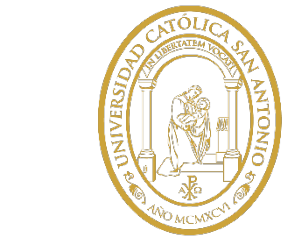
TESIS DOCTORAL



UCAM
UNIVERSIDAD CATÓLICA
DE MURCIA

ESCUELA INTERNACIONAL DE DOCTORADO

Programa de Doctorado Tecnologías de la Computación e Ingeniería Ambiental

Análisis y desarrollo de medidas de centralidad aplicadas a la formación de metales vítreos

Autor/a:

Rocío Rodríguez Gómez

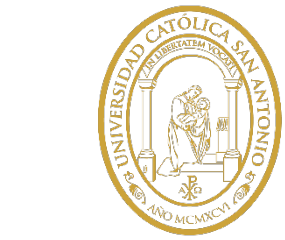
Directores/as:

Dr. D. Manuel Curado Navarro

Dr. D. José Francisco Vicent Frances

Murcia, marzo 2023

TESIS DOCTORAL



UCAM
UNIVERSIDAD CATÓLICA
DE MURCIA

ESCUELA INTERNACIONAL DE DOCTORADO

Programa de Doctorado Tecnologías de la Computación e Ingeniería Ambiental

Análisis y desarrollo de medidas de centralidad aplicadas a la formación de metales vítreos

Autor/a:

Rocío Rodríguez Gómez

Directores/as:

Dr. D. Manuel Curado Navarro

Dr. D. José Francisco Vicent Frances

Murcia, marzo de 2023

AUTORIZACIÓN DEL DIRECTOR DE LA TESIS
PARA SU PRESENTACIÓN

Los doctores D. Manuel curado Navarro y D. José Francisco Vicent Frances como directores de la Tesis Doctoral titulada "*Análisis y desarrollo de medidas de centralidad aplicadas a la formación de metales vítreos*" realizada por Dña. Rocío Rodríguez Gómez en el Programa de Doctorado Tecnologías de la Computación e Ingeniería Ambiental autorizan su presentación a trámite dado que reúne las condiciones necesarias para su defensa.

Lo que firmo, para dar cumplimiento al Real Decreto 99/2011 de 28 de enero, en Murcia a 1 de marzo de 2023.

CURADO NAVARRO MANUEL - 45747088L
Firmado digitalmente por CURADO NAVARRO MANUEL - 45747088L
Fecha: 2023.03.18 10:06:41 +01'00'

Fdo.: Manuel Curado Navarro

JOSE FRANCISCO|VICENT|FRANCES
Firmado digitalmente por JOSE FRANCISCO|VICENT|FRANCES
Fecha: 2023.03.18 11:23:36 +01'00'

Fdo.: José F. Vicent Francés

RESUMEN

La presente Tesis Doctoral tiene como objetivo general contribuir a la comprensión de la formación de metales vítreos a través del uso de redes complejas, mediante la implementación de una nueva medida de centralidad adaptada. Para lograr este objetivo general, el inicio del presente estudio se centra en la introducción y estudio de las medidas de centralidad existentes, así como la implementación de las mismas, en casos de estudio como pueden ser las redes urbanas. A continuación, se propone una nueva medida de centralidad y la adaptación de la misma para posteriormente implementarla en el caso particular de formación de metales vítreos, objetivo principal de la presente Tesis Doctoral.

La primera parte de la presente Tesis se centra en el estudio de las medidas de centralidad clásicas para la comprensión de las mismas así como de su utilización. En esta parte introductoria del estudio se aplican estas medidas de centralidad clásicas en redes urbanas con el objetivo de dilucidar mejoras concretas y eficientes de restauración que proporcionen un aumento sustancial en la accesibilidad en una ciudad histórica.

El uso de las medidas clásicas de centralidad resulta ser muchas veces insuficiente para el estudio de problemas específicos como es el caso de la formación de metales vítreos. Por ello, en la actualidad existen diferentes medidas de centralidad adaptadas a cada problema concreto, siendo unas más generales y otras más específicas. En esta tesis se propone una medida de centralidad adaptada, con cierta generalidad para poder ser aplicada a diferentes ámbitos, pero pensada especialmente para resolver el problema de la comprensión de la formación de los metales vítreos.

La medidas de centralidad realizadas con la medida clásica de intermediación o *betweenness* (CBT), están basados en los caminos más cortos y aleatorios, es decir, miden la importancia global de un nodo como nodo intermedio o de transición, pero tienen la característica común de no tener en cuenta la densidad de clúster de cada nodo. Para solucionar esta carencia se ha propuesto una nueva medida de centralidad basada en los caminos aleatorios de retorno tipo *betweenness* (2RW). Desde el punto de vista de las redes densas, está medida realiza una cuantificación de la importancia de un nodo a través de las relaciones entre cuatro nodos diferentes conectados. En el análisis de la implementación de la nueva medida de

centralidad de la red, ésta se ha aplicado desde una perspectiva orientada a clasificar nodos, reforzando comunidades densas mediante la evaluación de grafos y utilizando una matriz de probabilidad de transición de dos posibles caminos. Por tanto, se ha desarrollado una nueva medida de centralidad 2RW que combina la idea de la centralidad *betweenness* y el algoritmo de predicción de enlaces *Return Random Walk*. En concreto, la métrica propuesta aumenta la posición del ranking de relevancia en la clasificación de los nodos que pertenecen a clústeres densos con un grado medio superior al del resto de clústeres, observando que funciona mejor en redes densas. Además, podemos detectar la debilidad de una red comparando el método CBT con nuestra propuesta (2RW).

La aplicación de la medida de centralidad inicial propuesta (2RW) no se adapta completamente al problema planteado inicialmente debido fundamentalmente a la direccionalidad natural del mismo y a la falta de comprensión del comportamiento de cada nodo dentro de la red, es decir, su papel o rol dentro de la misma. Por tanto, en la presente Tesis se ha propuesto el Algoritmo de Centralidad de intermediación aleatorio bidireccional dirigido (D2RWBT), un nuevo modelo de centralidad para redes dirigidas. Mediante este modelo se han obtenido un ranking de los nodos en redes dirigidas para describir su relevancia dentro de la red como nodos de transición, teniendo en cuenta el comportamiento de los mismos en la red. Más en detalle, el modelo describe un nodo mediante cuatro índices que proporcionan información sobre la densidad de su clúster/comunidad (denso o disperso), la fuerza de sus conexiones, la importancia relativa y absoluta en la red, o la relevancia como nodo intra o interclúster.

De la aplicación de la nueva medida de centralidad (D2RWBT) para la comprensión de la formación de los metales vítreos a través de una de las variables de mayor relevancia, la temperatura de transición vítrea reducida (Trg), se han obtenido resultados que han sido ratificados por una buena correlación entre estos y la obtención real de vidrios metálicos en recientes investigaciones. En concreto se ha podido extraer la relevancia de los elementos químicos que componen las redes densas y los elementos que forman redes dispersas, así como comparar estos resultados con los obtenidos con la medida clásica de centralidad clásica *betweenness*. Además, se ha obtenido el rol de los elementos químicos formadores de metales vítreos dentro de la red y cuáles de ellos desempeñan funciones semejantes dentro de la misma y, por tanto, pueden ser posibles elementos de sustitución. Como otra parte del estudio final se han analizado los elementos con funciones de relevancia en la composición de ambos tipos de redes. Por último, se

han cotejado los resultados y se ha realizado una síntesis de la compresión del problema de formación de metales vítreos abordando el objetivo general descrito.

PALABRAS CLAVE: Medidas de centralidad; metales vítreos; redes complejas; informática; propiedades de materiales

ABSTRACT

The general objective of this doctoral thesis is to contribute to the understanding of glassy metal formation through the use of complex networks by implementing a new centrality measure. In order to achieve this general objective, the beginning of the present study focuses on the introduction of centrality measures, as well as their implementation, in case studies with a type of application of complex networks, such as urban networks. Then, a new centrality measure is proposed and adapted to be implemented in the particular case of glassy metal formation, which is the main objective of this Doctoral Thesis.

The first part of this thesis focuses on the study of classical centrality measures in order to understand them and their use. In this introductory part of the study, these classical centrality measures are applied to urban networks with the aim of elucidating concrete and efficient restoration improvements that provide a substantial increase in accessibility in a historic city.

The use of classical centrality measures often proves to be insufficient for the study of specific problems such as glassy metal formation, the main objective of this Doctoral Thesis. For this reason, there are currently different centrality measures adapted to each specific problem, some being more general and others more specific. In this thesis, an adapted centrality measure is proposed, with a certain generality to be applied to different fields, but specially designed to solve the problem of understanding the formation of glassy metals.

The centrality measures performed with the classical betweenness measure (CBT) are based on the shortest paths and the random walk, i.e. they measure the global importance of a node as an intermediate node, but they have the common feature of not taking into account the cluster density of each node. To address this shortcoming, a new centrality measure based on betweenness random walks (2RW) has been proposed. From the point of view of dense networks, this measure quantifies the importance of a node through the relationships between four different neighbouring nodes. In the analysis of the implementation of the new network centrality measure, it has been applied from a node ranking perspective, reinforcing dense communities by evaluating graphs and using a transition

probability matrix of two possible paths. Therefore, a new 2RW centrality measure has been developed that combines the idea of betweenness centrality of the random walk and the link prediction algorithm Return Random Walk. Specifically, the proposed metric increases the ranking of nodes belonging to dense clusters with a higher average degree than the rest of the clusters, observing that it performs better in dense networks. Moreover, we can detect the weakness of a network by comparing the CBT method with our proposal (2RW).

The application of the original proposed centrality measure 2RW was not completely adapted to the problem initially posed, mainly due to the natural directionality of the problem. Therefore, in this Thesis, the Directed Bidirectional Randomised Bidirectional Brokerage Centrality Algorithm (D2RWBT), a new centrality model for directed networks, has been proposed. Using this model, rankings of nodes in directed networks have been obtained to describe their relevance within the network as transition nodes. In more detail, the model describes a node by means of four indices that provide information about the density of its cluster/community (dense or sparse), the strength of its connections, the relative and absolute importance in the network, or the relevance as an intra- or inter-cluster node.

From the application of the new centrality measure (D2RWBT) for the understanding of the formation of glassy metals through one of the most relevant variables, the reduced glass transition temperature, results have been obtained that have been ratified by a good correlation between these and the real obtaining of metallic glasses in recent research. In particular, it has been possible to extract the relevance of the chemical elements that make up the dense networks and the elements that form dispersed networks, as well as to compare these results with those obtained with the classical measure of classical centrality, Betweenness (BT). In addition, the role of the glassy metal-forming chemical elements within the lattice and which of them play similar functions within the lattice and can therefore be possible substitution elements has been obtained from the study. As a further part of the final study, the elements with relevant functions in the composition of both types of lattices have been analysed. Finally, the results have been collated and a synthesis of the understanding of the glassy metal formation problem has been made, addressing the general objective described above.

KEYWORDS: Centrality measures; metallic glass; complex networks; computer science; material properties.

AGRADECIMIENTOS

Agradecer a mi familia y amigos su apoyo incondicional a cada paso en cada proyecto.

A mis directores de Tesis por toda su enseñanza, confianza, apoyo y paciencia.

CITA

"Somos como nodos orquestados por la vida y conectados
por nuestros afectos”

ÍNDICE GENERAL

ÍNDICE DE FIGURAS

ÍNDICE DE TABLAS

SIGLAS Y ABREVIATURAS

MG, Metales vítreos
BGM, Bulk Metallic glasses
GFA, Capacidad formación de vidrio
Trg, Temperatura de transición vítrea reducida
APA, PageRank adaptado
2RW, Two-way random walk betweenness centrality
D2RWBT, Directed Two-Way Random Walk Betweenness Centrality

0 - RESULTADOS DE INVESTIGACIÓN

0-RESULTADOS DE INVESTIGACIÓN

La Tesis doctoral adopta el formato de compendio de publicaciones y contiene una síntesis de los resultados obtenidos durante el desarrollo de la investigación. El compendio de publicaciones incluye 4 artículos científicos publicados en revistas indexadas de la base Journal Citation Report (JCR). Los artículos publicados son:

1.- Curado, M., Rodríguez, R., Jiménez, M., Tortosa, L., & Vicent, J. F. (2021). A New Methodology to Study Street Accessibility: A Case Study of Avila (Spain). *ISPRS International Journal of Geo-Information*, *10*(7), 491.

DOI: https://doi.org/10.3390/ijgi10070491

Factor de impacto 2021 (JCR): 3.099; Categoría (JCR): COMPUTER SCIENCE, INFORMATION SYSTEMS; Clasificación: 88/164 (Q3); Fecha de publicación 2021.

2.- Curado, M., Rodríguez, R., Tortosa, L., & Vicent, J. F. (2022). Anew centrality measure in dense networks based on two-way random walk betweenness. *Applied Mathematics and Computation*, *412*, 126560.

DOI: https://doi.org/10.1016/j.amc.2021.126560

Factor de impacto 2021 (JCR): 4.397; Categoría (JCR): MATHEMATICS, APPLIED; Clasificación: 7/267 (Q1); Fecha de publicación 2022.

3.- Curado, M., Rodríguez, R., Terroso-Sáenz, F., Tortosa, L., & Vicent, J. F. (2022). A centrality model for directed graphs based on the two-way-random path and associated indices for characterizing the nodes. *Journal of Computational Science*, *63*, 101819.

DOI: https://doi.org/10.1016/j.jocs.2022.101819

Factor de impacto 2021 (JCR): 3.817; Categoría (JCR): COMPUTER SCIENCE, THEORY & METHODS; Clasificación: 25/110 (Q1); Fecha de publicación 2022.

4.- Rodríguez, R., Curado, M., Tortosa, L., & Vicent, J. F. (2023). Understanding the metallic glasses formation by applying a centrality measure based on betweenness. *Computational Materials Science*, *218*, 111986.

DOI: https://doi.org/10.1016/j.commatsci.2022.111986

Factor de impacto 2021 (JCR): 3.572; Categoría (JCR): MATERIALS SCIENCE, MULTIDISCIPLINARY; Clasificación: 184/345 (Q3); Fecha de publicación 2023.

Adicionalmente, resultados complementarios de la investigación se han presentado al siguiente congreso internacional:

1. XI Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA 2023), Alicante, España. Contribución: Influential Yield Strength of the Steel Materials with Return Random Walk Gravity Centrality.

I – INTRODUCCIÓN

I - INTRODUCCIÓN

En las últimas décadas, el estudio de redes complejas ha ganado notoriedad en el avance de la comprensión de los sistemas interconectados como redes biológicas, sociales, de transporte y de comunicación [1-6]. Además, en sistemas complejos los nodos interactúan entre sí siguiendo diferentes tipos de patrones que pueden abarcar distintos tipos de relaciones [7]. La transferencia de ideas o enfermedades, los flujos de información y la influencia entre nodos son una función clave de las redes complejas [8-11]. Por lo tanto, evaluar la influencia de los nodos en una red es un tema de especial relevancia para comprender puntos como la propagación de la información en una red en función de la influencia de un nodo dentro de la misma [12-15].

Las tecnologías digitales modernas y la accesibilidad de los conjuntos de datos, obtenidos a través de diferentes fuentes, han aumentado el interés en estudiar las propiedades estructurales de las redes que representan problemas reales y que se caracterizan por propiedades no triviales. Los diferentes métodos, conceptos, algoritmos y modelos propuestos para analizar la estructura de las redes reales han formado el marco de la llamada teoría de redes complejas [16,17].

Es habitual que las características de una red compleja revelen información muy importante sobre su comportamiento, tanto a nivel local como global. Por ejemplo, relacionar la posición ocupada para un nodo en un grafo y el papel de éste en procesos dinámicos. Por lo tanto, no todos los nodos tienen la misma importancia en problemas como la propagación de rumores o enfermedades en una red porque las relaciones entre ellos guardan información importante.

Para caracterizar y comprender el papel de un nodo en una red compleja es muy habitual hacer uso de medidas de centralidad. Existe una gran diversidad de estas métricas para el estudio, tanto a nivel local como global, de las características e importancia de un elemento (nodo) dentro de una red compleja. Se puede clasificar en dos tipos: las más clásicas, como centralidad de grado, de cercanía, de intermediación o *PageRank*, entre otras, hasta otras medidas más personalizadas, para analizar la red desde un punto de vista más concreto y específico. Debido a esta personalización, puede analizarse problemas de diferente índole, siempre que puedan ser representadas la información en forma de red compleja.

En el campo de la ciencia de los materiales, los metales vítreos o *metallic glasses* (MG) son materiales no cristalinos formados por metales puros o combinaciones de metales y metaloides [18]. Son materiales relativamente nuevos (década de los 60), y consisten en aleaciones sólidas no cristalinas que tienen una disposición atómica heredada directamente del estado líquido. Su estructura atómica se ha entendido básicamente mediante el uso de un esquema de empaquetamiento aleatorio denso, y el descubrimiento de aleaciones que se pueden obtener en secciones de hasta unos pocos centímetros de espesor ha despertado gran interés en esta categoría de materiales. Debido a su estabilidad y a propiedades mecánicas como su dureza (puede ser hasta tres veces más duros que el acero), elasticidad (mejor que los materiales cerámicos) y resistencia (menos frágiles que el vidrio transparente a base de óxido) [19], estos nuevos compuestos abren nuevas áreas de aplicación y facilitan estudios fundamentales.

A pesar del gran potencial de estos materiales, hay una evidente falta de comprensión en la base científica de cómo se forman a partir de los elementos que los componen.

Es por ello que, en la presente Tesis se propone el estudio de los metales vítreos existentes actualmente a través de una medida de centralidad sobre redes complejas adaptada para la comprensión del comportamiento de cada elemento dentro de un compuesto. Además, se pretende que sirva como base para el descubrimiento de nuevos metales vítreos.

1.1. FUNDAMENTACIÓN DE LAS MEDIDAS DE CENTRALIDAD

Las medidas de centralidad conforman uno de los conceptos más importantes dentro del moderno análisis de redes complejas. La centralidad de un nodo evalúa la importancia del nodo dentro de la red desde diferentes perspectivas. Por tanto, existen diversos tipos de medidas de centralidad para obtener distintas interpretaciones y objetivos dependiendo de las características topológicas de la red y de la aplicación concreta que se esté estudiando [20].

El concepto de importancia de un nodo es por tanto muy relativo puesto que según su aplicación puede ser más relevante sí está conectado con muchos nodos dentro de esa red, si es el nodo es de paso en muchos caminos que conectan

con otros nodos, o sí está conectado con otros nodos importantes, entre otros. Todas las interpretaciones son posibles según la aplicación de la red. Con dependencia de la interpretación se definen por tanto distintas medidas de centralidad. Por ejemplo, entre las medidas de centralidad denominadas clásicas o más utilizadas destacamos: a) la centralidad del grado (*degree*), b) centralidad de cercanía (*closeness*), c) centralidad de intermediación (betweennes), d) centralidad de vector propio (*eigenvector*), y e) centralidad PageRank (ver resumen en Tabla 1).

Tabla 1. Resumen de las medidas de centralidad clásicas

Medida de centralidad	***Descripción breve***
Degree	*Importancia del nodo en función del número/peso de los enlaces.*
Closeness	*Importancia en función de la cercanía del nodo respecto al resto de la red (alcanzabilidad).*
Betweenness	*Importancia del nodo en función de la conectividad con otros nodos de la red.*
Eigenvector	*La medición indica la centralidad de grado más la relevancia de las uniones de los nodos cercanos.*
PageRank	*Importancia del nodo en función de la importancia de los nodos con los que conecta.*

En la construcción de redes complejas hay que tener en cuenta diferentes aspectos, destacando dos puntos: dirección de los enlaces (dirigido y no dirigido) y el peso de los mismos (*weighted* y *unweighted*).

En redes no dirigidas, la centralidad de grado evalúa el número de enlaces (aristas) que posee un nodo o la suma del peso de las mismas para el caso *unweighted*, mientras que en redes dirigidas, se distinguen entre aristas de entrada y de salida, y por tanto, tiene centralidad de grado de entrada y de salida (in and *out-degree*). La centralidad de cercanía es la más conocida y una de las más utilizadas. Se basa en calcular la suma o bien el promedio de las longitudes de los caminos más cortos desde un nodo hacia todos los demás. Por tanto, mientras mayor sea la distancia entre dos vértices, menor será la cercanía entre estos. La

cercanía se definirá como el inverso multiplicativo de la lejanía entre dos vértices. La medida de centralidad de intermediación o *betweenness* es una medida de centralidad que cuantifica la frecuencia o el número de veces que un nodo se encuentra entre las longitudes del camino más corto entre los nodos. Un nodo tendrá una alta intermediación si es un vértice de paso en la conexión de muchos pares de nodos. Respecto a la medida de centralidad del vector propio, se trata de una medida de centralidad que cuantifica el nivel de influencia, prestigio o estatus de un nodo en una red. Los nodos que poseen un valor alto de esta medida de centralidad están conectados a otros nodos que a su vez son muy relevantes o bien a muchos otros nodos menos relevantes, teniendo en cuenta la interpretación de relevancia de la propia medida. Por último, en la medida de centralidad de PageRank, al igual que con el vector propio, la relevancia de una entidad refleja no solo el número de conexiones, sino también cuántas conexiones tienen esas entidades vecinas dentro de la red. Un nodo tendrá una relevancia mayor cuando tiene muchos nodos relevantes conectados que, a su vez, tienen muchas conexiones importantes.

El uso de estas medidas clásicas ha resultado ser un poco rígidas para alcanzar la gran cantidad de problemas diferentes en el mundo real. Por ello, en la actualidad existen diferentes medidas de centralidad adaptadas a cada problema concreto, siendo unas más generales y otras más específicas. En esta tesis se propone una medida de centralidad adaptada, con cierta generalidad para poder ser aplicada a diferentes ámbitos, pero pensado especialmente para el problema que se plantea.

1.2. FUNDAMENTACIÓN DE LOS METALES VÍTREOS

Los denominados metales vítreos o también vidrios metálicos son materiales metálicos con una estructura desordenada a escala atómica. Estos metales no son cristalinos como la mayoría de los metales y por tanto su estructura atómica es compleja y no sigue leyes de orden conocidas. Por tanto, poseen en su nombre el adjetivo de vidrios puesto que hace referencia a la condición de los materiales en los cuales se produce una estructura así de desordenada en forma directa desde el estado líquido durante la solidificación. La formación de los metales vítreos se realiza por tanto a partir de un proceso de solidificación extremadamente rápido,

como son la deposición física de vapores, las reacciones de estado sólido, la implantación de iones, hilado en estado de fusión (*melt spinning*) y aleación mecánica. Existe controversia entre los autores en la denominación de algunos metales amorfos como vítreos si se han obtenido a partir de estas técnicas, pero la norma general incluye a todos los materiales de este tipo en la clasificación de metales vítreos con independencia del proceso de obtención.

Si enfriamiento es del orden de millones de grados por segundo, es demasiado rápido para permitir la formación de cristales y el material entonces queda "atrapado" en estado vítreo. Más recientemente, se han obtenido una serie de aleaciones con tasa de enfriamiento crítica lo suficientemente baja como para permitir la formación de estructuras amorfas en capas gruesas (más de 1 milímetro). Estos se conocen como vidrios metálicos masivos (*bulk metallic glasses*, BMG).

El estudio de la formación de los materiales vitreos se ha llevado a cabo a lo largo de la historia a través de criterios empíricos basados en temperatura térmica característica (temperatura de transición vítrea, cristalización y temperatura de líquidos), así se han propuesto en varios trabajos para evaluar y predecir la formación la habilidad de capacidad de formación de vidrio (GFA) de los metales vítreos [22-23]. Sin embargo, en cierto modo, los parámetros empíricos propuestos no son universalmente aplicables a la familia de los metales vítreos, aunque algunos hayan sido razonablemente eficaces para evaluar la capacidad de formación de vidrio [24,25]. Además, actualmente se están estudiando nuevos parámetros empíricos como se corrobora en [26,27]. El principal problema de la aplicación de estos estudios, basados en métodos empíricos, es que se han basado principalmente en pruebas de ensayo y error [28]. Este enfoque es bien conocido por ser ineficiente en términos de tiempo y coste. Por lo tanto, el desarrollo de nuevos enfoques para el diseño de los metales vítreos, en particular los químicamente complejos con buenas propiedades de formación de vidrio sigue siendo un tema abierto para la comunidad científica.

En la última década se han producido avances significativos en la ciencia de datos y los modelos de aprendizaje automático basados en el aprendizaje supervisado para intentar dar respuesta a los problemas relacionados con el diseño de metales vítreos, estableciendo parámetros clave como la capacidad de formación de vidrio (GFA) de las aleaciones, las temperaturas características relacionadas con

dicho parámetros y otras propiedades mecánicas y magnéticas. Así pues, un algoritmo clasificador como el clasificador de vectores de soporte (SVC) [29] o el clasificador de bosque aleatorio (RFC) [30,31]. (o redes neuronales de retropropagación (BPNN) [32] o redes neuronales gráficas [33], entre otros, se han utilizado para etiquetar las MG en diferentes categorías. Cabe señalar que en algunas investigaciones los autores entrenaron un algoritmo de red neuronal de retropropagación construyendo un conjunto de datos compuesto por 3.277 composiciones de aleaciones ternarias, a través de las cuales se pueden identificar con precisión las clases de vidrio metálico y vidrio no metálico [32]. Del mismo modo, algoritmos de regresión como la regresión vectorial (SVR) [34], la regresión de bosque aleatorio (RFG) [35,36], red neuronal basada en la correlación (CBNN) [37], se han utilizado para construir los modelos de predicción composición-propiedades. Sin embargo, los investigadores se enfrentan a retos relacionados con el desarrollo de modelos de aprendizaje automático fiable y adaptable debido a la falta de suficientes datos sobre la capacidad de formación de vidrio de alta fiabilidad para los diferentes tipos de vidrios metálicos [38].

