



Universitat Rovira i Virgili (URV) y Universitat Oberta de Catalunya (UOC)

Máster en Ingeniería Computacional y Matemática

Trabajo Final de Máster

Centralidad de intermediación flexible sobre redes de transporte urbanas

Nombre Estudiante: Martín González Soto

Nombre Director/a: Sergio Gómez, Albert Solé-Ribalta

Fecha de Entrega: 20/06/2021

El Dr. Sergio Gómez y el Dr. Albert Solé-Ribalta, certifica que el estudiante Martín González Soto ha elaborado el trabajo bajo su dirección y autoriza la presentación de esta memoria para la su evaluación.

Firma del director/a:



Esta obra está sujeta a una licencia de Reconocimiento-No Comercial-SinObraDerivada [3.0 España de Creative Commons](https://creativecommons.org/licenses/by-nc-nd/3.0/es/)

Licencias alternativas

(elegir alguna de las siguientes y sustituir la de la página anterior)

A) Creative Commons:



Esta obra está sujeta a una licencia de Reconocimiento-NoComercial-SinObraDerivada [3.0 España de Creative Commons](#)



Esta obra está sujeta a una licencia de Reconocimiento-NoComercial-CompartirIgual [3.0 España de Creative Commons](#)



Esta obra está sujeta a una licencia de Reconocimiento-NoComercial [3.0 España de Creative Commons](#)



Esta obra está sujeta a una licencia de Reconocimiento-SinObraDerivada [3.0 España de Creative Commons](#)



Esta obra está sujeta a una licencia de Reconocimiento-CompartirIgual [3.0 España de Creative Commons](#)



Esta obra está sujeta a una licencia de Reconocimiento [3.0 España de Creative Commons](#)

B) GNU Free Documentation License (GNU FDL)

Copyright © AÑO TU-NOMBRE.

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.

A copy of the license is included in the section entitled "GNU Free Documentation License".

C) Copyright

© (el autor/a)

Reservados todos los derechos. Está prohibido la reproducción total o parcial de esta obra por cualquier medio o procedimiento, comprendidos la impresión, la reprografía, el microfilme, el tratamiento informático o cualquier otro sistema, así como la distribución de ejemplares mediante alquiler y préstamo, sin la autorización escrita del autor o de los límites que autorice la Ley de Propiedad Intelectual.

Título del trabajo:	<i>Centralidad de intermediación flexible sobre redes de transporte</i>
Nombre del autor:	Martín González Soto
Nombre del director/a:	Dr Albert Solé-Ribalta Dr. Sergio Gómez
Fecha (mm/aaaa):	20/06/2021
Titulación:	Máster en Ingeniería Computacional y Matemática
Área del Trabajo Final:	<i>Redes Complejas</i>
Idioma del trabajo:	<i>Castellano</i>
Palabras clave: <small>Máximo 3 palabras clave (propuestas por el autor del TFM o en base a listados, tesauros, etc. y validadas por el director)</small>	<ol style="list-style-type: none"> 1. Redes Complejas 2. Sistemas Urbanos 3. Sistemas Complejos
Resumen del Trabajo (máximo 250 palabras): <i>Con la finalidad, contexto de aplicación, metodología, resultados y conclusiones del trabajo.</i>	
<p>El diseño de sistemas urbanos sostenibles es un objetivo de la mayoría de las grandes ciudades. El crecimiento demográfico y los nuevos modelos de consumo han afectado al modo en que nos relacionamos con los ecosistemas urbanos. En este contexto, se han propuesto diferentes medidas que reduzcan los efectos derivados del tráfico de vehículos tanto a nivel económico como sanitario provocadas por la contaminación. Ante esta necesidad y la complejidad existente en este tipo de sistemas, surge la ciencia de las ciudades como una disciplina cuyo objetivo es entender la estructura y comportamiento de los sistemas urbanos.</p> <p>En este Trabajo de Fin de Máster se busca desarrollar un algoritmo que logre establecer qué puntos dentro de una ciudad presentan un mayor riesgo de congestión a partir de asunciones que permitan modelar de manera más realista el comportamiento de los usuarios de las redes de transporte. Para ello se ha modificado la definición y cálculo de la centralidad de intermediación con el objetivo de que este considere rutas subóptimas próximas a la mínima como rutas aceptables. Finalmente, se ha estudiado cómo esta nueva medida de centralidad se comporta aplicada sobre redes de transporte reales.</p>	

Abstract (in English, 250 words or less):

The design of sustainable urban systems is a goal of most large cities. Population growth and new consumption patterns have affected the way we relate to urban ecosystems. In this context, different measures have been proposed to reduce the economic and health effects of vehicle traffic caused by pollution. Given this need and the existing complexity in this type of system, the science of cities arises as a discipline whose objective is to understand the structure and behavior of urban systems.

In this Master's Thesis, the aim is to develop an algorithm to establish which points within a city present a higher risk of congestion based on assumptions that allow more realistic modeling of the behavior of users of transport networks. To this end, the definition and calculation of the betweenness centrality have been modified so that it considers suboptimal routes close to the minimum as acceptable routes. Finally, it has been studied how this new centrality measure behaves when applied to real transport networks.

Índice

Capítulo 1. Introducción.....	1
1.1 Contexto y justificación del trabajo.....	1
1.2 Objetivos del trabajo.....	2
1.3 Enfoque y método seguido.....	2
1.4 Breve resumen de productos obtenidos.....	3
1.5 Breve descripción de los otros capítulos de la memoria.....	3
Capítulo 2. Marco teórico.....	4
2.1 Teoría de redes.....	4
2.1.1 Definición.....	4
2.1.2 Algunas clases de redes relevantes para este trabajo.....	4
2.1.3 Recorridos, caminos y distancias.....	6
2.2 Centralidad.....	6
2.2.1 Definición.....	6
2.2.2 Algunas medidas de centralidad de uso extendido.....	6
2.2.3 Centralidad de intermediación.....	8
2.2.4 Aplicación de la centralidad de intermediación en sistemas urbanos.....	9
Capítulo 3. Desarrollo.....	10
3.1 Algoritmo de Brandes para la centralidad de intermediación.....	10
3.2 Algoritmo de Dijkstra.....	11
3.3 <i>Backtracking</i>	13
3.4 Modificación del algoritmo de Dijkstra.....	14
3.5 Validación del método propuesto.....	16
3.5.1 Descripción del experimento computacional.....	16
3.5.2 Proceso de evaluación.....	16
3.5.3 Resultados.....	17
3.6 Tiempo de ejecución.....	21
Capítulo 4. Resultados experimentales.....	22
4.1 Bases de datos y herramientas.....	22
4.2 Centralidad de intermediación flexible sobre redes de transporte.....	23
4.2.1 Descripción del experimento computacional.....	23
4.2.2 Proceso de evaluación.....	24
4.2.3 Ponderación de aristas por distancias.....	25
4.2.4 Ponderación de aristas por tiempo de viaje.....	30
4.2.5 Comparación con los valores de centralidad <i>random walk</i>	35
Capítulo 5. Conclusiones.....	37
Bibliografía.....	38
Anexos.....	41
Anexo A: Centralidad de intermediación flexible en Rabat, Marruecos.....	41
Anexo B: Centralidad de intermediación flexible en Ann Arbor, EEUU.....	49
Anexo C: Centralidad de intermediación flexible en Tallin, Estonia.....	57

Lista de figuras

Índice de figuras

Ilustración 1: Comparación de valores de centralidad para distintas medidas.....	7
Ilustración 2: Centralidad de intermediación.....	8
Ilustración 3: Comparación entre los valores esperados de centralidad y los valores obtenidos por el método propuesto para varios tamaños de grid 2d $\ell \times \ell$	17
Ilustración 4: Relación entre la cantidad de nodos por lado y la tolerancia límite a partir de la cual se comiencen a dar caminos mínimos de mayor número de aristas a las esperadas en el caso estándar.....	19
Ilustración 5: Comparación entre los valores esperados de centralidad y los valores obtenidos por el método propuesto para varios tamaños de grid 2d $\ell \times \ell$ corregidos. En la fila superior se muestran los resultados observados con la tolerancia anterior y en la fila inferior se muestran los resultados cuando se establece una tolerancia límite.....	20
Ilustración 6: Segundos necesarios para computar la centralidad de intermediación en un grid $\ell \times \ell$ con diferentes métodos y tolerancias.....	21
Ilustración 7: Relaciones entre los valores de centralidad de intermediación estándar frente a los valores de centralidad de intermediación flexible para varios valores de tolerancia ponderado las aristas con distancias entre nodos.....	25
Ilustración 8: Comparación de la distribución de los valores de centralidad obtenidos mediante el método estándar y el método propuesto para varios valores de tolerancia. .	25
Ilustración 9: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.....	27
Ilustración 10: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.....	28
Ilustración 11: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.....	29
Ilustración 12: Relaciones entre los valores de centralidad de intermediación estándar frente a los valores de centralidad de intermediación flexible para varios valores de tolerancia ponderado las aristas con tiempos de viaje entre nodos.....	30
Ilustración 13: Comparación de la distribución de los valores de centralidad obtenidos mediante el método estándar y el método propuesto para varios valores de tolerancia. .	30
Ilustración 14: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.....	32
Ilustración 15: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos	

nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.....	33
Ilustración 16: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.....	34
Ilustración 17: Relaciones entre los valores de centralidad de intermediación estándar frente a los valores de centralidad de intermediación flexible para varios valores de tolerancia ponderado las aristas con distancias entre nodos.....	35
Ilustración 18: Comparación de la distribución de los valores de centralidad obtenidos mediante el método estándar y el método propuesto para varios valores de tolerancia. Ponderado con distancias.....	35
Ilustración 19: Relaciones entre los valores de centralidad de intermediación estándar frente a los valores de centralidad de intermediación flexible para varios valores de tolerancia ponderado las aristas con tiempo de viaje entre nodos.....	36
Ilustración 20: Comparación de la distribución de los valores de centralidad obtenidos mediante el método estándar y el método propuesto para varios valores de tolerancia. Ponderado con tiempos de viaje.....	36
Ilustración 21: Relaciones entre los valores de centralidad de intermediación estándar frente a los valores de centralidad de intermediación flexible para varios valores de tolerancia ponderado las aristas con distancias entre nodos.....	41
Ilustración 22: Comparación de la distribución de los valores de centralidad obtenidos mediante el método estándar y el método propuesto para varios valores de tolerancia. .	41
Ilustración 23: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.....	42
Ilustración 24: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.....	43
Ilustración 25: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.....	44
Ilustración 26: Relaciones entre los valores de centralidad de intermediación estándar frente a los valores de centralidad de intermediación flexible para varios valores de tolerancia ponderado las aristas con tiempo de viaje entre nodos.....	45
Ilustración 27: Comparación de la distribución de los valores de centralidad obtenidos mediante el método estándar y el método propuesto para varios valores de tolerancia. .	45
Ilustración 28: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.....	46

Ilustración 29: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.....	47
Ilustración 30: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.....	48
Ilustración 31: Relaciones entre los valores de centralidad de intermediación estándar frente a los valores de centralidad de intermediación flexible para varios valores de tolerancia ponderado las aristas con distancias entre nodos.....	49
Ilustración 32: Comparación de la distribución de los valores de centralidad obtenidos mediante el método estándar y el método propuesto para varios valores de tolerancia. .	49
Ilustración 33: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.....	50
Ilustración 34: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.....	51
Ilustración 35: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.....	52
Ilustración 36: Relaciones entre los valores de centralidad de intermediación estándar frente a los valores de centralidad de intermediación flexible para varios valores de tolerancia ponderado las aristas con tiempos de viaje entre nodos.....	53
Ilustración 37: Comparación de la distribución de los valores de centralidad obtenidos mediante el método estándar y el método propuesto para varios valores de tolerancia. .	53
Ilustración 38: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.....	54
Ilustración 39: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.....	55
Ilustración 40: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.....	56

Ilustración 41: Relaciones entre los valores de centralidad de intermediación estándar frente a los valores de centralidad de intermediación flexible para varios valores de tolerancia ponderado las aristas con distancias entre nodos.....	57
Ilustración 42: Comparación de la distribución de los valores de centralidad obtenidos mediante el método estándar y el método propuesto para varios valores de tolerancia. .	57
Ilustración 43: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.....	58
Ilustración 44: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.....	59
Ilustración 45: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.....	60
Ilustración 46: Relaciones entre los valores de centralidad de intermediación estándar frente a los valores de centralidad de intermediación flexible para varios valores de tolerancia ponderado las aristas con distancias entre nodos.....	61
Ilustración 47: Comparación de la distribución de los valores de centralidad obtenidos mediante el método estándar y el método propuesto para varios valores de tolerancia. .	61
Ilustración 48: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.....	62
Ilustración 49: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.....	63
Ilustración 50: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.....	64

Capítulo 1

Introducción

1.1 Contexto y justificación del trabajo

En la actualidad, la sostenibilidad de los sistemas urbanos representa una preocupación a nivel global. En este contexto nace la ciencia de las ciudades, una disciplina que busca entender las ciudades y su diseño no como artefactos sino como sistemas complejos constituidos por flujos y redes (Batty, 2017). En el proceso de discernimiento de los sistemas urbanos la optimización de las redes de transporte urbanas se posiciona como una tarea prioritaria y de especial interés hoy en día.

Ante esta nueva perspectiva, la teoría de redes emerge como la solución natural para analizar las interacciones y estudiar el comportamiento de los agentes que operan dentro de la red. La teoría de redes complejas permite modelar, entender cómo y porqué algunos fenómenos típicos del tráfico de vehículos se produce en las redes de carreteras urbanas.

El análisis de la centralidad de las redes de transporte urbanas es clave a la hora de entender el comportamiento del tráfico (Guimerà et al., 2002). La centralidad de intermediación o *betweenness centrality* se posiciona como un indicador valioso para medir la propensión de un punto a congestionarse. La importancia de esta medida radica en su naturaleza, la *betweenness centrality* se centra en estudiar la importancia de un nodo en función de cuántos caminos mínimos lo atraviesan. Debido a la capacidad de modelar de forma precisa la heurística que rige cómo se desplazan las personas en las ciudades, esta medida de centralidad ha sido ampliamente empleada para estudiar el tráfico en redes urbanas.

Pese a ser un buen indicador de riesgo de embotellamiento, se ha observado que algunos usuarios de las redes de transporte toman rutas alternativas de mayor longitud (Quercia et al., 2014). Junto a este comportamiento es posible incluir que los usuarios de las redes de transporte no toman decisiones perfectamente racionales en lo relativo a la planificación de rutas.

Un problema que presenta la centralidad de *betweenness* es que no contempla esa flexibilidad para modelar la toma de decisiones de los conductores. La centralidad de intermediación únicamente considera un camino mínimo entre dos puntos, permitiendo la posibilidad de que existan varios caminos mínimos si y sólo si todos ellos tienen exactamente la misma longitud, algo extremadamente poco usual.

Pese a que los conductores toman decisiones racionales a la hora de planificar rutas, existe cierta variabilidad sobre la ruta óptima a escoger que desde un punto de vista práctico, tanto de tiempo como de distancia, no es significativo. Como se ha mencionado, la centralidad de intermediación no es capaz de capturar esa variabilidad y por tanto no consigue modelar una toma de decisiones de los usuarios de las redes de transporte más acorde con la realidad.

1.2 Objetivos del trabajo

Presentada la problemática existente en la modelización del comportamiento de los usuarios de las redes de carreteras urbanas, se establece como objetivo para este trabajo de fin de máster el desarrollo de una nueva medida de centralidad que permita considerar como caminos válidos ciertos caminos que estén dentro de un margen de aceptación definido por una tolerancia y el camino mínimo encontrado.

De esta forma se dispondría de una medida de centralidad flexible con la capacidad de otorgar una mayor importancia a caminos subóptimos que tengan una distancia próxima al camino mínimo encontrado.

1.3 Enfoque y método seguido

Para lograr el objetivo principal se establecer una serie de metas parciales que permitan desglosar las tareas para alcanzar el objetivo por etapas.

- i. Modificar la definición y el cálculo de la centralidad de intermediación para que este considere a un camino como mínimo si cae dentro de un margen de aceptación respecto al camino mínimo observado.
- ii. Validar los resultados de centralidad sobre un *toy model*.
- iii. Estudiar los efectos de introducir flexibilidad en la medición de la centralidad.
- iv. Comparar los valores de centralidad obtenidos para la nueva medida de centralidad con otros métodos existentes.

1.4 Breve resumen de productos obtenidos

El resultado obtenido de este trabajo es una medida de centralidad apta para la predicción del riesgo de embotellamiento de puntos de una red de carreteras urbanas. Como se ha comentado, el producto obtenido del trabajo ofrece la posibilidad de ajustar un valor de tolerancia para permitir cierta flexibilidad en la consideración de caminos mínimos dentro de una red.

Este producto obtenido tiene la ventaja potencial sobre otras medidas de centralidad en redes, la capacidad de modelar la toma de decisiones de los usuarios de las redes de transporte en ciudades de una forma más realista.

