VAN PERSIE Wesley SNEIJDER Arjen ROBBEN Edson BRAAFHEID Eljero ELIA Rafael VAN DER VAART																										
From	TP	1	2	3	4	5	6	7	8	9	10	11	15	17	23	Total			Long			Medium			Short			
Maarten STEKELENBURG	1	129'20"		4	8	13	5	-	-	-	3	1	1	-	-	-	35	54	65%	13	30	20	21	2	3			
Gregory VAN DER WIEL	2	129'20"	8		4	-	3	7	-	1	1	2	5	-	-	-	31	54	57%	7	8	16	27	8	19			
John HEITINGA	3	115'08"	5	2		4	-	10	-	1	3	-	2	-	-	-	27	45	60%	4	12	19	26	4	7			
Joris MATHJUSEN	4	129'20"	8	5	2		4	1	3	5	1	2	1	-	1	-	33	48	69%	6	16	21	25	6	7			
Giovanni VAN BRONCKHORST	5	109'59"	3	2	2	4		1	9	3	2	-	1	-	2	-	29	39	74%	3	8	20	24	6	7			
Mark VAN BOMMEL	6	129'20"	2	3	3	3	1		1	3	4	3	1	-	1	2	27	46	59%	2	6	20	34	5	6			
Dirk KUYT	7	72'32"	3	1	-	-	2	-		1	1	2	1	-	-	-	11	18	61%	5	5	5	8	1	5			
Nigel DE JONG	8	104'19"	2	4	-	1	1	2	4		1	6	3	-	1	-	25	37	68%	7	8	14	24	4	5			
Robin VAN PERSIE	9	129'20"	-	2	-	-	2	1	2	-		6	4	1	1	-	19	31	61%	4	6	11	15	4	10			
Wesley SNEIJDER	10	129'20"	-	-	4	2	-	2	1	3	3		8	1	3	1	28	46	61%	5	9	18	26	5	11			
Arjen ROBBEN	11	129'20"	2	3	-	-	-	-	-	3	2	3		-	-	1	14	31	45%	4	8	5	12	5	11			
Edson BRAAFHEID	15	19'21"	1	-	-	-	-	-	-	-	-	-	-		1	2	4	7	57%	0	1	3	5	1	1			
Eljero ELIA	17	56'48"	1	-	-	-	-	-	-	1	-	-	-	2		-	4	11	36%	1	2	2	5	1	4			
Rafael VAN DER VAART	23	25'01"	-	-	1	1	-	1	-	-	-	4	-	-	-	-	7	8	88%	0	0	5	6	2	2			
Total passes received:			35	26	24	28	18	25	20	21	21	29	27	4	10	6	Team total:			294	475	62%	61	119	179	258	54	98

		To																										
		Iker CASILLAS Gerard PIQUE Carles PUYOL Andres INIESTA David VILLA XAVI Joan CAPDEVILA XABI ALONSO SERGIO RAMOS Sergio BUSQUETS PEDRO Fernando TORRES Cesc FABREGAS Jesus NAVAS																										
From	TP	1	3	5	6	7	8	11	14	15	16	18	9	10	22	Total			Long			Medium			Short			
Iker CASILLAS	1	129'20"		6	5	-	-	-	2	-	4	2	1	1	1	22	35	63%	9	20	12	13	1	2				
Gerard PIQUE	3	129'20"	3		7	4	4	4	3	1	6	7	4	-	4	51	60	85%	12	16	29	34	10	10				
Carles PUYOL	5	129'20"	6	8		1	-	12	7	5	1	5	2	-	6	1	54	62	87%	2	5	34	38	18	19			
Andres INIESTA	6	129'20"	-	-	1		1	7	8	7	3	6	2	-	5	1	41	70	59%	2	8	22	39	17	23			
David VILLA	7	111'29"	-	-	-	4		2	1	-	2	-	2	-	2	13	25	52%	0	1	7	14	6	10				
XAVI	8	129'20"	-	11	7	8	10		9	8	7	10	4	-	4	2	80	99	81%	11	20	51	59	18	20			
Joan CAPDEVILA	11	129'20"	1	-	6	11	-	12		3	1	12	2	-	5	-	53	71	75%	3	10	35	42	15	19			
XABI ALONSO	14	88'30"	-	4	3	11	1	8	1		3	4	4	-	2	-	41	53	77%	5	7	21	27	15	19			
SERGIO RAMOS	15	129'20"	-	1	5	4	4	8	2	6		10	4	1	2	7	54	73	74%	6	9	34	47	14	17			
Sergio BUSQUETS	16	129'20"	-	10	9	7	2	11	6	8	6		2	-	6	3	70	80	88%	5	6	49	56	16	18			
PEDRO	18	61'55"	-	2	-	-	2	-	1	3	4	1		-	-	-	13	20	65%	1	1	9	14	3	5			
Fernando TORRES	9	17'51"	-	-	-	1	-	-	-	-	-	1	-	-	-	-	2	3	67%	0	0	2	2	0	1			
Cesc FABREGAS	10	40'50"	-	4	2	9	1	7	2	-	2	3	-	-	1	-	31	38	82%	3	4	24	29	4	5			
Jesus NAVAS	22	67'25"	1	1	-	2	1	3	1	1	5	-	-	1	1	-	17	26	65%	2	5	11	15	4	6			
Total passes received:			11	47	45	62	26	74	43	42	44	61	27	3	34	23	Team total:			542	715	76%	61	112	340	429	141	174