La capacidad de formación de metales vítreos depende de numerosos parámetros metálicos, así como de la combinación de los mismos. No obstante, no se ha encontrado un parámetro determinante para la formación de metales vítreos que explique en su totalidad la formación de los mismos. El parámetro temperatura de transición vítrea reducida (Trg) es uno de los parámetros empíricos más analizados para predecir la formación de metales vítreos a pesar de sus limitaciones [39,40] siendo uno de los parámetros que más se siguen utilizando para la predicción de los mismos [41-43]. Para la mayoría de los vidrios metálicos, el intervalo de formación vítrea coincide con la región eutéctica donde el metal fundido es estable a una temperatura más baja que en las otras regiones del diagrama de fases. Por lo tanto, la temperatura de transición vítrea reducida se calcula mediante la ecuación:

Ecuación 1. Temperatura de transición vítrea reducida

$$Trg = Tg/Tl$$

donde Tg es la temperatura de transición vítrea (la temperatura a la que un material vítreo metálico pierde su estado vítreo) y Tl es la temperatura de liquidus

(la temperatura a la que un material es totalmente líquido, y la temperatura máxima a la que los cristales pueden coexistir con la fusión en equilibrio termodinámico).

Este parámetro se utiliza para evaluar la capacidad de formación de vidrio (GFA) de una aleación [44]. Los grandes valores de la temperatura de transición vítrea reducida en los vidrios metálicos corresponden a la mayor capacidad de formación vidrio [45]. La investigación actual confirma el papel de *Trg* como un parámetro fundamental en la capacidad de formación de vidrio, utilizando el mismo como uno de los parámetros a determinar. La influencia del zinc en la capacidad de formación de vidrio en aleaciones Al-Cu-Mg (Zn) [46], la influencia del estaño en la capacidad de deformación plástica de los vidrios metálicos de Fe-Si-B-P-Sn, [47] o la estructura y propiedades mecánicas de las aleaciones de Cu45Zr48Al, [48], por lo que en este estudio se considera el factor de relación más relevante entre los diferentes elementos que componen los metales vítreos. Estos estudios y notación nos permiten establecer la base de la construcción de la red a partir de la implementación en la misma del parámetro empírico de formación de metales vítreos como es la temperatura de transición vítrea reducida.

Teniendo en cuenta las investigaciones realizadas, sobre la formación de vítreos metálicos, en la presente Tesis Doctoral se ha llevado a cabo el estudio y análisis de la formación de los metales vítreos conocidos mediante la aplicación de una medida de centralidad adaptada. El objetivo principal es conocer las características y la importancia de cada elemento en la composición del material, así como las relaciones entre todos los elementos que lo componen. Este conocimiento no sólo permite comprender el papel de cada elemento en la formación de vidrios metálicos, sino que también es útil como hoja de ruta para futuras investigaciones sobre la generación de nuevos metales vítreos o la mejora de los ya existentes.

II –JUSTIFICACIÓN

II - JUSTIFICACIÓN

La presente Tesis doctoral se ha desarrollado en distintas fases de ejecución para poder abordar los objetivos específicos propuestos. Para ello, el primer paso ha sido el estudio e implementación de medidas de centralidad tradicionales en la resolución de problemas relacionados con redes complejas, aplicándolo en un problema urbanístico (accesibilidad de las rutas turísticas). En la segunda fase del desarrollo de la Tesis se han propuesto el desarrollo de una nueva medida de centralidad adaptada al problema expuesto en el objetivo general (metales vítreos), que permitan solucionar las debilidades de las medidas de centralidad tradicionales en este caso concreto. En la tercera fase se ha perfeccionado y propuesto la medida de centralidad adaptada al tipo de redes complejas que representan los metales vítreos para poder analizar el comportamiento de cada nodo a nivel local y global. Finalmente, se analizan los metales vítreos existentes a partir de esta medida de centralidad propuesta, y se proponen posibles nuevas composiciones a partir de dicho análisis.

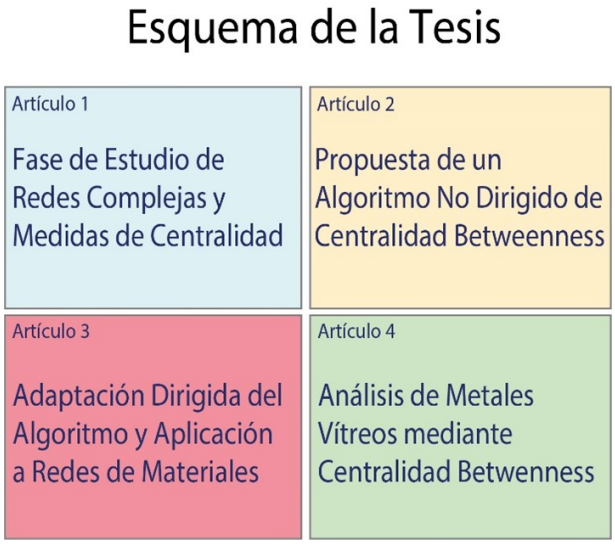


Figura 1. Esquema de los artículos que comprenden la presente Tesis Doctoral.

La tesis está compuesta por cuatro artículos (ver Figura 1), los cuáles serán detallados en capítulos siguientes.

En las siguientes subsecciones se describe brevemente cada artículo, y se expone la relevancia de cada uno dentro de la consecución de los objetivos de la tesis.

2.1. FASE DE ESTUDIO DE REDES COMPLEJAS Y MEDIDAS DE CENTRALIDAD

En este artículo se analiza la accesibilidad de una red urbana para la propuesta de rutas turísticas adaptadas. En concreto, se realiza un caso de estudio de la accesibilidad turística de la ciudad de Ávila (España), estudiando e implementado las medidas clásicas para conocer el alcance de los mismos en un problema concreto. En la siguiente Figura 2 se puede observar el esquema global del artículo o *graphical abstract*.

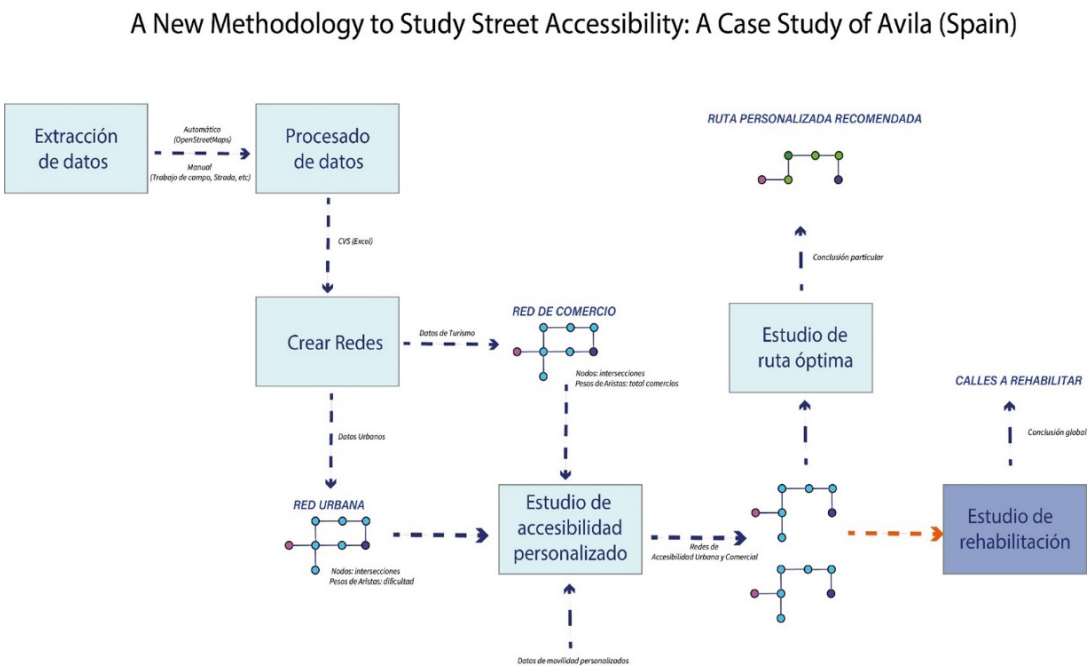


Figura 2. Esquema de la metodología seguida para la realización del estudio perteneciente al artículo 1.

Se trata de un artículo preliminar de base teórica para la tesis. En primera instancia, se hace una toma de contacto con la representación de una base de datos con información diversa sobre accesibilidad urbana.

A continuación, se propone un nuevo algoritmo para cuantificar y etiquetar la dificultad de las calles de una ciudad mediante diferentes parámetros de accesibilidad, como la pendiente de una calle, la existencia de barandillas, aceras o escaleras, etc.

Además, se muestra sobre el ejemplo del casco histórico de la ciudad de Ávila, en Castilla y León (España), la identificación de las zonas más inaccesibles de la ciudad utilizando medidas de centralidad y analizando los efectos en términos de accesibilidad al comercio y servicios de la ciudad.

En esta fase, se han estudiado, implementado y aplicado diferentes medidas de centralidad clásicas, como grado, cercanía, intermediación o *PageRank*. Después de una comparativa de resultados, se toma la decisión de usar un algoritmo de *PageRank* adaptado (APA). El motivo es que el citado algoritmo tiene en cuenta tanto la topología de la red como los datos asociados a la misma.

Finalmente, adicionalmente a las redes de turismo, se aplica dicha medida de centralidad para analizar el comercio, y se propone la restauración de ciertas calles para mejorar la accesibilidad a un mayor número de calles y comercio de la ciudad, y optimizar las rutas turísticas de la ciudad.

A raíz de este estudio de medidas de centralidad, se estudia la información disponible sobre una base de datos con todos los materiales vítreos conocidos hasta la fecha (más de 500) para poder representar dicho problema como red compleja, y se llega a la conclusión de la necesidad de **adaptar una medida de centralidad de intermediación para analizar el comportamiento y las características de cada**

elemento dentro de la red, ya que la importancia de un material radica en la conectividad de los diferentes elementos químicos que la componen.

2.2. PROPUESTA DE UN ALGORITMO NO DIRIGIDO DE CENTRALIDAD BETWEENNESS

Una vez se ha decidido que la medida de centralidad buscada para analizar los metales vítreos debe estar basada en intermediación o *betweenness*, se estudian las medidas de centralidad existentes, llegando a la conclusión de que ninguna centra el foco en una característica importante de la red, como es la densidad de cada nodo dentro de forma dinámica, tanto a nivel local (su clúster o comunidad) como global.

Por ello, con la finalidad de adaptar esta centralidad al problema, se propone una nueva medida de centralidad basada en caminos aleatorios o *random walks* desde la perspectiva de redes densas, suponiendo que éstas están totalmente conectadas y libres de bucles.

La perspectiva usada se basa en el algoritmo de predicción de enlaces *Return Random Walk* [21], el cual da más relevancia a los nodos que pertenecen a comunidades o clústeres densos mediante una evaluación de probabilidad de transición en dos pasos (ida y vuelta). En otras palabras, un nodo será relevante a nivel de conectividad si enlaza dos nodos cualesquiera y a su vez no depende de él mismo para retornar o deshacer el camino hecho (ver Figura 1 del artículo 2, y el ejemplo de la sección 3.2 del mismo).

La medida de centralidad propuesta, denominada *Two-way random walk betweenness centrality* (2RW), proporciona una clasificación de los nodos de una red centrando la importancia en el refuerzo de las comunidades densas en función de su papel (local o global): relevancia intra-clúster (cómo el nodo refuerza la red) o inter-clúster (cómo de importante es el nodo en la conectividad global de la red y qué implicaciones supone si se elimina). En la Figura 3 se muestra un resumen del método.

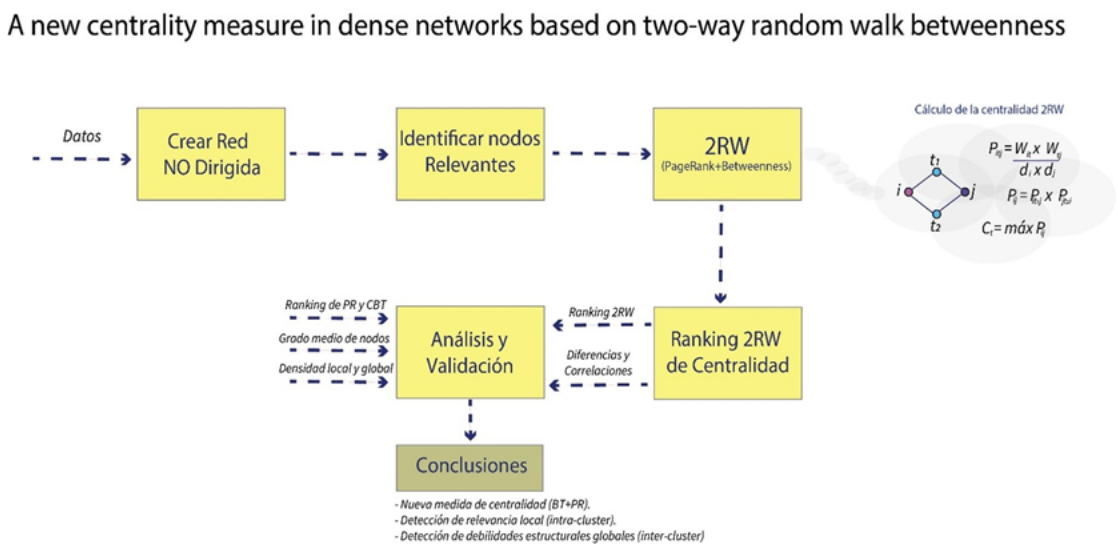


Figura 3. Esquema de la metodología de la medida de centralidad Two-way random walk betweenness centrality (2RW) y su aplicación en el artículo 2

Aunque la medida resulta muy útil para diferentes aplicaciones (por ejemplo, redes sociales), está orientada inicialmente para grafos no dirigidos, y no da más información sobre el comportamiento de un nodo que la conectividad global y relevancia local. No se obtienen grandes resultados en materiales debido al carácter dirigido de algunas de las propiedades más relevantes de los elementos químicos, como la temperatura de transición vítrea reducida (Trg). Es por ello que se necesita una adaptación de la medida de centralidad a redes dirigidas con la finalidad de aportar información sobre el comportamiento específico de cada nodo dentro de la red.

2.3. ADAPTACIÓN DIRIGIDA DEL ALGORITMO Y APLICACIÓN A REDES DE MATERIALES

Para adaptar la medida de centralidad 2RW al problema planteado, se desarrolla una nueva medida de centralidad para redes dirigidas: el algoritmo de centralidad de intermediación mediante caminos aleatorios bidireccional o *Directed*

Two-Way Random Walk Betweenness Centrality (D2RWBT). Esta nueva propuesta caracteriza a los nodos de una red mediante cuatro índices que miden la importancia de un nodo dentro de su comunidad o clúster, su conectividad a nivel global, así como otras propiedades que describen el comportamiento del mismo dentro de la red, como la densidad del clúster al que pertenece, el rol que juegan en el mismo, etc. En la figura 4 se muestra el *graphical abstract* del artículo.

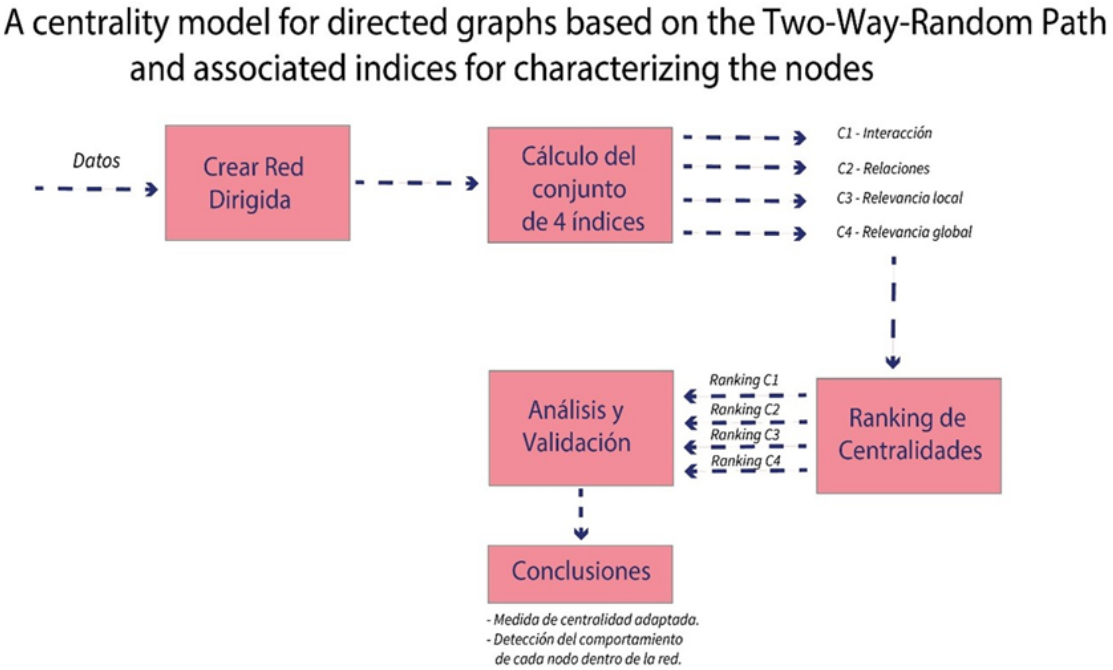


Figura 4. Esquema gráfico de la metodología usada para el algoritmo D2RWBT y su aplicación en el artículo 3

Por tanto, con esta nueva medida de centralidad D2RWBT, no solo se puede analizar la importancia de cada elemento o nodo en su red, sino que se puede analizar y caracterizar el comportamiento del mismo dentro de dicha red, obteniendo información sobre características relevantes como si el nodo está situado en un cluster denso o disperso, si el nodo es relevante desde el punto de vista de su centralidad, si el nodo desempeña un papel importante en la conectividad entre clústeres (inter-clúster) o la relevancia en la cohesión del mismo clúster al que pertenece (intra-clúster).

En resumen, se ha desarrollado un algoritmo para redes dirigidas, basado en la centralidad de intermediación, que comprende diferentes índices que ayudan en la caracterización y comprensión del comportamiento de los nodos dentro de una red, tanto a nivel local como global, aplicable a diferentes problemas de redes complejas, y especialmente diseñada para redes de materiales. Se trata de un algoritmo general que puede aplicarse a problemas de diferentes ámbitos, pero pensado con detalle para su aplicación en las redes objeto de estudio en esta tesis.

2.4. ANÁLISIS DE METALES VÍTREOS MEDIANTE EL ALGORITMO DE CENTRALIDAD BETWEENNESS ADAPTADO

Finalmente, se aplica el algoritmo sobre la red de materiales vítreos. Antes de ello, hay que tener en cuenta dos aspectos relevantes:

- Decisión de la información relevante a estudiar en lo relativo a materiales vítreos.
- Construcción de la red a partir de la base de datos.

El primer punto se basa en un estudio de las diferentes características de las composiciones vítreas, teniendo en consideración que los pesos de las aristas deben de ser factores empíricos que influyen en la capacidad de formación de los metales vítreos. Se analizan los parámetros empíricos más utilizados y ratificados en la formación de metales vítreos y pese a las limitaciones que pueden tener cada uno de estos parámetros se decide implementar uno de los más verificados como es la temperatura de transición vítrea (*Trg*), tal y como se muestra en la Figura 5.

En segundo lugar, se parte de una base de datos con todos los metales vítreos conocidos (hoy en día más de 500). La construcción de la red se centra en el estudio de las uniones de los elementos que conforman cada uno de los metales vítreos que existen. Cada elemento de la red será considerado como un nodo y el peso de su arista recae en el porcentaje de cada elemento junto con el valor de Trg del material. Las uniones entre los nodos se realizan desde el elemento principal en cuanto a porcentaje de composición hacia los secundarios, conformándose una red dirigida, tal y como se puede ver mediante un ejemplo de formación de la red en la figura 2 del artículo 4.

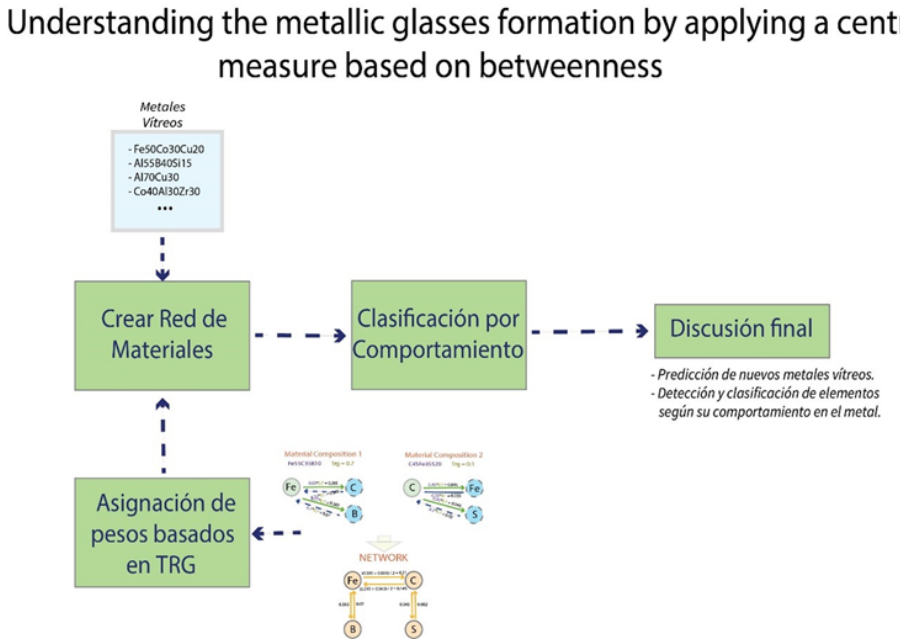


Figura 5. Esquema gráfico resumen de la metodología llevada a cabo para la elaboración del artículo 4

Una vez generada la red compleja se calcula su centralidad, basándose en las posiciones de relevancia que ocupan cada uno de los elementos dentro de la red. Con la finalidad de comprender el papel que estos elementos juegan en la formación de metales vítreos dentro de redes densas y dispersas, se secuencian los resultados atendiendo a la comparativa con medidas de centralidad clásica. Como resultado se obtiene sí el elemento actúa como un importante nodo de transición entre clústeres (inter-clúster) o si el nodo es importante dentro del mismo cluster (intra-clúster). La clasificación de los elementos se realiza siguiendo el esquema de decisión mostrado en la figura 2 del artículo 3.

Finalmente se exponen las conclusiones en base a los resultados obtenidos de cada elemento desde las cuales se puede no sólo predecir la formación de posibles nuevos metales vítreos sino la capacidad de sustitución de cada elemento por uno afín, es decir que juega el mismo papel dentro del mismo tipo de red.

III – OBJETIVOS

III - OBJETIVOS

El objetivo general (OG) de esta tesis es el estudio de los materiales vítreos existentes haciendo uso de medidas de centralidad de redes complejas adaptadas para entender el comportamiento de cada elemento dentro de un compuesto y la propuesta de nuevos metales vítreos.

Para ello, se tienen en cuenta los siguientes objetivos específicos:

- OE1: Estudio de las medidas de centralidad clásicas.
- OE2: Adaptación y estudio de la representación de metales vítreos como red compleja.
- OE3: Propuesta de una medida de centralidad personalizada para analizar el comportamiento de cada elemento que forman los compuestos.
- OE4: Análisis de los metales vítreos existentes y propuesta de nuevos metales mediante una redes complejas y medidas de centralidad personalizadas.

Una vez estudiadas e implementadas las medidas de centralidad clásicas en el problema de estudio de las redes urbanas se analiza el problema de formación de metales vítreos con profusión entendiendo como se relacionan los elementos y cuáles pueden ser los parámetros clave en la relación de los mismos. En base a este estudio preliminar se analizan las deficiencias de las medidas de centralidad clásicas para dar respuesta al problema clave, el entendimiento de formación de metales vítreos.

Las medidas de centralidad basadas en la centralidad de la intermediación (*betweenness*) no tienen en cuenta una característica importante de la red, como es la densidad del clúster que define donde se encuentra cada nodo de la red con respecto a los otros. Con la finalidad de solventar esta debilidad se ha desarrollado una nueva medida de centralidad, basada en caminos aleatorios (*random walk*), desde la perspectiva de redes densas, suponiendo estás libres de bucles.

La nueva medida de centralidad propuesta, denominada *Two-way random walk betweenness centrality* (2RW), cuantifica la importancia de un nodo, en la

transmisión de información, a través de las relaciones entre cuatro nodos diferentes –nodo origen *i*, nodo destino *j*, nodo de transición a la ida *t* y nodo de transición a la vuelta *k*-.

La base de esta propuesta es el algoritmo *Return Random Walk* [21], que refuerza las comunidades densas mediante la evaluación de grafos utilizando una matriz de probabilidad de transición de dos viajes. Se ha tenido en cuenta que trabajamos con grafos sin nodos aislados ni bucles propios.

La medida de centralidad 2RW propuesta, en resumen, proporciona la clasificación de los nodos de una red, con la característica principal de reforzar las comunidades densas en función de su papel: relevancia intra-clúster (cómo el nodo refuerza la red) o inter-clúster (cuánto se desconecta la red se desconecta cuando eliminamos el nodo objetivo).

Una vez propuesta la medida de centralidad nueva 2RW se analiza que el nuevo método sólo se propone para grafos no dirigidos, y no da más información sobre el comportamiento de un nodo. Para solventar esta problemática se propuso otra nueva medida de centralidad adaptada la cual sea relevante también para grafos dirigidos: algoritmo de centralidad de camino aleatorio bidireccional (D2RWBT). Este nuevo modelo de centralidad se basa en cuatro índices que miden la importancia de un nodo como enlace intermedio estudiando algunas propiedades de comportamiento. Por tanto, con esta nueva medida D2RWBT, se describe el comportamiento del nodo dentro de una red obteniendo información sobre algunas características relevantes como; si el nodo está situado en un cluster denso o no, si desde el punto de vista de su centralidad el nodo es importante, si el nodo desempeña un papel importante, como nodo de transición, entre clúster (inter-clúster) o si el nodo es importante en el mismo cluster (intra-clúster).

IV – ARTÍCULOS

IV -ARTÍCULOS

En esta sección se muestran las contribuciones de esta tesis por compendio tal y como aparecen en las revistas.

4.1. ARTÍCULO 1

Este artículo ha sido publicado en una edición especial titulada *The Application of Artificial Intelligence Techniques on Geo-Information Systems*, en la revista *International Journal of Geo-Information*, en la editorial *Multidisciplinary Digital Publishing Institute* (MDPI).

Se enmarca dentro de un proyecto regional competitivo de la Diputación de Ávila, titulado *Generación de rutas turísticas personalizadas accesibles de forma dinámica*.

A continuación, se muestra algunos indicios de calidad de la revista, y otros indicios adicionales de otras plataformas.