Aunque esta medida se haya diseñado específicamente para el ámbito del estudio de las redes de carreteras urbanas mediante redes complejas, el producto obtenido podría tener aplicación en otros campos de estudio.

Relativo a los productos obtenidos de este trabajo se pretende contribuir al paquete de código abierto de Python NetworkX, proporcionando la funcionalidad de desarrollada en este trabajo, una vez exista una publicación en una revista indexada que la avale.

1.5 Breve descripción de los otros capítulos de la memoria

El resto de la memoria se estructura de la siguiente forma.

- En el capítulo 2 se hace una introducción al marco teórico de la teoría de redes complejas sobre los cuales se asienta el trabajo. Se describen algunos conceptos fundamentales de redes y la propiedad de centralidad de estas.
- En el capítulo 3 detalla las características del algoritmo propuesto para obtener medidas de centralidad flexibles.
- En el capítulo 4 se presentan los resultados de los experimentos computacionales donde se estudia la capacidad del algoritmo presentado para medir la centralidad de los nodos en redes de transporte urbanas.
- En el capítulo 5 se exponen las conclusiones extraídas de los experimentos computacionales efectuados en el apartado anterior, así como las posibles líneas de trabajo futuras que emergen de este trabajo.

Capítulo 2

Marco teórico

Este primer capítulo está orientado a introducir conceptos fundamentales que serán necesarios para el desarrollo de la propuesta que se presenta en esta memoria. Primeramente se hará una introducción a la teoría de redes complejas tratando los temas de concepto de red, tipos de redes relevantes para este trabajo, recorridos y fenómenos observados en redes reales.

A continuación, se tratará el concepto de centralidad de una red y se enumeran algunas medidas de centralidad típicas para después introducir el concepto de centralidad de intermediación o *betweenness centrality*.

2.1 Teoría de redes

2.1.1 Definición

La teoría de redes es una disciplina que se centra en el estudio de las redes, abstracciones que representan relaciones entre objetos discretos. Desde el punto de la teoría de grafos un grafo se puede describir formalmente de la siguiente manera.

Una red o grafo $G = (V,E)$ es un par ordenado en el que V es un conjunto finito no vacío y E es un conjunto de pares no ordenados $\{u,v\}$ de elementos de V y $u \neq v$. Los elementos de V se denominan vértices, o nodos, de G y los de E se denominan aristas de G .

La teoría de redes dirige sus esfuerzos al estudio de redes complejas. Las redes complejas son redes con propiedades topológicas no triviales que son capaces de replicar fenómenos que ocurren en algunos sistemas complejos reales.

Un sistema complejo es un sistema compuesto por múltiples agentes que interactúan entre sí. Los sistemas complejos son sistemas cuyo comportamiento es difícil de modelar debido a las dependencias entre sus partes. Así, los sistemas complejos tienen propiedades que surgen de estas conexiones como la no linealidad, los bucles de retroalimentación y fenómenos emergentes.

2.1.2 Algunas clases de redes relevantes para este trabajo

A la hora de hablar de tipologías de redes es necesario establecer una serie de definiciones para distinguir entre tres tipos de redes a un nivel básico en función de la estructura de sus aristas. En el apartado anterior hemos definido el concepto de red también conocido como grafo o grafo simple.

Un multigrafo $G = (V,E)$ es un par ordenado en el que V es un conjunto finito no vacío, y E es un multiconjunto de pares no ordenados $\{u,v\}$ de elementos de V con $u \neq v$. Los multigrafos son grafos que permiten más de una arista entre pares de vértices.

Un pseudografo $G = (V,E)$ es un par ordenado en el que V es un conjunto finito no vacío y E es un multiconjunto de pares no ordenados $\{u,v\}$ de elementos de V . Un pseudografo es un grafo que además de permitir aristas múltiples entre pares de vértices permite aristas entre un vértice y el mismo. A estos vértices que tienen origen y destino sobre el mismo vértice se les denomina bucles.

Una red dirigida $G = (V,E)$ es un par ordenado en el que V es un conjunto finito no vacío y $E \subseteq V \times V$. Es decir, las aristas son pares ordenados de elementos de V llamados arcos

Una red ponderada es un par (G,W) donde $G = (V,E)$ es un grafo y W es una función $W : E \rightarrow \mathbb{R}$ que asigna pesos a las aristas del grafo.

Una vez definidas tres tipologías de red en función de la naturaleza de las aristas existentes en la red, es necesario definir cuál es la naturaleza de las redes de transporte urbanas. Desde un punto de vista estricto, una red de carreteras urbanas es un pseudografo dirigido ponderado. Esto es así ya que las estructuras de carreteras pueden resultar en bucles al modelar algunas construcciones viales como glorietas u otras estructuras en las autopistas como intercambiadores o cruces. Así mismo, es posible que surjan varias aristas para conectar dos puntos de una red.

Por lo general, al trasladar las redes urbanas reales a una abstracción observamos que una gran parte de las vías dentro de una ciudad tienen un sentido definido, con lo que es lógico que el producto de esta abstracción sea una red dirigida.

Por último, las redes urbanas tienen una naturaleza marcadamente ponderada ya que cada calle en un sistema urbano tiene un conjunto de propiedades como longitud, velocidad límite o tiempo de viaje esperado. Estas propiedades tienen una importancia decisiva en la toma de decisiones de los conductores a la hora de planificar rutas.

Sin embargo, en el proceso de abstracción de redes de transporte urbanas la forma más extendida de representación de redes de transporte es aquella que adopta las intersecciones entre vías como nodos y las aristas como las vías (Marshall et al., 2018), (Steadman, 2004). Esta representación se denomina *junction graph* y simplifica algunas propiedades de una red real generando un grafo no dirigido que puede contener bucles.

De aquí en adelante se usará el término general red o grafo, aunque en ocasiones puedan existir bucles en las redes con lo que bajo las definiciones establecidas serían considerados estrictamente pseudografos.

2.1.3 Recorridos, caminos y distancias

Un recorrido en un grafo $G = (V,E)$ es una secuencia de vértices v_1, v_2, \dots, v_n con la propiedad de que $\{v_i, v_{i+1}\} \in E$. Un recorrido de extremos v_1, v_n se denomina v_1-v_n recorrido.

Un recorrido es un itinerario si todas las aristas son diferentes. Se pueden destacar los siguientes tipos de itinerario:

- Un camino, si no se repiten los vértices.
- Un circuito, si es cerrado.
- Un ciclo es un circuito que, eliminando el primer vértice, también es un camino.

Dado un grafo conexo $G = (V,E)$, la distancia entre dos vértices $u, v \in V$ es la mínima de las longitudes de los caminos que conectan u y v :

$$d_{ij} = \min \{ \ell(C) \mid C \text{ es un } u-v \text{ camino} \}$$

El camino mínimo entre dos nodos es el camino asociado a la distancia entre ellos d_{ij} .

2.2 Centralidad

2.2.1 Definición

Dado un grafo $G = (V,E)$, una medida de centralidad para G es un par (f, R) compuesto de una función $f : V \rightarrow R^+$ y una relación de orden $R \subset I_f \times I_f$, en la que I_f es la imagen de f . Dado un vértice $v \in V$, nos referiremos a $f(v)$ como a la centralidad de v .

La centralidad busca determinar qué nodo o conjunto de nodos son los que tienen más relevancia dentro de la red. El problema es que la definición de importancia no es absoluta y está abierta a interpretaciones (Newman, 2018). A razón de esto existen varias medidas de centralidad que adoptan diferentes estrategias a la hora de definir la importancia dentro de una red.

2.2.2 Algunas medidas de centralidad de uso extendido

Como se ha mencionado, existen diferentes estrategias para medir la importancia de un nodo en la red. Quizá una de las más básicas es la centralidad de grado (Freeman, 1978), (Davis and Hornseth, 1967). Esta medida de centralidad define la importancia de un nodo en función de las conexiones que este tenga. Así, un nodo será más relevante dentro de la red cuantas más conexiones tenga.

Sea A la matriz de adyacencia del grafo $G = (V, E)$ donde $a_{ij} = 1$ si $\{i, j\} \in E$ y $a_{ij} = 0$ en otro caso, entonces definimos la centralidad de grado del nodo i como

$$C_c(i) = \sum_j a_{ij}$$

Esta medida de centralidad pese a ser bastante simple funciona como un buen indicador de importancia en ciertos contextos como es en el caso de las redes sociales. Otras medidas como la centralidad de Katz (Katz, 1953) o la centralidad de Hubbell (Hubbell, 1965) también exploran este enfoque de la centralidad.

Un problema que presenta la centralidad de grado es que evalúa todas las conexiones de manera uniforme, cuando en muchos contextos ciertas conexiones son más relevantes que otras. Bajo este enfoque se enmarca la centralidad de autovector, la centralidad de autovalor (Bonacich, 1972), (Bonacich, 1972), (Bonacich, 1987) considera más importantes las conexiones con nodos que son a su vez importantes. Este enfoque es compartido con PageRank (Brin and Page, 1998) originalmente empleado por Google para medir la relevancia de las páginas webs en función del número e importancia de los links que dirigen a esta.

Otro enfoque posible para medir la importancia de un nodo en una red es la distancia a la que este se encuentra del resto de nodos de la red. Así, un nodo será más importante dentro de la red cuando más cerca se encuentre del resto de nodos. Esta estrategia es empleada por la centralidad de cercanía (Sabidussi, 1966), (Bavelas, 1950).

Sea d_{ij} la distancia mínima entre el nodo i y el nodo j . La distancia mínima media entre i y el resto de nodos de una red

$$\ell_i = \frac{1}{n} \sum_j d_{ij}$$

Para obtener la centralidad de cercanía se emplea la inversa de la distancia mínima media. La centralidad de cercanía es una medida de centralidad que mide de forma natural la importancia de un nodo en una red en ciertos contextos (Newman, 2018). Puede ser un buen indicador de cara a medir con qué facilidad un rumor se transmite más rápidamente en una red social.

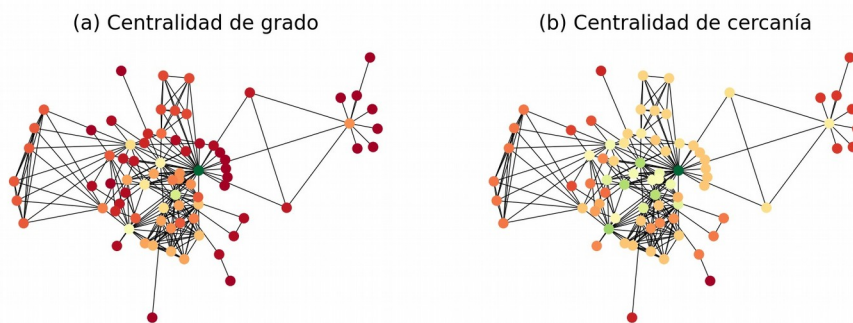


Ilustración 1: Comparación de valores de centralidad para distintas medidas

2.2.3 Centralidad de intermediación

Propuesta por Freeman (Freeman, 1977), la centralidad de intermediación o *betweenness centrality* es una medida de centralidad que adopta otra estrategia de cara a considerar la importancia en una red. El enfoque de esta medida se basa en medir la importancia de un nodo en la red en función de cuántos caminos mínimos entre otros dos nodos cualesquiera pasan por él. De este modo, en cuantos más caminos mínimos participe un nodo más importante será desde el punto de vista de la *betweenness centrality*.

Sea un grafo $G = (V, E)$ se define la centralidad de intermediación de un nodo $v \in V$ como

$$C_B(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

Donde σ_{st} se define como el número de caminos mínimos entre s y t . Por su parte, $\sigma_{st}(v)$ se define como el número de caminos mínimos entre s y t en los que participa v . Los valores obtenidos de la centralidad son relativos al número de caminos mínimos entre nodos de los que un nodo forma parte pero es posible normalizar estos valores.

La importancia de la *betweenness centrality* el contexto del estudio de las redes de carreteras urbanas viene marcada por la naturaleza de esta medida, la relación entre la centralidad y caminos mínimos. Desde el punto de vista de planificación de rutas, el objetivo es minimizar el tiempo o la distancia de viaje, por lo que es lógico pensar que una medida de centralidad que valore la importancia en función del número de caminos mínimos que atraviesan dicho punto será un buen indicador de la afluencia esperada por ese punto debido a su capacidad de aglutinar rutas rápidas.

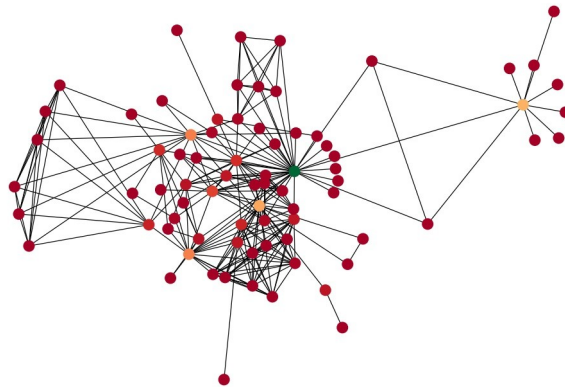


Ilustración 2: Centralidad de intermediación

Frente a los métodos anteriores, los nodos que actúan como puente tienen una centralidad más relevante que en los casos anteriores ya que los caminos mínimos entre zonas o polos del grafo se ven forzados a atravesarlos.

2.2.4 Aplicación de la centralidad de intermediación en sistemas urbanos

El uso de la teoría de redes como herramienta para modelar y estudiar las propiedades estructurales y dinámicas de las ciudades ha experimentado un crecimiento en los últimos años (Newman, 2003),(Boccaletti et al., 2006) debido a la disponibilidad de infraestructuras de cómputo más potentes, disponibilidad de repositorios de datos (Vieira and Haddad, 2020),(OpenStreetMap contributors, 2017) y el interés que suscita este campo desde el punto de vista socio-económico (Levy et al., 2010),(Currie and Walker, 2011),(Weisbrod et al., 2003),(Goodwin, 2004).

Bajo esta nuevo campo muchos autores se han centrado en el estudio de la relación entre la topología de las redes de carreteras urbanas y la dinámica relativa a la congestión. (Wen et al., 2017),(Holme, 2003),(Tadic et al., 2004),(Tadić et al., 2007), (Guo et al., 2019),(Solé-Ribalta et al., 2018).

La medida de centralidad de intermediación ha sido estudiada como un predictor del riesgo de congestión de un nodo dentro de una red que modela la red de carreteras urbanas de una ciudad debido a su naturaleza (Kazerani and Winter, 2009),(Xue et al., 2012),(Henry et al., 2019). Algunos autores han desarrollado modificaciones sobre esta medida con el fin de obtener resultados más realistas que permitiesen explicar la fenomenología observadas en las redes, modelando el comportamiento complejo de la toma de decisiones (Altshuler et al., 2011),(Gao et al., 2013),(Yan et al., 2006).

Capítulo 3

Desarrollo

En este capítulo se presenta la modificación propuesta para el cálculo de una medida de centralidad de intermediación flexible. El método propuesto en esta memoria se basa en la modificación del algoritmo de Brandes (Brandes, 2001) para el cálculo de la *betweenness centrality* de forma que se pueda establecer un margen de aceptabilidad para que aquellos caminos cercanos al valor mínimo sean considerados como tal.

Mediante este algoritmo se busca desarrollar métodos que permitan una valoración más realista de la tendencia de un punto dentro de una red de carreteras urbanas de sufrir embotellamientos. Esto es posible gracias a la capacidad de este algoritmo para considerar rutas próximas a la ruta óptima como rutas subóptimas que podrían ser consideradas por los conductores a la hora de planificar sus viajes.

3.1 Algoritmo de Brandes para la centralidad de intermediación

Previo al inicio de la introducción de los algoritmos para el cálculo de la centralidad de intermediación, se debe aclarar que se realizará la modificación del algoritmo de Brandes para redes ponderadas. Por tanto, se excluye del alcance del trabajo la modificación del algoritmo para redes no ponderadas debido a la naturaleza ponderada de las redes con las que se trabajan en este proyecto.

Sea un grafo $G = (V, E)$

- (1) Para cada nodo $v \in V$, se establece $C_B(v) = 0$.
- (2) Para cada nodo origen s en V .
 - (1) Computar los caminos mínimos entre el nodo origen s y el resto de nodos en V mediante el algoritmo de Dijkstra. El algoritmo de Dijkstra devuelve para cada nodo tomado como origen tres estructuras de datos: una tabla hash P que almacena los nodos predecesores de cada nodo en el camino mínimo entre el nodo de origen y el propio nodo, un vector S que almacena los nodos visitados y una tabla hash σ que almacena el número de caminos mínimos que atraviesan un nodo (consultar apartado 3.2).
 - (2) Una vez se realice la computación del algoritmo de Dijkstra para un nodo s se realiza el *backtracking* o acumulación, que con los outputs del paso previo se actualizarán los valores de centralidad, de cada nodo, en función de los caminos que pasen por el (consultar apartado 3.3).
- (3) Una vez se completen las iteraciones sobre todos los nodos del grafo se dispondrán de los valores de centralidad finales para cada nodo. Es posible escalar estos valores.


```

Brandes(G, s, escalar = False)
    centralidad = {}
    para cada v ∈ V hacer
        centralidad[v] = 0.0
    para cada s ∈ V hacer
        S, P, sigma = Dijkstra(G, s)
        centralidad = Backtracking(centralidad, S, P, sigma, s)

    devolver centralidad

```

3.2 Algoritmo de Dijkstra

Debido a que el interés del proyecto se centra en la centralidad de intermediación teniendo en cuenta aristas ponderadas, el algoritmo de Dijkstra tiene un papel determinante en la obtención de los caminos mínimos entre puntos. La idea del algoritmo de Dijkstra es recorrer todos los nodos que pertenecen a un grafo conexo partiendo de un nodo inicial sobre el que se calculan las distancias al resto. Para cada nodo visitado se obtendrán los nodos adyacentes y las distancias de estos respecto al nodo de partida. Una vez completado este proceso, se repetirá con el nodo no visitado con menor distancia al nodo de partida registrado hasta que no queden nodos sin visitar.