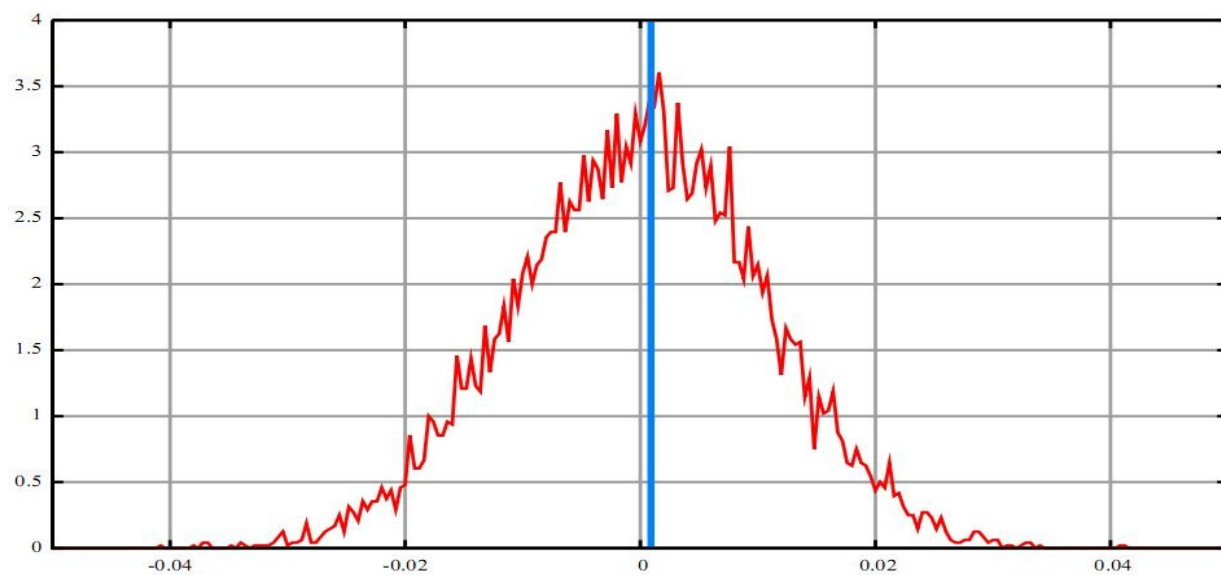
(Figura 2: distribución de pases entre los jugadores de Holanda y España)

Para cada equipo de futbol se representó a los jugadores como nodos y las aristas entre ellos fueron ponderadas con la cantidad de pases entre este par de jugadores. Las aristas fueron normalizadas con la cantidad máxima de pases para ese equipo en ese partido. Para calcular el camino más corto para cada par de nodos se utilizó el algoritmo de Floyd-Warshall, posterior a esto se calculó la eficiencia para cada equipo en cada partido. La Figura 3 muestra las eficiencias de los equipos ganadores y las de los perdedores para cada partido.