Tabla 2. Otros indicios de calidad del artículo 1

Fecha de publicación	Citescore	5- Year Impact Factor	Lecturas en ResearchGate
Julio 2021	5.0	3.165	50

International Journal of Geo-Information

MDPI

Article

A New Methodology to Study Street Accessibility: A Case Study of Avila (Spain)

Manuel Curado [1,*], Rocio Rodriguez [2], Manuel Jimenez [1], Leandro Tortosa [3] and Jose F. Vicent [3]

[1] Campus Los Jeronimos, Polytechnic School, Catholic University of Murcia, E-30107 Murcia, Spain; mjimenez6@alu.ucam.edu
[2] Department of Mechanical Engineering, Catholic University of Ávila, Calle Canteros, E-05005 Ávila, Spain; rocio.rodriguez@ucavila.es
[3] Department of Computer Science and Artificial Intelligence, Campus de San Vicente, University of Alicante, Ap. Correos 99, E-03080 Alicante, Spain; tortosa@ua.es (L.T.); jvicent@ua.es (J.F.V.)
* Correspondence: mcurado@ucam.edu

Abstract: Taking into account that accessibility is one of the most strategic and determining factors in economic models and that accessibility and tourism affect each other, we can say that the study and improvement of one of them involved the development of the other. Using network analysis, this study presents an algorithm for labeling the difficulty of the streets of a city using different accessibility parameters. We combine network structure and accessibility factors to explore the association between innovative behavior within the street network, and the relationships with the commercial activity in a city. Finally, we present a case study of the city of Avila, locating the most inaccessible areas of the city using centrality measures and analyzing the effects, in terms of accessibility, on the commerce and services of the city.

Keywords: centrality measures; urban networks; accessibility

Citation: Curado, M.; Rodriguez, R.; Jimenez, M.; Tortosa, L.; Vicent, J.F. A New Methodology to Study Street Accessibility: A Case Study of Avila (Spain). *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 491. https://doi.org/10.3390/ijgi10070491

Academic Editor: Wolfgang Kainz

Received: 29 May 2021
Accepted: 19 July 2021
Published: 20 July 2021

1. Introduction

Urban tourism has been expanding globally over the last few decades [1]. The current mobility [2] and the world tourism dynamics are in constant change and, at the same time, they are heavily structuring the trajectories of cities [3]. In this context, governments have relied on strategic planning to shape cities to be more touristic. However, these changes usually focus on a healthy tourist without mobility problems.

Accessibility and tourism affect each other, and therefore the study and improvement of one of them involves the development of the other [4]. The concept of urban accessibility can be understood as the possibility that a person with any ability can reach and use places, as well as carry out activities of interest without barriers [5] and it is one of the most strategic factors in economic models. Some of research has managed to build accessibility indexes that can generally capture transport access to an area [6], understanding that accessibility is driven by good infrastructure [7]. As already mentioned, accessibility has great impact on a key economic driving sector such as tourism [8]. Many scientific studies focus on tourist trends, proposing theories about their behavior to predict even their seasonality [9,10]. Within the different typologies of tourism, the following research refers to cultural tourism, studying behavior through images of geolocation techniques to even propose tourist routes within the different cities [11].

Extrapolation of the concept of network and its application for the understanding of urban networks is widely accepted and occupies a prominent place. Thus, there are examples about the study of networks in the categorization of the structures of the city [12] and analysis of connectivity in cities [13,14] among others.

The study of complex network theory continues to be a challenge in research given a large number of applications in many fields, such as economics, biology, communication,

transportation. An increasing amount of research supports the implementation of this theory, reaffirming the great importance that nodes within a network can have and the significance of the possible paths that prevail in it [15–17]. Complex network theory continues to be developed with the concepts, algorithms, methods, and models proposed in real network analyses [18,19].

For the analysis of the relevance of the role of nodes within the network in network theory, the most commonly used tools have been the centrality measures [20,21]. There exist many investigations on the concept of centrality however, the presence of data in the study of centrality in networks is relatively recent [22–24]. In [22], the authors defined a novel centrality measure, called the Adapted PageRank Algorithm (APA) that establishes a node classification that takes into account the topology and data associated with the nodes.

Studies concerning the impact of accessibility on the organization of the pedestrian network, the location and distribution of activities, as well as urban services, focusing on the elderly [25,26] should be noted. Currently, studies of this relationship are generally divided into two aspects. The first one is based on concepts of gravitation/spatial interaction [27] and the other, on its usefulness for consumption [28–30]. With some of these studies, networks have been carried out that are capable of modeling and predicting tourist activity to analyze deficiencies and propose improvements in it [31].

Moreover, the result of having a World Heritage status has boosted tourism in many cities around the world. A large number of tourists and the side effects, such as the overcrowding of the main spaces in the city, are resulting in the need to develop new strategies to protect heritage assets. One of these cities is Avila (Spain). This city won the first-ever European award for disabled-friendly cities (Access City Award in 2011) with different policies to improve the city through accessible resources [32] or transport cards [33].

In [34], the author analyzes the accessibility of the city attending all special conditions of disabled people to improve access to hotels, restaurants, monuments, etc. However, it is interesting to complement this work with the accessibility of the streets of the city, because there is a lack of strategies about this topic, and is personalized to all people with mobility restrictions. Moreover, these strategies have to incorporate solutions to some problems such as over-tourism, conflicts between residential and tourist uses, and mobility problems. In this paper we propose a novel algorithm for labeling the difficulty of the streets of a city using accessibility parameters such as the slope of a street, the existence of railings, sidewalks, or stairs, etc. In addition, we show a numerical example of the city of Avila, locating the most inaccessible areas of the city using centrality measures and analyzing the effects in terms of accessibility to the commerce and services of the city. Finally, we show how the restoration of some of the city streets could affect the accessibility and commercial impact.

2. Materials

2.1. Data Acquisition

Data was obtained using OpenStretMap (OSM) [35] and Overpass-turbo [36]. The extracted data, such as roads, secondary roads, etc., are tied to the highway tag. From OSM, we extract information about the city as the location of stairs, commercial business, length of all streets, type of pavement, and other useful information. All this data is reviewed with manual labeling from fieldwork activities. With Microsoft's Bing Maps [37] service, the heights of a location given by its coordinates have been extracted. With this information, the slope of the streets in all its sections has been self-calculated. In addiction, the Shuttle Radar Topography Mission (SRTM) data was used to obtain the height of a point in the city, and the efficiency was compared using information obtained with data extracted from a third source, Strava, a popular website used to track users activity via GPS-enabled devices. The result is similar and we decided to use Bing Maps data.

2.2. Accessibility: Parameters

Allowing all people to move around the city (e.g., walking or on wheelchairs), which reduces the use of cars and promotes public transport, are policies that can improve life in a city. In regards to tourism, making the city more accessible could positively affect the economy so as to help all people arrive at a business easily.

However, there are some obstacles or situations that entail mobility problems in different streets. For example, irregular pavement or an excessive slope are architectural barriers to access the different areas in a city. In this paper, we attend to the following problems: (i) Stairs, which pose a terminal obstacle for people with several mobility problems, (ii) slopes or ramps, that do not exceed 1:20, even preferably less, (iii) railings, whose existence will make a slope or a set of stairs more accessible, (iv) sidewalk, whose width should be at least 2.6 m plus wheelchair dimensions, and (v) ground or pavement, where bad conditions, or an irregular state could affect people with mobility problems.

2.3. Commercial and Tourism: Parameters

The data collected for the realization of this study belong to the number and geolocation of shops, hostels, hotels, bars, and businesses related to the tourism sector. The database has been carried out in detail for each of the streets belonging to the historical part of the city center. The accessibility measures for the different streets analyzed have also been integrated into the database. As previously described, the objective is to find the implication of the accessibility of different streets for a city's economic and tourist development.

3. Methods and Procedure

In this section, we explain the proposed algorithm for labeling the difficulty of a city's streets using accessibility parameters such as the slope of a street, the existence of railings, sidewalks, or stairs, etc. First, we show the APA Centrality measure to classify the streets of a city. Then, we explain the labeling of the difficulty estimation value of each street.

3.1. Centrality Measure

In this subsection, we briefly describe the measure of centrality used to classify the streets of the city.

The Adapted PageRank Algorithm (APA [22]) was proposed in 2012. This algorithm is an adaptation of the PageRank model for spatial networks with data, and it was initially thought of for urban street networks. The APA model is based on the construction of a matrix M_{APA} that has two terms, one related to the topology of the graph and the second one related to the information of the network:

$$M_{APA} = (1 - \alpha)P - \alpha V.$$

The eigenvector $\vec{x}$ of the matrix M_{APA} provides a classification of the network nodes according to the connectivity and presence of data.

3.2. Accessibility Limits

A city has different problems attending to accessibility limits, such as big slopes, stairs, uneven ground, cobblestone floor, inadequate, or nonexistent sidewalks, etc. All of these types add to the mobility problems of each tourist (use of wheelchairs, baby carriages, and older people with different mobility limitations) producing constraint paths or streets that certain types of people can not visit. To identify and classify the streets of a city according to some types of limitations or parameters, an algorithm is proposed.

In Figure 1, we show the decisions made about the labeling of street difficulty:

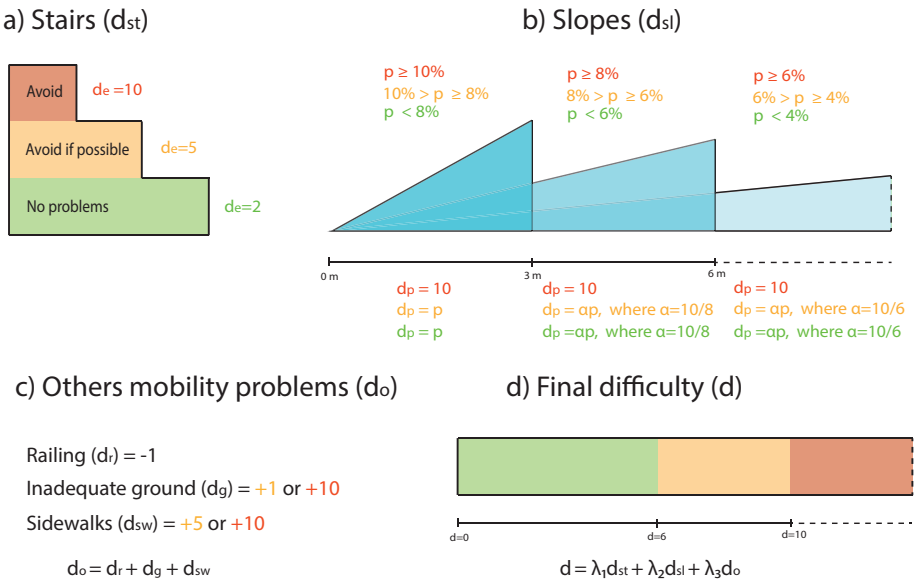


Figure 1. Accessibility scheme: Calculation of the difficulty of a street.

3.2.1. Stairs Difficulty Estimation (d_{st})

Following the specifications of stairs in [38], we estimate three levels of stairs difficulty ordered by color, using the indications of the users (see Figure 1a):

- Avoid stairs: This difficulty level is the most restrictive. It indicates all people that can not use stairs (e.g., people in wheelchairs, or people who prefer the option of avoiding all stairs). All streets in this level have been registered as inaccessible paths (the difficulty of the street is $d_{st} = 10$);
- Avoid if possible: This difficulty level means that a person can use stairs, but it is better not to use them if possible, due to mobility problems. All people belonging to this level prefer not to use stairs (e.g., people with a baby carriage or people who have selected this option). All streets in this level have been registered as moderate-difficulty paths (the difficulty of the street is $d_{st} = 5$);
- No mobility problems: This difficulty level means that there are no stairs constraints. People can use all kind of stairs. All streets in this level are low difficulty paths (the difficulty of the street is $d_{st} = 2$).

3.2.2. Slope Difficulty Estimation (d_{sl})

When attending to architectural design considerations, ramps are a problem because building or monument entrances could be inaccessible due to the difference between indoor and outdoor levels. Governments and organizations work on policies to improve the accessibility of these places. However, the slopes o streets affect people with mobility constraints.

Many streets with an elevated slope result in inaccessible routes due to differences in level. In Table 1, we show the maximum recommended slope of a ramp, following the specifications of Technical Accessibility Code of Spain Buildings [38], and their associated difficulty based on a parameter α that represents the ration between length and slope obtained from the maximum ramp permitted for this code (see Figure 1b).

With respect to the ratio α, we select different values according to the maximum value of the range of the slope in each level. If we select the minimum value of the range (0.6 and 0.4 in orange and green levels, respectively) we will reduce the threshold of the difficulty of a street and more streets will labeled with less difficulty.

Table 1. The difficulty of a street attending to its recommended slope and length.

Length (*len*)	Slope (*sl*)	Difficulty (d_{sl})
$len < 3$ m	$sl \geq 10\%$	$d_{sl} = 10$
	$10\% > sl \geq 8\%$	$d_{sl} = 0.8sl$
	$sl < 8\%$	$d_{sl} = 0.6sl$
$3\text{ m} \leq len < 6$ m	$sl \geq 8\%$	$d_{sl} = 10$
	$8\% > sl \geq 6\%$	$d_{sl} = 0.8sl$
	$sl < 6\%$	$d_{sl} = 0.6sl$
$len \geq 6$ m	$sl \geq 6\%$	$d_{sl} = 10$
	$6\% > sl \geq 4\%$	$d_{sl} = 0.8sl$
	$sl < 4\%$	$d_{sl} = 0.6sl$

3.2.3. Others Mobility Problems

There are other obstacles in streets that directly affect the mobility of people with limitations of movements. Regarding the ground, an inadequate floor (cobblestone floor, bad condition floor, etc.) could limit the access of many people, such as a person with a wheelchair. Moreover, the conditions of sidewalks are crucial. If a street has no sidewalks, or their conditions are irregular, the use of these streets is restricted. Moreover, we have labeled as problematic streets other cases such as the kerb height of sidewalks, which can be a major obstacle for wheelchair users. In the case of a pedestrian street or a living street, there is no difficulty in terms of sidewalks (if the pavement is irregular, the difficulty is applied in this category). If a street where there are stairs or a ramp has a railing, that helps to improve the mobility of the people and the difficulty of the street is reduced.

In Table 2, we show the assigned values of these obstacles. We can observe that railings are the only element that reduces the accumulative difficulty (see Figure 1c).

Table 2. Difficulty of the different obstacles in the street.

Obstacle	Moderate State	Extreme State
Ground	$d_o+ = 1$	$d_o+ = 10$
Sidewalks	$d_o+ = 5$	$d_o+ = 10$
Railing	$d_o- = 1$	$d_o- = 1$

3.2.4. Final Difficulty Estimation (D)

To calculate the final difficulty of a street, we sum all partial values of mobility problems as follows:

$$D = \lambda_1 d_{st} + \lambda_2 d_{sl} + \lambda_3 d_o. \quad (1)$$

The parameters λ_1, λ_2, and λ_3 represent a weighting of the different restoration levels. With these parameters we could control the state of different elements of a city (e.g., if a sidewalk can not be restored, we assume a value of 1).

We propose 3 accessibility levels, labeled by color (see Figure 1d), to represent the accessibility of the streets. With respect to the mobility restrictions for a specific person, we specify another three levels from a low restrictive level, when a person has few mobility problems (e.g., a person who can use stairs), to a high restrictive level, when a person has several mobility restrictions (e.g., a person who can not use stairs). In Table 3, we can see the accessibility levels, the final decision, and users characteristics of different streets. For instance, a street labeled with a final difficulty of 12, will be an inaccessible street that only healthy people can use.

Table 3. Three accessibility levels of a street.

Difficulty	Colour Level	Street Label	Users
$D \geq 10$	Red	Inaccessible	Healthy people
$10 > D \geq 6$	Orange	Problematic	People with moderate mobility
$D < 6$	Green	Accessible	All people

4. Experiments and Results

4.1. Constructing Urban City Networks

In this section, we study a case study: The city of Avila, Spain. Is an awarded city in terms of accessibility.

The construction of its accessibility and commerce networks is shown. Note that both networks are undirected because we focus on the difficulty of the streets independent of the direction of the path, and we are interested in the commerce information of them. Testing different values of lambda, as $\lambda = 0.8$, we observe that few streets have been affected (e.g., only seven streets will change from a red level color to orange, and studying these streets, the accessibility labeled could be wrong). For that, we assume the lambda parameter is equal to the unit because the city can not be strongly restored due to its historical restrictions (UNESCO World Heritage City). With respect to the ratio α, we have tested the maximum value (0.8 and 0.6 in orange and green levels, respectively) and the minimum value of the range (0.6 and 0.4). However, in our case study, there are no differences (only two streets will change from the red to orange level, and we see as the state of both streets could be ambiguous).

4.1.1. Accessibility Network

Given a list of final difficulties of streets or partial sections of these streets, we construct a network through a primary graph as follows. Let $G_A(V, E, W)$, an undirected weighted graph where V is the set of n nodes, representing the intersections between streets, E is the set of e edges or streets, and W is a weighted matrix, which contains the difficulty values of all streets (D).

We create three different graphs representing the accessibility according to the different constrained mobility levels: (i) Low, a graph with the accessible streets that all people could use (see Figure 2), (ii) moderate, adding to accessible streets some problematic streets (see Figure 3) that people with moderate mobility could use, and (iii) high, where all inaccessible streets (see Figure 4) have been removed, and only people without mobility constraints could access.

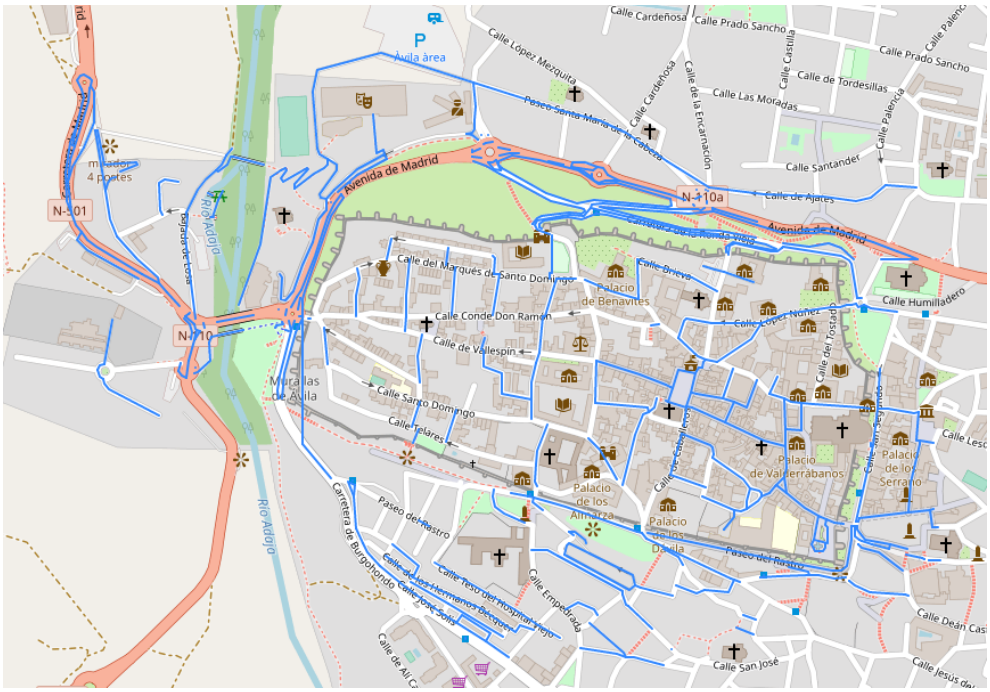


Figure 2. Accessible streets. In blue, we show all streets that are accessible to everyone, regardless of their mobility. This map of the city represents the accessible streets (correspond with the green level color in Table 3) after removing the inaccessible streets for a person with high mobility restrictions is presented.

Figure 3. Problematic streets. In blue, we show all streets labeled as intermediate difficulty (see the orange color level in Table 3).

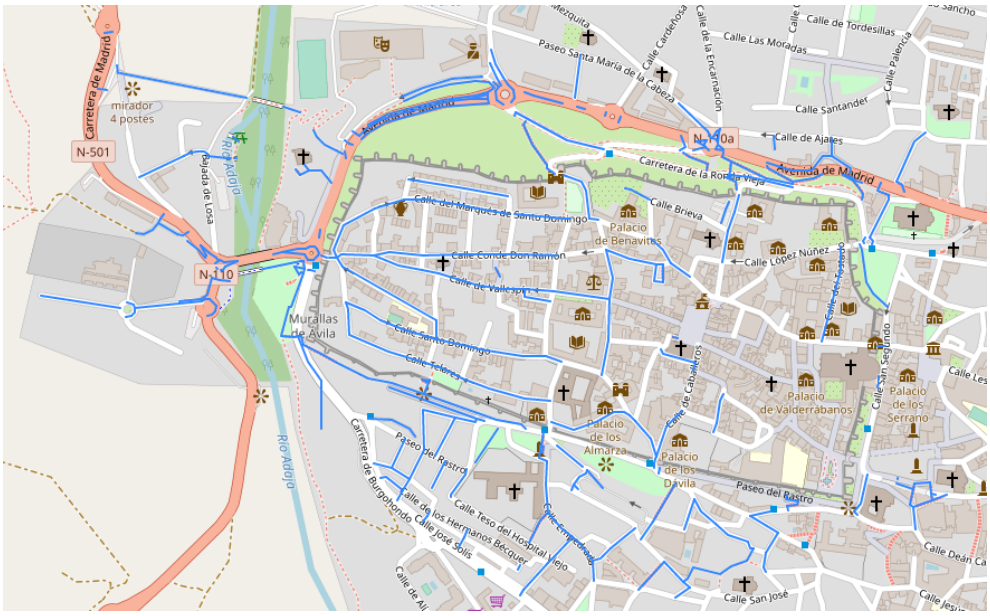


Figure 4. Inaccessible streets. In blue, we show all streets that can not be accessed by people with important mobility problems (e.g., people with wheelchairs), that corresponds with the red color level in Table 3.

4.1.2. Commercial Network

In Figure 5, we show the street with the most commercial businesses in the studied area of the city. We list, in Table 4, the top streets regarding the number of businesses.

Given a list of the commercial businesses of every street, we construct a primary graph $G_C(V, E, W)$ (undirected weighted graph) where V is the set of n nodes, representing the intersections between streets, E is the set of e edges (streets), and W is a weighted matrix, which contains the sum of all commercial businesses of each street (explained in Section 2.2).

Figure 5. Top commercial business streets of the city.

Table 4. Main streets in terms of the number of commercial businesses.

Ranking	Street Name	Number of Commercial Business
1	C/de los Reyes Catolicos	20
2	C/Don Geronimo	14
3	Plaza Mercado Chico	14
4	C/Comuneros de Castilla	10
5	Plaza de la Catedral	9
6	C/Alemania	8
7	C/Tomas Luis de Victoria	8
8	C/de Vallespin	7
9	Plaza Jose Tome	7
10	C/de la Cruz Vieja	6
11	C/Pedro de Lagasca	4
12	C/Bracamonte	4

4.2. Detecting Main Commercial and Touristic Areas

First, from the commercial graph $G_C(V, E, W)$, we applied the APA algorithm with the aim to classify the nodes of the network. We rank all nodes in terms of this centrality to detect the main areas of the city. Table 5 presents the top 10 APA ranking of commercial areas from the point of view of the APA centrality index. Note that we calculate the importance of a street from the centrality values of all included nodes in each street. In addiction, we can show the map of the city (Figure 6), drawing in blue the nodes with a high centrality value, that represent the main commercial streets of the city in terms of this measure, and highlighting, in yellow, the main areas of commercial business.

Table 5. Top 10 classification of the commercial areas using the APA centrality index.

APA Ranking Area	Street Name
1	Plaza de la Catedral
2	C/de los Reyes Catolicos
3	C/Alemania
4	Plaza de Santa Teresa
5	C/de la Cruz Vieja
6	Plaza del Mercado Chico
7	C/Don Geronimo
8	C/Estrada
9	Plaza Jose Tome
10	C/Comuneros de Castilla

4.3. Detecting Accessibility Problems

In this section, we construct three different networks according to the three levels of difficulty of urban accessibility (see Table 3). In Figure 7, we show the difficulty of all the streets keeping in mind the mobility restrictions of the people. From the original network, we remove all edges in the primal graph that exceed a certain threshold (5 or 10).

With the aim to evaluate the impact of the accessibility we present, in Table 6, different configurations in terms of different mobility constraints and the accessibility level. The changes in the characteristics of the network result in less density and possible paths (worse connectivity, a reduced average degree) than the original graph. Note that, in this table, the first row shows the network with all streets that a person with low mobility restrictions could access.

Figure 6. Map of most important areas according to the APA centrality measure.

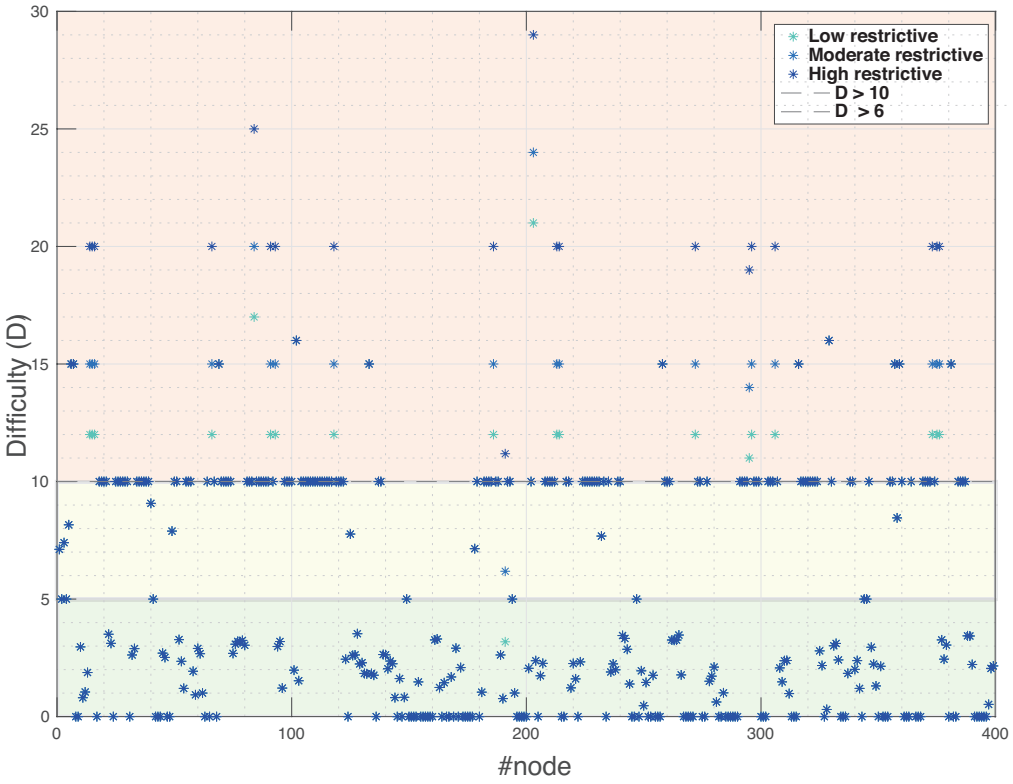


Figure 7. Difficulty values of the main streets of the city. The background color means the street accessibility levels for three different mobility restrictions.