Para el caso que se presenta es necesario incluir algunas estructuras de datos adicionales para adaptar la idea del algoritmo de Dijkstra a las necesidades de la centralidad de intermediación.

Sea un grafo $G = (V, E)$ no dirigido ponderado y sea $s \in V$ el nodo de origen del cual se parte para obtener las distancias del resto de nodos $\{v \in V \mid v \neq s\}$. Es necesario considerar las siguientes estructuras de datos para adaptarlo a las demandas de la centralidad de intermediación.

- i. Una tabla hash P que almacena para cada nodo en la red la lista de nodos que preceden a este en el camino mínimo desde s al mencionado nodo. Inicialmente, esta tabla hash asignará a cada nodo presente en la red un vector vacío.
- ii. Un vector S que almacena los nodos ya visitados. Este vector se inicializa vacío.
- iii. Una tabla hash sigma que almacena el número de caminos mínimos que existen hasta un nodo desde s. Esta tabla hash se inicializa asignando a todos los nodos de la red 0 excepto a s que se le asigna 1.
- iv. Una tabla hash D que almacena la distancia mínima definitiva desde s hasta un determinado nodo. Esta tabla hash se inicializa vacía.
- v. Una tabla hash seen que almacena para cada nodo en la red la distancia mínima observada provisional a este desde el nodo inicial s. Esta tabla hash se iniciará con el par clave valor (s, 0) estableciendo que la distancia a s desde el nodo inicial s es cero.
- vi. Una cola de prioridad Q que almacena el orden de los nodos a visitar en función de la distancia de este al nodo s. Esta cola se inicializa con el par (0, s, s).

S, P y sigma se emplearán en los siguientes pasos para incrementar la centralidad de los nodos predecesores presentes entre un nodo n y s. El número de caminos sigma se empleará para ponderar el incremento de esta centralidad. Una vez inicializada las estructuras a emplear se procede a realizar las siguientes operaciones mientras la cola Q de prioridad no esté vacía.

- (1) Se extraerá el nodo de menor distancia provisional seen almacenado en la cola de prioridad. Si este nodo no ha sido visitado se continuará con el proceso, en otro caso de extraerá el siguiente nodo de Q.
- (2) Se añade el nodo al vector S de nodos visitados.
- (3) Para cada nodo adyacente al nodo extraído:
 - (1) Computar la distancia al nodo adyacente.
 - (2) Si el nodo no ha sido visitado y la distancia provisional registrada es mayor a la que se acaba de calcular entonces se guarda esta como nueva distancia minima provisional, se coloca el nodo en la cola Q y se reinician la lista de predecesores y sigma.
 - (3) Si el nodo está a la misma distancia que la provisional, entonces se registra como predecesor alternativo y se incrementa el valor de sigma.

Dijkstra (G, s)

```

S = {}
D = {}
para cada v ∈ V hacer
    P[v] = {}
    sigma[v] = 0
seen[s] = 0
sigma[s] = 1
añadir(Q, (D[s], s, s))

mientras Q no esta vacía hacer
    distancia, predecesor, v = extraer_mínimo(Q)
    si v ∈ D hacer
        saltar
        sigma[v] = sigma[v] + sigma[predecesor]
        D[v] = distancia
        S = S U {v}

    para cada w ∈ {w ∈ V | (v,w) ∈ E } hacer // nodos adyacentes
        dist_w = D[v] + peso(v, w)

        // predecesor mínimo
        si w ∉ D y (w ∉ seen o dist_w < seen[w]) hacer
            seen[w] = dist_w
            añadir(Q, (D[s], v, w))
            P[w] = {v}
            sigma[w] = 0

        o si dist_w == seen[w] hacer // predecesor alternativo
            P[w] = P[w] U {v}
            sigma[w] = sigma[w] + sigma[v]

devolver S, P, sigma
  
```

3.3 Backtracking

Una vez obtenidas las estructuras de datos que contienen la información de los predecesores, nodos visitados y los caminos mínimos que pasan cada nodo, se procede a realizar la acumulación para computar la centralidad de intermediación parcial de cada nodo que se irá actualizando hasta recorrer todos los nodos.

Sea un grafo $G = (V, E)$ no dirigido ponderado y sea $s \in V$ el nodo de origen sobre el cual se parte para obtener las distancias del resto de nodos $\{v \in V \mid v \neq s\}$. Diponinedo de las siguientes estructuras de datos ya inicializadas con los valores de los pasos previos.

- i. Una tabla hash P que almacena para cada nodo en la red la lista de nodos que preceden a este en el camino mínimo desde s al mencionado nodo.
- ii. Un vector S que almacena los nodos ya visitados.
- iii. Una tabla hash σ que almacena el número de caminos mínimos que existen hasta un nodo desde s .

(1) Se inicializa una tabla de hash δ para cada uno de los nodos contenidos en S con valor cero.

(2) Mientras el conjunto S no esté vacío:

(1) Se extrae un nodo w

(2) Se calcula el coeficiente de dicho nodo. El coeficiente dividirá la importancia de dicho nodo en función del número de predecesores que este tenga.

(3) Para cada predecesor de w se actualiza el valor de δ

(4) Si el nodo w que se estudia no es el nodo de origen se actualiza su centralidad

```
Backtracking (centralidad, S, P, sigma, s)
  delta = {}
  para cada v ∈ S hacer
    delta[v] = 0
  mientras S no esta vacía hacer
    w = extraer(S)
    coeficiente = (1 + delta[w]) / sigma[w]
    para cada predecesor ∈ P[w] hacer
      delta[v] = delta[v] + sigma[v] * coeficiente
    si w ≠ s hacer
      centralidad[w] = centralidad[w] + delta[w]

  devolver centralidad
```

3.4 Modificación del algoritmo de Dijkstra

Para lograr obtener una medida de centralidad flexible es necesario considerar como caminos mínimos caminos que tengan una distancia cercana a la mínima, para ello se propone el uso de tolerancias. De este modo, un camino no mínimo se considerará como tal si entra dentro de un rango de aceptación. Para lograr esto es necesario modificar el algoritmo de Dijkstra.

Sea un grafo $G = (V, E)$ no dirigido ponderado, sea $s \in V$ el nodo de origen sobre el cual se parte para obtener las distancias del resto de nodos $\{ v \in V \mid v \neq s \}$ y sea t un valor de tolerancia para el cual cualquier camino entre dos puntos será aceptable si

$$d_{s,e} \leq d_{\min(s,e)} \cdot (1+t)$$

Es necesario considerar las siguientes estructuras de datos para adaptarlo a las demandas de la centralidad de intermediación flexible.

- i. Una tabla hash P que almacena para cada nodo en la red un vector de pares (predecesor, distancia). Inicialmente se asignará a cada nodo un vector vacío.
- ii. Un vector S que almacena los nodos ya visitados. Este vector se inicializa vacío.
- iii. Una tabla hash σ que almacena un conjunto de pares (predecesor, caminos mínimos aportados) que existen hasta un nodo desde s . Esta tabla hash se inicializa asignando a todos los nodos de la red un conjunto vacío excepto a s que se le asigna un par $(s, 0)$.
- iv. Una tabla hash D que almacena la distancia mínima definitiva desde s hasta un determinado nodo. Esta tabla hash se inicializa vacía.
- v. Una tabla hash seen que almacena para cada nodo en la red la distancia mínima observada provisional a este desde el nodo inicial s . Esta tabla hash se iniciará con el par clave valor $(s, 0)$ estableciendo que la distancia a s desde el nodo inicial s es cero.
- vi. Una cola de prioridad Q que almacena el orden de los nodos a visitar en función de la distancia de este al nodo s . Esta cola se inicializa con $(0, s, s)$.

Las estructuras de datos empleadas son análogas a las empleadas en la versión original, las diferencias principales son las añadidas en P y σ . La razón de estos cambios se debe a que para filtrar los predecesores no válidos al añadir un nuevo predecesor mínimo, es necesario descartar las aportaciones de caminos mínimos de estos predecesores. Por tanto, es necesario mantener esa información para poder trazar su origen.

Una vez inicializadas las estructuras a emplear se procede a realizar las siguientes operaciones mientras la cola Q de prioridad no esté vacía.

- (1) Se extraerá el nodo de menor distancia provisional seen almacenado en la cola de prioridad. Si este nodo no ha sido visitado se continuará con el proceso, en otro caso de extraerá el siguiente nodo de Q.
- (2) Se añade el nodo al vector S de nodos visitados.
- (3) Para cada nodo adyacente al nodo extraído:
 - (1) Computar la distancia al nodo adyacente.
 - (2) Si el nodo no ha sido visitado y la distancia provisional registrada es mayor a la que se acaba de calcular entonces se guarda esta como nueva distancia mínima provisional, se coloca el nodo en la cola Q. Se filtran y se mantienen aquellos predecesores que aportan un camino mínimo con una distancia dentro de tolerancia. Se eliminan aquellos caminos aportados por nodos que ya no son predecesores.
 - (3) Si el nodo no es mínimo pero la distancia aportada está dentro de la tolerancia se añade al conjunto de predecesores y se anotan los caminos aportados.

Dijkstra (G, s, t)

S = {}

D = {}

para cada v ∈ V **hacer**

 P[v] = {}

 sigma[v] = {}

seen[s] = 0

sigma[s] = (s, 1)

añadir(Q, (D[s], s, s))

mientras Q no esta vacía **hacer**

 distancia, predecesor, v = extraer_mínimo(Q)

si v ∈ D **hacer**

saltar

 D[v] = distancia

 S = S U {v}

para cada w ∈ {w ∈ V | (v,w) ∈ E } **hacer** // nodos adyacentes a v
 dist_w = D[v] + peso(v, w)

 // predecesor mínimo

si w ∉ D y (w ∉ seen o dist_w < seen[w]) **hacer**

 seen[w] = dist_w

 añadir(Q, (D[s], v, w))

 P[w] = {p ∈ P[w] | D[p] + peso(p, w) ≤ dist_w * (1+t)}

 sigma[w] = {s ∈ sigma[w] | s ∈ P[w]}

 sigma[w] = sigma[w] U {(v, suma_caminos(sigma[v]))}

 // predecesor aceptable

o si dist_w ≤ seen[w] * (1+t) **hacer**

 P[w] = P[w] U {(v, dist_w)}

 sigma[w] = sigma[w] U {(v, suma_caminos(sigma[v]))}

devolver S, P, sigma

3.5 Validación del método propuesto

3.5.1 Descripción del experimento computacional

Tras el desarrollo de la modificación del algoritmo, se procede a diseñar un experimento de validación para evaluar el funcionamiento del algoritmo sobre un entorno controlado. El test de validación busca comparar los resultados de centralidad estándar teóricos y los resultados obtenidos mediante el algoritmo desarrollado en este proyecto.

Para ello, se parte de una red grid 2d de tamaño $\ell \times \ell$ sin ponderar sobre la cual se procede a obtener los valores de centralidad para sus nodos. Una vez obtenidos los valores de centralidad de intermediación con el algoritmo estándar, se procede a introducir ruido en el peso de las aristas del grafo entre unos valores aleatorios dentro de un rango conocido $\forall e \in E, w(e) \in \{1.00, 1.01, \dots, 1.05\}$.

Una vez introducido el ruido, se procede a computar la centralidad del grafo mediante el algoritmo propuesto estableciendo una tolerancia del 5%. En estas circunstancias, debido a que los valores introducidos de los pesos de las aristas no son mayores al 5% respecto al valor mínimo, todas las aristas serán consideradas de igual forma por el algoritmo propuesto.

3.5.2 Proceso de evaluación

Para evaluar la capacidad del método propuesto a la hora de eliminar el ruido que se introduce a la red, y de este modo computar las centralidades permitiendo cierta flexibilidad, es necesario establecer algún modo de medir su desempeño.

Para medir el desempeño del método propuesto se miden dos métricas. Por un lado se dispone el error relativo medio entre los valores obtenidos por el método propuesto y los valores de centralidad estándar de cada nodo en la red.

$$\varepsilon_v = \frac{|B_c(v)_{flexible} - B_c(v)_{estandar}|}{B_c(v)_{estandar}}$$

$$\varepsilon = \frac{1}{n} \sum_{v \in V} \varepsilon_v$$

Por otro lado se establece la métrica R^2 , que no es concluyente pero nos permite evaluar de un modo visual como se distribuyen los valores de centralidad computados frente a los esperados. Estos valores deben estar ubicados sobre la recta $y=x$ ya que el resultado esperado es la igualdad para ambos valores. Por tanto, el ajuste medido por R^2 a dicha recta es un indicador no concluyente del desempeño. Esto es así ya que por si solo no refleja el ajuste a la recta considerada, por ello se acompaña esto con un gráfico de dispersión que proporciona dicha información.

3.5.3 Resultados

A continuación, se muestran los resultados obtenidos para los tamaños de grid siguientes

- 2d grid 10x10
- 2d grid 20x20
- 2d grid 30x30
- 2d grid 40x40
- 2d grid 50x50

En la siguiente figura (Ilustración 3) se muestra la comparación entre la centralidad de cada nodo obtenida mediante la centralidad de intermediación estándar y la centralidad de intermediación flexible.

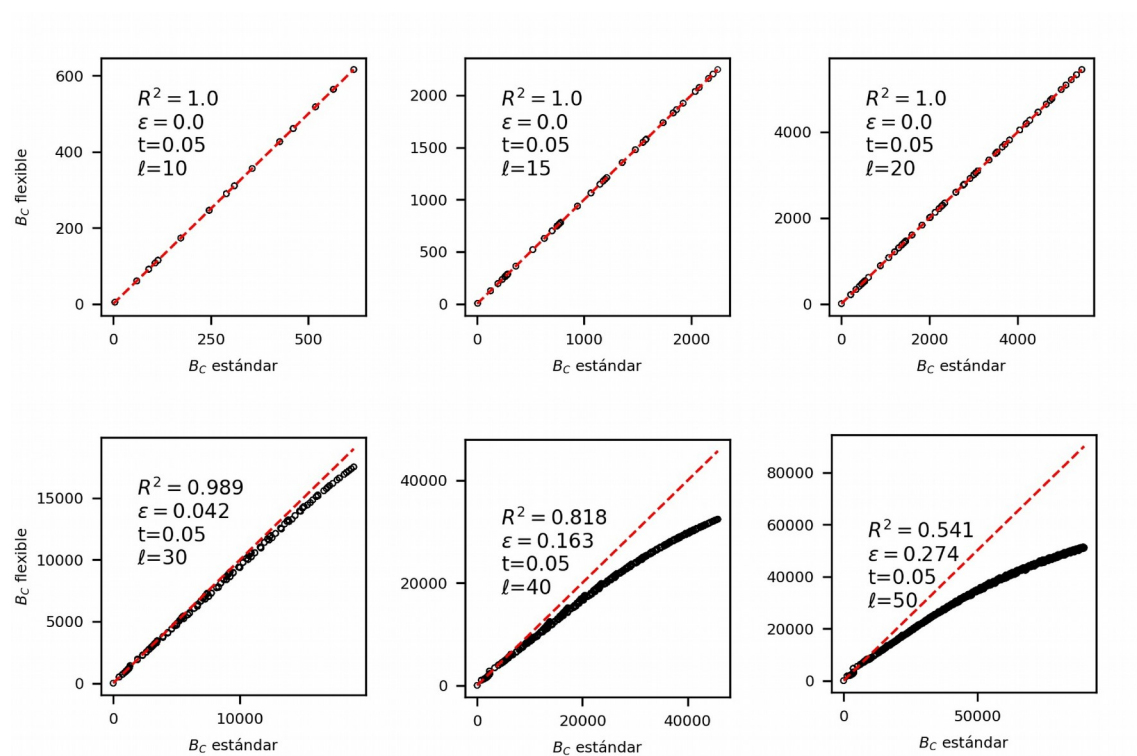


Ilustración 3: Comparación entre los valores esperados de centralidad y los valores obtenidos por el método propuesto para varios tamaños de grid 2d $l \times l$

En las comparaciones mostradas se observa un buen desempeño que reporta el método para descartar el ruido introducido en la red. Esta capacidad de eliminar el ruido y considerar todas las aristas como iguales es una propiedad deseable para el algoritmo propuesto. Esta propiedad se traduciría en una controlada flexibilidad para considerar caminos mínimos aquellos que no lo son estrictamente, pero desde un punto de vista práctico podrían considerarse así.

