

Universitat Politècnica de Catalunya (UPC)  
Facultat d'Informàtica de Barcelona (FIB)



---

# Estudio experimental del Forward Linear Threshold Rank

---

*Autor:*  
Pau García Rodríguez

*Directora:*  
Maria José Serna Iglesias

*Especialización:* Computación

27 de junio de 2019

## Resumen

La centralidad y la difusión de influencia son dos de los conceptos más estudiados en el análisis de redes sociales. La mayoría de medidas de centralidad propuestas, se basan en características topológicas de la red. Recientemente se han propuesto nuevas medidas de centralidad basadas en los dos principales modelos de difusión de influencia, el *Independent Cascade Model* y el *Linear Threshold Model*. Una de estas medidas basadas en estos modelos es el *Linear Threshold Rank*, la cual se define como el número de nodos influenciados cuando el conjunto de activación inicial es formado por el mismo nodo y sus vecinos inmediatos. En este proyecto, introducimos nuevas medidas de centralidad basadas en esta medida y continuaremos con el análisis de otras medidas basadas en este mismo modelo y en otras variantes de él. Compararemos estas medidas con otras ya conocidas, utilizando diferentes herramientas para el análisis de desigualdad en redes. Propondremos un algoritmo simple y eficiente, para calcularlas, y lo implementaremos en C++ para reducir los tiempos de ejecución. Finalmente, publicaremos los resultados en una plataforma online, para que puedan ser consultados y contrastados fácilmente.

# Índice general

<b>1. Contexto y alcance del proyecto</b>	<b>5</b>
1.1. Contexto y formulación del problema . . . . .	5
1.2. Actores implicados . . . . .	7
1.2.1. Investigador . . . . .	7
1.2.2. Directora . . . . .	7
1.2.3. Beneficiarios . . . . .	7
1.3. Estado del arte . . . . .	7
1.4. Alcance . . . . .	8
1.5. Metodología . . . . .	9
1.5.1. Herramientas de Desarrollo . . . . .	9
1.5.2. Herramientas de Seguimiento . . . . .	9
1.5.3. Métodos de validación . . . . .	9
1.5.4. Obstáculos y riesgos en el proyecto . . . . .	9
<b>2. Planificación y organización</b>	<b>11</b>
2.1. Descripción de tareas . . . . .	11
2.1.1. Obtención de redes sociales . . . . .	11
2.1.2. Optimización de las centralidades previamente introducidas	12
2.1.3. Programación de las nuevas centralidades . . . . .	12
2.1.4. Familiarización con el entorno de <i>rdlab</i> . . . . .	12
2.1.5. Análisis estadístico de los resultados . . . . .	12
2.1.6. Publicación de los resultados . . . . .	12
2.2. Tiempo estimado . . . . .	13
2.3. Gráfica Gannt . . . . .	13
2.4. Recursos utilizados . . . . .	14
2.4.1. Hardware . . . . .	14
2.4.2. Software . . . . .	14
2.5. Alternativas y plan de acción . . . . .	15
2.6. Cambios en el plan de acción . . . . .	15
<b>3. Presupuesto y Sostenibilidad</b>	<b>16</b>
3.1. Autoevaluación de la sostenibilidad . . . . .	16
3.2. Presupuesto . . . . .	17

3.2.1.	Costes Directos . . . . .	17
3.2.2.	Costes Indirectos . . . . .	18
3.2.3.	Costes imprevistos . . . . .	19
3.2.4.	Coste total final . . . . .	19
3.2.5.	Control de costes . . . . .	19
3.3.	Sostenibilidad . . . . .	20
3.3.1.	Dimensión Ambiental . . . . .	20
3.3.2.	Dimensión Económica . . . . .	20
3.3.3.	Dimensión Social . . . . .	21
<b>4.</b>	<b>Diseño y desarrollo</b>	<b>22</b>
4.1.	Introducción . . . . .	22
4.1.1.	Redes Sociales . . . . .	22
4.1.2.	Linear Threshold Model . . . . .	22
4.1.3.	Linear Threshold Rank . . . . .	25
4.1.4.	Forward / Backward Linear Threshold Rank . . . . .	26
4.1.5.	Fading Linear Threshold Model . . . . .	27
4.1.6.	Bounded Linear Threshold Model . . . . .	29
4.1.7.	Otras medidas de centralidad . . . . .	30
4.1.8.	Estadísticos . . . . .	32
4.1.9.	Migración de actores . . . . .	33
4.2.	Calculo de las medidas introducidas . . . . .	35
4.2.1.	Diseño del algoritmo . . . . .	35
4.2.2.	Implementación . . . . .	41
<b>5.</b>	<b>Experimentos</b>	<b>44</b>
5.1.	Redes utilizadas . . . . .	44
5.2.	Resultados BWLTR y FWLTR . . . . .	48
5.2.1.	Resumen de los resultados . . . . .	48
5.2.2.	Correlación de medidas . . . . .	50
5.2.3.	Migraciones de actores . . . . .	54
5.3.	Resultados del FLTM . . . . .	60
5.3.1.	Correlación de medidas . . . . .	60
5.3.2.	Tiempos de ejecución . . . . .	70
5.4.	Resultados del BLTM . . . . .	76
5.4.1.	Correlación de medidas . . . . .	76
5.4.2.	Tiempos de ejecución . . . . .	86
<b>6.</b>	<b>Conclusiones y Trabajo Futuro</b>	<b>90</b>
6.1.	Conclusiones . . . . .	90
6.2.	Trabajo Futuro . . . . .	92
	<b>Bibliografía</b>	<b>94</b>
<b>A.</b>	<b>Códigos</b>	<b>97</b>
A.1.	Graph . . . . .	97



A.1.1.	Graph.h . . . . .	97
A.1.2.	Graph.cpp . . . . .	99
A.2.	Otras medidas de Centralidad . . . . .	102
A.2.1.	Degree . . . . .	102
A.2.2.	Closeness . . . . .	102
A.2.3.	Betweenness . . . . .	104
A.2.4.	Katz . . . . .	105
A.2.5.	PageRank . . . . .	106
A.2.6.	ICR . . . . .	107
<b>B.</b>	<b>Migraciones de nodos</b>	<b>108</b>
B.1.	Clasificación por posición . . . . .	108
B.1.1.	ArXiv . . . . .	108
B.1.2.	Caida . . . . .	112
B.1.3.	Dining_Table . . . . .	115
B.1.4.	Dolphins . . . . .	118
B.1.5.	Epinions . . . . .	121
B.1.6.	Higgs . . . . .	124
B.1.7.	Human_Brain . . . . .	127
B.1.8.	Wikipedia . . . . .	130
B.2.	Clasificación por valor . . . . .	133
B.2.1.	ArXiv . . . . .	133
B.2.2.	Caida . . . . .	136
B.2.3.	Dining_Table . . . . .	139
B.2.4.	Dolphins . . . . .	142
B.2.5.	Epinions . . . . .	145
B.2.6.	Higgs . . . . .	148
B.2.7.	Human_Brain . . . . .	151
B.2.8.	Wikipedia . . . . .	154
<b>C.</b>	<b>Tiempos de ejecución</b>	<b>157</b>
C.1.	Fading Linear Threshold Model . . . . .	157
C.1.1.	ArXiv . . . . .	158
C.1.2.	Caida . . . . .	159
C.1.3.	Dining_Table . . . . .	160
C.1.4.	Dolphins . . . . .	161
C.1.5.	Epinions . . . . .	162
C.1.6.	Higgs . . . . .	163
C.1.7.	Human_Brain . . . . .	164
C.1.8.	Texas . . . . .	165
C.1.9.	Wikipedia . . . . .	166
C.2.	Bounded Linear Threshold Model . . . . .	167
C.2.1.	ArXiv . . . . .	168
C.2.2.	Caida . . . . .	169
C.2.3.	Dining_Table . . . . .	170
C.2.4.	Dolphins . . . . .	171

C.2.5. Epinions . . . . .	172
C.2.6. Higgs . . . . .	173
C.2.7. Human_Brain . . . . .	174
C.2.8. Texas . . . . .	175
C.2.9. Wikipedia . . . . .	176

# Capítulo 1

## Contexto y alcance del proyecto

### 1.1. Contexto y formulación del problema

En los últimos años, la influencia y el uso de las redes sociales no ha hecho más que aumentar. Las redes sociales son cada vez más integradas en nuestro día a día, en ellas compartimos imágenes, nos comunicamos con nuestros amigos, expresamos nuestras opiniones y leemos y comentamos las de otros. Este incremento masivo en el uso de las redes sociales ha comportado la existencia de una extensa variedad y complejidad de redes sociales, lo cual ha generado una necesidad de nuevas herramientas y medidas para el estudio y análisis de esas.

Un concepto esencial en las redes sociales es saber como de relevante es un actor dentro de la red social, es decir su importancia relativa respecto a los otros actores en la red. Este concepto se conoce como la centralidad de un actor y es uno de los conceptos más estudiados en sociología. Fue introducido inicialmente por Bavelas a finales de los años 40 [1], y se empezó a utilizar en redes sociales a finales de los años 70 [2].

Identificar actores relevantes en un red social es particularmente útil para muchas aplicaciones, como por ejemplo: Detección de regiones funcionales en el cerebro [3], búsqueda de investigadores clave en una red de investigación [4], los usuarios más activos y populares en Twitter [5], en la transmisión de enfermedades [6] o en el estudio de las calles urbanas [7].

Representaremos una red social con un grafo, donde los actores son los vértices,

y los arcos las relaciones interpersonales dentro de la red. Las medidas de centralidad más tradicionales como, **Degree**, **Closeness** y **Betweenness**, que están relacionadas con la topología del grafo. El reciente incremento masivo en el uso de internet, y la aparición de nuevas redes sociales más variadas y complejas, ha creado la necesidad de nuevas medidas de centralidad más sofisticadas. Debido al inmenso tamaño de esas redes, las nuevas medidas de centralidad han de poder computarse de manera eficiente. Dos de las medidas más populares que cumplen estos requisitos son, **Katz** [8] y **PageRank** [9]. Ambas están basadas en la centralidad del vector propio [10]. Muchas de las nuevas centralidades que han ido surgiendo están basadas en las anteriores.

Existe un problema en sociología, de como modelizar la expansión de ciertas ideas a través de un red social. Los dos modelos más conocidos y más usados son el ICM (*Independent Cascade Model*) [11] y el LTM (*Linear Threshold Model*) [11]. La mayoría de estudios se han centrado en utilizar estos dos modelos, para resolver el problema de maximizar la difusión de influencia. Este problema consiste en encontrar un conjunto de actores de tamaño  $k$ , que maximice la difusión de influencia respecto a todos los otros conjuntos posibles de tamaño  $k$ . Estos modelos también se pueden utilizar para obtener nuevas medidas de centralidad. Existen pocos estudios que usen estos modelos para nuevas centralidades, y solo se usa el ICM como modelo [12–14].

Recientemente se introdujo una nueva medida de centralidad basada en el LTM [15]. Esta nueva centralidad, LTR (*Linear Threshold Rank*), usa como modelo de influencia el LTM. En el modelo LTM cada nodo tiene un valor de umbral asignado y todas las aristas tienen un peso asignado, un nodo será activado (influenciado) si la suma de los pesos de los arcos de los nodos adyacentes previamente activados es igual o superior a su umbral. La centralidad LTR consiste en que, dado un nodo  $i$  y un conjunto de nodos determinados, simular cuantos nodos se activan en total, si partiésemos con ellos como conjunto de nodos activados inicialmente. La centralidad de cada nodo, será el numero de nodos activados en total dividido entre el numero total de nodos. En esta primera versión de LTR el conjunto inicial es el mismo nodo  $i$  más todos sus vecinos en cualquier dirección. Más tarde, en [16], se crearon nuevos tipos de LTR con diferentes conjuntos iniciales. Principalmente, en vez de usar los vecinos del nodo  $i$  en todas las direcciones, se usan solo los vecinos sucesores de  $i$ , el FLTR (*Forward Linear Threshold*). También se estudian diferentes comportamientos en la dispersión de influencia.

En este trabajo vamos a continuar el estudio realizado descrito en anteriormente. Propondremos nuevas formas de obtener el conjunto inicial, nuevas formas de asignar el umbral en cada nodo, y compararemos los resultados usando diferentes redes sociales usadas anteriormente, las cuales podremos modelar como

grafos de influencia. Con los resultados obtenidos montaremos una plataforma *online* sencilla para que se puedan contrastar y comparar. También vamos a implementar las nuevas centralidades de forma eficiente, paralelizando el código, en C++, ya que en los otros dos trabajos mencionados, con redes de varios miles de nodos, los tiempos de ejecución eran de varias horas o incluso días.

## 1.2. Actores implicados

Los actores implicados en el desarrollo del proyecto son los siguientes:

### 1.2.1. Investigador

Es el encargado principal del proyecto. Es responsable de la planificación del proyecto, la documentación, la búsqueda de información y la programación de los diferentes códigos.

### 1.2.2. Directora

La directora de este proyecto es Maria José Serna Iglesias profesora del departamento de computación de la UPC. Su rol es guiar al investigador y dar consejo.

### 1.2.3. Beneficiarios

Existen muchas centralidades diferentes porque medir como de *importante* es un actor en una red es un concepto muy amplio. Diferentes centralidades pueden dar resultados muy diferentes en un mismo grafo. La elección de una centralidad dependerá de lo que el usuario considere que es más importante medir. Así pues, los investigadores que estudien la centralidad de los actores en las redes sociales, relacionada con la dispersión de influencia, serán beneficiados por las nuevas centralidades que presentamos.

## 1.3. Estado del arte

Existe mucha investigación en centralidad de redes sociales, y sigue atrayendo a una gran variedad de investigadores, debido al uso masivo de la redes sociales y a la necesidad de herramientas para poder analizarlas. Si bien existen muchos

estudios sobre la centralidad en redes sociales, pocos se basan en los modelos de dispersión de influencia.

Como hemos mencionado antes, hasta hace poco, los pocos estudios que se basan en la dispersión de influencia, se basan en el modelo ICM. Su principal utilidad es la de resolver el problema de maximizar la dispersión de influencia en una red social. El problema consiste en encontrar un conjunto de actores de tamaño  $k$  que maximice la dispersión de influencia respecto a los otros posibles conjuntos. En [17] se demostró que este problema es *NP-Hard*. Mediante la centralidad ICR (*Independent Cascade Rank*), se consigue una aproximación de este problema, que da mejores resultados respecto al resto de aproximaciones que usan otras centralidades y diferentes heurísticas [12].

Recientemente se introdujo una nueva centralidad LTR, basada en el modelo de dispersión LTM [15]. En [16] se continua la investigación anterior y se introducen nuevas centralidades basadas en LTR y se comparan con otras centralidades más usadas con diferentes medidas estadísticas. Que nosotros sepamos, hasta ahora, no existen más estudios de centralidades basadas en el modelo de dispersión LTM. Es por eso que vamos a continuar con el estudio de este tipo de centralidad, añadiendo variaciones, compáralas con otras ya existentes y publicando los resultados en un plataforma online para que se puedan consultar y contrastar fácilmente.

## 1.4. Alcance

Primero optimizaremos y paralelizaremos en C++ las medidas de centralidad vistas en [15, 16] y las nuevas que introduzcamos, ya que en redes sociales con un tamaño de nodos y arcos de varios miles, los tiempos de ejecución pueden ser de varias horas o incluso días si los códigos no están bien optimizados. Estos códigos se ejecutarán en el cluster del *rddlab* para poder reducir los tiempos de ejecución lo máximo posible.

Introduciremos nuevas centralidades basadas en el modelo de dispersión LTM, buscando diferentes maneras de generar el conjunto de actores inicial y la asignación de diferentes umbrales de activación a cada actor. Para comparar los resultados de las diferentes centralidades usaremos los coeficientes de **Kendall** y **Spearman**.

Ejecutaremos los experimentos sobre diferentes redes sociales, las cuales podremos modelar fácilmente como grafos de influencia, de varios tamaños y propie-

dades. De esta forma podremos ver el comportamiento de las centralidades en una variedad de diferentes redes sociales.

Todos los resultados que obtengamos, así como los códigos para implementar las centralidades, serán publicados en una plataforma online (*rdlab*) para que se puedan consultar y contrastar. Si el tiempo lo permite, intentaremos que los experimentos se puedan ejecutar de forma interactiva en la misma plataforma.

## 1.5. Metodología

Para poder desarrollar correctamente el proyecto nos hemos organizado de la siguiente forma. Nos reuniremos cada 7 o 15 días, para discutir los resultados, comentar el estado del proyecto y acordar los siguientes pasos a realizar.

### 1.5.1. Herramientas de Desarrollo

Programaremos las centralidades en el lenguaje `C++` y la librería `OpenMP`. Para el análisis de los resultados utilizaremos tanto el lenguaje `R` como `Python`. Finalmente para guardar y organizar los resultados en diferentes equipos usaremos el servicio de alojamiento de archivos en la nube `Dropbox`.

### 1.5.2. Herramientas de Seguimiento

Para el seguimiento del proyecto se usaran `Git` y la plataforma `GitHub`, para poder tener un seguimiento de todas las versiones realizadas de los códigos y las tareas a realizar.

### 1.5.3. Métodos de validación

Para las medidas de centralidad vistas en [15, 16], comparemos los resultados obtenidos mediante los códigos optimizados con los resultados de los trabajos previos.

### 1.5.4. Obstáculos y riesgos en el proyecto

La ejecución de las medidas de centralidad en redes sociales con miles de actores y relaciones puede ser excesivamente larga. En [15, 16] los tiempos de ejecución

podían llegar a horas o incluso días en algunos casos. Es por eso que reescribiremos las centralidades en `C++` paralelizando el código lo máximo posible y ejecutaremos los experimentos en el cluster del *rdlab*, y así asegurarnos que tendremos tiempo a ejecutar todos los experimentos necesarios.



## Capítulo 2

# Planificación y organización

El proyecto empezó el 24 de enero, 2019 con fecha límite el 24 de junio. Es decir, que la duración del proyecto es de unos 5 meses. La planificación descrita aquí puede modificarse durante el desarrollo del proyecto.

### 2.1. Descripción de tareas

Vamos a describir las diferentes tareas que se realizaran en el proyecto.

#### 2.1.1. Obtención de redes sociales

Para los diferentes experimentos, usaremos diferentes redes sociales con varios tamaños y de diferentes fuentes. Debido a la diversa procedencia de las redes sociales y que estas están representadas en diferentes formatos, para un uso correcto y fácil uso de ellas debemos convertirlas en un mismo formato. Así pues, convertiremos cada red social en un grafo de influencia. Todos los grafos de influencia tendrán el mismo formato, para su posterior uso en los códigos de las centralidades.

### **2.1.2. Optimización de las centralidades previamente introducidas**

Optimizaremos las centralidades que se introdujeron y se usaron en los dos estudios [15,16]. Esas centralidades están escritas en el lenguaje *Python* y usan la librería *NetworkX*. Nosotros vamos a reescribirlas usando el lenguaje *C++* y la librería *OpenMP* para poder paralelizar los códigos. De esta forma, reduciremos los tiempos de ejecución y podremos ejecutar los experimentos dentro del plazo del proyecto. Para verificar que están bien implementadas, compararemos los resultados con los previamente obtenidos usando *Python*.

### **2.1.3. Programación de las nuevas centralidades**

Una vez hayamos optimizado las anteriores centralidades, programaremos las nuevas centralidades que vayamos introduciendo en *C++* usando *OpenMP* para paralelizar el código, y así obtener el rendimiento deseado.

### **2.1.4. Familiarización con el entorno de *rdlab***

Debido al tamaño de algunas de las redes sociales que usamos, un código bien optimizado no es suficiente para su correcta ejecución en el tiempo del proyecto, se requiere una buena potencia computacional. También necesitaremos una plataforma en donde subir nuestros resultados. El cluster de *rdlab* nos proporciona la potencia computacional necesaria. Para su correcto uso, es necesario un proceso de aprendizaje y familiarización de su entorno y sus múltiples herramientas.

### **2.1.5. Análisis estadístico de los resultados**

Mediante el lenguaje *R* y el IDE *RStudio* analizaremos y compararemos los resultados obtenidos usando diferentes estadísticos. También obtendremos diferentes gráficas para representar nuestros resultados.

### **2.1.6. Publicación de los resultados**

Finalmente, para que se puedan comprobar y contrastar fácilmente los resultados obtenidos, estos serán publicados en una plataforma online. La cual será probablemente *ALBCOM*.

## 2.2. Tiempo estimado

En la siguiente tabla podemos ver los tiempos de ejecución estimados de cada tarea en horas.

Tarea	Duración Estimada (h)
Obtención de redes sociales	25
Optimización de las centralidades previamente introducidas	100
Programación de las nuevas centralidades	100
Familiarización con el entorno de rdlab	25
Análisis estadístico de los resultados	75
Publicación de los resultados	50
Total	375

## 2.3. Gráfica Gannt

En la siguiente figura se observa la planificación de las diferentes tareas en una gráfica de Gannt.

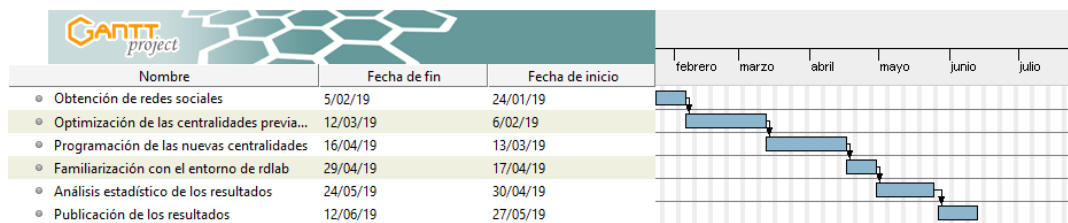


Figura 2.1: Gráfica de Gannt

## 2.4. Recursos utilizados

Vamos a describir los recursos utilizados.

### 2.4.1. Hardware

- HP Envy 13-AB002NS Intel Core i5-7200U/8GB/256GB SSD
- Cluster del *rddlab*: HP ProLiant DL380p server 2x Intel(R) Xeon(R) E5-2660

### 2.4.2. Software

- C++
- Python
- R
- OpenMP
- NetworkX
- RStudio
- Latex
- Atom
- Dropbox

## 2.5. Alternativas y plan de acción

Como se ha comentado anteriormente, nos reuniremos cada 7 o 15 días y pondremos en común todo lo que se haya hecho y las dudas y problemas que hayan podido surgir.

Lo primero que se va a hacer es optimizar las centralidades previamente introducidas. Seguidamente, programaremos y optimizaremos las nuevas centralidades que vayamos introduciendo. Ejecutaremos todos los experimentos en el cluster del *rllab* para asegurarnos que terminan dentro de los plazos del proyecto. Analizaremos estos resultados con el lenguaje *R* utilizando *RStudio*, mediante diferentes estadísticos. Finalmente subiremos nuestros resultados en una plataforma online para su consulta.

Nuestra principal fuente de retrasos es la indisponibilidad del cluster. Aunque es poco probable, si el cluster no estuviera disponible, seguramente no tendríamos suficiente tiempo para ejecutar los experimentos en las redes sociales más grandes, así que si esto sucediera tendríamos que sustituir estas redes por otras con un tamaño más asequible.

Esta posible desviación, supondría un incremento de la duración total del proyecto sobre unas 20 horas, buscar redes más asequibles en tamaño que sustituyan a las más grandes, ejecutar los experimentos en ellas y el posterior análisis y publicación de los resultados.

## 2.6. Cambios en el plan de acción

Por causas ajenas a nosotros, la plataforma *ALBCOM* no esta disponible para subir los resultados, por problemas técnicos, y en el momento de escribir esta memoria aún no se han arreglado. Este inconveniente no nos ha supuesto ningún contratiempo adicional, simplemente cuando la plataforma vuelva a estar disponible subiremos los resultados.

## Capítulo 3

# Presupuesto y Sostenibilidad

### 3.1. Autoevaluación de la sostenibilidad

En este cuestionario se nos hacen diversas preguntas para que nos auto-evaluemos nuestro conocimiento sobre la sostenibilidad.

En el cuestionario he visto que, tengo poco conocimiento sobre las causas, las soluciones y las consecuencias sobre la sostenibilidad. Hasta ahora, pocas veces he analizado un problema desde la perspectiva de la sostenibilidad. En la resolución de problemas, identifico las causas, sus posibles consecuencias y de relacionar el problema con otros conocidos y ya resueltos. Conozco poco los conceptos de creatividad e innovación y no comprendo del todo las técnicas de innovación.

Tengo poco en cuenta los efectos ambientales en mis proyectos, tampoco sé medir muy bien el impacto social y ambiental en el planeta mediante los indicadores adecuados. Esto podría llegar a repercutir negativamente en el impacto social y medioambiental de mis proyectos, ya que al desconocer indicadores y herramientas para medir ese impacto, en caso de ser un impacto negativo, no sería consciente de este hecho.

En general, en la planificación de proyectos y en su ejecución tengo poca idea sobre el impacto social que tiene. Ni si este contribuye a mejorar el bien común de la sociedad. Si bien sé valorar la viabilidad económica de un proyecto, no

valoro la compatibilidad con las dimensiones ambiental y social de la sostenibilidad.

En proyectos con diferente personal, conozco y uso herramientas de trabajo colaborativo, y sé valorar el trabajo con diferentes personas.

En conclusión, el cuestionario me ha hecho ver y reflexionar, que tengo poco en cuenta los principios de la sostenibilidad y sus herramientas para medir el impacto que tienen mis proyectos. En definitiva, he visto que necesito informarme y concienciarme más sobre el tema de la sostenibilidad, para así aplicar y tener en cuenta estos conceptos en la planificación y desarrollo de proyectos futuros.

## 3.2. Presupuesto

### 3.2.1. Costes Directos

A continuación se exponen los costes directos estimados en cuanto a recursos humanos y software:

- **Recursos Humanos:** Para la estimación de costes en cuanto a recursos humanos se refiere, se ha tomado como referencia el importe, de 7.32 € por hora, de la beca para la ayuda a la contratación de personal investigador novel (FI-2019).

Actividad	Duración Estimada (h)	Coste investigador
Obtención de redes sociales	25	183 €
Optimización de las centralidades previamente introducidas	100	732 €
Programación de las nuevas centralidades	100	732 €
Familiarización con el entorno de rdlab	25	183 €
Análisis estadístico de los resultados	75	549 €
Publicación de los resultados	50	366 €
Total	375	2745 €

- **Software:** Para el desarrollo del proyecto usaremos software gratuito:

Software	Coste
Rstudio	0 €
LaTeX	0 €
Atom	0 €
Dropbox	0 €
Total	0 €

### 3.2.2. Costes Indirectos

Los costes indirectos estimados son los siguientes:

Producto	Precio	Vida Útil	Amortización
HP Envy 13-AB002NS	900€	5 años	40€
Total	900€	-	40€

Los costes de amortización de cada producto, suponiendo que hay 251 días laborables y se trabajan 8 horas diarias, se han calculado mediante la siguiente formula:

$$\frac{\text{Precio}}{\text{Vida\_util} \cdot \text{Dias\_laborables}(251) \cdot \text{Horas\_trabajo\_dia}(8)} \cdot \text{Horas\_TFG}$$

#### Coste total

Así pues, el coste total del proyecto sin contar los imprevistos es de :

Concepto	Coste estimado
Recursos Humanos	2745 €
Software	0 €
Costes Indirectos	40 €
Subtotal	2815 €
Contingencia (5%)	140.75 €
<b>Total</b>	<b>2955.75 €</b>



### 3.2.3. Costes imprevistos

Como hemos comentado anteriormente, nuestra principal fuente de retrasos es la indisponibilidad del *rddlab*. Esta indisponibilidad supondría buscar nuevas redes sociales con un tamaño más asequible y ejecutar los experimentos en ellas. Estimamos que la probabilidad de que esto suceda es bastante baja, sobre un 5%. Así pues, el coste adicional sería el siguiente:

Actividad	Duración Estimada (h)	Coste	Coste con riesgo (5%)
Obtención de redes	5	36.6 €	1.83 €
Análisis estadístico	15	109.8 €	5.49 €
<b>Total</b>	20	146.4 €	7.32 €

### 3.2.4. Coste total final

Teniendo en cuenta todos los costes anteriores, el coste final total estimado es el siguiente:

Concepto	Precio
Coste total	2955.75 €
Costes imprevisto	7.32 €
<b>Total</b>	2963.07 €

### 3.2.5. Control de costes

Para controlar el coste del proyecto, al final de la ejecución de cada tarea, actualizaremos el coste temporal y los recursos utilizados, así como los costes adicionales imprevistos que hayan surgido. Compararemos estos números con los estimados, para ver la desviación con respecto al presupuesto estimado. Aplicaremos las siguientes formulas:

$$\begin{aligned} \text{Desviación de costes} &= (CE - CR) \cdot HR \\ \text{Desviación de consumo} &= (HE - HR) \cdot CE \end{aligned}$$

Donde:

$$\begin{aligned} CE &= \text{Coste estimado} \\ CR &= \text{Coste real} \\ HE &= \text{Horas estimadas} \\ HR &= \text{Horas reales} \end{aligned}$$

Como actualizaremos el presupuesto al final de cada tarea, podremos ver si hay desviaciones en el presupuesto y donde han ocurrido. De esta forma podremos ver si el presupuesto de contingencias e imprevistos cubre estas desviaciones, o por otra parte el coste real supera al coste presupuestado.

### 3.3. Sostenibilidad

#### 3.3.1. Dimensión Ambiental

El único impacto medioambiental de nuestro proyecto es el consumo de electricidad asociado al computo de las medidas de centralidad. Al tratarse de un estudio experimental sobre algoritmos, no creamos ningún tipo de producto y no usamos ningún material, es por eso que no nos hemos planteado el reciclaje de materiales ni la reutilización de recursos de otros proyectos.

Una parte del proyecto se dedica a optimizar los diferentes códigos que utilizamos. En [15], los tiempos de ejecución para a una de las redes más grandes es del orden de horas, nosotros en la misma red tardamos unos 28 segundos. Supongamos que anteriormente el tiempo de ejecución total era de unas 2 horas (en el paper no se especifica exactamente). En los dos casos utilizamos los mismos procesadores, dos Intel(R) Xeon(R) E5-2660, que consumen cada uno 95 W. Así pues tendríamos que en la ejecución de las medidas sobre esta red, cada uno gastamos:

$$2 \cdot 95W/procesador \cdot \frac{1J}{segundo} \cdot 28segundos = 5.32KJ$$
$$2 \cdot 95W/procesador \cdot \frac{1J}{segundo} \cdot (2 \cdot 3600)segundos = 1.367MJ$$

Vemos pues, que al optimizar los códigos, reducimos drásticamente el consumo energético.

No consideramos que existan escenarios, que hagan que nuestro aumento su huella ecológica, ya que en este proyecto solo introducimos unas medidas de centralidad, y no vemos como podrían incrementar su huella ecológica.

#### 3.3.2. Dimensión Económica

Como se ha visto en anterioridad en el documento, hemos hecho una detallada cuantificación de los costes involucrados en proyecto tanto en recursos humanos,

y en software. Creemos que los costes son los necesarios para llevar a cabo el proyecto y no vemos como ahorrar más en él. Al no haber desviaciones temporales, los costes se han mantenido igual que los previstos.

Los principales costes del proyecto son en recursos humanos, los cuales se han estimado a partir del importe de la beca de investigación (FI-2019). Así pues, comparado con otros estudios, a nivel económico estaríamos en el mismo nivel.

Como nuestro proyecto no presenta ningún producto, solo un estudio sobre medidas de centralidad, no van a ver costes adicionales en ajustes ni reparaciones durante su vida útil.

### **3.3.3. Dimensión Social**

A nivel personal, la realización del proyecto no tiene ningún efecto sobre mí, más allá de la experiencia y conocimientos que me aporte, ya que se trata del estudio de unas medidas de relevancia de actores en una red social. Tampoco afecta en ningún aspecto sobre la situación política y social de donde se ha realizado. Ni ningún colectivo ni segmento de la población puede verse afectado, positiva o negativamente, directamente por este proyecto.

El principal beneficio que aporta este proyecto a la sociedad, es la puesta en conocimiento de nuevas centralidades para redes sociales.

## Capítulo 4

# Diseño y desarrollo

### 4.1. Introducción

En esta sección vamos a explicar y mostrar las múltiples definiciones y conceptos que usaremos en nuestro proyecto.

#### 4.1.1. Redes Sociales

Representaremos las redes sociales como un grafo dirigido, en el cual los vértices representan los actores de la red y los arcos, las relaciones entre esos actores. Así pues, dado un grafo  $G = (V, E)$ , tenemos a  $V(G)$  como el conjunto de actores y  $E(G)$  como el conjunto de relaciones.

Existen redes sociales donde además de indicarse las relaciones entre actores, se nos indica la *fuera* de esta relación. Para representar este tipo de redes usaremos un grafo ponderado, donde  $G$  es el grafo definido anteriormente y  $w : E(G) \rightarrow \mathbb{N}$  la función de ponderación, la cual nos asigna los pesos de cada arco en el grafo.

Finalmente, para simplificar la notación usaremos  $w_{ij} = w((i, j))$  para cualquier arco  $(i, j) \in E(G)$ ,  $|V| = n$  y  $|E| = m$ .

#### 4.1.2. Linear Threshold Model

La expansión de influencia en una red social nos permite modelizar como los actores en ella, se influyen los unos a los otros mediante sus relaciones. Un

ejemplo de esto sería el mecanismo mediante el cual se extiende un rumor o una opinión a través de la red, o como afecta una campaña publicitaria a los actores de la red. En nuestro caso sería ver como los vértices ejercen su influencia a través del grafo.

Los dos modelos más conocidos de expansión de influencia en redes son el LTM (*Linear Threshold Model*) y el ICM (*Independent Cascade Model*) [11]. En ambos modelos se empieza con un conjunto inicial de actores, decimos que hemos activado estos actores. Luego, este mismo conjunto de actores intentara activar los actores que estén relacionados con ellos, ejerciendo su influencia en la red. Los actores que estén activados ya no podrán desactivarse. Los actores activados repetirán el proceso de intentar activar a los otros actores, ejerciendo influencia a través la red, hasta que ya no se puedan activar más.

El LTM es un modelo determinista, en el cual cada actor tiene una resistencia a ser influido y los actores tienen diferentes capacidades a influirse entre ellos. En este modelo, las redes se representan como un grafo dirigido, donde cada arco tiene un peso asociado  $w_{ij}$ , el cual se interpreta como que el actor  $i$  puede influenciar al actor  $j$  con una fuerza  $w_{ij}$ . Además cada vértice tiene asociado un valor que nos indica como de susceptible es de ser influenciado por el resto de actores. Este valor se conoce como el umbral (*threshold*) del actor.

La definición formal de un grafo de influencia (definido en [15]) es la siguiente:

**Definición 1.** *Un grafo de influencia es una tupla  $(G, w, f)$ , donde  $G = (V, E)$  es un grafo dirigido formado por un conjunto de actores  $V$  y un conjunto de arcos  $E$ ;  $w : E \rightarrow \mathbb{N}$  es una **función de ponderación** que asigna un peso a cada arco, y  $f : V \rightarrow \mathbb{N}$  una **función de etiquetado** que nos cuantifica como de influenciable es un actor. Un actor  $i \in V$  ejerce influencia sobre otro actor  $j \in V$  si y solo si  $(i, j) \in E$ .*

Así pues, dado un grafo de influencia  $(G, w, f)$  y un conjunto inicial de activación  $X \subseteq V$ , vamos a considerar el siguiente proceso iterativo de activación. Sea  $F_t(X) \subseteq V$  el conjunto de nodos activados en una iteración  $t$ . Inicialmente, en el paso  $t = 0$ , solo los nodos en  $X$  están activados, eso es  $F_0(X) = X$ . En la iteración  $t + 1$  un nodo  $i \in V$  será activado si y solo si

$$\sum_{j \in F_t(X)} w_{ij} \geq f(i)$$

Es decir, un actor  $i$  es activado si y solo si, la suma de los pesos de los actores

activados incidentes a  $i$  es superior o igual a su valor de umbral. Terminamos ese proceso cuando dejan de ocurrir nuevas activaciones. Este proceso se conoce como expansión de influencia, y se define formalmente en [15] como:

**Definición 2.** Sea  $(G, w, f)$  un grafo de influencia, la **expansión de influencia** de  $X$  es

$$F(X) = \bigcup_{t=0}^k F_t(X) = F_0(X) \cup \dots \cup F_k(X)$$

donde  $k = \min\{t \in \mathbb{N} | F_t(X) = F_{t+1}(X)\} \leq n$ . El valor  $t$  denota el actual **nivel de expansión** (la iteración) de  $X$ .

En la siguiente figura vemos un ejemplo de expansión de influencia, en un grafo con los vértices etiquetados y los arcos ponderados. Usaremos como conjunto inicial de nodos activados  $X = \{a\}$ .

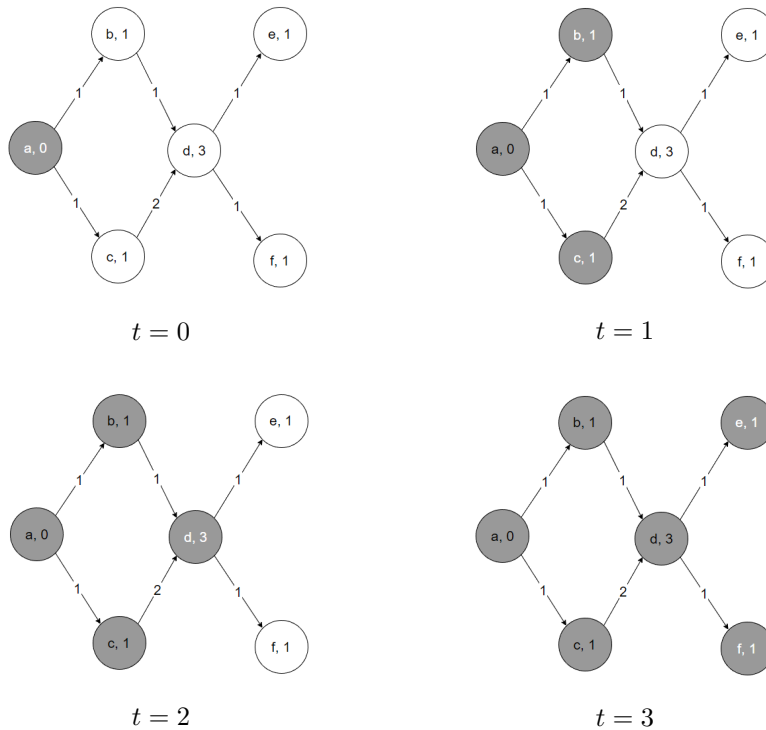


Figura 4.1: Ejemplo de expansión de influencia en el LTM.

### 4.1.3. Linear Threshold Rank

Como ya hemos mencionado anteriormente, una medida de centralidad asigna un valor a cada actor para medir su relevancia en la red. En [15] se introdujo una medida de centralidad basada en el LTM, el LTR (*Linear Threshold Rank*).

Esta centralidad nos indica qué actores son los más importantes en una red social, basándose en su capacidad de ejercer influencia bajo el modelo LTM. Su definición formal en [15] es la siguiente:

**Definición 3.** Sea  $(G, w, f)$  un grafo de influencia, con  $G = (V, E)$ ,  $n = |V|$ , y  $i \in V$  un actor. Denotamos el **Linear Threshold Rank** de  $i$  por  $LTR(i)$ , y es dado por

$$LTR(i) = \frac{|F(\{i\} \cup N(i))|}{n}$$

donde  $N(i) = \{j \in V | (i, j) \in E \vee (j, i) \in E\}$ , los vecinos del actor  $i$ .

En otras palabras, el LTR de un actor es el número total de actores influenciados bajo el modelo LTM, usando como conjunto inicial de actores activados a sí mismo y a sus vecinos directos (tanto *sucesores* como *predecesores*), dividido entre el número total de actores en la red.

Los valores que obtienen los actores en el grafo de la Figura 4.1 usando LTR son los siguientes:

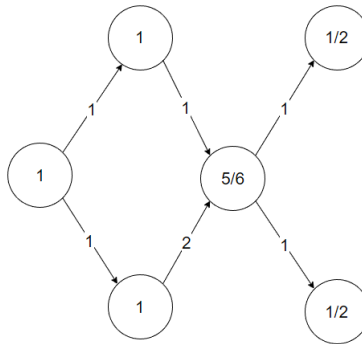


Figura 4.2: Rankings obtenidos con LTR

Vemos como los nodos  $\{a, b, c\}$  obtienen el valor mayor, mientras que los nodos  $\{e, f\}$  el menor.

#### 4.1.4. Forward / Backward Linear Threshold Rank

Las dos primeras medidas de centralidad que introducimos son el FWLTR (*Forward Linear Threshold Rank*) y el BWLTR (*Backward Linear Threshold Rank*). Están basadas en el LTR, modificando ligeramente su definición de vecindad. Sus definiciones formales son las siguientes:

**Definición 4.** Sea  $(G, w, f)$  un grafo de influencia, con  $G = (V, E)$ ,  $n = |V|$ , y  $i \in V$  un actor. Denotamos el **Forward Linear Threshold Rank** de  $i$  por  $FWLTR(i)$ , y es dado por

$$FWLTR(i) = \frac{|F(\{i\} \cup N_f(i))|}{n}$$

donde  $N_f(i) = \{j \in V | (i, j) \in E\}$ .

**Definición 5.** Sea  $(G, w, f)$  un grafo de influencia, con  $G = (V, E)$ ,  $n = |V|$ , y  $i \in V$  un actor. Denotamos el **Backward Linear Threshold Rank** de  $i$  por  $BWLTR(i)$ , y es dado por

$$BWLTR(i) = \frac{|F(\{i\} \cup N_b(i))|}{n}$$

donde  $N_b(i) = \{j \in V | (j, i) \in E\}$ .

Básicamente en FWLTR y BWLTR usamos como conjunto inicial, respectivamente, o bien los vecinos sucesores directos, o bien los vecinos predecesores directos.

Los valores que obtienen los actores en el grafo de la Figura 4.1 usando esas dos medidas son los siguientes:

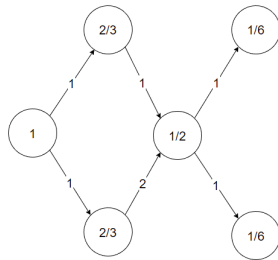


Figura 4.3: FWLTR

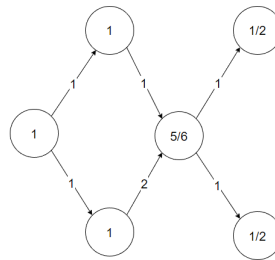


Figura 4.4: BWLTR

Podemos ver como el FWLTR y el BWLTR asignan valores diferentes, pero que el BWLTR coincide con el LTR. Nótese que en un grafo no dirigido las tres medidas siempre coinciden.



### 4.1.5. Fading Linear Threshold Model

En [16] se introduce un nuevo modelo de expansión de influencia basado en el LTM, el FLTM (*Fading Linear Threshold Model*). En este modelo, la influencia que ejerce un actor se devalúa en el tiempo. En el primer intento de influir a un actor, se ejerce la máxima influencia posible, pero conforme avanza el tiempo, decrece esa capacidad influencia. En el modelo esa capacidad decrece exponencialmente con  $\gamma \in [0, 1]$ .

Así pues, dado un grafo de influencia  $(G, w, f)$  y  $X \subseteq V$  el conjunto inicial de activación, vamos a considerar el siguiente proceso iterativo de activación. Sea  $F_t(X) \subseteq V$  el conjunto de nodos activados en una iteración  $t$ . Con la misma definición que en el LTM. En la iteración  $t + 1$  un nodo  $i \in V$  será activado si y solo si

$$\sum_{j \in F_t(X)} \gamma^{(t-t_j)} w_{ji} \geq f(i)$$

siendo  $t_j$  la iteración en que el actor  $j$  es activado. Igual que en el LTM terminamos el proceso anterior cuando dejan de ocurrir nuevas activaciones. Es el mismo proceso de expansión de influencia definido en la Definición 2, cambiando la formula de activación de los actores.

Nótese que si  $\gamma = 1$  obtenemos el LTM. Por otra parte, si  $\gamma = 0$ , solo en la iteración  $t_j + 1$ , una iteración después de su activación, un actor podrá ejercer su influencia. Es decir, cada actor solo tiene una oportunidad de activar a otro actor vecino.

Este modelo nos permite representar situaciones en las que el tiempo que se tarde en influenciar un actor sea algo relevante, atenuando la influencia que ejerce cada actor en el tiempo. En las siguientes dos figuras, vemos dos ejemplos triviales de expansión de influencia de un actor sobre un mismo grafo, utilizando los modelos LTM y FLTM con  $\gamma = 0$ , empezando en ambos casos con el conjunto inicial  $X = \{a\}$ .

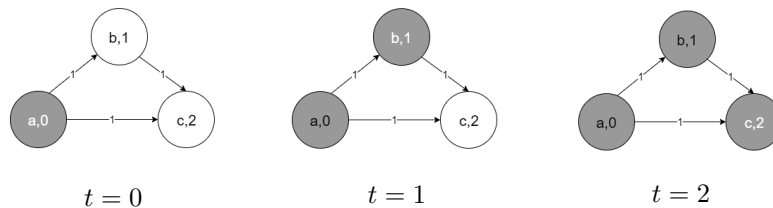


Figura 4.5: LTM.

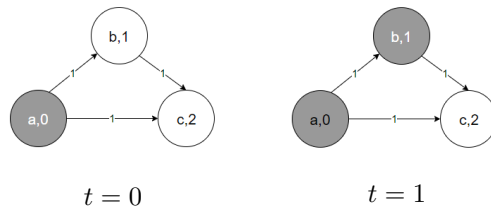


Figura 4.6: FLTM.

Como puede verse, en determinados grafos de influencia, la influencia que ejerce un actor puede cambiar según el modelo de expansión y los parámetros que se usen.

Modificaremos las medidas de centralidad vistas anteriormente, utilizando como modelo el FLTM en vez del LTM. Las cuales definimos como: FLTR (*Fading Linear Threshold Rank*), FWFLTR (*Forward Fading Linear Threshold Rank*) y BWFLTR (*Backward Fading Linear Threshold Rank*). Más adelante, observaremos experimentalmente que diferencias hay entre las medidas que usan un modelo u otro, usando diferentes valores de  $\gamma \in [0, 1]$ .

#### 4.1.6. Bounded Linear Threshold Model

Vamos a introducir una generalización del LTM, el BLTM (*Bounded Linear Threshold Model*). En este modelo, los actores tienen un tiempo máximo para influenciar al resto de actores. Para ello, limitamos el nivel de expansión a una constante  $p$  como parámetro. Definimos formalmente la expansión de influencia del BLTM como:

**Definición 6.** Sea  $(G, w, f)$  un grafo de influencia, y la constante  $p$  el nivel máximo de influencia, la **expansión de influencia** de  $X$  es

$$F(X) = \bigcup_{t=0}^k F_t(X) = F_0 \cup \dots \cup F_k(X)$$

donde  $k = \min\{\min\{t \in \mathbb{N} | F_t(X) = F_{t+1}(X)\}, p\} \leq n$ .

En la siguiente figura vemos el resultado de expandir la influencia, usando como conjunto inicial  $X = \{a\}$ , con diferentes valores de  $p$ .

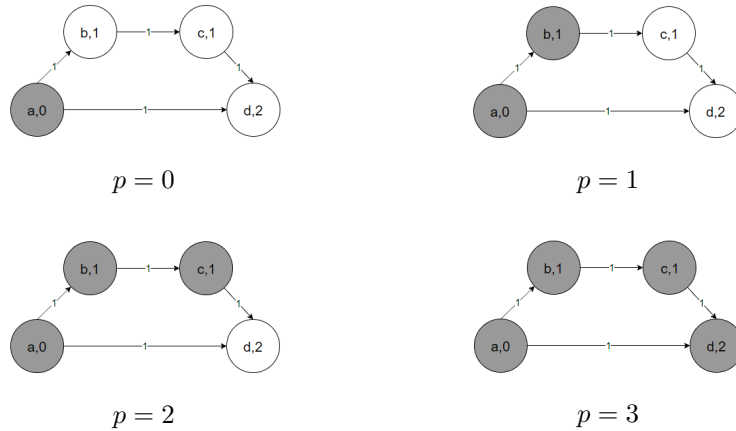


Figura 4.7: Nodos influenciados bajo el modelo BLTM, con diferentes valores de  $p$ .

Como podemos ver, el nodo  $a$  es capaz de influenciar al resto de nodos en el grafo en condiciones normales. Pero al limitar el tiempo para influenciar al resto de nodos, limitando el nivel de expansión, obtenemos resultados diferentes.

Igual que en el FLTM, modificaremos las medidas de centralidad vistas anteriormente, utilizando como modelo el BLTM en vez del LTM. Las cuales definimos como: BLTR (*Bounded Linear Threshold Rank*), BFWLTR (*Bounded Forward Linear Threshold Rank*) y BBWLTR (*Bounded Backward Threshold Rank*). Más adelante, observaremos experimentalmente que diferencias hay entre las medidas que usan un modelo u otro, usando diferentes valores de  $p \in [1, 10]$ .

### 4.1.7. Otras medidas de centralidad

Vamos a comparar las medidas de centralidad que hemos introducido, con otras ya bien conocidas, para así poder ver sí y qué diferencias hay entre ellas. Así pues, vamos a definir las que hemos usado, empezando por una de las más simples, la de grado **Degree** [2].

**Definición 7.** Dado un grafo dirigido  $G = (V, E)$ , para un nodo cualquiera  $v \in V$ , la **centralidad de grado** de  $v$  es el propio grado de  $v$ , dividido entre el número total de nodos menos uno. Es decir

$$\text{Degree}(v) = \frac{|\{(v, u) \in E\} \cup \{(u, v) \in E\}|}{n - 1}$$

La siguiente es la cercanía **Closeness** [2]. En esta se da más importancia a los nodos que están a menos distancia del resto de nodos.

**Definición 8.** Dado un grafo dirigido  $G = (V, E)$ , para un nodo cualquiera  $v \in V$ , la **cercanía** es la suma recíproca de las distancias a todo el resto de nodos  $u \in V \setminus \{v\}$ . Es decir

$$\text{Closeness}(v) = \frac{nr - 1}{n - 1} \frac{nr - 1}{\sum_{u=1}^{n-1} d(v, u)}$$

donde  $nr$  es el número de nodos accesibles desde  $v$  y  $d(v, u)$  la longitud del camino más corto entre  $v$  y  $u$ .

Para calcularla usaremos la fórmula de *Wasserman* y *Faust* [18], la cual tiene en cuenta los grafos con más de una componente conexa.

En la siguiente, la intermediación **Betweenness** [2], un nodo es más importante según la frecuencia en que actúa como un puente en el camino más corto entre dos nodos cualquiera.

**Definición 9.** Dado un grafo dirigido  $G = (V, E)$ , para un nodo cualquiera  $v \in V$ , definimos la **intermediación** es la suma, de todos los pares posibles  $s, t \in V$ , de la fracción de los caminos más cortos entre  $s, t$  que pasan por  $v$ , partido entre el número total de caminos más cortos entre  $s$  y  $t$ . Es decir

$$\text{Betweenness}(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

donde  $\sigma_{st}$  es el numero total de caminos más cortos entre  $s$  y  $t$ , y  $\sigma_{st}(v)$  el numero de caminos más cortos entre  $s$  y  $t$  que pasan por  $v$ .

La medida de centralidad de **Katz** [8] se basa en la medida de centralidad del vector propio. A partir de dos parámetros constantes  $\alpha$  y  $\beta$ , Katz calcula la influencia relativa de un nodo en la red, midiendo el numero de vecinos inmediatos (nodos de primer grado) y el resto de nodos en la red que se conectan al nodo en cuestión, a través de estos vecinos inmediatos. Sin embargo, las conexiones entre vecinos distantes se penalizan con un factor de atenuación constante  $\alpha$ . A cada ruta o conexión entre un par de nodos se le asigna un peso determinado por  $\alpha$  y la distancia  $d$  entre los nodos como  $\alpha^d$ . Además, se puede proporcionar un peso extra a los vecinos directos a partir del parámetro  $\beta$ .

**Definición 10.** Sea  $A$  la matriz de adyacencia de un grafo dirigido  $G = (V, E)$ , con los elementos  $a_{vu} = 1$  si el nodo  $v$  está conectado con el nodo  $u$ , y  $a_{vu} = 0$  en caso contrario. El **Katz** de un nodo  $v \in V$  es

$$Katz(v) = \alpha \sum_{u \in V} A_{vu} Katz(u) + \beta$$

donde  $\alpha < \frac{1}{\lambda_{max}}$  representa el factor de atenuación, siendo  $\lambda_{max}$  el valor propio más grande de  $A$ , y  $\beta$  la centralidad inicial de cada nodo.

Otra centralidad basada en la medida de centralidad del vector propio es el famoso **PageRank** [19]. El **PageRank** se diseñó para determinar la relevancia de una página web en internet. Una web será importante si contiene información relevante o si da acceso a otras páginas con información relevante. **PageRank** utiliza un parámetro  $\alpha \in [0, 1]$ , el cual representa la probabilidad de que un usuario continúe navegando por la red a través de los enlaces de la página actual, mientras que  $1 - \alpha$  representa la probabilidad de que el usuario salte a otra página aleatoria en la red.

**Definición 11.** Dado un grafo dirigido  $G = (V, E)$ , el **PageRank** de un actor  $v \in V$  es dado por

$$PR(v) = (1 - \alpha) + \alpha \sum_{(u,v) \in E} \frac{PR(u)}{\delta^+(u)}$$

donde  $\delta^+(v)$  es el grado de salida de un nodo  $v \in V$  y  $\alpha$  el factor de amortiguación tal que  $0 < \alpha \leq 1$ .

Finalmente, la medida de centralidad ICR (*Independent Cascade Rank*) [11], esta basada en el otro modelo de difusión de influencia, el ICM (*Independent Cascade Model*). En este modelo, cada actor activado tiene una única oportunidad de activar a un actor vecino no activado (con una probabilidad  $p$  de éxito). Cuando un actor es activado se repite el proceso para ese actor. El proceso termina cuando no quedan más actores que puedan influenciar a otros. Así pues, el ICR nos indica a cuantos actores es capaz de influenciar un actor bajo el ICM con una probabilidad  $p$ .

**Definición 12.** Dado un grafo dirigido  $G = (V, E)$  y una probabilidad  $p \in [0, 1]$ , el **ICR** de un actor  $v \in V$   $ICR(v, p)$ , es dado por

$$ICR(v, p) = \frac{|F'(v, p)|}{\max\{|F'(u, p)| : u \in V\}}$$

donde  $F'(v, p)$  es el conjunto de actores influenciados bajo el modelo ICM empezando por el actor  $v$ .

#### 4.1.8. Estadísticos

Para poder comparar dos medidas de centralidad distintas, utilizaremos diferentes estadísticos. La correlación estadística es una de las técnicas más usadas comúnmente en estos casos para comparar dos clasificaciones. Nosotros, igual que en [15, 16], utilizaremos los coeficientes de correlación de Spearman  $\rho$  y Kendall  $\tau$ .

**Definición 13.** Sean  $x$  e  $y$  dos listas de  $n$  actores, siendo  $x_i, y_i$  el ranking del actor  $i$  obtenido en las listas  $x, y$ , los coeficientes se definen como:

$$\rho = 1 - \frac{6 \sum_{i=1}^n (x_i - y_i)^2}{n(n^2 - 1)}$$

$$\tau = \frac{n_c - n_d}{n(n-1)/2}$$

donde  $n_c$  representa el numero de pares concordantes

$$n_c = |\{i, j \in [1, n] : (i < j) \wedge ((x_i > x_j \wedge y_i > y_j) \vee (x_i < x_j \wedge y_i < y_j))\}|$$

y  $n_d$  el numero de pares discordantes

$$n_d = |\{i, j \in [1, n] : (i < j) \wedge ((x_i > x_j \wedge y_i < y_j) \vee (x_i < x_j \wedge y_i > y_j))\}|$$

Estos coeficientes nos dan unos valores dentro del intervalo  $[-1, 1]$ , donde 1 significa que las dos listas son iguales y -1 que son inversas. Mientras que **Spearman** considera las distancias de los rankings, **Kendall** solo tiene en cuenta su orden.