Table 6. Network attributes.

Mobility Constraints	Accessibility Level	Accessible Streets (%)	Avg. Degree
Low	All network	793 (100%)	29.14
Low	Green	452 (56.99%)	3.85
Low	Orange	464 (58.51%)	4.51
Moderate	Green	449 (56.62%)	3.85
Moderate	Orange	464 (58.51%)	4.65
High	Green	445 (56.11%)	3.71
High	Orange	458 (57.75%)	4.43

4.4. How to Increase the Touristic Impact: Densifying the Network

Following the results obtained in Table 6, it is possible to evaluate the inaccessible businesses in a case study where the mobility constraints are low.

In Table 7, we can see that a quarter (25.42%) of the businesses of the city are located on accessible streets for people without major restrictions in mobility, but people with high mobility problems can not access them. Moreover, 8.75% of them are inaccessible for all people with any mobility problem (they are only accessible for healthy people).

Table 7. Commercial accessibility before restoration.

Accessibility Level	Inaccessible Business (%)
Green	25.42%
Orange	8.75%

Some streets could be made more accessible through different actions such as improving the pavement ($D = D - 5$), adding a railing ($D = D - 1$), and improving or adding the sidewalks ($D = D - 5$). In Table 8, we show the number of streets whose accessibility difficulty is slightly higher than the thresholds. For example, for a person with low mobility restrictions, if we restore a street by adding a railing, we can make four streets more accessible, even one of the streets will be fully accessible. Improving the pavement or the sidewalks of 282 streets will be more accessible (orange accessibility level), and 58 of them will be completely accessible (green accessibility level).

Table 8. Restoration streets candidates.

Mob. Constraints	Thresholds Exceeded (α)	#Streets	#Streets
Low	$\alpha < 1$	4	1
	$\alpha < 5$	282	58
Moderate	$\alpha < 1$	2	0
	$\alpha < 5$	281	28
High	$\alpha < 1$	1	1
	$\alpha < 5$	282	25

After the restoration of previous streets, the network is denser than the original, and the number of the commercial inaccessible businesses is reduced from 25.42% to 21.32% with a slight restoration and to 8.76% with a moderate restoration in the best case (see Table 9). With a strong restoration of the city, almost all business will be accessible by most of the people (0.71%).

Table 9. Commercial accessibility after restoration.

Accessibility Level	Thresholds Exceeded (α)	Inaccessible Business (%)
Green	$\alpha < 1$	21.32%
	$\alpha < 5$	8.76%
Orange	$\alpha < 1$	8.61%
	$\alpha < 5$	0.71%

Finally, we show a case study of the restoration of three streets: *Calle del Tostado*, *Calle de Caballeros*, and *Calle de la Cruz Vieja*. After studying these streets, we propose the following actions:

- To widen a sidewalk in *Calle del Tostado*, reducing its accessibility difficulty to 0. There are two restaurants and a hotel located on the street;
- To widen a sidewalk in a small stretch of street *Calle de Caballeros*, reducing its accessibility difficulty to 0. There are seven businesses, a restaurant and a hotel located on the street;
- To add appropriate sidewalks in *Calle de la Cruz Vieja*, reducing its accessibility difficulty to 2.14. There are two restaurants, a hostel, and a business located on the street.

According to these changes, we recalculate the connectivity of the city with APA centrality measure, and we show how the average centrality of different streets changes and the impact that this produces in terms of accessibility on different areas of the city: Historic and commercial streets of the city. Then, in Table 10, a comparison of the average of the APA centrality of main streets of the city in terms of accessibility before and after the restoration of the selected three streets are presented. The second column is the impact of the restoration calculating the difference of the averages between APA centrality measures of the main streets (historical area) before and after restoration and the third column is similar to the second column however by taking the main commercial streets (see Table 5).

For instance, the impact of the restoration (for people with low mobility problems) of these three streets is an improvement of 12.9% (centrality) in the streets belonging to the historic area and 4.71% in the streets more important in terms of business.

Table 10. Comparison of the average of the APA centrality of main streets of the city in terms of accessibility before and after the restoration of the selected three streets.

Mob. Constraints	Historic Center Impact	Commercial Impact
Low	3.6436×10^{-4} (+12.90%)	9.0455×10^{-5} (+4.71%)
Moderate	3.6432×10^{-4} (+12.89%)	9.0446×10^{-5} (+4.69%)
High	3.3467×10^{-4} (+11.41%)	1.8928×10^{-5} (+0.95%)

From these results, we can see the advantages of the restoration of some streets. Moreover, if we adapt several streets of the city we achieve an improvement of their accessibility and reduce the number of inaccessible businesses. We have a more connected city (a denser graph, and a uniform degree of all streets), which means more equal accessibility to the main streets of the city, where the tourist could take more possible paths and access more businesses in the city.

5. Discussion

In this paper, we focus on studying the accessibility of all people that visit a city with some mobility restrictions and its effects in other aspects, such as the economy (accessibility to all businesses of a city). For that, the proposed algorithm measures the difficulty of each street attending to the existence of stairs, slopes, railings, and the state of sidewalks and the pavement. This difficulty has different parameters to allow change for the restoration

possibilities of the streets in a city, and we choose the values of these parameters following the specifications of the Technical Accessibility Code of Spain Buildings.

Concerning parameters that represent the restoration levels (lambda in Equation (1)), these values help to control the state of restoration of different elements of a city. If we select a low value, it means that we relax the difficulty with respect to the thresholds. In our case study, in theory, the best value is the unit because it is difficult to restore a historical area of the city of Avila. To test this hypothesis, we probe a smaller value of lambda (0.8) and we observe few changes (only seven streets will change from the red level color to orange), and comparing with manual labeling, these streets could be wrong. For that, the best value in our case study is 1.

After constructing the urban and commercial networks, we applied the APA algorithm to the commercial graph, with the aim of ranking the nodes. As a result, Table 5 presents the top 10 APA ranking of commercial areas.

Focusing on the impact of accessibility, Table 6 shows different configurations in terms of different mobility constraints and accessibility levels. It can be observed that changes in the characteristics of the network result in less density than the original graph. Then, following the results obtained in that Table, we evaluate the inaccessible businesses where the mobility constraints are low. Table 7 show that a quarter (25.42%) of the businesses of the city are located on inaccessible streets for people with high mobility problems. Moreover, 8.75% of them are only accessible for healthy people.

Our following step is to study how to increase the accessibility of some streets through different actions such as improving the pavement, adding a railing, and improving or adding the sidewalks. It can be seen in Table 9 that after the restoration of some streets, the network is denser than the original, and the number of commercial inaccessible businesses is reduced from 25.42% to 21.32% with a slight restoration and to 8.76% with a moderate restoration in the best case. With a strong restoration of the city, almost all businesses will be accessible.

Finally, we go into more detail by analyzing how the improving of the accessibility of three streets could affect the entire city. To demonstrate this, we recalculate the connectivity of the city with APA centrality measure of the main streets of the city in terms of accessibility before and after the restoration (see Table 10). An analysis of the aforementioned table shows that adapting several streets of the city, it is possible to have a more connected city in terms of accessibility.

6. Conclusions

The methodology proposed in this work highlights the importance of restoring some streets in a city, which produces an improvement in the accessibility of a city, and in addition, commerce and tourism benefit. Although the concept is clear, the main problem is the limitations in changes on areas of a historic city, because of heritage laws which make a strong restoration impossible. As a result of that, we focus on demonstrating how a simple improvement of two or three streets (e.g., installing railings) could positively affect the accessibility of a city. Another detected problem is the length of a path: Our method can obtain a personalized path for each person with accessibility problems, but this path could be different to the shortest route. That is, this methodology gives priority to a less difficult path even if it is longer than the shortest path.

We propose a methodology for labeling the estimated difficulty of all streets of a city in terms of different parameters of accessibility, such as stairs, the slopes of a street, the existence of railings, the state of a sidewalk, or irregular grounds.

In our experiments, focused on a case study of the city of Avila (Spain). By labeling the difficulty, we generate the network of accessibility for different mobility constraints. We study the most important areas in terms of the number of businesses using the APA centrality index, and compare it with the network of accessibility to show the inaccessible zones of a city, wherein the worse case (43.01%) shows the impossibility of people with mobility problems accessing 25.42% of commerce in the city.

Finally, we show a simulated situation where three streets are restored. The results show how accessibility (12.90% in the city center) and commercial impact (4.71%) are improved.

Author Contributions: All authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Diputación de Ávila (Convocatoria de Ayudas a la Investigación sobre Temas Abulenses 2019).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Russo, A.P.; Scarnato, A. "Barcelona in common": A new urban regime for the 21st-century tourist city? *J. Urban Aff.* **2018**, *40*, 455–474. [CrossRef]
2. Urry, J. *Mobilities: New Perspectives on Transport and Society*; Routledge: London, UK, 2016
3. Clavé, S.A. Urban tourism and walkability. In *The Future of Tourism*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 195–211.
4. Russo, A.P. The "vicious circle" of tourism development in heritage cities. *Ann. Tour. Res.* **2002**, *29*, 165–182. [CrossRef]
5. Gaglione, F.; Gargiulo, C.; Zucaro, F. Elders' quality of life. A method to optimize pedestrian accessibility to urban services. *TeMA J. Land Use Mobil. Environ.* **2019**, *12*, 295–312.
6. Allen, W.B.; Liu, D.; Singer, S. Accesibility measures of US metropolitan areas. *Transp. Res. Part B Methodol.* **1993**, *27*, 439–449. [CrossRef]
7. Maroto, A.; Zofío, J.L. Accessibility gains and road transport infrastructure in Spain: A productivity approach based on the Malmquist index. *J. Transp. Geogr.* **2016**, *52*, 143–152. [CrossRef]
8. Domínguez, T.; Fraiz, J.A.; Alén, E. Economic profitability of accessible tourism for the tourism sector in Spain. *Tour. Econ.* **2013**, *19*, 1385–1399. [CrossRef]
9. Rodríguez-Echeverría, J.; Semanjski, I.; Van Gheluwe, C.; Ochoa, D.; IJben, H.; Gautama, S. Density-Based Spatial Clustering and Ordering Points Approach for Characterizations of Tourist Behaviour. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 686. [CrossRef]
10. Zaragozí, B.; Trilles, S.; Gutiérrez, A. Passive Mobile Data for Studying Seasonal Tourism Mobilities: An Application in a Mediterranean Coastal Destination. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 98. [CrossRef]
11. Domènech, A.; Mohino, I.; Moya-Gómez, B. Using Flickr Geotagged Photos to Estimate Visitor Trajectories in World Heritage Cities. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 646. [CrossRef]
12. Crucitti, P.; Latora, V.; Porta, S. Centrality measures in spatial networks of urban streets. *Phys. Rev. E* **2006**, *73*, 036125. [CrossRef]
13. Hennemann, S.; Derudder, B. An alternative approach to the calculation and analysis of connectivity in the world city network. *Environ. Plan. B Plan. Des.* **2014**, *41*, 392–412. [CrossRef]
14. Curado, M.; Tortosa, L.; Vicent, J.F.; Yeghikyan, G. Understanding mobility in Rome by means of a multiplex network with data. *J. Comput. Sci.* **2021**, *51*, 101305. [CrossRef]
15. Kim, Y.A.; Song, H.S. Strategies for predicting local trust based on trust propagation in social networks. *Knowl.-Based Syst.* **2011**, *24*, 1360–1371. [CrossRef]
16. Wang, S.; Du, Y.; Deng, Y. A new measure of identifying influential nodes: Efficiency centrality. *Commun. Nonlinear Sci. Numer. Simul.* **2017**, *47*, 151–163. [CrossRef]
17. Zhang, X.; Zhu, J.; Wang, Q.; Zhao, H. Identifying influential nodes in complex networks with community structure. *Knowl.-Based Syst.* **2013**, *42*, 74–84. [CrossRef]
18. Albert, R.; Jeong, H.; Barabási, A.-L. Diameter of the world-wide web. *Nature* **1999**, *401*, 130–131. [CrossRef]
19. Newman, M.E. The structure and function of complex networks. *SIAM Rev.* **2003**, *45*, 167–256. [CrossRef]
20. Estrada, E.; Bodin, Ö. Using network centrality measures to manage landscape connectivity. *Ecol. Appl.* **2008**, *18*, 1810–1825. [CrossRef]
21. Freeman, L.C. Centrality in social networks conceptual clarification. *Soc. Netw.* **1978**, *1*, 215–239. [CrossRef]
22. Agryzkov, T.; Oliver, J.L.; Tortosa, L.; Vicent, J.F. An algorithm for ranking the nodes of an urban network based on the concept of PageRank vector. *Appl. Math. Comput.* **2012**, *219*, 2186–2193. [CrossRef]
23. Agryzkov, T.; Tortosa, L.; Vicent, J.F. New highlights and a new centrality measure based on the Adapted PageRank Algorithm for urban networks. *Appl. Math. Comput.* **2016**, *291*, 14–29. [CrossRef]
24. Agryzkov, T.; Curado, M.; Tortosa, L.; Vicent, J.F. Extending the Adapted PageRank Algorithm Centrality to Multiplex Networks with Data Using the PageRank Two-Layer Approach. *Symmetry* **2019**, *11*, 284. [CrossRef]
25. Gaglione, F.; Cottrill, C.; Gargiulo, C. Urban services, pedestrian networks and behaviors to measure elderly accessibility. *Transp. Res. Part D Transp. Environ.* **2021**, *90*, 102687. [CrossRef]
26. Handy, S.L.; Niemeier, D.A. Measuring accessibility: An exploration of issues and alternatives. *Environ. Plan. A* **1997**, *29*, 1175–1194. [CrossRef]
27. Ben-Akiva, M. Disaggregate travel and mobility choice models and measures of accessibility. In *Behavioural Travel Modelling*; Routledge: London, UK, 1979.

28. Kwan, M.P. Space-time and integral measures of individual accessibility: A comparative analysis using a point-based framework. *Geogr. Anal.* **1998**, *30*, 191–216. [CrossRef]
29. Páez, A.; Scott, D.M.; Morency, C. Measuring accessibility: Positive and normative implementations of various accessibility indicators. *J. Transp. Geogr.* **2012**, *25*, 141–153. [CrossRef]
30. Sarlas, G.; Páez, A.; Axhausen, K.W. Betweenness-accessibility: Estimating impacts of accessibility on networks. *J. Transp. Geogr.* **2020**, *84*, 102680. [CrossRef]
31. Asero, V.; Gozzo, S.; Tomaselli, V. Building tourism networks through tourist mobility. *J. Travel Res.* **2016**, *55*, 751–763. [CrossRef]
32. Ayuntamiento, D.Á. *Guía de Recursos Accesibles de la Ciudad de Ávila*; Ayuntamiento de Ávila: Ávila, Spain, 2008
33. Oficina Municipal de Accesibilidad (Avila, Spain). 2007. Available online: http://www.avila.es/articles/accesibilidad (accessed on 29 May 2021).
34. Arcos, F.J.M. Turismo accesible, turismo para todos en una ciudad patrimonio de la humanidad: El caso de Ávila. *Rev. Tur. Desenvolv.* **2009**, *11*, 73–79.
35. OpenStreetMap. Available online: https://www.openstreetmap.org (accessed on 29 May 2021).
36. Overpass Turbo. Available online: https://overpass-turbo.eu (accessed on 29 May 2021).
37. Microsoft Bing. Available online: https://www.bing.com/maps (accessed on 29 May 2021).
38. de la Edificación, L.d.O.; Parte, I. *Código Técnico de la Edificación*; Minvivienda: Madrid, Spain, 2006.

4.2. ARTÍCULO 2

Este artículo ha sido publicado en la revista *Applied Mathematics and Computation*, en la editorial *ElSevier*.

A continuación, se muestra algunos indicios de calidad de la revista, y otros indicios adicionales de otras plataformas.

Tabla 3. Otros indicios de calidad del artículo 2

Fecha de publicación	Citescore	5- Year Impact Factor	Lecturas en ResearchGate
Julio 2021	7.6	3.767	46

Applied Mathematics and Computation 412 (2022) 126560

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

A new centrality measure in dense networks based on two-way random walk betweenness

Manuel Curado[a,*], Rocio Rodriguez[b], Leandro Tortosa[c], Jose F. Vicent[c]

[a] Polytechnic School, Catholic University of Murcia, Campus Los Jeronimos, s/n, E-30107 Murcia, Spain
[b] Department of Mechanical Engineering, Catholic University of Ávila, Calle Canteros, s/n, E-05005, Ávila, Spain
[c] Department of Computer Science and Artificial Intelligence, University of Alicante, Campus de San Vicente, Ap. Correos 99, E-03080, Alicante, Spain

ARTICLE INFO

Article history:
Received 4 May 2021
Revised 25 July 2021
Accepted 28 July 2021

Keywords:
Centrality measure
Betweenness centrality
Random walks
Densification

ABSTRACT

Many scholars have tried to address the identification of critical nodes in complex networks from different perspectives. For instance, by means of the betweenness methods based on shortest paths and random walk, it is possible to measure the global importance of a node as an intermediate node. All these metrics have the common characteristic of not taking into account the density of the clusters. In this paper, we apply an analysis of network centrality, from a perspective oriented to ranking nodes, reinforcing dense communities using evaluating graphs using a two-trip transition probability matrix. We define a new centrality measure based on random walk betweenness. We study and analyse the new metric as a betweenness centrality measure with common characteristics with Pagerank, presenting through its practical implementation in some examples based on synthetic, and testing with well-known real-world networks. This method helps to increase the ranking of nodes belonging to dense clusters with a higher average degree than the remaining clusters, and it can detect the weakness of a network comparing it with the classical betweenness centrality measure.

1. Introduction

In the last decades, complex network theory has gained much attention [1–6], especially with the advances in the understanding of the interconnected systems as biological, social, transport and communication networks. Similarly, in complex systems the nodes interact with each other following different types of patterns that can encompass multiple types of relationships [7].

The transfer of ideas or diseases, the flows of information and the influence between nodes are a key function of complex networks [8–11]. Therefore, assessing the influence of nodes in a network is a very important issue, as evidence by the fact, that the information can spread more quickly in a network the grater is the influence of the node [12–15].

The modern digital technologies and the accessibility of data sets, obtained through different sources, have increased the interest of study structural properties of networks that represent real systems and that are characterized by non-trivial properties. The different methods, concepts, algorithms and models proposed to analyze the structure of real networks have formed the framework of the so-called theory of complex network [16,17]. It is usual, that the characteristics of a complex

* Corresponding author.
E-mail addresses: mcurado@ucam.edu (M. Curado), rocio.rodriguez@ucavila.es (R. Rodriguez), tortosa@ua.es (L. Tortosa), jvicent@ua.es (J.F. Vicent).

https://doi.org/10.1016/j.amc.2021.126560
0096-3003/

network disclose very important information about its function and dynamics. For instance, relating the position occupied for a node in a graph and the role of this node in dynamic processes. So, not all nodes have the same importance in spreading rumors or diseases in a network because the number and the organization of the edges have an impact on the process.

As it is shown in [18,19], the dynamic and stability of complex networks are determined by the topology of the network. This can be clearly seen by removing the important nodes from a network, that is, the nodes with the highest centrality. The breakdown of the network structure is faster in this case than if other nodes are removed [20,21]. To characterize nodes in a network, centrality measures are the most used tools. To identify the vertices that have a strategic role in the network, these metrics are fundamental and it has been an active research field [22,23].

A measure with specific applications in transportation, communications or infrastructural networks is betweenness centrality [1,24–26]. The use in these types of applications is because this metric measures the importance of a node for the flow of information between pairs of them. In the beginning, the betweenness centralities metric were applied to static and single-layer networks. Nowadays, however, they are used in all types of networks, static or dynamic, single layer or multilayer [4,6,27] because of their important effects on the network functioning.

Several models to analyze the participation of a node in its communications have been proposed. Fremman presented the shortest-path betweenness centrality considering that the communications flow only along the shortest path [1,28]. In general terms, a weakness of measures based on shortest paths is that they are not appropriate where link content distribution is governed by rules different to shortest paths. On the other hand, the centrality in the current flow betweenness is measured as the total sum of electrical current that flows through it [26,29]. Another simpler definition of betweenness centrality is how often a node is traversed by random walks between two other nodes because it counts all possible paths crossing it.

All these betweenness variants have in common that they do not take into account the cluster density. To solve this weakness, a novel betweenness centrality based on random walks from a dense-networks perspective is proposed. A quantification of the importance of a node, in the information transmission, through the relationships between four different nodes is presented in this new measure.

The goal of this paper is the application of network centrality analysis, based on betweenness centrality, to complex networks from a perspective oriented to ranking nodes by means the reinforcing dense communities. More specifically, we propose a centrality measure that combines the idea of the random-walk betweenness centrality and the link predictor algorithm Return Random Walk [30].

To achieve this objective, the paper is organized as follows: Section 2 describes two classical centrality measures, PageRank and betweenness, and it serves as a starting point to establish our metric. Then, in Section 3, the new centrality measure based on random walk betweenness centrality are described in detail. The applications of the proposed algorithm are illustrated and discussed in the Numerical Results section. Finally, the conclusions are summarized in Section 5.

2. Preliminary

2.1. Pagerank centrality

PageRank algorithm ranks the importance of a web page using the hyperlink structure of web system [31]. The PageRank value (PR) of a web page p_i ($i = 1, \cdots, n$) is calculated as follows

$$PR(p_i) = \sum_{p_j \in D(p_i)} \frac{PR(p_j)}{N(p_j)},$$

where $PR(p_j)$ is the set of pages pointing to p_i and $N(p_j)$ is the number of out-links from page p_j.

Since the $PR(p_j)$ is unknown, an iterative process is introduced

$$PR_{k+1}(p_i) = \sum_{p_j \in D(p_i)} \frac{PR_k(p_j)}{N(p_j)}.$$

Using a compact formula we have

$$R^{k+1} = TR^k,$$

where R^k is the PageRank vector of all pages after the iteration k and T is a normalized link matrix defined as

$$T_{ij} = \begin{cases} \frac{1}{N(p_j)} & \text{if } p_j \in D(p_i) \\ 0 & \text{otherwise.} \end{cases}$$

Matrix T is a sparse and stochastic transition probability matrix, but it can exist zero rows which mean that those nodes have no out-links (dangling nodes). To solve this and ensure the convergence, a new matrix G is proposed

$$R^{k+1} = GR^k,$$

where

$$G = \alpha H + (1-\alpha)ee^T/n,$$

with $\alpha \in [0,1]$, e a column vector of all ones and

$$H = T + ae^T/n.$$

The vector a indicates the dangling node vector, $a_i = 1$ if i is a dangling node and 0 otherwise.

To ensure that the iteration process can still converge to a unique positive vector when the initial value is arbitrarily chosen, the Google $n \times n$ matrix G is verified to satisfy the properties of stochastic, irreducible and aperiodic.

2.2. Betweenness centrality

In this section, the classic definition of betweenness centrality is presented. Let $G = (V, E, W)$ an undirected graph, where V is the set of nodes, E the set of edges and W the weight of each node (in unweighted graphs, we assume that the weight of each edge is 1).

If σ_{ij} is the number of shortest paths from a node i to node j and σ_{itj} is the number of shortest paths from node i to node j, through a transition node t. The formal definition of betweenness centrality measure of a node t is [1]:

$$C_t = \sum_{i \neq t} \sum_{j \neq i,t} \frac{\sigma_{itj}}{\sigma_{ij}}. \tag{1}$$

The complexity of this algorithm is $O(n^3)$ and takes space $O(n^2)$ on any graph G with n nodes.

It exists an important number of betweenness centrality measures defined in the literature. For instance, the shortest paths betweenness centrality measures the flow of a network where edges or links represent canals that transmit a unit amount of flow [32]. The value of the flow is calculated as the sum of flow generated at i, and the amount of flow through a node t is the sum of flow leaving this node. In other words, the betweenness centrality (or flow) of a node i is defined as the average of the maximum flow through t over all possible pairs (i, j). The complexity of this algorithm is $O(e^2 n)$ (see [26]) and it only works in directed networks.

In [26] the authors proposed a modification of the original measure based on random walks. This adjustment supposes that information transmission between any pairs of nodes in a network follows a random path and, does not need to calculate the shortest path. The principle underpinning the random walk betweenness of a node t is the expected flow of a random walk from an origin node i to a different destination node j through an intermediate node t, having in count all possible pairs (i, j). Moreover, this method shows that the random-walk betweenness is identical to current-flow betweenness. Several measures work following this idea [24,25]. To optimize random-walk betweenness for computing all nodes, [29] propose an efficient algorithm that takes time $O(I(n-1) + mn \log n)$, where $I(n) = O(n^3)$ is the time for computing the inverse of a matrix with n nodes.

3. Two-way random walk betweenness centrality (2RW)

All centrality measures based on betweenness centrality do not take into account an important network characteristic, such as the density of the cluster where each node of the network is located. Because of this, a novel betweenness centrality based on random walks from a dense, fully connected and self-loop free networks perspective is proposed.

We quantify the importance of a node, in the information transmission, through the relationships between four different nodes –from a origin node i to a destination node j through a node t and returning to i crossing by a different node k–. The basis of our proposal is the link predictor algorithm, Return Random Walk [30], which reinforces the dense communities by means of evaluating graphs using a two-trip transition probability matrix.

Remark that we work with networks without isolated nodes and self-loops.

3.1. Formal analysis

Given an undirected weighted graph $G(V, E, W)$, where V is the set of n nodes, E is the set of e edges (e_{ij} is the edge linking nodes i and j) and W is the weighted matrix with $w_{ij} = w_{ji} > 0$ if $e_{ij} \in E$ and 0 otherwise. This means that not allowing self-loops, $w_{ii} = 0$ for any i.