Puede observarse un efecto a medida que se incrementa el tamaño de la red. El algoritmo introduce error al obtener la centralidad de intermediación estándar mediante el algoritmo flexible. A medida que se incrementa el tamaño del grafo, el número de nodos crece de forma exponencial así como el número de aristas. Debido a que los pesos de las aristas son generados aleatoriamente, pueden aparecer caminos mínimos que atraviesan un número mayor de aristas de las que atravesaría en el caso de que todas las aristas pesaran igual, en cuyo caso esto no podría suceder.

Sea $W \sim U(w_{\min}, \dots, w_{\max})$ una variable aleatoria discreta con distribución uniforme que modela los pesos de las aristas de un grafo $G=(V,E)$ y sea $E[W]$ el valor esperado de W . Se necesita que un camino atraviese λ_0 aristas de peso w_{\min} para igualar un camino de $(\lambda - \delta)$ aristas de peso $w_{\min} + e$, $0 < e \leq w_{\max} - w_{\min}$

$$\lambda_0(\delta, e) = \frac{\delta(w_{\min} + e)}{e}$$

Es posible obtener cual es el peso medio de la arista necesario para que un camino que atraviesa λ aristas iguale a un camino que atraviesa $(\lambda - \delta)$ aristas de peso $E[W]$

$$w_i(\delta, \lambda) = (1 - \frac{\delta}{\lambda}) E[W]$$

Por tanto, se necesitan λ_{\min} aristas para igualar una distancia con una diferencia de δ para un camino de w_i de peso medio

$$\lambda_{\min}(\delta, w_i) = \frac{\delta E[W]}{E[W] - w_i}$$

Se puede obtener cual es el número mínimo de aristas que debe atravesar para que este fenómeno se empiece a producir. Estableciendo que el fenómeno comenzará a ocurrir cuando exista un camino mínimo que iguale a un camino mínimo recorriendo una arista más se establece que $\delta = 1$. A su vez el primer camino que iguale a un camino mínimo recorriendo $\delta+1$ tendrá el peso mínimo por arista w_{\min} permitido. Con esto es posible generalizar el tamaño límite a partir del conjunto de pesos de un grid

$$\lambda_{\text{threshold}}(W) = \frac{E[W]}{E[W] - w_{\min}}$$

Dado que $G=(V,E)$ es un grid, el máximo camino mínimo posible es $2(\ell-1)$ con lo que es posible obtener cual es el tamaño del grid máximo permitido antes de que comience a reproducirse este fenómeno

$$\ell_{threshold}(W) < \frac{\lambda_{threshold}}{2} - 1 = \frac{E[W]}{2(E[W] - w_{min})} - 1$$

Es posible obtener una expresión que relacione la tolerancia límite con el tamaño del grafo. De este modo se establece la tolerancia máxima permitida, a partir de la cual existen caminos mínimos que recorren un mayor número de aristas que las esperadas en el caso no ponderado. Sea $G=(V,E)$ un grid de tamaño $\ell \times \ell$ la tolerancia máxima permitida se define como:

$$t_{threshold}(\ell) = \frac{2}{2\ell + 1}$$

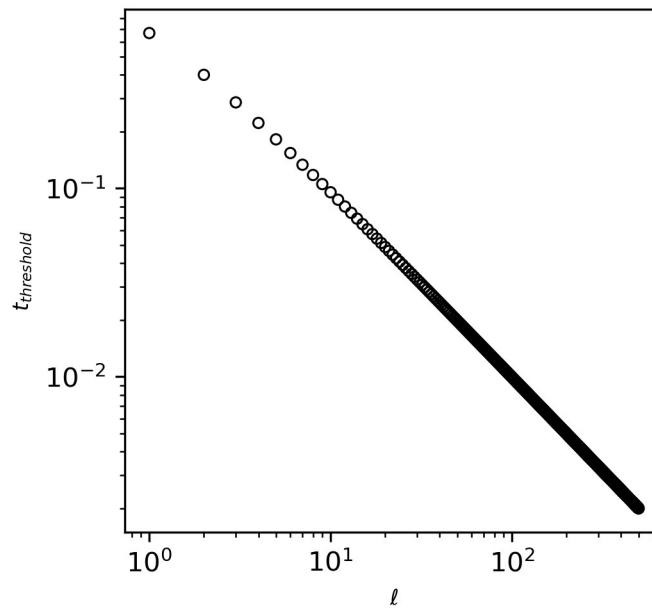


Ilustración 4: Relación entre la cantidad de nodos por lado y la tolerancia límite a partir de la cual se cominezan a dar caminos mínimos de mayor número de aristas a las esperadas en el caso estándar

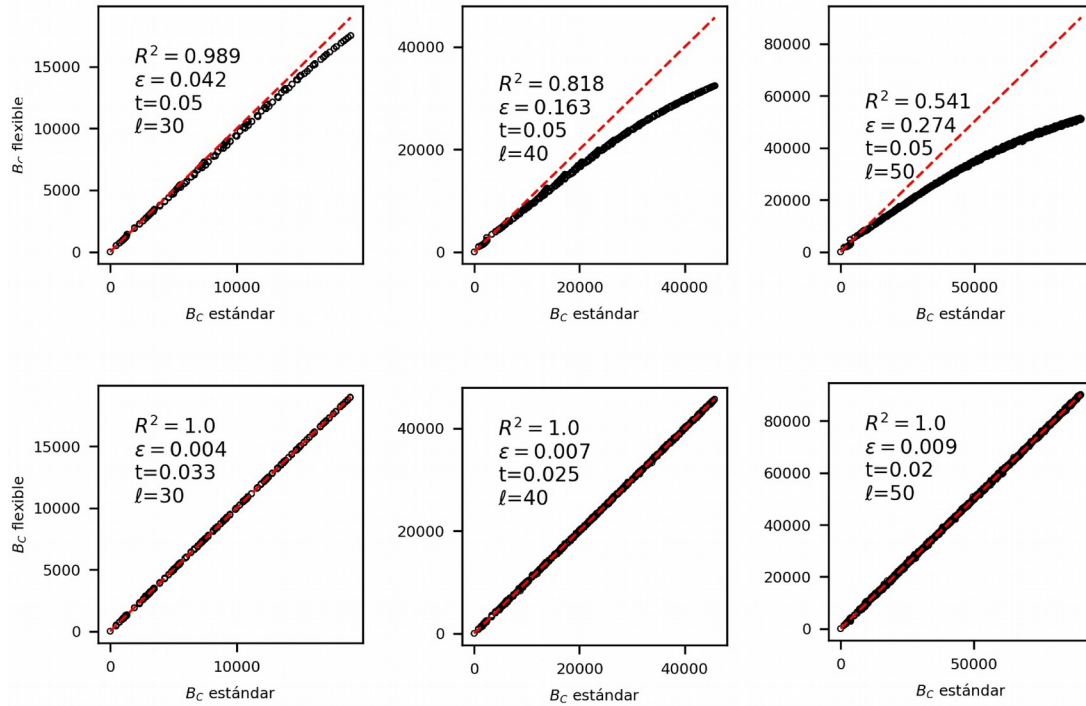


Ilustración 5: Comparación entre los valores esperados de centralidad y los valores obtenidos por el método propuesto para varios tamaños de grid 2d l x l corregidos. En la fila superior se muestran los resultados observados con la tolerancia anterior y en la fila inferior se muestran los resultados cuando se establece una tolerancia límite.

Puede considerarse como un valor de referencia para establecer un valor de tolerancia en una red de transporte urbana. Sea $G=(V,E)$ una red de transporte de n nodos se establece la tolerancia límite de nodos de una red como

$$t_{node\ threshold}(n) = \frac{2}{2\sqrt{n}+1}$$

Es posible llegar a otro valor de la tolerancia a través de las aristas de una red, sea m el número de aristas de una red se establece el valor de tolerancia límite de aristas como

$$t_{edge\ threshold}(m) = \frac{2}{2+\sqrt{1+2m}}$$

$$t_{threshold} = \min(t_{node\ threshold}, t_{edge\ threshold})$$

Nótese que la tolerancia límite extrapolada asume que los pesos de la red se distribuyen de forma uniforme. Por tanto, esta extrapolación puede no ser adecuada para ciertas topologías de red de carreteras urbana. Este valor de tolerancia ha de tomarse como un valor orientativo para realizar el ajuste de tolerancia deseado.

3.6 Tiempo de ejecución

En este apartado se presenta un análisis de los tiempos de ejecución para varios grids $2d$ $l \times l$ (Ilustración 6). El objetivo de este análisis es observar cómo se incrementa el tiempo de ejecución al incrementar la tolerancia considerada en el cómputo. Al incrementar la tolerancia entregada al método, más nodos serán considerados como predecesores aceptables de un nodo y en consecuencia más operaciones serán necesarias realizar a medida que se incrementa la tolerancia.

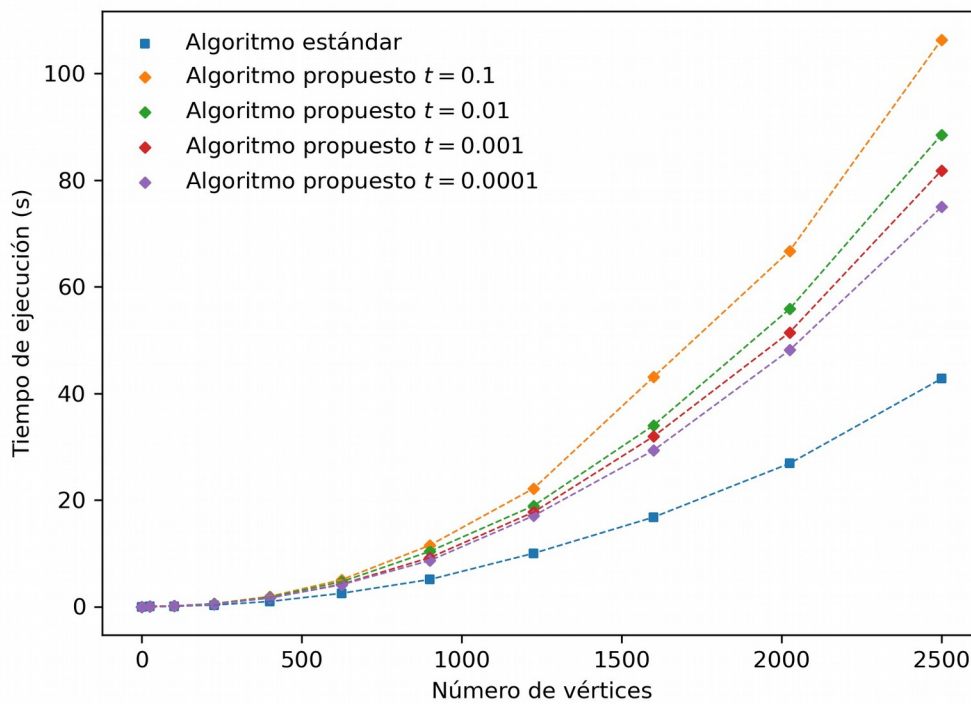


Ilustración 6: Segundos necesarios para computar la centralidad de intermediación en un grid $l \times l$ con diferentes métodos y tolerancias.

Capítulo 4

Resultados experimentales

En el siguiente capítulo se describen y justifican los experimentos computacionales llevados a cabo en el presente trabajo. Primeramente se introduce un experimento de validación del método propuesto que tiene por objetivo comprobar el funcionamiento del mismo sobre *toy models*. Seguidamente se procederá a evaluar el funcionamiento de la modificación propuesta sobre redes de transporte urbanas reales.

4.1 Bases de datos y herramientas

Para llevar a cabo los experimentos computacionales realizados en este trabajo ha sido necesario disponer de una serie de conjuntos de datos de redes reales de transporte urbanas con el fin de poder ejecutar sobre ellas el procedimiento diseñado y estudiar los resultados de dichos experimentos computacionales.

Para la obtención de datos de tipo geoespaciales se empleará la base de datos OpenStreetMap. OpenStreetMap es un proyecto de colaboración para crear un mapa editable libre global (OpenStreetMap contributors, 2017).

Para la obtención de los datos de OpenStreetMap se empleará OSMnx. OSMnx es un paquete de Python que permite descargar datos geoespaciales de OpenStreetMap y modelar, proyectar, visualizar y analizar redes de calles del mundo real y cualquier otra geometría geoespacial (Boeing, 2017).

La herramienta que se empleará en este proyecto para manipular redes complejas es NetworkX. NetworkX es un paquete del lenguaje Python para la exploración y el análisis de redes y algoritmos de redes (A. Hagberg et al., 2008). NetworkX implementa algoritmos para el cálculo de la centralidad de intermediación así como otras muchas herramientas que serán de gran utilidad para este proyecto.

4.2 Centralidad de intermediación flexible sobre redes de transporte

4.2.1 Descripción del experimento computacional

Tras la validación del método propuesto se procede a diseñar un segundo experimento donde se estudia el comportamiento del algoritmo sobre redes de transporte de ciudades reales. Este experimento busca observar cómo varían los valores calculados por la nueva medida de centralidad de intermediación flexible respecto a los valores estándares de centralidad.

El experimento computacional propuesto estudiará la variación de los valores de centralidad teniendo en cuenta dos posibles pesos a la hora de medir la importancia de un nodo desde un punto de vista del tráfico.

Por un lado, se ponderará cada arista por su distancia real, de modo que se calcule qué nodos son los que son atravesados por caminos mínimos en distancia. Por otro lado, se ponderará cada arista el tiempo que se toma en recorrerla. De esta segunda forma se calculará cuales son los nodos que son atravesados por más caminos rápidos. Ambas opciones son, a priori, buenas indicadoras de que nodos colapsarán antes por ser atravesados por rutas cortas o bien, rutas rápidas.

Los datos de longitud y tiempo de viaje de cada una de las aristas de la red de una ciudad son obtenidos mediante OSMnx, que como se mencionó en el primer apartado de este capítulo, se extraen de OpenStreetMap.

Para cada una de estas alternativas se computa la centralidad de intermediación estándar que servirá de referencia para evaluar la variación de la centralidad de cada nodo computada por el método propuesto. También se computa la centralidad de intermediación flexible para un rango de valores de tolerancia.

Con los resultados obtenidos se compara la relación entre los valores de centralidad estándar con los valores de centralidad de intermediación estándar mediante un diagrama de dispersión. También se muestra la distribución geográfica de los valores de centralidad de intermediación estándar frente a los de centralidad obtenidos mediante el método propuesto.

Adicionalmente, se compara de la misma forma los resultados obtenidos con los valores de centralidad aleatoria.

4.2.2 Proceso de evaluación

El objetivo del experimento es estudiar la capacidad del método propuesto para flexibilizar la consideración de caminos mínimos sobre ciudades reales. Para ello debemos establecer alguna forma de comprobar la capacidad de realizar esta corrección sobre la medida de centralidad de intermediación estándar.

Primeramente se describe el efecto que se debe observar al aplicar el algoritmo propuesto para calcular los valores de centralidad flexible. Debido a que se le entrega un margen para considerar caminos algo más pesados que el óptimo como caminos mínimos, se espera una mayor uniformidad en la importancia de los nodos en la red. Esto es debido a que existirán más caminos mínimos y los nodos previamente considerados más importantes no tendrán esa exclusividad.

Por tanto, se espera que los nodos más importantes en la red verán su importancia relativa disminuida debido al incremento de rutas alternativas subóptimas. Por otra parte, se podrá observar que ciertos nodos que inicialmente no eran considerados importantes verán su centralidad relativa incrementada por el mismo motivo.

Para estudiar la capacidad de corrección del método propuesto, se propone el uso de gráficos de dispersión de la medida de centralidad de intermediación estándar frente a la medida de centralidad propuesta. De esta forma se podrá estudiar de una forma visual si existe alguna variación en cada nodo respecto a la medida inicialmente considerada.

Debido a la reconsideración de otros posibles caminos subóptimos dentro de la red se espera que en el gráfico de dispersión, algunos nodos de baja *betweenness*, se posicionen por encima de la diagonal, lo que representaría un incremento respecto al valor anterior.

4.2.3 Ponderación de aristas por distancias

Se obtienen resultados de centralidad para la ciudad de Barcelona. Barcelona dispone de una red de transporte urbana con $n = 8687$ y $m = 13639$. Al ejecutar el algoritmo sobre la red de carreteras de Barcelona con varios valores de tolerancia, se obtiene el siguiente resultado (Ilustración 7). En el gráfico se muestra la relación entre el valor de centralidad mediante ambos métodos. Al entregar un valor de tolerancia relativamente alto, según las observaciones experimentales, se generaran muchos caminos mínimos nuevos con lo que debido a esto la mayoría de nodos incrementarán su centralidad en varios órdenes de magnitud.