A parte de esos dos coeficientes de correlación, usaremos el coeficiente **Gini**, como en [16]. El coeficiente de **Gini** es una medida de desigualdad utilizado en economía. El coeficiente nos indica del 0 al 1, como de desigual es un conjunto de datos. A valores más grandes, más desigual es el conjunto de datos.

**Definición 14.** Dado una lista  $x$  con  $n$  elementos, definimos el coeficiente de *Gini* como

$$Gini(x) = \frac{\sum_{i=1}^n \sum_{j=1}^n |x_i - x_j|}{2n \sum_{i=1}^n x_i}$$

#### 4.1.9. Migración de actores

Para profundizar en el análisis comparativo de medidas de centralidad, clasificaremos los actores en diferentes categorías según el la posición o valor que tengan en una medida de centralidad determinada. Al cambiar de medida, los actores pueden cambiar de categoría. Vamos a observar esa migración de actores entre categorías, al pasar de una centralidad a otra. Este tipo de análisis es común en estudios de sociología y economía sobre desigualdad [20]. En nuestro caso utilizaremos dos clasificaciones distintas.

Fijada una medida de centralidad, en la primera clasificación, clasificaremos los actores de la misma manera que en [16], ordenándolos de forma descendente según su ranking. Clasificaremos los actores en tres categorías diferentes (A,B y C). Partiendo de la lista de actores ordenada anterior, la primera categoría A consiste en el 10 % de los primeros actores, la categoría B el 40 % de los siguientes actores, finalmente la categoría C el 50 % de los actores restantes.

En la siguiente figura vemos un ejemplo de esta clasificación:

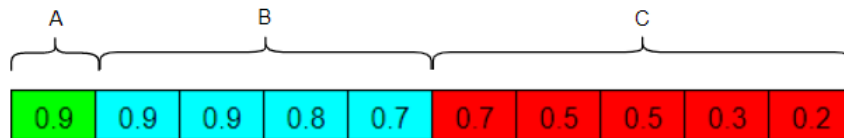


Figura 4.8: Clasificación por posición.

En la segunda, también clasificamos los actores en tres categorías (A,B y C), pero ahora según su valor en el ranking. En la primera tendremos los actores con un ranking mayor o igual al cuantil  $p = 0.9$ , en la siguiente aquellos con un valor inferior al cuantil  $p = 0.9$  y mayor al  $p = 0.5$ , en la última tendremos todos aquellos con un valor menor o igual al cuantil  $p = 0.5$ .

En la siguiente figura vemos un ejemplo de esta clasificación usando los mismos datos que en el ejemplo anterior:

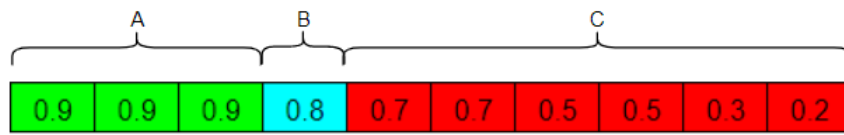


Figura 4.9: Clasificación por valor.

Nótese que a diferencia de la primera clasificación, el tamaño de las categorías puede cambiar.



## 4.2. Cálculo de las medidas introducidas

En esta sección explicaremos el algoritmo que hemos diseñado para poder calcular las diferentes medidas de centralidad introducidas, y como se han implementado.

### 4.2.1. Diseño del algoritmo

Dado un grafo de influencia  $(G, w, f)$ , un subconjunto inicial de actores  $X \subseteq V$  y una función de influencia  $f_i$ , nuestro algoritmo expande la influencia ejercida por el conjunto inicial de actores  $X$ , y nos devuelve el número total de actores que han sido activados. La función de influencia nos permitirá expandir la influencia bajo los tres modelos que hemos visto anteriormente (LTM, FLTM y BLTM).

Para calcular la expansión de influencia, haremos un recorrido en anchura partiendo del conjunto inicial  $X$ . Para ello, asociaremos diferentes atributos a cada actor  $v \in V$ . En el atributo  $v.influence$  guardaremos la influencia total que recibe por parte de sus vecinos, es decir, la suma de los pesos de las aristas  $(u, v) \in E$ , siendo  $u$  un actor activado. Por otra parte, el atributo  $v.lsl$  nos indica, en caso de que no este activado, el nivel de expansión en el que lo han intentado influenciar por última vez, y en caso contrario, en que nivel ha sido activado. Finalmente, el atributo  $v.influenced$  nos indica si el actor ha sido activado o no.

Así pues, para cada actor  $u \in V$  activado, recorreremos todos los actores  $v \in V$  adyacentes a él. Si el actor  $v$  no está activado lo influenciamos con la función  $f_i$ , si una vez influenciado, la influencia que recibe  $v.influence$  es mayor o igual a su valor umbral  $f(v)$  entonces el actor  $v$  pasa a estar activado. Repetiremos este proceso hasta que no se puedan activar más nodos.

Nuestro algoritmo *Expand\_Influence* descrito más abajo, asume que el grafo de influencia  $(G, w, f)$  de entrada esta representado por una lista de adyacencia. El algoritmo utiliza una cola  $Q$  FIFO para manejar el conjunto de actores activos que aún no han sido tratados.

```

Expand_Influence( $G, w, f, X, f_i$ )
1:  $total = 0$ 
2:  $Q = \emptyset$ 
3: for each  $v \in G.V \setminus X$  do
4:    $v.influenced = \text{FALSE}$ 
5:    $v.influence = 0$ 
6:    $v.lsl = -1$ 
7: end for
8: for each  $v \in X$  do
9:    $v.influenced = \text{TRUE}$ 
10:   $v.influence = 0$ 
11:   $v.lsl = 0$ 
12:   $total = total + 1$ 
13:  ENQUEUE( $Q, v$ )
14: end for
15: while  $Q \neq \emptyset$  do
16:   $v = \text{DEQUEUE}(Q)$ 
17:  for each  $u \in G.Adj[v]$  do
18:    if not  $u.influenced$  then
19:       $u.influence = u.influence + f_i(w(v, u), v, u)$ 
20:      if  $u.influence \geq f(u)$  then
21:         $u.influenced = \text{TRUE}$ 
22:         $total = total + 1$ 
23:        ENQUEUE( $Q, u$ )
24:      end if
25:       $u.lsl = v.lsl + 1$ 
26:    end if
27:  end for
28: end while
29: return  $total$ 

```

El funcionamiento del algoritmo es el siguiente. La línea 1 inicializa el contador actores activados totales a 0. La línea 2 inicializa  $Q$  como una cola vacía. Las líneas 3-7 inicializan los actores que no forman parte del conjunto inicial de activación, indicando que no están activados, que reciben 0 influencia y que el nivel de expansión del último actor que lo ha intentado activar es -1. En las líneas 8-14 inicializamos los actores del conjunto inicial, indicando que están activados, que reciben 0 influencia y que han sido activados en el nivel de expansión 0. También incrementamos el contador de actores activados totales y añadimos a cada uno de ellos a la cola de actores activos  $Q$ .

El bucle **while** de las líneas 15-28 itera mientras haya actores activos en la cola, los cuales aún no han intentado influenciar a sus actores vecinos no activos. Este bucle **while** mantiene el siguiente invariante:

En el test de la línea 15, la cola  $Q$  consiste en los actores activados que aún no han intentado influenciar a sus vecinos no activos.

Es fácil ver que mantenemos el invariante antes de la primera iteración y que cada iteración del bucle sigue manteniendo el invariante. Antes de la primera iteración, los actores que están en  $Q$  son los del conjunto inicial de activación, de los cuales ninguno de ellos ha intentado influenciar a sus vecinos no activos. En la línea 16 obtenemos el actor  $v$  que está en la cabeza de la cola y lo removemos de ella. El bucle **for** de las líneas 17-27 tratan a cada actor  $u$  adyacente de  $v$ . Si el actor  $u$  ya está activado no hacemos nada. Por otra parte, si el actor no ha sido activado aún, en la línea 19 lo intentamos influenciar incrementando la influencia que recibe mediante la función de influencia  $f_i$ . Si entonces resulta que la influencia que recibe el actor  $u$  en total, es superior o igual al valor de umbral que le ha asignado la función  $f$ , las líneas 20-24 activan el actor  $u$ , incrementan el contador de actores activados totales y lo añaden al final de la cola  $Q$ . Finalmente, la línea 25 actualiza el nivel de expansión  $u$  en que ha sido activado o lo han intentado activar. El invariante del bucle se mantiene, ya que cuando activamos un actor (en la línea 21) también lo añadimos a la cola (línea 23), y cuando lo sacamos de la cola, intentamos activar a todos sus vecinos no activos, y no vuelve a entrar en la cola durante el resto de la ejecución.

Finalmente, una vez salimos del bucle **while**, en la línea 29, devolvemos el número total de actores activados.

La búsqueda por profundidad (BFS) tiene la propiedad de que conforme añadimos los vértices visitados a la cola, su distancia respecto al vértice inicial crece monótonamente. Análogamente, en nuestro algoritmo, el nivel de expansión de los actores activados crece monótonamente conforme se añaden a la cola. Como veremos más adelante, esta propiedad nos servirá para poder calcular eficientemente las medidas de centralidad que utilizan el modelo FLTM, ya que no tendremos que recorrer a todos los actores que anteriormente han intentado influenciar a un actor y reducir la influencia que ejerce.

Para demostrar la propiedad anterior, primero demostraremos que en el test de la línea 15, los actores activos en la cola  $Q$  tienen como mucho dos valores distintos de nivel de expansión.

**Lema 1.** *Supongamos que durante la ejecución de `Expand_Influence` en un grafo de influencia  $(G, w, f)$ , la cola  $Q$  contiene los actores activados  $\langle v_1, v_2, \dots, v_r \rangle$ , donde  $v_1$  es la cabeza de  $Q$  y  $v_r$  la cola. Entonces tenemos que,  $v_r.lsl \leq v_1.lsl + 1$  y  $v_i.lsl \leq v_{i+1}.lsl$  para  $i = 1, 2, \dots, r - 1$ .*

**Demostración** Demostraremos el lema anterior por inducción sobre el numero de operaciones de la cola. Inicialmente, cuando la cola solo contiene el conjunto de activación inicial, el lema se mantiene ( $v.lsl = 0$  para todos los actores en  $Q$ ).

Para el paso inductivo, debemos demostrar que el lema se mantiene después de añadir o sacar un actor de la cola. Si sacamos de la cabeza de la cola el actor  $v_1$ , el actor  $v_2$  se convierte en la nueva cabeza de la cola. Si la cola pasa a estar vacía el lema se mantiene. En caso contrario, por hipótesis de inducción tenemos que  $v_1.lsl \leq v_2.lsl$ . Luego tenemos que  $v_r.lsl \leq v_1.lsl + 1 \leq v_2.lsl + 1$ , con el resto de igualdades cumpliéndose. Así pues, con  $v_2$  como cabeza de la cola el lema se sigue cumpliendo.

Cuando añadimos un actor  $u$  recién activado a la cola en la línea 23, el actor pasa a ser el  $v_{r+1}$  de la cola. En ese momento ya hemos sacado de la cola al actor  $v$ , del cual en este momento estamos intentando influenciar a todos sus actores vecinos no activos. De la cola  $Q$ , y por hipótesis de inducción, tenemos que la nueva cabeza  $v_1.lsl \geq v.lsl$ . Entonces, después de actualizar el nivel de expansión en la línea 25, tenemos que  $v_{r+1}.lsl = u.lsl = v.lsl + 1 \leq v_1.lsl + 1$ . Luego también por hipótesis de inducción, tenemos que  $v_r.lsl \leq v.lsl + 1$ , es decir  $v_r.lsl \leq v.lsl + 1 = u.lsl = v_{r+1}.lsl$ , con el resto de igualdades cumpliéndose. Así pues, el lema se mantiene después de añadir el actor  $u$  a la cola y actualizar el nivel de expansión en la línea 25. ■

Finalmente, el siguiente corolario demuestra la propiedad anterior.

**Corolario 1.** *Supongamos que los actores  $v_i$  y  $v_j$  son añadidos a la cola durante la ejecución, y que  $v_i$  se añade antes que  $v_j$ . Entonces tenemos que  $v_i.lsl \leq v_j.lsl$  después de añadir  $v_j$  a la cola y actualizar su nivel de expansión.*

**Demostración** Inmediata, a partir del Lema 1 y de que un actor activo solo se añade una única vez a la cola durante la ejecución del algoritmo. ■

Vamos a ver las funciones de influencia  $f_i$  para los tres modelos diferentes de expansión de influencia. Para el modelo LTM usaremos la siguiente función trivial *normal\_influence*.

Normal\_influence( $w_{vu}, v, u$ )

1: **return**  $w_{vu}$

Para el modelo BLTM también utilizaremos una función sencilla. Simplemente si el nivel de expansión del actor  $v$  es mayor que el valor  $p$  el actor no ejercerá ninguna influencia al actor  $u$ , en caso contrario, ejercerá su influencia de forma normal.

Bounded\_influence( $w_{vu}, v, u$ )

```

1: if  $v.lsl > p$  then
2:   return 0
3: end if
4: return  $w_{vu}$ 

```

Finalmente, para el FLTM usaremos la siguiente función.

Fading\_influence( $w_{vu}, v, u$ )

```

1: if  $u.lsl \neq v.lsl + 1$  then
2:    $u.influence = u.influence \cdot \gamma$ 
3: end if
4: return  $w_{vu}$ 

```

Vamos a ver que, efectivamente, con la función anterior, calculamos la expansión de influencia del FLTM correctamente. De la definición de activación de nodos bajo el modelo FLTM vista en la sección 4.1.5, se deduce que la influencia que recibe un nodo  $i$  en el nivel de expansión  $t + 1$  es

$$\sum_{j \in F_t(X)} \gamma^{(t-t_j)} w_{ji}$$

dado un nodo  $i$  y un nivel de expansión  $t$ , definimos  $w_i^t$  como

$$w_i^t = \sum_{j \in F_t(X) \setminus F_{t-1}(X)} w_{ji}$$

es decir, la influencia total que recibe el nodo  $i$  de los nodos activados en el nivel de expansión  $t$ . Entonces, podemos expresar la expresión anterior de influencia, con el siguiente polinomio.

$$\begin{aligned}
& w_i^{(0)} \cdot \gamma^{(t)} + w_i^{(1)} \cdot \gamma^{(t-1)} + \dots + w_i^{(t-1)} \cdot \gamma^{(1)} + w_i^{(t)} \cdot \gamma^{(0)} = \\
& = \sum_{x=0}^t w_i^x \cdot \gamma^{(t-x)}
\end{aligned}$$

Vamos a demostrar que después de ejecutar la línea 2 de la función, la influencia que recibe un actor no activo  $i$  en el nivel de expansión  $t + 1$  es

$$\sum_{x=0}^{t-1} w_i^x \cdot \gamma^{(t-x)}$$

**Demostración** Lo demostraremos por inducción sobre el nivel de expansión de influencia. Inicialmente, en el nivel  $0 + 1$ , justo después de añadir los actores iniciales a la cola, para todos los nodos activados iniciales  $v$  y todos los vecinos actores no activados  $u$ . Al ejecutar la condición de la línea 1, tenemos que, o bien el actor  $v$  es el primero en intentar influenciar al actor  $u$  en este nivel de expansión, o bien otro actor ha intentado influenciar al actor  $u$  anteriormente. Por el corolario 1, podemos asegurar que para cada actor no activado  $u$ , solo un nodo por nivel de expansión podrá ejecutar la línea 2. Así pues solo si el actor  $v$  es el primero en intentar influenciar al nodo  $u$  en este nivel de expansión, podrá ejecutar la línea 2. En este caso la influencia que recibe hasta ahora el actor  $u$  es 0, y al ejecutar la línea 2 tenemos que  $u.influence = 0 \cdot \gamma = 0$ , cumpliéndose

$$\sum_{x=0}^{0-1} w_u^x \cdot \gamma^{(t-x)} = 0$$

Al actualizar el nivel de influencia en la línea 25 de *Expand\_influence*, garantizamos que ningún otro actor del mismo nivel vuelva a ejecutar la línea 2 de la función.

Para el paso inductivo, tenemos que en el nivel de expansión  $t + 1$ , por hipótesis de inducción, la influencia que recibe cada actor  $u$  no activado es

$$\sum_{x=0}^{t-2} w_u^x \cdot \gamma^{(t-1-x)}$$

Como en nuestra función devolvemos siempre la influencia  $w_{vu}$  sin alterar, todos los actores ejercen influencia de manera normal. Por lo tanto, añadimos a la influencia anterior la influencia  $w_u^{t-1}$  que han ejercido los actores activos en el nivel anterior  $t$ . Así pues, tenemos

$$\begin{aligned} & \sum_{x=0}^{t-2} w_u^x \cdot \gamma^{(t-1-x)} + w_u^{t-1} = \\ & = w_u^{(0)} \cdot \gamma^{(t-1)} + \dots + w_u^{(t-2)} \cdot \gamma^{(1)} + w_u^{(t-1)} \end{aligned}$$

Finalmente, como la línea 2 se ejecuta una vez para cada actor no activo  $u$  y cada nivel de expansión. Después de ejecutar la línea 2, la influencia será

$$\begin{aligned} & \gamma \cdot (w_u^{(0)} \cdot \gamma^{(t-1)} + \dots + w_u^{(t-2)} \cdot \gamma^{(1)} + w_u^{(t-1)}) = \\ & = w_u^{(0)} \cdot \gamma^{(t)} + \dots + w_u^{(t-2)} \cdot \gamma^{(2)} + w_u^{(t-1)} \cdot \gamma^{(1)} = \\ & = \sum_{x=0}^{t-1} w_u^x \cdot \gamma^{(t-x)} \end{aligned}$$

■

En cuanto a costes temporales, es fácil ver que las funciones descritas arriba tienen un coste temporal constante  $O(1)$ . Así pues, igual que en el BFS, nuestro algoritmo *Expand\_Influence* tiene un coste temporal  $O(E + V)$ . Es decir, el algoritmo *Expand\_Influence* se ejecuta en tiempo lineal respecto al tamaño de la lista de adyacencia que representa al grafo de influencia  $(G, w, f)$  de entrada.

#### 4.2.2. Implementación

Para implementar el algoritmo anterior, hemos decidido usar el lenguaje de programación C++, para poder ejecutar los experimentos de manera eficiente y en poco tiempo. Además, hemos paralelizado la ejecución de las medidas con la API `OpenMP`. En el anexo A.2 se encuentra la implementación del resto de medidas de centralidad.

Para representar un grafo de influencia, utilizamos la clase `Graph`. En el anexo A.1 adjuntamos su implementación. El resto de estructuras de datos que usamos son las siguientes:

```
typedef queue<int> QI;
typedef list<int> LI;
typedef vector<double> VD;

typedef struct
{
    double influence;
    bool influenced;
    int last_spread_level;

    Info(): influence(0), influenced(false), last_spread_level(-1) {}
}Info;
typedef vector<Info> VInfo;
```

Donde en el `struct Info` almacenamos los atributos de los actores necesarios para *Expand\_Influence*.

La implementación del algoritmo *Expand\_Influence* en C++ es bastante directa:

```
int Expand_Influence(const Graph &G, const LI &X, double (*fi)(double, int, int, VInfo &))
{
    int rank = X.size();
    int n = G.N;
    VInfo info(n, Info());
    QI Q;

    for (int v : X)
    {
        info[v].last_spread_level = 0;
        info[v].influenced = true;
        Q.push(v);
    }

    while (not Q.empty())
    {
        int v = Q.front(); Q.pop();
        for (PID arc : G.adjency[v])
        {
            int u = arc.first;
            double w = arc.second;
            if (not info[u].influenced)
            {
                info[u].influence += fi(w, v, u, info);
                if (info[u].influence >= G.threshold[u])
                {
                    info[u].influenced = true;
                    Q.push(u);
                    ++rank;
                }
                info[u].last_spread_level = info[v].last_spread_level + 1;
            }
        }
    }

    return rank;
}
```



La implementación de las funciones de influencia es aún más simple:

```
double Normal_influence(double w, int v, int u, VInfo &influence)
{
    return w;
}
double Bounded_influence(double w, int v, int u, VInfo &influence)
{
    if (influence[v].last_spread_level > p) return 0;
    else return w;
}
double Fading_influence(double w, int v, int u, VInfo &influence)
{
    if (influence[u].last_spread_level != influence[v].last_spread_level + 1)
    {
        influence[u].influence *= Gamma;
    }
    return w;
}
```

Donde las constantes  $p$  y  $\gamma$  son previamente definidas como variables globales.

Finalmente, para obtener el ranking de cada actor en una red social, con las medidas de centralidad que hemos introducido. Utilizamos la función *Rank\_Actors*, la cual recibe como parámetros un vector de `double` para almacenar las puntuaciones obtenidas de cada actor, el grafo de influencia  $G$  que representa la red social, una función `fn` que genera el conjunto inicial de actores a partir de un actor concreto y el grafo de influencia, y finalmente la función de influencia. En esta función, ejecutaremos de forma paralela la obtención de los rankings de cada actor en la red. Para ello, simplemente utilizamos la directiva de `OpenMP`, `#pragma omp parallel for`, en el bucle `for` que itera sobre todos los actores de la red.

```
void Rank_Actors(VD &ranking, const Graph &G, void (*fn)(LI&, const Graph &, int),
                double (*fi)(double, int, int, VInfo&))
{
    int n = G.N;
    ranking = VD(n, 0);
    #pragma omp parallel for
    for (int i = 0; i < n; ++i)
    {
        LI X;fn(X, G, i);
        ranking[i] = (double) Expand_Influence(G, X, fi) /n;
    }
}
```

## Capítulo 5

# Experimentos

En este capítulo, vamos a mostrar los resultados de los diferentes experimentos que hemos realizado con las medidas de centralidad introducidas previamente. Todos los experimentos han sido ejecutados en una maquina HP ProLiant DL380p server con dos Intel(R) Xeon(R) E5-2660. Primero de todo, introducimos las redes que se han utilizado.

### 5.1. Redes utilizadas

Tabla 5.1: Resumen de la redes. ACC = Average Clustering Coefficient, MC = tamaño del Main Core, D = dirigida o no, W = con pesos o no, Diam = diámetro de la red. Cuando la red no es conexa, indicamos que su diámetro es infinito y añadimos entre paréntesis el diámetro de la componente conexa más grande.

Red	V	E	D	W	ACC	Diam	MC
ArXiv	5242	14496	no	no	0.5296	inf(17)	44
Caida	26475	106762	si	si	0.2082	17	50
Dining Table	26	52	si	si	0.1178	inf(6)	20
Dolphins	62	159	no	no	0.2590	8	36
Epinions	75879	508837	si	no	0.1378	14	422
Higgs	256491	328132	si	si	0.0156	19	10
Human Brain	480	1000	no	si	0.3004	inf(20)	11
Wikipedia	7115	106762	si	no	0.1409	7	336
Texas	1379917	1921660	no	no	0.0470	1054	-

Inicialmente, utilizaremos las mismas redes que se usaron en [15, 16] para nuestros experimentos. En la tabla 5.1 están resumidas las características de estas

redes. El ACC (*Average Clustering Coefficient*), nos indica la media de *local clustering coefficient* de cada nodo, siendo el *local clustering coefficient* la proporción de las aristas entre sus vecinos inmediatos, entre el número de aristas totales que podrían existir entre ellos. El diámetro es la mayor distancia entre dos pares de nodos cualquiera en la red. El diámetro de un grafo es infinito cuando este no es (fuertemente) conexo. En este caso indicamos el diámetro de la componente conexas más grande. Finalmente, definimos el MC (*Main Core*) como el *k-core* con el grado más grande. Donde el *k-core* de un grafo, es el subgrafo maximal donde todos los vértices tienen un grado mínimo  $k$ .

Hemos utilizado las mismas redes que en [15, 16], ya que son fáciles de modelar como un grafo de influencia y tienen una gran variedad de tamaños y propiedades. Las redes ArXiv<sup>1</sup> [21], Caida<sup>2</sup> [22], Epinions<sup>3</sup> [23], Higgs<sup>4</sup> [24] Wikipedia<sup>5</sup> [25, 26] y Texas<sup>6</sup> [27] las hemos obtenido de la colección de *Stanford Large Network Dataset Collection* [28]. La red Texas, es una nueva incorporación que ha sido seleccionada por su gran diámetro y tamaño. En cuanto a las otras redes más pequeñas, la red Dining Table<sup>7</sup> [29] la obtenemos en la colección de Pajek, la red Dolphins<sup>8</sup> [30] está disponible en el repositorio *The Network Data Repository* [31]. Finalmente, la red Human Brain<sup>9</sup> [32, 33] la conseguimos en el *PIT Bioinformatics Group Department of Computer Science, Eötvös University*.

La primera red, ArXiv, contiene las colaboraciones entre autores, de papers sobre relatividad general, subidos en la plataforma ArXiv. Si un par de actores  $i, j$  han colaborado en un paper, se añade una arista en el grafo relacionando los dos actores.

La segunda red, Caida, representa un red de proveedores de servicios de internet (ISP) conectados entre si. Los datos vienen del proyecto CAIDA, recolectados en 2007. Los nodos de la red representan a ISPs, y los arcos entre ellos representan los contratos y relaciones de negocios entre ellos. Existen cuatro tipos de relaciones: 'es cliente de', 'es proveedor de', 'son compañeros', 'son hermanos'. Esta última relación se refiere a los ISP que son de una misma compañía. Hemos asignado un peso y una dirección a cada relación de acuerdo con la tabla 5.2

<sup>1</sup><https://snap.stanford.edu/data/ca-GrQc.html>

<sup>2</sup><https://snap.stanford.edu/data/as-Caida.html>

<sup>3</sup><https://snap.stanford.edu/data/soc-Epinions1.html>

<sup>4</sup><https://snap.stanford.edu/data/higgs-twitter.html>

<sup>5</sup><https://snap.stanford.edu/data/wiki-Vote.html>

<sup>6</sup><http://snap.stanford.edu/data/roadNet-TX.html>

<sup>7</sup><http://vlado.fmf.uni-lj.si/pub/networks/data/esna/dining.htm>

<sup>8</sup><http://networkrepository.com/soc-dolphins.php>

<sup>9</sup><https://pitgroup.org/connectome/>

tipo de relación	dirección del arco	peso
“a es cliente de b”	de b hacia a	1
“a es proveedor de b”	de a hacia b	1
“a y b son compañero”	ambas direcciones	1
“a y b son hermanos”	ambas direcciones	2

Tabla 5.2: Traducción de los datos de CAIDA a un grafo

La red de **Epinions**, representa las relaciones de confianza de los usuarios en la pagina web Epinions. Los usuarios de la web pueden decidir confiar en las reseñas de otros usuarios. De esta forma se decide que reseñas mostrar al usuario al consultar un producto.

La red de **Higgs**, contiene todos los tweets relacionados con el descubrimiento del Boson de Higgs, que fueron retweeteados entre el 1 y 7 de julio de 2012. En este grafo dirigido, un arco  $(i, j) \in E$  representa a una persona  $i$  retweeteando a un tweet de la persona  $j$ . El peso representa la frecuencia del retweet.

La quinta red, **Wikipedia**, representa la elección de los nuevos administradores de Wikipedia por parte de la comunidad. Los datos contienen todas las elecciones des de la creación de Wikipedia hasta en enero de 2008. Los nodos en la red representan a los usuarios de Wikipedia, y un arco de un nodo  $i$  a un nodo  $j$ , representa el voto del usuario  $i$  al usuario  $j$ .

La siguiente red, **Dining Table**, representa a las preferencias de compañeras de mesa de 26 chicas viviendo en una casa de campo en el *New York State Training School*. A cada chica se le pide que compañera prefiere tener al lado, indicando sus dos primeras opciones. Los nodos del grafo representan a las chicas. Habrá un arco  $(i, j)$ , si la chica  $j$  tiene preferencia por la chica  $i$ . Si es su primera opción, tendrá un peso de 2, en caso contrario tendrá un peso de 1.

La red **Dolphins**, es una red no dirigida que representa las relaciones entre delfines en una comunidad que vive en Nueva Zelanda.

La red de **Texas**, representa las conexiones no dirigidas entre carreteras del estado de Texas. Inicialmente no teníamos pensado usar esta red, pero por motivos que veremos más adelante, en un experimento necesitamos una red con un diámetro excepcionalmente grande, y la red de **Texas** tiene esta característica.

Finalmente, la última red, **Human Brain**, representa el grafo no dirigido de un

cerebro. Para crear este grafo, se han unificado 477 conectomas de distintas personas, computados a partir de las resonancias magnéticas del *Human Connectome Project*. Las regiones del cerebro se representan como nodos. El peso de una arista se calcula como el número de caminos entre dos regiones, dividido por la longitud media de estos caminos.

Representamos cada red como un grafo de influencia  $(G, w, f)$ . Para redes sin pesos, le asignamos a todos los arcos peso 1. Transformamos las redes no dirigidas, en redes dirigidas simétricas. Ninguna de las redes tiene los nodos etiquetados, igual que en [15], para cada actor  $i \in V$  le asignaremos el valor de umbral  $f(i) = \lfloor \bar{w}/2 \rfloor + 1$ , donde  $\bar{w} = \sum_{(j,i) \in E} w_{ji}$ . En este modelo, un actor es activado si pertenece al conjunto inicial de activación, o si la influencia de los nodos activos que apuntan a él, suma más que la mitad del peso total que apunta a él. Cuando todos los pesos valen 1, esto se puede interpretar como que 'un actor en la red es activado si la mayoría de sus vecinos lo están'.

## 5.2. Resultados BWLTR y FWLTR

En esta sección vamos a ver los resultados obtenidos de las medidas de centralidad BWLTR y FWLTR. También compararemos estos resultados con los obtenidos por el resto de centralidades con los coeficientes de correlación. Finalmente, observaremos las migraciones de nodos que ocurren al pasar de una medida a otra.

Para calcular el ICR usamos  $p = 0.1$  y lo ejecutamos 100 veces. Para el PageRank, usamos un factor de atenuación  $\alpha = 0.85$ . Finalmente, para Katz, calcular el valor propio máximo con un algoritmo sencillo tiene un coste temporal de  $O(n^3)$ , este coste resulta prohibitivo para redes con centenares de miles de nodos. Al evitar el cálculo de  $\lambda_{max}$ , podemos calcular Katz en tiempo  $O(n + m)$  [34]. Así pues, fijaremos  $\alpha$  y  $\beta$  con los valores  $\alpha = 0.1, \beta = 1.0$ . Además, para asegurarnos que Katz converge el mayor número de veces posibles, reduciremos la tolerancia de seis dígitos significativos a solo dos.

### 5.2.1. Resumen de los resultados

En la tabla 5.3 podemos ver un resumen de los resultados obtenidos. En ella vemos la desviación estándar de las asignaciones, el número de valores asignados diferentes, y el coeficiente de desigualdad gini. Como en [16], con el coeficiente de gini, buscamos ver como de equitativas son las medidas a la hora de asignar valores a los actores. En el coeficiente de gini como más cercano sea a 0, más equitativamente se reparten los valores entre los actores, por otra parte, como más cercano sea a 1, más desigualdades habrá entre los actores.

En redes no dirigidas, como era de esperar, obtenemos los mismos resultados para LTR, FWLTR y BWLTR, ya que en este tipo de redes las tres medidas asignan los mismos valores, al tener los conjuntos iniciales idénticos. Estas tres medidas, se encuentran entre las que asignan menos valores distintos y tienen una desviación estándar más baja, pero por otra parte, excepto en ArXiv, también son las que de forma más desigual asignan.

En redes dirigidas, el coeficiente de gini de FWLTR es más bajo que el de LTR y BWLTR, especialmente en el caso de Higgs. También en las redes dirigidas, excepto en Epinions, podemos observar que el FWLTR es el que asigna menos valores distintos y tiene una desviación estándar más pequeña. Parece ser pues, que FWLTR reparte de manera más igualitaria, pero que por otra parte distingue a menos actores. Curiosamente, en Caida, siendo una red dirigida, FWLTR y BWLTR coinciden, pero no con LTR. Seguramente esto se deba a la topología de la propia red, siendo Caida prácticamente simétrica.

Tabla 5.3: Desviación estándar ( $\sigma$ ), coeficiente de gini (gini), numero de valores distintos (#) para las medidas de centralidad sobre las diferentes redes. El símbolo '-' para la medida de centralidad de Katz, indica que el algoritmo no ha convergido.

network measure	ArXiv			Caida			Dining Table			Dolphins		
	$\sigma$	#	gini	$\sigma$	#	gini	$\sigma$	#	gini	$\sigma$	#	gini
BWLTR	0.006	169	0.540	0.015	1234	0.705	0.080	7	0.280	0.080	16	0.331
FWLTR	0.006	169	0.540	0.015	1234	0.705	0.038	4	0.128	0.080	16	0.331
LTR	0.006	169	0.540	0.048	2728	0.690	0.077	8	0.171	0.080	16	0.331
Betweenness	0.004	1569	0.885	0.002	9213	0.977	0.036	21	0.537	0.102	54	0.614
Closeness	0.057	2850	0.269	0.035	10939	0.075	0.140	19	0.460	0.052	43	0.095
Degree	0.003	65	0.555	0.003	158	0.628	0.066	6	0.215	0.048	12	0.325
ICR	0.244	3956	0.609	0.129	16834	0.447	0.038	13	0.022	0.182	51	0.170
Katz	-	-	-	-	-	-	0.032	20	0.086	0.047	60	0.221
Pagerank	1.32e-04	3526	0.322	2.69e-04	13329	0.545	0.051	21	0.559	0.008	60	0.274
network measure	Epinions			Higgs			Human Brain			Wikipedia		
	$\sigma$	#	gini	$\sigma$	#	gini	$\sigma$	#	gini	$\sigma$	#	gini
BWLTR	0.003	1363	0.855	1.65e-04	396	0.578	0.088	86	0.589	0.007	250	0.799
FWLTR	0.001	828	0.830	6.33e-06	52	0.218	0.088	86	0.589	0.006	234	0.766
LTR	0.004	1768	0.843	1.66e-04	392	0.442	0.088	86	0.589	0.014	393	0.767
Betweenness	8.11e-04	24976	0.941	1.59e-04	5134	0.998	0.061	240	0.884	4.61e-04	1376	0.942
Closeness	0.079	34130	0.423	0.004	4335	0.993	0.056	276	0.142	0.104	2158	0.694
Degree	6.94e-04	575	0.814	1.52e-04	375	0.569	0.016	25	0.524	0.008	308	0.752
ICR	0.265	32514	0.689	0.014	534	0.081	0.150	321	0.422	0.306	4417	0.663
Katz	-	-	-	0.002	4858	0.145	-	-	-	-	-	-
Pagerank	5.54e-05	35745	0.682	8.30e-05	9996	0.590	0.004	430	0.505	2.33e-04	2353	0.543

### 5.2.2. Correlación de medidas

Igual que en [15,16], comparamos dos medidas de centralidad distintas, utilizando los coeficientes de correlación **Spearman** y **Kendall** (véase la sección 4.1.8). En las tablas 5.4, 5.5 y 5.6 podemos ver los resultados obtenidos. Hemos comparado las medidas de centralidad **LTR**, **BWLTR**, **FWLTR** y el resto de medidas, para ver como de similares son entre ellas. Los valores en color rojo, son aquellos con un  $p$ -valor mayor que 0.05. Mientras que el coeficiente **Spearman** depende de la distancia relativa entre los rankings, el coeficiente de **Kendall** solo tiene en consideración el orden del ranking. Puede pasar que las distancias relativas cambien entre rankings, pero que el orden se mantenga, y que por tanto **Spearman** y **Kendall** den valores distintos.

Primero de todo vemos como en redes no dirigidas, las medidas **LTR**, **BWLTR** y **FWLTR** tienen una correlación positiva perfecta, ya que sus asignaciones de valores son las mismas en este tipo de redes. También vemos como en este tipo de redes, excepto en **Human Brain**, hay una correlación bastante alta entre el **LTR** y sus variantes con la centralidad de grado, con índices mayores que 0.75 tanto en **Spearman** como en **Kendall**.

En cuanto a las dirigidas, en general observamos como hay bastante correlación entre las medidas **LTR** y **BWLTR**, en todas las redes excepto en **Higgs**. También podemos ver que hay bastante correlación entre el **BWLTR** y **Closeness**, especialmente en la red de **Wikipedia**. En el caso de **FWLTR**, parece ser que la medida en que se parece más es el **ICR**, aunque no llega a obtener los índices de correlación tan altos vistos entre **BWLTR** y **Closeness**. Finalmente, en **Caida** se confirma el fenómeno que hemos visto en la tabla anterior, las medidas **FWLTR** y **BWLTR** tienen un índice de correlación perfecto en ambos coeficientes, es decir que sus asignaciones coinciden perfectamente. Como hemos dicho antes, esto probablemente se deba a características topológicas de la red **Caida**.



Tabla 5.4: Coeficientes de correlación para las medidas de centralidad aplicadas a las redes **ArXiv**, **Caida** y **Dining Table**. Los coeficientes de **Kendall** ( $\tau$ ) se muestran en la región triangular superior, los coeficientes de **Spearman** ( $\rho$ ) se muestran en la triangular inferior. Los valores con un  $p$ -valor mayor que 0.05 se muestran en rojo. Para **Katz**, el símbolo '-' indica que el algoritmo no ha convergido.

$\rho / \tau$	<b>ArXiv</b>									
	BWLTR	FWLTR	LTR	Betweenness	Closeness	Degree	ICR	Katz	Pagerank	
BWLTR	1	1	1	0.513	0.580	0.763	0.667	-	0.359	
FWLTR	1	1	1	0.513	0.580	0.763	0.667	-	0.359	
LTR	1	1	1	0.513	0.580	0.763	0.667	-	0.359	
Betweenness	0.638	0.638	0.638	1	0.430	0.572	0.404	-	0.529	
Closeness	0.748	0.748	0.748	0.554	1	0.452	0.778	-	0.070	
Degree	0.888	0.888	0.888	0.676	0.591	1	0.582	-	0.577	
ICR	0.840	0.840	0.840	0.529	0.922	0.735	1	-	0.146	
Katz	-	-	-	-	-	-	-	-	-	
Pagerank	0.511	0.511	0.511	0.668	0.130	0.724	0.243	-	1	
$\rho / \tau$	<b>Caida</b>									
	BWLTR	FWLTR	LTR	Betweenness	Closeness	Degree	ICR	Katz	Pagerank	
BWLTR	1	1	0.896	0.231	0.673	0.364	0.538	-	0.253	
FWLTR	1	1	0.896	0.231	0.673	0.364	0.538	-	0.253	
LTR	0.982	0.982	1	0.227	0.724	0.362	0.566	-	0.226	
Betweenness	0.289	0.289	0.284	1	0.232	0.719	0.338	-	0.652	
Closeness	0.862	0.862	0.902	0.299	1	0.399	0.625	-	0.193	
Degree	0.470	0.470	0.469	0.803	0.510	1	0.555	-	0.812	
ICR	0.735	0.735	0.765	0.445	0.816	0.683	1	-	0.326	
Katz	-	-	-	-	-	-	-	-	-	
Pagerank	0.376	0.376	0.339	0.808	0.308	0.926	0.491	-	1	
$\rho / \tau$	<b>Dining Table</b>									
	BWLTR	FWLTR	LTR	Betweenness	Closeness	Degree	ICR	Katz	Pagerank	
BWLTR	1	0.092	0.785	0.463	0.485	0.701	0.097	0.610	0.453	
FWLTR	0.097	1	0.197	-0.203	-0.212	-0.209	0.068	-0.208	-0.153	
LTR	0.857	0.229	1	0.402	0.471	0.583	0.083	0.542	0.425	
Betweenness	0.604	-0.235	0.511	1	0.751	0.768	-0.129	0.753	0.748	
Closeness	0.615	-0.240	0.587	0.890	1	0.721	-0.039	0.794	0.914	
Degree	0.788	-0.228	0.674	0.878	0.843	1	0.021	0.918	0.674	
ICR	0.102	0.078	0.096	-0.197	-0.062	0.037	1	0.010	-0.071	
Katz	0.735	-0.236	0.654	0.897	0.904	0.975	0.027	1	0.724	
Pagerank	0.582	-0.178	0.551	0.882	0.980	0.811	-0.113	0.875	1	

Tabla 5.5: Coeficientes de correlación para las medidas de centralidad aplicadas a las redes **Dolphins**, **Epinions** y **Higgs**. Los coeficientes de Kendall ( $\tau$ ) se muestran en la región triangular superior, los coeficientes de Spearman ( $\rho$ ) se muestran en la triangular inferior. Los valores con un  $p$ -valor mayor que 0.05 se muestran en rojo. Para Katz, el símbolo '-' indica que el algoritmo no ha convergido.

$\rho / \tau$	<b>Dolphins</b>								
	BWLTR	FWLTR	LTR	Betweenness	Closeness	Degree	ICR	Katz	Pagerank
BWLTR	1	1	1	0.568	0.412	0.797	0.665	0.723	0.776
FWLTR	1	1	1	0.568	0.412	0.797	0.665	0.723	0.776
LTR	1	1	1	0.568	0.412	0.797	0.665	0.723	0.776
Betweenness	0.724	0.724	0.724	1	0.578	0.665	0.597	0.570	0.658
Closeness	0.563	0.563	0.563	0.748	1	0.548	0.593	0.593	0.458
Degree	0.912	0.912	0.912	0.813	0.711	1	0.802	0.857	0.930
ICR	0.831	0.831	0.831	0.790	0.780	0.926	1	0.812	0.719
Katz	0.876	0.876	0.876	0.743	0.773	0.955	0.944	1	0.740
Pagerank	0.919	0.919	0.919	0.819	0.640	0.988	0.891	0.912	1
$\rho / \tau$	<b>Epinions</b>								
	BWLTR	FWLTR	LTR	Betweenness	Closeness	Degree	ICR	Katz	Pagerank
BWLTR	1	0.251	0.779	0.525	0.776	0.550	0.165	-	0.595
FWLTR	0.323	1	0.511	0.573	0.250	0.638	0.744	-	0.272
LTR	0.884	0.586	1	0.542	0.651	0.636	0.357	-	0.490
Betweenness	0.637	0.691	0.666	1	0.532	0.731	0.530	-	0.615
Closeness	0.916	0.350	0.810	0.653	1	0.555	0.207	-	0.619
Degree	0.659	0.755	0.749	0.834	0.679	1	0.584	-	0.632
ICR	0.240	0.883	0.472	0.661	0.310	0.712	1	-	0.211
Katz	-	-	-	-	-	-	-	-	-
Pagerank	0.766	0.389	0.647	0.733	0.798	0.767	0.306	-	1
$\rho / \tau$	<b>Higgs</b>								
	BWLTR	FWLTR	LTR	Betweenness	Closeness	Degree	ICR	Katz	Pagerank
BWLTR	1	-0.270	0.380	0.428	0.803	0.374	-0.214	0.805	0.788
FWLTR	-0.295	1	0.536	0.183	-0.390	0.499	0.660	-0.390	-0.391
LTR	0.402	0.544	1	0.384	0.381	0.923	0.324	0.383	0.370
Betweenness	0.445	0.195	0.405	1	0.519	0.407	0.139	0.518	0.508
Closeness	0.818	-0.423	0.405	0.535	1	0.379	-0.286	0.996	0.977
Degree	0.395	0.513	0.936	0.425	0.399	1	0.332	0.381	0.372
ICR	-0.264	0.764	0.382	0.168	-0.350	0.391	1	-0.286	-0.287
Katz	0.818	-0.423	0.406	0.535	1.000	0.401	-0.351	1	0.979
Pagerank	0.816	-0.428	0.397	0.528	0.998	0.395	-0.355	0.998	1

Tabla 5.6: Coeficientes de correlación para las medidas de centralidad aplicadas a las redes **Human Brain** y **Wikipedia**. Los coeficientes de **Kendall** ( $\tau$ ) se muestran en la región triangular superior, los coeficientes de **Spearman** ( $\rho$ ) se muestran en la triangular inferior. Los valores con un  $p$ -valor mayor que 0.05 se muestran en rojo. Para **Katz**, el símbolo '-' indica que el algoritmo no ha convergido.

$\rho / \tau$	<b>Human Brain</b>								
	BWLTR	FWLTR	LTR	Betweenness	Closeness	Degree	ICR	Katz	Pagerank
BWLTR	1	1	1	0.240	0.743	0.491	0.730	-	0.254
FWLTR	1	1	1	0.240	0.743	0.491	0.730	-	0.254
LTR	1	1	1	0.240	0.743	0.491	0.730	-	0.254
Betweenness	0.319	0.319	0.319	1	0.176	0.662	0.265	-	0.645
Closeness	0.897	0.897	0.897	0.229	1	0.400	0.744	-	0.137
Degree	0.619	0.619	0.619	0.813	0.522	1	0.536	-	0.712
ICR	0.898	0.898	0.898	0.359	0.916	0.682	1	-	0.259
Katz	-	-	-	-	-	-	-	-	-
Pagerank	0.355	0.355	0.355	0.825	0.201	0.857	0.372	-	1
$\rho / \tau$	<b>Wikipedia</b>								
	BWLTR	FWLTR	LTR	Betweenness	Closeness	Degree	ICR	Katz	Pagerank
BWLTR	1	-0.004	0.640	0.674	0.936	0.628	-0.016	-	0.937
FWLTR	0.012	1	0.503	0.414	-0.016	0.508	0.830	-	-0.013
LTR	0.729	0.531	1	0.531	0.613	0.983	0.390	-	0.615
Betweenness	0.735	0.482	0.615	1	0.659	0.530	0.372	-	0.676
Closeness	0.988	-0.006	0.720	0.725	1	0.604	-0.026	-	0.924
Degree	0.718	0.541	0.998	0.612	0.709	1	0.397	-	0.607
ICR	-0.004	0.932	0.469	0.457	-0.020	0.479	1	-	-0.023
Katz	-	-	-	-	-	-	-	-	-
Pagerank	0.988	-0.001	0.721	0.734	0.990	0.711	-0.015	-	1

### 5.2.3. Migraciones de actores

Para profundizar en el análisis comparativo de las medidas, aplicamos un estudio de desigualdad a los valores asignados por las medidas de centralidad a cada actor en la red. Este método es el descrito en la sección 4.1.9. En resumen, clasificamos a los actores en tres categorías diferentes, según su valor o posición en una medida de centralidad determinado. Luego observamos los movimientos de actores entre estas categorías, al pasar de una centralidad a otra.

En nuestro caso, utilizaremos dos criterios diferentes para clasificar los actores. El primer criterio es el mismo que se usa en [16], el cual clasifica los actores según la posición que obtengan, al ordenarlos de forma descendente según su valor asignado por una medida de centralidad. En el otro criterio, clasificamos los actores según el valor asignado y en que cuantil pertenezcan.

Representamos los resultados con un grafo dirigido con pesos. Cada grafo representa la migración de actores a categorías diferentes, al pasar de una medida de centralidad  $x$  a otra medida  $y$ . En estos grafos, los nodos representan las categorías y las aristas las migraciones. En las aristas se indica en porcentaje cuantos actores han migrado de una categoría a otra. Las aristas que 'bajan' representan a los actores que pasan de una categoría concreta en la medida  $x$  a otra categoría en la medida  $y$ . Similarmente, las aristas que 'suben' representan a los actores que pasan de una categoría en la medida  $y$  a otra categoría en la medida  $x$ . Para el segundo criterio de clasificación, añadimos el porcentaje de actores que hay en cada categoría en las dos medidas distintas, ya que a diferencia del primer criterio, los tamaños de las categorías varían según la medida. En la siguiente figura vemos un ejemplo de migración actores, al pasar de la medida de centralidad LTR a la medida de centralidad de Degree, en la red Caída, clasificando los actores según su valor (segundo criterio).

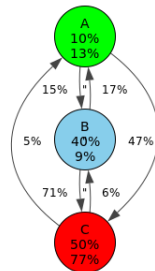


Figura 5.1: LTR - Degree

En la figura anterior tenemos que, las categorías A,B y C tienen un 10%, 40% y 50% de los actores respectivamente, con la medida de centralidad LTR. Para Degree, A, B y C tienen un 13%, 9% y 77% de los actores respectivamente. En el grafo podemos ver como al pasar de LTR a Degree, un 47% de los actores en la categoría A de LTR pasan a la categoría C de Degree, un 15% pasan del A de LTR al B de Degree, un 5% del C de Degree al A de LTR, etc.

Vamos a comentar solo los resultados más relevantes, el resto de resultados los incluimos en el apéndice B para su consulta.

Empezando por el primero criterio de clasificación, vemos que la mayoría de migraciones se producen entre nodos adyacentes y en porcentajes no muy altos. Las migraciones directas entre A y C con un porcentaje alto de actores solo ocurren en muy pocas redes y entre pocas medidas. En la red Dining Table, vemos varios ejemplos de estas migraciones.

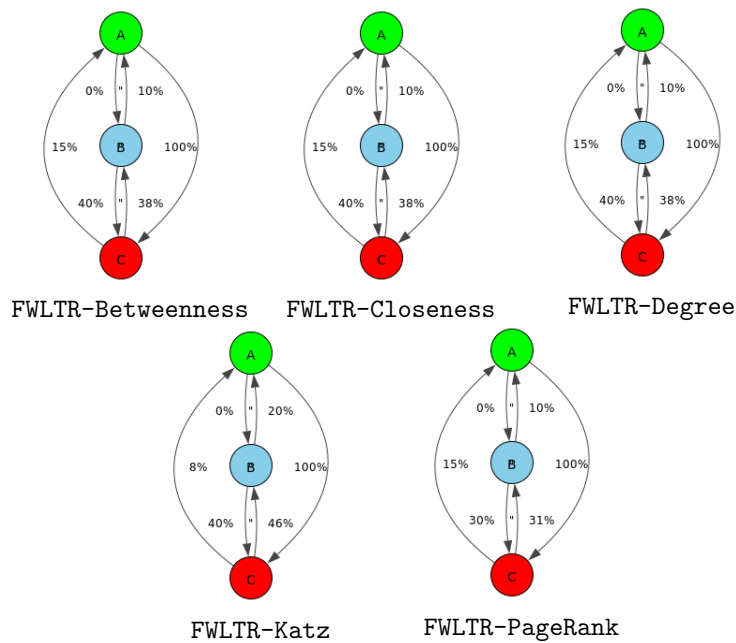


Figura 5.2: Migraciones de actores en Dining Table

Estas migraciones extremas, probablemente se den debido al reducido tamaño

de **Dining Table** (la red más pequeña de la colección), y que por lo tanto sea más susceptible a cambios que el resto de redes.

Justamente, la otra red en que observamos este comportamiento, **Higgs**, es una de las más grandes de la colección. En esta red solo observamos este comportamiento al pasar de **BWLTR** a **FWLTR**, y de **BWLTR** a **ICR**.

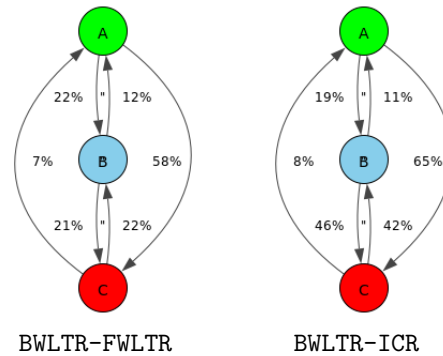


Figura 5.3: Migraciones de actores en **Higgs**

El hecho que el coeficiente de desigualdad **gini** sea más bajo que el de **BWLTR**, y que tanto **ICR** como **FWLTR** tengan poca correlación con **BWLTR**, podría explicar esta gran migración de actores de la categoría superior a la inferior. Además como más bajo es **gini** (el **gini** de **ICR** es muy cercano a 0) más actores parecen moverse.

Pasamos ahora a ver las migraciones entre categorías, con el criterio de clasificación por valor. Como veremos más abajo, en este criterio de clasificación, al ser más flexible en el tamaño de las categorías (en la clasificación por posición los tamaños de las categorías están fijados a 10%, 40% y 50% por definición), habrá más migraciones entre categorías extremas y con una mayor cantidad de actores.

La primera red en que vemos esta migración extrema entre categorías es **Caida**. En ella observamos este comportamiento entre las tres medidas basadas en el **LTM** al pasar a la medida de **Degree**, y a la de **Betweenness**.

Tanto en **BWLTR** como en **FWLTR** obtenemos las mismas migraciones de actores, independientemente de la medida con que se compare, ya que como hemos visto

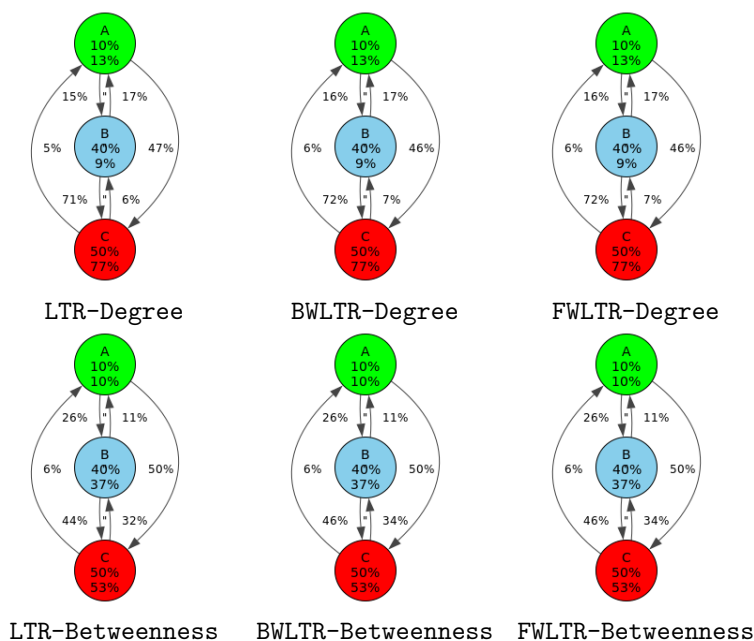


Figura 5.4: Migraciones de actores en Caida

anteriormente, en **Caida** las asignaciones de **BWLTR** y **BWLTR** coinciden perfectamente.

En las migraciones a **Degree**, vemos como hay un movimiento muy grande de actores hacia la categoría más baja **C**. Esto probablemente se deba a que, al clasificar los actores por valor, a diferencia del primer criterio de clasificación, los actores con un mismo valor van a la misma categoría. Entonces cuando una medida asigna pocos valores distintos, inevitablemente una categoría queda superpoblada. Este es el caso con **Degree**, tenemos que el 77% de los actores son clasificados en la categoría más baja. Entonces, si la medida de origen no le ocurre este mismo fenómeno, veremos una gran migración de actores a la última categoría al pasar de la medida de origen a **Degree**.

En el caso de **Betweenness** es otro, aquí las proporciones de las categorías se mantienen más o menos iguales entre medidas. Probablemente nos encontremos en un caso similar al de **Higgs**, visto anteriormente, con el primer criterio de clasificación. Igual que en ese caso, tenemos que **LTR**, **BWLTR** y **FWLTR** tienen muy poca correlación con **Betweenness**, pero a diferencia de **Higgs**, el coeficiente

gini de `Betweenness` es muy cercano a 1 y no a 0. Así pues, la migración de actores que estamos viendo, probablemente sean los que tienen más influencia en `Caida` bajo `Betweenness` intercambiando su sitio con el resto de actores subiendo a la categoría más alta.

En la red `Dining Table` también observamos estas migraciones extremas, pero igual que hemos visto antes con el criterio de clasificación por posición. Es muy probable que estas migraciones extremas se deban a su tamaño reducido y que por lo tanto sea más sensible a cambios de asignación.

Finalmente, la red de `Higgs` es la última red en mostrar estas migraciones extremas. En ella observamos estas migraciones al pasar de la medidas `LTR`, `BWLTR`, `FWLTR` a `Betweenness`, y de `FWLTR` a `Closeness`, `Katz`, `PageRank` y `BWLTR`. Al tener más flexibilidad en el tamaño de las categorías, vemos más movimiento de actores que en las migraciones de `Higgs` con la clasificación por posición.