The process of calculating the centrality measure follows these steps:

Step 1: with the matrix $W = \{w_{ij}\}_{i,j=1,2,\ldots,n} \in \mathbb{R}^{n\times n}$, and all-ones vector $\boldsymbol{u} \in \mathbb{R}^{n\times 1}$, we calculate the matrix $D = W\boldsymbol{u} = \{d_i\} \in \mathbb{R}^{n\times 1}$, containing the sum of all weights of the edges of each node.

Step 2: we calculate the matrix P by exploring the 2-step random walk between 3 different nodes $i, j, t \in V$.

$$P_{itj} = \frac{w_{it} w_{tj}}{d_i d_j} \tag{2}$$

where $d_i = \sum_{j=1}^{n} w_{ij}$. Then, P_{itj} represents the probability to reach a node j from a node i through any transition node t.

Step 3: for a pair of nodes $i, j \in V$, we calculate a vector $V_{P_{itj}}$ with all possible 2-step paths between nodes i and j through any different transition node t:

$$V_{P_{itj}} = \begin{pmatrix} P_{i1j} & P_{i2j} & P_{i3j} & \cdots & P_{inj} \end{pmatrix} \tag{3}$$

Step 4: we construct the real matrix $V_P \subseteq M_{e\times n}$ representing the probability of all possible 2-step paths in the network. Remark that $e = \frac{n(n-1)}{2} + n$ is the maximum number of edges of a graph with n nodes:

$$V_P = \begin{pmatrix} V_{P_{1t1}} \\ V_{P_{1t2}} \\ V_{P_{1t3}} \\ \vdots \\ V_{P_{1tn}} \\ V_{P_{2t2}} \\ V_{P_{2t3}} \\ \vdots \\ V_{P_{2tn}} \\ V_{P_{3t3}} \\ V_{P_{3t4}} \\ \vdots \\ V_{P_{n_{-1}tn}} \\ V_{P_{ntn}} \end{pmatrix}_{e\times n} = \begin{pmatrix} P_{111} & P_{121} & P_{131} & \cdots & P_{1n_{-1}1} & P_{1n1} \\ P_{112} & P_{122} & P_{132} & \cdots & P_{1n_{-1}2} & P_{1n2} \\ P_{113} & P_{123} & P_{133} & \cdots & P_{1n_{-1}3} & P_{1n3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ P_{11n} & P_{12n} & P_{13n} & \cdots & P_{1n_{-1}n} & P_{1nn} \\ P_{212} & P_{222} & P_{232} & \cdots & P_{2n_{-1}2} & P_{2n2} \\ P_{213} & P_{223} & P_{233} & \cdots & P_{2n_{-1}3} & P_{2n3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ P_{21n} & P_{22n} & P_{23n} & \cdots & P_{2n_{-1}n} & P_{2nn} \\ P_{313} & P_{323} & P_{333} & \cdots & P_{3n_{-1}3} & P_{3n3} \\ P_{314} & P_{324} & P_{334} & \cdots & P_{3n_{-1}4} & P_{3n4} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ P_{n_{-1}1n} & P_{n_{-1}2n} & P_{n_{-1}3n} & \cdots & P_{n_{-1}n_{-1}n} & P_{n_{-1}nn} \\ P_{n1n} & P_{n2n} & P_{n3n} & \cdots & P_{nn_{-1}n} & P_{nnn} \end{pmatrix}_{e\times n} \tag{4}$$

Step 5: for each pair of nodes i, j in the network, we generate a transition nodes matrix T_{ij}.

$$T_{ij} = V_{P_{itj}}^T V_{P_{jti}} = \begin{pmatrix} P_{i1j} \\ P_{i2j} \\ P_{i3j} \\ \cdots \\ P_{inj} \end{pmatrix} \begin{pmatrix} P_{j1i} & P_{j2i} & P_{j3i} & \cdots & P_{jni} \end{pmatrix} =$$

$$= \begin{pmatrix} P_{i1j}P_{j1i} & P_{i1j}P_{j2i} & P_{i1j}P_{j3i} & \cdots & P_{i1j}P_{jni} \\ P_{i2j}P_{j1i} & P_{i2j}P_{j2i} & P_{i2j}P_{j3i} & \cdots & P_{i2j}P_{jni} \\ P_{i3j}P_{j1i} & P_{i3j}P_{j2i} & P_{i3j}P_{j3i} & \cdots & P_{i3j}P_{jni} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ P_{inj}P_{j1i} & P_{inj}P_{j2i} & P_{inj}P_{j3i} & \cdots & P_{inj}P_{jni} \end{pmatrix} \tag{5}$$

These matrices represent the probabilities to go from an origin node i to a destination node j through a transition node t and return to the origin crossing another transition node k different to t. Thus, we assign to zero the diagonal values of the resulting matrix T_{ij} because it means the use of the same transition node ($t = k$). In other words, this matrix shows the random walks reaching a node j from a node i through any transition node t, and returning to i through a node k different of t (note that $i \neq j \neq t \neq k$, in other case the probability will be 0). In this way, we represent relationships between 4 different nodes in the network and, because of this, if exists any community or cluster smaller than 4 nodes, their centrality values will be 0.

Step 6: for all the transition matrices T_{ij} we select the maximum element:

$$\max(T_{ij}) = P_{itj}P_{jki}. \tag{6}$$

These maximum values represent the maximum probability to go from node i to node j and return, through two transition nodes t (go) and k (return).

Step 7: for all the maximum vales obtained in Step 6 the corresponding counter is increased:

$$\forall T_{ij} : \; C_t = C_t + 1, \;\; C_k = C_k + 1. \tag{7}$$

Currently, a novel metric for each node, measuring the centrality of all nodes, is obtained:

$$C = \{C_i\}_{i=1,\cdots,n}.$$

We summarize the described method in the following algorithm:
The code of the algorithm is available in a repository[1].

[1] https://github.com/manucurado/2RW

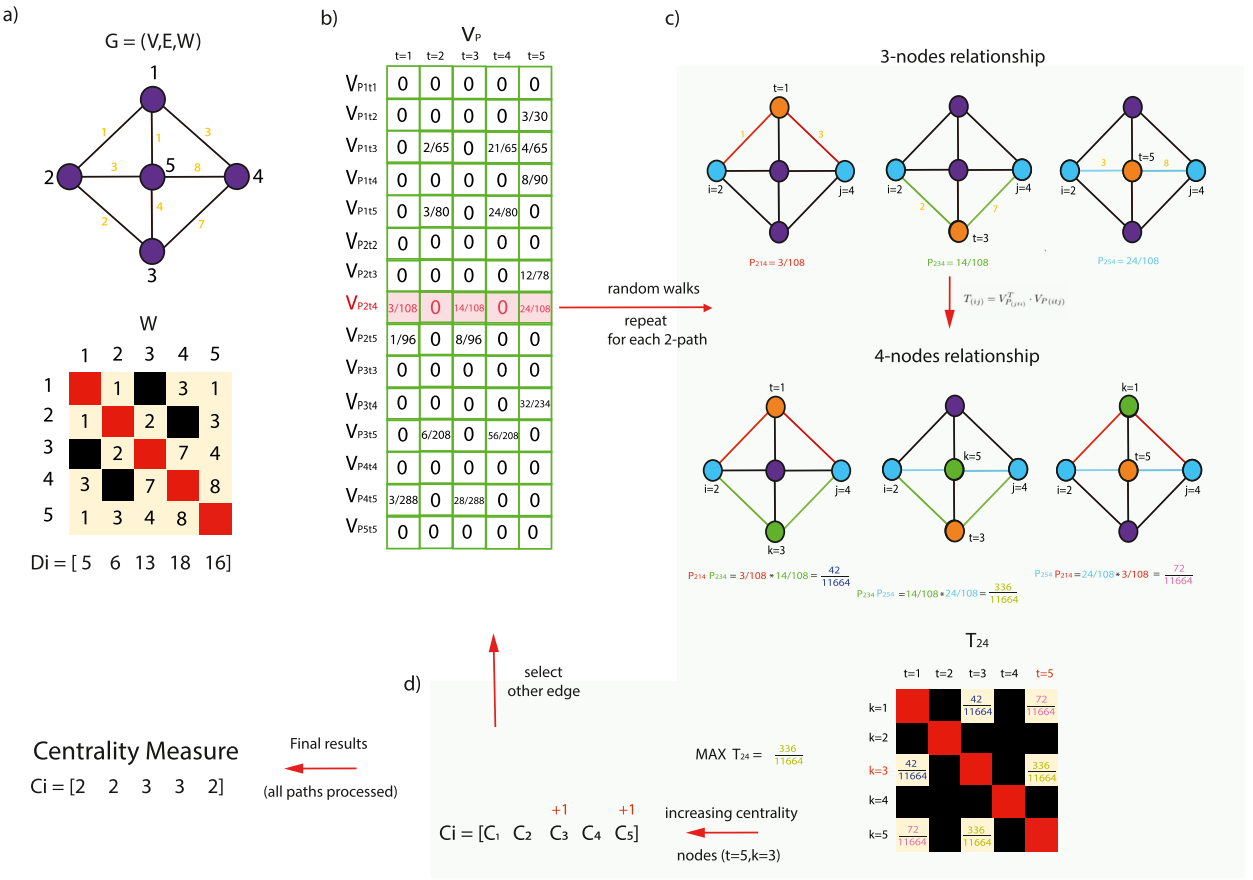


Fig. 1. Toy sample. This undirected network consists of 5 nodes and 8 edges. We calculated the 2RW centrality of each node iterating all possible 2-step paths connecting a pair of nodes. In b-d. we show an iteration of the 2RW that connecting the nodes 2 and 4.

3.2. Toy example

To illustrate the algorithm and its formulation, we show a toy sample with a weighted graph $G = (V, E, W)$ with 5 nodes and 8 edges (see Fig. 1a).

$$T_{24} = V_{P_{2t4}}^{T} V_{P_{4t2}} = \begin{pmatrix} P_{214} \\ P_{224} \\ P_{234} \\ \cdots \\ P_{2n4} \end{pmatrix} \begin{pmatrix} P_{412} & P_{422} & P_{432} & \cdots & P_{4n2} \end{pmatrix} \tag{8}$$

In Fig. 1b, we can see its associated V_P matrix with 15 rows (all possible edges in an undirected graph of 5 nodes) and 5 columns. In Fig. 1c, we show all possible 2-step paths connecting nodes 2 and 4 highlighted in red, green and blue colour. The matrix T_{24} (Eq. 8) represents the probability of all possible random walks to go from node 2 to node 4 (in blue), passing for two different transition nodes (in orange and green). The maximum element of the matrix T_{24} is $P_{234}P_{254} = \frac{336}{11664}$. We increment the centrality of nodes 3 and 5, which play the role of transition nodes, green and orange, in the maximum probability path to go from node 2 to node 4 (see Fig. 1d).

Finally, a vector with the values of the new centrality measure for each node ($C = \{2, 2, 3, 3, 2\}$) is obtained. The comparison of this metric with the classical betweenness centrality, in which node 5 is the most important, allow us to see the differences.

In the next section, other examples, using syntactic and real networks, are analyzed.

4. Numerical results

In this section, we analyze and discuss the proposed centrality measure, to compare and verify the differences between this metric and other centrality measures. From numerical analysis, we consider two synthetic networks and two real networks.

Table 1
Communities Information. The global density of the network is 13%.

Cluster	Nodes	Local Density	Avg. Degree	Max. Degree
Red	23	48.62%	113.43	226
Green	30	40.00%	120.33	259
Blue	39	31.44%	125.46	254

Table 2
The 10-top ranking of classic centrality measures in the contact network: PR (columns 1–2), DG (columns 3–4), CBT (columns 5–6) and our method (columns 7–8). In color, we show the community of each node (see Fig. 2).

PR	#Node	DG	#Node	CBT	#Node	2RW	#Node
1	55	1	55	1	55	1	54
2	54	2	54	2	8	2	25
3	25	3	25	3	49	3	55
4	23	3	23	4	50	4	59
5	56	3	56	5	1	5	56
6	59	3	59	6	54	6	23
7	29	7	29	7	51	7	30
8	4	8	4	8	29	8	27
9	1	8	1	9	33	9	0
10	58	8	58	10	73	10	58

4.1. Synthetic networks

4.1.1. Virus spreading

In Fig. 2, we show a random graph G with 92 nodes and 619 edges. The weights of the nodes have been randomly generated with values between 1 and 20. This network has been constructed using an interactive tool named *GraphVis* [33]. Looking at the graph, it can be assumed that this example simulates a network in which the nodes represent the set of people distributed in 3 different communities or clusters (coloured in red, green and blue) and the edges represent the a link between two people. Similarly, the weight of each edge (values between 1 and 20) can represent the number of physical interactions between 2 nodes.

Some basic community properties, local density, average degree and maximum degree are presented in Table 1. Analyzing the density of the different clusters we can clearly see that red is the densest.

Remark that to quantify the local density, we adopt a classic notion used in the densest-subgraph problem, where density is defined as the ratio between the edges and the vertices of a subgraph.

We applied some classical centrality measures, PageRank (PR), classical Betweenness (CBT) and Degree (DG), intending to establish an initial comparison with the new measure to determine the main characteristics and properties of this measure. By viewing the information about the importance of some nodes shown in Table 2, we can see as the most important nodes, in the case of PR and DG, belong to blue and green communities. This is due to the combination of lower local density and a higher average degree. Regarding the CBT, there are nodes of the red cluster with a best ranking (nodes 8 and 1) in comparison with the previous measures, because of this metric measures the extent to which a node lies on paths between other nodes. Nodes with high CBT rank may have considerable influence within a network under their control over information passing between others. Likewise, they are also the ones whose removal from the network will most disrupt communications between other vertices because they lie on the largest number of paths taken by messages. The nodes' ranking obtained by 2RW centrality is similar to PR and DG, respectively. This similarity leads us to believe that both 2RW and PR centralities present common characteristics, although the idea behind the proposed centrality is closer to the CBT measure.

Fig. 3 shows a comparison of centrality values for CBT and 2RW. Remark that all the values are normalized and the different background colors are associated to different network clusters. Observing this figure we can appreciate by simple inspection that clusters blue and green have nodes with higher centrality than red cluster, despite its smaller local density. The proposed measure increases the centrality value of the nodes belonging to dense clusters with a higher average degree than the remaining clusters ($+0.679$ in average for nodes in red cluster, $+0.1185$ in green cluster and $+0.1645$ in blue cluster).

Moreover, note the differences of both measures in some particular nodes. For instance, the node 8 has a high CBT centrality in a hyper-connected cluster (local density or intra-links ratio, is 48.52%) with only 4 weak inter-links but it has a ranking position of 25 in our method. In contrast, several nodes in green and blue clusters ascending to the 20-top ranking of our measure. Summarizing, in CBT ranking, the proportions are 20% of nodes in red cluster, 50% of nodes in green cluster, and 30% in blue ones. However, taking into account our ranking, we obtain the following proportions: 10% of nodes in red

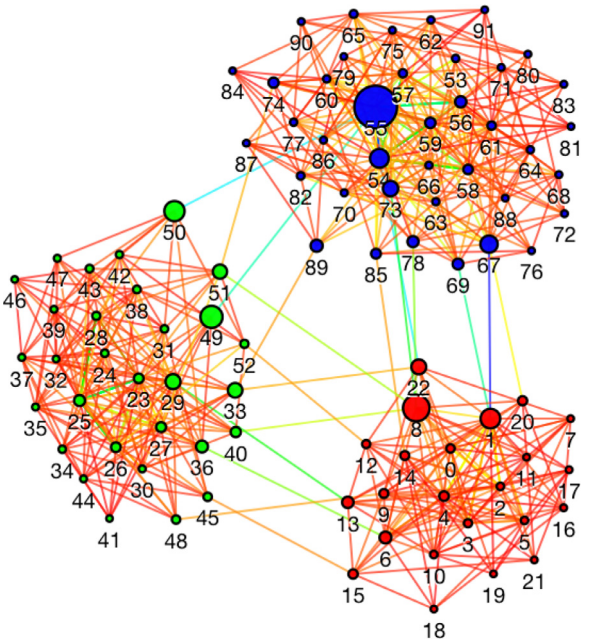


Fig. 2. Synthetic experiment: a contact network with 92 nodes and 619 edges. This network is clustered in 3 communities (coloured in red, green and blue). The edges represent the number of contacts between a pair of nodes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

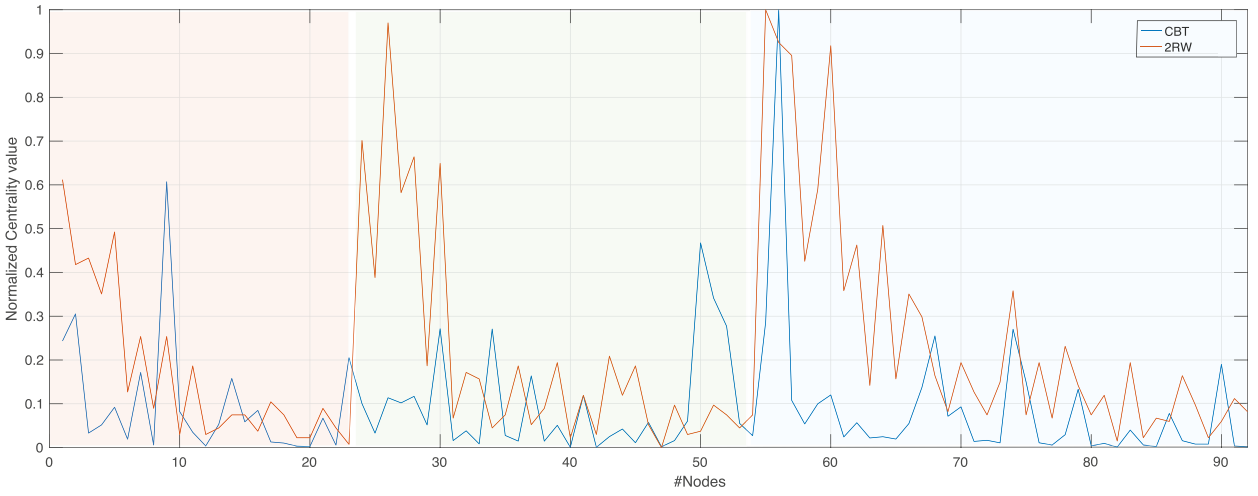


Fig. 3. Numerical comparison between normalized values of CBT and 2RW where the background colors are associated to the three network clusters. The proposed measure (red) increases, in general, the centrality value with respect to CBT measure (blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

cluster, 40% of nodes in green cluster, and 50% of nodes in blue cluster. Summarizing, the 2RW centrality is, in general, higher with density intra-clusters and high average degree.

To carry out the comparison between CBT and 2RW is interesting to study the correlation between them. For this purpose, the Pearson correlation index is useful since it indicates a predictive relationship that can be exploited in practice.

To perform a study about correlation between CBT and 2RW, we use a network with the three clusters (red, green and blue) but with different global and local densities. More exactly, we take networks with global densities of 10%, 13% and 18%, changing the local densities in each of the clusters. There, we proceed to compute the correlation including our measure 2RW and both CBT and PR.

These data are summarized in Table 3. In these table, the first column shows the values of the global densities, while the second one shows the local densities values. The correlation between CBT and 2RW is in column 3 and finally, column 4 shows us the correlation between PR and 2RW.

Table 3
Top- Correlations between Pagerank (PR), Classic Betweenness Ranking (CBT) and our measure (2RW) for different global and local densities network configurations. Down- Correlations between PR and CBT, and the average Degree.

Global Density	Local densities	Corr. CBT-2RW	Corr. PR-2RW
10% (random)	34.39%-31.95%-23.62%	51.10%	92.29%
10% (-3% red)	0.67%-31.95%-23.62%	79.89%	81.09%
10% (-3% green)	34.39%-11.03%-23.62%	31.43%	68.47%
10% (-3% blue)	34.39%-40%-14.84%	58.16%	87.59%
13% (original)	48.62%-40%-31.44%	40.61%	85.85%
18% (random)	50.59%-43.22%-35.22%	89.78%	93.48%
18% (+5% red)	58.10%-40%-31.44%	84.94%	92.82%
18% (+5% green)	48.62%-50.11%-31.44%	85.80%	94.06%
18% (+5% blue)	48.62%-40%-39.27%	84.33%	94.03%

Table 4
Correlation between CBT-2RW, PR-2RW and PR-CBT. First row is the original network and it is compared with two denser versions of this network.

Global Density	Corr CBT-2RW	Corr. PR-2RW	Corr. PR-CBT	Avg. Degree
0.39% (Fig. 4)	32.62%	83.76%	59.27%	328.55
2%	70.74%	88.02%	92.52%	500.12
5%	100%	99.11%	99.11%	1255.2

We can observe that the highest values of correlation are given for dense networks, specially when PR and 2RW are compared. However, for low dense networks, correlation values are not as high as in the case of dense networks, specially when comparing with CBT measure.

Global Density	Local densities	Corr. PR-CBT	Avg. Degree
10% (random)	34.39%-31.95%-23.62%	66.35%	96.26
10% (-3% red)	0.67%-31.95%-23.62%	71%	91.84
10% (-3% green)	34.39%-11.03%-23.62%	73.08%	91.10
10% (-3% blue)	34.39%-40%-14.84%	66.90%	369.30
13% (original)	48.62%-40%-31.44%	57.95%	120.78
18% (random)	50.59%-43.22%-35.22%	94.19%	360.69
18% (+5% red)	58.10%-40%-31.44%	86.24%	377.71
18% (+5% green)	48.62%-50.11%-31.44%	88.8%	374.06
18% (+5% blue)	48.62%-40%-39.27%	86.58%	369.30

4.1.2. Social network

Another synthetic network with 1000 nodes and 6612 edges is studied in this subsection (see Fig. 4. This network can simulate different profiles of a social network where the weight of each edge represents the number of interactions between 2 nodes, in a range $[1, 100]$. As can be seen in Fig. 4, the network has 20 communities (nodes colour), with a high local density, where the size of a node is proportional to the classic betweenness index.

As in previous subsection, to carry out the comparison is interesting to study the correlation between the different measures. Table 4 shows the results of the analysis of centrality measures taking into account that to study the importance of a node we focus on Pagerank, but to analyze the interactions between nodes betweenness measure are used.

The study is performed comparing the original network (global density of 0.39%) with two denser versions of the same network with 2% and 5% of global density respectively. As we can see, the average degree ofproposal and these three networks increases with the density and, as density is increased, the correlation between our 2RW proposal and the CBT is near to a perfect correlation. An explanation of this fact is because of dense networks give a high set of random walks to go and return, and it is comparable to shortest paths. Dense local communities in a network let us obtain a better correlation between Pagerank and the 2RW measure because of several random walks crossing a node as a transition step means that this node is an influential node into its community.

Increasing the number of inter-cluster links and consequently, the density of the network, the correlation between Pagerank and 2RW is worse than classic betweenness because our measure focuses on the importance of local densities while CBT finds new inter-communities shortest paths.

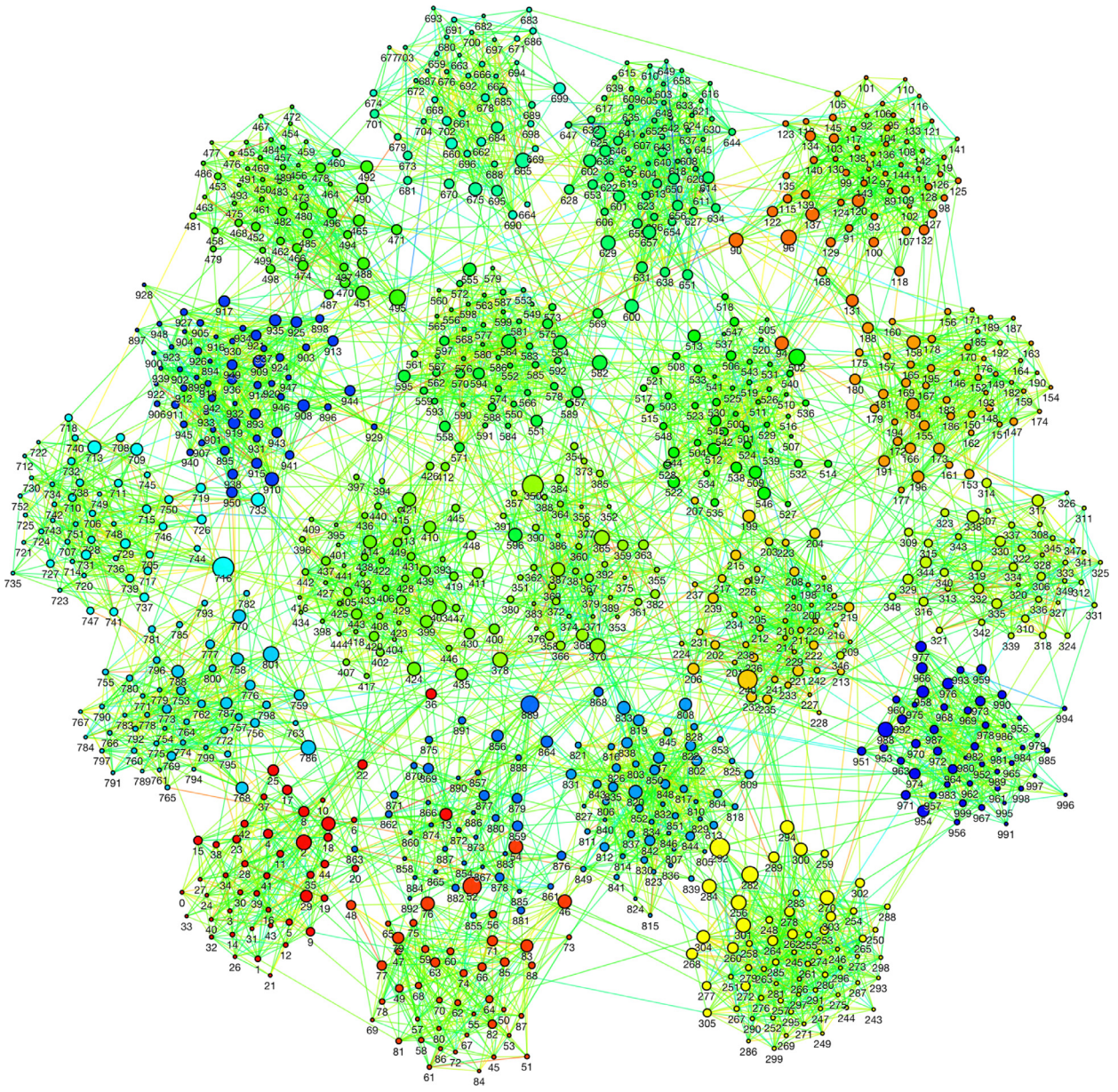


Fig. 4. Synthetic network representing a social network with 1000 nodes and 6612 edges, distributed in 20 clusters. The global density of this network is 0.39%.