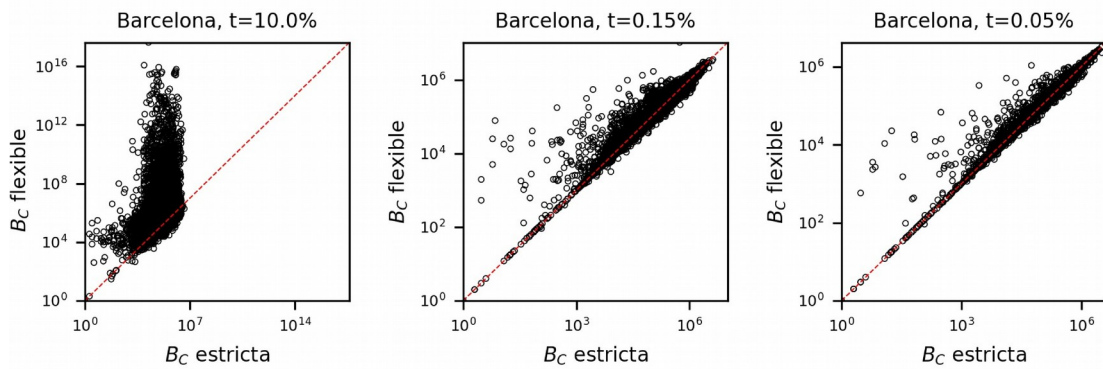


Ilustración 7: Relaciones entre los valores de centralidad de intermediación estándar frente a los valores de centralidad de intermediación flexible para varios valores de tolerancia ponderado las aristas con distancias entre nodos.

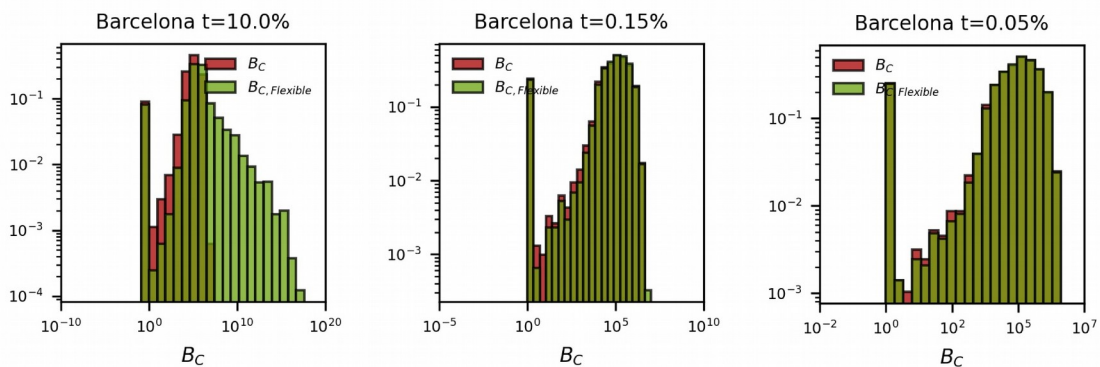


Ilustración 8: Comparación de la distribución de los valores de centralidad obtenidos mediante el método estándar y el método propuesto para varios valores de tolerancia

En el primer caso (Ilustración 7, $t = 10.0\%$) es posible observar un incremento súbito de los valores de centralidad para nodos en un tramo de centralidad media-alta sin observar un incremento de la centralidad en algunos de los nodos de mayor centralidad. Cabe destacar que el incremento de la centralidad al considerar cierta flexibilidad es normal.

Esto es debido a que al considerar la flexibilidad en el cálculo van a existir más caminos, con lo que el número de caminos que van a atravesar un nodo es mayor. Es necesario fijarse en la relación de incremento entre los nodos de menor centralidad frente a los nodos de mayor centralidad.

Volviendo al primer caso (Ilustración 7, $t = 10.0\%$), se observa un incremento en los valores mencionados acompañados del histograma (Ilustración 8, $t = 10.0\%$) podemos confirmar estos resultados. El problema que presenta este valor de tolerancia es que se pierde la jerarquía de nodos y la estructura previa anterior.

Un incremento más sutil de forma que se suavice la diferencia entre los nodos de mayor centralidad frente a los de menor centralidad es más deseable. Esto es debido a que la razón de esta métrica es para reducir la inflexibilidad que entrega la centralidad de intermediación estándar al únicamente considerar uno o varios caminos mínimos que tienen una distancia mínima.

Al reducir la tolerancia (Ilustración 7, $t = 0.15\%$, $t = 0.05\%$) se consigue el resultado buscado. La relación entre las medidas guarda el comportamiento previo, los nodos más importantes siguen siendo los más importantes pero su impacto ya no es tal y se ha suavizado la diferencia.

Si observamos el histograma (Ilustración 8, $t = 0.15\%$, $t = 0.05\%$) de los casos de menor tolerancia se observa que existen disminuciones en las frecuencias de algunos valores de centralidad de valor inferior con ciertos incrementos en los valores de centralidad más altos.

A continuación, también se muestran las distribuciones geoespaciales de las diferencias de valores de centralidad (Ilustraciones 9, 10 y 11). Los nodos rojos representan nodos que han visto reducida su centralidad al aplicar esta nueva medida, en verde los que han sufrido un incremento. Barcelona resulta un caso interesante debido al grid que presenta. Es posible observar que existe un conjunto de rutas preferidas o de mayor congestión en L'Eixample, al considerar la nueva medida de centralidad es posible diferenciar esta ruta debido a que un conjunto de nodos conectados sufren una disminución en su centralidad. Sin embargo, nodos próximos a estos sufren un incremento.

Esto es debido a que en un grid existen varias distancias de Manhattan de la misma longitud que en la práctica se traduce como una pequeña diferencia. Esta diferencia al suavizarse mediante el algoritmo presentado provoca que los nodos próximos sufran un incremento.

Distribución geográfica de la diferencia $\log_{10}(|B_{C, Flexible} - B_C|)$, $t = 0.1$



Ilustración 9: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.

Distribución geográfica de la diferencia $\log_{10}(|B_{C, Flexible} - B_C|)$, $t = 0.0015$



Ilustración 10: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.

Distribución geográfica de la diferencia $\log_{10}(|B_{C, Flexible} - B_C|)$, $t = 0.0005$



Ilustración 11: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.

4.2.4 Ponderación de aristas por tiempo de viaje

De la misma forma que en el caso anterior, se procede a obtener la variación de la centralidad de los nodos de la ciudad de Barcelona a través de los gráficos de dispersión para varias tolerancias (Ilustración 12).

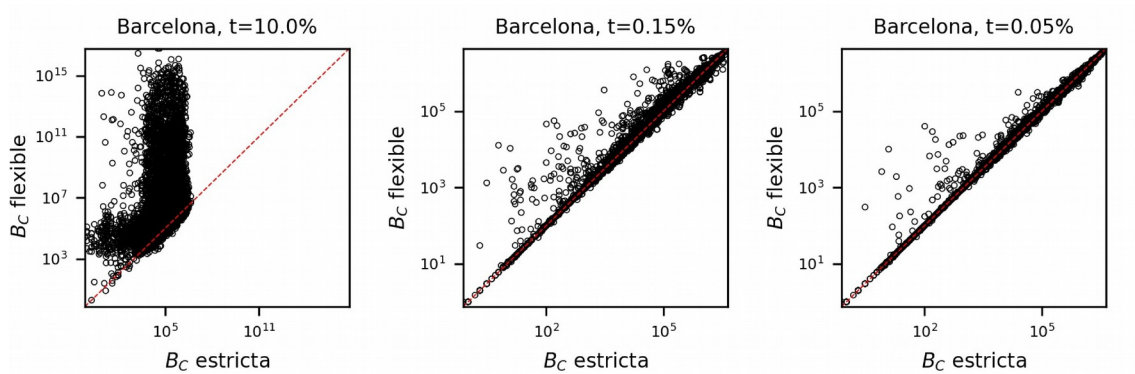


Ilustración 12: Relaciones entre los valores de centralidad de intermediación estándar frente a los valores de centralidad de intermediación flexible para varios valores de tolerancia ponderado las aristas con tiempos de viaje entre nodos.

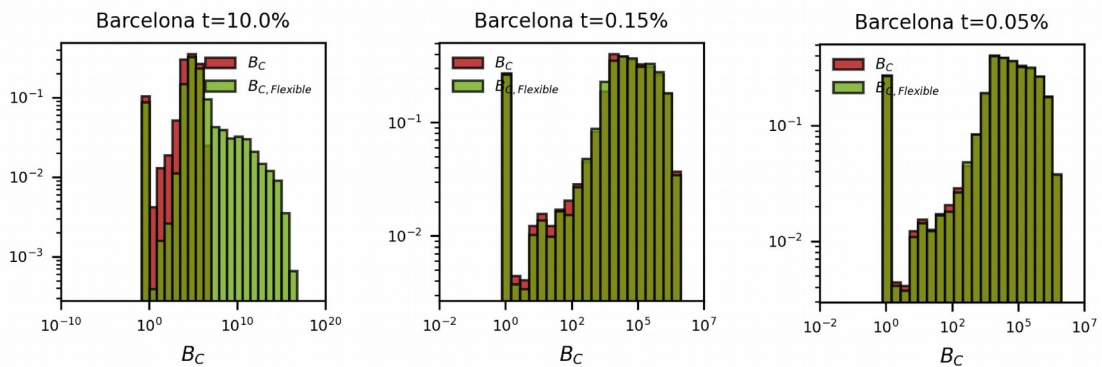


Ilustración 13: Comparación de la distribución de los valores de centralidad obtenidos mediante el método estándar y el método propuesto para varios valores de tolerancia

De modo análogo al caso de estudio de la centralidad empleando las distancias en lugar de los tiempos de viaje para ponderar las aristas se obtienen los resultados de comparación de la centralidad.

Los resultados pese a variar la ponderación de las aristas mantienen un comportamiento similar. Para el primer valor de tolerancia se obtiene un cambio abrupto en los valores de centralidad de intermediación para un determinado grupo de nodos.

Mientras que los resultados con menor tolerancia mantienen la jerarquía de importancia de los nodos suavizando la diferencia entre ellos. De la misma forma que para el caso anterior a medida que se disminuye la tolerancia los valores tienden a una correspondencia perfecta como cabría esperar.

Los histogramas (Ilustración 13) confirman la relación observada en los gráficos de dispersión. De la misma forma que en el caso anterior se proceden a mostrar las diferencias de las distribuciones geoespaciales obtenidas.

Distribución geográfica de la diferencia $\log_{10}(|B_{C, Flexible} - B_C|)$, $t = 0.1$



Ilustración 14: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.

Distribución geográfica de la diferencia $\log_{10}(|B_{C, Flexible} - B_C|)$, $t = 0.0015$



Ilustración 15: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.

Distribución geográfica de la diferencia $\log_{10}(|B_{C, Flexible} - B_C|)$, $t = 0.0005$



Ilustración 16: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.

4.2.5 Comparación con los valores de centralidad *random walk*

En este apartado se busca comparar los valores de centralidad obtenidos mediante el método propuesto y los obtenidos mediante el algoritmo de *random walk betweenness centrality*. Esta medida de centralidad valora la importancia de los nodos en una red de la misma forma que la centralidad de intermediación, la diferencia es que en lugar de emplear caminos mínimos entre pares de nodos considera caminos aleatorios.

La medida propuesta al considerar cierta flexibilidad para establecer caminos mínimos entre nodos se encuentra en un punto entre la centralidad de intermediación estándar y la centralidad de intermediación con caminos aleatorios.

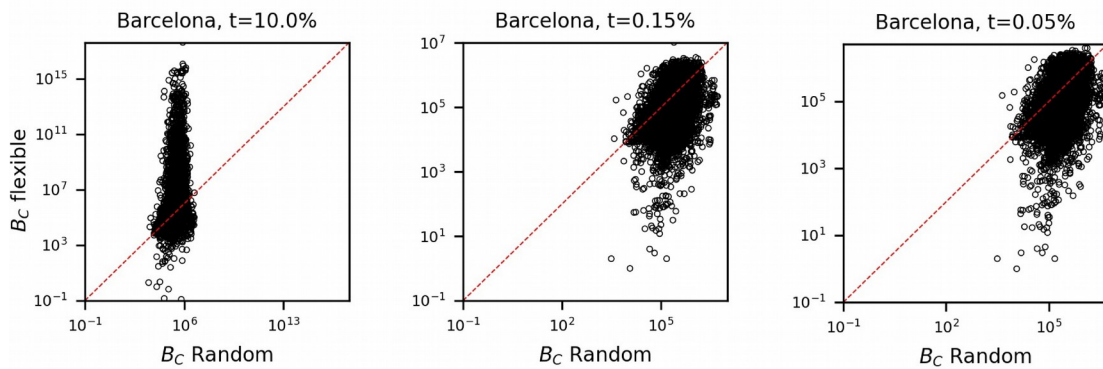


Ilustración 17: Relaciones entre los valores de centralidad de intermediación estándar frente a los valores de centralidad de intermediación flexible para varios valores de tolerancia ponderado las aristas con distancias entre nodos.

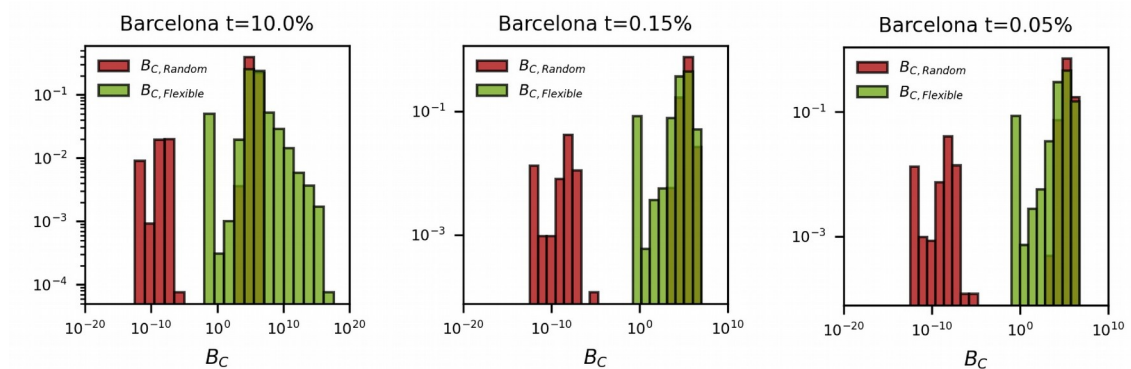


Ilustración 18: Comparación de la distribución de los valores de centralidad obtenidos mediante el método estándar y el método propuesto para varios valores de tolerancia. Ponderado con distancias.

En el gráfico de dispersión (Ilustración 19) no aparecen valores para la centralidad con caminos aleatorios para valores que aparecen del orden de 10^{-10} ya que estos valores tienen centralidad cero para la medida propuesta. Al considerar caminos aleatorios a ciertos nodos se les entrega una importancia que no se contemplaría en la centralidad de intermediación estándar o en la centralidad de intermediación flexible.

Si se realiza la misma comparación para los valores obtenidos ponderando con tiempo de viaje las aristas se obtiene lo siguiente.

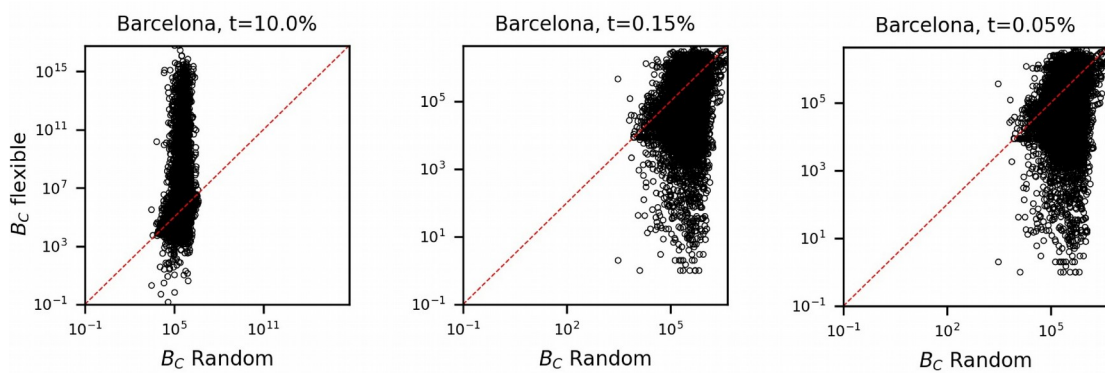


Ilustración 19: Relaciones entre los valores de centralidad de intermediación estándar frente a los valores de centralidad de intermediación flexible para varios valores de tolerancia ponderado las aristas con tiempo de viaje entre nodos.