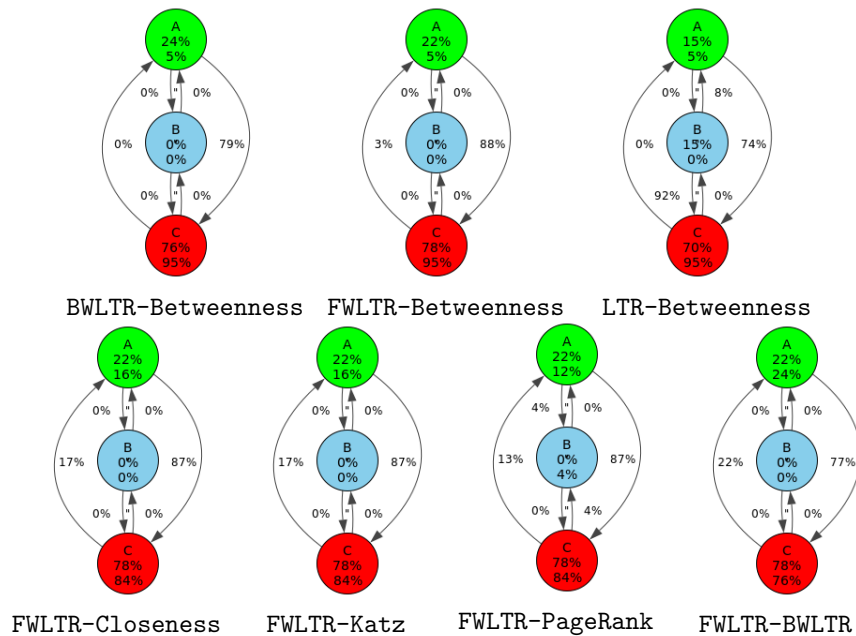


Figura 5.5: Migraciones de actores en `Higgs`

En el primer caso, donde pasamos de una medida a `Betweenness`, tenemos una combinación de los dos fenómenos que hemos visto anteriormente en `Caida`. Por una parte, las proporciones no son iguales, si bien las medidas reparten



los actores entre prácticamente solo dos categorías, en **Betweenness** la mayor parte de los actores entran en la última categoría. Por otra parte, esta extrema desigualdad se ve reflejada en el coeficiente de **gini** de **Betweenness** el cual es prácticamente 1. Esta diferencia de proporciones, y la desigualdad extrema en **Betweenness**, producen esta migración masiva de actores al pasar de una medida a **Betweenness**.

En el segundo caso, más o menos mantenemos las proporciones al cambiar de medida. El cambio en el coeficiente de **gini** al pasar de **FWLTR** (0.218 de **gini**) a otra medida, podría explicar la migración entre categorías al pasar a **Closeness** (0.993 de **gini**), y apurando también podría explicar la migración al pasar a **BWLTR** (0.578 de **gini**) o **PageRank** (0.590 de **gini**). Pero el coeficiente de **gini** no explica la migración de actores al pasar **Katz**, ya que la diferencia es muy poca. Si miramos las correlaciones entre **FWLTR** y **Katz** en **Higgs**, vemos que obtenemos una ligera correlación negativa en ambos coeficientes de correlación, aun así son demasiado bajos como para explicar esta migración.

Finalmente, para un trabajo futuro, sería interesante ver como estas migraciones ayudan a caracterizar una red. Ya que como hemos visto, al cambiar la red, pero utilizando las mismas medidas de centralidad, las migraciones varían considerablemente.

## 5.3. Resultados del FLTM

En esta sección vamos a ver los resultados obtenidos de las medias de centralidad basadas en el modelo FLTM definido previamente en la sección 4.1.5. Usaremos distintos valores de  $\gamma \in \{0.0, 0.1, \dots, 0.9, 1.0\}$ , y compararemos los resultados entre sí mediante los coeficientes de correlación introducidos en la sección 4.1.8. Finalmente compararemos los tiempos de ejecución obtenidos con las medidas basadas en el modelo FLTM para los mismos valores de  $\gamma$ , respecto a los tiempos de ejecución de las medidas basadas en el LTM.

### 5.3.1. Correlación de medidas

Primero de todo, veremos la correlación que hay entre las medidas de centralidad con diferentes valores de  $\gamma \in \{0.0, 0.1, \dots, 0.9, 1.0\}$  y ver como de similares son.

En las tablas 5.7-5.15, podemos ver los resultados obtenidos. Rápidamente vemos que, en todas las redes, las medidas de centralidad tienen una correlación casi perfecta positiva con todos los valores de  $\gamma$ . Además, cuando  $\gamma = 1$  tenemos que las medidas se comportan igual que las medidas basadas en el LTM, por definición del FLTM. Así pues, las medidas basadas en el FLTM dan una asignación casi idéntica a las medidas basada en LTM.

Inicialmente, pensamos que esta similitud se debía a que las redes que usamos en los experimentos no tienen un diámetro muy grande (como mucho 20), y que por lo tanto podría ser que no diese tiempo a que la influencia se devalué. Así pues, y aprovechando que, como veremos más adelante, los tiempos de ejecución para redes más grandes no son muy grandes, hemos buscado una nueva red que tenga un diámetro mucho más grande y un tamaño mayor a las redes utilizadas hasta ahora.

Esta red es la de **Texas**, la cual hemos definido anteriormente. Con un diámetro de 1054 creemos que es más que suficiente para que la influencia tenga tiempo suficiente en diluirse.

Tabla 5.7: Coeficientes de correlación para las medidas de centralidad BWFLTR con diferentes valores de  $\gamma$ , aplicadas a las redes ArXiv, Caida y Dining Table. Los coeficientes de Kendall ( $\tau$ ) se muestran en la región triangular superior, los coeficientes de Spearman ( $\rho$ ) se muestran en la triangular inferior. Los valores con un  $p$ -valor mayor que 0.05 se muestran en rojo.

<b>BWFLTR</b>	<b>ArXiv</b>										
$\rho / \tau$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1	1.000	1.000	1.000	1.000	1.000	0.999	0.998	0.997	0.963
0.1	1	1	1.000	0.998	0.998	0.996	0.996	0.995	0.982	0.978	0.963
0.2	1.000	1.000	1	0.998	0.998	0.996	0.996	0.995	0.982	0.978	0.963
0.3	0.998	1.000	1.000	1	1.000	0.998	0.998	0.997	0.983	0.980	0.965
0.4	0.998	1.000	1.000	1.000	1	0.998	0.998	0.997	0.983	0.980	0.965
0.5	0.996	1.000	1.000	1.000	1.000	1	1.000	0.999	0.985	0.982	0.966
0.6	0.996	1.000	1.000	1.000	1.000	1.000	1	0.999	0.985	0.982	0.966
0.7	0.995	0.999	0.999	1.000	1.000	1.000	1.000	1	0.986	0.983	0.967
0.8	0.982	0.998	0.998	0.998	0.998	0.998	0.998	0.998	1	0.996	0.980
0.9	0.978	0.997	0.997	0.997	0.997	0.997	0.997	0.998	1.000	1	0.983
1	0.994	0.994	0.994	0.994	0.994	0.995	0.995	0.995	0.997	0.997	1
<b>BWFLTR</b>	<b>Caida</b>										
$\rho / \tau$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1	1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.1	1	1	1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	1	1	1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	1	1	1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000
0.5	1.000	1.000	1.000	1.000	1.000	1	1	1.000	1.000	1.000	1.000
0.6	1.000	1.000	1.000	1.000	1.000	1	1	1.000	1.000	1.000	1.000
0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1	1.000	1.000
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1	1.000	1.000
0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
<b>BWFLTR</b>	<b>Dining Table</b>										
$\rho / \tau$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1	1	1	1	1	1	1	1	1	0.973
0.1	1	1	1	1	1	1	1	1	1	1	0.973
0.2	1	1	1	1	1	1	1	1	1	1	0.973
0.3	1	1	1	1	1	1	1	1	1	1	0.973
0.4	1	1	1	1	1	1	1	1	1	1	0.973
0.5	1	1	1	1	1	1	1	1	1	1	0.973
0.6	1	1	1	1	1	1	1	1	1	1	0.973
0.7	1	1	1	1	1	1	1	1	1	1	0.973
0.8	1	1	1	1	1	1	1	1	1	1	0.973
0.9	1	1	1	1	1	1	1	1	1	1	0.973
1	0.990	0.990	0.990	0.990	0.990	0.990	0.990	0.990	0.990	0.990	1

Tabla 5.8: Coeficientes de correlación para las medidas de centralidad BWFLTR con diferentes valores de  $\gamma$ , aplicadas a las redes Dolphins, Epinions y Higgs. Los coeficientes de Kendall ( $\tau$ ) se muestran en la región triangular superior, los coeficientes de Spearman ( $\rho$ ) se muestran en la triangular inferior. Los valores con un  $p$ -valor mayor que 0.05 se muestran en rojo.

BWFLTR $\rho / \tau$	Dolphins										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1	1	1	1	1	1	1	1	1	0.957
0.1	1	1	1	1	1	1	1	1	1	1	0.957
0.2	1	1	1	1	1	1	1	1	1	1	0.957
0.3	1	1	1	1	1	1	1	1	1	1	0.957
0.4	1	1	1	1	1	1	1	1	1	1	0.957
0.5	1	1	1	1	1	1	1	1	1	1	0.957
0.6	1	1	1	1	1	1	1	1	1	1	0.957
0.7	1	1	1	1	1	1	1	1	1	1	0.957
0.8	1	1	1	1	1	1	1	1	1	1	0.957
0.9	1	1	1	1	1	1	1	1	1	1	0.957
1	0.982	0.982	0.982	0.982	0.982	0.982	0.982	0.982	0.982	0.982	1
BWFLTR $\rho / \tau$	Epinions										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998
0.1	1	1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998
0.2	1	1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998
0.3	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	0.998
0.4	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	0.998
0.5	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	0.998
0.6	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	0.998
0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	0.998
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	0.998
0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	0.998
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
BWFLTR $\rho / \tau$	Higgs										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1	1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.1	1	1	1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	1	1	1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	1	1	1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000
0.5	1.000	1.000	1.000	1.000	1.000	1	1	1	1.000	1.000	1.000
0.6	1.000	1.000	1.000	1.000	1.000	1	1	1	1.000	1.000	1.000
0.7	1.000	1.000	1.000	1.000	1.000	1	1	1	1.000	1.000	1.000
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1	1.000
0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1	1.000
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1

Tabla 5.9: Coeficientes de correlación para las medidas de centralidad BWFLTR con diferentes valores de  $\gamma$ , aplicadas a las redes **Human Brain** y **Wikipedia**. Los coeficientes de Kendall ( $\tau$ ) se muestran en la región triangular superior, los coeficientes de Spearman ( $\rho$ ) se muestran en la triangular inferior. Los valores con un  $p$ -valor mayor que 0.05 se muestran en rojo.

<b>BWFLTR</b>	<b>Human Brain</b>										
$\rho / \tau$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.963
0.1	1	1	1.000	0.999	0.999	0.995	0.994	0.992	0.989	0.987	0.963
0.2	1.000	1.000	1	0.999	0.999	0.995	0.994	0.992	0.989	0.987	0.963
0.3	0.999	1.000	1.000	1	0.999	0.995	0.994	0.992	0.989	0.987	0.962
0.4	0.999	1.000	1.000	1.000	1	0.996	0.995	0.993	0.989	0.988	0.963
0.5	0.995	1.000	1.000	1.000	1.000	1	0.999	0.996	0.992	0.990	0.965
0.6	0.994	1.000	1.000	1.000	1.000	1.000	1	0.996	0.992	0.989	0.965
0.7	0.992	1.000	1.000	1.000	1.000	1.000	1.000	1	0.996	0.992	0.964
0.8	0.989	0.999	0.999	0.999	0.999	1.000	1.000	1.000	1	0.993	0.964
0.9	0.987	0.999	0.999	0.999	0.999	0.999	0.999	1.000	1.000	1	0.966
1	0.996	0.996	0.996	0.995	0.996	0.996	0.996	0.996	0.996	0.996	1
<b>BWFLTR</b>	<b>Wikipedia</b>										
$\rho / \tau$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1	1	1	1	1	1.000	1.000	1.000	1.000	0.999
0.1	1	1	1	1	1	1	1.000	1.000	1.000	1.000	0.999
0.2	1	1	1	1	1	1	1.000	1.000	1.000	1.000	0.999
0.3	1	1	1	1	1	1	1.000	1.000	1.000	1.000	0.999
0.4	1	1	1	1	1	1	1.000	1.000	1.000	1.000	0.999
0.5	1	1	1	1	1	1	1.000	1.000	1.000	1.000	0.999
0.6	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	0.999
0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	0.999
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	0.999
0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	0.999
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1

Tabla 5.10: Coeficientes de correlación para las medidas de centralidad FLTR con diferentes valores de  $\gamma$ , aplicadas a las redes ArXiv, Caida y Dining Table. Los coeficientes de Kendall ( $\tau$ ) se muestran en la región triangular superior, los coeficientes de Spearman ( $\rho$ ) se muestran en la triangular inferior. Los valores con un  $p$ -valor mayor que 0.05 se muestran en rojo.

<b>FLTR</b>	<b>ArXiv</b>										
$\rho / \tau$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1	1.000	1.000	1.000	1.000	1.000	0.999	0.998	0.997	0.963
0.1	1	1	1.000	0.998	0.998	0.996	0.996	0.995	0.982	0.978	0.963
0.2	1.000	1.000	1	0.998	0.998	0.996	0.996	0.995	0.982	0.978	0.963
0.3	0.998	1.000	1.000	1	1.000	0.998	0.998	0.997	0.983	0.980	0.965
0.4	0.998	1.000	1.000	1.000	1	0.998	0.998	0.997	0.983	0.980	0.965
0.5	0.996	1.000	1.000	1.000	1.000	1	1.000	0.999	0.985	0.982	0.966
0.6	0.996	1.000	1.000	1.000	1.000	1.000	1	0.999	0.985	0.982	0.966
0.7	0.995	0.999	0.999	1.000	1.000	1.000	1.000	1	0.986	0.983	0.967
0.8	0.982	0.998	0.998	0.998	0.998	0.998	0.998	0.998	1	0.996	0.980
0.9	0.978	0.997	0.997	0.997	0.997	0.997	0.997	0.998	1.000	1	0.983
1	0.994	0.994	0.994	0.994	0.994	0.995	0.995	0.995	0.997	0.997	1
<b>FLTR</b>	<b>Caida</b>										
$\rho / \tau$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.996
0.1	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.996
0.2	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.996
0.3	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	0.999	0.999	0.996
0.4	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	0.999	0.999	0.996
0.5	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	0.997
0.6	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	0.997
0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	0.997
0.8	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	0.997
0.9	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	0.997
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
<b>FLTR</b>	<b>Dining Table</b>										
$\rho / \tau$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1	1	1	1	1	1	1	1	1	0.936
0.1	1	1	1	1	1	1	1	1	1	1	0.936
0.2	1	1	1	1	1	1	1	1	1	1	0.936
0.3	1	1	1	1	1	1	1	1	1	1	0.936
0.4	1	1	1	1	1	1	1	1	1	1	0.936
0.5	1	1	1	1	1	1	1	1	1	1	0.936
0.6	1	1	1	1	1	1	1	1	1	1	0.936
0.7	1	1	1	1	1	1	1	1	1	1	0.936
0.8	1	1	1	1	1	1	1	1	1	1	0.936
0.9	1	1	1	1	1	1	1	1	1	1	0.936
1	0.963	0.963	0.963	0.963	0.963	0.963	0.963	0.963	0.963	0.963	1

Tabla 5.11: Coeficientes de correlación para las medidas de centralidad FLTR con diferentes valores de  $\gamma$ , aplicadas a las redes **Dolphins**, **Epinions** y **Higgs**. Los coeficientes de **Kendall** ( $\tau$ ) se muestran en la región triangular superior, los coeficientes de **Spearman** ( $\rho$ ) se muestran en la triangular inferior. Los valores con un  $p$ -valor mayor que 0.05 se muestran en rojo.

<b>FLTR</b>	<b>Dolphins</b>										
$\rho / \tau$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1	1	1	1	1	1	1	1	1	0.957
0.1	1	1	1	1	1	1	1	1	1	1	0.957
0.2	1	1	1	1	1	1	1	1	1	1	0.957
0.3	1	1	1	1	1	1	1	1	1	1	0.957
0.4	1	1	1	1	1	1	1	1	1	1	0.957
0.5	1	1	1	1	1	1	1	1	1	1	0.957
0.6	1	1	1	1	1	1	1	1	1	1	0.957
0.7	1	1	1	1	1	1	1	1	1	1	0.957
0.8	1	1	1	1	1	1	1	1	1	1	0.957
0.9	1	1	1	1	1	1	1	1	1	1	0.957
1	0.982	0.982	0.982	0.982	0.982	0.982	0.982	0.982	0.982	0.982	1
<b>FLTR</b>	<b>Epinions</b>										
$\rho / \tau$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.997
0.1	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.997
0.2	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.997
0.3	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	0.997
0.4	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	0.997
0.5	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	0.997
0.6	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	0.997
0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	0.997
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	0.997
0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	0.997
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
<b>FLTR</b>	<b>Higgs</b>										
$\rho / \tau$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1	1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.1	1	1	1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	1	1	1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	1	1	1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000
0.5	1.000	1.000	1.000	1.000	1.000	1	1	1	1.000	1.000	1.000
0.6	1.000	1.000	1.000	1.000	1.000	1	1	1	1.000	1.000	1.000
0.7	1.000	1.000	1.000	1.000	1.000	1	1	1	1.000	1.000	1.000
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1	1.000
0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1	1.000
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1





Tabla 5.13: Coeficientes de correlación para las medidas de centralidad FWFLTR con diferentes valores de  $\gamma$ , aplicadas a las redes ArXiv, Caida y Dining Table. Los coeficientes de Kendall ( $\tau$ ) se muestran en la región triangular superior, los coeficientes de Spearman ( $\rho$ ) se muestran en la triangular inferior. Los valores con un  $p$ -valor mayor que 0.05 se muestran en rojo.

<b>FWFLTR</b>	<b>ArXiv</b>										
$\rho / \tau$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1	1.000	1.000	1.000	1.000	1.000	0.999	0.998	0.997	0.963
0.1	1	1	1.000	0.998	0.998	0.996	0.996	0.995	0.982	0.978	0.963
0.2	1.000	1.000	1	0.998	0.998	0.996	0.996	0.995	0.982	0.978	0.963
0.3	0.998	1.000	1.000	1	1.000	0.998	0.998	0.997	0.983	0.980	0.965
0.4	0.998	1.000	1.000	1.000	1	0.998	0.998	0.997	0.983	0.980	0.965
0.5	0.996	1.000	1.000	1.000	1.000	1	1.000	0.999	0.985	0.982	0.966
0.6	0.996	1.000	1.000	1.000	1.000	1.000	1	0.999	0.985	0.982	0.966
0.7	0.995	0.999	0.999	1.000	1.000	1.000	1.000	1	0.986	0.983	0.967
0.8	0.982	0.998	0.998	0.998	0.998	0.998	0.998	0.998	1	0.996	0.980
0.9	0.978	0.997	0.997	0.997	0.997	0.997	0.997	0.997	1.000	1	0.983
1	0.994	0.994	0.994	0.994	0.994	0.995	0.995	0.995	0.997	0.997	1
<b>FWFLTR</b>	<b>Caida</b>										
$\rho / \tau$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1	1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.1	1	1	1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	1	1	1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	1	1	1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000
0.5	1.000	1.000	1.000	1.000	1.000	1	1	1.000	1.000	1.000	1.000
0.6	1.000	1.000	1.000	1.000	1.000	1	1	1.000	1.000	1.000	1.000
0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1	1.000	1.000
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1	1.000	1.000
0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
<b>FWFLTR</b>	<b>Dining Table</b>										
$\rho / \tau$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1	1	1	1	1	1	1	1	1	0.973
0.1	1	1	1	1	1	1	1	1	1	1	0.973
0.2	1	1	1	1	1	1	1	1	1	1	0.973
0.3	1	1	1	1	1	1	1	1	1	1	0.973
0.4	1	1	1	1	1	1	1	1	1	1	0.973
0.5	1	1	1	1	1	1	1	1	1	1	0.973
0.6	1	1	1	1	1	1	1	1	1	1	0.973
0.7	1	1	1	1	1	1	1	1	1	1	0.973
0.8	1	1	1	1	1	1	1	1	1	1	0.973
0.9	1	1	1	1	1	1	1	1	1	1	0.973
1	0.990	0.990	0.990	0.990	0.990	0.990	0.990	0.990	0.990	0.990	1

Tabla 5.14: Coeficientes de correlación para las medidas de centralidad FWFLTR con diferentes valores de  $\gamma$ , aplicadas a las redes Dolphins, Epinions y Higgs. Los coeficientes de Kendall ( $\tau$ ) se muestran en la región triangular superior, los coeficientes de Spearman ( $\rho$ ) se muestran en la triangular inferior. Los valores con un  $p$ -valor mayor que 0.05 se muestran en rojo.

FWFLTR $\rho / \tau$	Dolphins										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1	1	1	1	1	1	1	1	1	0.957
0.1	1	1	1	1	1	1	1	1	1	1	0.957
0.2	1	1	1	1	1	1	1	1	1	1	0.957
0.3	1	1	1	1	1	1	1	1	1	1	0.957
0.4	1	1	1	1	1	1	1	1	1	1	0.957
0.5	1	1	1	1	1	1	1	1	1	1	0.957
0.6	1	1	1	1	1	1	1	1	1	1	0.957
0.7	1	1	1	1	1	1	1	1	1	1	0.957
0.8	1	1	1	1	1	1	1	1	1	1	0.957
0.9	1	1	1	1	1	1	1	1	1	1	0.957
1	0.982	0.982	0.982	0.982	0.982	0.982	0.982	0.982	0.982	0.982	1
FWFLTR $\rho / \tau$	Epinions										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998
0.1	1	1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998
0.2	1	1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998
0.3	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	0.998
0.4	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	0.998
0.5	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	0.998
0.6	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	0.998
0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	0.998
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	0.998
0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	0.998
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
FWFLTR $\rho / \tau$	Higgs										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1	1	1	1	1	1	1	1	1	1.000
0.1	1	1	1	1	1	1	1	1	1	1	1.000
0.2	1	1	1	1	1	1	1	1	1	1	1.000
0.3	1	1	1	1	1	1	1	1	1	1	1.000
0.4	1	1	1	1	1	1	1	1	1	1	1.000
0.5	1	1	1	1	1	1	1	1	1	1	1.000
0.6	1	1	1	1	1	1	1	1	1	1	1.000
0.7	1	1	1	1	1	1	1	1	1	1	1.000
0.8	1	1	1	1	1	1	1	1	1	1	1.000
0.9	1	1	1	1	1	1	1	1	1	1	1.000
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1

Tabla 5.15: Coeficientes de correlación para las medidas de centralidad FWFLTR con diferentes valores de  $\gamma$ , aplicadas a las redes **Human Brain** y **Wikipedia**. Los coeficientes de Kendall ( $\tau$ ) se muestran en la región triangular superior, los coeficientes de Spearman ( $\rho$ ) se muestran en la triangular inferior. Los valores con un  $p$ -valor mayor que 0.05 se muestran en rojo.

<b>FWFLTR</b> $\rho / \tau$	<b>Human Brain</b>										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.963
0.1	1	1	1.000	0.999	0.999	0.995	0.994	0.992	0.989	0.987	0.963
0.2	1.000	1.000	1	0.999	0.999	0.995	0.994	0.992	0.989	0.987	0.963
0.3	0.999	1.000	1.000	1	0.999	0.995	0.994	0.992	0.989	0.987	0.962
0.4	0.999	1.000	1.000	1.000	1	0.996	0.995	0.993	0.989	0.988	0.963
0.5	0.995	1.000	1.000	1.000	1.000	1	0.999	0.996	0.992	0.990	0.965
0.6	0.994	1.000	1.000	1.000	1.000	1.000	1	0.996	0.992	0.989	0.965
0.7	0.992	1.000	1.000	1.000	1.000	1.000	1.000	1	0.996	0.992	0.964
0.8	0.989	0.999	0.999	0.999	0.999	1.000	1.000	1.000	1	0.993	0.964
0.9	0.987	0.999	0.999	0.999	0.999	0.999	0.999	1.000	1.000	1	0.966
1	0.996	0.996	0.996	0.995	0.996	0.996	0.996	0.996	0.996	0.996	1
<b>FWFLTR</b> $\rho / \tau$	<b>Wikipedia</b>										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1	1	1	1	1	1	1	1	1	1.000
0.1	1	1	1	1	1	1	1	1	1	1	1.000
0.2	1	1	1	1	1	1	1	1	1	1	1.000
0.3	1	1	1	1	1	1	1	1	1	1	1.000
0.4	1	1	1	1	1	1	1	1	1	1	1.000
0.5	1	1	1	1	1	1	1	1	1	1	1.000
0.6	1	1	1	1	1	1	1	1	1	1	1.000
0.7	1	1	1	1	1	1	1	1	1	1	1.000
0.8	1	1	1	1	1	1	1	1	1	1	1.000
0.9	1	1	1	1	1	1	1	1	1	1	1.000
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1

En la tabla 5.16 vemos los resultados obtenidos. Vemos como las medidas siguen comportándose igual, independientemente de que hayamos utilizado una red con un diámetro mayor. Así pues, parece ser que las medidas basadas en el modelo FLTM se comportan prácticamente igual que las basadas en el modelo LTM.

### 5.3.2. Tiempos de ejecución

Al ver que las medidas del FLTM se comportan igual que las del LTM, pensamos que sería interesante comparar los tiempos de ejecución de los dos tipos de medidas. Si bien tienen la misma complejidad temporal  $O(m+n)$ , en las medidas basadas en el FLTM, al devaluar la influencia en el tiempo, hay actores que no llegan a activarse, con lo cual los tiempos de ejecución podrían ser algo inferiores a los de las medidas del LTM, y al mismo tiempo darnos resultados prácticamente idénticos.

Para medir los tiempos de ejecución, ejecutamos cada medida 5 veces, y obtenemos en cada ejecución el *elapsed time* de la medida. Estos tiempos han sido obtenidos con la librería estándar de C++, **chrono**, la cual nos permite tomar una captura del tiempo actual de sistema durante la ejecución del programa. Con dos capturas, una antes de ejecutar la medida, y otra justo después de terminar, podremos determinar el tiempo de ejecución de la medida.

En cada red mostramos los resultados con dos gráficas para cada medida de centralidad. En la primera gráfica de la izquierda, mostramos la media de los tiempos de ejecución en segundos de las dos versiones de las medidas, con diferentes valores de  $\gamma$  para las medidas FLTM, las barras verticales representan la desviación estándar de las cinco 5 ejecuciones. En la segunda gráfica de la derecha, vemos el *speedup* de las medidas FLTM respecto a las medidas LTM. Comentaremos los resultados más relevantes, en el apéndice C.1 incluimos el resto de resultados para su consulta.

En las figuras 5.6 - 5.14, vemos las gráficas de los tiempos de ejecución para redes **Dining Table**, **Caida** y **Texas**.

En las redes pequeñas como **Dining Table**, los tiempos de ejecución varían sin importar el valor de  $\gamma$ , la mayoría de medidas del FLTM obtienen unos tiempos de ejecución ligeramente más altos que los de LTM. En **Caida**, vemos que conforme aumenta el tamaño de la red los tiempos de ejecución se estabilizan, independientemente del valor de  $\gamma$ . En esta red los tiempos de ejecución de FLTM son ligeramente más lentos que los de LTM. Finalmente, en **Texas**, vemos que como más grande sea la red más similares son los tiempos de ejecución, con *speedups* de  $\pm 0.1\%$ . Así pues, parece ser que los tiempos de ejecución de FLTM también son prácticamente iguales a los de LTM.

Cabe destacar, que al programar las medidas en `C++`, los tiempos de ejecución han disminuido drásticamente respecto a los vistos en [15], usando una máquina similar. Al usar `C++` en vez de `Python` hemos pasado a ejecutar `Higgs` con tiempos del orden de horas, a ejecutarlo con tiempos del orden de varios segundos (véase el apéndice C.1.6). Al paralelizar la obtención de las medidas, con 32 threads, obtenemos speedups de entre 2.5x a 13x respecto a las versiones secuenciales, según la red.

Tabla 5.16: Coeficientes de correlación para las medidas de centralidad basadas en el modelo FLTM, con diferentes valores de  $\gamma$ , aplicadas a la red de Texas. Los coeficientes de Kendall ( $\tau$ ) se muestran en la región triangular superior, los coeficientes de Spearman ( $\rho$ ) se muestran en la triangular inferior. Los valores con un  $p$ -valor mayor que 0.05 se muestran en rojo.

<b>BWFLTR</b>	<b>Texas</b>										
$\rho / \tau$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1	1	0.999	0.999	0.999	0.999	0.999	0.983	0.983	0.939
0.1	1	1	1	0.998	0.998	0.998	0.998	0.998	0.949	0.949	0.939
0.2	1	1	1	0.998	0.998	0.998	0.998	0.998	0.949	0.949	0.939
0.3	0.998	0.999	0.999	1	1	1.000	1.000	1.000	0.951	0.951	0.940
0.4	0.998	0.999	0.999	1	1	1.000	1.000	1.000	0.951	0.951	0.940
0.5	0.998	0.999	0.999	1.000	1.000	1	1	1.000	0.951	0.951	0.940
0.6	0.998	0.999	0.999	1.000	1.000	1	1	1.000	0.951	0.951	0.940
0.7	0.998	0.999	0.999	1.000	1.000	1.000	1.000	1	0.951	0.951	0.940
0.8	0.949	0.983	0.983	0.984	0.984	0.984	0.984	0.984	1	1.000	0.990
0.9	0.949	0.983	0.983	0.984	0.984	0.984	0.984	0.984	1.000	1	0.990
1	0.978	0.978	0.978	0.979	0.979	0.979	0.979	0.979	0.997	0.997	1
<b>FLTR</b>	<b>Texas</b>										
$\rho / \tau$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1	1	0.999	0.999	0.999	0.999	0.999	0.983	0.983	0.939
0.1	1	1	1	0.998	0.998	0.998	0.998	0.998	0.949	0.949	0.939
0.2	1	1	1	0.998	0.998	0.998	0.998	0.998	0.949	0.949	0.939
0.3	0.998	0.999	0.999	1	1	1.000	1.000	1.000	0.951	0.951	0.940
0.4	0.998	0.999	0.999	1	1	1.000	1.000	1.000	0.951	0.951	0.940
0.5	0.998	0.999	0.999	1.000	1.000	1	1	1.000	0.951	0.951	0.940
0.6	0.998	0.999	0.999	1.000	1.000	1	1	1.000	0.951	0.951	0.940
0.7	0.998	0.999	0.999	1.000	1.000	1.000	1.000	1	0.951	0.951	0.940
0.8	0.949	0.983	0.983	0.984	0.984	0.984	0.984	0.984	1	1.000	0.990
0.9	0.949	0.983	0.983	0.984	0.984	0.984	0.984	0.984	1.000	1	0.990
1	0.978	0.978	0.978	0.979	0.979	0.979	0.979	0.979	0.997	0.997	1
<b>FWFLTR</b>	<b>Texas</b>										
$\rho / \tau$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	1	1	1	0.999	0.999	0.999	0.999	0.999	0.983	0.983	0.939
0.1	1	1	1	0.998	0.998	0.998	0.998	0.998	0.949	0.949	0.939
0.2	1	1	1	0.998	0.998	0.998	0.998	0.998	0.949	0.949	0.939
0.3	0.998	0.999	0.999	1	1	1.000	1.000	1.000	0.951	0.951	0.940
0.4	0.998	0.999	0.999	1	1	1.000	1.000	1.000	0.951	0.951	0.940
0.5	0.998	0.999	0.999	1.000	1.000	1	1	1.000	0.951	0.951	0.940
0.6	0.998	0.999	0.999	1.000	1.000	1	1	1.000	0.951	0.951	0.940
0.7	0.998	0.999	0.999	1.000	1.000	1.000	1.000	1	0.951	0.951	0.940
0.8	0.949	0.983	0.983	0.984	0.984	0.984	0.984	0.984	1	1.000	0.990
0.9	0.949	0.983	0.983	0.984	0.984	0.984	0.984	0.984	1.000	1	0.990
1	0.978	0.978	0.978	0.979	0.979	0.979	0.979	0.979	0.997	0.997	1

# Dining\_Table

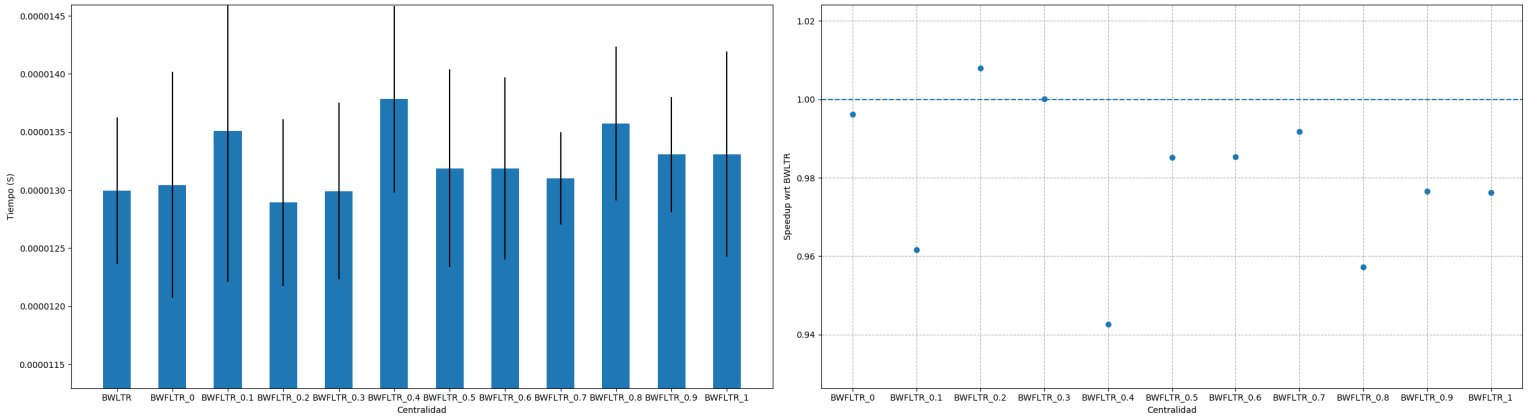


Figura 5.6: BWLTR vs BWFLTR

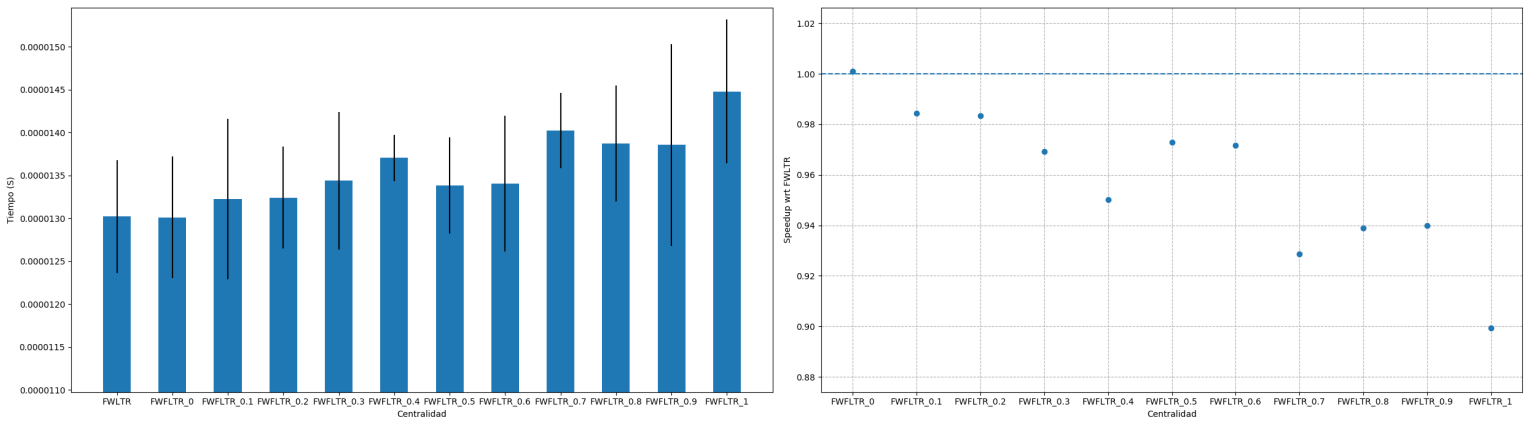


Figura 5.7: FWLTR vs FWFLTR

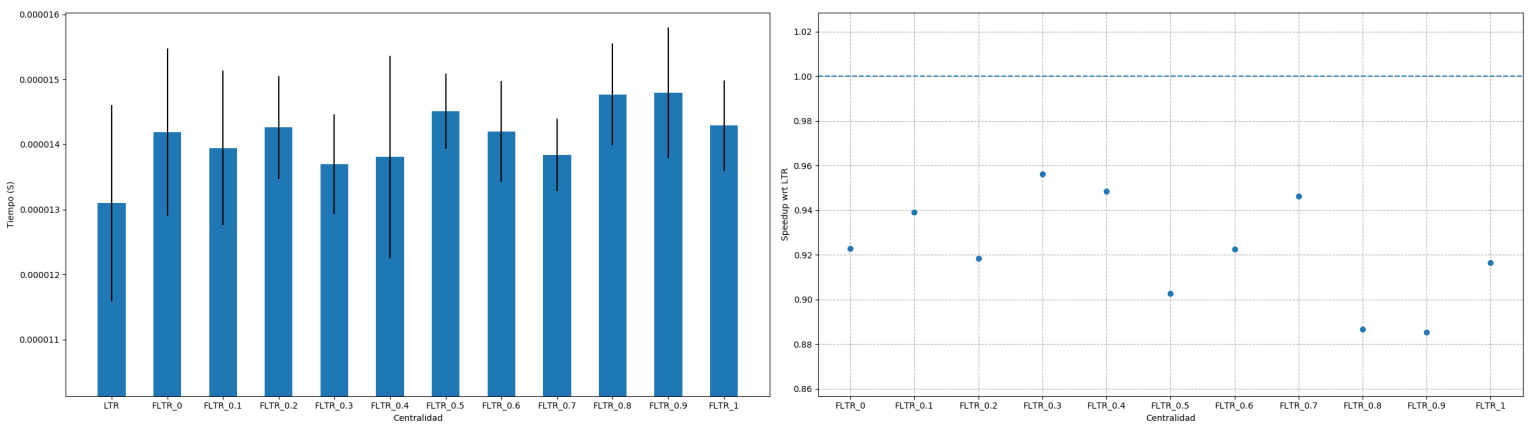


Figura 5.8: LTR vs FLTR

# Caida

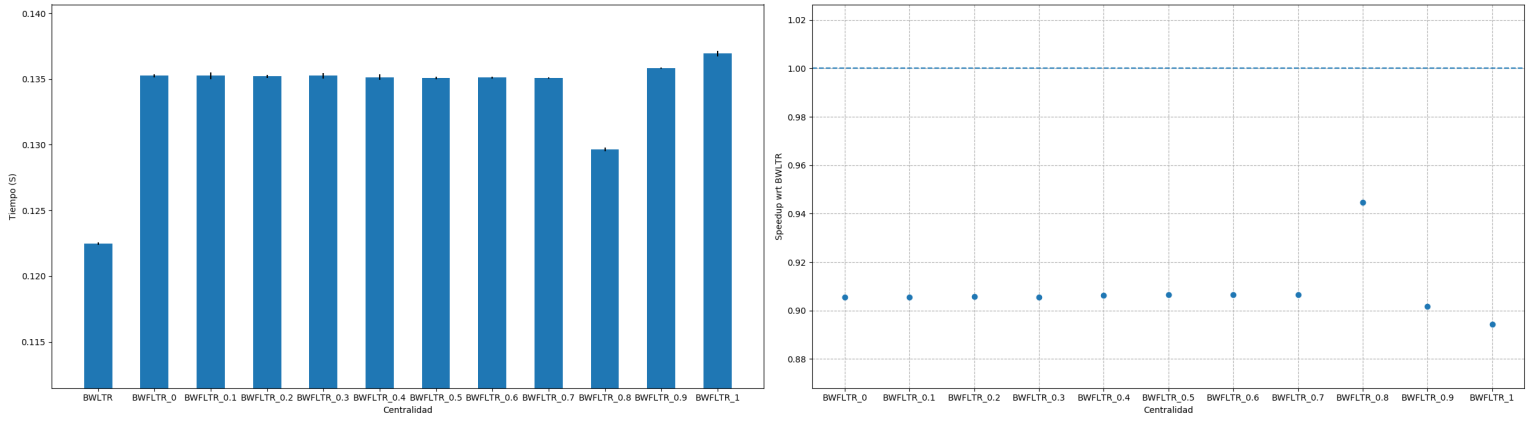


Figura 5.9: BWLTR vs BWFLTR

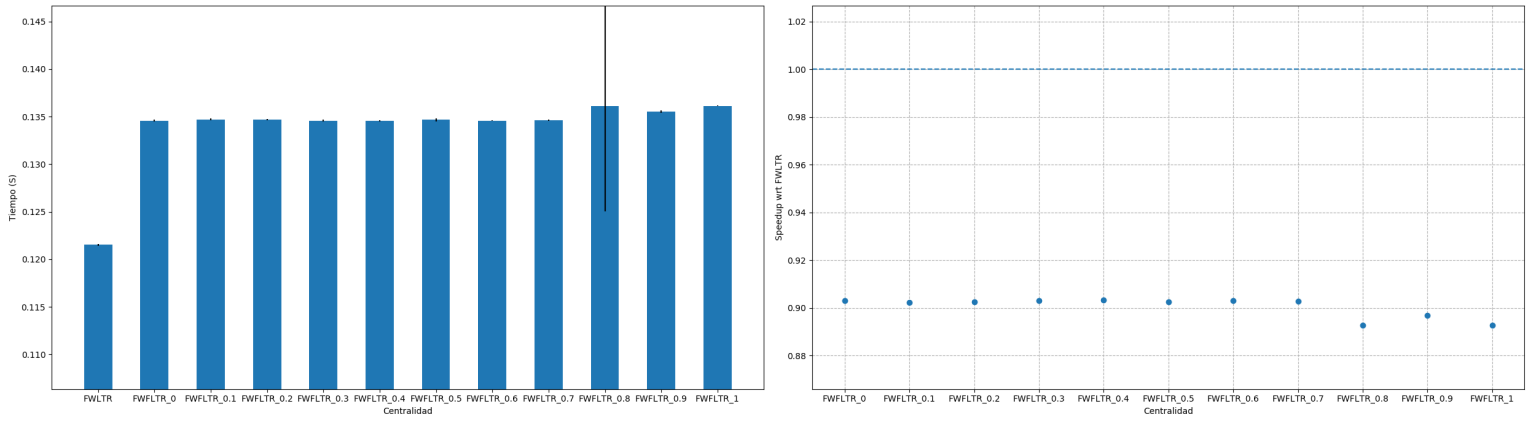


Figura 5.10: FWLTR vs FWFLTR

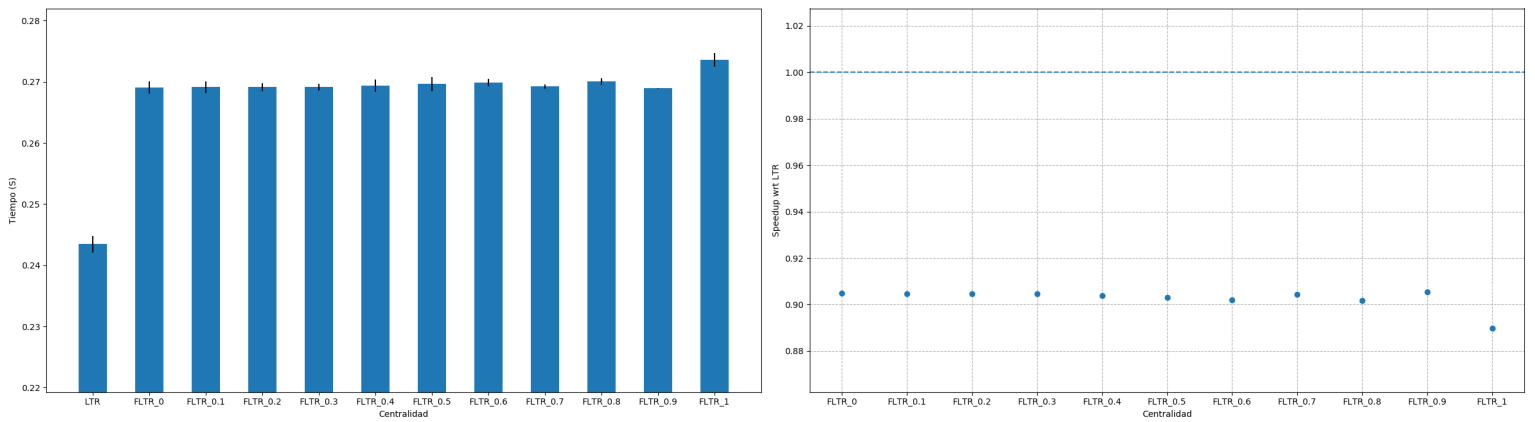


Figura 5.11: LTR vs FLTR



# Texas

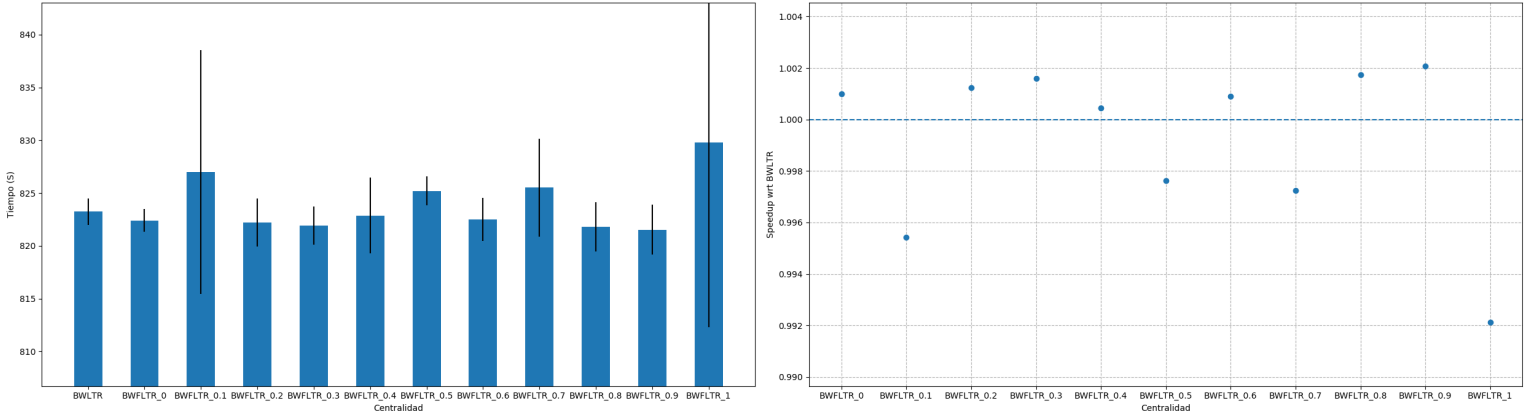


Figura 5.12: BWLTR vs BWFLTR

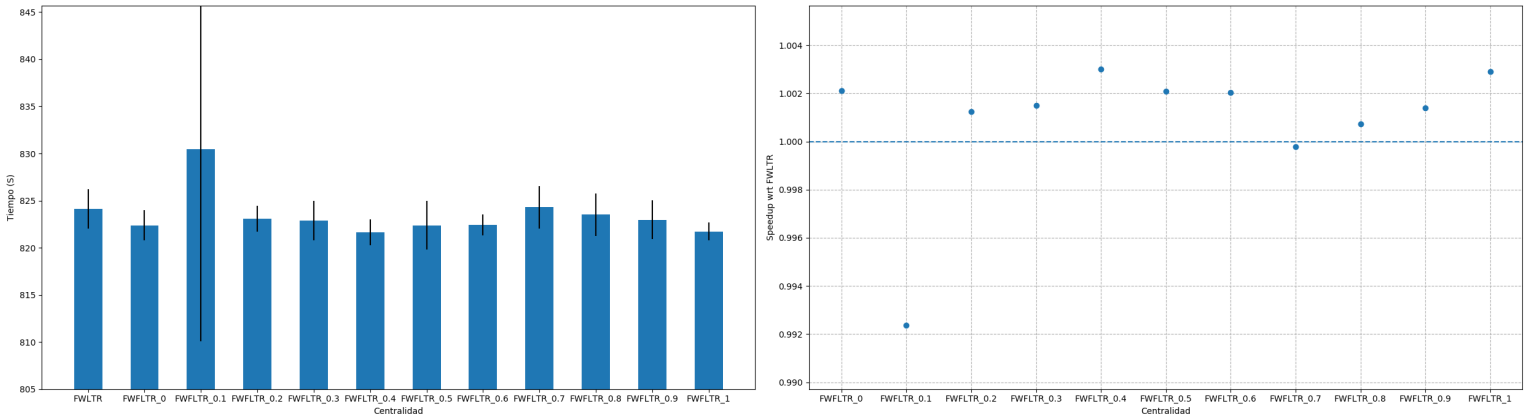


Figura 5.13: FWLTR vs FWFLTR

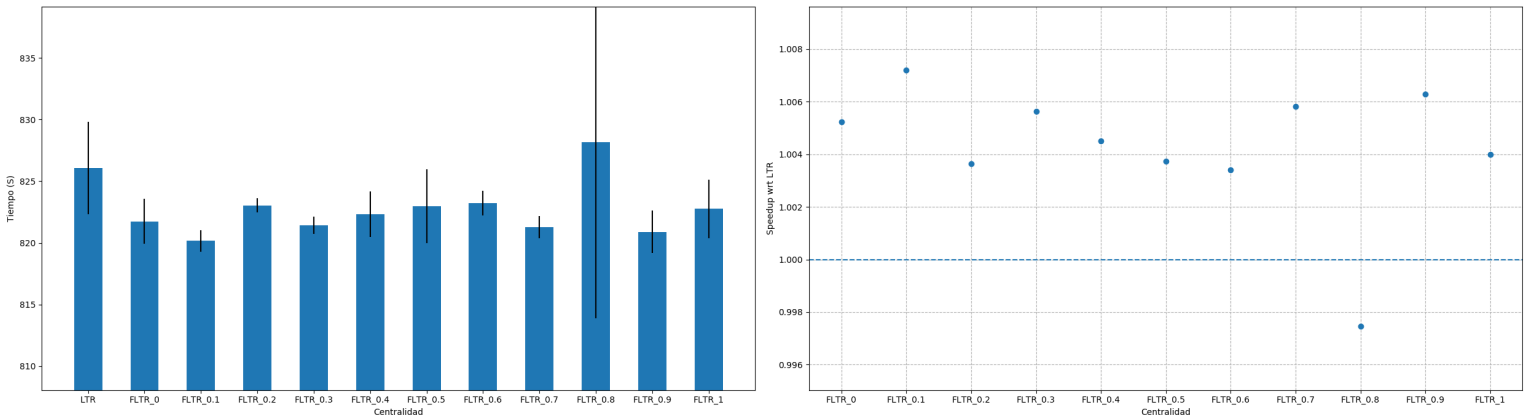


Figura 5.14: LTR vs FLTR

## 5.4. Resultados del BLTM

En esta sección, igual que con FLTM, vamos a ver los resultados obtenidos de las medias de centralidad basadas en el modelo BLTM definido previamente en la sección 4.1.6. Usaremos distintos valores de  $p \in \{1, 2, \dots, 10\}$ , y compararemos los resultados entre si mediante los coeficientes de correlación introducidos en la sección 4.1.8. Finalmente compararemos los tiempos de ejecución obtenidos con las medidas basadas en el modelo BLTM para los mismos valores de  $p$ , respecto a los tiempos de ejecución de las medidas basadas en el LTM.

### 5.4.1. Correlación de medidas

Empezamos viendo las similitudes que hay entre las medidas basadas en el BLTM, con distintos valores de  $p \in \{1, 2, \dots, 10\}$ , y las medidas basadas en el LTM. También utilizaremos la red **Texas**, definida en la sección 5.3.1.

En las tablas 5.17-5.25, podemos ver los resultados obtenidos. Observamos que se repite el mismo fenómeno que en FLTM, las medidas basadas en BLTM se comportan prácticamente igual que las basadas en LTM, independientemente del valor que tenga el parámetro  $p$ . Si bien cuando el parámetro  $p$  es igual o más grande que el diámetro de la red, los valores que asigna la medida coinciden completamente con los de LTM, vemos que en nuestro caso, esta similitud se produce mucho antes que al llegar al diámetro de la red.

Vemos pues, que los actores en nuestras redes solo son capaces de expandir su influencia hasta niveles bajos de expansión. Esto explicaría, los resultados vistos anteriormente en FLTM, como los actores solo son capaces de expandir su influencia unos pocos niveles de expansión, no hay suficiente tiempo a que su influencia se devalué, y por lo tanto obtenemos los mismos resultados que LTM.

Tabla 5.17: Coeficientes de correlación para las medidas de centralidad BBWLTR con diferentes valores de  $p$ , aplicadas a las redes ArXiv, Caida y Dining Table. Los coeficientes de Kendall ( $\tau$ ) se muestran en la región triangular superior, los coeficientes de Spearman ( $\rho$ ) se muestran en la triangular inferior. Los valores con un  $p$ -valor mayor que 0.05 se muestran en rojo.

<b>BBWLTR</b>	<b>ArXiv</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	BWLTR
1	1	0.992	0.988	0.987	0.986	0.986	0.986	0.986	0.986	0.986	0.986
2	0.999	1	0.997	0.996	0.996	0.996	0.995	0.995	0.995	1.000	0.995
3	0.999	1.000	1	0.999	0.999	0.999	0.999	0.999	0.999	1.000	0.999
4	0.998	1.000	1.000	1	1.000	1.000	1.000	1.000	0.999	1.000	0.999
5	0.998	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000
6	0.998	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000
7	0.998	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000
8	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000
9	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1	1
10	0.998	0.995	0.999	0.999	1.000	1.000	1.000	1.000	1	1	1
BWLTR	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1	1
<b>BBWLTR</b>	<b>Caida</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	BWLTR
1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995
2	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995
3	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995
4	1.000	1.000	1.000	1	1	1	1	1	1	1	0.995
5	1.000	1.000	1.000	1	1	1	1	1	1	1	0.995
6	1.000	1.000	1.000	1	1	1	1	1	1	1	0.995
7	1.000	1.000	1.000	1	1	1	1	1	1	1	0.995
8	1.000	1.000	1.000	1	1	1	1	1	1	1	0.995
9	1.000	1.000	1.000	1	1	1	1	1	1	1	0.995
10	1.000	1.000	1.000	1	1	1	1	1	1	1	0.995
BWLTR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
<b>BBWLTR</b>	<b>Dining Table</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	BWLTR
1	1	0.975	0.969	0.969	0.969	0.969	0.969	0.969	0.969	0.969	0.969
2	0.990	1	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.999	0.995
3	0.988	0.999	1	1	1	1	1	1	1	1	1
4	0.988	0.999	1	1	1	1	1	1	1	1	1
5	0.988	0.999	1	1	1	1	1	1	1	1	1
6	0.988	0.999	1	1	1	1	1	1	1	1	1
7	0.988	0.999	1	1	1	1	1	1	1	1	1
8	0.988	0.999	1	1	1	1	1	1	1	1	1
9	0.988	0.999	1	1	1	1	1	1	1	1	1
10	0.988	0.995	1	1	1	1	1	1	1	1	1
BWLTR	0.988	0.999	1	1	1	1	1	1	1	1	1

Tabla 5.18: Coeficientes de correlación para las medidas de centralidad BBWLTR con diferentes valores de  $p$ , aplicadas a las redes Dolphins, Epinions y Higgs. Los coeficientes de Kendall ( $\tau$ ) se muestran en la región triangular superior, los coeficientes de Spearman ( $\rho$ ) se muestran en la triangular inferior. Los valores con un  $p$ -valor mayor que 0.05 se muestran en rojo.