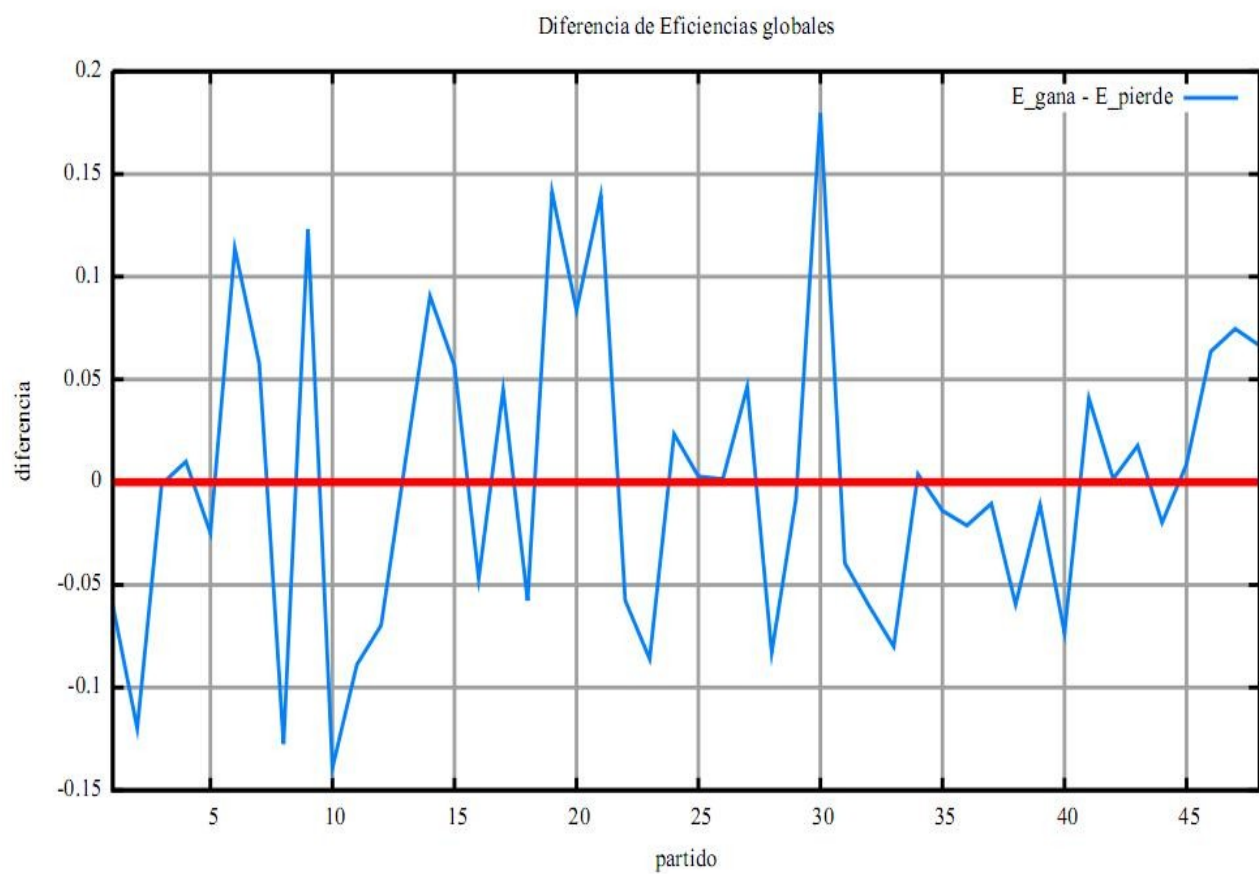


(Figura 3: eficiencias de los equipos ganadores y perdedores)

Se simuló el mundial aleatoriamente 10,000 veces de la siguiente manera. De las 128 eficiencias del mundial (hubo 64 partidos, lo que nos dio 128 eficiencias) se tomaron 48 al azar para los equipos ganadores y 48 al azar para los perdedores; sólo se tomaron 48 porque hubo 16 empates. Luego se calculó el promedio de sus diferencias. La Figura 4 muestra en el eje horizontal el promedio de las diferencias y en el eje vertical la probabilidad de aparición de estas. La línea azul es el promedio de las diferencias en el mundial real. En la Figura 5 se muestran las diferencias para cada partido en el mundial real.