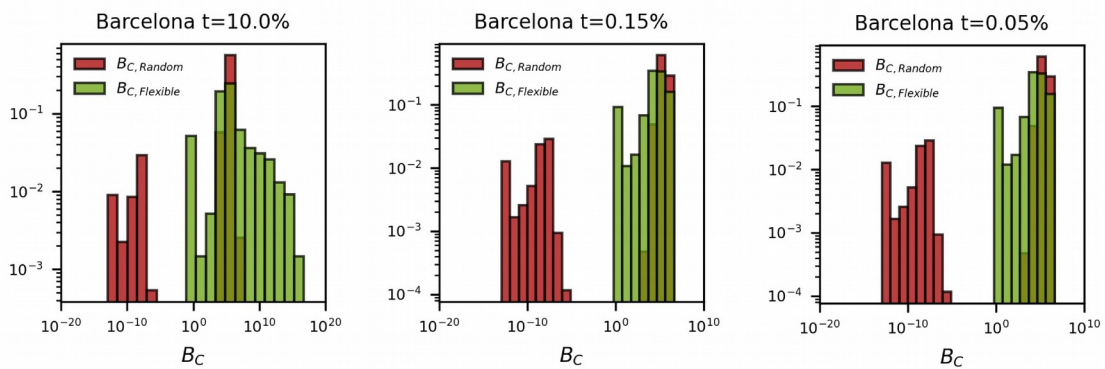


Ilustración 20: Comparación de la distribución de los valores de centralidad obtenidos mediante el método estándar y el método propuesto para varios valores de tolerancia. Ponderado con tiempos de viaje.

Capítulo 5

Conclusiones

Se ha presentado la medida de centralidad de intermediación flexible, una medida diseñada para obtener la relevancia de los nodos en redes de carreteras urbanas. Mediante esta nueva medida de centralidad, que es capaz de eliminar pequeñas diferencias entre caminos mínimos, es posible obtener una centralidad para los nodos de una red de carreteras urbanas con potencial para modelar de una forma más realista el comportamiento de los usuarios de las vías urbanas.

Al considerar pequeñas diferencias globales entre caminos mínimos es posible equiparar conjuntos de rutas alternativas que un conductor podría considerar a la hora de la toma de decisiones en la planificación de rutas.

La introducción de esta nueva medida puede resultar en un resultado deseado en ciertas circunstancias. Es importante destacar que, como se ha visto en los resultados de los experimentos sobre las redes de carreteras, la selección del valor de la tolerancia es un proceso clave para obtener resultados válidos para posteriores aplicaciones. Como se ha visto la elección de este valor depende del tamaño de la red y su topología. Como una posible línea de trabajo futuro que pueda surgir a partir de este trabajo, se propone el cálculo de límites o valores de referencia de tolerancia basados en modelos de redes de carreteras urbanas como el modelo Grid-Tree (Lampo et al., 2021) en lugar de en redes grid presentadas.

Adicionalmente, se considera que es necesario seguir profundizando en las consecuencias de introducir este tipo de márgenes de tolerancia a la hora de considerar caminos mínimos. Esta propuesta de línea de trabajo futura se entiende al observar los cambios que se producen al introducir una variación mínima en la tolerancia. Es por ello que se considera necesario seguir estudiando este tipo de fenómenos que suceden al introducir esta flexibilidad en sistemas complejos.

Desde el punto de vista de la planificación y metodología, se considera que se han logrado alcanzar tanto el objetivo principal como los objetivos parciales del trabajo. Las principales razones de haber alcanzado los objetivos propuestos al inicio del trabajo son la de una definición clara del alcance del proyecto y de su procedimiento tanto en el desarrollo como en la fase de verificación y experimentación. Estas directrices marcadas han sido de gran ayuda a la hora de solventar los problemas surgidos durante la realización del trabajo.

Bibliografía

A. Hagberg, A., A. Schult, D., J. Swart, P., 2008. Exploring Network Structure, Dynamics, and Function using NetworkX. Proceedings of the 7th Python in Science conference (SciPy 2008) 11–15.

Altshuler, Y., Puzis, R., Elovici, Y., Bekhor, S., 2011. Augmented *betweenness centrality* for Mobility Prediction in Transportation Networks. Undefined.

Batty, M., 2017. The New Science of Cities, Reprint edition. ed. The MIT Press, Cambridge, Massachusetts London, England.

Bavelas, A., 1950. Communication Patterns in Task Oriented Groups. The Journal of the Acoustical Society of America 22, 725–730. <https://doi.org/10.1121/1.1906679>

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.-U., 2006. Complex networks: Structure and dynamics. Physics Reports 424, 175–308. <https://doi.org/10.1016/j.physrep.2005.10.009>

Boeing, G., 2017. OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. Computers, Environment and Urban Systems 65, 126–139. <https://doi.org/10.1016/j.compenvurbsys.2017.05.004>

Bonacich, P., 1987. Power and Centrality: A Family of Measures. American Journal of Sociology 92, 1170–1182. <https://doi.org/10.1086/228631>

Bonacich, P., 1972a. Technique for Analyzing Overlapping Memberships. Sociological Methodology 4, 176–185. <https://doi.org/10.2307/270732>

Bonacich, P., 1972b. Factoring and weighting approaches to status scores and clique identification. The Journal of Mathematical Sociology 2, 113–120. <https://doi.org/10.1080/0022250X.1972.9989806>

Brandes, U., 2001. A faster algorithm for *betweenness centrality*. The Journal of Mathematical Sociology 25, 163–177. <https://doi.org/10.1080/0022250X.2001.9990249>

Brin, S., Page, L., 1998. The anatomy of a large-scale hypertextual Web search engine. Computer Networks and ISDN Systems, Proceedings of the Seventh International World Wide Web Conference 30, 107–117. [https://doi.org/10.1016/S0169-7552\(98\)00110-X](https://doi.org/10.1016/S0169-7552(98)00110-X)

Currie, J., Walker, R., 2011. Traffic Congestion and Infant Health: Evidence from E-ZPass. American Economic Journal: Applied Economics 3, 65–90. <https://doi.org/10.1257/app.3.1.65>

Davis, J.H., Hornseth, J., 1967. Discussion Patterns and Word Problems. Sociometry 30, 91–103. <https://doi.org/10.2307/2786441>

- Freeman, L.C., 1978. Centrality in social networks conceptual clarification. *Social Networks* 1, 215–239. [https://doi.org/10.1016/0378-8733\(78\)90021-7](https://doi.org/10.1016/0378-8733(78)90021-7)
- Freeman, L.C., 1977. A Set of Measures of Centrality Based on Betweenness. *Sociometry* 40, 35–41. <https://doi.org/10.2307/3033543>
- Gao, S., Wang, Y., Gao, Y., Liu, Y., 2013. Understanding Urban Traffic-Flow Characteristics: A Rethinking of *betweenness centrality*. *Environ Plann B Plann Des* 40, 135–153. <https://doi.org/10.1068/b38141>
- Goodwin, P., 2004. The economic costs of road traffic congestion (Report). UCL (University College London), The Rail Freight Group, London, UK.
- Guimerà, R., Díaz-Guilera, A., Vega-Redondo, F., Cabrales, A., Arenas, A., 2002. Optimal network topologies for local search with congestion. *Phys Rev Lett* 89, 248701. <https://doi.org/10.1103/PhysRevLett.89.248701>
- Guo, S., Zhou, D., Fan, J., Tong, Q., Zhu, T., Lv, W., Li, D., Havlin, S., 2019. Identifying the most influential roads based on traffic correlation networks. *EPJ Data Sci.* 8, 1–17. <https://doi.org/10.1140/epjds/s13688-019-0207-7>
- Hubbell, C.H., 1965. An Input-Output Approach to Clique Identification. *Sociometry* 28, 377–399. <https://doi.org/10.2307/2785990>
- Katz, L., 1953. A new status index derived from sociometric analysis. *Psychometrika* 18, 39–43. <https://doi.org/10.1007/BF02289026>
- Lampo, A., Borge-Holthoefer, J., Gómez, S., Solé-Ribalta, A., 2021. Multiple abrupt phase transitions in urban transport congestion. [arXiv:2005.12902](https://arxiv.org/abs/2005.12902) [nlin, physics:physics].
- Levy, J.I., Buonocore, J.J., von Stackelberg, K., 2010. Evaluation of the public health impacts of traffic congestion: a health risk assessment. *Environ Health* 9, 65. <https://doi.org/10.1186/1476-069X-9-65>
- Marshall, S., Gil, J., Kropf, K., Tomko, M., Figueiredo, L., 2018. Street Network Studies: from Networks to Models and their Representations. *Netw Spat Econ* 18, 735–749. <https://doi.org/10.1007/s11067-018-9427-9>
- Newman, M., 2018. *Networks*, 2nd edition. ed. Oxford University Press, Oxford, United Kingdom ; New York, NY, United States of America.
- Newman, M.E.J., 2003. The Structure and Function of Complex Networks. *SIAM Rev.* 45, 167–256. <https://doi.org/10.1137/S003614450342480>
- OpenStreetMap contributors, 2017. OpenStreetMap.
- Quercia, D., Schifanella, R., Aiello, L.M., 2014. The Shortest Path to Happiness: Recommending Beautiful, Quiet, and Happy Routes in the City. [arXiv:1407.1031](https://arxiv.org/abs/1407.1031) [physics].

Sabidussi, G., 1966. The centrality index of a graph. *Psychometrika* 31, 581–603. <https://doi.org/10.1007/BF02289527>

Vieira, R.S., Haddad, E.A., 2020. A weighted travel time index based on data from Uber Movement. *EPJ Data Sci.* 9, 1–15. <https://doi.org/10.1140/epjds/s13688-020-00241-y>

Weisbrod, G., Vary, D., Treyz, G., 2003. Measuring Economic Costs of Urban Traffic Congestion to Business. *Transportation Research Record* 1839, 98–106. <https://doi.org/10.3141/1839-10>

Yan, G., Zhou, T., Hu, B., Fu, Z.-Q., Wang, B.-H., 2006. Efficient routing on complex networks. *Phys. Rev. E* 73, 046108. <https://doi.org/10.1103/PhysRevE.73.046108>

Anexos

Anexo A: Centralidad de intermediación flexible en Rabat, Marruecos

Ponderación de las aristas por distancia entre nodos

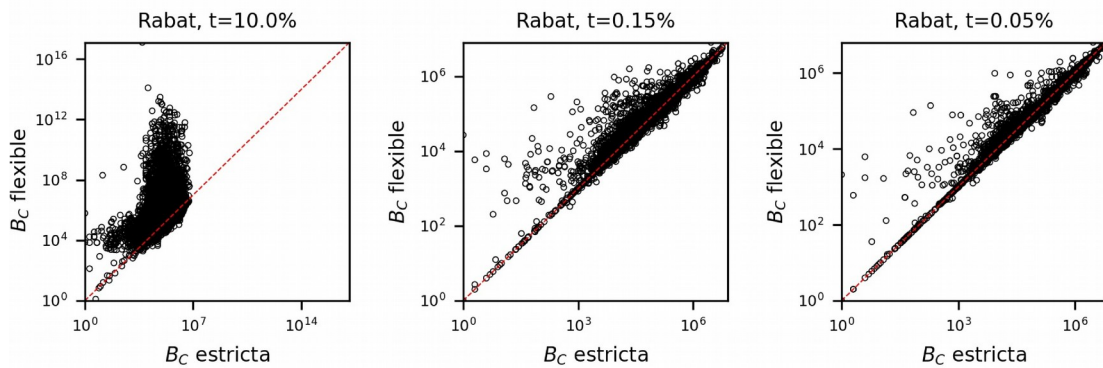


Ilustración 21: Relaciones entre los valores de centralidad de intermediación estándar frente a los valores de centralidad de intermediación flexible para varios valores de tolerancia ponderado las aristas con distancias entre nodos.

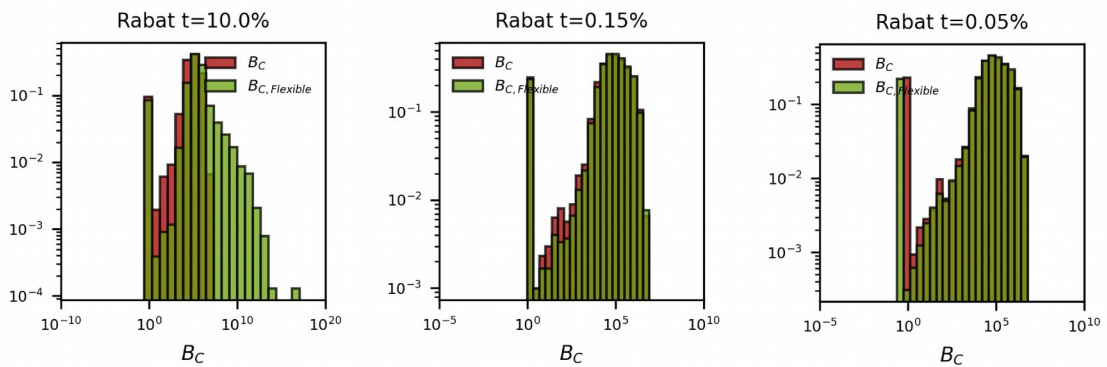


Ilustración 22: Comparación de la distribución de los valores de centralidad obtenidos mediante el método estándar y el método propuesto para varios valores de tolerancia

Distribución geográfica de la diferencia $\log_{10}(|B_{C, Flexible} - B_C|)$, $t = 0.1$



Ilustración 23: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.

Distribución geográfica de la diferencia $\log_{10}(|B_{C, Flexible} - B_C|)$, $t = 0.0015$



Ilustración 24: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.

Distribución geográfica de la diferencia $\log_{10}(|B_{C, Flexible} - B_C|)$, $t = 0.0005$



Ilustración 25: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.

Ponderación de las aristas por tiempo de viaje entre nodos

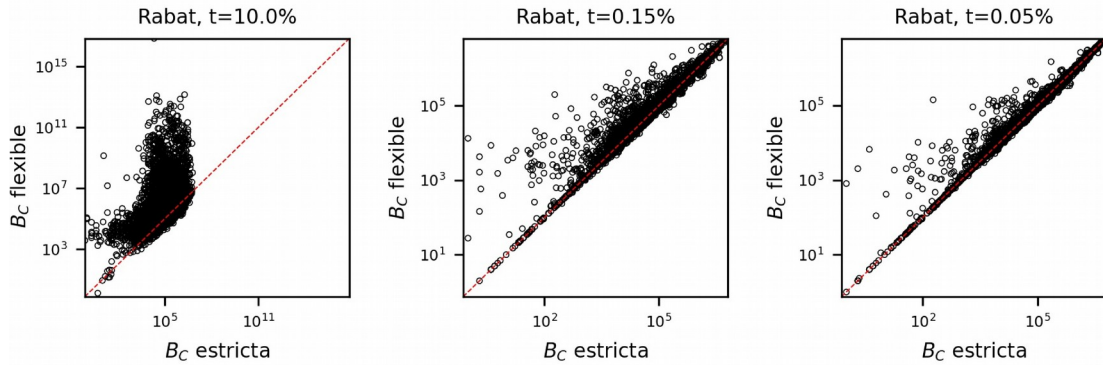


Ilustración 26: Relaciones entre los valores de centralidad de intermediación estándar frente a los valores de centralidad de intermediación flexible para varios valores de tolerancia ponderado las aristas con tiempo de viaje entre nodos.

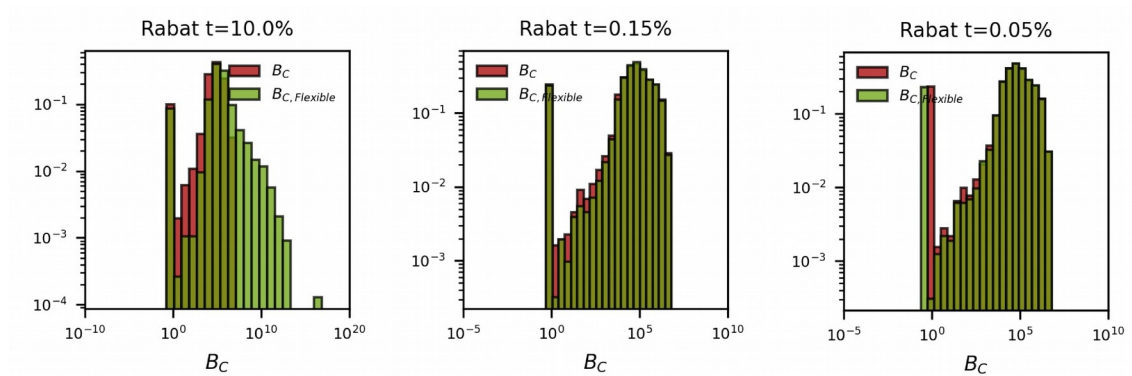


Ilustración 27: Comparación de la distribución de los valores de centralidad obtenidos mediante el método estándar y el método propuesto para varios valores de tolerancia

Distribución geográfica de la diferencia $\log_{10}(|B_{C, Flexible} - B_C|)$, $t = 0.1$



Ilustración 28: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.

Distribución geográfica de la diferencia $\log_{10}(|B_{C, Flexible} - B_C|)$, $t = 0.0015$



Ilustración 29: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.

Distribución geográfica de la diferencia $\log_{10}(|B_{C, Flexible} - B_C|)$, $t = 0.0005$



Ilustración 30: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.