<b>BBWLTR</b>	<b>Dolphins</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	BWLTR
1	1	0.982	0.980	0.980	0.980	0.980	0.980	0.980	0.980	0.980	0.980
2	0.994	1	0.999	0.999	0.999	0.999	0.999	0.999	0.999	1.000	0.999
3	0.993	1.000	1	1	1	1	1	1	1	1	1
4	0.993	1.000	1	1	1	1	1	1	1	1	1
5	0.993	1.000	1	1	1	1	1	1	1	1	1
6	0.993	1.000	1	1	1	1	1	1	1	1	1
7	0.993	1.000	1	1	1	1	1	1	1	1	1
8	0.993	1.000	1	1	1	1	1	1	1	1	1
9	0.993	1.000	1	1	1	1	1	1	1	1	1
10	0.993	0.999	1	1	1	1	1	1	1	1	1
BWLTR	0.993	1.000	1	1	1	1	1	1	1	1	1
<b>BBWLTR</b>	<b>Epinions</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	BWLTR
1	1	1.000	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
2	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000
6	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000
7	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000
8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000
9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000
10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000
BWLTR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
<b>BBWLTR</b>	<b>Higgs</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	BWLTR
1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	1.000	1.000	1.000	1	1	1	1	1	1	1	1
5	1.000	1.000	1.000	1	1	1	1	1	1	1	1
6	1.000	1.000	1.000	1	1	1	1	1	1	1	1
7	1.000	1.000	1.000	1	1	1	1	1	1	1	1
8	1.000	1.000	1.000	1	1	1	1	1	1	1	1
9	1.000	1.000	1.000	1	1	1	1	1	1	1	1
10	1.000	1.000	1.000	1	1	1	1	1	1	1	1
BWLTR	1.000	1.000	1.000	1	1	1	1	1	1	1	1

Tabla 5.19: Coeficientes de correlación para las medidas de centralidad BBWLTR con diferentes valores de  $p$ , aplicadas a las redes **Human Brain**, **Texas** y **Wikipedia**. Los coeficientes de Kendall ( $\tau$ ) se muestran en la región triangular superior, los coeficientes de Spearman ( $\rho$ ) se muestran en la triangular inferior. Los valores con un  $p$ -valor mayor que 0.05 se muestran en rojo.

<b>BBWLTR</b>	<b>Human Brain</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	BWLTR
1	1	0.992	0.986	0.976	0.973	0.972	0.972	0.972	0.972	0.972	0.972
2	1.000	1	0.992	0.982	0.979	0.977	0.977	0.977	0.977	0.998	0.977
3	0.999	1.000	1	0.990	0.986	0.985	0.985	0.984	0.984	0.999	0.984
4	0.998	0.998	0.999	1	0.996	0.994	0.994	0.994	0.994	1.000	0.994
5	0.997	0.998	0.999	1.000	1	0.999	0.999	0.998	0.998	1.000	0.998
6	0.997	0.998	0.999	1.000	1.000	1	1.000	0.999	0.999	1.000	0.999
7	0.997	0.998	0.999	1.000	1.000	1.000	1	0.999	0.999	1.000	0.999
8	0.997	0.998	0.999	1.000	1.000	1.000	1.000	1	1	1	1
9	0.997	0.998	0.999	1.000	1.000	1.000	1.000	1	1	1	1
10	0.997	0.977	0.984	0.994	0.998	0.999	0.999	1	1	1	1
BWLTR	0.997	0.998	0.999	1.000	1.000	1.000	1.000	1	1	1	1
<b>BBWLTR</b>	<b>Texas</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	BWLTR
1	1	0.985	0.983	0.982	0.982	0.982	0.982	0.982	0.982	0.982	0.982
2	0.996	1	0.999	0.998	0.998	0.998	0.998	0.998	0.998	1.000	0.998
3	0.995	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	0.995	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0.995	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000
6	0.995	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000
7	0.995	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000
8	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000
9	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000
10	0.995	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000
BWLTR	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
<b>BBWLTR</b>	<b>Wikipedia</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	BWLTR
1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000
6	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000
7	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000
8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1	1.000	1.000
9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000
10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000
BWLTR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1

Tabla 5.20: Coeficientes de correlación para las medidas de centralidad BFWLTR con diferentes valores de  $p$ , aplicadas a las redes ArXiv, Caida y Dining Table. Los coeficientes de Kendall ( $\tau$ ) se muestran en la región triangular superior, los coeficientes de Spearman ( $\rho$ ) se muestran en la triangular inferior. Los valores con un  $p$ -valor mayor que 0.05 se muestran en rojo.

<b>BFWLTR</b>	<b>ArXiv</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	FWLTR
1	1	0.992	0.988	0.987	0.986	0.986	0.986	0.986	0.986	0.986	0.986
2	0.999	1	0.997	0.996	0.996	0.996	0.995	0.995	0.995	1.000	0.995
3	0.999	1.000	1	0.999	0.999	0.999	0.999	0.999	0.999	1.000	0.999
4	0.998	1.000	1.000	1	1.000	1.000	1.000	1.000	0.999	1.000	0.999
5	0.998	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000
6	0.998	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000
7	0.998	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000
8	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000
9	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1	1
10	0.998	0.995	0.999	0.999	1.000	1.000	1.000	1.000	1	1	1
FWLTR	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1	1
<b>BFWLTR</b>	<b>Caida</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	FWLTR
1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995
2	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995
3	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995
4	1.000	1.000	1.000	1	1	1	1	1	1	1	0.995
5	1.000	1.000	1.000	1	1	1	1	1	1	1	0.995
6	1.000	1.000	1.000	1	1	1	1	1	1	1	0.995
7	1.000	1.000	1.000	1	1	1	1	1	1	1	0.995
8	1.000	1.000	1.000	1	1	1	1	1	1	1	0.995
9	1.000	1.000	1.000	1	1	1	1	1	1	1	0.995
10	1.000	1.000	1.000	1	1	1	1	1	1	1	0.995
FWLTR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
<b>BFWLTR</b>	<b>Dining Table</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	FWLTR
1	1	0.993	0.993	0.993	0.993	0.993	0.993	0.993	0.993	0.993	0.993
2	0.999	1	1	1	1	1	1	1	1	1	1
3	0.999	1	1	1	1	1	1	1	1	1	1
4	0.999	1	1	1	1	1	1	1	1	1	1
5	0.999	1	1	1	1	1	1	1	1	1	1
6	0.999	1	1	1	1	1	1	1	1	1	1
7	0.999	1	1	1	1	1	1	1	1	1	1
8	0.999	1	1	1	1	1	1	1	1	1	1
9	0.999	1	1	1	1	1	1	1	1	1	1
10	0.999	1	1	1	1	1	1	1	1	1	1
FWLTR	0.999	1	1	1	1	1	1	1	1	1	1

Tabla 5.21: Coeficientes de correlación para las medidas de centralidad BFWLTR con diferentes valores de  $p$ , aplicadas a las redes **Dolphins**, **Epinions** y **Higgs**. Los coeficientes de **Kendall** ( $\tau$ ) se muestran en la región triangular superior, los coeficientes de **Spearman** ( $\rho$ ) se muestran en la triangular inferior. Los valores con un  $p$ -valor mayor que 0.05 se muestran en rojo.

<b>BFWLTR</b>	<b>Dolphins</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	FWLTR
1	1	0.982	0.980	0.980	0.980	0.980	0.980	0.980	0.980	0.980	0.980
2	0.994	1	0.999	0.999	0.999	0.999	0.999	0.999	0.999	1.000	0.999
3	0.993	1.000	1	1	1	1	1	1	1	1	1
4	0.993	1.000	1	1	1	1	1	1	1	1	1
5	0.993	1.000	1	1	1	1	1	1	1	1	1
6	0.993	1.000	1	1	1	1	1	1	1	1	1
7	0.993	1.000	1	1	1	1	1	1	1	1	1
8	0.993	1.000	1	1	1	1	1	1	1	1	1
9	0.993	1.000	1	1	1	1	1	1	1	1	1
10	0.993	0.999	1	1	1	1	1	1	1	1	1
FWLTR	0.993	1.000	1	1	1	1	1	1	1	1	1
<b>BFWLTR</b>	<b>Epinions</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	FWLTR
1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000
6	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000
7	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000
8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000
9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000
10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000
FWLTR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
<b>BFWLTR</b>	<b>Higgs</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	FWLTR
1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3	1.000	1.000	1	1	1	1	1	1	1	1	1
4	1.000	1.000	1	1	1	1	1	1	1	1	1
5	1.000	1.000	1	1	1	1	1	1	1	1	1
6	1.000	1.000	1	1	1	1	1	1	1	1	1
7	1.000	1.000	1	1	1	1	1	1	1	1	1
8	1.000	1.000	1	1	1	1	1	1	1	1	1
9	1.000	1.000	1	1	1	1	1	1	1	1	1
10	1.000	1.000	1	1	1	1	1	1	1	1	1
FWLTR	1.000	1.000	1	1	1	1	1	1	1	1	1

Tabla 5.22: Coeficientes de correlación para las medidas de centralidad BFWLTR con diferentes valores de  $p$ , aplicadas a las redes Human Brain, Texas y Wikipedia. Los coeficientes de Kendall ( $\tau$ ) se muestran en la región triangular superior, los coeficientes de Spearman ( $\rho$ ) se muestran en la triangular inferior. Los valores con un  $p$ -valor mayor que 0.05 se muestran en rojo.

<b>BFWLTR</b>	<b>Human Brain</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	FWLTR
1	1	0.992	0.986	0.976	0.973	0.972	0.972	0.972	0.972	0.972	0.972
2	1.000	1	0.992	0.982	0.979	0.977	0.977	0.977	0.977	0.998	0.977
3	0.999	1.000	1	0.990	0.986	0.985	0.985	0.984	0.984	0.999	0.984
4	0.998	0.998	0.999	1	0.996	0.994	0.994	0.994	0.994	1.000	0.994
5	0.997	0.998	0.999	1.000	1	0.999	0.999	0.998	0.998	1.000	0.998
6	0.997	0.998	0.999	1.000	1.000	1	1.000	0.999	0.999	1.000	0.999
7	0.997	0.998	0.999	1.000	1.000	1.000	1	0.999	0.999	1.000	0.999
8	0.997	0.998	0.999	1.000	1.000	1.000	1.000	1	1	1	1
9	0.997	0.998	0.999	1.000	1.000	1.000	1.000	1	1	1	1
10	0.997	0.977	0.984	0.994	0.998	0.999	0.999	1	1	1	1
FWLTR	0.997	0.998	0.999	1.000	1.000	1.000	1.000	1	1	1	1
<b>BFWLTR</b>	<b>Texas</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	FWLTR
1	1	0.985	0.983	0.982	0.982	0.982	0.982	0.982	0.982	0.982	0.982
2	0.996	1	0.999	0.998	0.998	0.998	0.998	0.998	0.998	1.000	0.998
3	0.995	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	0.995	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0.995	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000
6	0.995	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000
7	0.995	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000
8	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000
9	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000
10	0.995	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000
FWLTR	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
<b>BFWLTR</b>	<b>Wikipedia</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	FWLTR
1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2	1.000	1	1	1	1	1	1	1	1	1	1
3	1.000	1	1	1	1	1	1	1	1	1	1
4	1.000	1	1	1	1	1	1	1	1	1	1
5	1.000	1	1	1	1	1	1	1	1	1	1
6	1.000	1	1	1	1	1	1	1	1	1	1
7	1.000	1	1	1	1	1	1	1	1	1	1
8	1.000	1	1	1	1	1	1	1	1	1	1
9	1.000	1	1	1	1	1	1	1	1	1	1
10	1.000	1	1	1	1	1	1	1	1	1	1
FWLTR	1.000	1	1	1	1	1	1	1	1	1	1



Tabla 5.23: Coeficientes de correlación para las medidas de centralidad BLTR con diferentes valores de  $p$ , aplicadas a las redes ArXiv, Caida y Dining Table. Los coeficientes de Kendall ( $\tau$ ) se muestran en la región triangular superior, los coeficientes de Spearman ( $\rho$ ) se muestran en la triangular inferior. Los valores con un  $p$ -valor mayor que 0.05 se muestran en rojo.

<b>BLTR</b>	<b>ArXiv</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	LTR
1	1	0.992	0.988	0.987	0.986	0.986	0.986	0.986	0.986	0.986	0.986
2	0.999	1	0.997	0.996	0.996	0.996	0.995	0.995	0.995	1.000	0.995
3	0.999	1.000	1	0.999	0.999	0.999	0.999	0.999	0.999	1.000	0.999
4	0.998	1.000	1.000	1	1.000	1.000	1.000	1.000	0.999	1.000	0.999
5	0.998	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000
6	0.998	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000
7	0.998	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000
8	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000
9	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1	1
10	0.998	0.995	0.999	0.999	1.000	1.000	1.000	1.000	1	1	1
LTR	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1	1
<b>BLTR</b>	<b>Caida</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	LTR
1	1	0.998	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.991
2	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.992
3	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.992
4	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	0.992
5	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	0.992
6	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	0.992
7	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	0.992
8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	0.992
9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	0.992
10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	0.992
LTR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
<b>BLTR</b>	<b>Dining Table</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	LTR
1	1	0.974	0.965	0.965	0.965	0.965	0.965	0.965	0.965	0.965	0.965
2	0.986	1	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.999	0.995
3	0.983	0.999	1	1	1	1	1	1	1	1	1
4	0.983	0.999	1	1	1	1	1	1	1	1	1
5	0.983	0.999	1	1	1	1	1	1	1	1	1
6	0.983	0.999	1	1	1	1	1	1	1	1	1
7	0.983	0.999	1	1	1	1	1	1	1	1	1
8	0.983	0.999	1	1	1	1	1	1	1	1	1
9	0.983	0.999	1	1	1	1	1	1	1	1	1
10	0.983	0.995	1	1	1	1	1	1	1	1	1
LTR	0.983	0.999	1	1	1	1	1	1	1	1	1

Tabla 5.24: Coeficientes de correlación para las medidas de centralidad BLTR con diferentes valores de  $p$ , aplicadas a las redes **Dolphins**, **Epinions** y **Higgs**. Los coeficientes de **Kendall** ( $\tau$ ) se muestran en la región triangular superior, los coeficientes de **Spearman** ( $\rho$ ) se muestran en la triangular inferior. Los valores con un  $p$ -valor mayor que 0.05 se muestran en rojo.

<b>BLTR</b>	<b>Dolphins</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	LTR
1	1	0.982	0.980	0.980	0.980	0.980	0.980	0.980	0.980	0.980	0.980
2	0.994	1	0.999	0.999	0.999	0.999	0.999	0.999	0.999	1.000	0.999
3	0.993	1.000	1	1	1	1	1	1	1	1	1
4	0.993	1.000	1	1	1	1	1	1	1	1	1
5	0.993	1.000	1	1	1	1	1	1	1	1	1
6	0.993	1.000	1	1	1	1	1	1	1	1	1
7	0.993	1.000	1	1	1	1	1	1	1	1	1
8	0.993	1.000	1	1	1	1	1	1	1	1	1
9	0.993	1.000	1	1	1	1	1	1	1	1	1
10	0.993	0.999	1	1	1	1	1	1	1	1	1
LTR	0.993	1.000	1	1	1	1	1	1	1	1	1
<b>BLTR</b>	<b>Epinions</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	LTR
1	1	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
2	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000
6	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000
7	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000
8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000
9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000
10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000
LTR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
<b>BLTR</b>	<b>Higgs</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	LTR
1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	1.000	1.000	1.000	1	1	1	1	1	1	1	1
5	1.000	1.000	1.000	1	1	1	1	1	1	1	1
6	1.000	1.000	1.000	1	1	1	1	1	1	1	1
7	1.000	1.000	1.000	1	1	1	1	1	1	1	1
8	1.000	1.000	1.000	1	1	1	1	1	1	1	1
9	1.000	1.000	1.000	1	1	1	1	1	1	1	1
10	1.000	1.000	1.000	1	1	1	1	1	1	1	1
LTR	1.000	1.000	1.000	1	1	1	1	1	1	1	1

Tabla 5.25: Coeficientes de correlación para las medidas de centralidad BLTR con diferentes valores de  $p$ , aplicadas a las redes Human Brain, Texas y Wikipedia. Los coeficientes de Kendall ( $\tau$ ) se muestran en la región triangular superior, los coeficientes de Spearman ( $\rho$ ) se muestran en la triangular inferior. Los valores con un  $p$ -valor mayor que 0.05 se muestran en rojo.

<b>BLTR</b>		<b>Human Brain</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	LTR	
1	1	0.992	0.986	0.976	0.973	0.972	0.972	0.972	0.972	0.972	0.972	
2	1.000	1	0.992	0.982	0.979	0.977	0.977	0.977	0.977	0.998	0.977	
3	0.999	1.000	1	0.990	0.986	0.985	0.985	0.984	0.984	0.999	0.984	
4	0.998	0.998	0.999	1	0.996	0.994	0.994	0.994	0.994	1.000	0.994	
5	0.997	0.998	0.999	1.000	1	0.999	0.999	0.998	0.998	1.000	0.998	
6	0.997	0.998	0.999	1.000	1.000	1	1.000	0.999	0.999	1.000	0.999	
7	0.997	0.998	0.999	1.000	1.000	1.000	1	0.999	0.999	1.000	0.999	
8	0.997	0.998	0.999	1.000	1.000	1.000	1.000	1	1	1	1	
9	0.997	0.998	0.999	1.000	1.000	1.000	1.000	1	1	1	1	
10	0.997	0.977	0.984	0.994	0.998	0.999	0.999	1	1	1	1	
LTR	0.997	0.998	0.999	1.000	1.000	1.000	1.000	1	1	1	1	
<b>BLTR</b>		<b>Texas</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	LTR	
1	1	0.985	0.983	0.982	0.982	0.982	0.982	0.982	0.982	0.982	0.982	
2	0.996	1	0.999	0.998	0.998	0.998	0.998	0.998	0.998	1.000	0.998	
3	0.995	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
4	0.995	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
5	0.995	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	
6	0.995	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	
7	0.995	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	
8	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	
9	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	
10	0.995	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	
LTR	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	
<b>BLTR</b>		<b>Wikipedia</b>										
$\rho / \tau$	1	2	3	4	5	6	7	8	9	10	LTR	
1	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
2	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
3	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
4	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
5	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	1.000	
6	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	1.000	
7	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	1.000	
8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	1.000	
9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	1.000	
10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	1.000	
LTR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	

### 5.4.2. Tiempos de ejecución

Igual que con FLTM, al ver que BLTM se comporta prácticamente igual que LTM, pensamos que sería interesante comparar los tiempos de ejecución de los dos tipos de medidas. Quizás al limitar el nivel de expansión máximo de las medidas, los tiempos de ejecución serán más bajos que las medidas LTM, y al mismo tiempo nos dan resultados prácticamente iguales a los de las medidas LTM.

Como en FLTM, ejecutaremos cada medida 5 veces en cada red. Para cada red mostraremos los resultados en dos gráficas diferentes. Estas gráficas son las mismas que definimos en la sección 5.3.2, para los tiempos de ejecución de las medidas FLTM, pero en vez de usar diferentes valores de  $\gamma$ , usamos  $p = \{1, 2, \dots, 10\}$ . También comentaremos solo los resultados más relevantes, dejando en el apéndice C.2 el resto de resultados para su consulta.

En las figuras 5.15-5.23, vemos las gráficas de los tiempos de ejecución para redes **Dining Table**, **Caida** y **Texas**.

En las redes pequeñas como **Dinning Table**, los tiempos de ejecución varían más, debido a su reducido tamaño. Aun así, vemos como no hay relación entre el parámetro  $p$  y el tiempo de ejecución. También vemos como en general, las medidas basadas en BLTM, van ligeramente más lentas que las medidas basadas en LTM. En **Caida**, al tener un tamaño mayor, vemos como los tiempos de ejecución empiezan a estabilizarse, y a ser más similares a los de las medidas LTM. Seguimos viendo como el parámetro  $p$  no parece influir de una manera concreta en los tiempos de ejecución. Finalmente, en redes más grandes, como en **Texas**, vemos que los tiempos de ejecución son muy parecidos entre ellos, y sin ninguna relación con  $p$ . Así pues, igual que en FLTM, los tiempos de ejecución de las medidas basadas en el BLTM son prácticamente iguales que las medidas basadas en el LTM.

# Dining\_Table

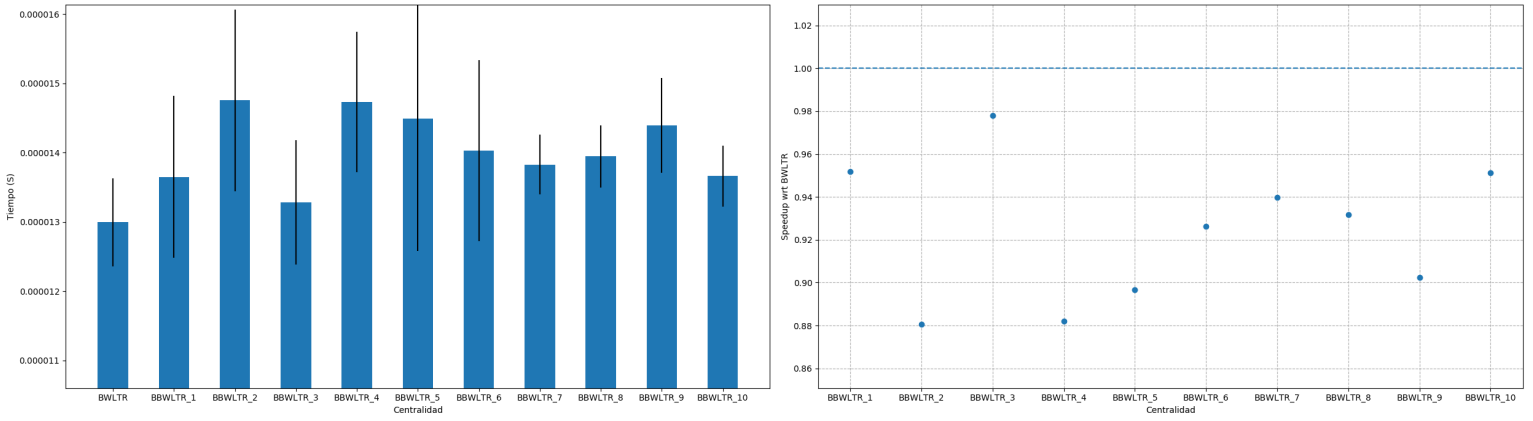


Figura 5.15: BWLTR vs BBWLTR

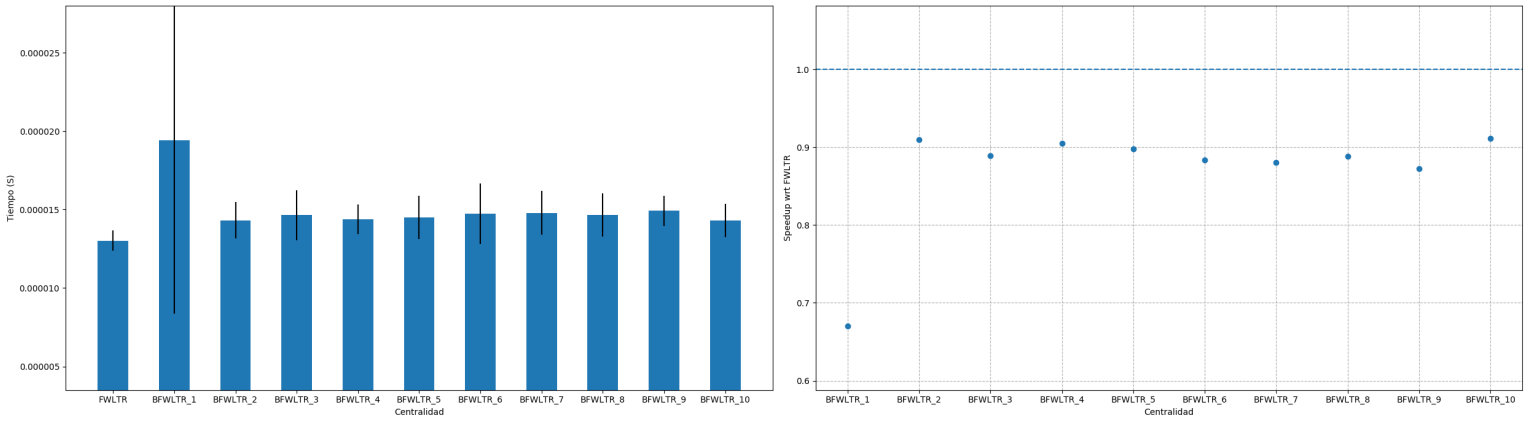


Figura 5.16: FWLTR vs BFWLTR

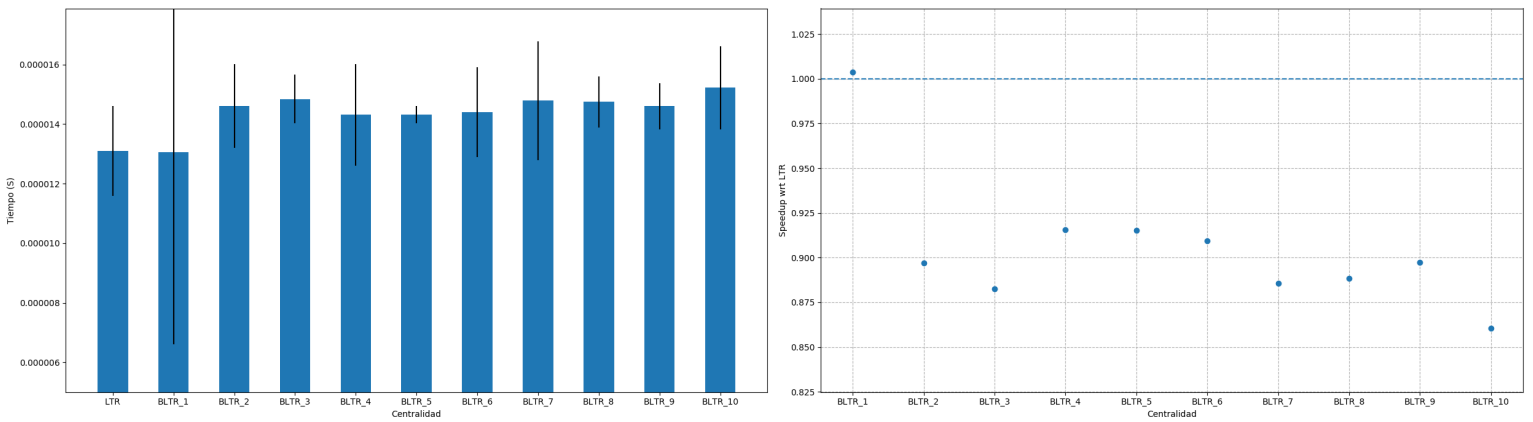


Figura 5.17: LTR vs BLTR

# Caida

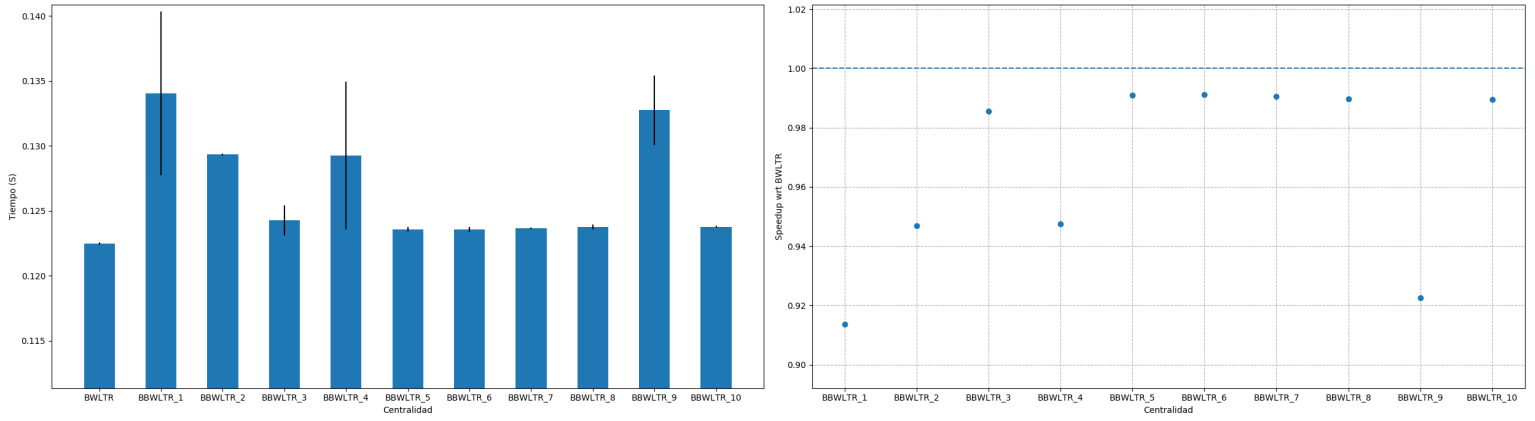


Figura 5.18: BWTTR vs BBWTTR

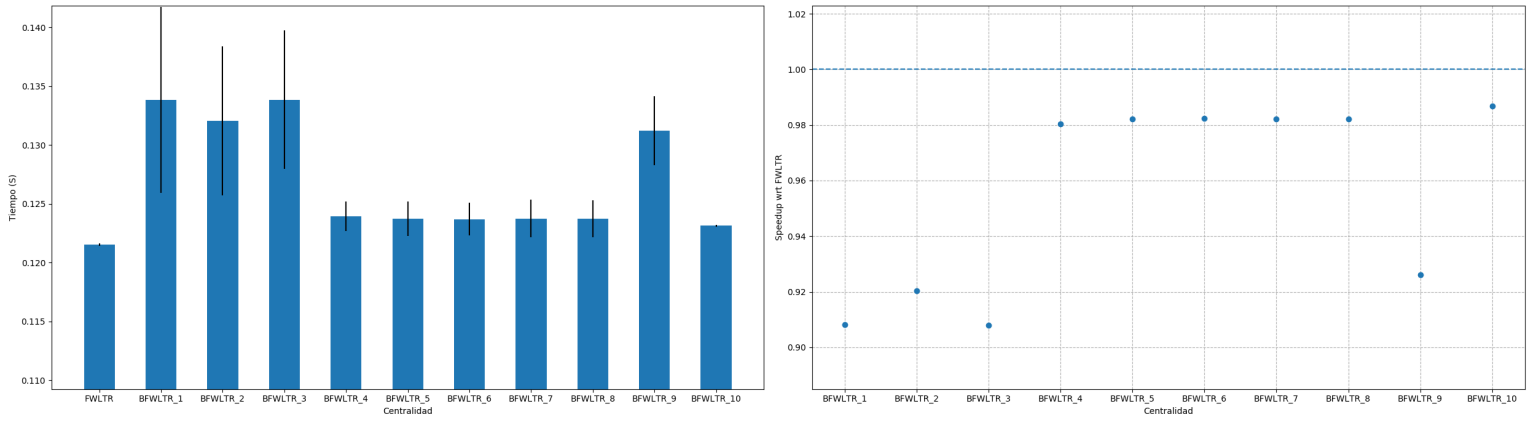


Figura 5.19: FWLTR vs BFWLTR

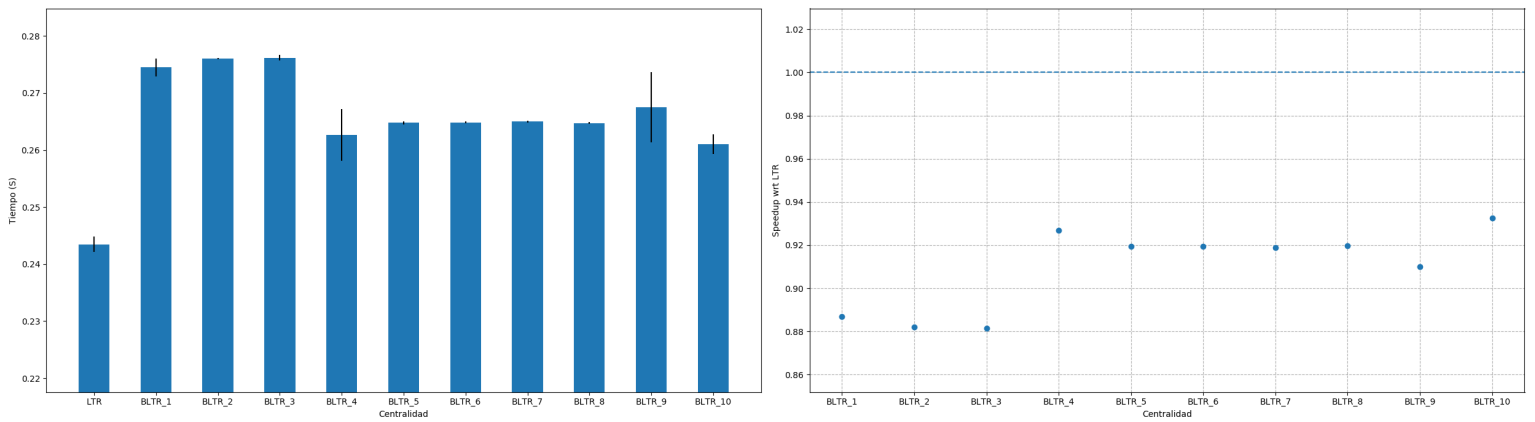


Figura 5.20: LTR vs BLTR

# Texas

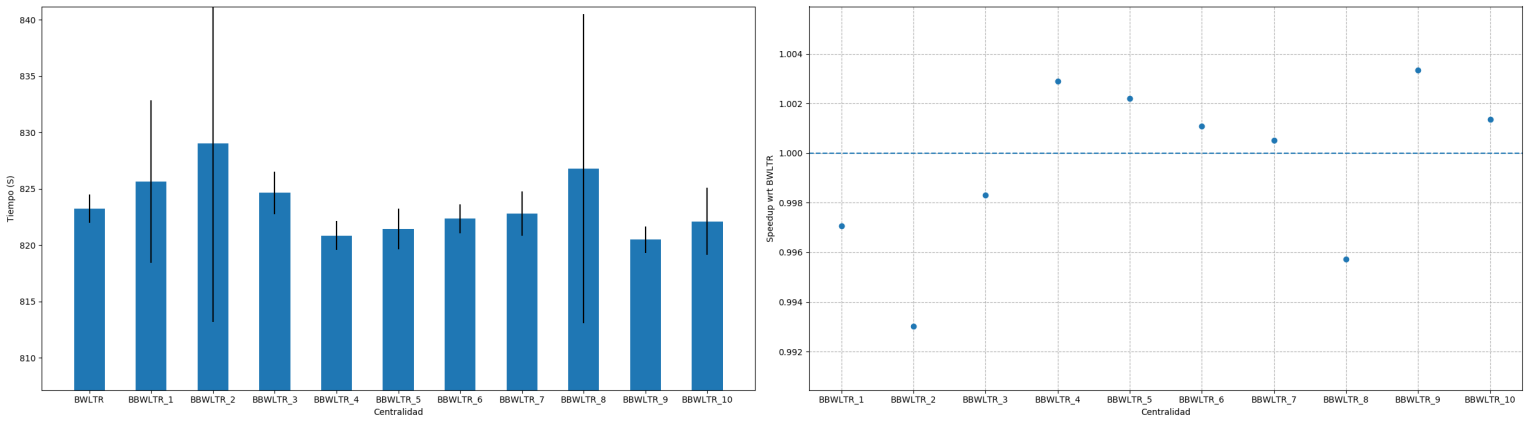


Figura 5.21: BWLTR vs BBWLTR

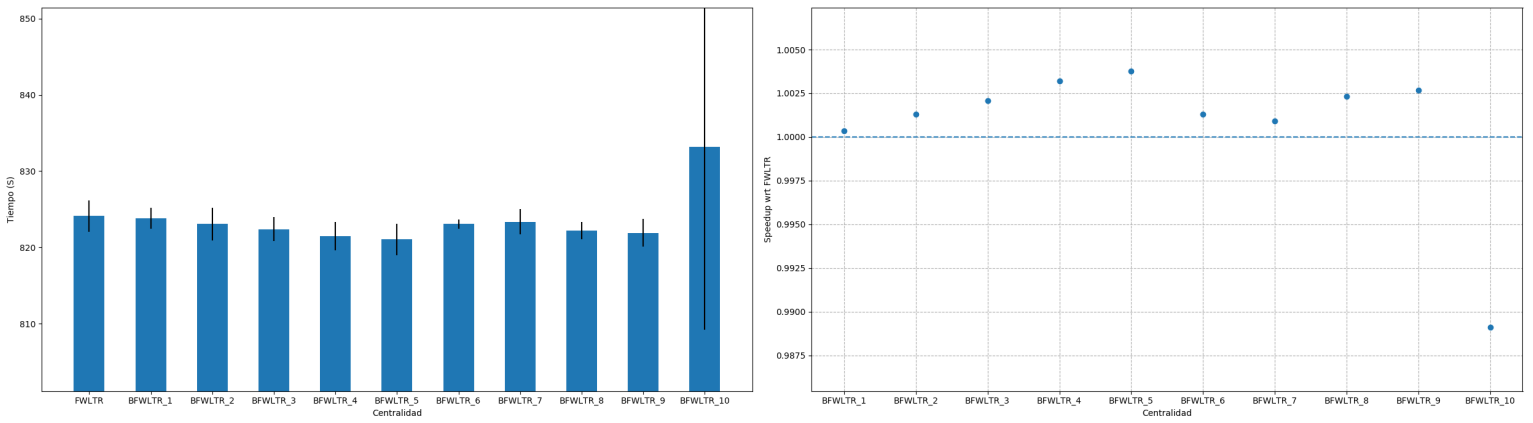


Figura 5.22: FWLTR vs BFWLTR

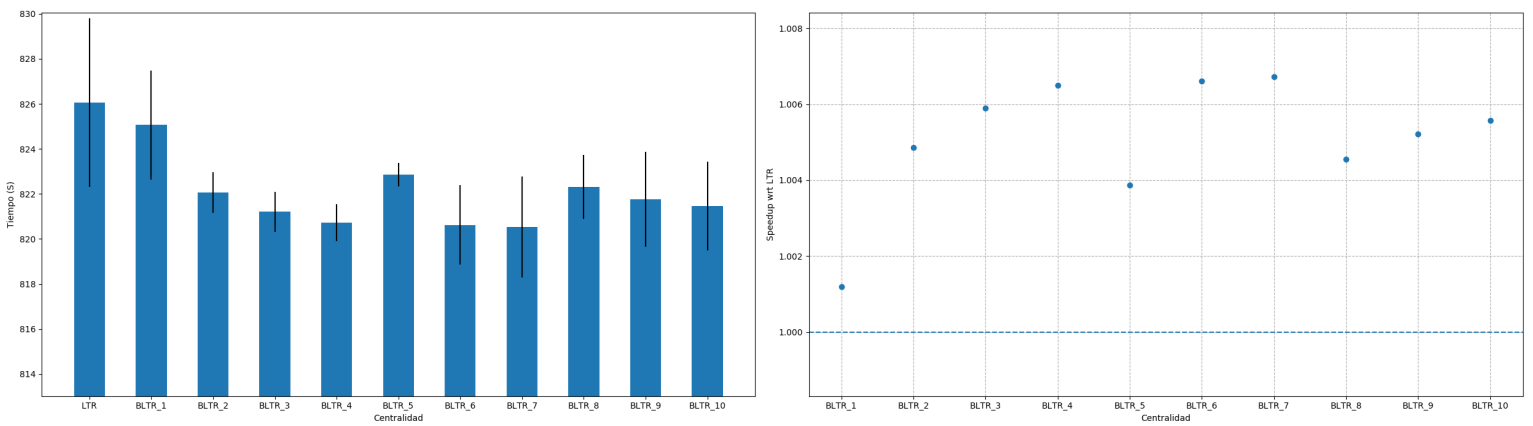


Figura 5.23: LTR vs BLTR

## Capítulo 6

# Conclusiones y Trabajo Futuro

### 6.1. Conclusiones

En este proyecto, hemos continuado el estudio realizado en [5, 16], sobre las medidas de centralidad basadas en el modelo de expansión de influencia *Linear Threshold Model* (LTM).

Hemos introducido dos medidas de centralidad basadas en el *Linear Threshold Rank* (LTR), el *Forward Linear Threshold Rank* (FWLTR) y el *Backward Linear Threshold Rank* (BWLTR). Las cuales modifican la definición de conjunto inicial del LTR. Usando como conjunto inicial, el propio actor al que se le mide la centralidad, más los actores sucesores directos de él en la red, para el FWLTR, y para el BWLTR, los actores predecesores directos.

Hemos profundizado el estudio del *Fading Linear Threshold Model* realizado en [16], utilizando las tres medidas del modelo LTM para definir tres nuevas medidas bajo este modelo. El *Fading Linear Threshold Rank* (FLTR), el *Backward Fading Linear Threshold Rank* (BWFLTR) y el *Forward Fading Linear Threshold Rank* (FWFLTR).

También se ha introducido una nueva variación del modelo LTM, el *Bounded Linear Threshold Model* (BLTM). El cual consiste en limitar la influencia que puede ejercer un actor hasta un cierto nivel de expansión  $p$ . Para este modelo, también hemos utilizado las tres medidas basadas en el LTM, para definir tres de nuevas bajo este modelo. El *Bounded Linear Threshold Rank* (BLTM), el *Bounded Back-*



*ward Linear Threshold Rank (BBWLTR)* y el *Bounded Forward Linear Threshold Rank (BFWLTR)*.

Para calcular todas estas medidas, hemos diseñado un algoritmo simple y eficiente. Este algoritmo ha sido implementado en C++, para poder ejecutarlo de forma eficiente sobre redes con centenares de miles de actores. Para reducir aún más los tiempos de ejecución, hemos paralelizado la asignación de valores de centralidad a los actores de una red, con la API `OpenMP`. Tal y como hemos visto en los tiempos de ejecución de los experimentos, esto ha sido así, pasando de tiempos de ejecución del orden de horas para redes más grandes en [15], a tiempos de ejecución del orden de varios segundos en las mismas redes, o minutos en una red mucho más grande (**Texas**) con millones de nodos.

Hemos comparado las medidas basadas en el modelo LTM, con otras medidas de centralidad bien conocidas, sobre varias redes de diferentes características y tamaños. En estos experimentos hemos visto como en la mayoría de redes dirigidas, las medidas LTR y BWLTR son bastante similares. Por otra parte, pensamos que la medida FWLTR es la más realista de las tres. Ya que tanto en LTR como en BWLTR, en el conjunto inicial añadimos los actores que influyen al actor que se le toma la medida. Esto puede hacer que si un actor, con una capacidad reducida para influenciar a pocos actores directamente, si recibe influencia de otro actor con mucha capacidad de influencia, esta capacidad se transmite al primero, y por lo tanto parece que tenga mucha más capacidad de influencia del que realmente tiene, lo cual creemos que no es muy realista.

Para estas mismas medidas, hemos observado como al clasificar los actores en tres categorías diferentes, al pasar de una medida de centralidad a otra, los actores migran entre estas tres categorías. Para ello, hemos utilizado dos criterios distintos de clasificación de actores, el primero se basa en la posición de los actores al estar ordenados de forma decreciente, y se introdujo en [16]. El segundo, clasifica a los actores según el valor que hayan obtenido en una medida, este ha sido introducido por nosotros. Hemos observado que parece haber alguna relación entre el coeficiente de desigualdad de **gini**, cuando hay migraciones extremas entre categorías lejanas. Además estas migraciones extremas se acentúan utilizando el segundo criterio de clasificación.

Finalmente, en las medidas basadas en el modelo FLTM, hemos visto que con varios valores de  $\gamma \in [0, 1]$  las medidas se comportan prácticamente igual que las medidas basadas en el modelo LTM, tanto en resultados, como en tiempos de ejecución, incluso en una red con un gran diámetro. Observamos este mismo fenómeno con medias de centralidad basadas en el modelo BLTM, utilizando distintos valores de  $p \in \{1, 2, \dots, 10\}$ . Los resultados y tiempos de ejecución son prácticamente idénticos a los de las medidas de centralidad basadas en el modelo

LTM. Si  $p$  tomara como valor el diámetro de la red, explicaría que los resultados fueran idénticos, pero vemos que los resultados son idénticos o casi idénticos, con valores de  $p$  bastante más pequeños que el diámetro de la red. Con esto deducimos que realmente, los actores solo son capaces de ejercer su influencia hasta niveles bajos de expansión. Esto explicaría que en las medidas basadas en el FLTM se obtengan los mismos resultados, ya que al no poder expandir la influencia a niveles más altos, no hay suficiente tiempo como para que la influencia ejercida en los primeros niveles de expansión se diluya.

## 6.2. Trabajo Futuro

En la sección 5.2.3, hemos visto como las migraciones de actores entre dos medidas distintas, varían considerablemente de una red a otra entre las mismas dos medidas. Sería interesante pues, ver si se puede usar esta migración de actores para caracterizar una red, y ver como de útil es al distinguir dos redes distintas.

El hecho de que los actores solo sean capaces de influenciar a otros actores relativamente cercanos a ellos, utilizando como conjunto inicial a sus vecinos (en diferentes direcciones), debería de investigarse si solo ocurre con las redes que hemos utilizado, o si por el contrario es algo común en la mayoría de redes. En caso de que así fuera, se debería de tener en cuenta este fenómeno al diseñar nuevas medidas basadas en el modelo LTM, ya que si no se tiene en cuenta, puede pasar lo mismo que nos ha pasado con las medidas basadas en los otros dos modelos, y obtener unas medidas que se comporten igual que las originales.

Recientemente, en [35], se introdujo una nueva medida de centralidad basada en el modelo de influencia LTM. Esta nueva medida, es una generalización del LTR. En ella en vez de usar solo los vecinos directos de un actor como conjunto inicial, dado un nivel  $l$ , con  $0 \leq l \leq d$  (siendo  $d$  el diámetro de la red), se usan los vecinos a distancia como mucho  $l$  del actor como conjunto inicial. En los experimentos que se realizan en este paper, se ve como a partir de un cierto valor  $l$ , el rango de valores que la medida asigna aumenta considerablemente, incrementando también, el número de actores que es capaz de distinguir. Con valores más grandes de  $l$ , la medida vuelve a asignar menos valores distintos y a distinguir a menos actores, ya que la mayoría de ellos obtienen un valor cercano a 1.

Los valores de  $l$  para los que se dan estos 'saltos' son bastante pequeños (entre 4 y 6 según el paper). Esos números podrían estar relacionados con el fenómeno que hemos visto en nuestros experimentos. Al usar los actores que están a esta distancia del actor, como conjunto de activación, somos capaces de pasar esta

especie de límite en el nivel de expansión que hemos observado, y poder influenciar a muchos más actores. Sería interesante ver si esto se reproduce en más redes, y ver hasta que punto es rentable 'comprar' más actores para el conjunto inicial de activación, teniendo en cuenta que con un cierto número podemos llegar a influenciar a muchos más.

# Bibliografía

- [1] Alex Bavelas. A mathematical model for group structures. *Human Organization*, 7(3):16–30, 1948.
- [2] Linton C. Freeman. Centrality in social networks conceptual clarification. *Social Networks*, 1(3):215 – 239, 1978.
- [3] Mikail Rubinov and Olaf Sporns. Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, 52(3):1059 – 1069, 2010. Computational Models of the Brain.
- [4] John Cardente. Using centrality measures to identify key members of an innovation collaboration network. Technical report, Stanford University, 2012.
- [5] Fabián Riquelme and Pablo González-Cantergiani. Measuring user influence on twitter: A survey. *Information Processing Management*, 52(5):949 – 975, 2016.
- [6] David C Bell, John S Atkinson, and Jerry W Carlson. Centrality measures for disease transmission networks. *Social Networks*, 21(1):1 – 21, 1999.
- [7] Paolo Crucitti, Vito Latora, and Sergio Porta. Centrality measures in spatial networks of urban streets. *Phys. Rev. E*, 73:036125, Mar 2006.
- [8] Leo Katz. A new status index derived from sociometric analysis. *Psychometrika*, 18(1):39–43, Mar 1953.
- [9] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. Technical Report 1999-66, Stanford InfoLab, November 1999. Previous number = SIDL-WP-1999-0120.
- [10] Phillip Bonacich. Factoring and weighting approaches to status scores and clique identification. *The Journal of Mathematical Sociology*, 2(1):113–120, 1972.

- [11] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. *Theory of Computing*, 11(4):105–147, 2015.
- [12] Suman Kundu, C. A. Murthy, and S. K. Pal. A new centrality measure for influence maximization in social networks. In Sergei O. Kuznetsov, Deba P. Mandal, Malay K. Kundu, and Sankar K. Pal, editors, *Pattern Recognition and Machine Intelligence*, pages 242–247, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.
- [13] S.K. Pal, S. Kundu, and C.A. Murthy. Centrality measures, upper bound, and influence maximization in large scale directed social networks. *Fundamenta Informaticae*, Vol. 130, nr 3:317–342, 2014.
- [14] Ibrahima Gaye, Gervais Mendy, Samuel Ouya, and Diaraf Seck. New centrality measure in social networks based on independent cascade (ic) model.
- [15] Fabián Riquelme, Pablo Gonzalez-Cantergiani, Xavier Molinero, and Maria Serna. Centrality measure in social networks based on linear threshold model. *Knowledge-Based Systems*, 140:92 – 102, 2018.
- [16] Eline Van Hove. The linear threshold rank as centrality measure in social networks, 2018.
- [17] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 137–146. ACM, 2003.
- [18] Stanley Wasserman and Katherine Faust. *Social network analysis: Methods and applications*, volume 8, page 201. Cambridge university press, 1994.
- [19] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. Technical report, Stanford InfoLab, 1999.
- [20] Thomas Piketty. About capital in the twenty-first century. *American Economic Review*, 105(5):48–53, May 2015.
- [21] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. Graph evolution: Densification and shrinking diameters. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 1(1):2, 2007.
- [22] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. Graphs over time: densification laws, shrinking diameters and possible explanations. In *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*, pages 177–187. ACM, 2005.
- [23] Matthew Richardson, Rakesh Agrawal, and Pedro Domingos. Trust management for the semantic web. In *International semantic Web conference*, pages 351–368. Springer, 2003.

- [24] Manlio De Domenico, Antonio Lima, Paul Mougél, and Mirco Musolesi. The anatomy of a scientific rumor. *Scientific reports*, 3:2980, 2013.
- [25] Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. Signed networks in social media. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1361–1370. ACM, 2010.
- [26] Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. Predicting positive and negative links in online social networks. In *Proceedings of the 19th international conference on World wide web*, pages 641–650. ACM, 2010.
- [27] Jure Leskovec, Kevin J Lang, Anirban Dasgupta, and Michael W Mahoney. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics*, 6(1):29–123, 2009.
- [28] Jure Leskovec and Andrej Krevl. SNAP Datasets: Stanford large network dataset collection. <http://snap.stanford.edu/data>, June 2014.
- [29] Wouter De Nooy, Andrej Mrvar, and Vladimir Batagelj. *Exploratory social network analysis with Pajek: Revised and expanded edition for updated software*, volume 46. Cambridge University Press, 2018.
- [30] David Lusseau, Karsten Schneider, Oliver J Boisseau, Patti Haase, Elisabeth Slooten, and Steve M Dawson. The bottlenose dolphin community of doubtful sound features a large proportion of long-lasting associations. *Behavioral Ecology and Sociobiology*, 54(4):396–405, 2003.
- [31] Ryan A. Rossi and Nesreen K. Ahmed. The network data repository with interactive graph analytics and visualization. In *AAAI*, 2015.
- [32] Balázs Szalkai, Csaba Kerepesi, Bálint Varga, and Vince Grolmusz. The budapest reference connectome server v2. 0. *Neuroscience Letters*, 595:60–62, 2015.
- [33] Balázs Szalkai, Csaba Kerepesi, Bálint Varga, and Vince Grolmusz. Parameterizable consensus connectomes from the human connectome project: The budapest reference connectome server v3. 0. *Cognitive neurodynamics*, 11(1):113–116, 2017.
- [34] Kurt C. Foster, Stephen Q. Muth, John J. Potterat, and Richard B. Rothenberg. A faster katz status score algorithm. *Computational & Mathematical Organization Theory*, 7(4):275–285, Dec 2001.
- [35] Fabián Riquelme, Pablo Gonzalez-Cantergiani, Xavier Molinero, and Maria Serna. The neighborhood role in the linear threshold rank on social networks. *Physica A: Statistical Mechanics and its Applications*, 528:121430, 2019.
- [36] Ulrik Brandes. A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology*, 25:163–177, 2001.

# Apéndice A

## Códigos

Adjuntamos los códigos más relevantes que hemos utilizado.

### A.1. Graph

Esta es la clase escrita en C++ que hemos utilizado para representar los grafos de influencia.

#### A.1.1. Graph.h

```
#include <vector>
#include <unordered_map>

using namespace std;

typedef vector<int> VI;
typedef vector<double> VD;
typedef pair<int,double> PID;
typedef vector<PID> VPID;
typedef vector<VPID> VVPID;

struct edge
{
    int v;
    int u;
    double w;
    edge(int v, int u, double w) : v(v), u(u), w(w){}
};
```

```

typedef edge edge;
typedef vector<edge> VE;

typedef unordered_map<int, int> UMII;

class Graph
{
public:
    int N = 0;
    int E = 0;
    bool directed = false;
    VD threshold;
    VI mapping;
    VVPID adjacency;
    VVPID predecessors;

    Graph(const VE &Edges, bool directed);
    Graph();

    Graph stochastic() const;

    VI dangling_nodes() const;
};

```



## A.1.2. Graph.cpp

```
#include "Graph.h"

Graph::Graph(const VE &Edges, bool directed)
{
    E = static_cast<int>(Edges.size());
    this->directed = directed;
    adjacency = VVPID();
    mapping = VI();
    threshold = VD();
    if (directed) predecessors = VVPID();

    int last_node = 0;
    UMII MP;

    for (edge e : Edges)
    {
        if (MP.find(e.v) == MP.end())
        {
            mapping.push_back(e.v);
            MP[e.v] = last_node++;
            adjacency.push_back(VPID());
            if (directed) predecessors.push_back(VPID());
            threshold.push_back(0);
        }
        if (MP.find(e.u) == MP.end())
        {
            mapping.push_back(e.u);
            MP[e.u] = last_node++;
            adjacency.push_back(VPID());
            if (directed) predecessors.push_back(VPID());
            threshold.push_back(0);
        }
        int v = MP[e.v];
        int u = MP[e.u];

        adjacency[v].push_back(make_pair(u, e.w));
        threshold[u] += e.w;

        if (not directed)
        {
            adjacency[u].push_back(make_pair(v, e.w));
            threshold[v] += e.w;
        }
    }
}
```

```

        else predecessors[u].push_back(make_pair(v, e.w));
    }

    N = static_cast<int>(threshold.size());

    for (double &v : threshold) v = int(v) / 2 + 1;
}

Graph Graph::stochastic() const
{
    Graph G = Graph();
    G.N = this->N;
    G.E = this->E;
    G.directed = this->directed;
    G.mapping = this->mapping;

    VVPID adj;
    for (const VPID &V : this->adjency)
    {
        VPID aux;
        double sw = 0;
        for (PID v : V) sw += v.second;

        if (sw > 0) for (PID v : V) aux.push_back(make_pair(v.first,v.second/sw));
        else for (PID v : V) aux.push_back(make_pair(v.first,0));
        adj.push_back(aux);
    }
    G.adjency = adj;

    if (G.directed)
    {
        VVPID pred;
        for (const VPID &V : this->predecessors)
        {
            VPID aux;
            double sw = 0;
            for (PID v : V) sw += v.second;

            if (sw > 0) for (PID v : V) aux.push_back(make_pair(v.first,v.second/sw));
            else for (PID v : V) aux.push_back(make_pair(v.first,0));

            pred.push_back(aux);
        }
        G.predecessors = pred;
    }
}

```

```
G.threshold = VD(G.N, 1/2 + 1);

return G;
}

VI Graph::dangling_nodes() const
{
    VI res;
    for (int v = 0; v < N; ++v) if (adjency[v].empty()) res.push_back(v);
    return res;
}

Graph::Graph() = default;
```

## A.2. Otras medidas de Centralidad

Adjuntamos las funciones que hemos utilizado para calcular las diferentes centralidades.