4.2. Real networks

In this section, we will study the behaviour of the proposed centrality applied to two real networks, such as the networks from the third book of the popular adventures of Game of Thrones and the train bombing terrorist network of Madrid attacks on March, 11 of 2004.

Table 5 show some basic properties of these networks:

4.2.1. Game of thrones network

The following example is a transformation of the third book *A Storm of Swords from the Game of Thrones* saga [34] into a social network (see Fig. 5). The network consists of 107 nodes, representing all the characters involved in this book. Each of these nodes joins another from an edge, which is larger or smaller, in thickness, depending on the number of interactions between the two characters. The methodology followed for the weights of the edges has been to count the number of interactions, which forms the edge, between characters such as the appearance of the names of these consecutively without

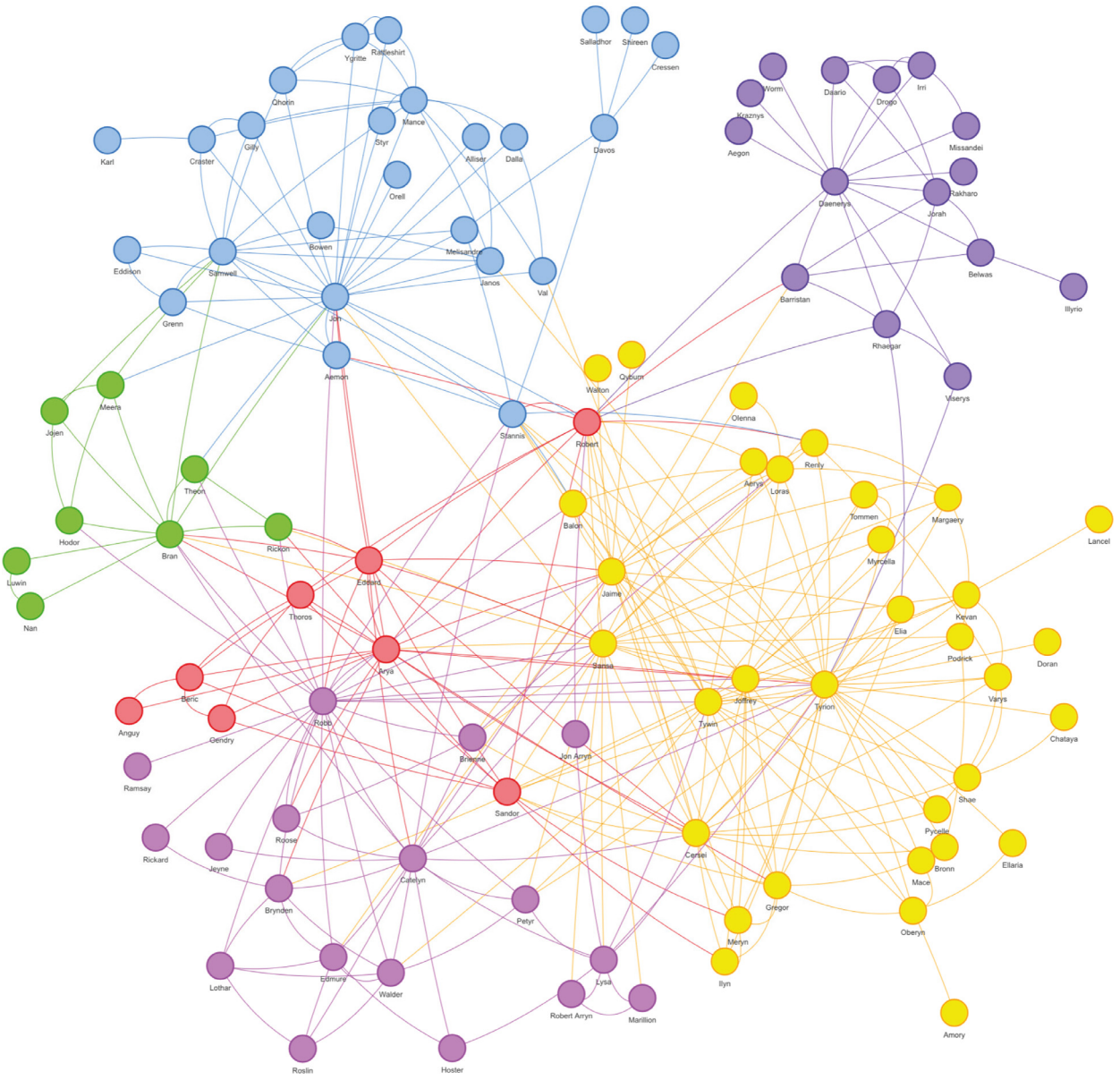


Fig. 5. Game of Thrones Network representing different plots (coloured clusters) of the third book titled *A Storm of Swords* from *The Game of Thrones* saga. Nodes are the characters and links represent the interactions between them.

having among them more than 15 words of the text of the book. In the figure each colour indicates the community to which the character belongs, there are seven communities, and the thickness of each node corresponds to its betweenness centrality.

Ranking the characters/nodes with classic betweenness centrality (CBT) and our method (2RW), it is possible to observe (see Table 6) the positions of the nodes regarding the CBT centrality. In the last column of this table, we can observe the difference in ranking between both measures. Focusing on the major differences, for instance, nodes 57 (*Samwell*), 14 (*Cersei*) and 30 (*Joffrey*) have better ranking in our method than CBT. This is because of these nodes belong to dense clusters (yellow and blue) and strengthen the importance of the main characters of their plots. Moreover, a dense cluster will have a global reinforcement of theirs betweenness indices.

Nodes 54 (*Robert*), 17 (*Daenerys*) and 62 (*Stannis*) suffering a relevant fall in the ranking. This drop in the ranking does not mean less importance of these nodes, but the weakness of the network. Remove these nodes represent an important cut in the network, where some nodes run the risk of becoming an isolated node. This is because of these nodes belong to

Table 5
Basic properties and characteristics of two real networks: number of nodes, number of nodes, number of edges, average degree with the standard deviation ($\overline{d} \pm \sigma$), random walk mixing time (τ), and global clustering coefficients (c).

Network	Nodes	Edges	$\overline{d} \pm \sigma$	τ	c
Game of Thrones	107	352	6.58 ± 6.58	14.10	0.33
Train Terrorists	64	243	7.59 ± 6.16	15.42	0.56

Table 6
Comparison between the 20-top rankings of CBT and 2RW methods (columns 1 and 3 represent the position of the ranking in terms of centrality values, and the last column is the difference in rankings). Colour represents their community (see Fig. 5).

CBT Rank	Character (node)	2RW Rank	Difference (CBT - 2RW)
1	Jon (32)	3	−2
2	Robert (54)	14	−12
3	Tyrion (64)	1	+2
4	Daenerys (17)	16	−12
5	Robb (53)	3	+2
6	Sansa (59)	2	+4
7	Stannis (62)	19	−12
8	Jaime (28)	5	+3
9	Arya (5)	10	−1
10	Tywin (65)	6	+4
11	Bran (9)	13	−2
12	Davos (18)	58	−46
13	Catelyn (13)	10	+3
14	Barristan (75)	25	−11
15	Samwell (57)	9	+6
16	Cersei (14)	7	+9
17	Joffrey (30)	8	+9
18	Janos (29)	34	−16
19	Oberyn (46)	28	−9
20	Eddard (19)	21	−1

Table 7
5-Top Ranking Differences between Classic Betweenness ranking (CBT) and 2RW (we show the highest ranking difference from 20-top ranking). Colour represents their community (see Fig. 5).

Character (node)	Difference (CBT → 2RW)
Mance (40)	+17 (29 → 12)
Gregor (24)	+15 (32 → 17)
Sandor (58)	+11 (26 → 15)
Cersei (14)	+9 (16 → 7)
Joffrey (30)	+9 (17 → 8)
Davos (18)	−46 (12 → 58)
Robert (54)	−12 (2 → 14)
Daenerys (17)	−12 (4 → 16)
Stannis (62)	−12 (7 → 19)
Barristan (75)	−11 (14 → 25)

sparse clusters or, what is the same, their communities are more hierarchical (for instance in star or tree topologies). In this kind of network, the nodes have not a set of random walks to choose.

In Table 7 we can show the highest differences between CBT and 2RW rankings. For instance, *Mance* reinforces to *Jon Snow* in *The Wall* plot (blue cluster), or *Gregor, Cersei* and *Joffrey* reinforce to main characters in *King's Landing* plot (yellow cluster). On the other hand, *Davos, Robert* or *Daenerys* belong the weak points of the network and if these characters die, their plots and characters could be isolated.

Finally, we calculate the correlation between CBT and 2RW measures, achieving a high correspondence (93.81%). Comparing Pagerank and both measures, our method has a better correlation than CBT (97.29% vs 93.58%).

Complementary, we try to remove the most important node for each measure (*Jon* in CBT and *Tyrion* in 2RW) to measure the importance of these nodes in the network. If we remove *Jon* node, the correlation between both indices is reduced to 90.36%, but comparing with Pagerank, our method improves CBT again (89.12% vs 87.33%).

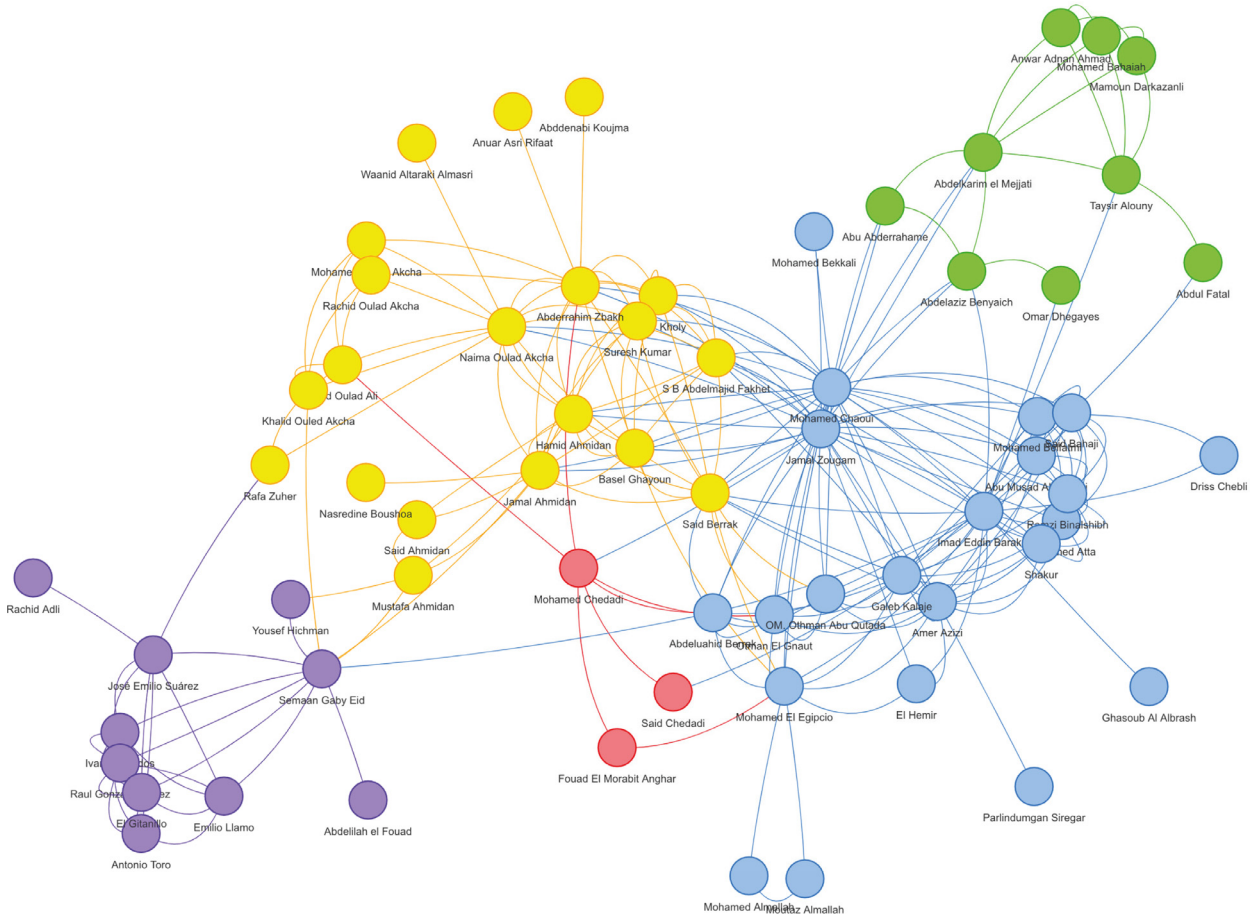


Fig. 6. Train Terrorist Network. A network with 70 terrorists, 342 edges and 6 clusters represented by different colours. Note that 6 of the terrorist are isolate nodes and its centrality values are zero, removing them from the representation.

On the other hand, if we remove *Tyrion*, the correlation is similar than the previous experiment (90.5%), but the resulting network highlight the importance of nodes in Pagerank at a high level (95.08% with our method, and 93.92% with CBT).

Summarizing, we can say that the 2RW measure not only remark the most intermediary characters (Betweenness centrality), but it point out the influential nodes in the story (Pagerank).

4.2.2. Madrid train bombing terrorist network

In Fig. 6 is the depiction of the network of terrorists involved in the March 11, 2004 bombing [35]. The news of the attack published in the two most important newspapers of the Spanish press, *"El Mundo"* and *"El Pais"*, were used for its elaboration. The nodes of the network represent suspicious people of having participated in the attack as well, as each of their relatives. The edges that join the different nodes represent any kind of linkage in the network between terrorists, and the connection strength represents the following: i) trust-friendship (contact, kinship, links in the telephone center), ii) ties to *al Qaeda* and to *Osama Bin Laden*, iii) co-participation in training camps and/or wars, and iv) co-participation in the previous terrorist. The dataset consists of 64 terrorists (nodes) and 243 relationships between them (links) after removing the six isolated nodes.

The core of a terrorist network is the field operations group, which includes who placed the explosives. This group is integrated by the following 10 terrorists: *Jamal Zougam, Mohamed Chaoui, Said Berrak, Basel Ghayoun, Hamid Admidan, SB Abdelmajid Fakhet, Rachid Oulad Akcha, Mohamed Oulad Akcha, Allekema Lamari, Abddenabi Koujma and Anuar Asri Rifaat*.

Table 9 show 14 nodes that represent the top 10 nodes of the CBT plus four nodes that belong to the top 10 2RW ranking. It is necessary to point out that the last column is the difference in the position of both rankings.

Observing the mentioned Table 9, it is possible to see some nodes as *Semaan Gaby Eid*, which is important in terms of betweenness in the network (first in CBT rank), has a worse ranking in our 2RW ranking in contrast to other nodes as *Jamal Zougam* or *Mohamed Chaoui*, that have a similar ranking. A strong drop in ranking does not mean less importance in the network, but these terrorist are a link in the network without communication alternatives. In other words, a node with a high 2RW value means a well-connected cluster, where the node has different communication channels.

Require: A weighted graph $G = (V, E, W)$
Ensure: A vector C with the centrality value for each node
for all $i, j \in V$ **do**
  for all $t \in V$ **do**
    if $t \neq i \neq j$ **then**
      $V_{P_{itj}} = \frac{w_{it} w_{tj}}{d_i d_j}$ [Eq. 3, 4]
    else
      $V_{P_{itj}} = 0$
    end if
  end for
end for
for all $i, j \in V$ **do**
  $T_{ij} = V_{P_{itj}}^T V_{P_{jti}}$ [Eq. 5]
  $\text{diag}(T_{ij}) = 0$
  $[t, k, maxValue] = \max(T_{ij})$[Eq. 6]
  $C_t = C_t + 1$ [Eq. 7]
  $C_k = C_k + 1$
end for

Algorithm 1: Two-way Random Walk Betweenness Centrality Algorithm.

Table 8
Correlations between Pagerank (PR), the Classic Betweenness (CBT) and our method (2RW). First row is the original network, and the remaining rows are different versions of this network removing a character.

Network	Corr. CBT-2RW	Corr. PR-2RW	Corr. PR-CBT
Original	93.81%	97.29%	93.58%
Without *Jon*	90.36%	89.12%	87.33%
Without *Tyrion*	90.5%	95.08%	93.92%

Table 9
10-Top Ranking Comparison between Classic Betweenness measure (CBT) and 2RW (four terrorists that belong to the 10-top 2RW Ranking have been added).

CBT Rank	Terrorist	2RW Rank	Difference (CBT - 2RW)
1	Semaan Gaby Eid	21	−20
2	Jamal Zougam	1	+1
3	Mohamed Chaoui	2	+1
4	Abdeluahid Berrak	15	−11
5	Imad Eddin Barakat	3	+2
6	Jamal Ahmidan	7	−1
7	Naima Oulad Akcha	5	−2
8	Abderrahim Zbakh	4	+4
9	Mohamed El Egipcio	28	−19
10	Abdelkarim el Mejjati	14	−4
11	Amer Azizi	6	+5
12	Jose Emilio Suarez	10	+2
16	Hamid Ahmidan	9	+7
25	Mohamed Belfatmi	8	+17

The arrest of a terrorist with high CBT and worse 2RW implies access to a person with much information of the network and a part of the network will be affected or disconnected without him. However, if we arrest a terrorist with low CBT and higher 2RW means a less influential person in the network, that has little information available, where his main role in the network is to reinforce a cluster or community to do stronger.

Attending to the correlation between CBT and our method, we obtain (85.08%), and if we compare Pagerank and both betweenness measures, the results are similar (82.88% vs 83.38%). However, the most relevant of our proposal is the search of the most dense-connected nodes. For instance, the centrality of *Abderrahim Zbakh* increases with our ranking (from 8th with CBT to 4th in 2RW). Although he does not belong to the field operation group, he is the unique connection of two of them with the remaining of the network (*Abddenabi Koujma and Anuar Asri Rifaat*). Moreover, he is a unique person with connection with all member of this group.

In Table 10, we show the maximum differences in ranking between CBT and 2RW. The positive changes in ranking represent the most hyper-connected terrorists. If we remove one of them, the network will continue working.

Table 10
5-Top Ranking Differences between Classic Betweenness measure (CBT) and 2RW (we show the highest ranking difference from top ranking or to top).

Terrorist (node)	Difference (CBT → 2RW)
Mohamed Bahaiah (55)	+43 (63 → 20)
Antonio Toro (35)	+37 (50 → 13)
Mohamed Belfatmi (15)	+17 (25 → 8)
Farid Oulad Ali (20)	+11 (23 → 12)
Hamid Ahmidan (33)	+7 (16 → 9)
Semaan Gaby Eid (63)	−20 (1 → 21)
Mohamed El Egipcio (61)	−19 (9 → 28)
Galeb Kalaje (18)	−15 (15 → 30)
Rafa Zuher (23)	−13 (19 → 32)
Abdeluahid Berrak (40)	−11 (4 → 15)

On the other hand, we represent the negative changes in ranking, that represent the weakness of the network. If we remove these nodes, we will disconnect a relevant part of the network or reduce the communication between several terrorists.

In other words, if we arrest to *Hamid Ahmidan*, who belong to the field operation group, we will not have much information and the network will not be very affected because our method point that there are alternative paths of communication. In contrast, if we arrest to *Mohamed El Egipcio*, who does not belong to the field operation group and his importance in the network is similar (Pagerank) than *Galeb Kalaje*, but we reduce the communication of the network because there are not many alternatives to this node.

5. Conclusions

The betweenness methods, based on shortest paths and random walk, measure the global importance of a node as an intermediate node, but they have the common characteristic of not taking into account the density of the cluster of each node. To solve this, a new centrality measure based on betweenness random walks has been proposed. And that is precisely the goal of the paper: defining a new centrality and presenting its main properties. From a dense-networks viewpoint, a quantification of the importance of a node through the relationships between four different nodes is presented in this new measure. We have applied analysis of network centrality, from a perspective oriented to ranking nodes, reinforcing dense communities by means of evaluating graphs using a two-trip transition probability matrix. Thus, a centrality measure combining the idea of the random-walk betweenness centrality and the link predictor algorithm Return Random Walk is presented.

The experimental results on two synthetic networks and two real networks show that the proposed measure perform relatively better than the traditional measures (as betweenness centrality) and also rank nodes differently. In other words, it ranks similar, presenting common characteristics, although the idea behind the proposed centrality is closer to the CBT measure. In detail, the proposed metric increases the ranking of nodes belonging to dense clusters with a higher average degree than the remaining clusters, observing that it works better in dense networks. Moreover, we can detect the weakness of a network comparing CBT method with our proposal (2RW).

References

[1] L.C. Freeman, A set of measures of centrality based on betweenness, Sociometry (1977) 35–41.
[2] M. Kurant, P. Thiran, Layered complex networks, Phys. Rev. Lett. 96 (13) (2006), doi:10.1103/physrevlett.96.138701.
[3] P. Holme, B.J. Kim, C.N. Yoon, S.K. Han, Attack vulnerability of complex networks, Phys. Rev. E 65 (2002) 056109, doi:10.1103/PhysRevE.65.056109.
[4] M. Kivela, A. Arenas, M. Barthelemy, J.P. Gleeson, Y. Moreno, M.A. Porter, Multilayer networks, J Complex Netw 2 (3) (2014) 203–271, doi:10.1093/comnet/cnu016.
[5] F. Battiston, V. Nicosia, V. Latora, The new challenges of multiplex networks: measures and models, The European Physical Journal Special Topics 226 (3) (2017) 401–416, doi:10.1140/epjst/e2016-60274-8.
[6] S. Boccaletti, G. Bianconi, R. Criado, C. del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendiña-Nadal, Z. Wang, M. Zanin, The structure and dynamics of multilayer networks, Phys Rep 544 (1) (2014) 1–122, doi:10.1016/j.physrep.2014.07.001. The structure and dynamics of multilayer networks
[7] Y. Moreno, M. Perc, Focus on multilayer networks, New J Phys 22 (1) (2019) 010201.
[8] Y.A. Kim, H.S. Song, Strategies for predicting local trust based on trust propagation in social networks, Knowl Based Syst 24 (8) (2011) 1360–1371, doi:10.1016/j.knosys.2011.06.009.
[9] Z. Lv, N. Zhao, F. Xiong, N. Chen, A new measure of identifying influential nodes: efficiency centrality, Commun. Nonlinear Sci. Numer. Simul. 47 (2017) 151–163, doi:10.1016/j.cnsns.2016.11.008.
[10] X. Zhang, J. Zhu, Q. Wang, H. Zhao, Identifying influential nodes in complex networks with community structure, Knowl Based Syst 42 (2013) 74–84, doi:10.1016/j.knosys.2013.01.017.
[11] M. Gosak, R. Markovič, J. Dolenšek, M.S. Rupnik, M. Marhl, A. Stožer, M. Perc, Network science of biological systems at different scales: a review, Phys Life Rev 24 (2018) 118–135.
[12] R. Yang, B.-H. Wang, J. Ren, W.-J. Bai, Z.-W. Shi, W.-X. Wang, T. Zhou, Epidemic spreading on heterogeneous networks with identical infectivity, Phys. Lett. A 364 (3) (2007) 189–193, doi:10.1016/j.physleta.2006.12.021.
[13] J. Borge-Holthoefer, Y. Moreno, Absence of influential spreaders in rumor dynamics, Phys. Rev. E 85 (2012) 026116, doi:10.1103/PhysRevE.85.026116.

[14] C. Gao, X. Lan, X. Zhang, Y. Deng, A bioinspired methodology of identifying influential nodes in complex networks, PLosOne 8 (2013) e66732, doi:10.1371/journal.pone.0066732.
[15] L. Lü, T. Zhou, Link prediction in complex networks: a survey, Physica A 390 (6) (2011) 1150–1170, doi:10.1016/j.physa.2010.11.027.
[16] R. Albert, H. Jeong, A.-L. Barabási, Diameter of the world-wide web, Nature 401 (6749) (1999) 130–131, doi:10.1038/43601.
[17] M.E.J. Newman, The structure and function of complex networks, SIAM Rev. 45 (2) (2003) 167–256, doi:10.1137/s003614450342480.
[18] S. Strogatz, Exploring complex networks, Nature 410 (6825) (2001) 268–276, doi:10.1038/35065725.
[19] J. Kolasa, Complexity, system integration, and susceptibility to change: biodiversity connection, Ecol. Complexity 2 (4) (2005) 431–442, doi:10.1016/j.ecocom.2005.05.002.
[20] P. Crucitti, V. Latora, M. Marchiori, A. Rapisarda, Error and attack tolerance of complex networks, Physica A 340 (1) (2004) 388–394, doi:10.1016/j.physa.2004.04.031. News and Expectations in Thermostatistics
[21] J. Memmott, N.M. Waser, M.V. Price, Tolerance of pollination networks to species extinctions, Proceedings: Biological Sciences 271 (1557) (2004) 2605–2611.
[22] E. Estrada, Órjan Órjan Bodin, Using network centrality measures to manage landscape connectivity, Ecol. Appl. 18 (7) (2008) 1810–1825.
[23] L.C. Freeman, Centrality in social networks conceptual clarification, Soc Networks 1 (1979) 215–239.
[24] E. Bozzo, M. Franceschet, Resistance distance, closeness, and betweenness, Soc Networks 35 (3) (2013) 460–469.
[25] I. Kivimäki, B. Lebichot, J. Saramäki, M. Saerens, Two betweenness centrality measures based on randomized shortest paths, Sci Rep 6 (1) (2016) 1–15.
[26] M.E. Newman, A measure of betweenness centrality based on random walks, Soc Networks 27 (1) (2005) 39–54.
[27] P. Holme, J. Saramäki, Temporal networks, Phys Rep 519 (3) (2012) 97–125, doi:10.1016/j.physrep.2012.03.001. Temporal Networks
[28] U. Brandes, A faster algorithm for betweenness centrality, J Math Sociol 25 (2001) 163–177.
[29] U. Brandes, D. Fleischer, Centrality measures based on current flow, in: Annual symposium on theoretical aspects of computer science, Springer, 2005, pp. 533–544.
[30] M. Curado, Return random walks for link prediction, Inf Sci (Ny) 510 (2020) 99–107.
[31] .M.C. Langville A., Google'sPageRank and Beyond: The Science of Search Engine Rankings, Princeton University Press, Oxford, UK, 2006, doi:10.2307/j.ctt7t8z9.
[32] L.C. Freeman, S.P. Borgatti, D.R. White, Centrality in valued graphs: a measure of betweenness based on network flow, Soc Networks 13 (2) (1991) 141–154.
[33] R.A. Rossi, N.K. Ahmed, The network data repository with interactive graph analytics and visualization, in: Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, 2015. http://networkrepository.com
[34] A. Beveridge, J. Shan, Network of thrones, Math Horizons 23 (4) (2016) 18–22.
[35] B. Hayes, Connecting the dots, Am. Sci. 94 (5) (2006) 400–404.