Anexo B: Centralidad de intermediación flexible en Ann Arbor, EEUU

Ponderación de las aristas por distancia entre nodos

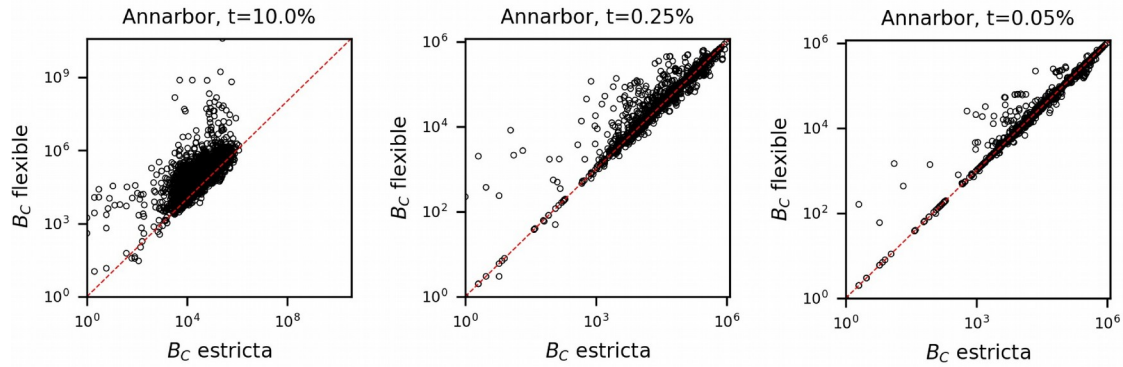


Ilustración 31: Relaciones entre los valores de centralidad de intermediación estándar frente a los valores de centralidad de intermediación flexible para varios valores de tolerancia ponderado las aristas con distancias entre nodos.

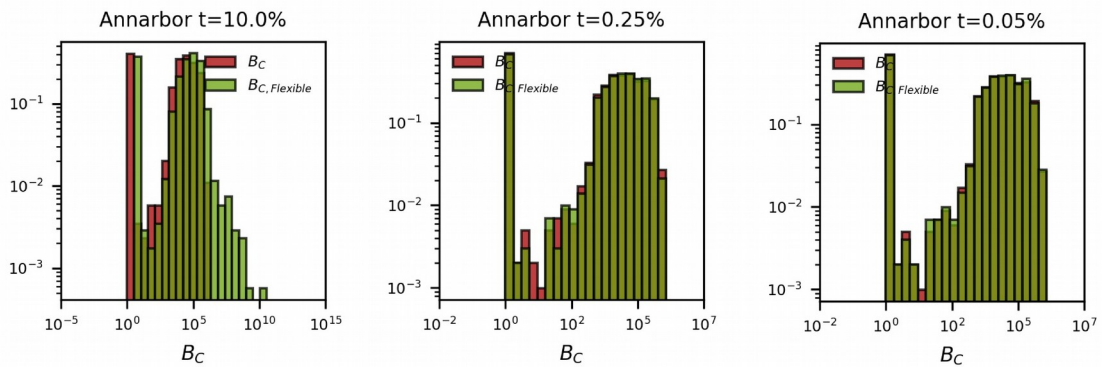


Ilustración 32: Comparación de la distribución de los valores de centralidad obtenidos mediante el método estándar y el método propuesto para varios valores de tolerancia

Distribución geográfica de la diferencia $\log_{10}(|B_{C, Flexible} - B_C|)$, $t = 0.1$



Ilustración 33: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.

Distribución geográfica de la diferencia $\log_{10}(|B_{C, Flexible} - B_C|)$, $t = 0.0025$



Ilustración 34: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.

Distribución geográfica de la diferencia $\log_{10}(|B_{C, Flexible} - B_C|)$, $t = 0.0005$

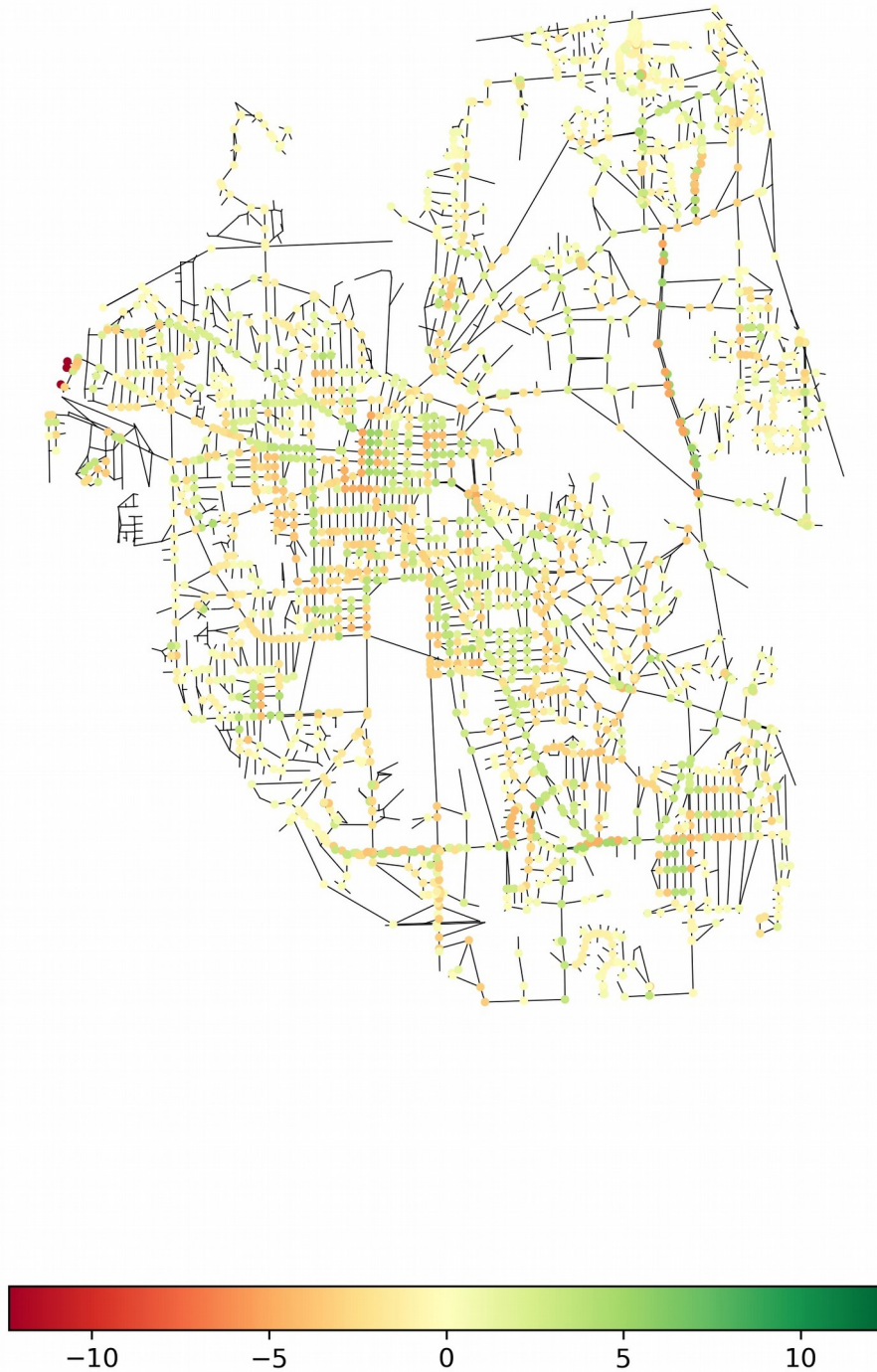


Ilustración 35: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.

Ponderación de las aristas por tiempo de viaje entre nodos

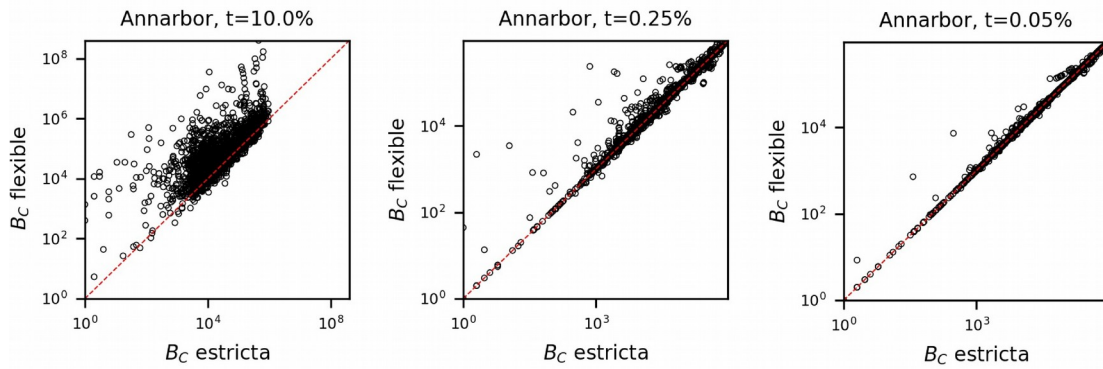


Ilustración 36: Relaciones entre los valores de centralidad de intermediación estándar frente a los valores de centralidad de intermediación flexible para varios valores de tolerancia ponderado las aristas con tiempos de viaje entre nodos.

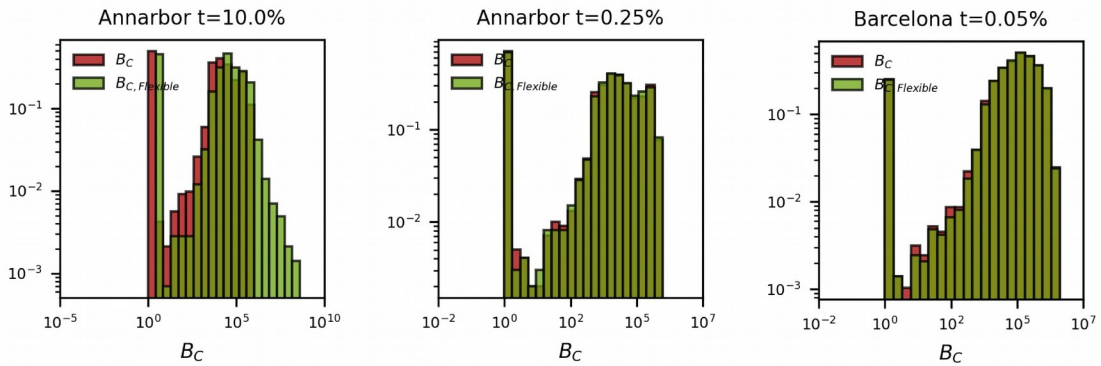


Ilustración 37: Comparación de la distribución de los valores de centralidad obtenidos mediante el método estándar y el método propuesto para varios valores de tolerancia

Distribución geográfica de la diferencia $\log_{10}(|B_{C, Flexible} - B_C|)$, $t = 0.1$

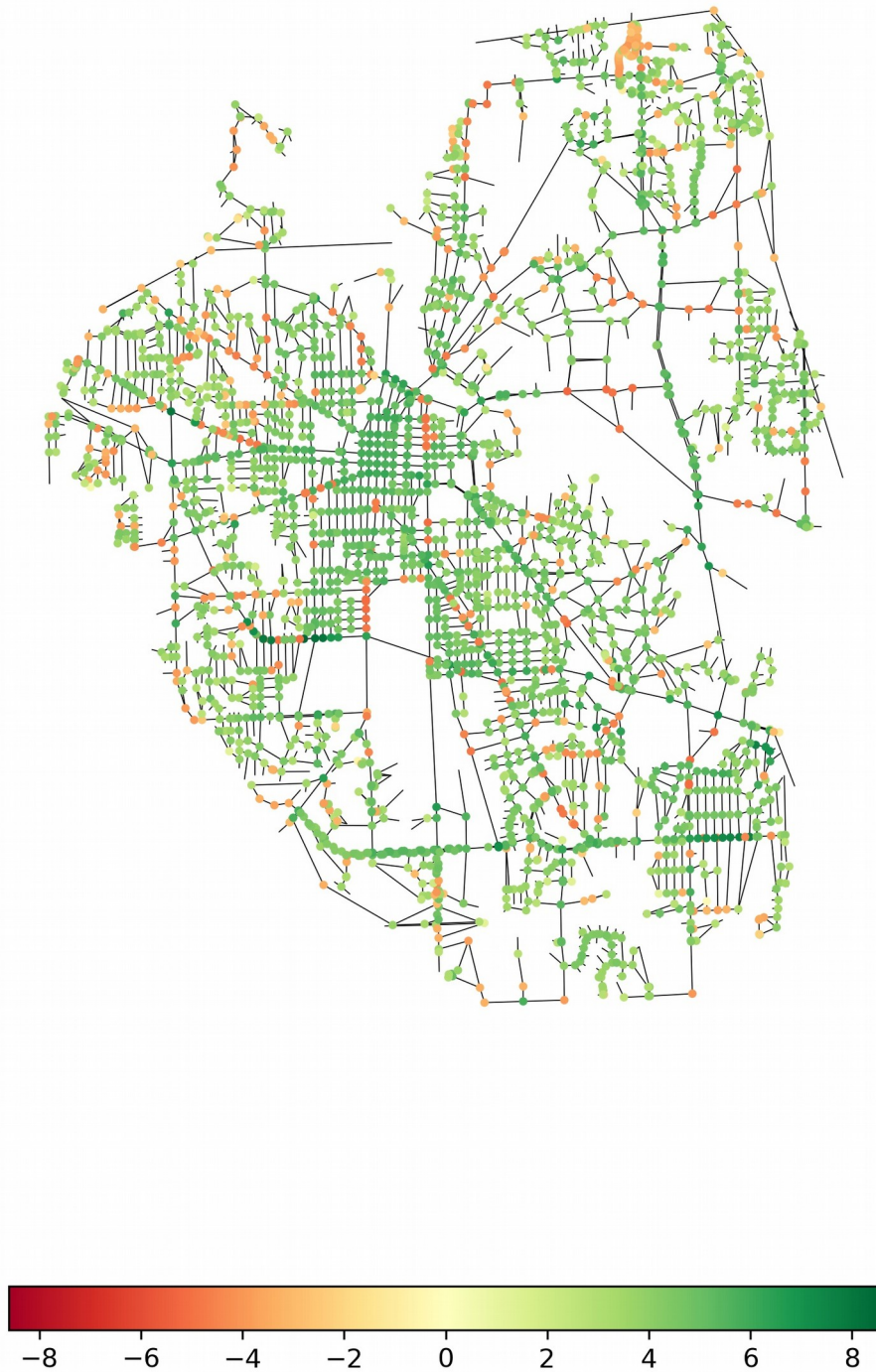


Ilustración 38: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.

Distribución geográfica de la diferencia $\log_{10}(|B_{C, Flexible} - B_C|)$, $t = 0.0025$

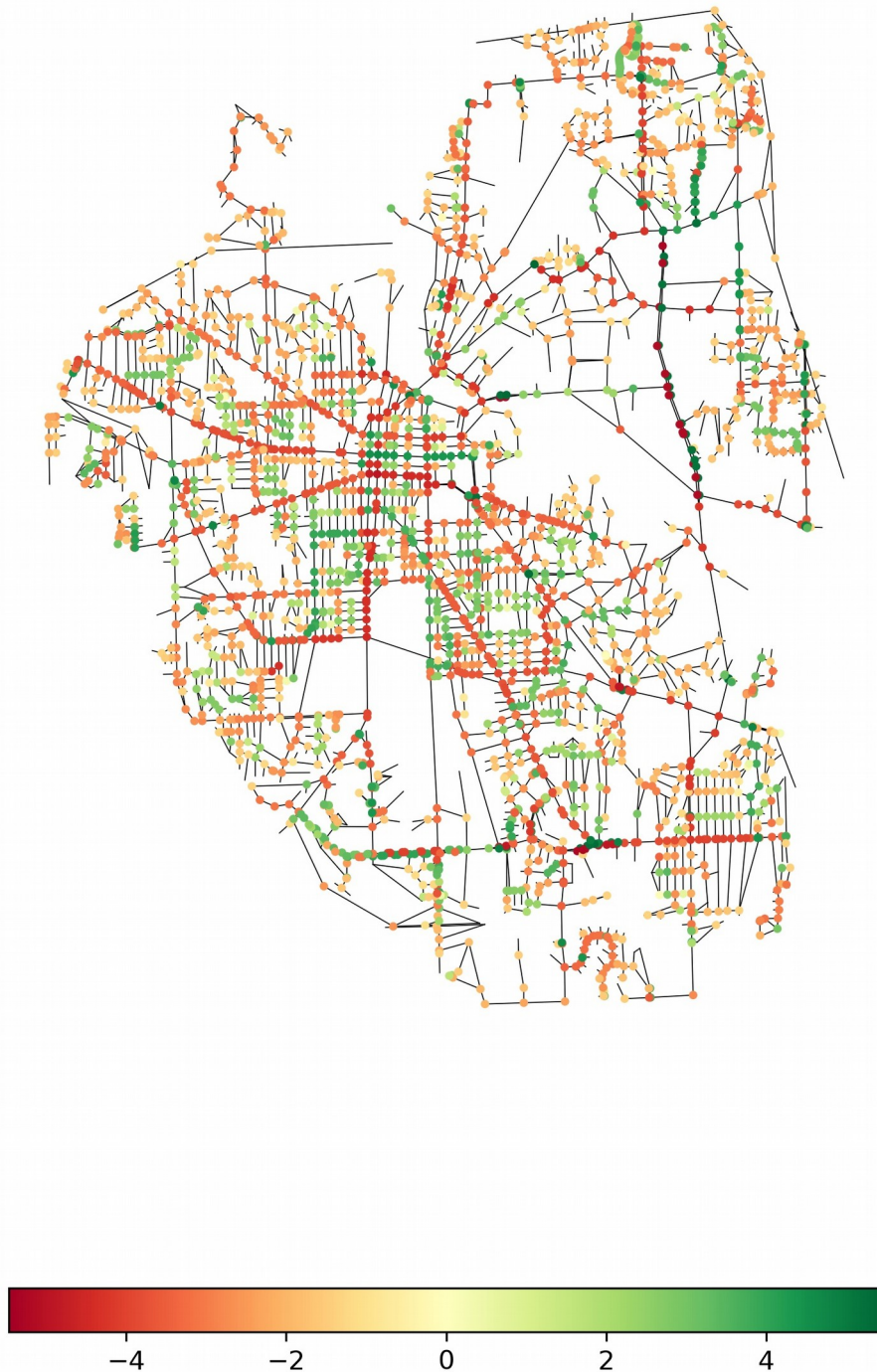


Ilustración 39: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.