### A.2.1. Degree

```
void rank_degree(VD &ranking, const Graph &G)
{
    int n = G.N;
    ranking = VD(n, 0);

    #pragma omp parallel for
    for (int v = 0; v < n; ++v)
    {
        if (G.directed)
            ranking[v] = (double) (G.adjacency[v].size() + G.predecessors[v].size()) / (n - 1);
        else ranking[v] = (double) G.adjacency[v].size() / (n - 1);
    }
}
```

### A.2.2. Closeness

```
void rank_closeness(VD &ranking, const Graph &G)
{
    int n = G.N;
    ranking = VD(n, 0);

    #pragma omp parallel for
    for (int s = 0; s < n; ++s)
    {
        double ncc = 0;
        double dt = 0;
        VI D(n, -1);
        D[s] = 0;
        QI Q;
        Q.push(s);

        while (not Q.empty())
        {
            int v = Q.front();
            Q.pop();
```

```

if (G.directed)
{
  for (PID e : G.predecessors[v])
  {
    int u = e.first;
    if (D[u] < 0)
    {
      ++ncc;
      D[u] = D[v] + 1;
      dt += D[u];
      Q.push(u);
    }
  }
}
else
{
  for (PID e : G.adjacency[v])
  {
    int u = e.first;
    if (D[u] < 0)
    {
      ++ncc;
      D[u] = D[v] + 1;
      dt += D[u];
      Q.push(u);
    }
  }
}
if (dt > 0 and n > 1) ranking[s] = (ncc / dt) * (ncc / (n - 1));
}
}

```

### A.2.3. Betweenness

Para calcular el Betweenness utilizamos el algoritmo de *Brandes* introducido en [36]

```
void rank_betweenness(VD &ranking, const Graph &G)
{
    int n = G.N;
    ranking = VD(n, 0);
    #pragma omp parallel for schedule(dynamic,25)
    for (int s = 0; s < n; ++s)
    {
        SI S;
        VLI P(n);
        VD R(n, 0);
        R[s] = 1;
        VI D(n, -1);
        D[s] = 0;
        QI Q;
        Q.push(s);
        while (not Q.empty())
        {
            int v = Q.front();
            Q.pop();
            S.push(v);
            for (PID e: G.adjacency[v])
            {
                int w = e.first;
                if (D[w] < 0)
                {
                    Q.push(w);
                    D[w] = D[v] + 1;
                }
                if (D[w] == D[v] + 1)
                {
                    R[w] += R[v];
                    P[w].push_back(v);
                }
            }
        }
        VD PD(n, 0);
        while (not S.empty())
        {
            int w = S.top();
            S.pop();
            for (int v : P[w]) PD[v] += (R[v] / R[w]) * (1 + PD[w]);
            if (w != s)
            {
                #pragma omp atomic update
                ranking[w] += PD[w];
            }
        }
    }
    if (G.directed) for (double &r : ranking) r /= (n - 1) * (n - 2);
    else for (double &r : ranking) r /= ((double) ((n - 1) * (n - 2)) / 2);
}
```

## A.2.4. Katz

```
void rank_katz(VD &ranking, const Graph &G, double alpha, double beta, double tol, int max_iter = 1000)
{
    int n = G.N;
    VD xlast(n, 1.0);
    bool converged = false;

    for (int i = 0; i < max_iter and not converged; ++i)
    {
        VD x(n, 0);

        #pragma omp parallel for schedule(dynamic, 25)
        for (int v = 0; v < n; ++v)
        {
            for (PID u : G.adjacency[v])
            {
                #pragma atomic update
                x[u.first] += xlast[v];
            }
        }

        #pragma omp parallel for
        for (int v = 0; v < n; ++v) x[v] = alpha * x[v] + beta;

        double error = 0;
        #pragma omp parallel for reduction(+:error)
        for (int v = 0; v < n; ++v) error += fabs(x[v] - xlast[v]);

        if (error < n * tol)
        {
            converged = true;
            double s = 0;
            #pragma omp parallel for reduction(+:s)
            for (int v = 0; v < n; ++v) s += x[v] * x[v];

            if (s > 0) s = 1.0 / sqrt(s);
            else s = 1;

            #pragma omp parallel for
            for (int v = 0; v < n; ++v) x[v] *= s;

            ranking = x;
        } else xlast = x;
    }

    if (not converged)
    {
        cout << "Failed to converge" << endl;
        ranking = VD(n, -1.0);
    }
}
```

## A.2.5. PageRank

```
void rank_page(VD &ranking, const Graph &G, double lambda = 0.85, double tol = 1.0e-6, int max_iter = 100)
{
    Graph W = G.stochastic();
    VI dangling_nodes = W.dangling_nodes();
    int n = W.N;

    VD xlast(n, 1.0 / n);
    bool converged = false;

    for (int i = 0; i < max_iter and not converged; ++i)
    {
        VD x(n, 0);

        double danglesum = 0;
        #pragma omp parallel for reduction(+:danglesum)
        for (int v = 0; v < dangling_nodes.size(); ++v) danglesum += xlast[dangling_nodes[v]];
        danglesum *= lambda;

        #pragma omp parallel for
        for (int v = 0; v < n; ++v)
        {
            for (PID u : W.adjacency[v])
            {
                #pragma omp atomic update
                x[u.first] += lambda * xlast[v] * u.second;
            }
            #pragma omp atomic update
            x[v] += danglesum * (1.0 / n) + (1.0 - lambda) * (1.0 / n);
        }

        long double error = 0;
        #pragma omp parallel for reduction(+:error)
        for (int v = 0; v < n; ++v) error += fabs(x[v] - xlast[v]);

        if (error < n * tol)
        {
            converged = true;
            ranking = x;
        } else xlast = x;
    }

    if (not converged)
    {
        cout << "Failed to converge" << endl;
        ranking = VD(n, -1.0);
    }
}
```



### A.2.6. ICR

```
void rank_icr(VD &ranking, const Graph &G, double p = 0.1, int iter = 100)
{
    int n = G.N;
    ranking = VD(n, iter);

    for (int i = 0; i < iter; ++i)
    {
        random_device rd;
        uniform_real_distribution <double> dst(0.0,1.0);
        cout << i << endl;
        #pragma omp parallel for schedule(dynamic,25)
        for (int s = 0; s < n; ++s)
        {
            VB Seen(n,false); Seen[s] = true;
            QI Q; Q.push(s);

            while (not Q.empty())
            {
                int v = Q.front(); Q.pop();
                for (PID u : G.adjency[v])
                {
                    if (not Seen[u.first] and dst(rd) <= p)
                    {
                        Seen[u.first] = true;
                        Q.push(u.first);
                        ++ranking[s];
                    }
                }
            }
        }
    }

    double max = 0;
    for (int i = 0; i < n; ++i)
    {
        ranking[i] /= iter;
        if (ranking[i] > max) max = ranking[i];
    }

    for (double &r : ranking) r /= max;
}
```