4.3. ARTÍCULO 3

Este artículo ha sido publicado en la revista *Journal of Computational Science*, en la editorial *ElSevier*.

Se enmarca dentro de un proyecto nacional competitivo de Ministerio de Ciencia e Innovación, titulado *ALLEGRO: Smart multi-modal crowdsensing-based system as a service oriented to the prediction of social problems*.

A continuación, se muestra algunos indicios de calidad de la revista, y otros indicios adicionales de otras plataformas.

Tabla 4. Otros indicios de calidad del artículo 3

Fecha de publicación	Citescore	5- Year Impact Factor	Lecturas en ResearchGate
Julio 2022	8.8	3.976	57

Journal of Computational Science 63 (2022) 101819

Contents lists available at ScienceDirect

Journal of Computational Science

journal homepage: www.elsevier.com/locate/jocs

A centrality model for directed graphs based on the Two-Way-Random Path and associated indices for characterizing the nodes

Manuel Curado [a,*,1], Rocio Rodriguez [b,1], Fernando Terroso-Saenz [a,1], Leandro Tortosa [c,1], Jose F. Vicent [c,1]

[a] Polytechnic School, Catholic University of Murcia, Campus Los Jeronimos, s/n, E-30107 Murcia, Spain
[b] Department of Mechanical Engineering, Catholic University of Ávila, Calle Canteros, s/n, E-05005, Ávila, Spain
[c] Department of Computer Science and Artificial Intelligence, University of Alicante, Campus de San Vicente, Ap. Correos 99, E-03080, Alicante, Spain

ARTICLE INFO

Keywords:
Centrality measure
Betweenness centrality
Random paths
Densification

ABSTRACT

Centrality metrics are one of the most meaningful features in a large number of real-world network systems. In that sense, the Betweenness centrality is a widely used measurement that quantifies the importance of a node in the information flow in a network. Moreover, there is a centrality measure, based on random-paths betweenness centrality, that provides a classification of the nodes of undirected networks, that are able to reinforce dense communities according to their role. In this paper, a new centrality model, based on random-paths betweenness centrality and applied on directed networks, is presented. This model, based on four indices, describes the behaviour of the nodes within the network in terms of its role, such as a transition node, in the same cluster or between clusters. Finally, we evaluate the model with several use cases based on real networks, two of them are proposed and created in this paper, giving insight into some interesting findings about the networks' features.

1. Introduction

Complex systems in real life are ubiquitous, as we can see in many types of real networks as social networks, air and rail networks, mobile phone networks, urban networks or power and electricity networks, among others. Generally, complex systems can be described by the interdisciplinary field of complex networks that, in recent decades, has provided different models and methods that have impacted our everyday life [1,2]. Complex networks (see [3,4]) can be used as frameworks for many real-world complex systems because of, on the one hand, the accessibility of datasets and, on the other hand, the digital technologies have increased the interest of study non-trivial structural properties of the networks. [5]. Thus, for instance, the spreading processes on networks, such as the spreading of epidemic disease, the information dissemination or the diffusion of ideas among community members, can be described by complex structures of networks in which the influence between nodes across the network is one of their fundamental functions [6,7].

In order to measure relevant properties of complex networks, that can be affected by some special nodes, many studies have been conducted. In particular, the topology of a complex network can determine its dynamic as we can see in [8,9] where the authors demonstrate these dependencies by removing important nodes of the network. Furthermore, different methods may consider different structural properties, which in turn, may lead to different rankings [10,11].

Hence, the question of how to measure the relative importance of nodes in a network is increasingly challenging in the field of complex networks [12,13]. To understand the role of nodes in a network, the node centrality analysis provides an efficient framework and, in that sense, the centrality measures are able to rank the nodes based on their influence or how they are affected by other nodes via their connection topology. One of the reasons why so many centrality measures have been defined is because all measures have limitations. For instance, in [14] the authors propose a new centrality measure to identify the influential nodes in directed complex networks, but this index is focused on the study of local structures, according to a limited amount of information contained in the inner structures of boxes calculated from shortest distances. Other measures, such as [15], consider the global and local information of the complex network using the effective distance to study the topological structure information of the real networks. However, it works with undirected networks. According

* Corresponding author.
E-mail addresses: mcurado@ucam.edu (M. Curado), rocio.rodriguez@ucavila.es (R. Rodriguez), fterroso@ucam.edu (F. Terroso-Saenz), tortosa@ua.es (L. Tortosa), jvicent@ua.es (J.F. Vicent).
[1] Contributed equally to the work.

https://doi.org/10.1016/j.jocs.2022.101819
Received 4 February 2022; Received in revised form 11 July 2022; Accepted 26 July 2022
Available online 10 August 2022

to the literature, each centrality measure works well to test certain phenomena, but at the same time, each measure also fails to capture other important structural characteristics of a network. Then, one of the limitations is that different measures can disagree on the most important vertex in a network. Although that can be viewed as a benefit since each centrality measure is just a way to quantify a particular type of importance (see [16,17]).

Regarding the connectivity of a network, a metric that evaluates the importance of a node from the perspective of the flow of information between pairs of them is the betweenness centrality [18]. This measure has applications in a significant number of areas such as transportation, communications or infrastructural networks [19,20]. Because of their important effects on the network functioning, the betweenness centrality has been applied in all types of networks, static or dynamic, single layer or multilayer [21] but, the initial shortest-path betweenness centrality is not appropriate in networks governed by rules different to shortest paths. There are several types of betweenness centrality [22], but in their simplest form, it can be explained as the number of times that a node is crossed, by random paths between two other nodes (counting all possible paths crossing it).

In [23], a novel centrality measure, based on random paths betweenness centrality, is proposed. This measure provides a ranking of the nodes of a network, with the main characteristic of reinforcing the dense communities according to their role: relevance intra-cluster (how the node reinforces the network) or inter-cluster (how much the network is disconnected when we remove the target node). However, this method is only proposed for undirected graphs, and it does not give more information about the behaviour of a node. In order to obtain new and relevant measures for directed networks, the *Directed Two-way Random Path Betweenness Centrality Algorithm (D2RWBT)* is proposed. This centrality model is based on four indices that measure the importance of a node as an intermediate link describing some behaviour properties. Therefore, with this model we describe the behaviour of the node within a network by obtaining information about some characteristics as whether the node is located in a dense cluster or not,if from the point of view of its centrality the node is important, if the node plays an important role as a transition node between clusters (inter-clusters) or within the node is important in the same cluster (intra-clusters).

To achieve this objective, the paper is organized as follows: Section 2 describes some preliminaries necessary for the correct understanding of the proposed centrality model. In Section 3, a detailed explanation of the composition of the model, including a toy example and the behaviour of the indices, is put forward. The evaluation the proposed model is discussed in Section 4 by means of three networks where two of them are based on real-world data. Finally, the conclusions are summarized in Section 5.

2. Preliminary

In order to capture important structural characteristics of a network that could not be fully described with the existing centrality measures, a new model focused on directed networks is proposed. This model is based on random paths betweenness centrality and because of this, some preliminary notes are presented in this section (for a more complete review see [24]). It can also be seen in this section some pros and cons of the centrality measures on which the proposed model is based (see Table 1).

2.1. In-degree and out-degree

Let $G(V, E, W)$ be a directed weighted graph, where V is the set of nodes, E is the set of edges $e_{ij} \in E$ starting at node i and ending at node j. Finally, W is the weighted matrix with $w_{ij} > 0$, if $e_{ij} \in E$ and 0 otherwise. The directed adjacency matrix A is defined as follows:

$$A_{ij} = \begin{cases} w_{ij} & \text{if } e_{ij} \in E \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

We measure the connectivity between nodes taking into account both in-degree and out-degree values of each node i. We will denote in the following the in-degree and out-degree values, for the node i, as d_i^- and d_i^+, respectively. Remember that in-degree of i refers to the number of arcs incident to i, while out-degree refers to the number of arcs incident from i.

The idea behind the centrality measure for undirected graphs is based on the number of paths between two nodes (origin i and destiny j) through a transition node t. In this path, we are interested in two degrees: out-degree of the origin node (d_i^+) to evaluate the probabilities to start at node i, and the in-degree of the destination node (d_i^-) to evaluate the probabilities to end at node j.

2.2. Betweenness centrality

Betweenness is a centrality measure of a vertex within a graph that quantifies the number of times a node acts as a bridge along the shortest path between two other nodes [18].

Observing the Betweenness formula in Table 1, it can be deduced that the total number of paths between source and target is counted differently for directed and undirected graphs.

The shortest paths betweenness centrality measures is one of the several methods based on the idea that edges or links represent channels that transmit *flow items*. The betweenness centrality of a node t is defined as the average of the maximum flow through t over all possible pairs (i, j).

There are some modifications to this centrality. For instance, one of them is based on random paths [20]. Compared with the original shortest-path version, the random-path betweenness changes in terms of the information transmission between any pairs of nodes that, now, follows a random path.

As it is shown in several works [19,25], the random-path betweenness is identical to current-flow betweenness. In order to optimize the random-path betweenness for computing all nodes, the work in [22] proposes an efficient algorithm that takes time $O(I(n-1) + mn\log n)$, taking into account that $I(n) = O(n^3)$ is the time for computing the inverse of a matrix with n nodes.

3. Directed Two-way Random Path Betweenness Centrality Algorithm (D2RWBT)

Based on the work described in [23] and with the intention of adapting this model to directed networks, we present a centrality model based on four indices whose main goal is to characterize the centrality betweenness of a node in a 4-relationship context (linking two nodes crossing other two transition nodes in random paths context [26]), located in a directed network. This approach is useful to award more importance to betweenness nodes in a dense cluster. In that sense, it is important to remark that, in this method, isolated nodes and self-loops have a null centrality.

3.1. Formal analysis

Before going into details about the mathematical formulation of the model, Table 2 includes the notation that will be used in the explanation.

Given a directed weighted graph $G(V, E, W)$, where V is the set of n nodes, E is the set of edges and $W \in \mathbb{R}^{n\times n}$ is the weighted matrix. We will denote a particular node as i, an each directed edge from an origin node i to a destination node j as e_{ij} and the weight of this arc e_{ij} as w_{ij}, with $e_{ij} \in E$.

With this matrix W, and all-ones vector $\boldsymbol{u} \in \mathbb{R}^{n\times 1}$, we calculate the vector

$$D^+ = W\boldsymbol{u} = \left(\sum_{l=1}^{n} w_{1l}, \sum_{l=1}^{n} w_{2l}, \ldots, \sum_{l=1}^{n} w_{nl}\right)^T$$

Table 1
Description of the centrality measures on which the presented model is based.

Centrality measure	Definition	Formula	Drawbacks
Degree	It is determined, according to in- and out-degree, the number of nodes with which it is connected.	$D^- = \sum_{e_{it} \in E} A_{it}$ $D^+ = \sum_{e_{ti} \in E} A_{ti}$	It can be considered as a local centrality measure, as a hub (densely connected node) may not be central.
Betweenness (Shortest-path) [18]	Number of times that a shortest path starting at i and ending at j passes through t along the way, averaged over all i and j.	$B_{SP} = \sum_{i \neq t} \sum_{j \neq i,t} \frac{\sigma_{itj}}{\sigma_{ij}}$	It assumes that a flow through a network moves along the shortest paths, but this is not always the case.
Betweenness (Random-walk)[20]	Number of times that a random walk starting at i and ending at j passes through t along the way, averaged over all i and j.	$B_{RW} = \sum_{i \neq t} \sum_{j \neq i,t} \frac{\sigma_{itj}}{\sigma_{ij}}$	It does not take into account the dense communities.

Table 2
Summary of the mathematical notation used.

G, V, E	Graph, Node set, Edge set.
W	Weighted matrix.
e_{ij}	Arc from i to j.
$w_{e_{ij}}$	Weight of the arc e_{ij}.
D^+	Node's weights of the out-degrees.
D^-	Node's weights of the in-degrees.
P_{itj}	Probability to reach a destination node j from an origin node i through transition node t.
$V_{P_{itj}}$	Vector whose elements are matrices representing all possible 2-step paths between nodes i and j through any different transition node t.
$T^{(i,j)}$	The transition matrix, for all i, j.
C^1	Interaction index.
C^2	Relationship importance index.
C^3	Individual Relevance index.
C^4	Global Relevance index.
ED	Effective distance.

Therefore, D^+ is a column vector containing the sum of the weights of the out-edges for each node. Similarly, the vector

$$D^- = \boldsymbol{u}^T W = \left(\sum_{l=1}^{n} w_{l1}, \sum_{l=1}^{n} w_{l2}, \ldots, \sum_{l=1}^{n} w_{ln} \right)$$

is a row vector with the sum of the weights of the in-edges for each node.

The goal of this algorithm is to find paths within G comprising 4 vertices and 4 edges. To do so, we firstly evaluate all the possible 2-step paths between each pair of nodes i and j in the graph G, where node t is the *transition* between both nodes ($i \to t \to j$). This is done by computing a matrix P by means of a 2-length random path between three different nodes $i, j, t \in V$, taking into account that i is the origin, t the betweenness node, and j the destination.

$$P_{itj} = \frac{w_{it} w_{tj}}{D^+(i) D^-(j)}, \tag{2}$$

where $D^+(i)$ is the ith component of the vector D^+ and $D^-(j)$ is the jth component of the vector D^-. Remark that w_{it} and w_{tj} are the weights corresponding to the edges e_{it} and e_{tj}, respectively.

The element P_{itj} represents the probability to reach a destination node j from an origin node i through the transition node t. Then, we calculate the vector $V_{P_{itj}}$, which represents all possible 2-length paths between nodes i and j through any different transition node t, as

$$V_{P_{itj}} = \begin{pmatrix} P_{i1j} & P_{i2j} & P_{i3j} & \cdots & P_{inj} \end{pmatrix}_{1 \times n}. \tag{3}$$

Remark that $P_{itj} = 0$ if $i = j$, $i = t$, or $j = t$.

Following the aforementioned approach but in reverse order, we evaluate all possible 2-step paths between each pair of nodes but now from origin node j to destination node i, with a transition node k. ($j \to k \to i$). Then, we calculate the random-path matrix of these three nodes as follows,

$$P_{jki} = \frac{w_{jk} w_{ki}}{D^+(j) D^-(i)}, \tag{4}$$

where $D^+(j)$ is the jth component of the vector D^+ and $D^-(i)$ is the ith component of the vector D^-. In this case, w_{jk} and w_{ki} are the weights corresponding to the edges e_{jk} and e_{ki}, respectively.

From P_{jki}, we can compute all possible 2-length paths between nodes j and i through any different transition node k, as

$$V_{P_{jki}} = \begin{pmatrix} P_{j1i} & P_{j2i} & P_{j3i} & \cdots & P_{jni} \end{pmatrix}_{1 \times n}.$$

At this point, it is important to remark that we actually define two different transition nodes denoted as t and k respectively

Finally we calculate each transition matrix $T^{(i,j)}$ for all possible pair of nodes $i, j \in V$ as follows:

$$T^{(i,j)} = V^T_{P_{itj}} V_{P_{jki}} = \begin{pmatrix} P_{i1j} \\ P_{i2j} \\ P_{i3j} \\ \cdots \\ P_{inj} \end{pmatrix} \begin{pmatrix} P_{j1i} & P_{j2i} & P_{j3i} & \cdots & P_{jni} \end{pmatrix}$$
$$= \begin{pmatrix} P_{i1j}P_{j1i} & P_{i1j}P_{j2i} & P_{i1j}P_{j3i} & \cdots & P_{i1j}P_{jni} \\ P_{i2j}P_{j1i} & P_{i2j}P_{j2i} & P_{i2j}P_{j3i} & \cdots & P_{i2j}P_{jni} \\ P_{i3j}P_{j1i} & P_{i3j}P_{j2i} & P_{i3j}P_{j3i} & \cdots & P_{i3j}P_{jni} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ P_{inj}P_{j1i} & P_{inj}P_{j2i} & P_{inj}P_{j3i} & \cdots & P_{inj}P_{jni} \end{pmatrix}. \tag{5}$$

With these matrices, we find the probabilities of a path starting in a node i, passing by a transition node t, reach a destination node j, and return to i passing by the transition node k ($i \to t \to j \to k \to i$). Note that, these matrices have assigned zero to the diagonal because both transition nodes must be different. This way, we force that two pairs of nodes cannot be connected the same intermediate go and return node.

Given this go-return approach, a random path always comprises four nodes and the probability will be different from zero if $i \neq j \neq t \neq k$, in other cases the resulting probability is zero.

A paramount aspect of this method is the fact that the representation of relationships will involve four different nodes. This can be regarded as a minimum-sized community. Hence, if a node is not included in any 4-step random path, its four indices will be zero.

The previous description of the model reflects the way in which the implementation has been carried out, which leads us directly to the definition of the indices that characterize the centrality measure. But before going into detail in the definition of these indices, it is convenient to carry out a mathematical interpretation of the process developed so far.

The mathematical interpretation of the process is based on a known property about the powers of the adjacency matrix of a graph. We

know that if we calculate the adjacency matrix squared, we obtain the number of paths between its vertices of length 2. Similarly, successive powers of higher order give us the number of paths of length equal to the order of the power.

More specifically, the matrix A^4 gives us the paths of length 4 between the different nodes of the network. That is, if we have that the element A^4_{ij} gives us the total number of paths from i to j of length 4. Following a similar reasoning, the elements A^4_{ii} gives us the total number of paths from i to i of length 4. Consequently, we can affirm that the diagonal of the matrix A^4 indicates the number of cycles of length 4 associated to every node. And that is precisely what we are looking for, cycles of length 4 of every node with the particularity that all the nodes that make up the path must be different from each other. As an example, if $A^4_{22} = 5$, we know that the number of cycles of length 4 for the node 2 is 5. From all of them, we want to find those cycles c_{2tjk2} going from $2 \to t \to j \to k \to 2$ such that $2 \neq t \neq j \neq k$.

In our case, as working with weighted adjacency matrices, it is necessary to assign a weight to each cycle, which is nothing more than the sum of the weights of each of the edges of the cycle. Once the weights of all the cycles of the node 2 have been calculated, we must calculate the maximum of them. Once obtained, we already know the intermediation nodes to which the counter we must increased (according to the process described above).

3.2. Calculating centrality model indices

Ones we have the transition matrices $T^{(i,j)}$ for all $i, j \in V$, we calculate the four indices that characterize the nodes in a network. Each index represents different properties or behaviour of each node in the network. Remark that we defined for each index a vector C_{1xn} representing the value of each index for all the nodes. This vector will be updated according to the importance of every node in the topology of the network. In this section, we explain how to calculate each index from a mathematical perspective. Next Section 3.4 describes the behaviour of a node based on these indices.

1. Interaction index (C^1). This index represents the interaction of a node in the network, that is to say, it indicates the number of times that a specific node acts as a transition one.

The interaction index may be seen as a vector with n components, in which each element represents a counter of times where the associated node is chosen as a transition node involved in a maximum probability walk. More in detail, for each matrix $T^{(i,j)}$, we select the maximum element of this matrix as shown in Eq. (6).

$$\max\left(T^{(i,j)}\right) = P_{irj}P_{jsi}, \tag{6}$$

This maximum value represents the maximum probability to go from node i to node j and return through two different transition nodes r (go random walk) and s (return random walk). Note that in previous section we describe t and k as each possible transition nodes for each 4-relationship in the network. Now, we define the chosen (maximum T_{ij}) transition nodes to connect a pair of nodes i and j. These transition nodes increment in a unit their respective counters in C^1 ($C^1(r)$ and $C^1(s)$), repeating this process for all $\max\left(T^{(i,j)}\right)$, for $i, j = 1, 2, \ldots, n$ (see Eq. (7))

$$\text{for all } \max\left(T^{(i,j)}\right) = P_{irj}P_{jsi} : \text{ then } C^1(r) = C^1(r) + 1 \text{ and } C^1(s) = C^1(s) + 1, \tag{7}$$

2. Relationship importance index (C^2). In the previous index (C^1), for each maximum value, one is added to the counter of the transition nodes (r and s), but without taking into account the value of the maximum probability itself (Eq. (6)). With the aim to analyse the importance of a node in the network, we calculate the importance of its relationships by adding this maximum value of the transition matrix. This point helps us to differentiate when a node is important in dense or sparse clusters/communities.

This index C^2 is also defined as a vector with n components associated to the n nodes in the graph. Next, we can see this formulation:

$$\text{for all } T^{(i,j)} : \; C^2(r) = C^2(r) + \max\left(T^{(i,j)}\right), \;\; C^2(s) = C^2(s) + \max\left(T^{(i,j)}\right), \tag{8}$$

where $C^2(r)$ and $C^2(s)$ are the rth and sth components of the index C^2.

Then, the *Relationship importance* index is given by updating the vector C^2, taking into account the maximum probability.

3. Individual relevance index (C^3). Following the same structure (a new null vector C^3 with n components associated to the n nodes in the graph), we define a new complementary index. In the previous indices, we assign the same value to the two involved transition nodes r and s. That is to say,

$$C^1(r) = C^1(s) \;\text{ and }\; C^2(r) = C^2(s).$$

We can further refine these indices to represent the different probabilities to go and return walks, evaluating the importance of a node in absolute terms

$$\text{for all } T^{(i,j)} : \quad \text{then} \quad C^3(r) = C^3(r) + P_{irj}, \;\; C^3(s) = C^3(s) + P_{jsi}, \tag{9}$$

where $\max\left(T^{(i,j)}\right) = P_{irj}P_{jsi}$.

As a result, we come up with an *Individual relevance* index defined by the vector C^3. This index let us differentiate if a transition node is relevant in terms of intra or inter-cluster.

4. Global relevance index (C^4). Last, we want to calculate the relative importance of a node. For that goal, the evaluation of the average probability of all the transition nodes is made. Regarding the previous index C^3, the main difference is that, in this case, we only take the involved maximum random paths into account

$$C^4(r) = \frac{\sum P_{irj}}{C^1(r)} = \frac{C^3(r)}{C^1(r)}, \tag{10}$$

for all r where $P_{irj}P_{jki} = \max\left(T^{(i,j)}\right)$.

$$C^4(s) = \frac{\sum P_{jsi}}{C^1(s)} = \frac{C^3(s)}{C^1(s)}, \tag{11}$$

for all s where $P_{itj}P_{jsi} = \max\left(T^{(i,j)}\right)$.

This index lets us know if a node is relevant as an inter-cluster or if it belongs to a dense cluster being relevant in terms of connectivity. For example, if a node m participates in 3 maximum 4-relationship walks with probabilities 0.6, 0.7 and 0.9, $C^4(m)$ will be the average of these probabilities – 0.733 –, being more relevant as inter-cluster level than other transition nodes with 20 maximum walks with an average of 0.4.

Consequently, *Global relevance* index is obtained by updating, as well in the cases before, the vector C^4. For the sake of clarity, Table 3 summarizes the meaning of the proposed indices.

All in all, the calculation of the four aforementioned indices can be summarized in the pseudo-code of the Algorithm 1.

Given a graph G with n nodes, the complexity of the algorithm is $O(n^3)$ with space $O(n^2)$. However, it is important to remark that the algorithm has been designed in a scalable way, in order to parallelize it across multiple CPUs or GPUs. Hence, it could be used in quite large networks.

3.3. Illustrative example

With the aim of understanding and clarifying the model, we calculate the centrality indices in a very simple directed network which is depicted in Fig. 1.

Table 3
Description of the centrality indices proposed in our model.

Notation	Physical significance
C^1	Is the centrality measure of the model (interaction). It represents the interaction of a node as a transition node in random walks of cycles of length 4.
C^2	Is the relationship importance measure of the model. It measures the importance of a node in its community or cluster taking into account the connectivity.
C^3	Is the individual importance measure. It states if a node is relevant as intra-cluster in dense communities or as inter-cluster in sparse ones.
C^4	Is the global importance measure. It measures the importance of the role assigned by index C^3

Require: A directed weighted graph $G = (V, E, W)$
Ensure: Four vector C^1, C^2, C^3 and C^4 with the centrality measures for all nodes
for all $i, j \in V$ **do**
  for all $t \in V$ **do**
    if $t \neq i \neq j$ **then**
      Calculate $P_{itj} = \frac{w_{it}w_{tj}}{D^+(i)D^-(j)}$ [Eq. (2)] and $V_{P_{itj}} = \begin{pmatrix} P_{i1j} & P_{i2j} & P_{i3j} & \cdots & P_{inj} \end{pmatrix}$ [Eq. (3)]
    else
      $V_{P_{itj}} = 0$
    end if
  end for
end for
for all $i, j \in V$ **do**
  $T^{(i,j)} = V^T_{P_{itj}} V_{P_{jti}}$ [Eq. (5)]
  $\text{diag}(T^{(i,j)}) = 0$
  $[r, s] = \max(T^{(i,j)})$
  $C^1(r) = C^1(r) + 1$ [Eq. (7)]
  $C^1(s) = C^1(s) + 1$
  $C^2(r) = C^2(r) + \max(T^{(i,j)})$ [Eq. (8)]
  $C^2(s) = C^2(s) + \max(T^{(i,j)})$
  $C^3(r) = C^3(r) + P_{irj}$ [Eq. (9)]
  $C^3(s) = C^3(s) + P_{jsi}$
end for
$C^4(r) = \frac{C^3(r)}{C^1(r)}$ [Eq. (10)]
$C^4(s) = \frac{C^3(s)}{C^1(s)}$ [Eq. (11)]

Algorithm 1: Directed Two-way Random Path Betweenness Centrality Algorithm.

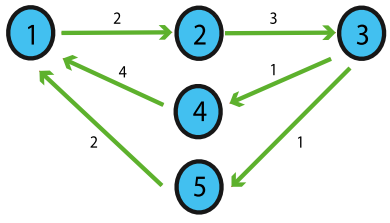


Fig. 1. Toy example: Each node is labelled with its identifier and each edge with its weight.