(Figura 4: Promedio de las diferencias aleatorias comparado con la probabilidad de aparición)



(Figura 5: Diferencia de eficiencias entre ganador y perdedor)

Resultados y Discusión

El promedio de las diferencias en el mundial real fue de 9.1×10^{-4} . Para conocer si el resultado era significativo se calcularon rangos de significancia para la Figura 4 al 90, 80 y 70 por ciento; un ejemplo sería que si queríamos que nuestro resultado tuviera una efectividad igual o mayor a 90% tomamos un 5% del lado izquierdo y derecho de nuestra Figura 4 y esto nos daba el límite inferior y superior respectivamente; nuestro promedio debía encontrarse fuera de estos límites.

Rangos al x%	Límite inferior	Límite Superior
90.00	-1.7×10^{-2}	1.7×10^{-2}
80.00	-1.3×10^{-2}	1.4×10^{-2}
70.00	-1.1×10^{-2}	1.1×10^{-2}

El promedio de las diferencias originales se encontró dentro de todos los rangos. Por lo tanto, el resultado no fue significativo.

Conclusiones

La eficiencia global no es un buen parámetro para medir el desempeño de un equipo de fútbol al representarlo como un grafo. Se necesita el uso de técnicas diferentes para encontrar qué equipo es mejor durante un partido.

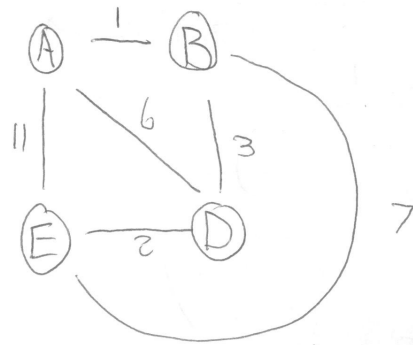
Agradecimientos

Gracias a Arlex Marín por ayudarnos en la implementación de los códigos y por habernos brindado todo el apoyo que necesitábamos. Gracias al Dr. Markus Mueller por permitirnos utilizar su oficina en la Facultad de Ciencias de la UAEM como centro de trabajo.

Bibliografía

1. <http://blogs.ua.es/matematicadiscrecion/2011/01/06/tema-7-grafos-ponderados-sesion-21122010/> imagen de grafos ponderado recuperada el 26 de mayo del 2011
2. Anonimo, Javier López y Hugo Touchette, Publicado el día 9 de Julio del 2010 consultado el 8/03/11 <http://ztfnews.wordpress.com/2010/07/09/el-pulpo-paul-y-la-teoria-de-grafos>
3. (Página de la Fifa) <http://es.fifa.com/worldcup/archive/southafrica2010/index.html> consultado 1/06/11

Anexo1:



C_{ij} = Camino de i a j

$\min(x, y)$ = El menor entre x y y

Grafo a evaluar

- ① Los caminos al principio son considerados igual de costosos que las aristas uniendolos nodos. Ejemplos

$$\begin{array}{ll} C_{BD} = 3 & C_{BE} = 7 \\ C_{AE} = 11 & C_{DA} = 6 \end{array}$$

- ② Luego se toma A como nodo intermedio y se intentan mejorar los caminos sin embargo en este caso no es posible.

Ejemplos

$$C_{EB} = \min(C_{EB}, C_{EA} + C_{AB}) = \min(7, 11 + 1) = 7$$

$$C_{DB} = \min(C_{DB}, C_{DA} + C_{AB}) = \min(3, 6 + 1) = 3$$

- ③ Luego se toma B como nodo intermedio y se intenta mejorar los caminos. En este caso ~~si~~ es posible.