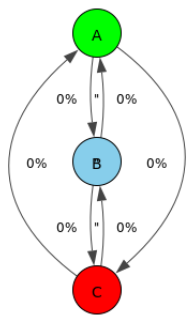
# Apéndice B

## Migraciones de nodos

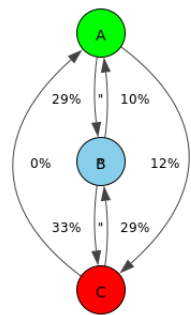
### B.1. Clasificación por posición

#### B.1.1. ArXiv

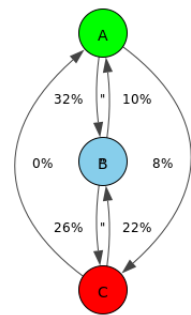
108



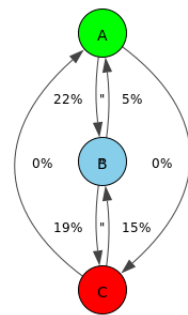
BWLTR-BWLTR



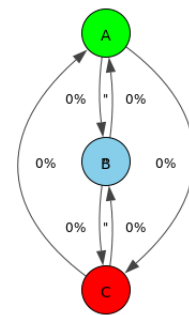
BWLTR-Betweenness



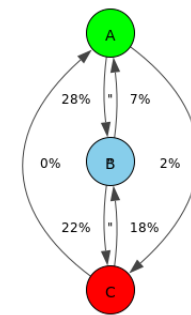
BWLTR-Closeness



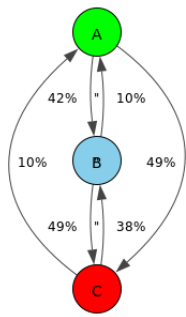
BWLTR-Degree



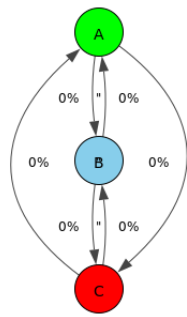
BWLTR-FWLTR



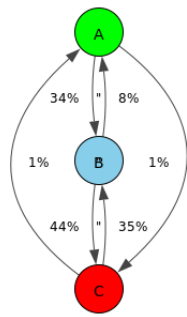
BWLTR-ICR



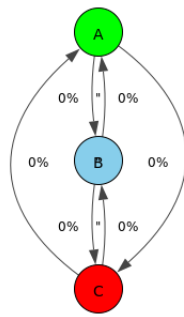
BWLTR-Katz



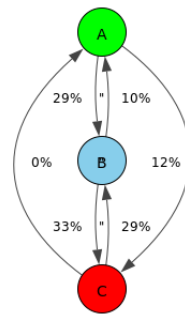
BWLTR-LTR



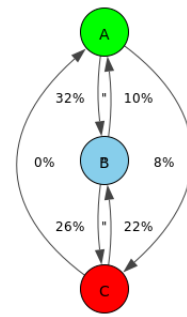
BWLTR-PageRank



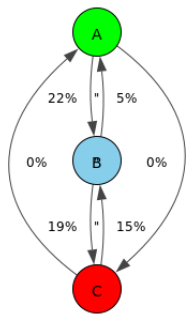
FWLTR-BWLTR



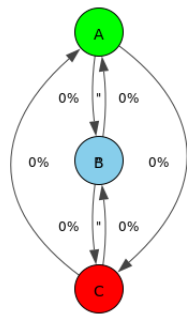
FWLTR-Betweenness



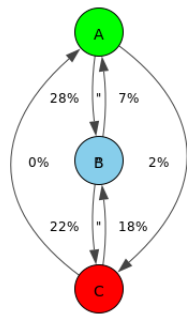
FWLTR-Closeness



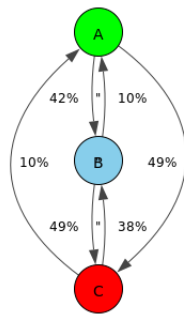
FWLTR-Degree



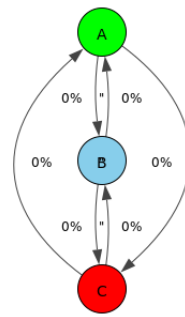
FWLTR-FWLTR



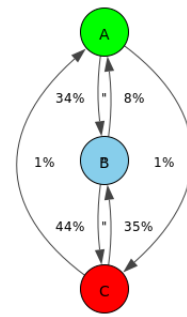
FWLTR-ICR



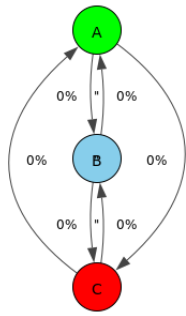
FWLTR-Katz



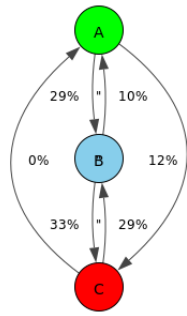
FWLTR-LTR



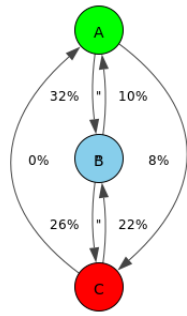
FWLTR-PageRank



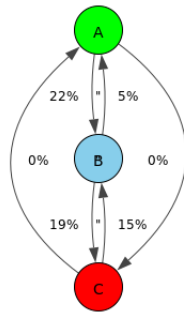
LTR-BWLTR



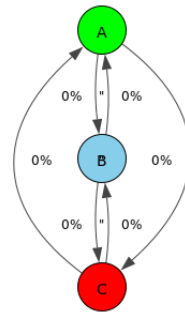
LTR-Betweenness



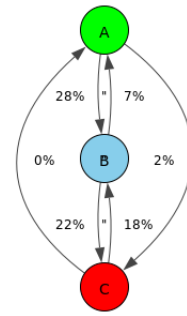
LTR-Closeness



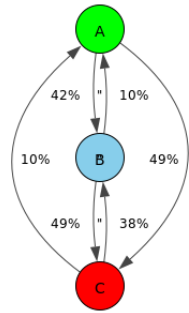
LTR-Degree



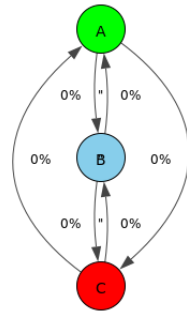
LTR-FWLTR



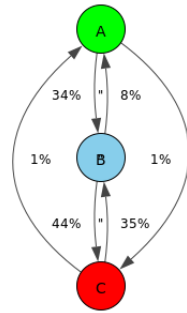
LTR-ICR



LTR-Katz

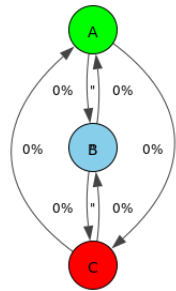


LTR-LTR

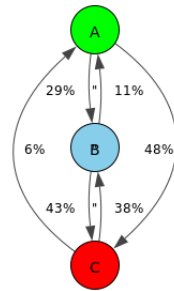


LTR-PageRank

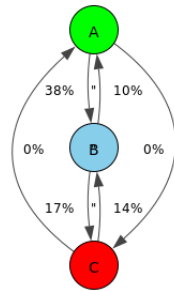
### B.1.2. Caida



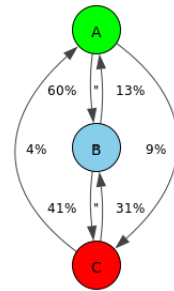
BWLTR-BWLTR



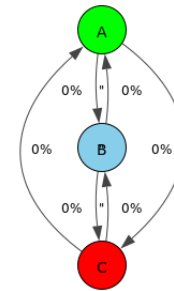
BWLTR-Betweenness



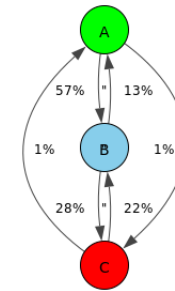
BWLTR-Closeness



BWLTR-Degree

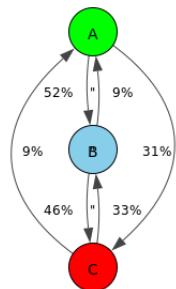


BWLTR-FWLTR

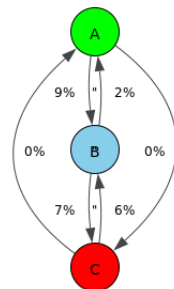


BWLTR-ICR

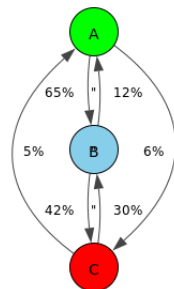
112



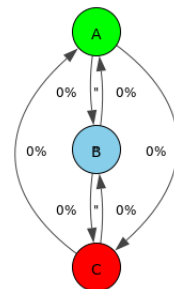
BWLTR-Katz



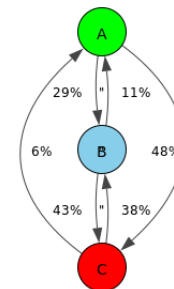
BWLTR-LTR



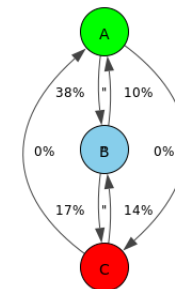
BWLTR-PageRank



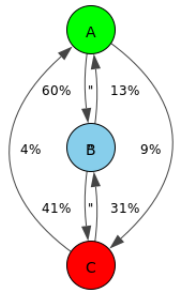
FWLTR-BWLTR



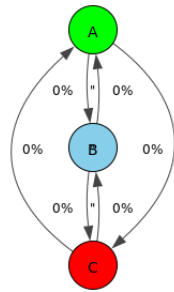
FWLTR-Betweenness



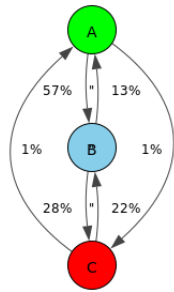
FWLTR-Closeness



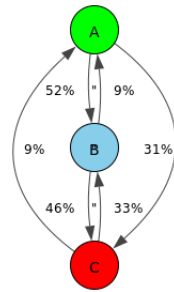
FWLTR-Degree



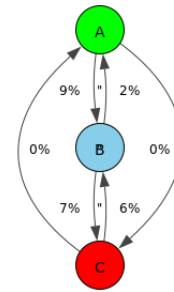
FWLTR-FWLTR



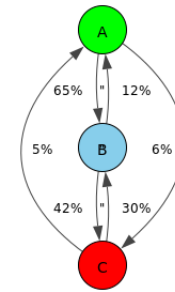
FWLTR-ICR



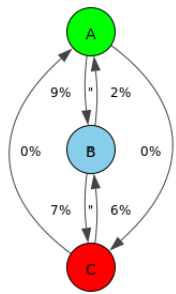
FWLTR-Katz



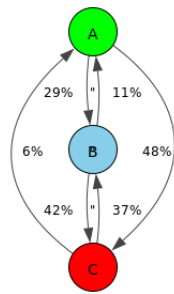
FWLTR-LTR



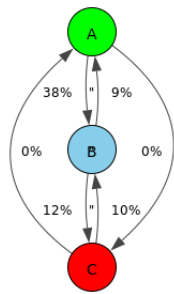
FWLTR-PageRank



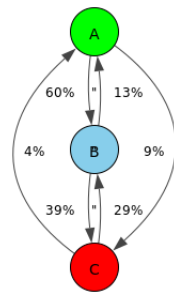
LTR-BWLTR



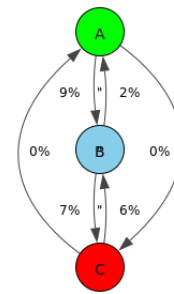
LTR-Betweenness



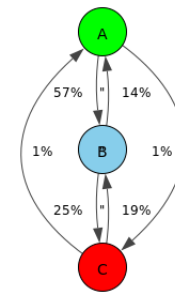
LTR-Closeness



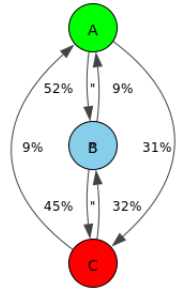
LTR-Degree



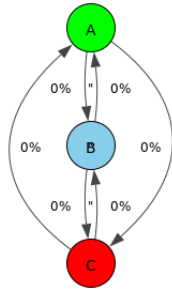
LTR-FWLTR



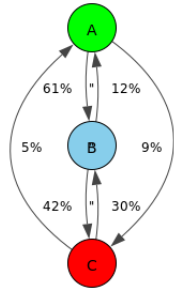
LTR-ICR



LTR-Katz



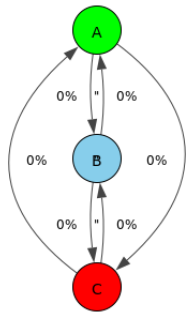
LTR-LTR



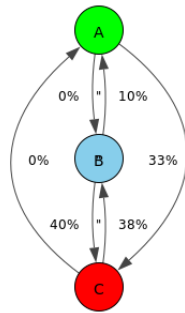
LTR-PageRank



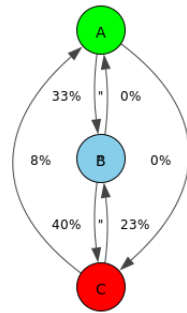
### B.1.3. Dining\_Table



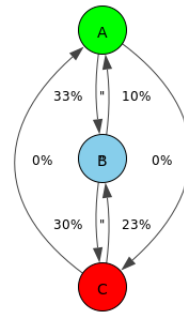
BWLTR-BWLTR



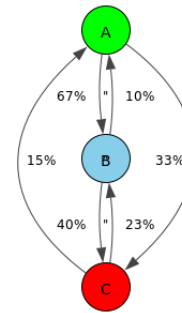
BWLTR-Betweenness



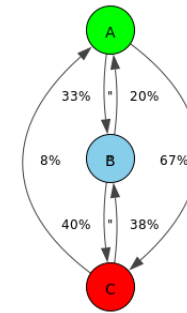
BWLTR-Closeness



BWLTR-Degree

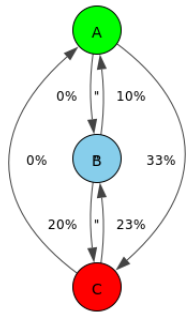


BWLTR-FWLTR

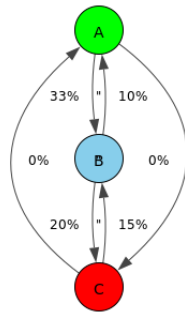


BWLTR-ICR

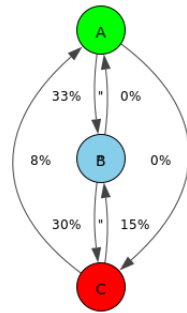
115



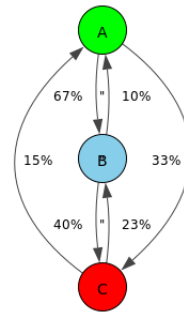
BWLTR-Katz



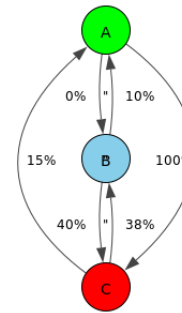
BWLTR-LTR



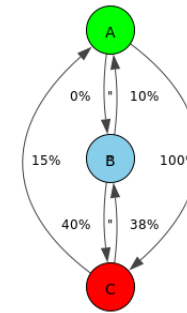
BWLTR-PageRank



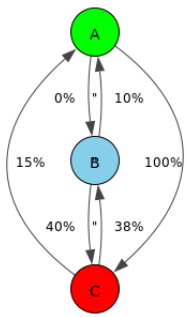
FWLTR-BWLTR



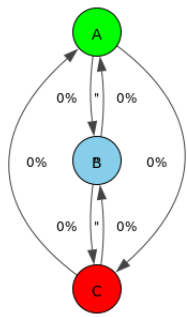
FWLTR-Betweenness



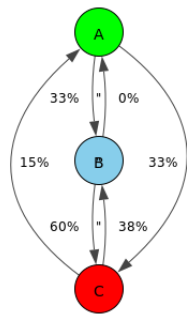
FWLTR-Closeness



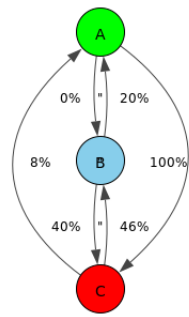
FWLTR-Degree



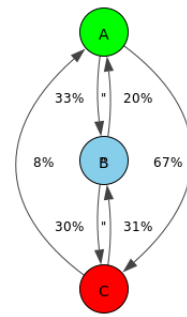
FWLTR-FWLTR



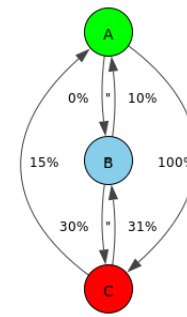
FWLTR-ICR



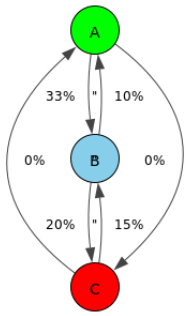
FWLTR-Katz



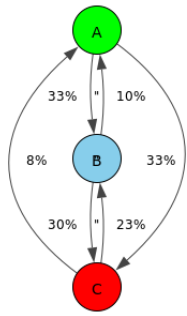
FWLTR-LTR



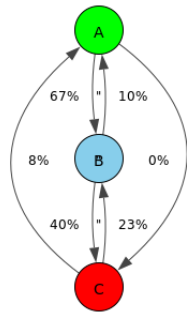
FWLTR-PageRank



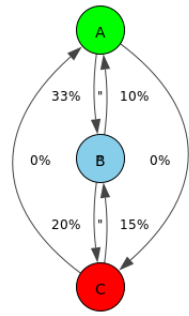
LTR-BWLTR



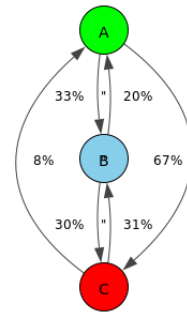
LTR-Betweenness



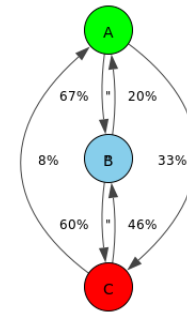
LTR-Closeness



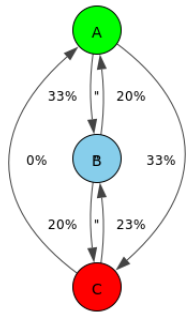
LTR-Degree



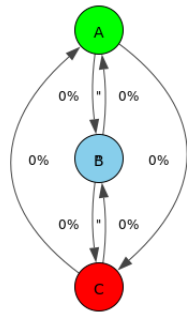
LTR-FWLTR



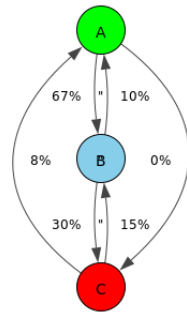
LTR-ICR



LTR-Katz

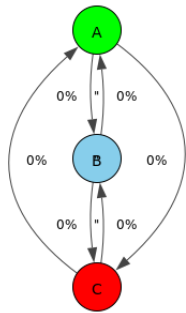


LTR-LTR

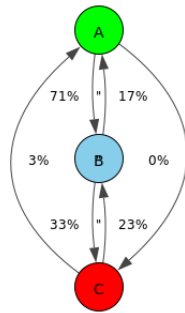


LTR-PageRank

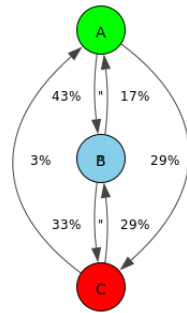
### B.1.4. Dolphins



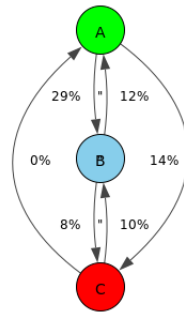
BWLTR-BWLTR



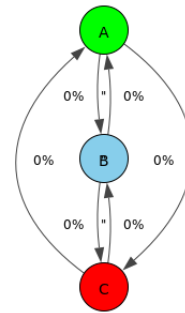
BWLTR-Betweenness



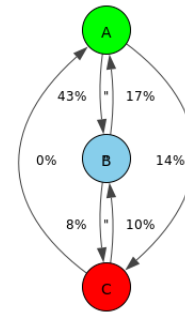
BWLTR-Closeness



BWLTR-Degree

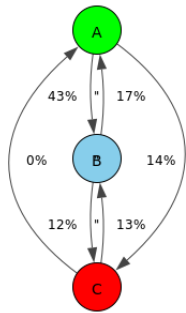


BWLTR-FWLTR

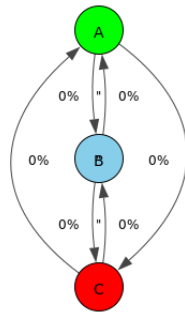


BWLTR-ICR

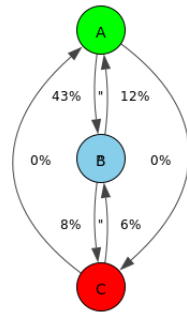
118



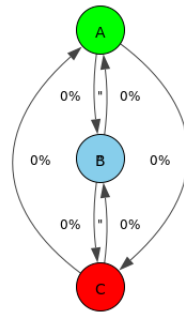
BWLTR-Katz



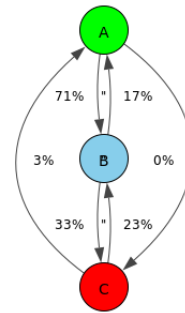
BWLTR-LTR



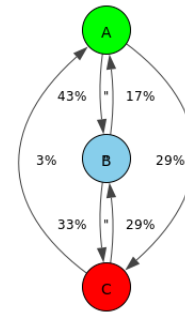
BWLTR-PageRank



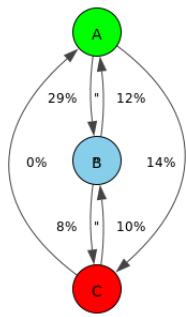
FWLTR-BWLTR



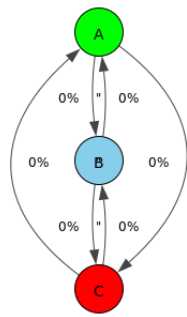
FWLTR-Betweenness



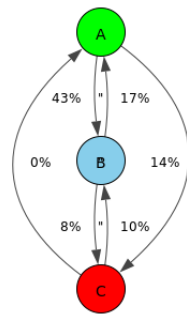
FWLTR-Closeness



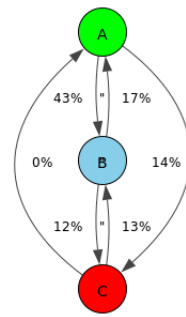
FWLTR-Degree



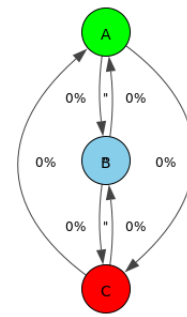
FWLTR-FWLTR



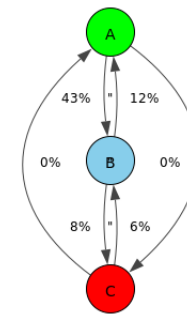
FWLTR-ICR



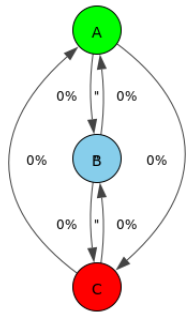
FWLTR-Katz



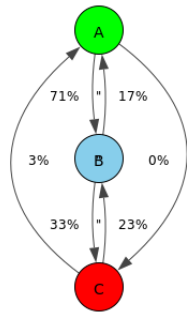
FWLTR-LTR



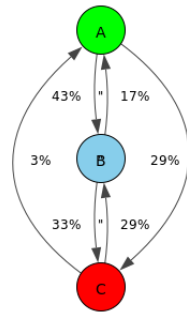
FWLTR-PageRank



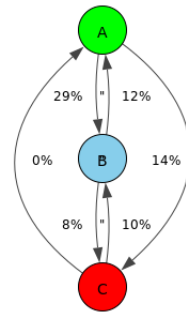
LTR-BWLTR



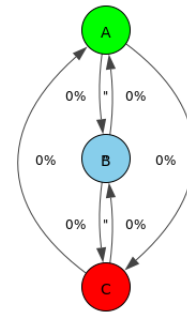
LTR-Betweenness



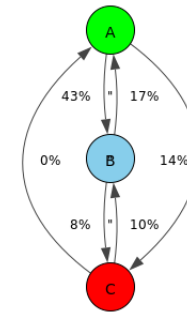
LTR-Closeness



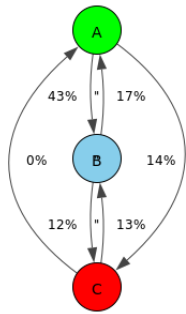
LTR-Degree



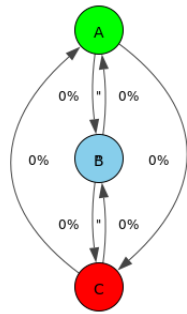
LTR-FWLTR



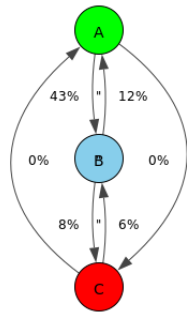
LTR-ICR



LTR-Katz

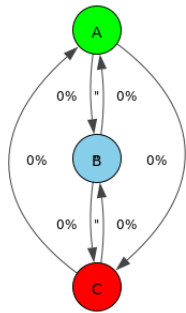


LTR-LTR

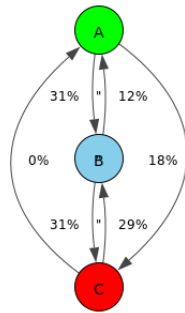


LTR-PageRank

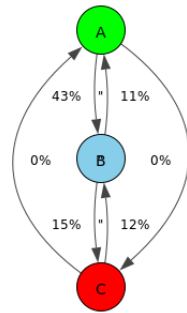
### B.1.5. Epinions



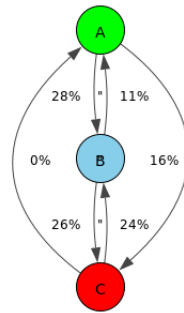
BWLTR-BWLTR



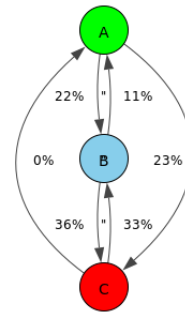
BWLTR-Betweenness



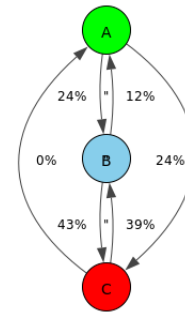
BWLTR-Closeness



BWLTR-Degree

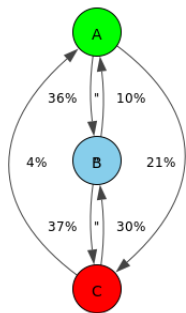


BWLTR-FWLTR

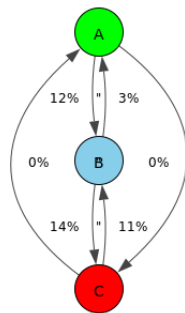


BWLTR-ICR

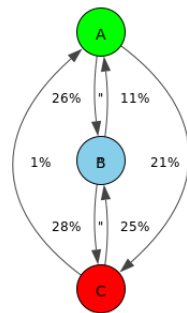
121



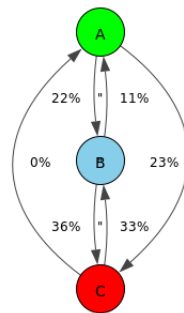
BWLTR-Katz



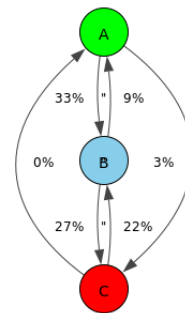
BWLTR-LTR



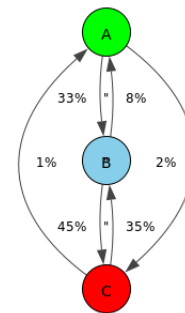
BWLTR-PageRank



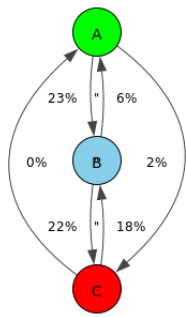
FWLTR-BWLTR



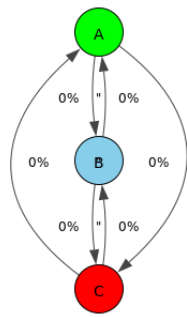
FWLTR-Betweenness



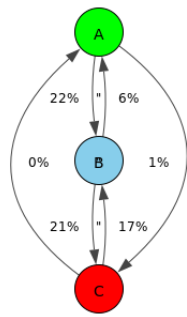
FWLTR-Closeness



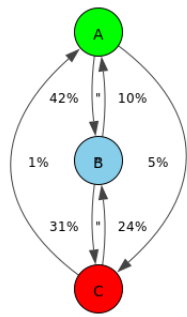
FWLTR-Degree



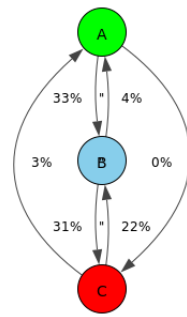
FWLTR-FWLTR



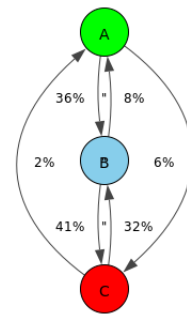
FWLTR-ICR



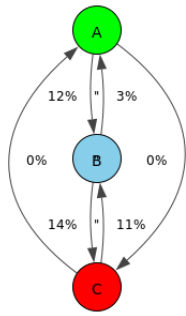
FWLTR-Katz



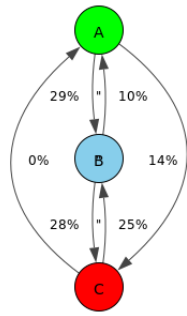
FWLTR-LTR



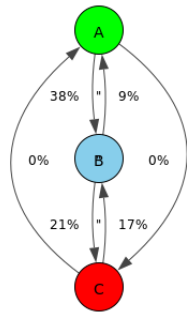
FWLTR-PageRank



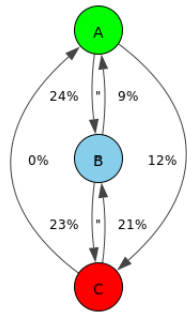
LTR-BWLTR



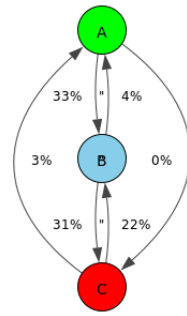
LTR-Betweenness



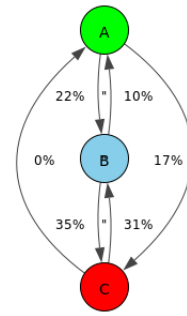
LTR-Closeness



LTR-Degree

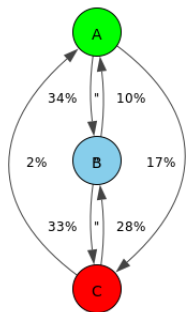


LTR-FWLTR

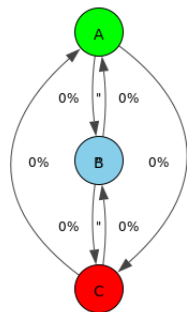


LTR-ICR

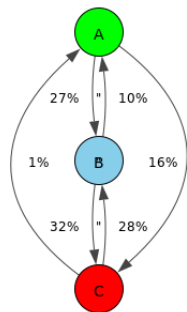




LTR-Katz

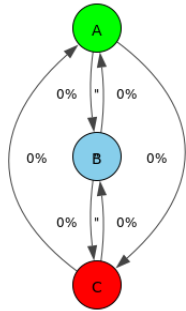


LTR-LTR

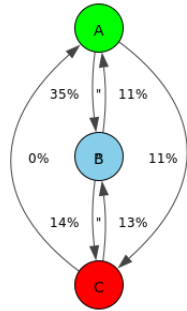


LTR-PageRank

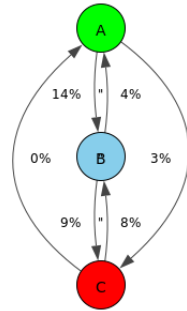
### B.1.6. Higgs



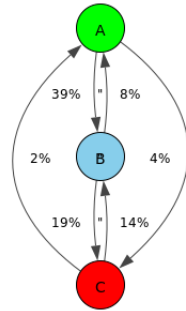
BWLTR-BWLTR



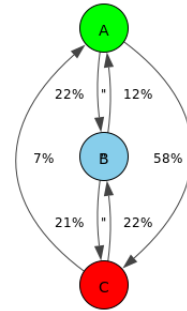
BWLTR-Betweenness



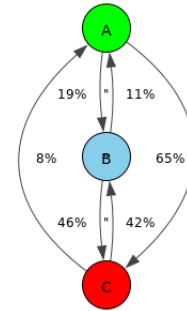
BWLTR-Closeness



BWLTR-Degree

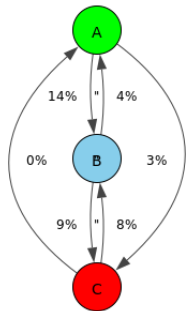


BWLTR-FWLTR

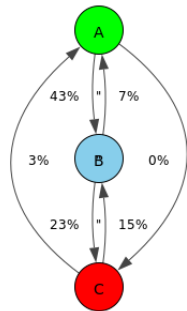


BWLTR-ICR

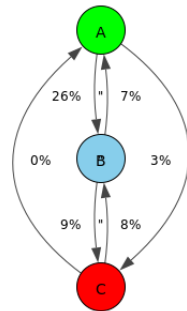
124



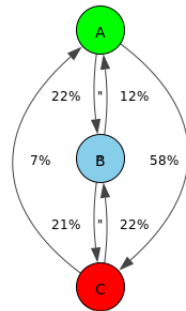
BWLTR-Katz



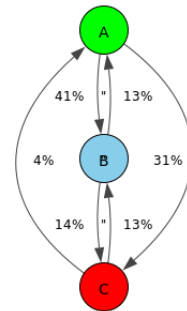
BWLTR-LTR



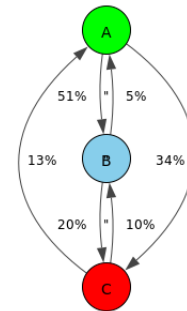
BWLTR-PageRank



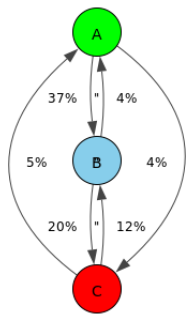
FWLTR-BWLTR



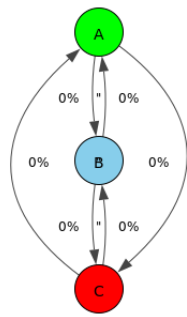
FWLTR-Betweenness



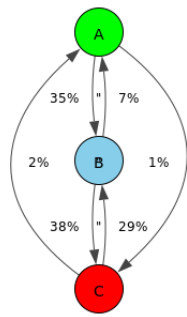
FWLTR-Closeness



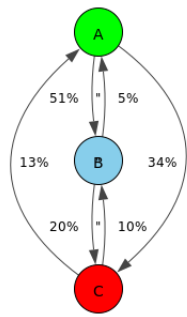
FWLTR-Degree



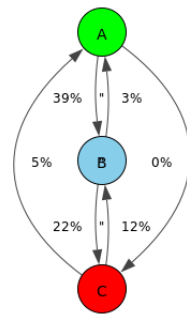
FWLTR-FWLTR



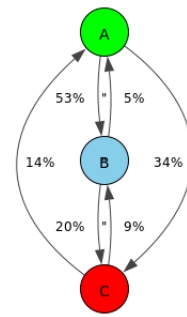
FWLTR-ICR



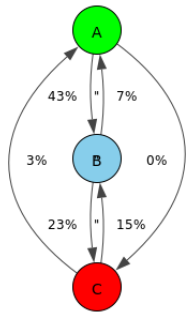
FWLTR-Katz



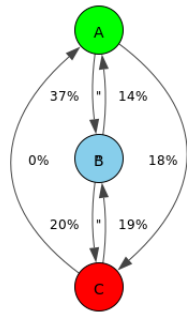
FWLTR-LTR



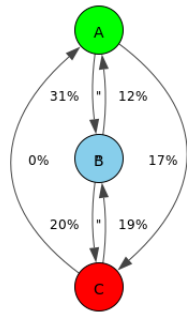
FWLTR-PageRank



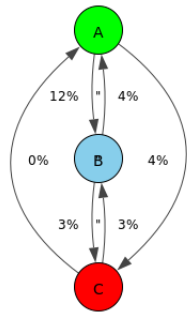
LTR-BWLTR



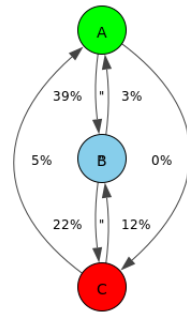
LTR-Betweenness



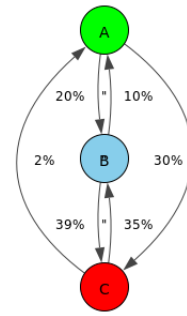
LTR-Closeness



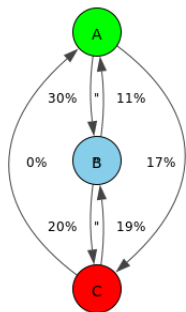
LTR-Degree



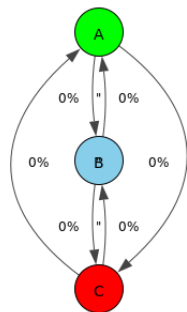
LTR-FWLTR



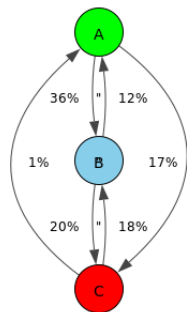
LTR-ICR



LTR-Katz

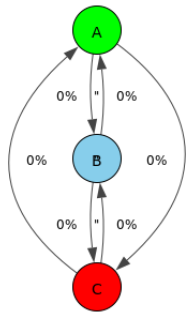


LTR-LTR

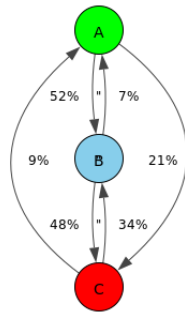


LTR-PageRank

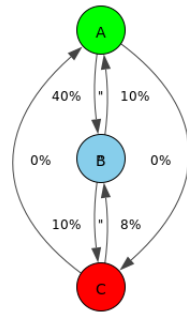
### B.1.7. Human\_Brain



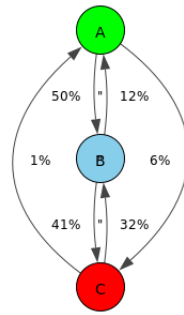
BWLTR-BWLTR



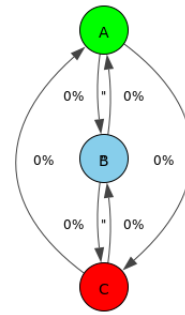
BWLTR-Betweenness



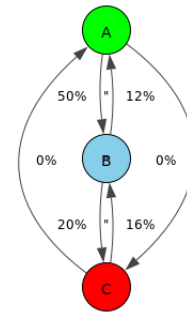
BWLTR-Closeness



BWLTR-Degree

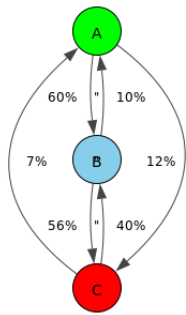


BWLTR-FWLTR

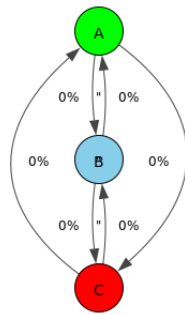


BWLTR-ICR

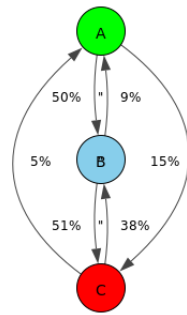
127



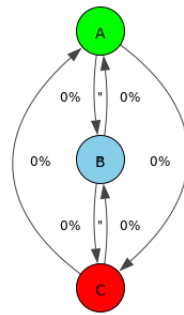
BWLTR-Katz



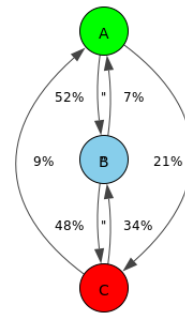
BWLTR-LTR



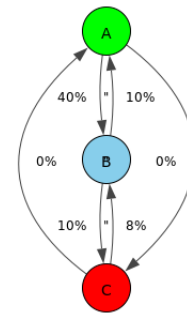
BWLTR-PageRank



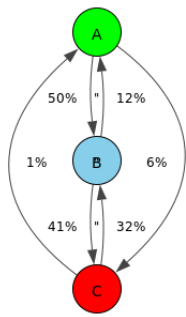
FWLTR-BWLTR



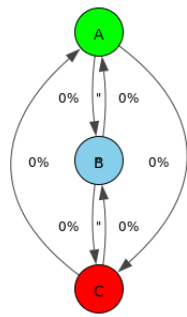
FWLTR-Betweenness



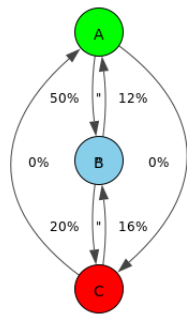
FWLTR-Closeness



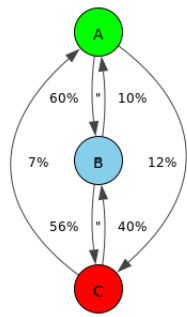
FWLTR-Degree



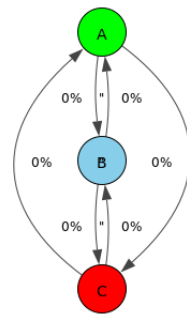
FWLTR-FWLTR



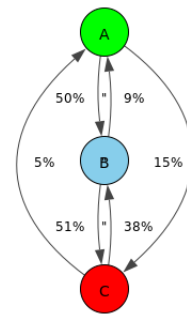
FWLTR-ICR



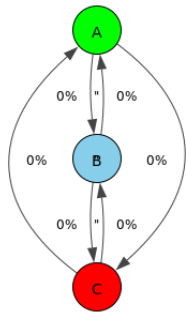
FWLTR-Katz



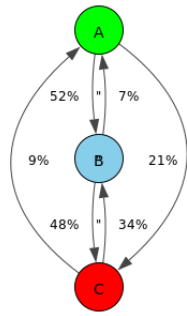
FWLTR-LTR



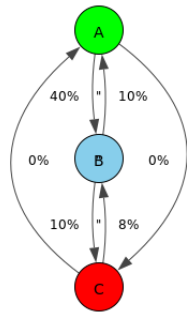
FWLTR-PageRank



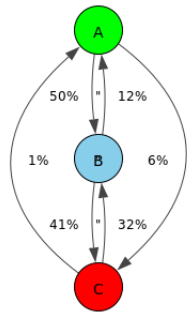
LTR-BWLTR



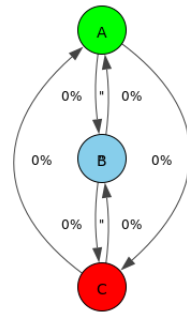
LTR-Betweenness



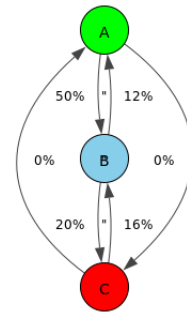
LTR-Closeness



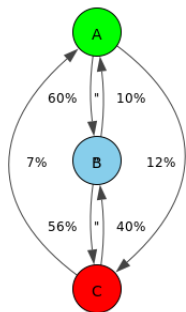
LTR-Degree



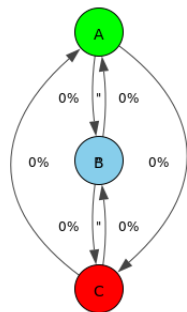
LTR-FWLTR



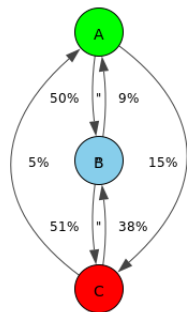
LTR-ICR



LTR-Katz

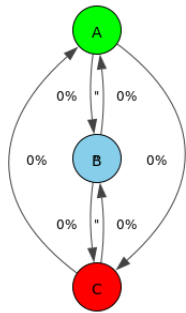


LTR-LTR

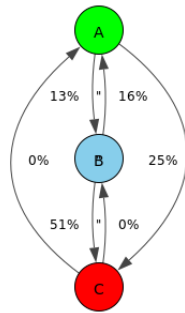


LTR-PageRank

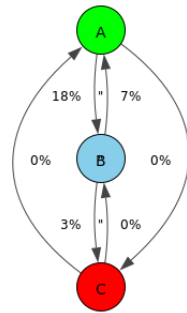
### B.1.8. Wikipedia



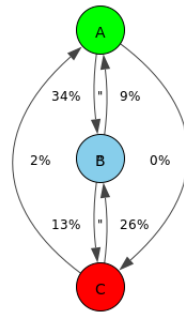
BWLTR-BWLTR



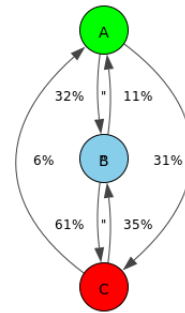
BWLTR-Betweenness



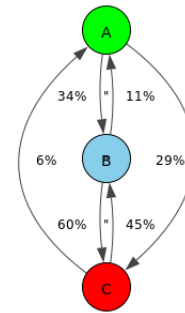
BWLTR-Closeness



BWLTR-Degree

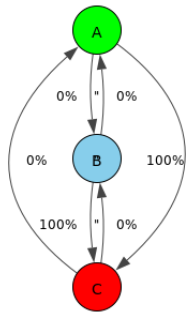


BWLTR-FWLTR

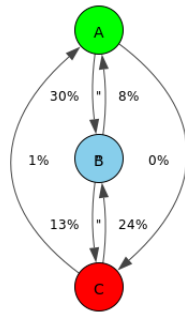


BWLTR-ICR

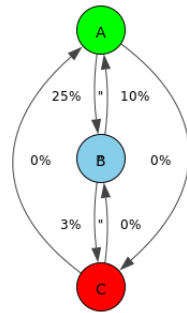
130



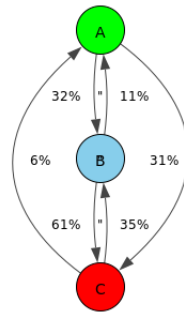
BWLTR-Katz



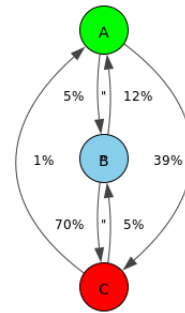
BWLTR-LTR



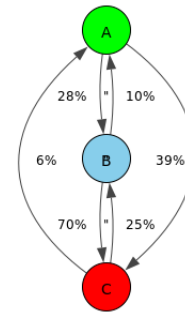
BWLTR-PageRank



FWLTR-BWLTR

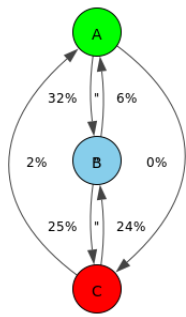


FWLTR-Betweenness

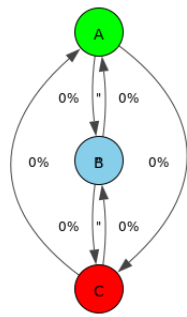


FWLTR-Closeness

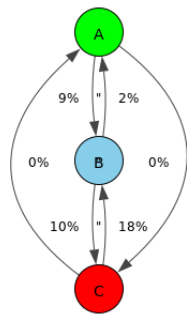




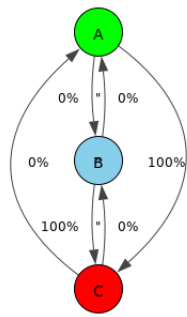
FWLTR-Degree



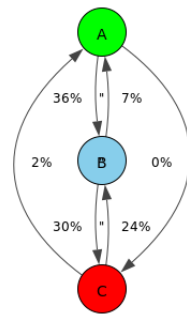
FWLTR-FWLTR



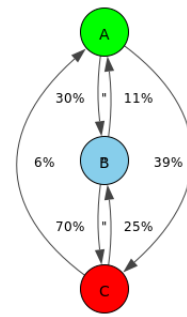
FWLTR-ICR



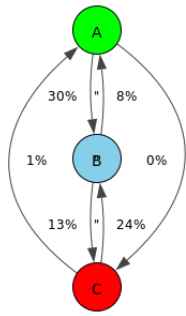
FWLTR-Katz



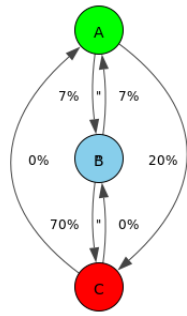
FWLTR-LTR



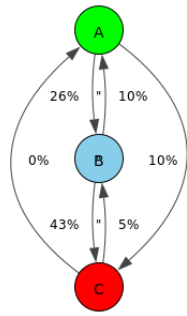
FWLTR-PageRank



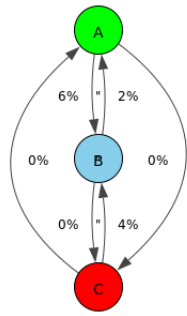
LTR-BWLTR



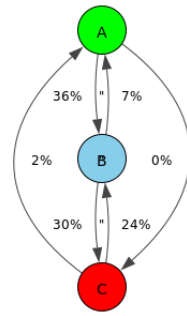
LTR-Betweenness



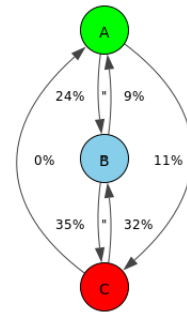
LTR-Closeness



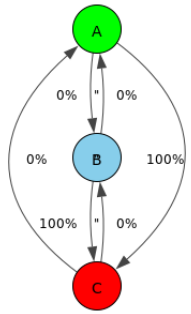
LTR-Degree



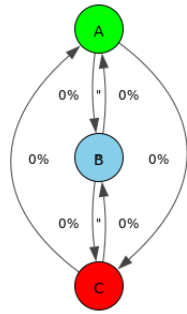
LTR-FWLTR



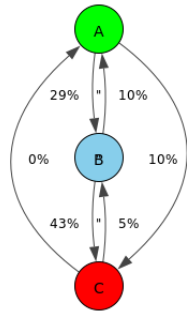
LTR-ICR



LTR-Katz



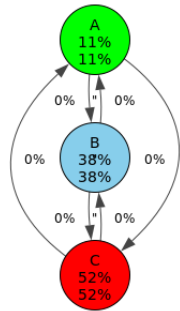
LTR-LTR



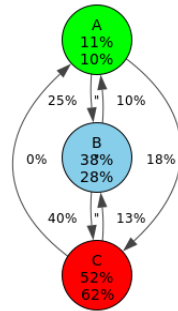
LTR-PageRank

## B.2. Clasificación por valor

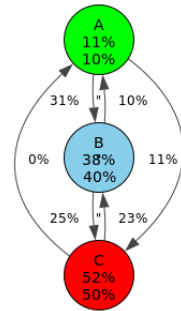
### B.2.1. ArXiv



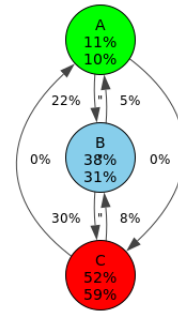
BWLTR-BWLTR



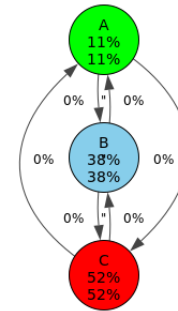
BWLTR-Betweenness



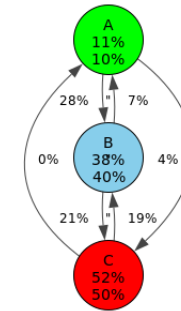
BWLTR-Closeness



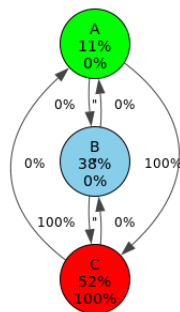
BWLTR-Degree



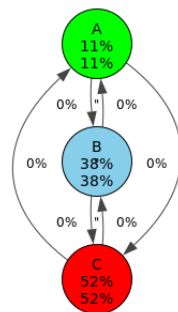
BWLTR-FWLTR



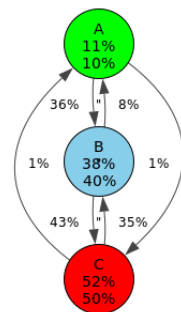
BWLTR-ICR



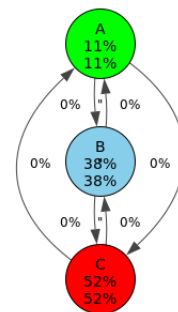
BWLTR-Katz



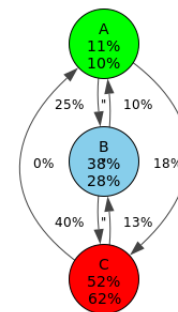
BWLTR-LTR



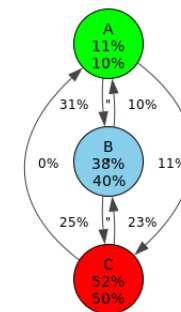
BWLTR-PageRank



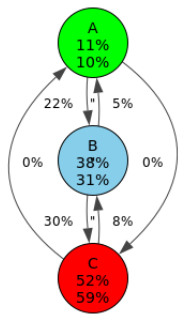
FWLTR-BWLTR



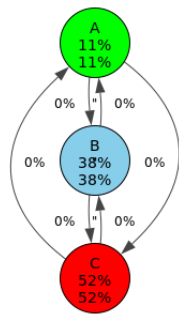
FWLTR-Betweenness



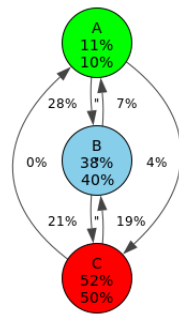
FWLTR-Closeness



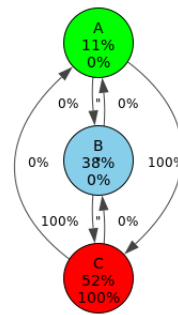
FWLTR-Degree



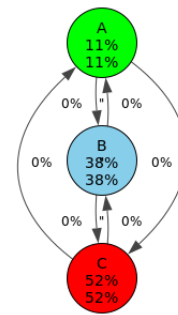
FWLTR-FWLTR



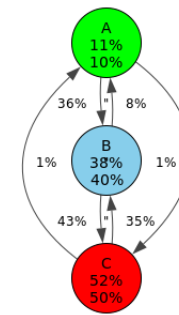
FWLTR-ICR



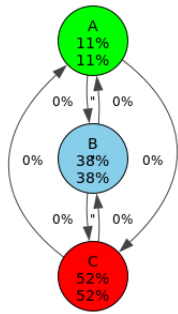
FWLTR-Katz



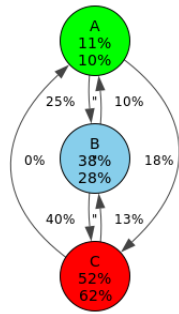
FWLTR-LTR



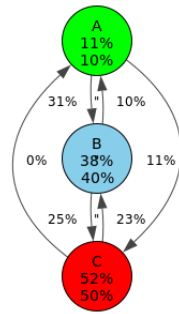
FWLTR-PageRank



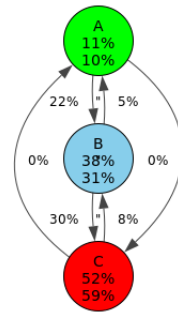
LTR-BWLTR



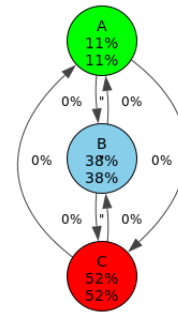
LTR-Betweenness



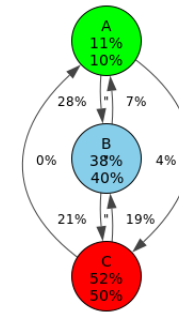
LTR-Closeness



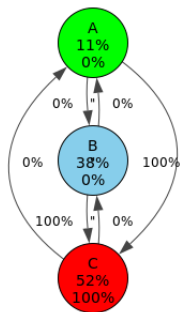
LTR-Degree



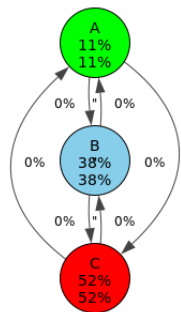
LTR-FWLTR



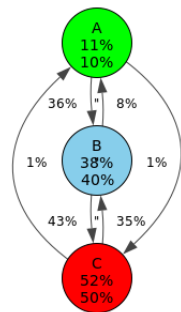
LTR-ICR



LTR-Katz

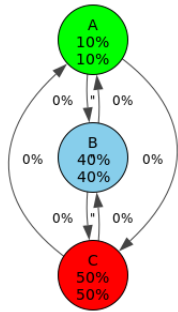


LTR-LTR

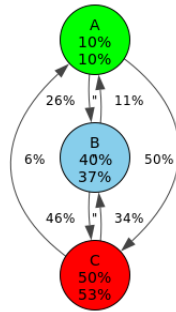


LTR-PageRank

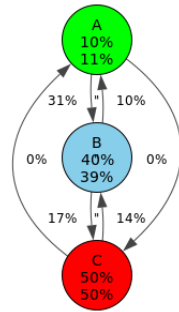
## B.2.2. Caida



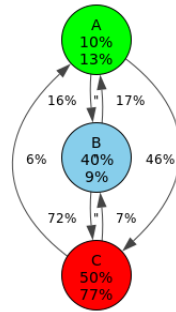
BWLTR-BWLTR



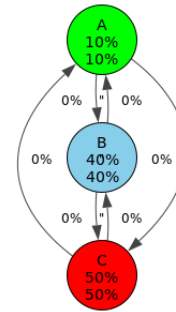
BWLTR-Betweenness



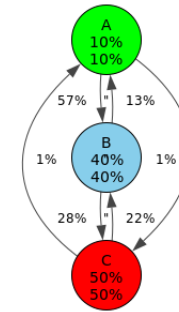
BWLTR-Closeness



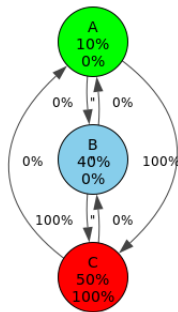
BWLTR-Degree



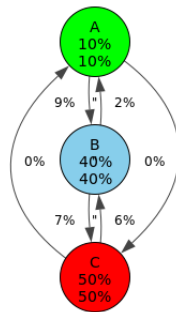
BWLTR-FWLTR



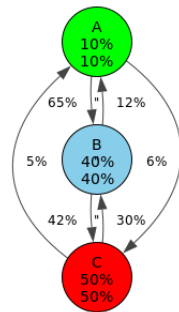
BWLTR-ICR



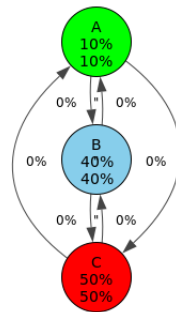
BWLTR-Katz



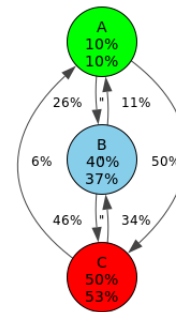
BWLTR-LTR



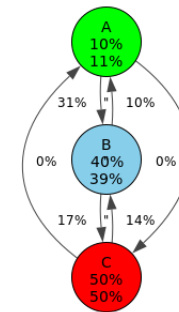
BWLTR-PageRank



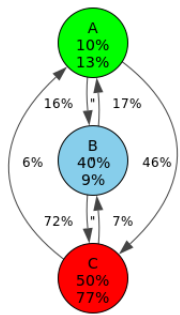
FWLTR-BWLTR



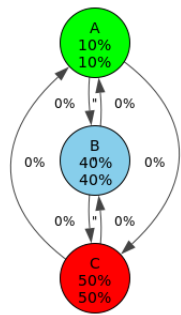
FWLTR-Betweenness



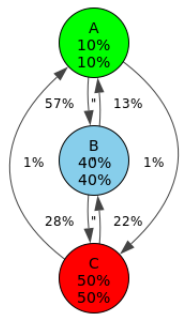
FWLTR-Closeness



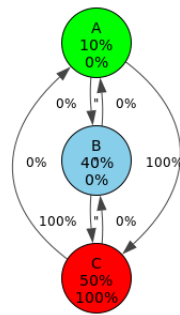
FWLTR-Degree



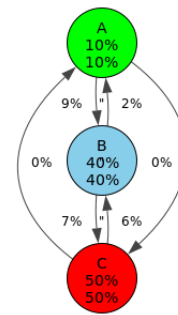
FWLTR-FWLTR



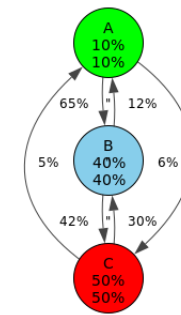
FWLTR-ICR



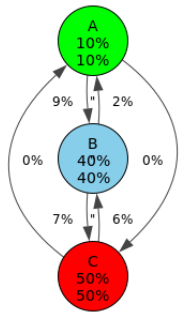
FWLTR-Katz



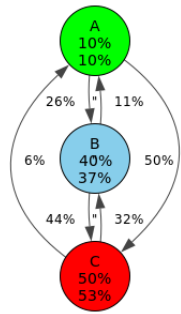
FWLTR-LTR



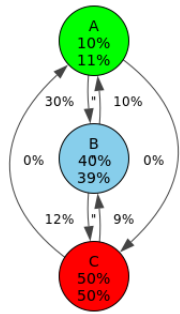
FWLTR-PageRank



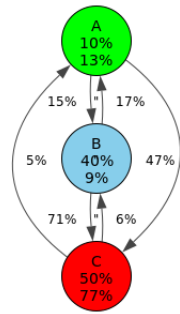
LTR-BWLTR



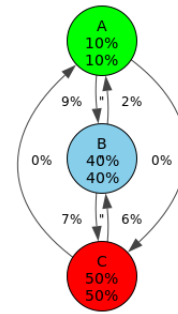
LTR-Betweenness



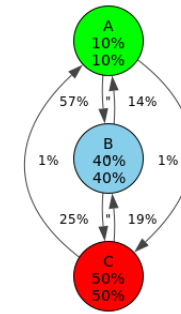
LTR-Closeness



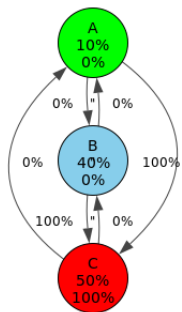
LTR-Degree



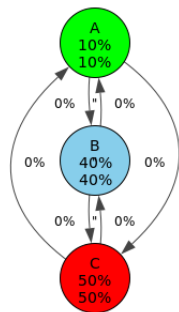
LTR-FWLTR



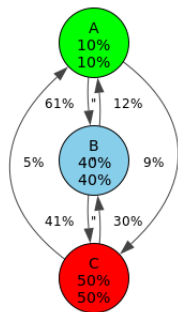
LTR-ICR



LTR-Katz



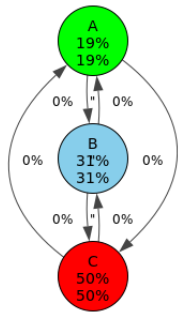
LTR-LTR



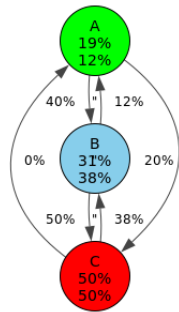
LTR-PageRank



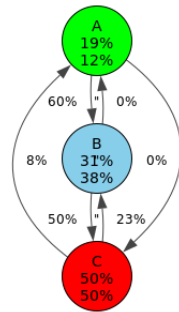
### B.2.3. Dining\_Table



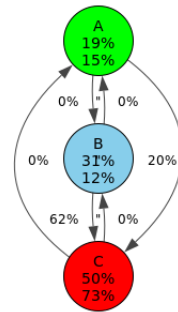
BWLTR-BWLTR



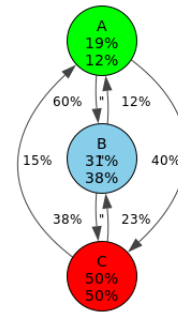
BWLTR-Betweenness



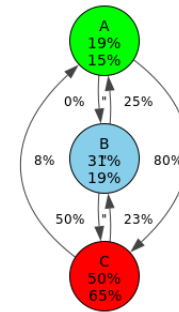
BWLTR-Closeness



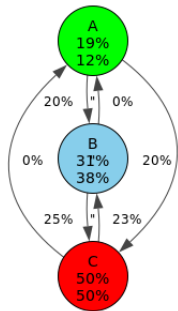
BWLTR-Degree



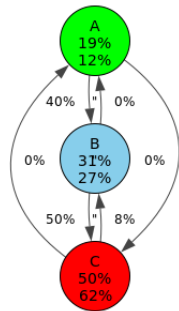
BWLTR-FWLTR



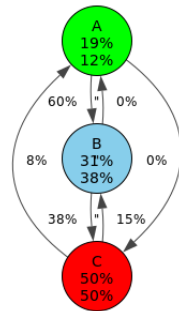
BWLTR-ICR



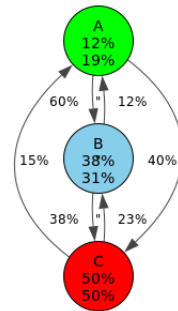
BWLTR-Katz



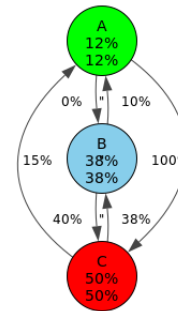
BWLTR-LTR



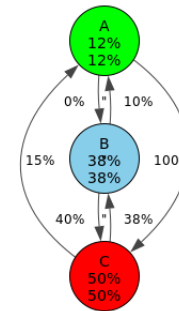
BWLTR-PageRank



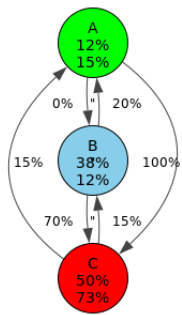
FWLTR-BWLTR



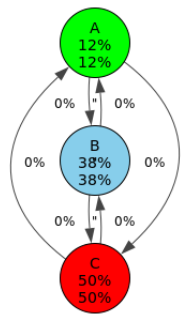
FWLTR-Betweenness



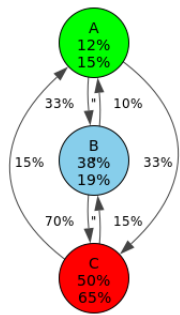
FWLTR-Closeness



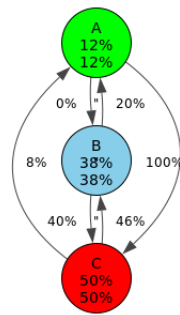
FWLTR-Degree



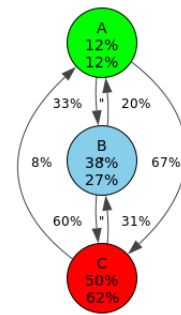
FWLTR-FWLTR



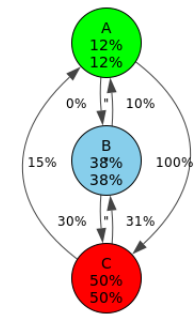
FWLTR-ICR



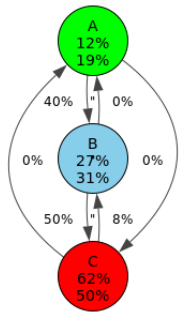
FWLTR-Katz



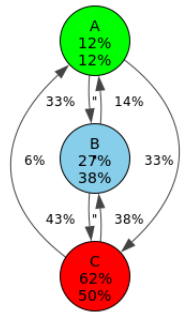
FWLTR-LTR



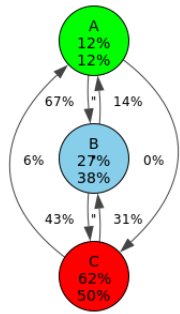
FWLTR-PageRank



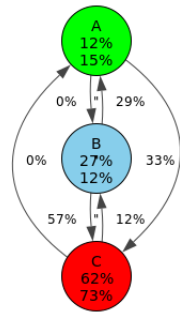
LTR-BWLTR



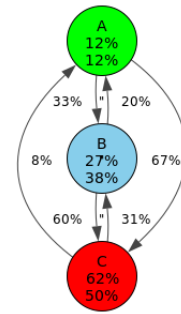
LTR-Betweenness



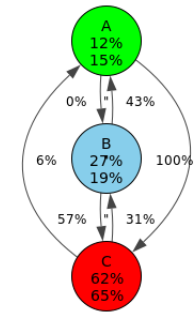
LTR-Closeness



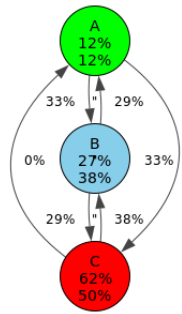
LTR-Degree



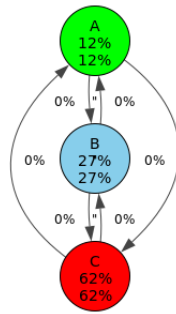
LTR-FWLTR



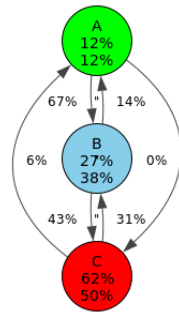
LTR-ICR



LTR-Katz

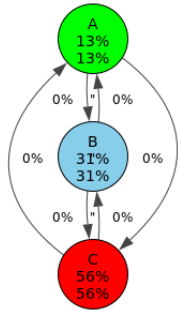


LTR-LTR

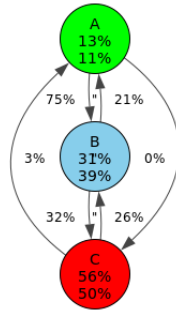


LTR-PageRank

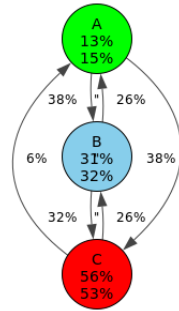
## B.2.4. Dolphins



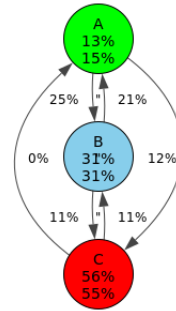
BWLTR-BWLTR



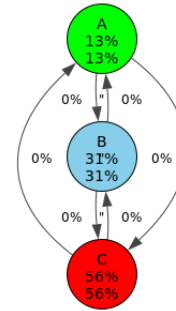
BWLTR-Betweenness



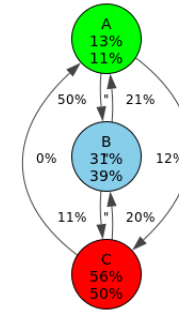
BWLTR-Closeness



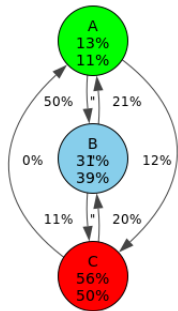
BWLTR-Degree



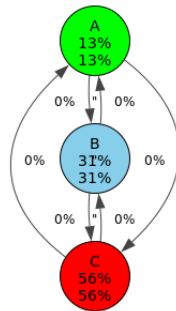
BWLTR-FWLTR



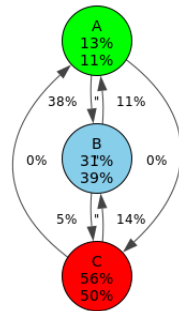
BWLTR-ICR



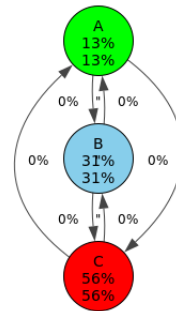
BWLTR-Katz



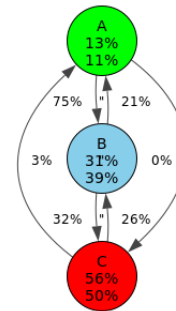
BWLTR-LTR



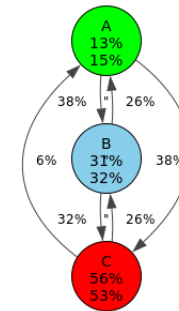
BWLTR-PageRank



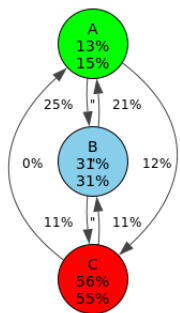
FWLTR-BWLTR



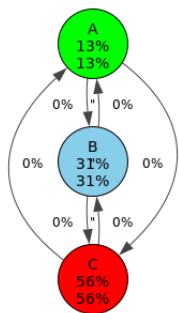
FWLTR-Betweenness



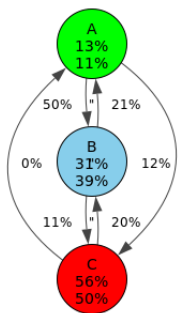
FWLTR-Closeness



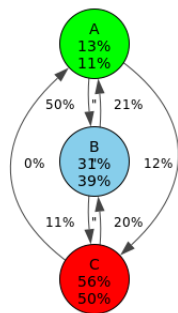
FWLTR-Degree



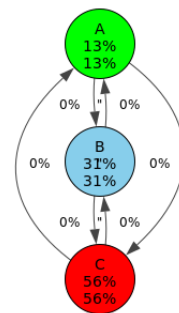
FWLTR-FWLTR



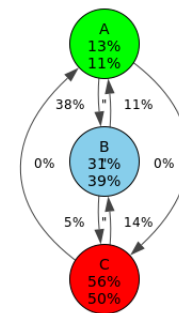
FWLTR-ICR



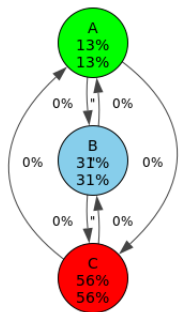
FWLTR-Katz



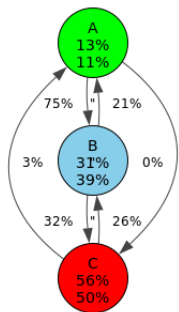
FWLTR-LTR



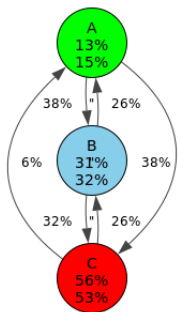
FWLTR-PageRank



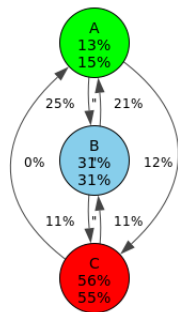
LTR-BWLTR



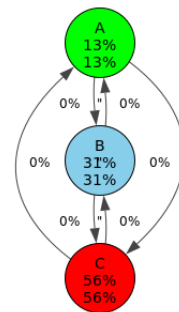
LTR-Betweenness



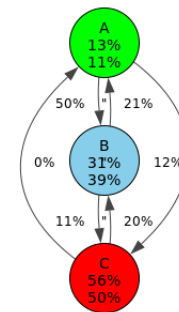
LTR-Closeness



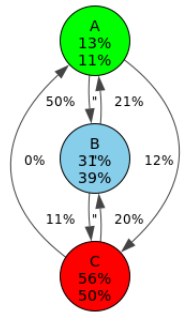
LTR-Degree



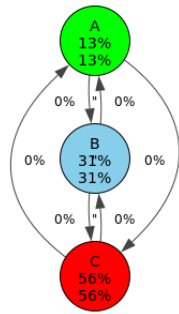
LTR-FWLTR



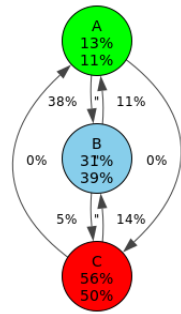
LTR-ICR



LTR-Katz

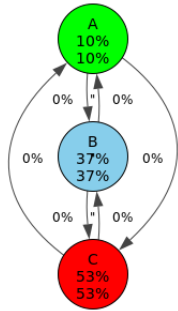


LTR-LTR

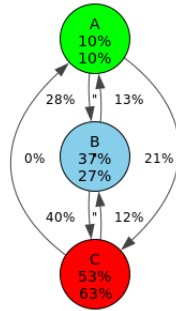


LTR-PageRank

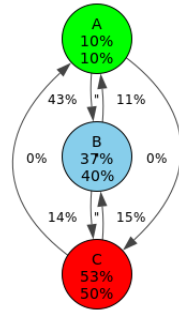
## B.2.5. Epinions



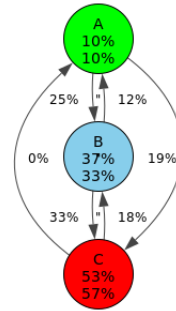
BWLTR-BWLTR



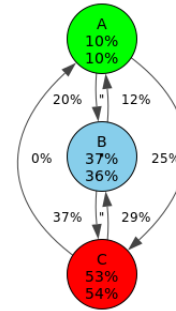
BWLTR-Betweenness



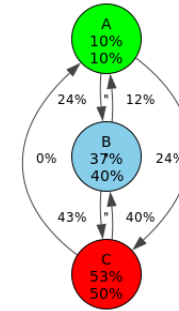
BWLTR-Closeness



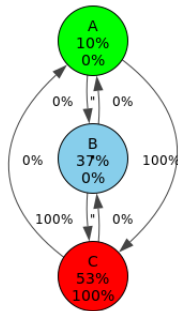
BWLTR-Degree



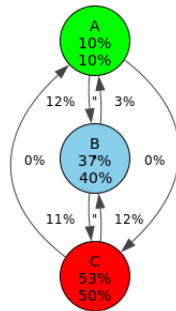
BWLTR-FWLTR



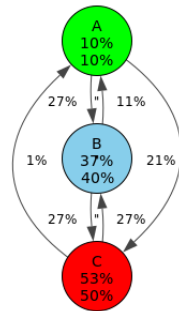
BWLTR-ICR



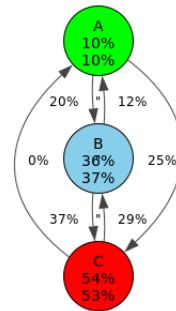
BWLTR-Katz



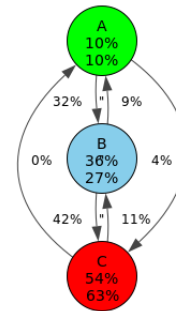
BWLTR-LTR



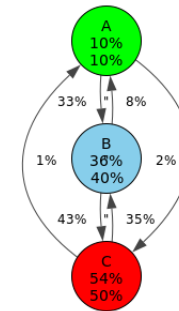
BWLTR-Pagerank



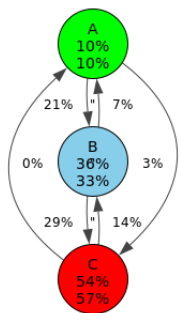
FWLTR-BWLTR



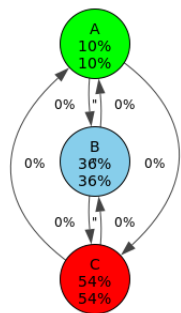
FWLTR-Betweenness



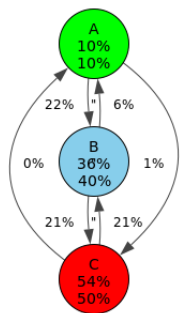
FWLTR-Closeness



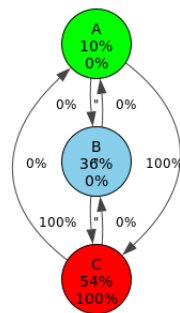
FWLTR-Degree



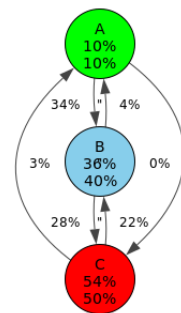
FWLTR-FWLTR



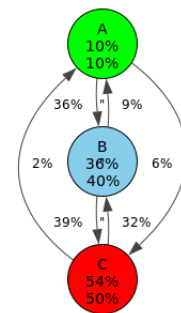
FWLTR-ICR



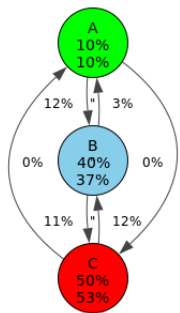
FWLTR-Katz



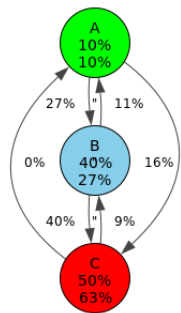
FWLTR-LTR



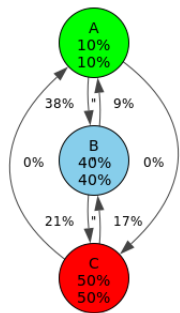
FWLTR-PageRank



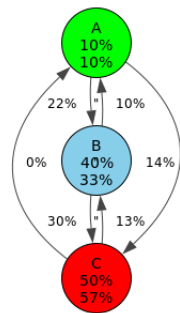
LTR-BWLTR



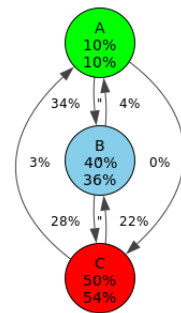
LTR-Betweenness



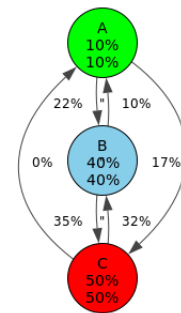
LTR-Closeness



LTR-Degree

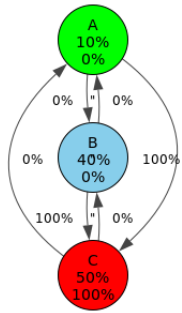


LTR-FWLTR

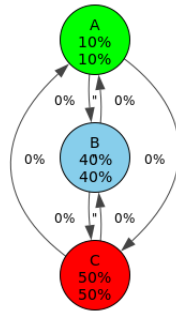


LTR-ICR

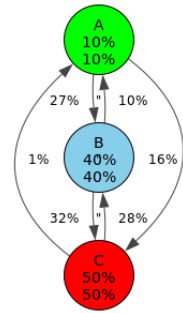




LTR-Katz

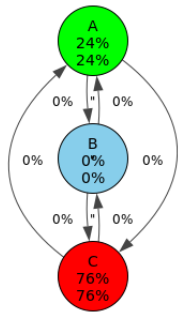


LTR-LTR

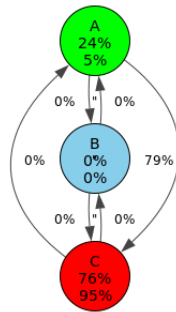


LTR-PageRank

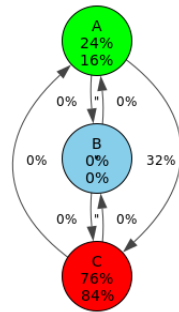
## B.2.6. Higgs



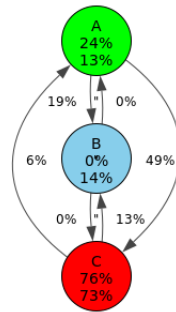
BWLTR-BWLTR



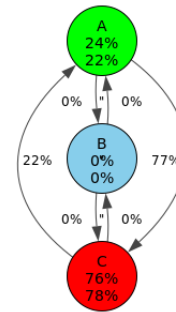
BWLTR-Betweenness



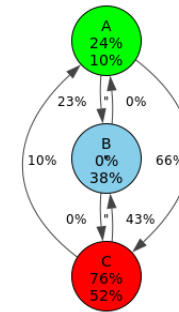
BWLTR-Closeness



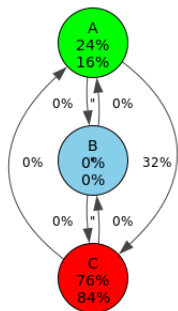
BWLTR-Degree



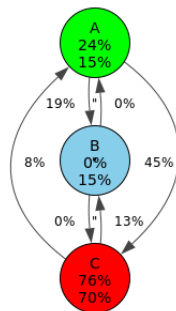
BWLTR-FWLTR



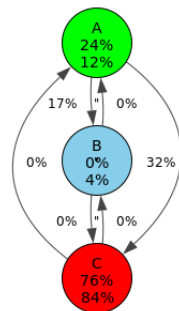
BWLTR-ICR



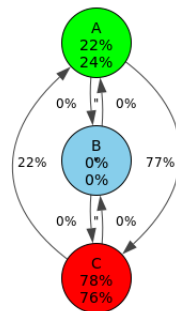
BWLTR-Katz



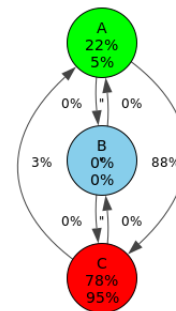
BWLTR-LTR



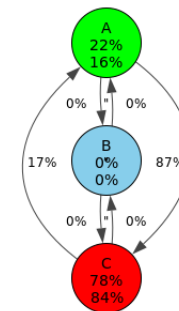
BWLTR-PageRank



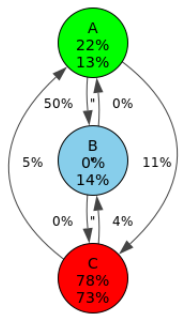
FWLTR-BWLTR



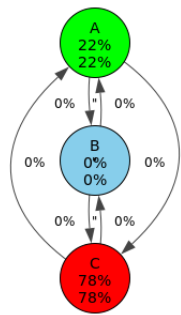
FWLTR-Betweenness



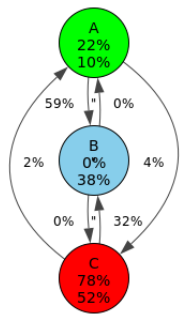
FWLTR-Closeness



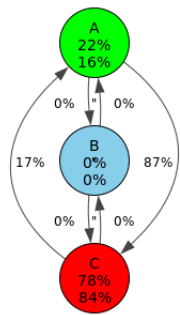
FWLTR-Degree



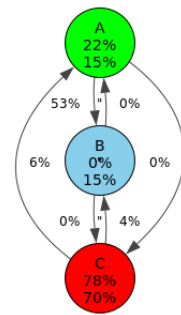
FWLTR-FWLTR



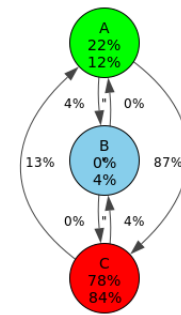
FWLTR-ICR



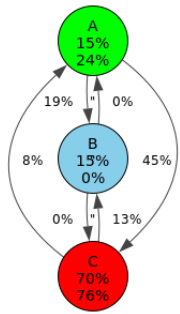
FWLTR-Katz



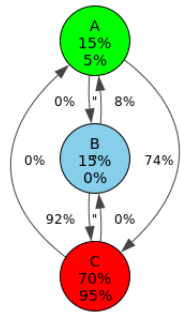
FWLTR-LTR



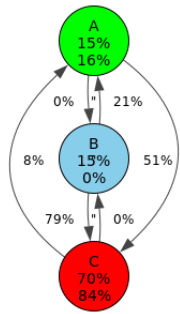
FWLTR-PageRank



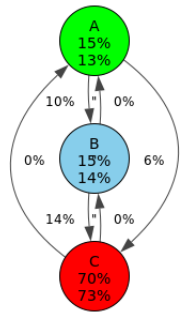
LTR-BWLTR



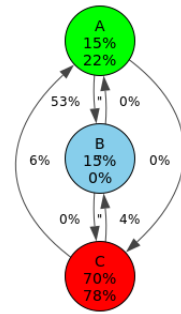
LTR-Betweenness



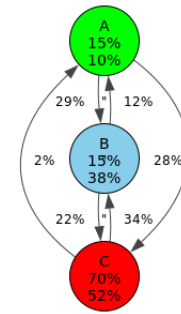
LTR-Closeness



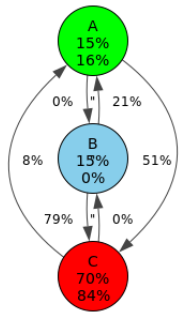
LTR-Degree



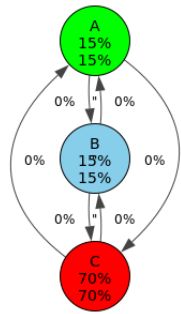
LTR-FWLTR



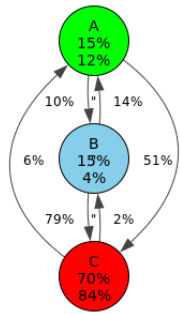
LTR-ICR



LTR-Katz

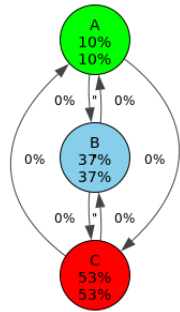


LTR-LTR

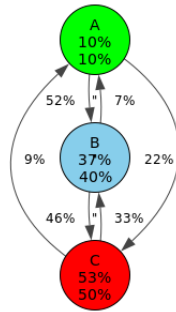


LTR-PageRank

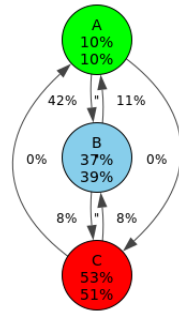
## B.2.7. Human\_Brain



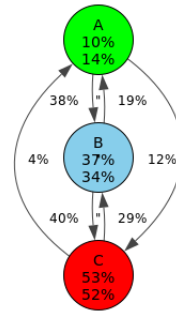
BWLTR-BWLTR



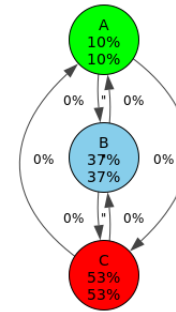
BWLTR-Betweenness



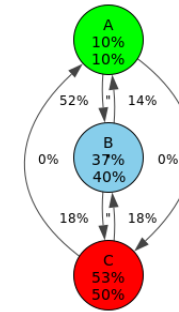
BWLTR-Closeness



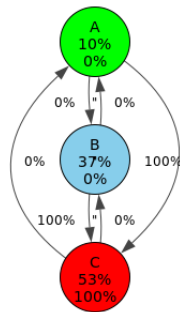
BWLTR-Degree



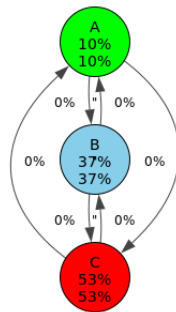
BWLTR-FWLTR



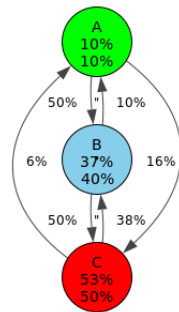
BWLTR-ICR



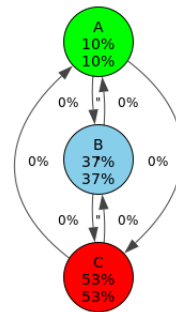
BWLTR-Katz



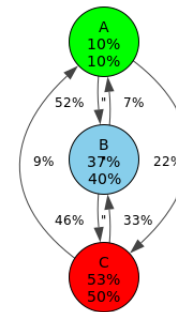
BWLTR-LTR



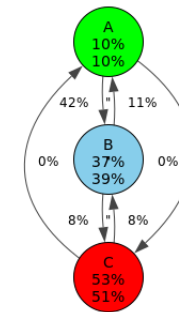
BWLTR-PageRank



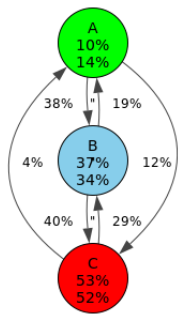
FWLTR-BWLTR



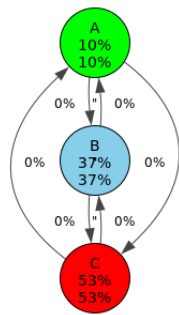
FWLTR-Betweenness



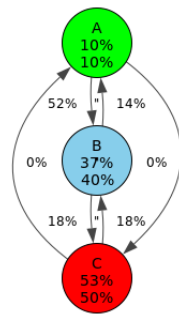
FWLTR-Closeness



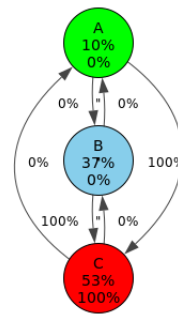
FWLTR-Degree



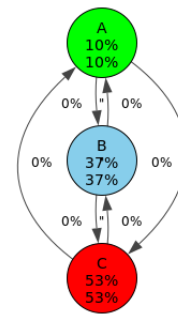
FWLTR-FWLTR



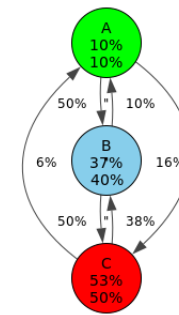
FWLTR-ICR



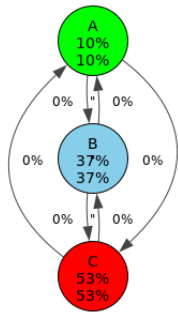
FWLTR-Katz



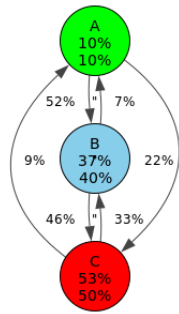
FWLTR-LTR



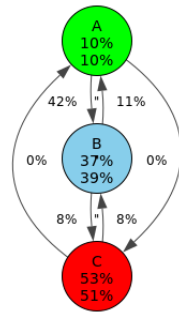
FWLTR-PageRank



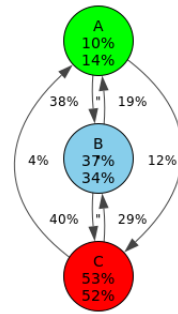
LTR-BWLTR



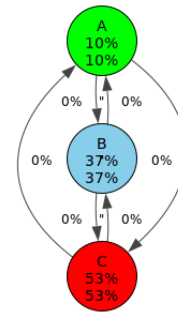
LTR-Betweenness



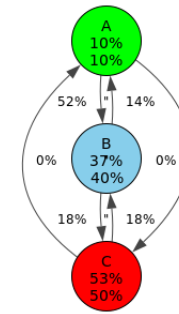
LTR-Closeness



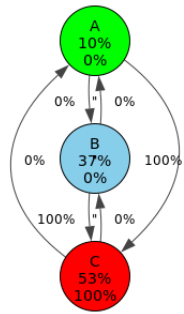
LTR-Degree



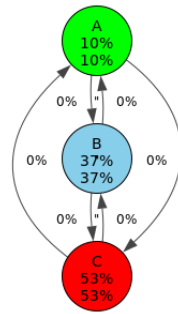
LTR-FWLTR



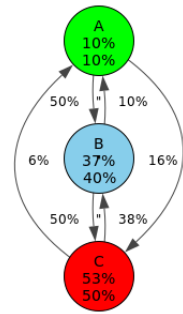
LTR-ICR



LTR-Katz

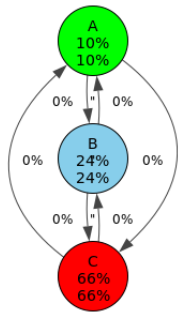


LTR-LTR

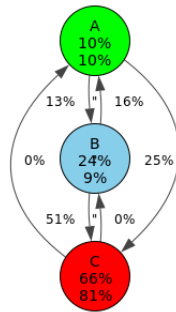


LTR-PageRank

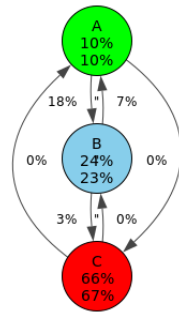
## B.2.8. Wikipedia



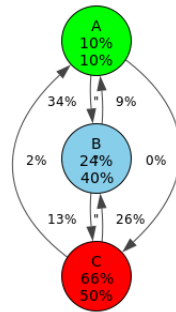
BWLTR-BWLTR



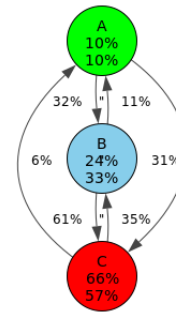
BWLTR-Betweenness



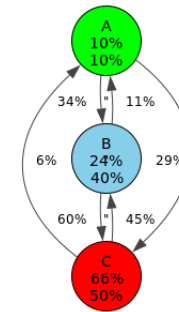
BWLTR-Closeness



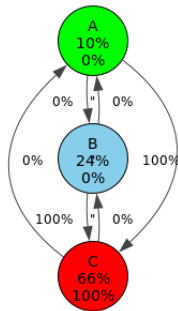
BWLTR-Degree



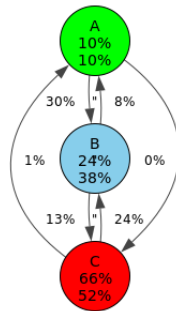
BWLTR-FWLTR



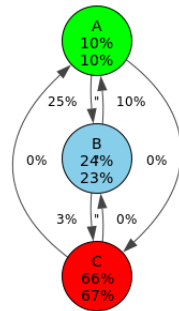
BWLTR-ICR



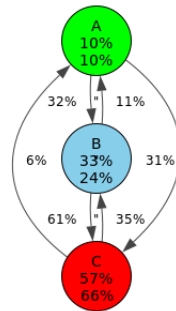
BWLTR-Katz



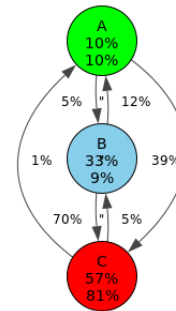
BWLTR-LTR



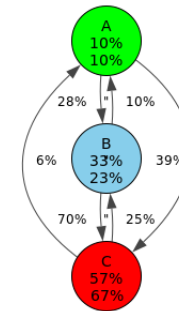
BWLTR-Pagerank



FWLTR-BWLTR

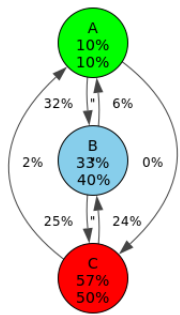


FWLTR-Betweenness

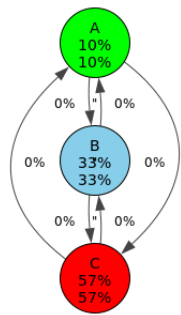


FWLTR-Closeness

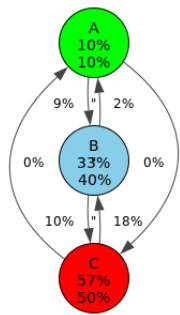




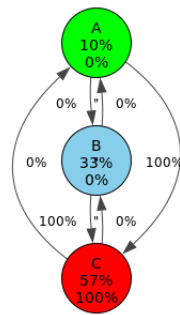
FWLTR-Degree



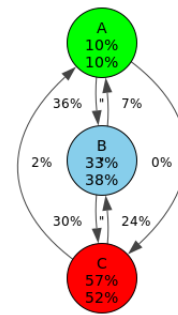
FWLTR-FWLTR



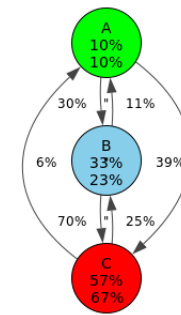
FWLTR-ICR



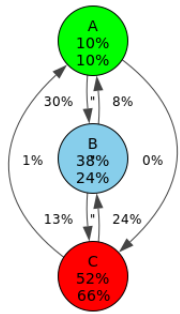
FWLTR-Katz



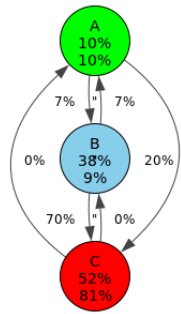
FWLTR-LTR



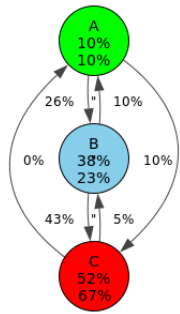
FWLTR-PageRank



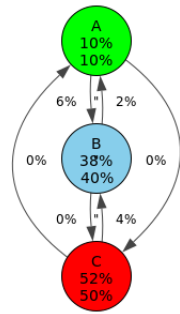
LTR-BWLTR



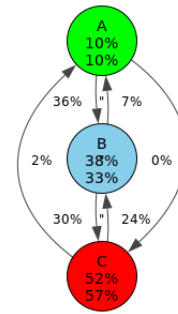
LTR-Betweenness



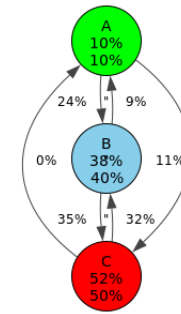
LTR-Closeness



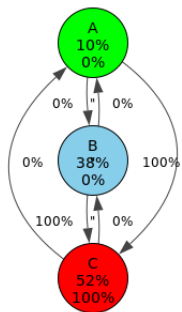
LTR-Degree



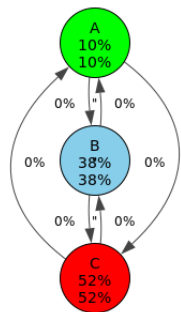
LTR-FWLTR



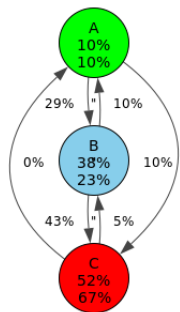
LTR-ICR



LTR-Katz



LTR-LTR



LTR-PageRank

## Apéndice C

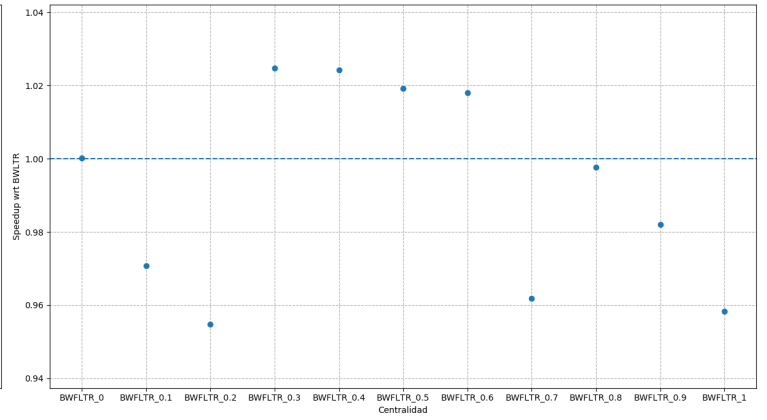
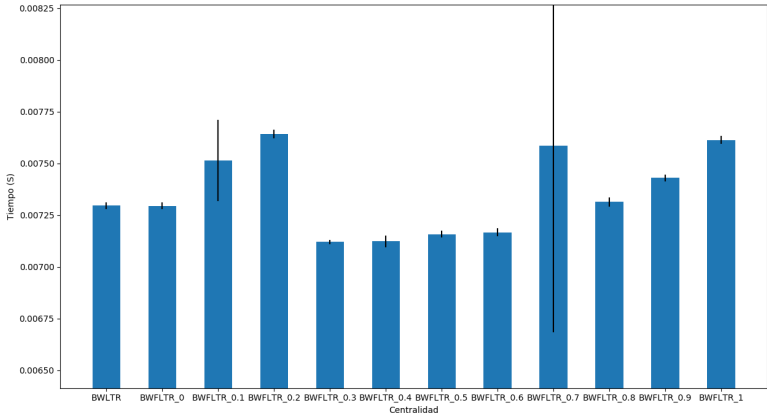
# Tiempos de ejecución

En este apéndice, incluimos los tiempos de ejecución de los diferentes experimentos.

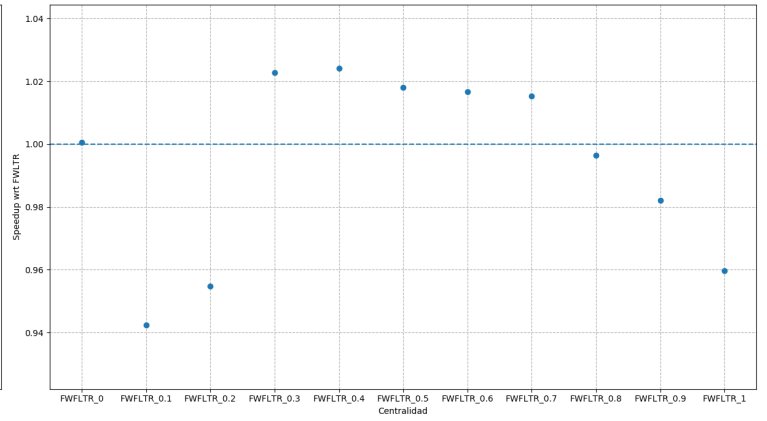
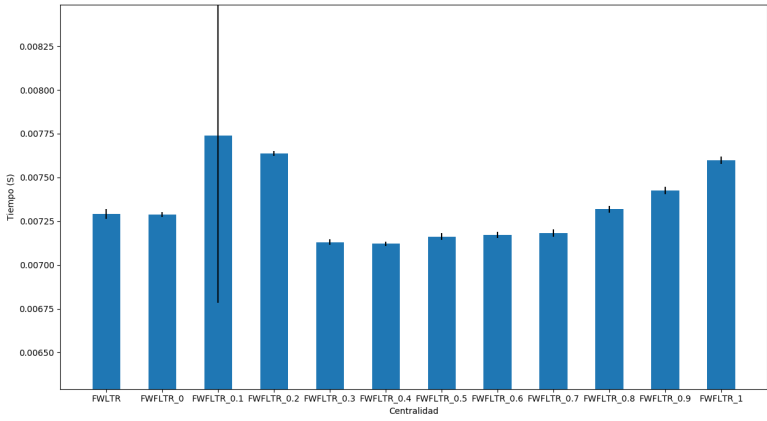
### C.1. Fading Linear Threshold Model

Añadimos aquí todos los resultados de los experimentos de la sección 5.3.2. En la primera gráfica de la izquierda, mostramos los tiempos de ejecución medios de las dos versiones de las medidas, con diferentes valores de  $\gamma$  para las medidas FLTM, los *bigotes* representan la desviación estándar de las cinco 5 ejecuciones. En la segunda gráfica de la derecha, vemos el *speedup* de las medidas FLTM respecto a las medidas LTM.