Distribución geográfica de la diferencia $\log_{10}(|B_{C, Flexible} - B_C|)$, $t = 0.0005$

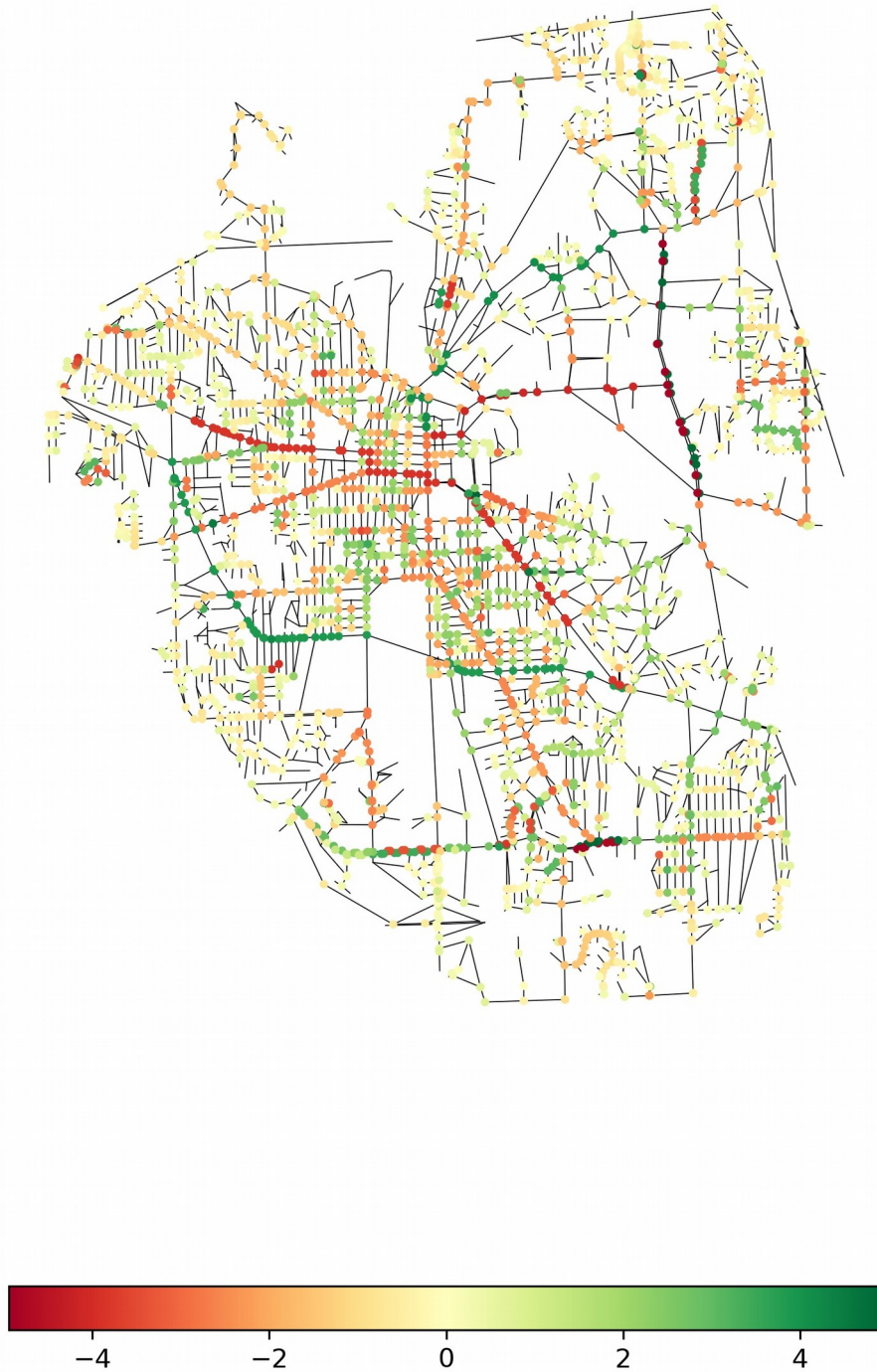


Ilustración 40: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.

Anexo C: Centralidad de intermediación flexible en Tallin, Estonia

Ponderación de las aristas por distancia entre nodos

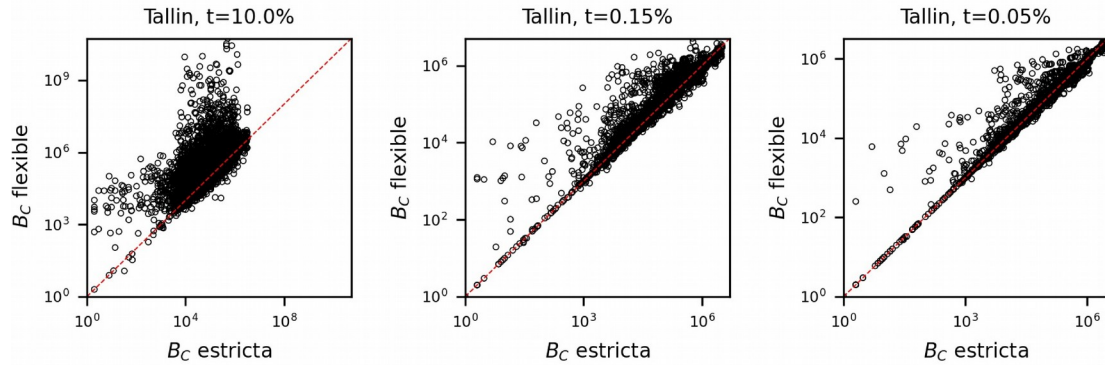


Ilustración 41: Relaciones entre los valores de centralidad de intermediación estándar frente a los valores de centralidad de intermediación flexible para varios valores de tolerancia ponderado las aristas con distancias entre nodos.

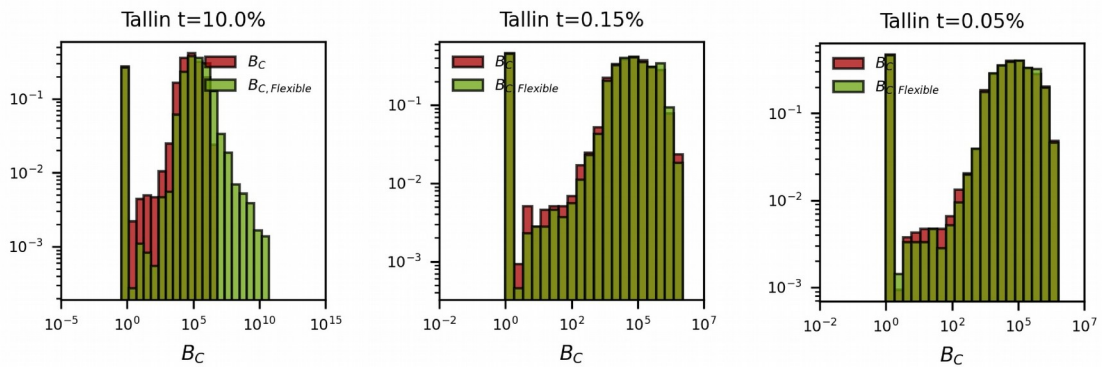


Ilustración 42: Comparación de la distribución de los valores de centralidad obtenidos mediante el método estándar y el método propuesto para varios valores de tolerancia

Distribución geográfica de la diferencia $\log_{10}(|B_{C, Flexible} - B_C|)$, $t = 0.1$



Ilustración 43: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.

Distribución geográfica de la diferencia $\log_{10}(|B_{C, Flexible} - B_C|)$, $t = 0.0015$



Ilustración 44: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.

Distribución geográfica de la diferencia $\log_{10}(|B_{C, Flexible} - B_C|)$, $t = 0.0005$



Ilustración 45: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.

Ponderación de las aristas por tiempos de viaje entre nodos

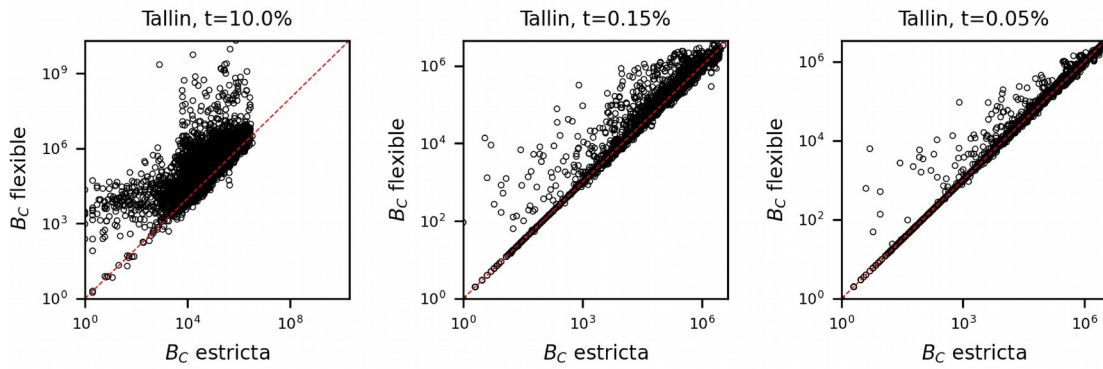


Ilustración 46: Relaciones entre los valores de centralidad de intermediación estándar frente a los valores de centralidad de intermediación flexible para varios valores de tolerancia ponderado las aristas con distancias entre nodos.

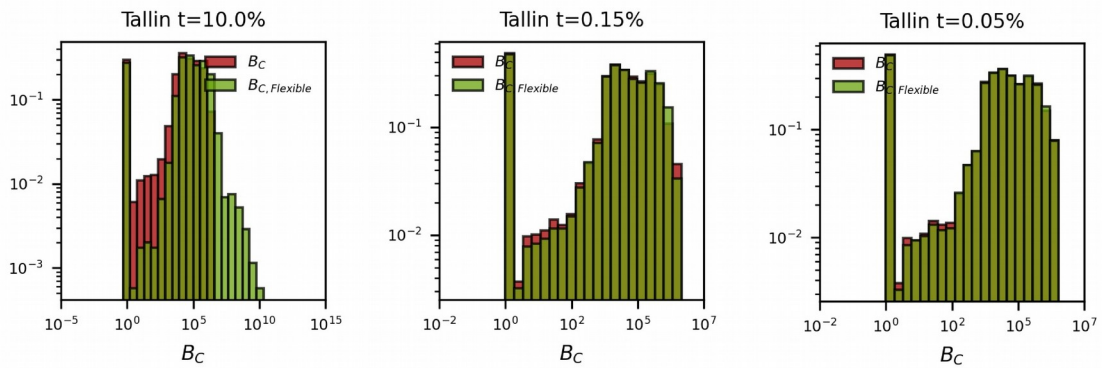


Ilustración 47: Comparación de la distribución de los valores de centralidad obtenidos mediante el método estándar y el método propuesto para varios valores de tolerancia

Distribución geográfica de la diferencia $\log_{10}(|B_{C, Flexible} - B_C|)$, $t = 0.1$



Ilustración 48: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.

Distribución geográfica de la diferencia $\log_{10}(|B_{C, Flexible} - B_C|)$, $t = 0.0015$



Ilustración 49: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.

Distribución geográfica de la diferencia $\log_{10}(|B_{C, Flexible} - B_C|)$, $t = 0.0005$

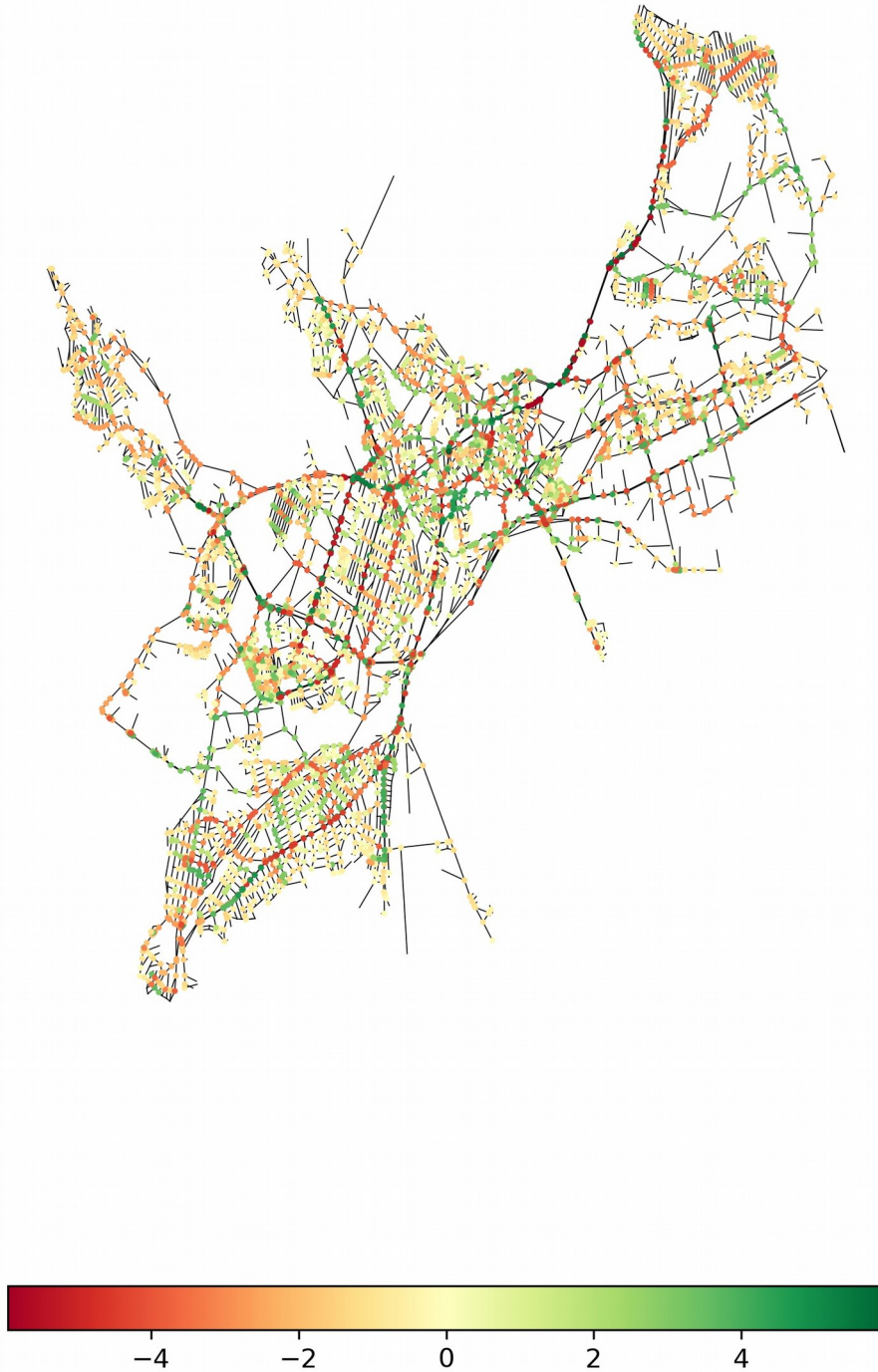


Ilustración 50: Variación de la centralidad para los nodos de la ciudad. Se muestra en rojo aquellos nodos que han visto disminuida su centralidad al obtenerla mediante la centralidad de intermediación flexible. En verde se muestra el caso opuesto, aquellos nodos que al introducir la medida de centralidad propuesta han visto incrementada su centralidad.