Ejemplos

$$C_{AE} = \min(C_{AE}, C_{AB} + C_{BE}) = \min(11, 1 + 7) = 8$$

$$C_{AD} = \min(C_{AD}, C_{AB} + C_{BD}) = \min(6, 1 + 3) = 4$$

- ④ Se hace lo mismo con los nodos D y E.



“Teoría de Grafos y Sudáfrica 2010”

15 X 7
espacio libre

PARTICIPANTES:

-Rodrigo Alfonsín de la Vega
-Ricardo Monter Aguilar

-Pablo Franquesalme Álvarez
-Armando Rocha Cebrián



Introducción:

La teoría de grafos explica las propiedades, y cualidades de los grafos. El grafo es el conjunto de objetos que llamamos vértices y una selección de pares de vértices que llamamos aristas. Cotidianamente un grafo se expresa mediante una serie de puntos que están conectados por líneas.

Resultados:

70% [0.0110051279 -0.011060257]

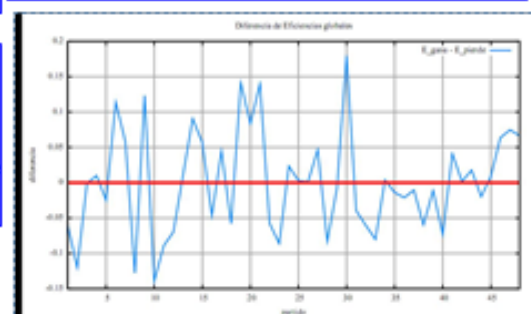
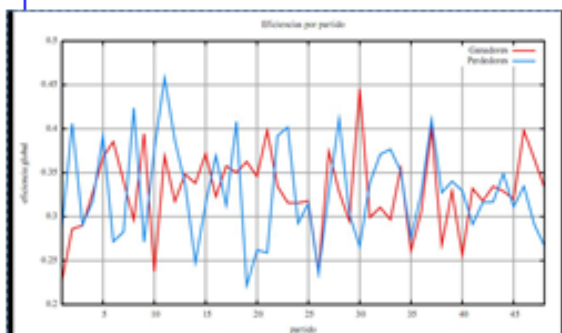
var=0.00530411479

Promedio de las diferencias=0.0009124375

No se obtuvo un resultado significativo ya que el promedio de las diferencias se encuentran dentro del rango del 70%.

Objetivo:

Crear una técnica para distinguir ganadores y perdedores de los encuentros del mundial 2010 usando la eficiencia global como criterio. Todo esto con un nivel de significancia definido.

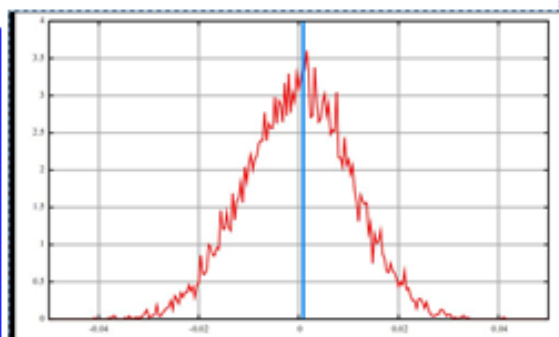


Conclusiones:

La eficiencia global no es una característica que permita distinguir el ganador del perdedor. Hay que buscar otro tipo de evaluación para lograr ese objetivo.

Materiales y Métodos:

Computadoras
UAEM
Recopilación de datos de países en la página de la FIFA.
Organización de países.
Evaluación de eficiencias
Graficar resultados.
Evaluación estadística de los resultados.



Agradecimientos: Agradecemos el apoyo del Doctor **Matthias Mader** por haberme dado la idea de este proyecto, la cual sin su ayuda no habría sido posible realizarla, a **Salim Omar Nijm** por su paciencia y también por su amable apoyo y orientación, por último a nuestros profesores el Doctor **Enrique Galindo** quien nos enseñó a hacer este proyecto.

Bibliografía:

Problemas y algoritmos, Luis M. Vargas, Addison Wesley, 1993-2011

www.fifa.com (Página de la FIFA, en donde se obtuvieron los datos para que el proyecto se hiciera posible.)