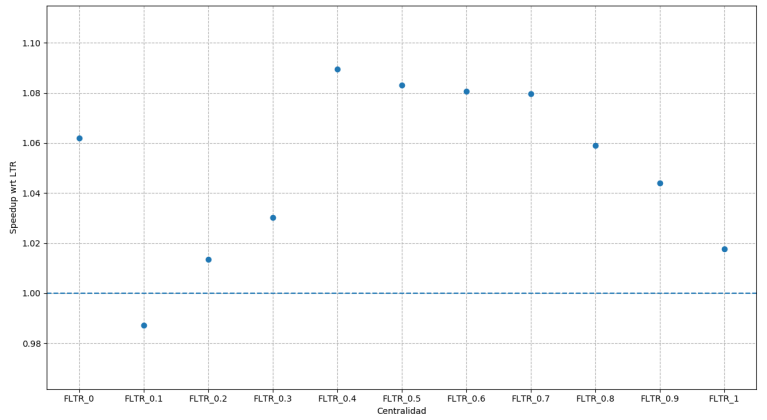
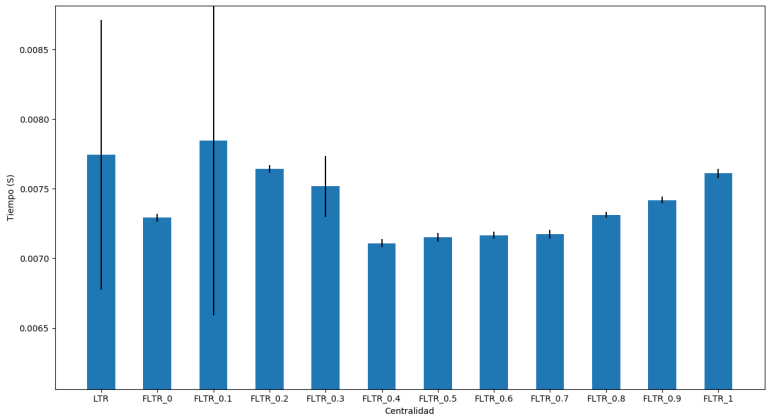
### C.1.1. ArXiv



BWLTR vs BWFLTR

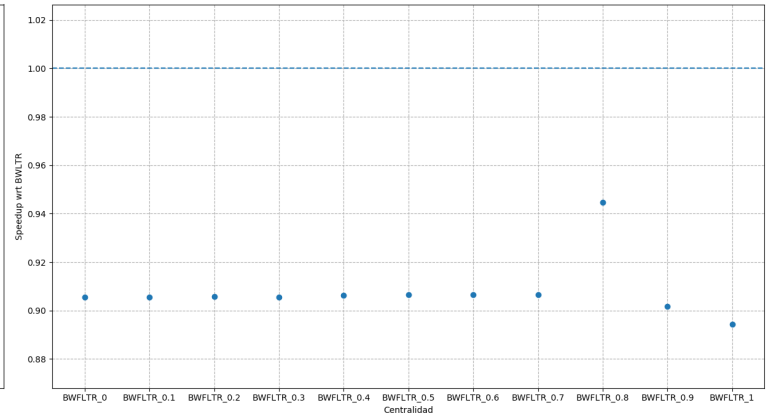
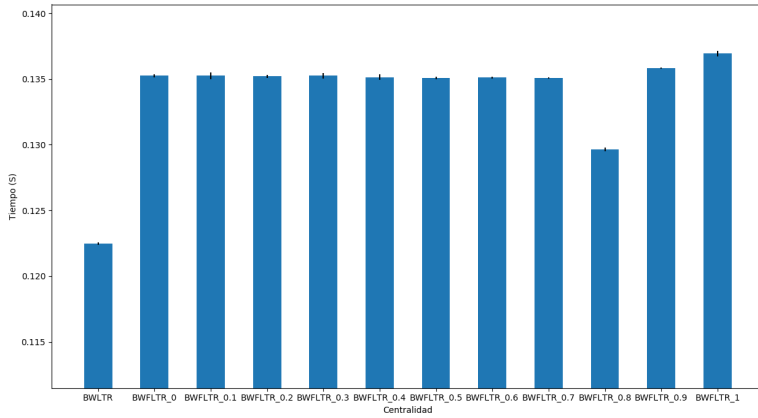


FWLTR vs FWFLTR

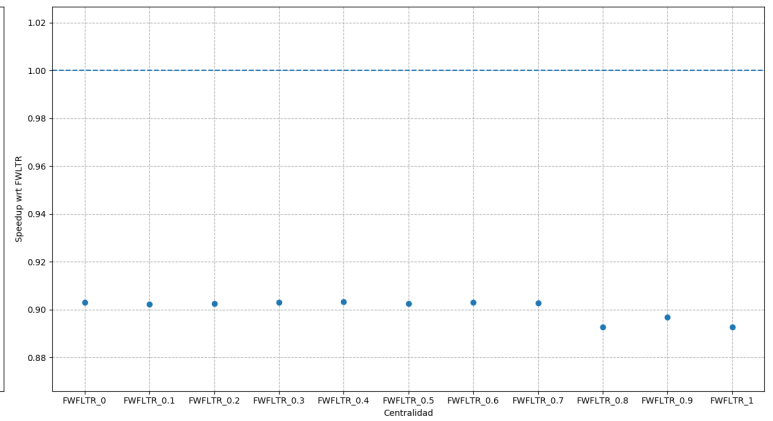
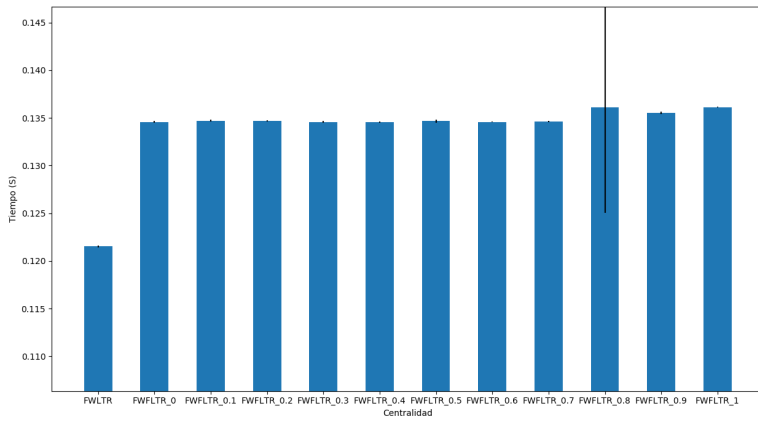


LTR vs FLTR

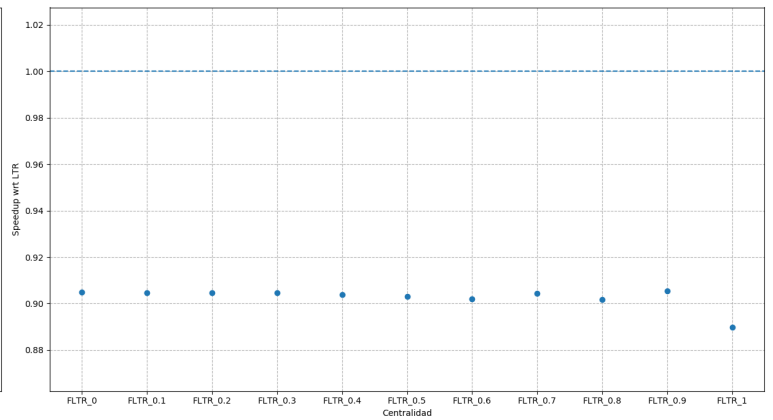
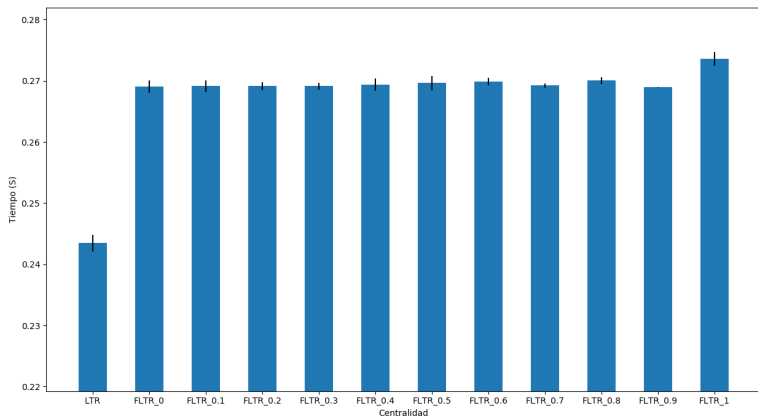
## C.1.2. Caida



BWLTR vs BWFLTR

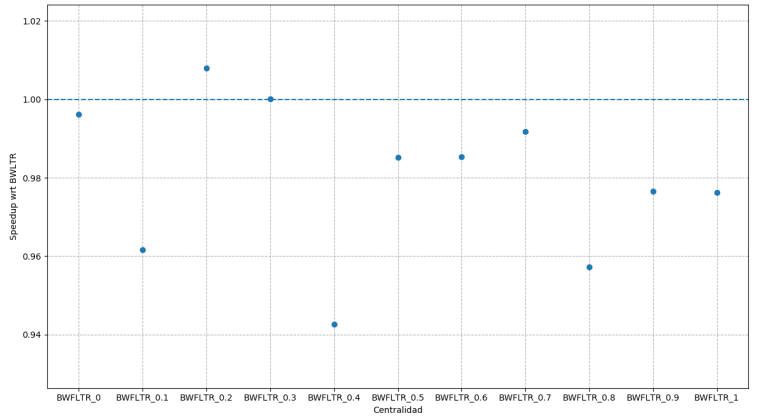
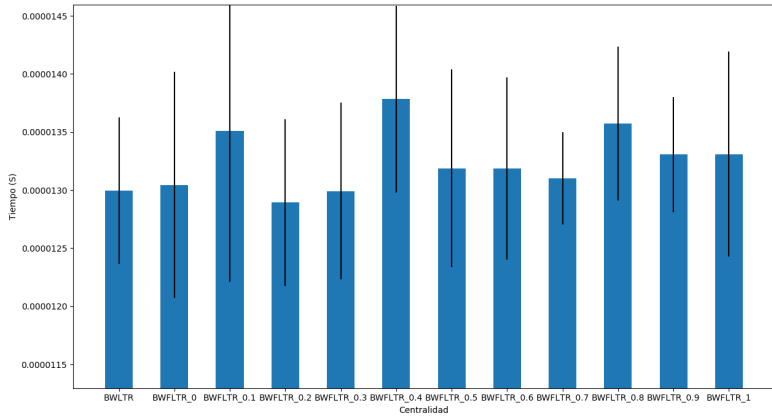


FWLTR vs FWFLTR

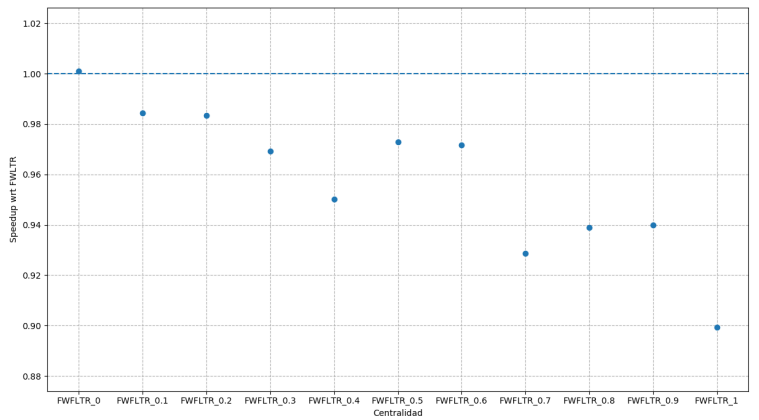
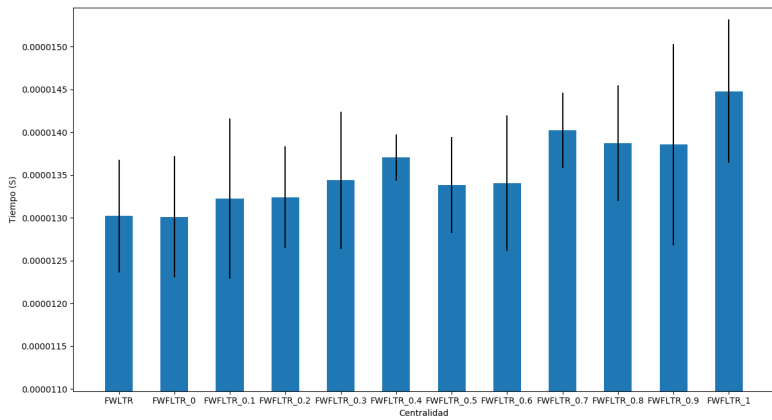


LTR vs FLTR

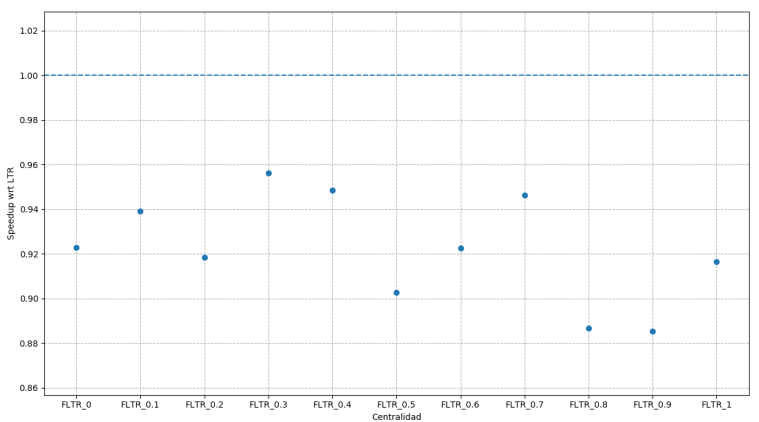
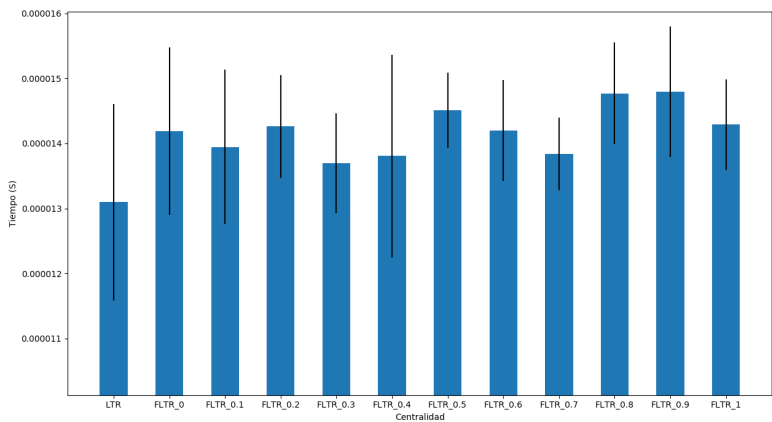
### C.1.3. Dining\_Table



BWLTR vs BWFLTR

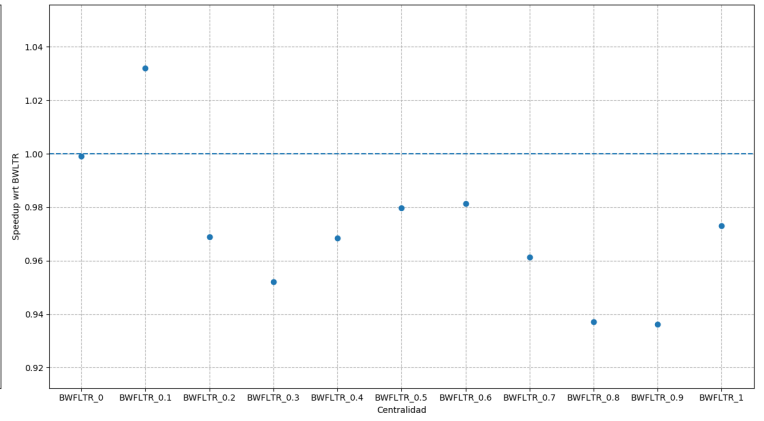
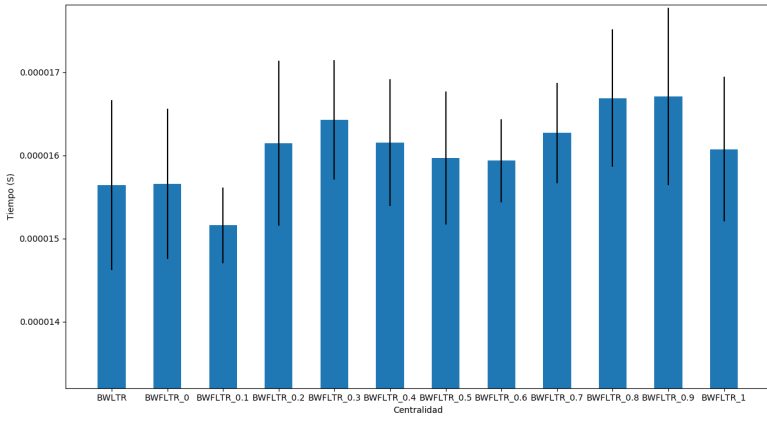


FWLTR vs FWFLTR

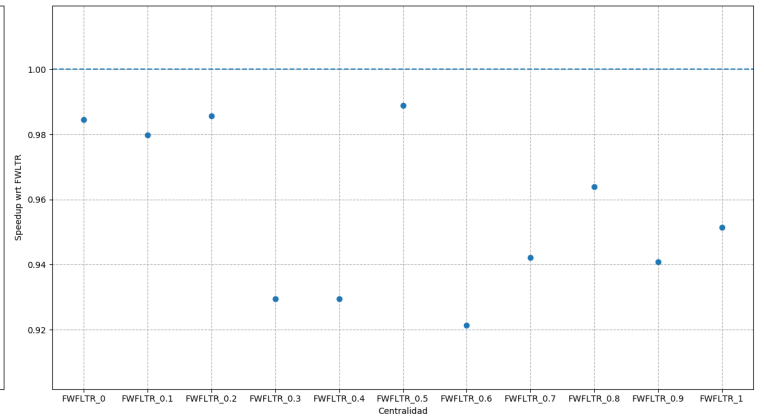
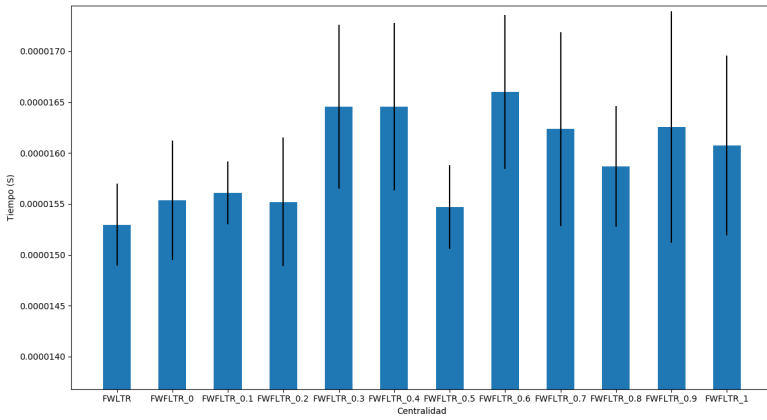


LTR vs FLTR

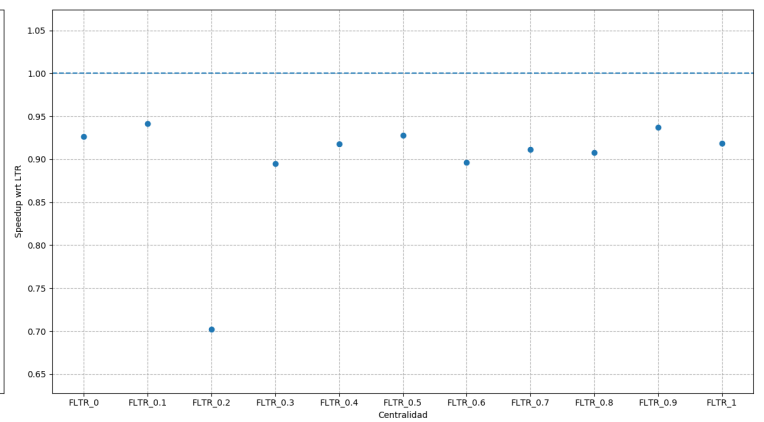
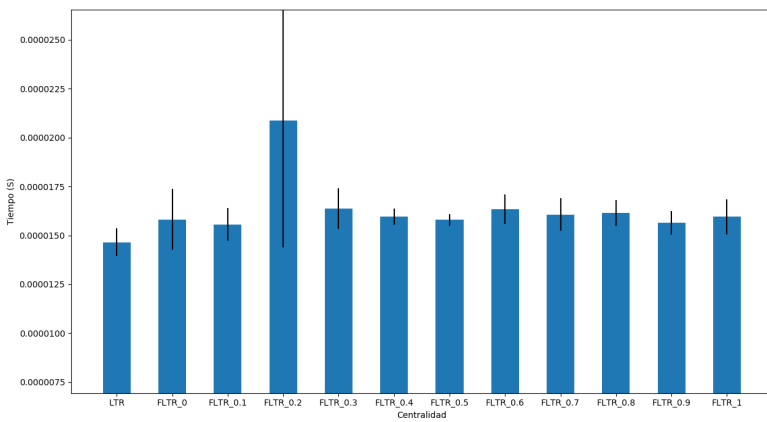
## C.1.4. Dolphins



BWLTR vs BWFLTR

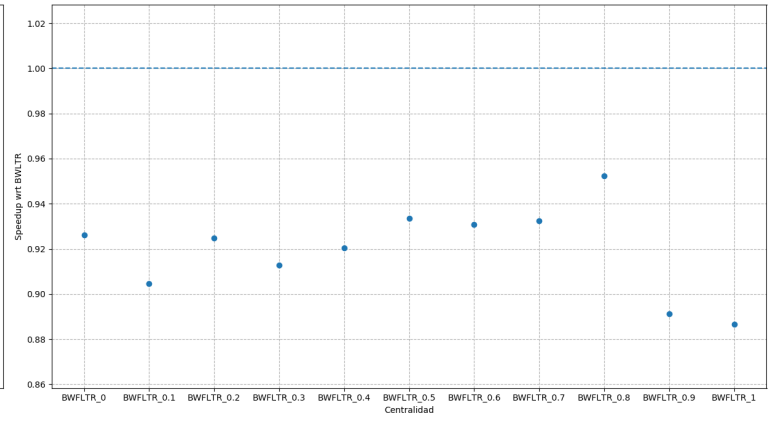
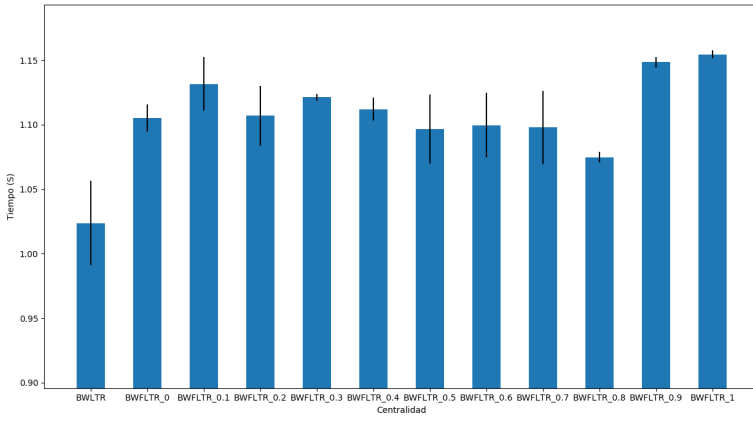


FWLTR vs FWFLTR

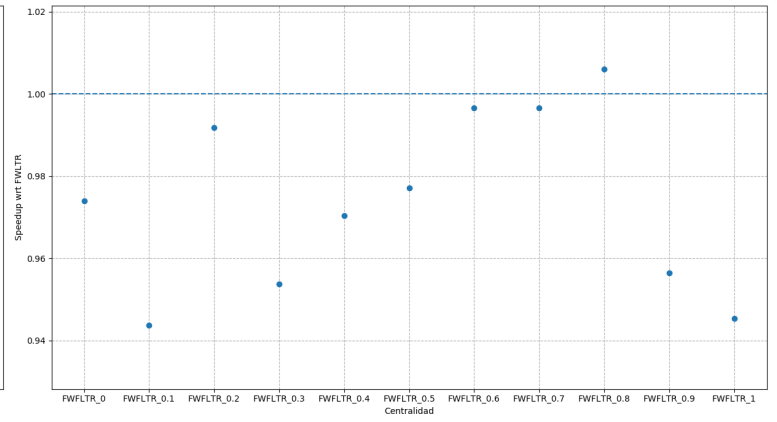
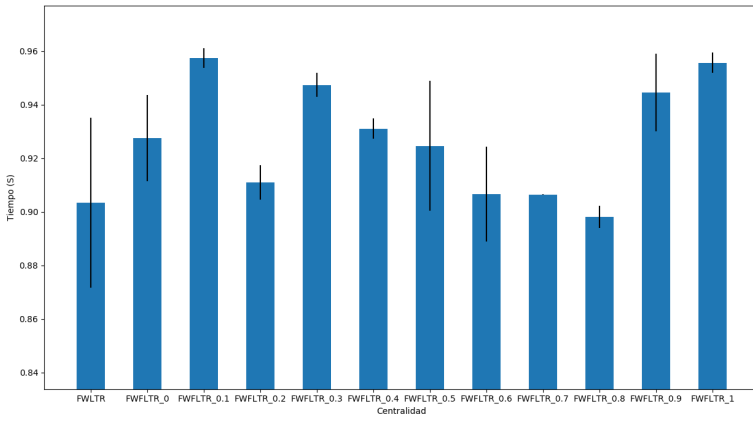


LTR vs FLTR

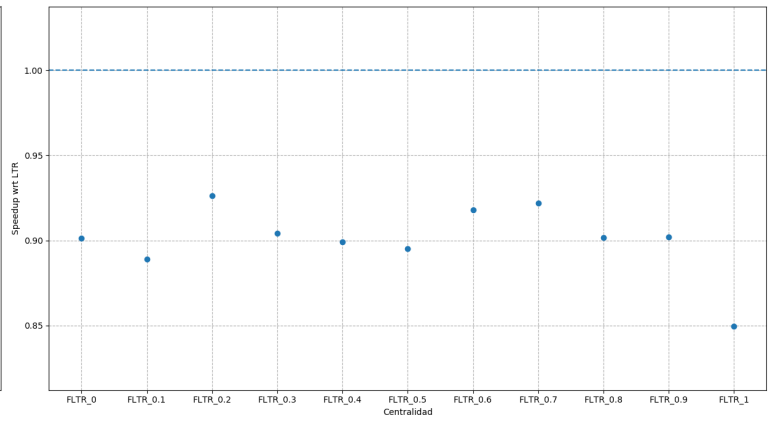
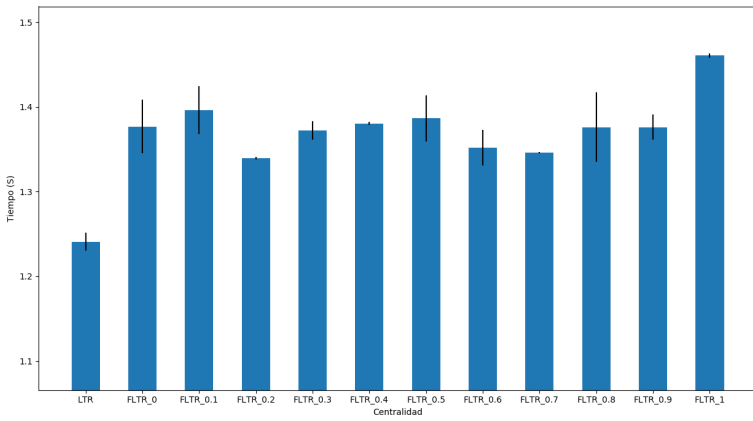
### C.1.5. Epinions



BWLTR vs BWFLTR



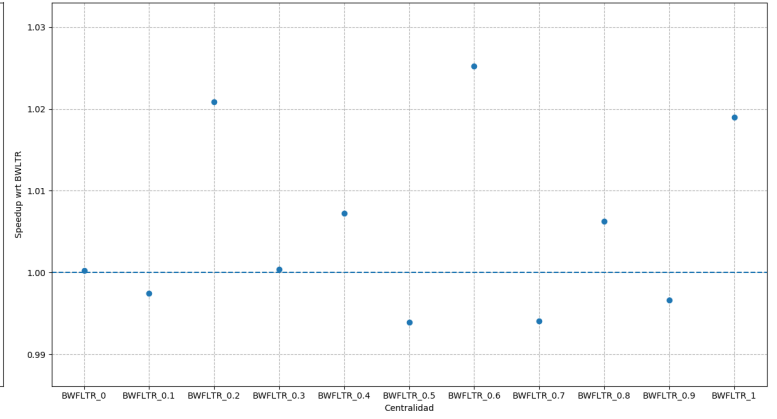
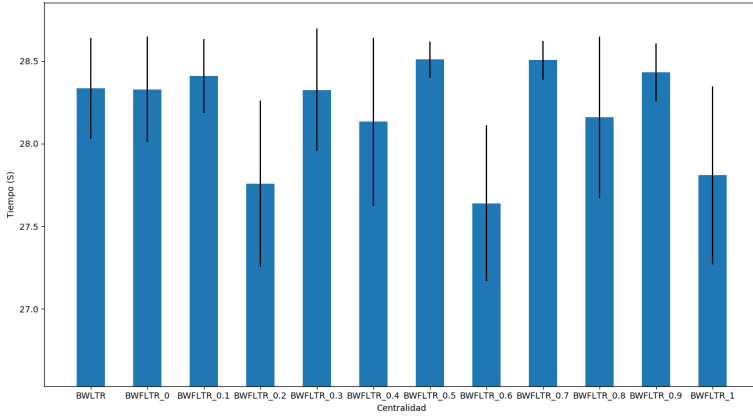
FWLTR vs FWFLTR



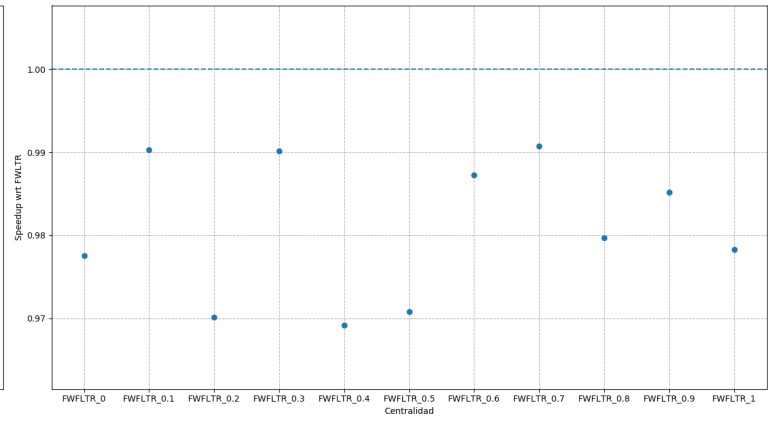
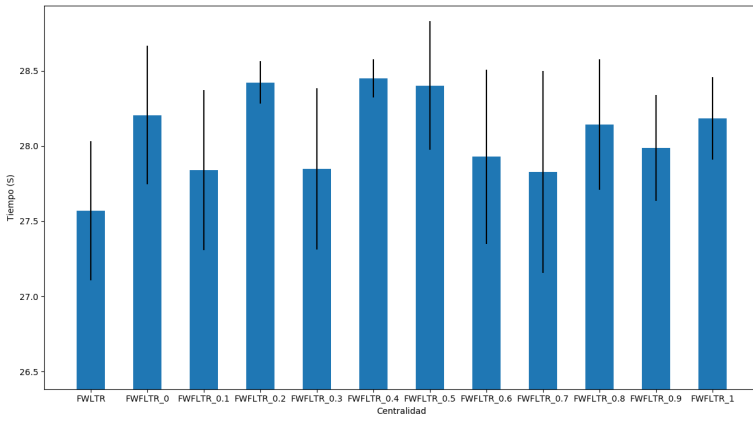
LTR vs FLTR



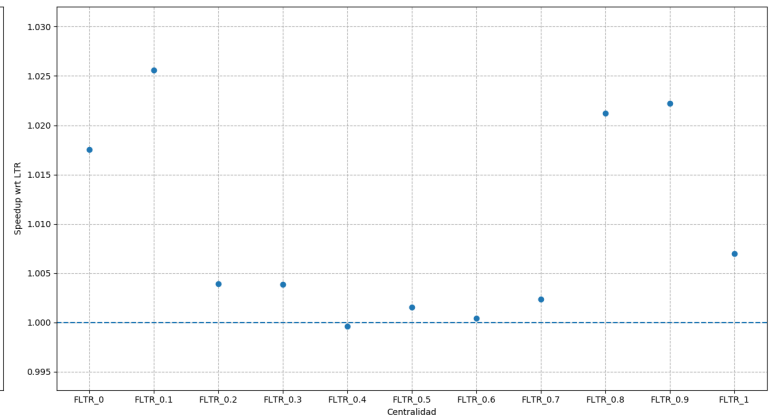
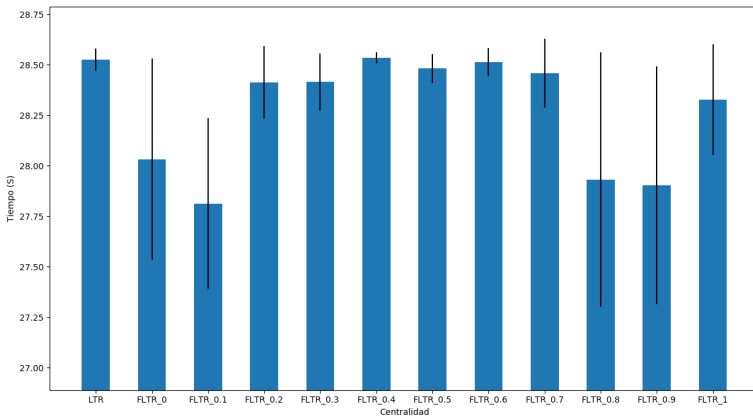
## C.1.6. Higgs



BWLTR vs BWFLTR

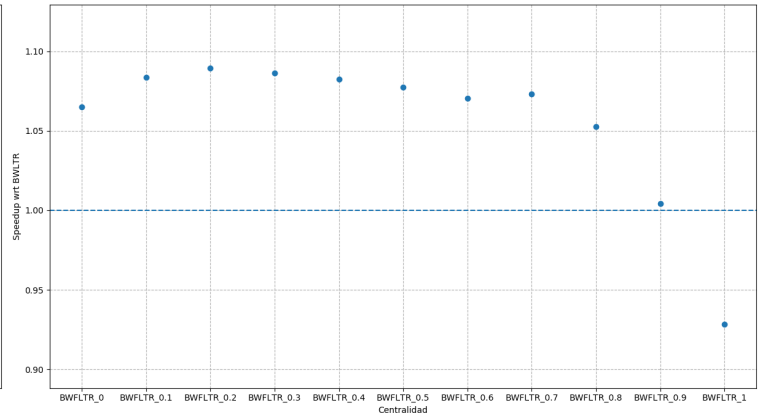
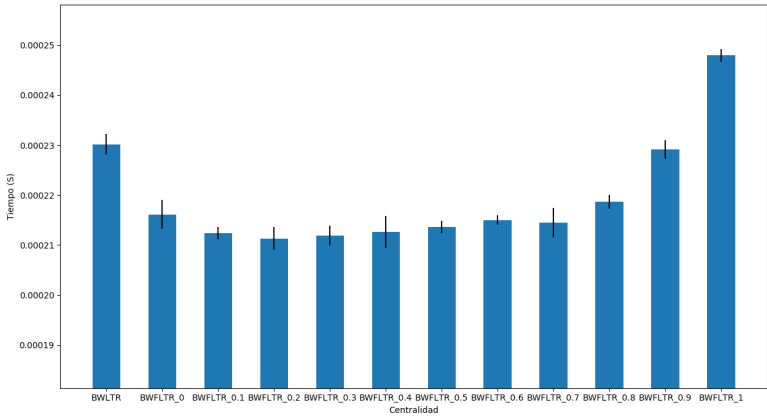


FWLTR vs FWFLTR

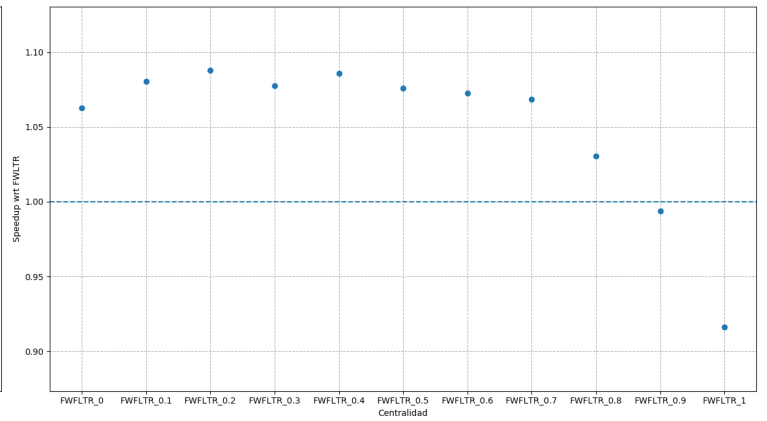
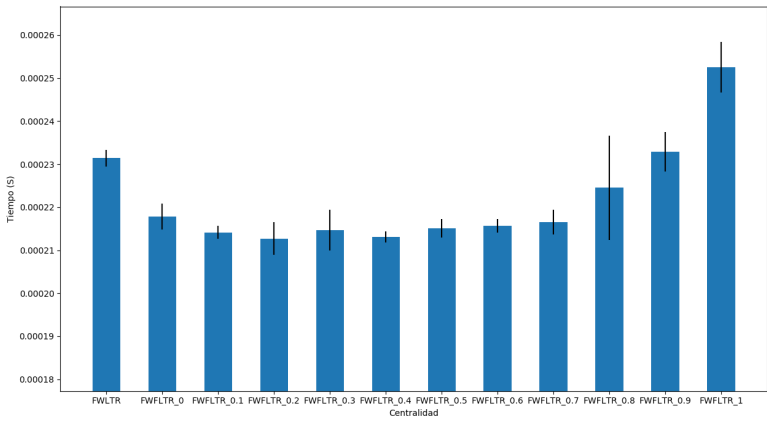


LTR vs FLTR

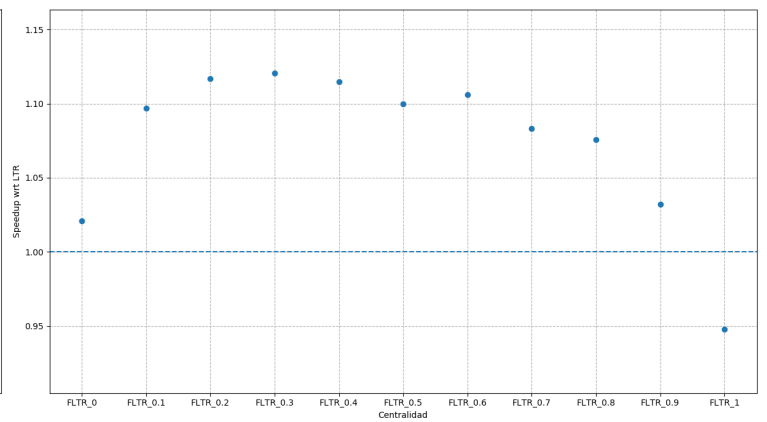
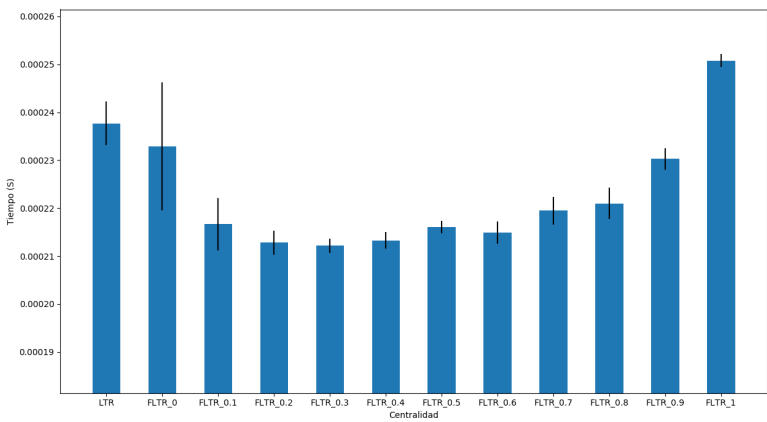
### C.1.7. Human Brain



BWLTR vs BWFLTR

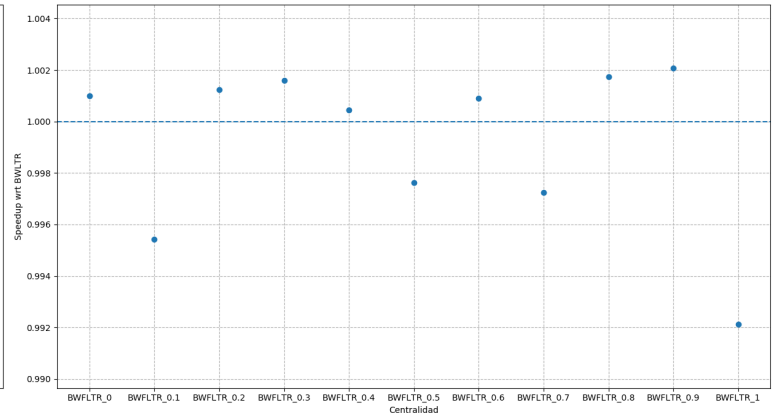
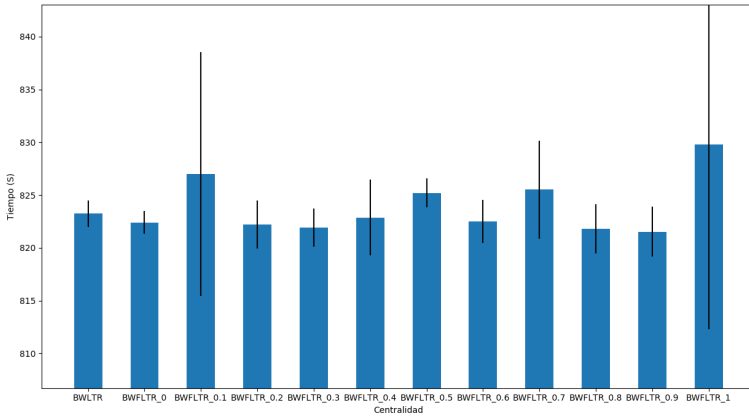


FWLTR vs FWFLTR

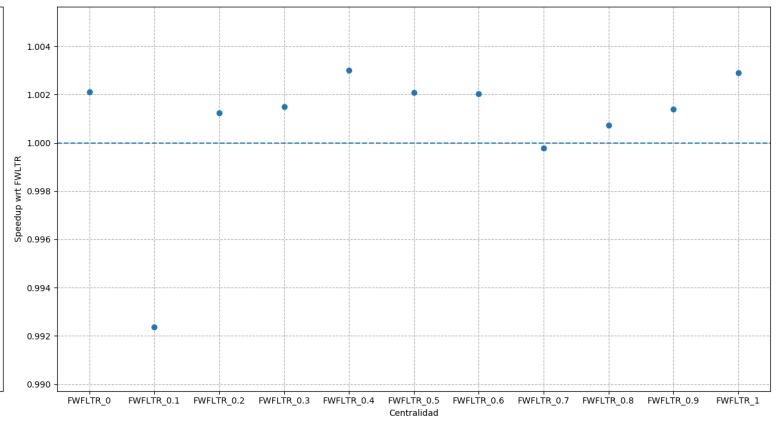
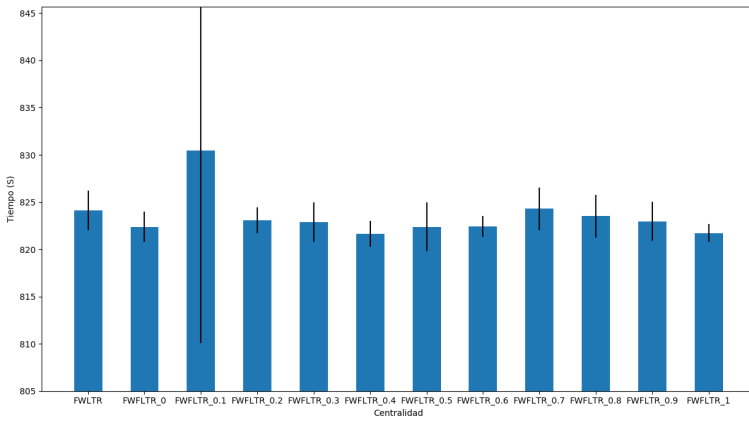


LTR vs FLTR

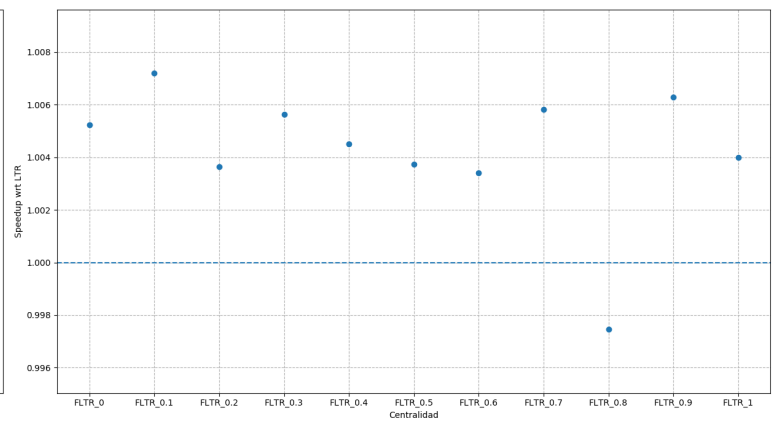
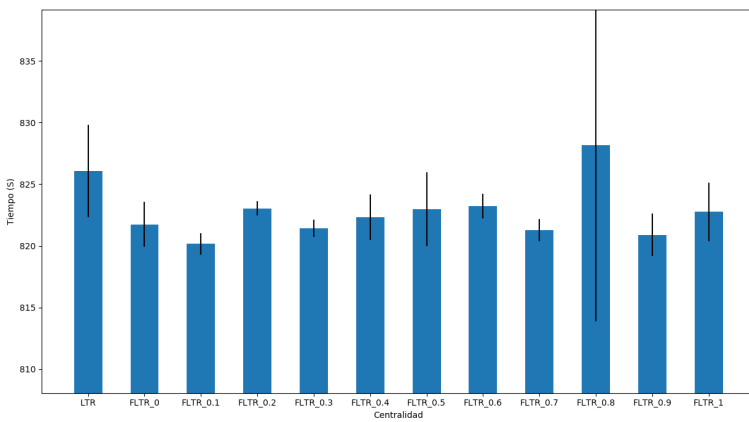
## C.1.8. Texas



BWLTR vs BWFLTR

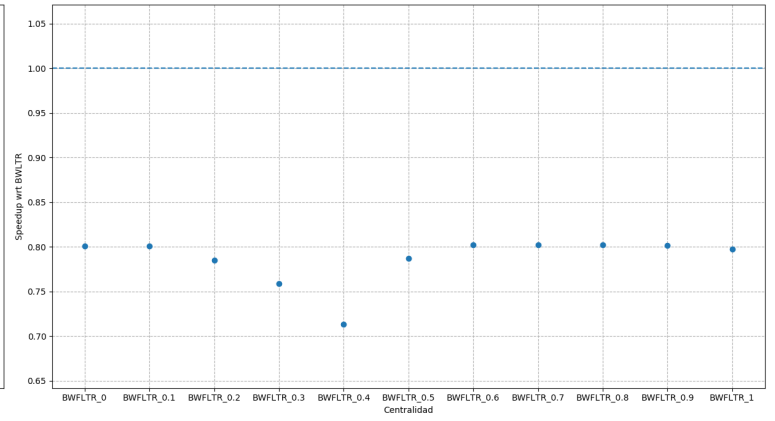
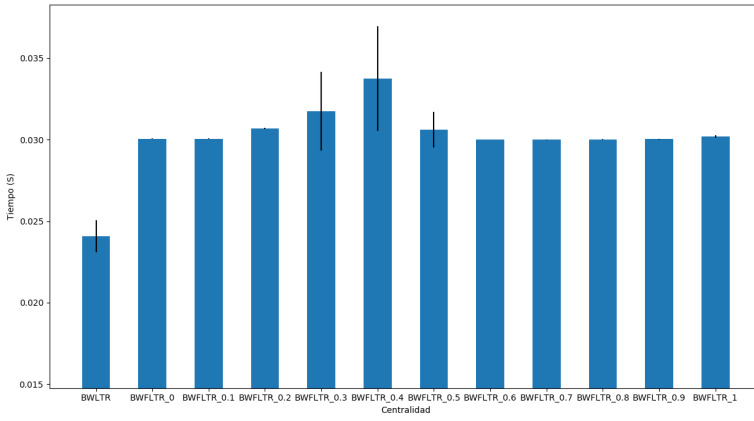


FWLTR vs FWFLTR

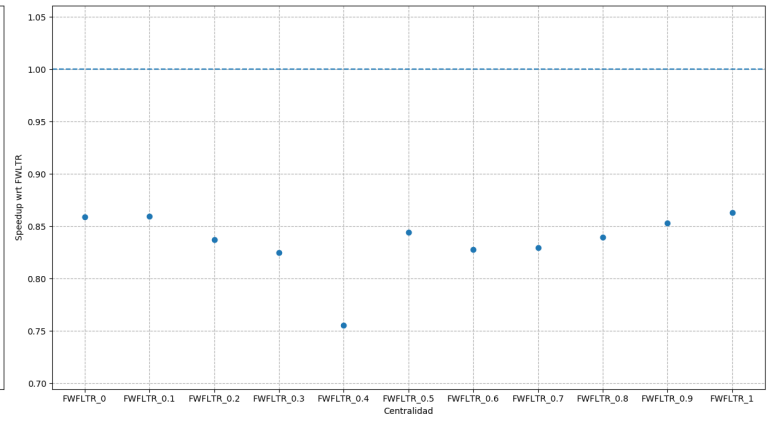
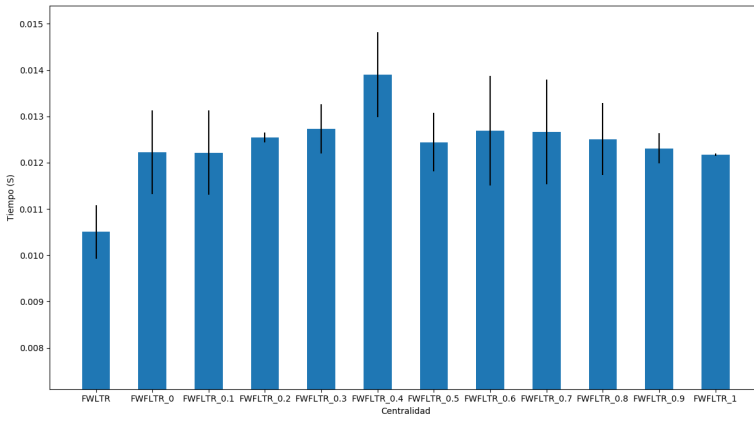


LTR vs FLTR

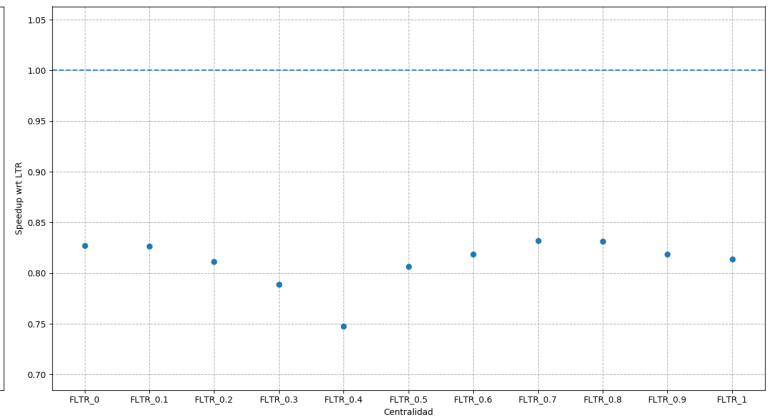
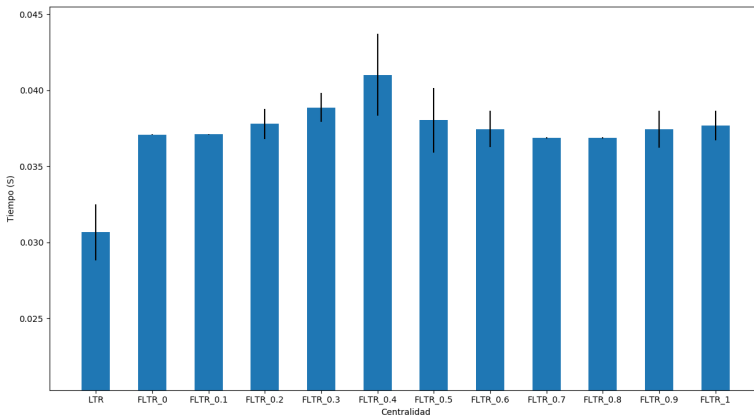
### C.1.9. Wikipedia



BWLTR vs BWFLTR



FWLTR vs FWFLTR

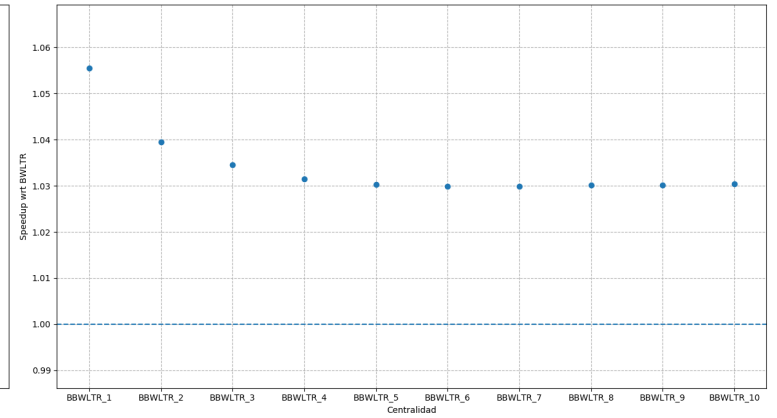
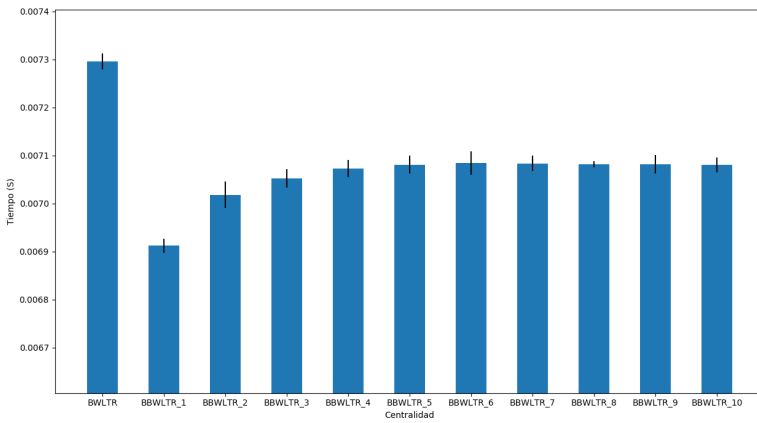


LTR vs FLTR

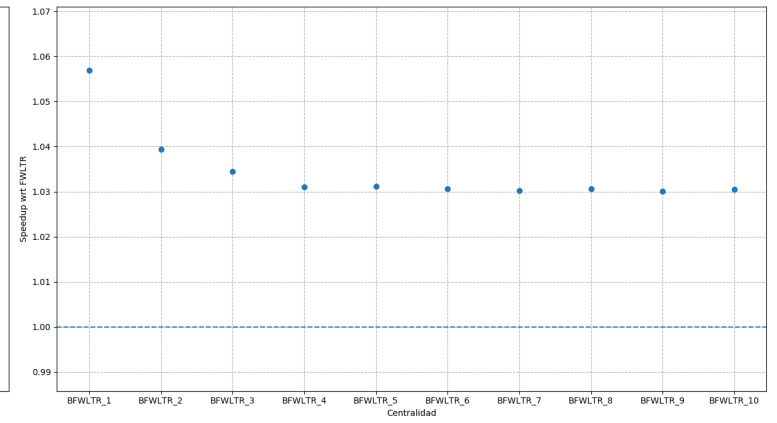
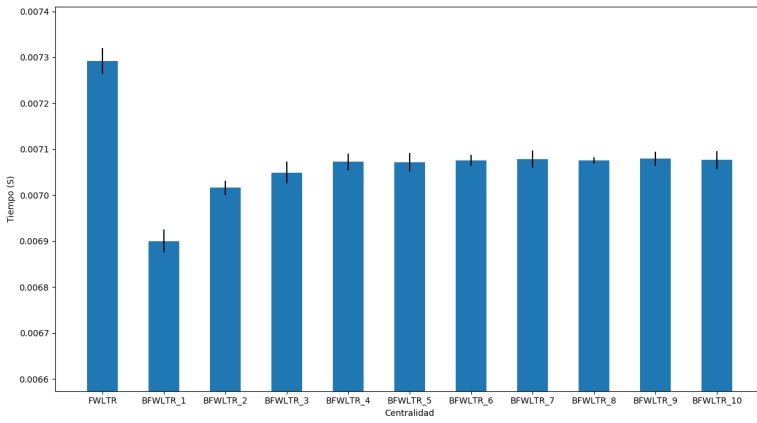
## C.2. Bounded Linear Threshold Model

Añadimos aquí todos los resultados de los experimentos de la sección 5.3.2. En la primera gráfica de la izquierda, mostramos los tiempos de ejecución medios de las dos versiones de las medidas, con diferentes valores de  $\gamma$  para las medidas BLTM, los *bigotes* representan la desviación estándar de las cinco 5 ejecuciones. En la segunda gráfica de la derecha, vemos el *speedup* de las medidas BLTM respecto a las medidas LTM.

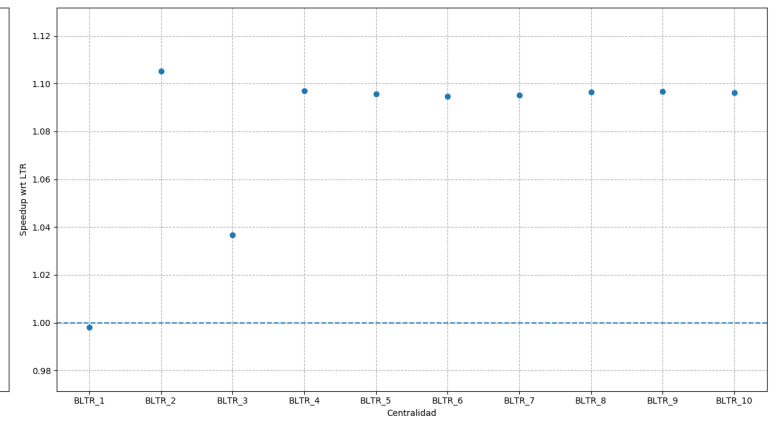
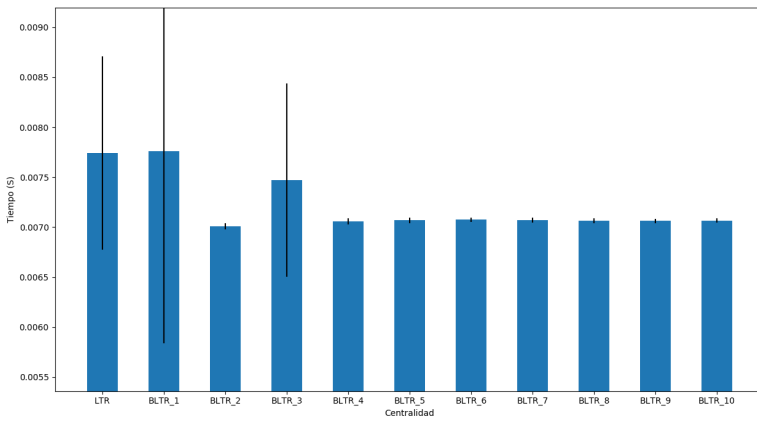
## C.2.1. ArXiv



BBWLTR vs BBWLTR

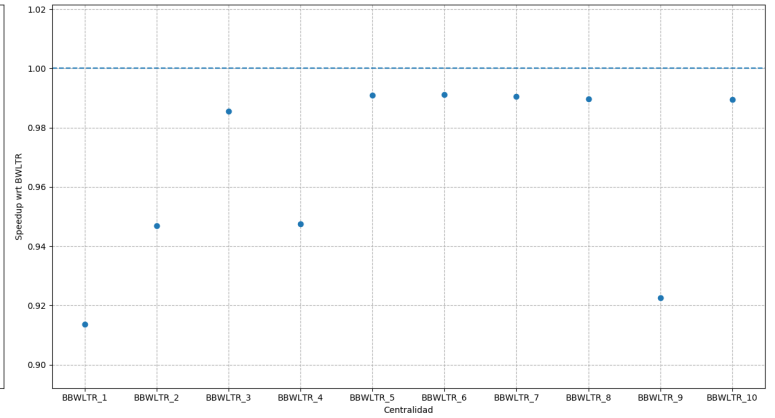
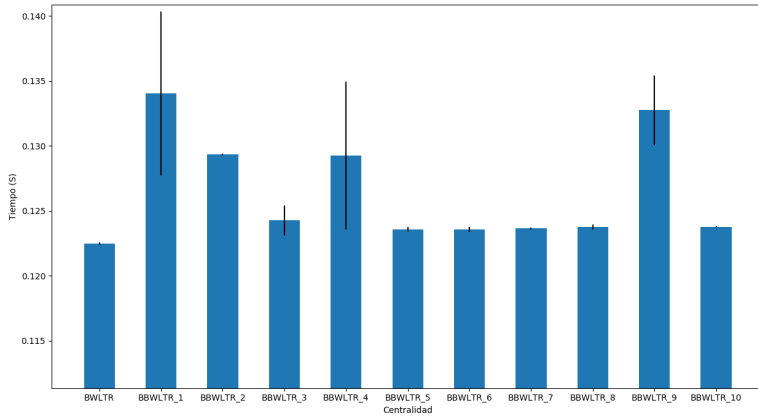


BFWLTR vs BFWLTR

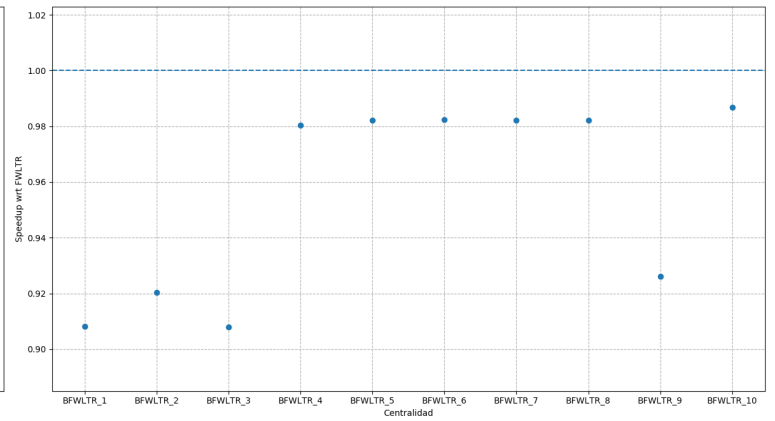
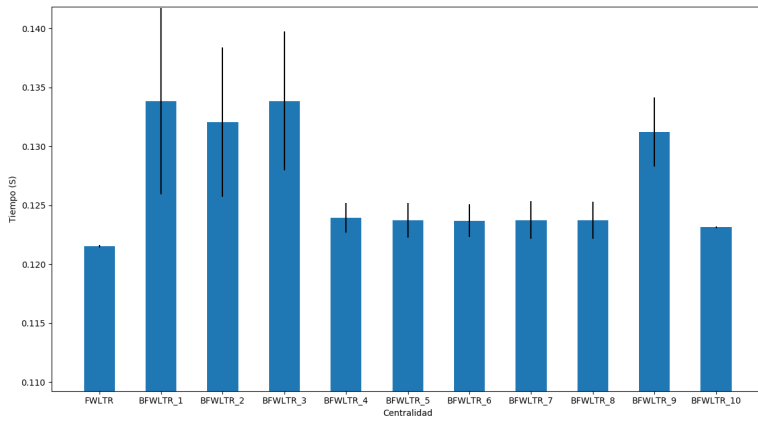


BLTR vs BLTR

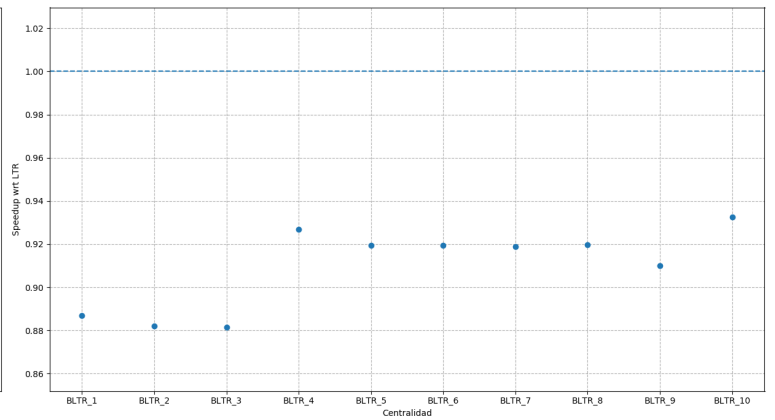
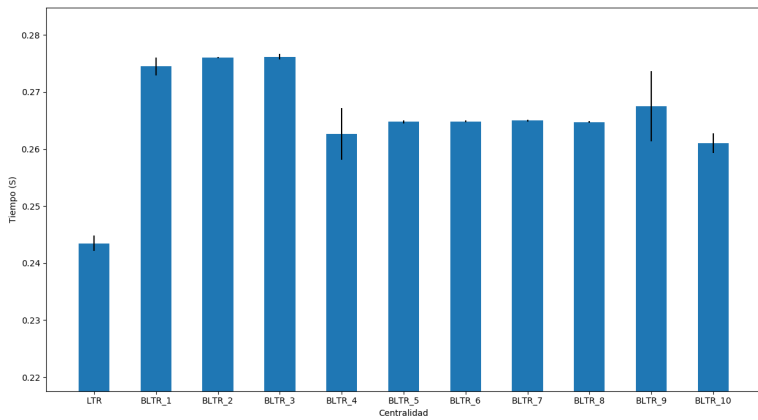
## C.2.2. Caida



BWLTR vs BBWLTR

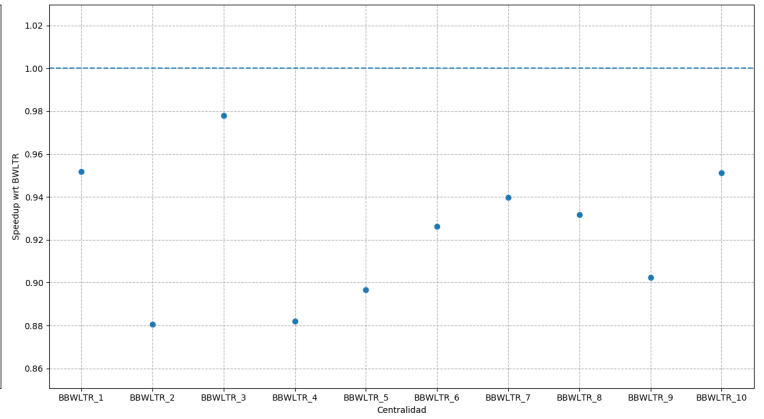
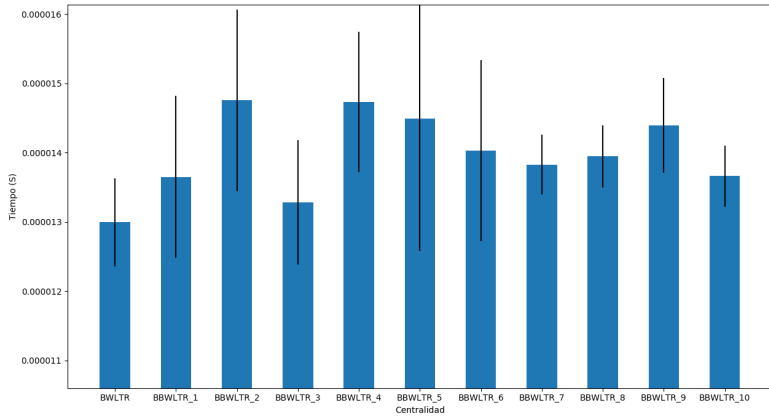


FWLTR vs BFWLTR

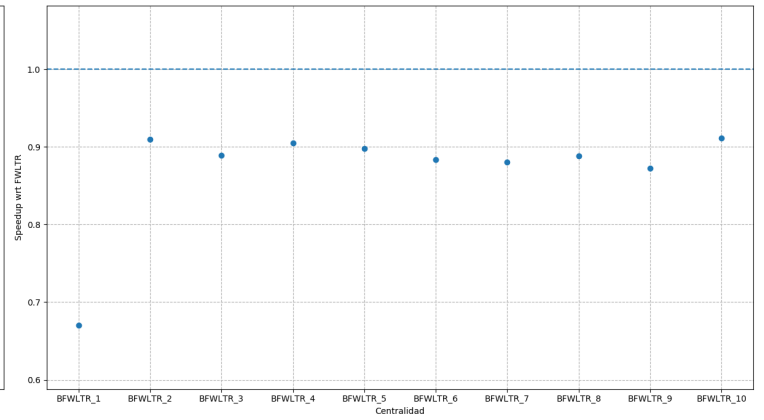
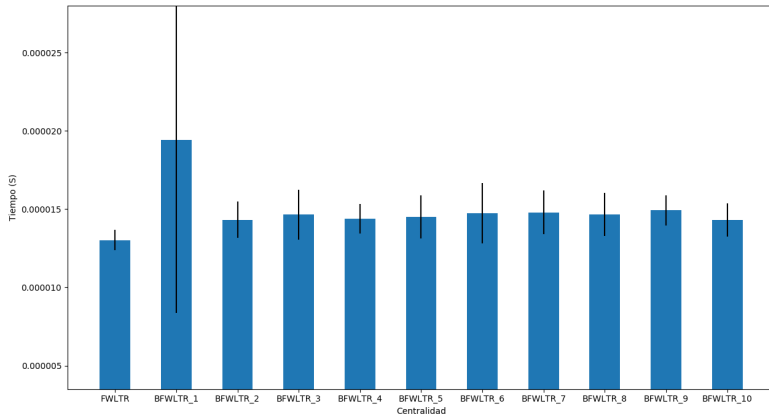


LTR vs BLTR

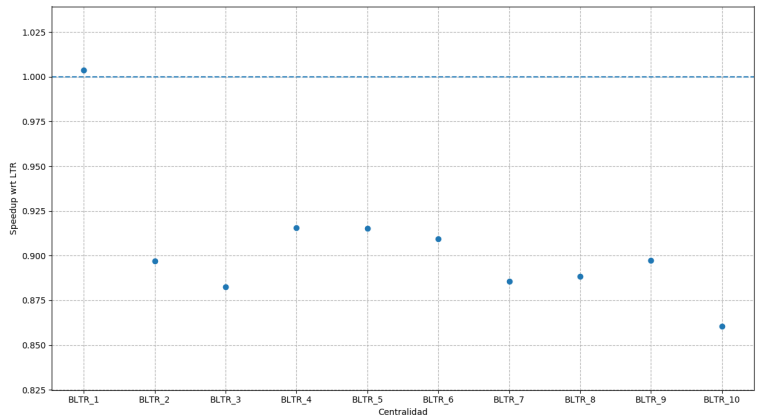
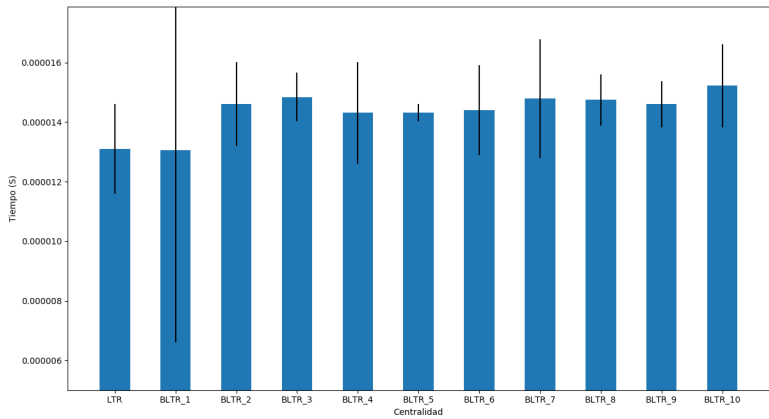
### C.2.3. Dining\_Table



BWLTR vs BBWLTR



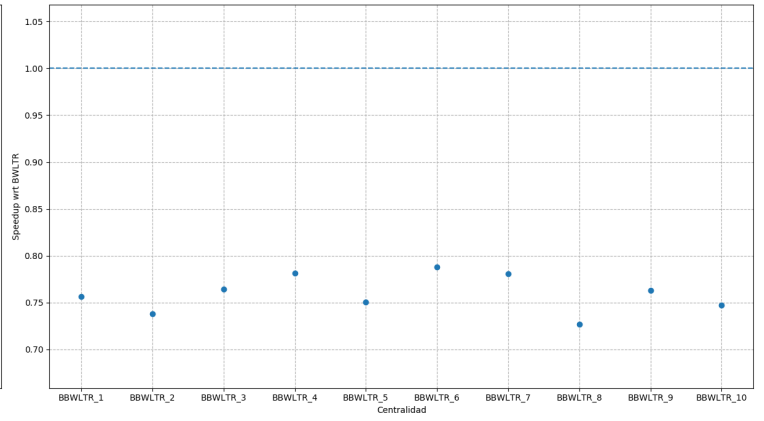
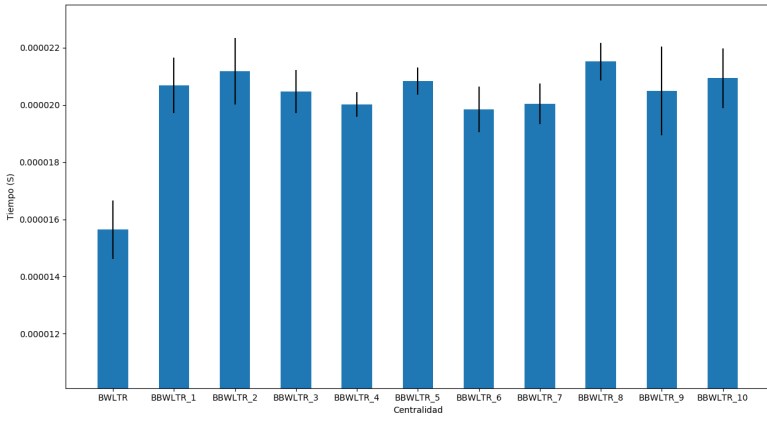
FWLTR vs BFWLTR



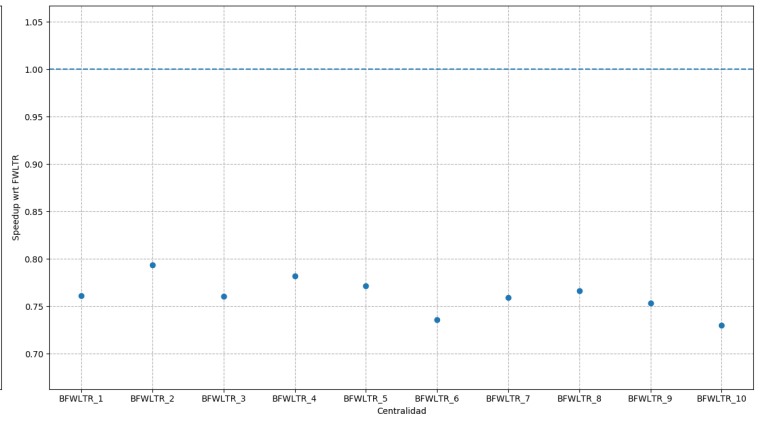
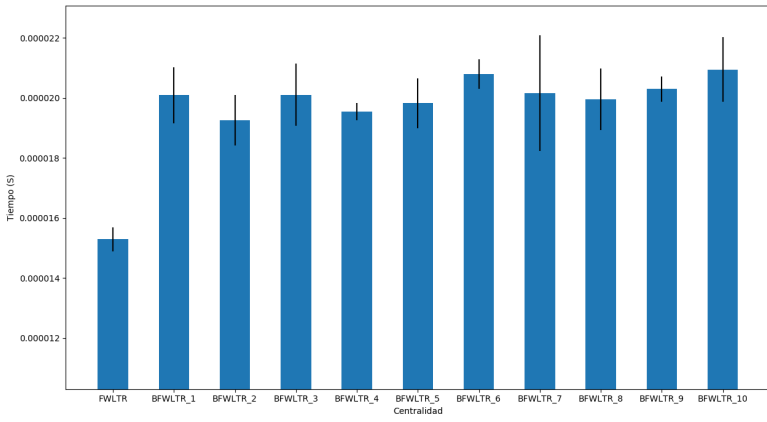
LTR vs BLTR



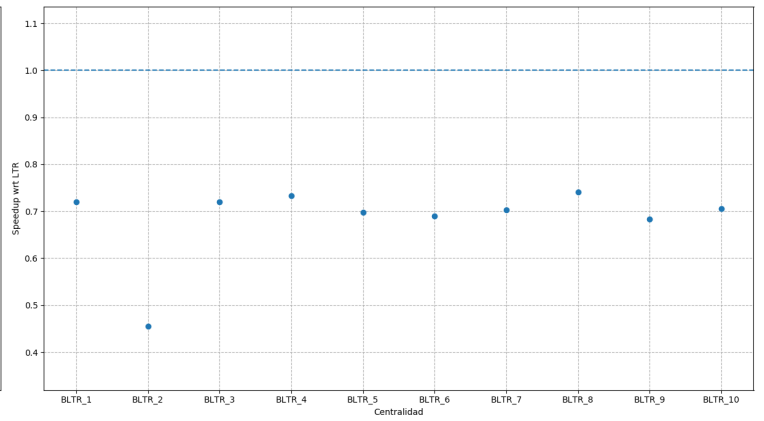
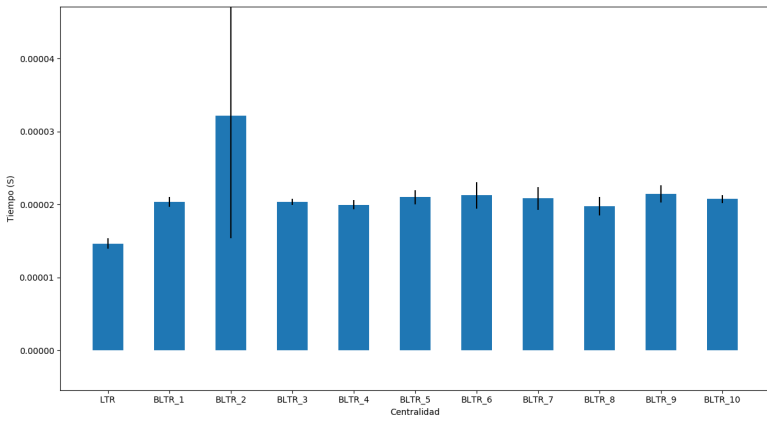
## C.2.4. Dolphins



BWLTR vs BBWLTR

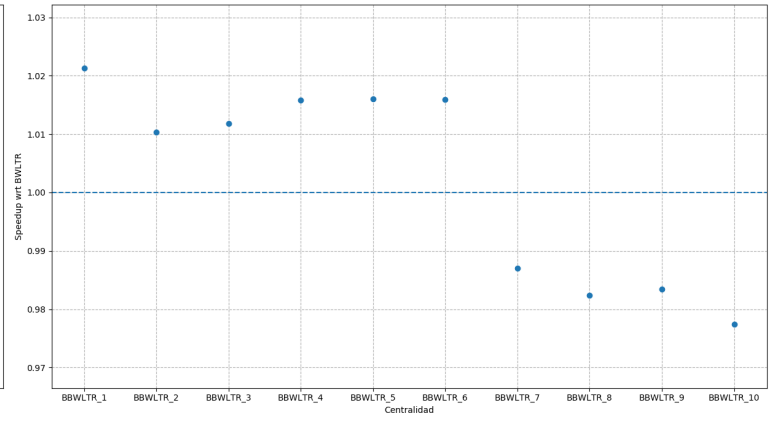
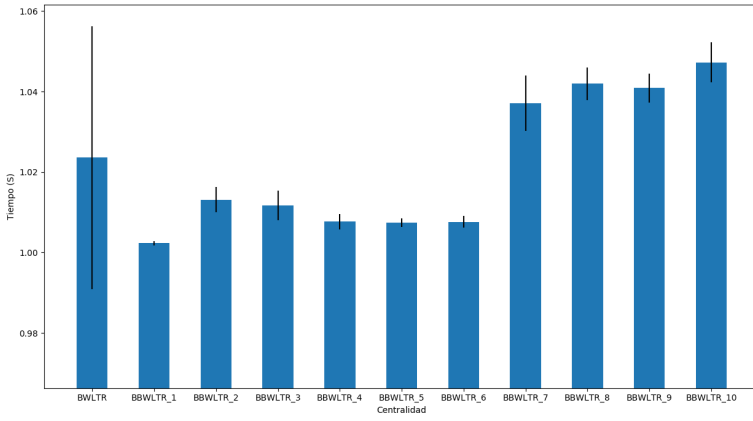


FWLTR vs BFWLTR

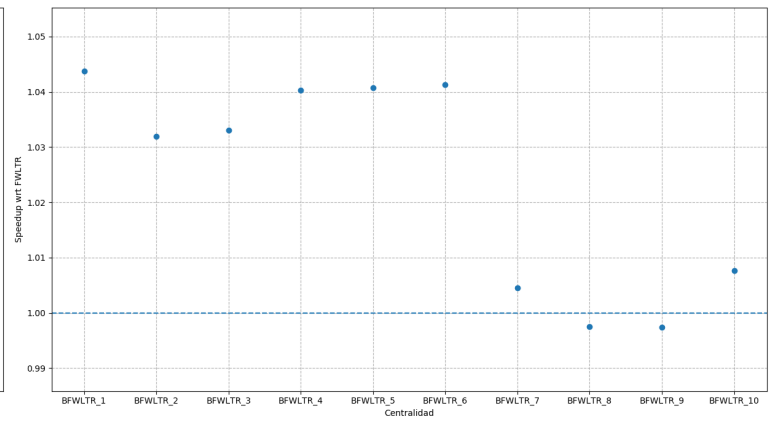
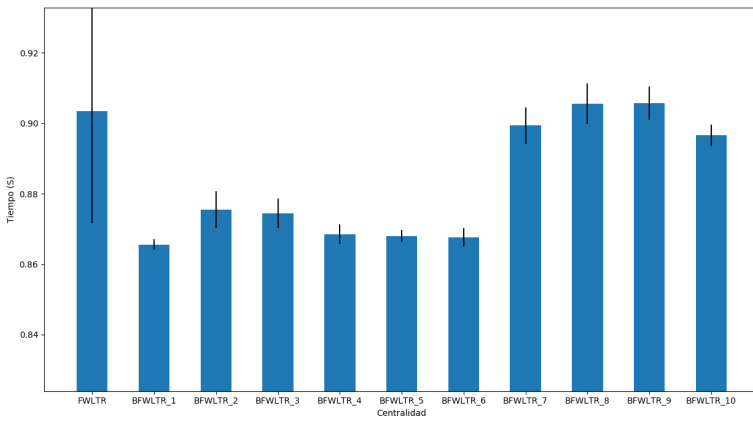


LTR vs BLTR

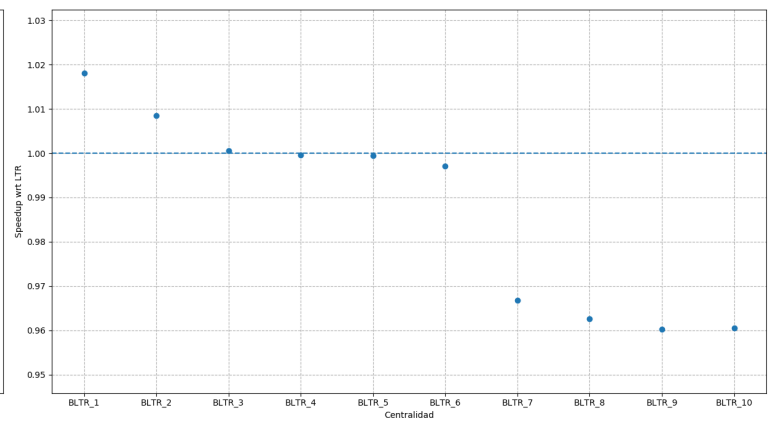
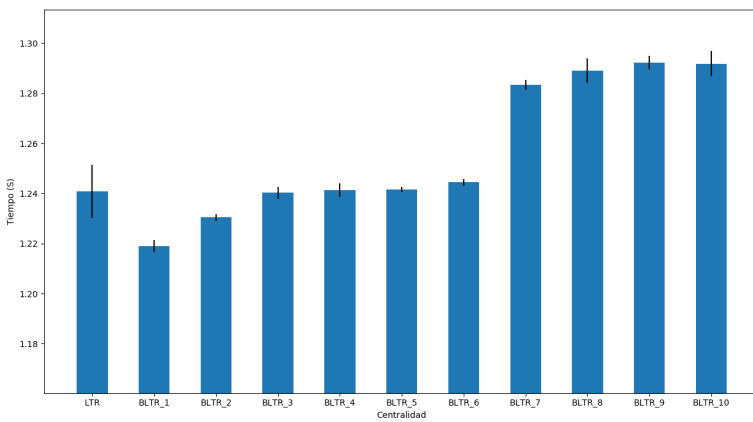
## C.2.5. Epinions



BBWLTR vs BBWLTR

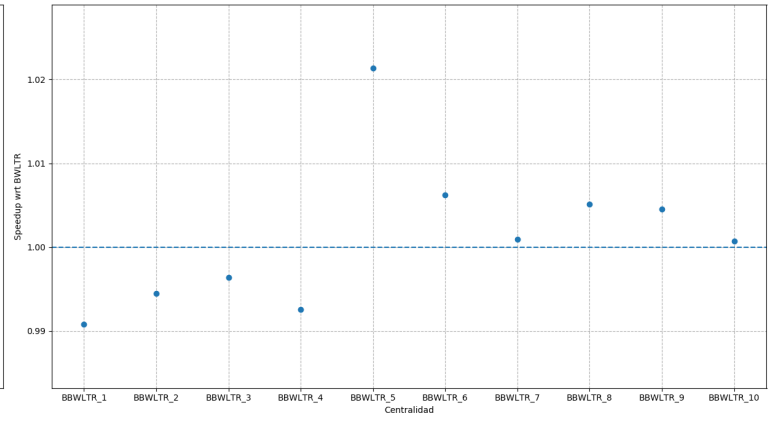
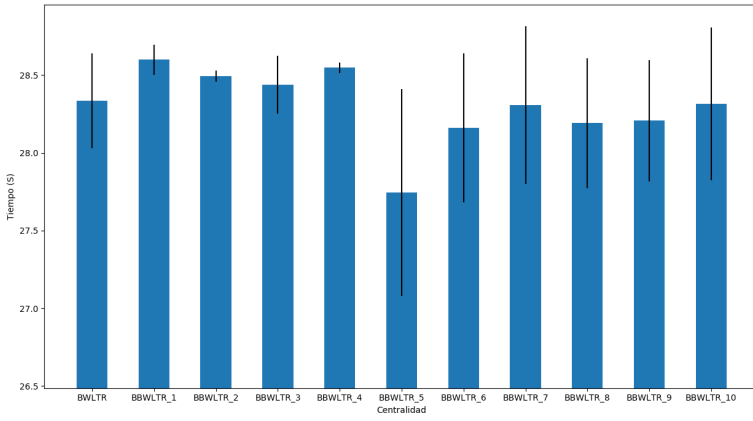


BFWLTR vs BFWLTR

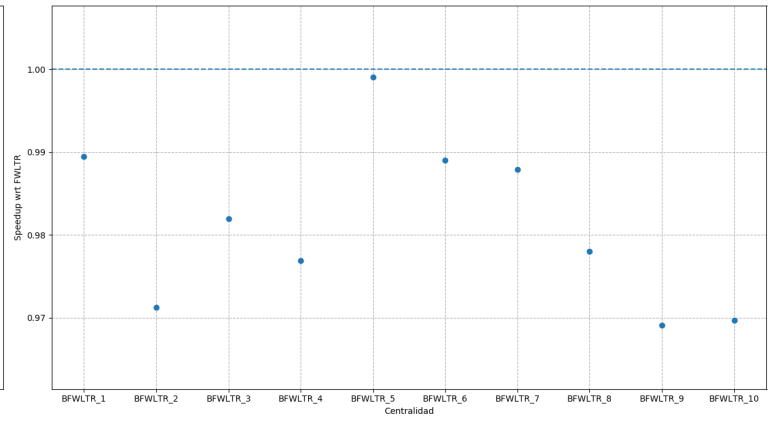
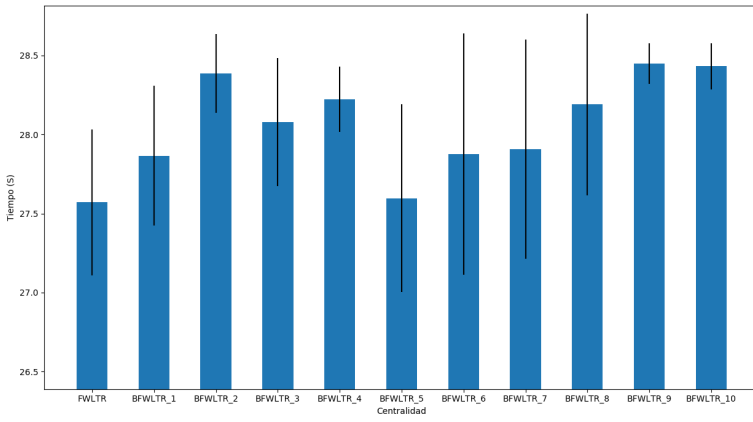


BLTR vs BLTR

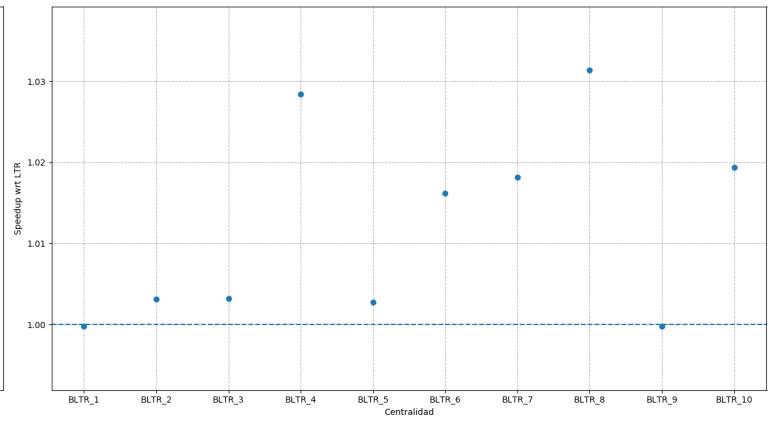
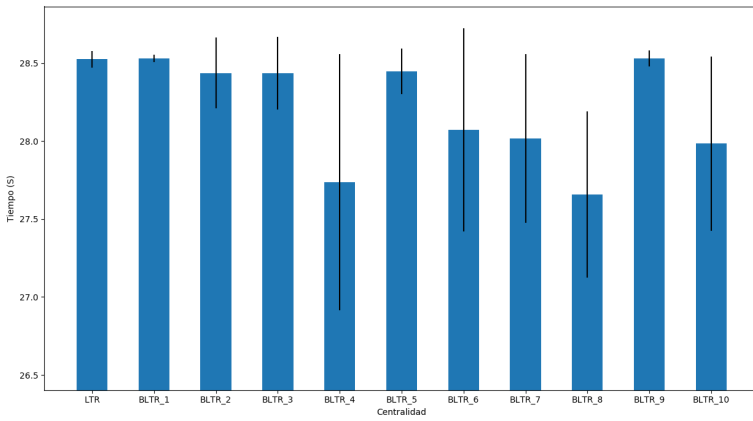
## C.2.6. Higgs



BWLTR vs BBWLTR

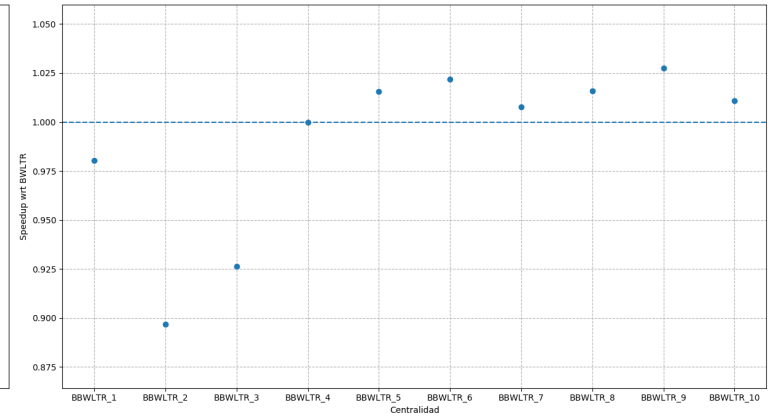
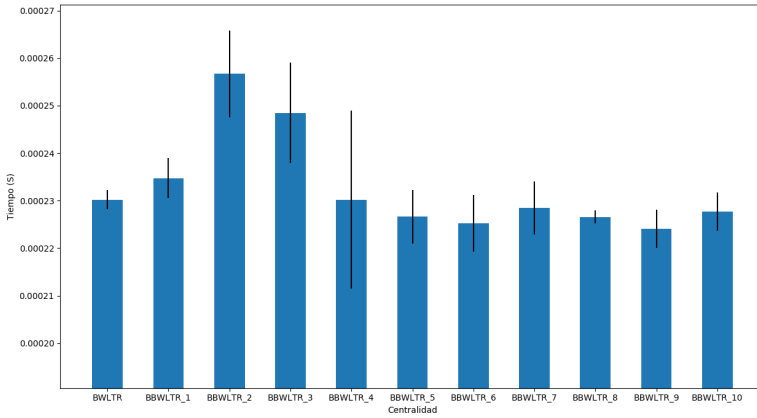


FWLTR vs BFWLTR

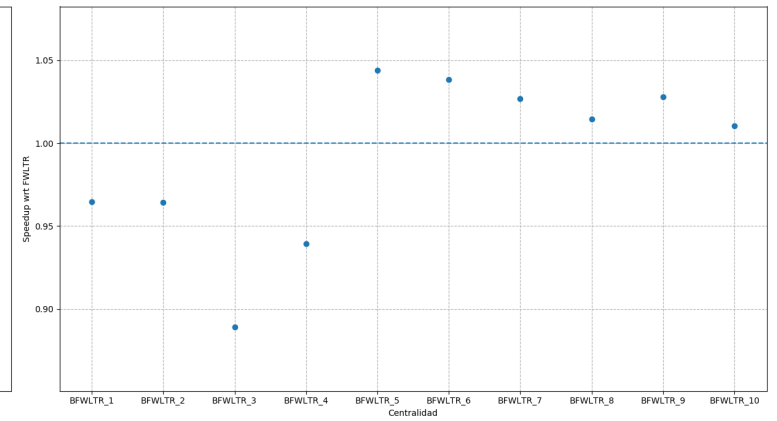
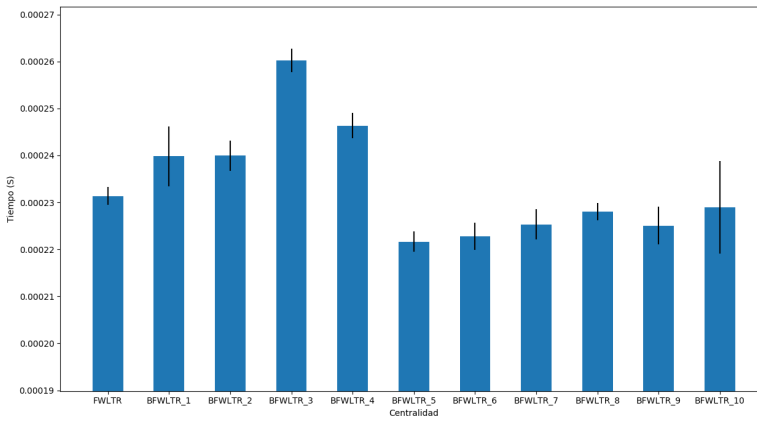


LTR vs BLTR

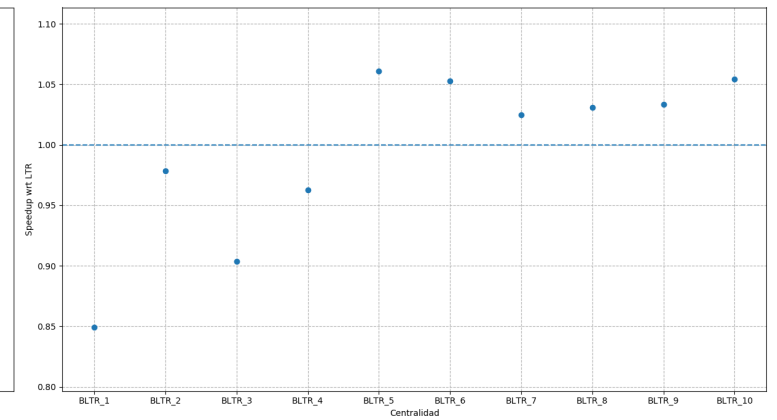
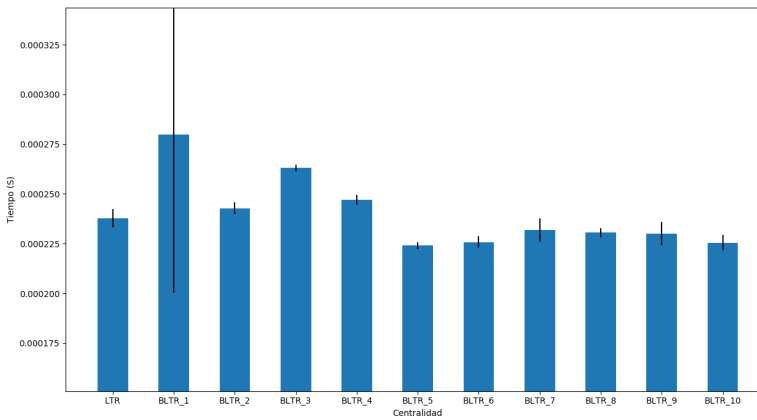
## C.2.7. Human Brain



BBWLTR vs BBWLTR

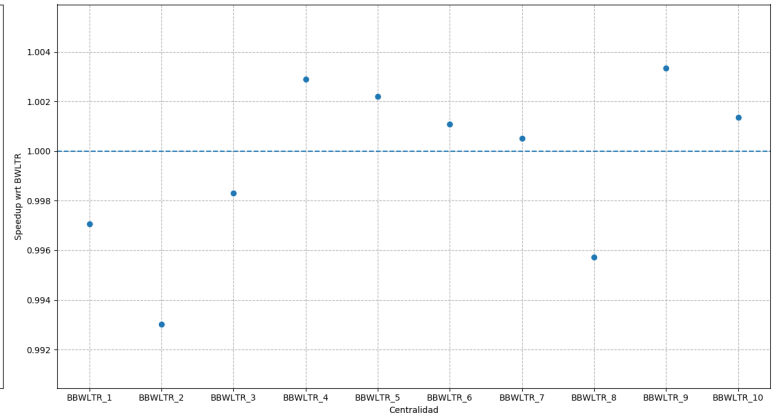
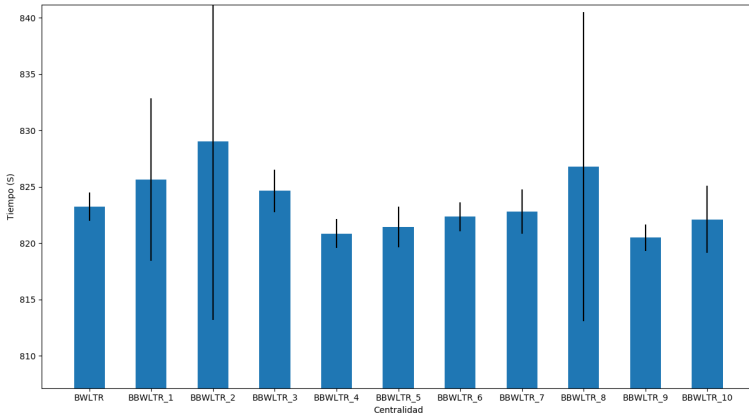


BFWLTR vs BFWLTR

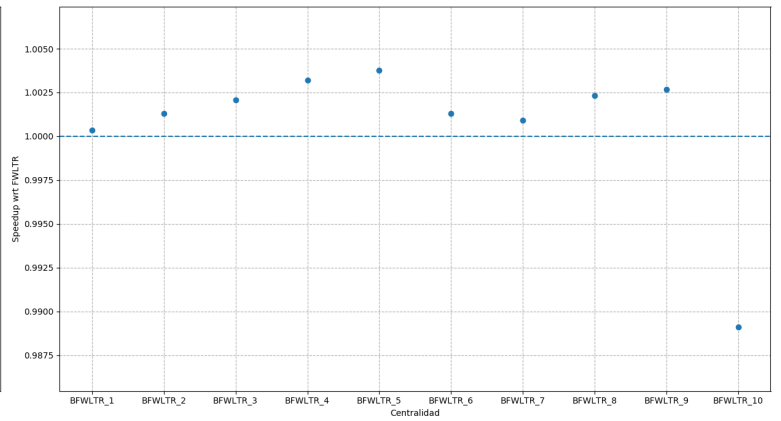
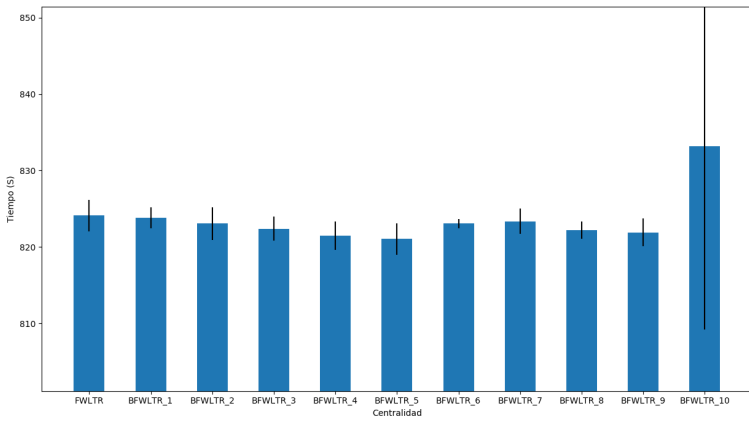


BLTR vs LTR

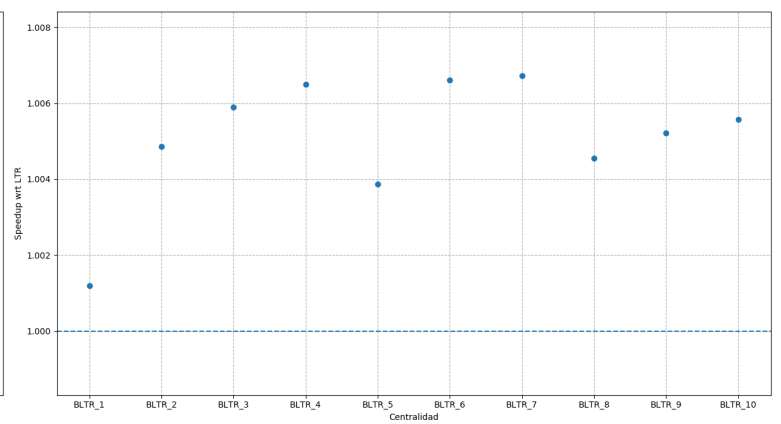
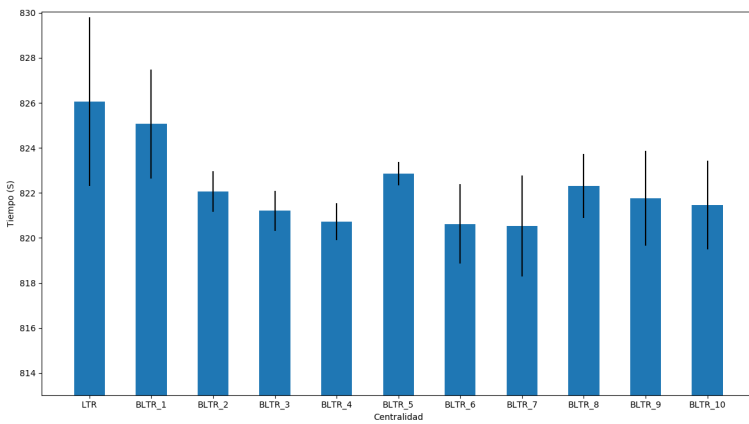
## C.2.8. Texas



BWLTR vs BBWLTR

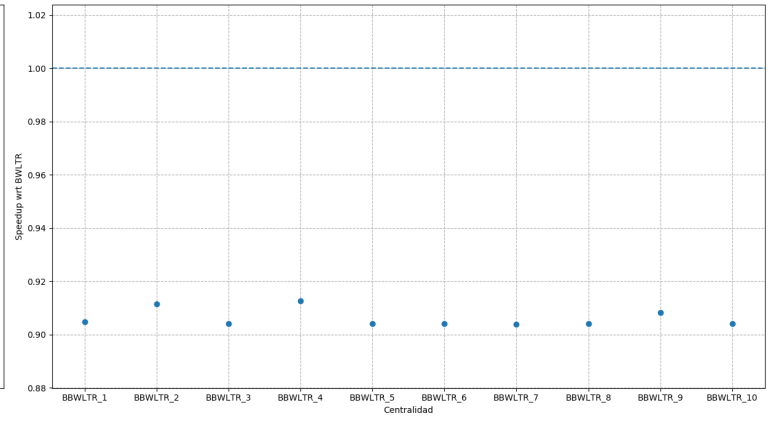
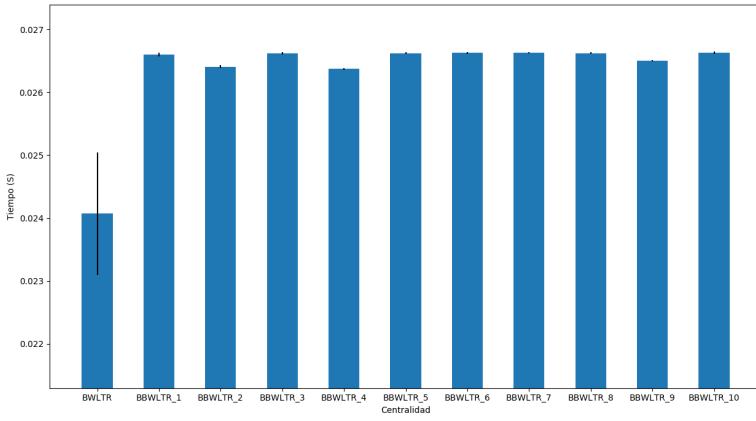


FWLTR vs BFWLTR

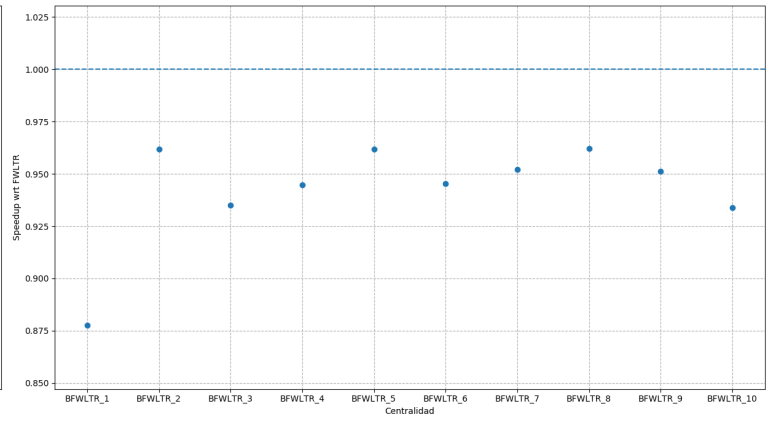
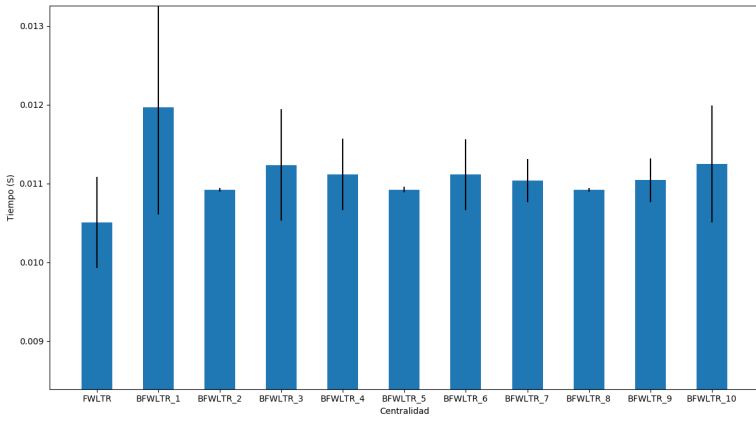


LTR vs BLTR

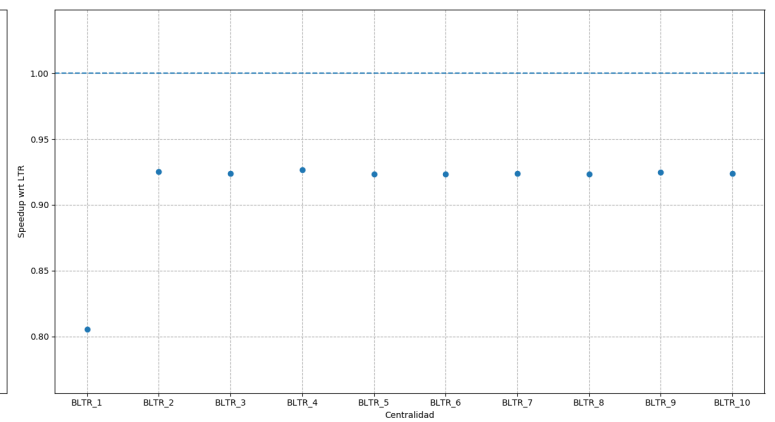
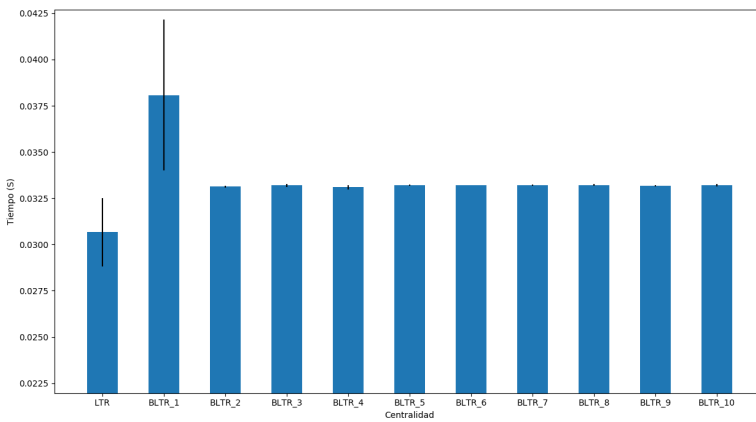
## C.2.9. Wikipedia



BWLTR vs BBWLTR



FWLTR vs BFWLTR



LTR vs BLTR