With the weighted matrix of the graph

$$W = \begin{pmatrix} 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 4 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \end{pmatrix},$$

we calculate the matrices

$$D^+ = W\boldsymbol{u} = \begin{pmatrix} 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 4 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 2 \\ 4 \\ 2 \end{pmatrix}$$

and

$$D^- = \boldsymbol{u}^T W = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 4 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \end{pmatrix}$$
$$= \begin{pmatrix} 6 & 2 & 3 & 1 & 1 \end{pmatrix}.$$

Remark that, D^+ is a column vector containing the sum of the weights of the out-edges of each node and D^- is a row vector with the sum of the weights of the in-edges of each node.

Taking into account that

$$P_{itj} = \frac{w_{it}w_{tj}}{D^+(i) \cdot D^-(j)},$$

is the probability to reach a destination node j from an origin node i through the transition node t and that $P_{itj} = 0$ if $i = j$, $i = t$, or $j = t$, we calculate by means Eq. (2) the vector

$$V_{P_{itj}} = \begin{pmatrix} P_{i1j} & P_{i2j} & P_{i3j} & P_{i4j} & P_{i5j} \end{pmatrix} \quad \forall i, j.$$

For instance, in the case $i = 1, j = 3$ or $i = 3, j = 1$

$$V_{P_{1t3}} = \begin{pmatrix} P_{113} & P_{123} & P_{133} & P_{143}P_{153} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \end{pmatrix},$$

$$V_{P_{3k1}} = \begin{pmatrix} P_{311} & P_{321} & P_{331} & P_{341}P_{351} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1/3 & 1/6 \end{pmatrix}.$$

With all these vectors $V_{P_{itj}}$, we compute the transition matrices using Eq. (5)

$$T^{(i,j)} = V^T_{P_{itj}} V_{P_{jki}}.$$

Then, in this toy example we have

$$T^{(1,3)} = V^T_{P_{1t3}} V_{P_{3k1}} = \begin{pmatrix} P_{113} \\ P_{123} \\ P_{133} \\ P_{143} \\ P_{153} \end{pmatrix} \begin{pmatrix} P_{311} & P_{321} & P_{331} & P_{341} & P_{351} \end{pmatrix}$$
$$= \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 1/3 & 1/6 \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/3 & 1/6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

$$T^{(3,1)} = V^T_{P_{3t1}} V_{P_{1k3}} = \begin{pmatrix} P_{311} \\ P_{321} \\ P_{331} \\ P_{341} \\ P_{351} \end{pmatrix} \begin{pmatrix} P_{113} & P_{123} & P_{133} & P_{143} & P_{153} \end{pmatrix}$$
$$= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1/3 \\ 1/6 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1/3 & 0 & 0 & 0 \\ 0 & 1/6 & 0 & 0 & 0 \end{pmatrix} = (T^{(1,3)})^T.$$

$$T^{(2,4)} = V_{P_{2i4}}^T V_{P_{4k2}} = \begin{pmatrix} P_{214} \\ P_{224} \\ P_{234} \\ P_{244} \\ P_{254} \end{pmatrix} \begin{pmatrix} P_{412} & P_{422} & P_{432} & P_{442} & P_{452} \end{pmatrix} =$$

$$= \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = (T^{(4,2)})^T.$$

$$T^{(2,5)} = V_{P_{2i5}}^T V_{P_{5k2}} = \begin{pmatrix} P_{215} \\ P_{225} \\ P_{235} \\ P_{245} \\ P_{255} \end{pmatrix} \begin{pmatrix} P_{512} & P_{522} & P_{532} & P_{542} & P_{552} \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = (T^{(5,2)})^T.$$

Note that, in this example,

$$T^{(1,3)} = (T^{(3,1)})^T, T^{(2,4)} = (T^{(4,2)})^T \text{ and } T^{(2,5)} = (T^{(5,2)})^T.$$

This is due to the simplicity of the network but, in fact, in most real networks this is not always the case.

As can be easily verified, the other transition matrices are $\mathbf{0}_{5\times5}$

$$T^{(1,2)} = T^{(1,4)} = T^{(1,5)} = T^{(2,3)} = T^{(3,4)} = T^{(3,5)} = T^{(4,5)} = \cdots = \mathbf{0}_{5\times5}.$$

Now, we can calculate the four indices.

- The *Interaction index* C^1 is calculated by means Eq. (7)

  $$\forall \max(T^{(i,j)}) = P_{irj}P_{jsi} : \; C^1(r) = C^1(r) + 1 \text{ and } C^1(s) = C^1(s) + 1.$$

  For instance, the maximum value for the transition matrix $T^{(1,3)}$ is 1/3 which corresponds to the position row 2 column 4. This means that we have to add one to the counters of nodes 2 and 4. Then, using only this transition matrix, we have $C^1 = (0, 1, 0, 1, 0)$. Repeating these steps for all non-zero transition matrices, we obtain

  $$C^1 = (4, 2, 4, 2, 0).$$

- The *Relationship importance index* C^2 is calculated by means Eq. (8)

  $$\forall T^{(i,j)} : \; C^2(r) = C^2(r) + \max(T^{(i,j)}), \;\; C^2(s) = C^2(s) + \max(T^{(i,j)}).$$

  In this case, is that the maximum value is added to the counters of the nodes and not one. For instance, the maximum value for the transition matrix $T^{(1,3)}$ is 1/3 which corresponds to the position row 2 column 4. Then, we add 1/3 to the counters of both nodes 2 and 4. Then, using only this transition matrix, we have the index $C^2 = (0, 1/3, 0, 1/3, 0)$.
  Repeating these steps for all non-zero transition matrices, the second index is obtained

  $$C^2 = (4, 2/3, 4, 2/3, 0).$$

- The *Individual relevance index* C^3 is calculated by means Eq. (9)

  $$\forall T^{(i,j)} : \; C^3(r) = C^3(r) + P_{irj}, \;\; C^3(s) = C^3(s) + P_{jsi},$$

  where $\max(T^{(i,j)}) = P_{irj}P_{jsi}$. This index takes into account the different probabilities to go and return. For instance, the maximum value for the transition matrix $T^{(1,3)}$ is 1/3 which corresponds to product $T^{(1,3)} = P_{123}P_{341} = 1 \cdot 1/3$. Then, we add 1 to the counter of node 2 and 1/3 to the counter of the node 4. Using only this transition matrix the index is $C^3 = (0, 1, 0, 1/3, 0)$.
  Repeating these steps for all non-zero transition matrices, we obtain

  $$C^3 = (4, 2, 4, 2/3, 0).$$

- The *Global relevance index* C^4 is calculated by means Eqs. (10) and (11)

  $$C^4(r) = \frac{\sum P_{irj}}{C^1(r)} = \frac{C^3(r)}{C^1(r)} \text{ and } C^4(s) = \frac{\sum P_{jsi}}{C^1(s)} = \frac{C^3(s)}{C^1(s)}.$$

  This index is the easiest to calculate if you have previously C^1 and C^3. In this example, the *Global relevance index* is

  $$C^4 = (4/4, 2/2, 4/4, (2/3)/2, 0) = (1, 1, 1, 1/3, 0).$$

3.4. Describing the behaviour of a node by means of the proposed indices

Our proposal is based on four indices that characterize a node playing the betweenness/intermediate role in dense networks. For that, we study the behaviour of each node as a transition node in small communities of 4 nodes.

Firstly, we evaluate how many maximum random paths involved a node as a transition node. That means to count the times where a node participates in the most probability 2-step random path (go or return walks). This approach may be similar to the classical betweenness, but we evaluate this intermediate role in a 4-relationship context to explore the characterize the nodes in different density communities. For instance, we find the maximum probability to go and return from a node i to a node j through two transition nodes (t and k). Both nodes play the role of transition nodes, and we measure this fact in the first index **Interaction index** C^1. A node with a high value of this index indicates the belonging to a **dense cluster** or community, and its relevance as a connectivity hub in the cluster. In contrast, a low value of this index means an irrelevance in terms of connectivity or belonging to a **sparse cluster**.

Furthermore, it is interesting to know the probability of the selection of a node as a transition node. For that, in **Relationship Importance index** C^2, we sum all the go and return probabilities which the node is evolved in the maximum 4-relationships. This index denotes the importance of a node in terms of **community/connectivity**. A high value means that the different pairs of nodes i, j connected through the transition nodes t, k have weak or low connections (low in and out degrees) but the links with the transition nodes are strong. For this reason, these nodes will be very important as transition nodes for their cluster or community.

The main weakness of the previous indices is the fact that they provide the same treatment to both the go and return nodes ($C^1(t) = C^1(k)$ and $C^2(t) = C^2(k)$). Given the **Individual Relevance index**, C^3, we want to determine the relative importance of each transition node separately, into each 4-node relationship, measuring the probability of go paths for node t and return paths for node k. Thus, we can

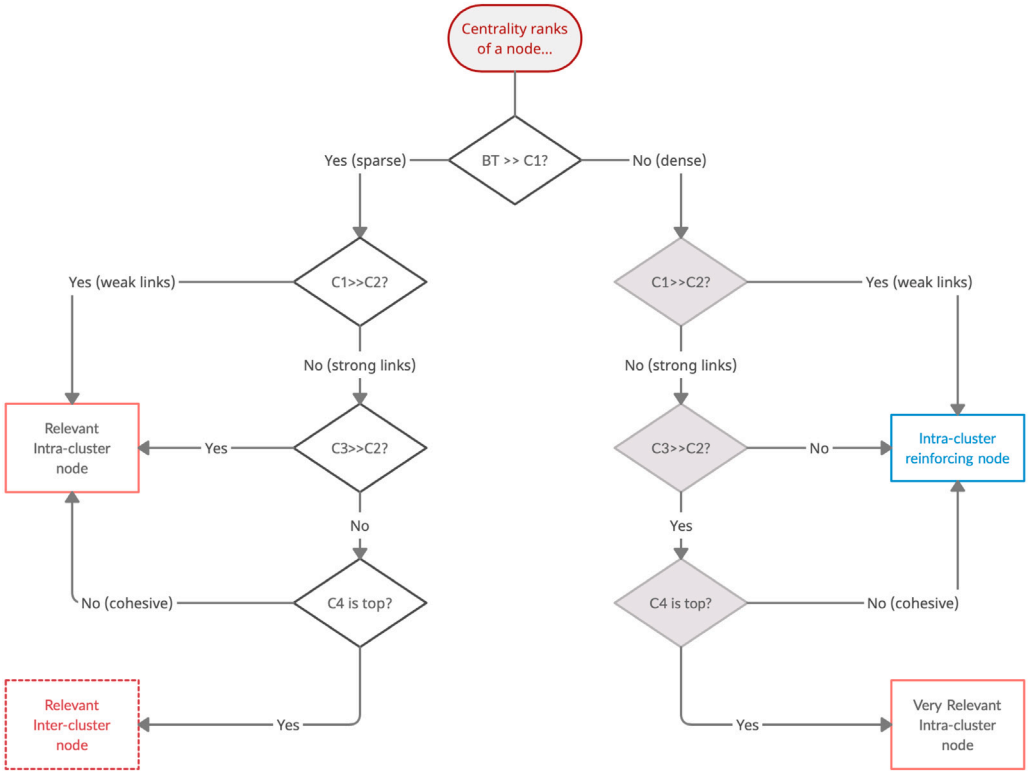


Fig. 2. Behaviour flowchart: analysis of the behaviour of a node in a network regarding the four indices. The comparative symbols ('≫' and '≪') mean that one ranking value is much better than the other, BT stands for Betweenness centrality, a grey decision box means 'in a dense context' and a white decision box means 'in a sparse context'.

differentiate the individual importance of a transition node in a dense context (intra-cluster), or the importance as inter-cluster node in sparse communities.

Finally, we want to distinguish between a transition node working as a hub (several random paths are better crossing by this node) independent of their probabilities, and a transition node involved in few random paths but stronger than others. In both cases, the transition nodes are relevant, but the characteristics of their clusters/communities are different. For that, in the **Global Relevance index** C^4, we calculate the sum of the probabilities of random paths where a transition node is involved in obtaining the average of these cases.

All these indices are complementary to other betweenness measures, since they could be used to analyse the behaviour of the nodes within a network by the comparison of the different indices (and Betweenness centrality) in terms of ranks.

For the sake of clarity, we summarize this analysis by a behaviour flowchart in Fig. 2. For instance, if we follow this flowchart by considering the indices obtained by the nodes in the toy sample described in Section 3.4, we can see that nodes 1, 2 and 3 are considered by our algorithm as relevant intra-cluster nodes, with weak links in a dense cluster. This is because the answer to the first question in the flowchart ($BT \gg C^1$?) would be *yes* for the three nodes and, next, the answer to the second question according to the flowchart ($C_1 \gg C^2$?) would be, again, *yes*. In addition to that, nodes 4 and 5 play a different role in the network according to their weights, node 4 is a relevant intra-cluster node whose main purpose is to give cohesion to its cluster (flowchart branch $BT \gg C^1 \leq C^2$ and C^4 is not top in ranking), and node 5 helps to reinforce the network (flowchart branch $BT \ll C^1 \gg C^2$).

4. Evaluation study

In this section, we evaluate our proposal by applying the four aforementioned indices in three different real networks. First, a network of relationships of the students in a residence hall is analysed. The second network (new proposal) describes the relationships of the components that form the existing vitreous-metal elements and the third one a nation-wide human mobility network in Spain created from real data.

4.1. A directed network sample: Residence hall

In this section we analyse the ANU Residence Hall Network [27], that represents the friendship between students living in a residence hall in Australia. If the student (or resident) i considers the student j as a friend, there is a link, and the weight of this link marks the level of friendship on a scale: 5 (best friend), 4 (close friend), 3 (friend), etc.

This network is represented by a directed graph $G = (V, E, W)$, where nodes (V) are the residents (217 nodes), links (E) are the friendship relationships (2672 edges), and W are the weighted matrix of friendship values.

After applying our method, we can characterize the behaviour of the nodes in the network. We start the study by comparing the Classical Betweenness (BT) with our Interaction index C^1, showing the results in Fig. 3. The top figure shows the Normalized Betweenness Centrality of all nodes, comparing BT and C^1. In this case, the correlation between both measures is 79.21%. The bottom figure represents a similar comparison but in ranking (obviously, a node has a better ranking close to $y = 0$ axis).

If we focus on the nodes with more variability between both metrics we can verify that the most relevant difference between these indices is the number of nodes implied in a betweenness relationship because there are three students or nodes in BT and four in C^1. Table 4 shows the six nodes whose differences (BT-C^1) are high. From this table we can observe that the first three nodes are relevant nodes in BT but not in C^1 and the remaining three nodes are relevant in C^1 in contrast to BT.

Then, we analyse the behaviour of these nodes in the network by the comparison of proposed indices in terms of ranks. For the sake

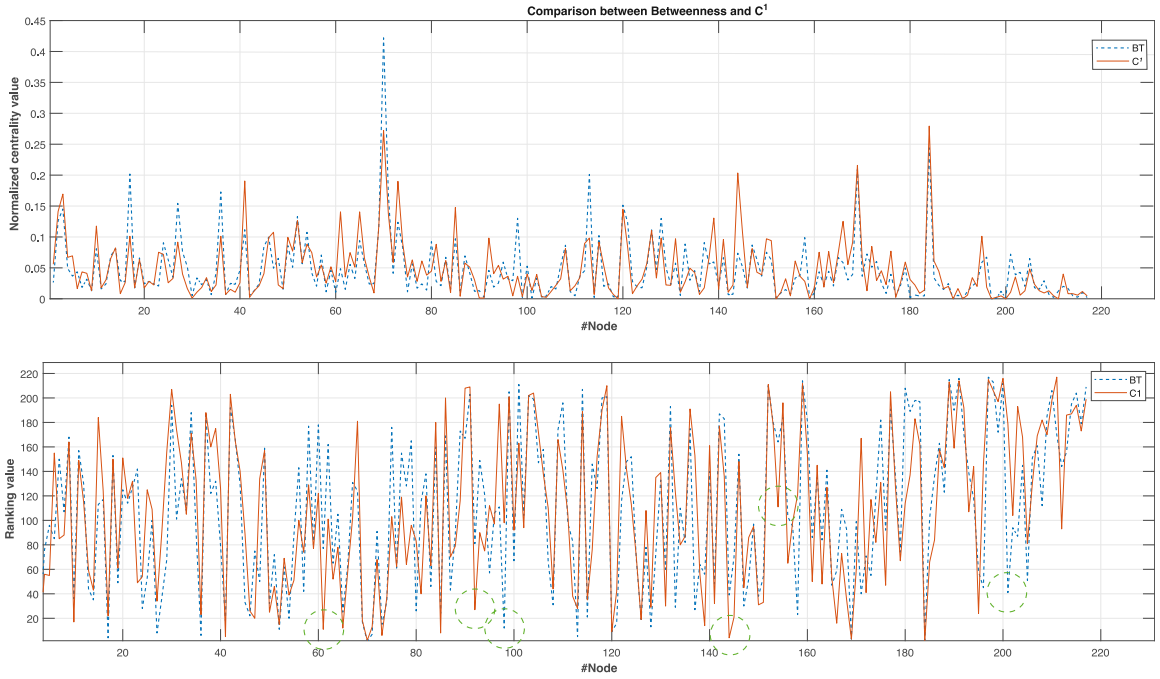


Fig. 3. Differences in value and rankings between classical betweenness centrality measure and Interaction index (C^1). In green dotted circle is marked the nodes with higher differences between both rankings (see Table 4).

of clarity, we show an example of the node 97, using the **behaviour flowchart** shown previously (Fig. 2).

Node 97. This node has a good ranking in classical betweenness (12th), but measuring the interaction in our approach (C^1), its ranking gets off to the 98 position (see Table 4). This degradation in ranking suppose that this node plays the role of transition node in a simple context (BT), but there are not enough return paths or they are weak (C^1). This indicates that the node 97 belongs to a sparse cluster or community. To know the behaviour of this node in a sparse context, we attend to the remaining measures of our approach (see Table 4). According to the **Relationship importance index** (C^2), we observe that node 97 has medium importance in the network: (i) in-degree: 20 nodes consider it 'friend' at different levels, (ii) out-degree: this node considers 13 nodes as a 'friend' (friendship average of 3.67), and (iii) friendship values or weights: in-degree friendship average of 2.95, and 3.67 in out-degree. Moreover, the C^2 ranking of the node is better than in C^1. This fact points to a global strong link in the cluster (the average friendship of the residents is high into the cluster).

Furthermore, if we analyse the Individual Relevance index (C^3) of the node 97 as a transition node in the network, we observe an important degradation in the ranking (153th out of 217 nodes). This fact could be interpreted as follows:

- The node is evolved in a significant number of 4-relationships into the network, but its role as a transition node is worse than others. Its role is relevant to densifying the connectivity into a sparse cluster.
- The node has few links, but their importance is significant. In this case, the node is relevant with respect to the inter-cluster connectivity.

Finally, we observe the Global Relevance index (C^4) to clarify these assumptions. This value indicates if the reduced individual relevance value of node 97 is motivated by a high number of random paths involved (C^1) with respect to the relevance of its participation in these relationships. As the node 97 is ranked as 172th, we could confirm the role of node 97 as a relevant inter-cluster node in the network.

To conclude, node 97 belongs to a sparse community or cluster, which links are strong. However, its links are weak in the cluster, but this relevance is high because it plays the role of a inter-cluster transition node.

Table 4
Rankings of different indices of 6 nodes with higher difference BT-C^1.

Node	BT	C^1	C^2	C^3	C^4
97	12	98	84	153	172
157	23	121	151	169	175
201	41	179	186	191	193
143	39	4	23	2	90
61	77	11	55	30	129
91	81	27	40	13	51

In order to evaluate the connectivity of the network and validate the presented model, we are going to measure the connectivity of the dataset by means of the metric introduced in [28].

$$Q = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} SP_{ij}}{N(N-1)} \quad \forall\, i \neq j, \tag{12}$$

where N is the number of nodes and SP_{ij} is the shortest distance between node i and node j. Moreover, we count the number of walks that are not possible to complete if remove a specific node.

To carry out this task, we apply Eq. (12) to the original graph and an alternative version without node 97. The connectivity Q of this second case was 8.3804, in contrast to the original connectivity including the node (8.3365). This means that removing this node makes the average of shortest paths in the network to increase, highlighting the importance of its links in a sparse cluster. Moreover, the number of unreachable walks increases with respect to the original network (+66.4%), pointing out its important role as inter-cluster node.

4.2. Vitreous metals dataset

Metallic glasses are non-crystalline materials made up of pure metals or combinations of metals and metalloids [29]. Non-crystalline solid alloys have an atomic arrangement inherited directly from the liquid state. They stand out for their particular mechanical properties: they can be up to three times harder than steels, more elastic than ceramic materials and are less fragile than transparent oxide-based glasses.

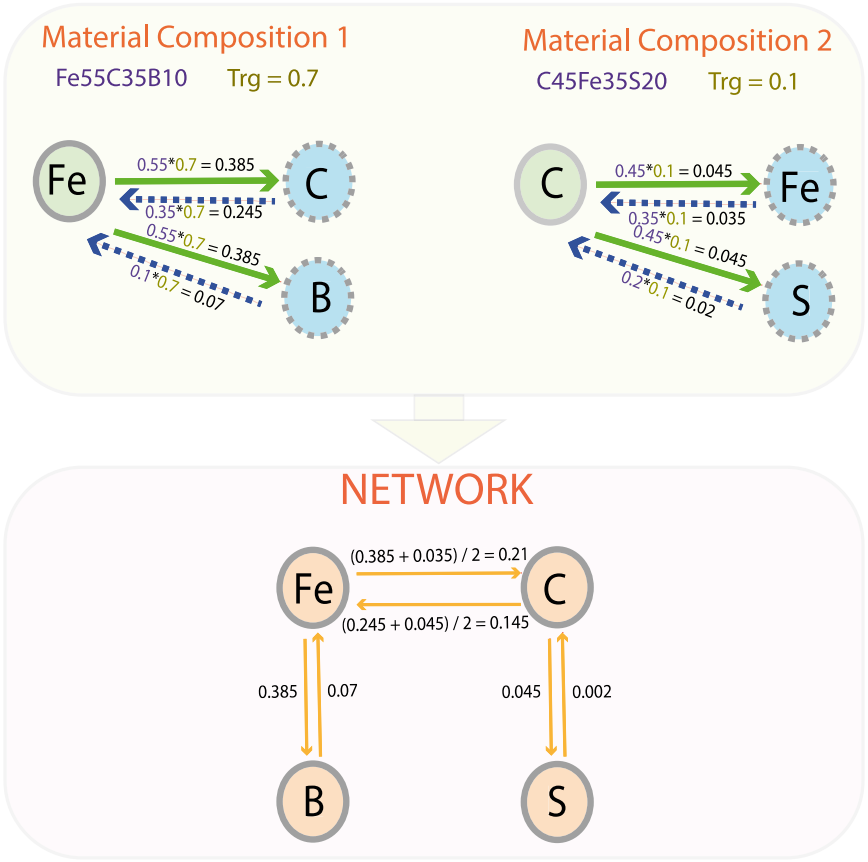


Fig. 4. Toy example representing a network based in two materials.

Nowadays, combinations of elements that make up vitreous metals continue to be discovered [30] and, because of this, the databases uses in current researches on the formation of vitreous metals range from 100 to 500 elements [31,32]. Nevertheless, in this paper, we are going to use a database of vitreous metals with the maximum number of these materials discovered so far (586), formed by 45 different chemical elements. By applying the proposed centrality model, we study the atomic formation capacity of these materials. This is done from the point of view of the connections that each chemical element, which makes them up, among the rest of the known vitreous metals, using the glass transition temperature parameter (Trg) as the relationship factor. The choice of this factor is due to the fact that it is considered, in many different studies, as one of the most influential parameters in the formation of vitreous metals [33,34].

We form a network by a directed weighted graph $G = (V, E, W)$ from a dataset with 586 different materials. These materials are formed by 45 different chemical elements, which comprise the nodes of the network (V), and the set of links (E) represent the relationship between them forming materials (219 edges). Finally, W is a ratio of the Trg parameter of the elements based on the materials they make up. For the sake of clarity, we show in Fig. 4 an example of how the directed network is represented by two materials. The numbers in the material label represent the amount of each element in the material, for instance, $Fe55C35B10$ has 55% of Iron, 35% of Carbon and 10% of Boron. Green nodes are the main element of its material composition, and green arrows are the link from the main element of the material (biggest amount in percentage) to the remaining elements. Blue elements (dotted lines) are the remaining elements. The weight of each link is calculated by the average ratio of Trg and the amount (in percentage) of the elements in all relationships between them (see Eq. (13)) (see Fig. 5).

The studied network represents the connections of the elements that make up most of the known vitreous metals by the relationship between the elements through the characteristic parameter of the reduced transition temperature.

$$W_{ij} = \frac{\sum_{k=1}^{N} Trg_k R_k^i}{N}, \tag{13}$$

where N is the number of materials where elements i and j are involved, Trg_k is the Trg of the material k and R_k^i is the percentage of the element i in that material k.

Results. The study has focused on the comparison between the classical Betweenness centrality measure (BT) and our D2RWBT algorithm. Again, we analyse the rankings of our proposed metrics (see Table 5) and we explain the following properties of this network,

Elements that form sparse networks are elements that usually participate in compounds with many elements, from highest to lowest importance: Silicium (Si), Gold (Au), Copper (Cu), Gallium (Ga), Praseodymium (Pr), Erbium (Er), Iron (Fe), Palladium (Pd), Silver (Ag), Strontium (Sr), Neodymium (Nd) and Nickel (Ni). On the contrary, the elements that form dense networks tend to participate in compounds with many few elements from highest to lowest importance: Hafnium (Hf), Cerium (Ce), Beryllium (Be), Samarium (Sm), Scandium (Sc), Platinum (Pt), Gadolinium (Gd), Phosphorus (P), Cobalt (Co), Aluminium (Al), Lanthanum (La), Yttrium (Y), Calcium (Ca), Titanium (Ti) and Zirconium (Zr).

Among the elements that make up dense networks, those that have weak connections and are relevant to reinforce the compound are the following, from highest to lowest in order of relevance: Y, Pt, Gd, Al, La, Ce, Zr and Hf. Besides, the elements that have strong connections are the following, in relevance order: Co and Ca. Finally, in the analysis of dense networks, we should remark the results of important elements

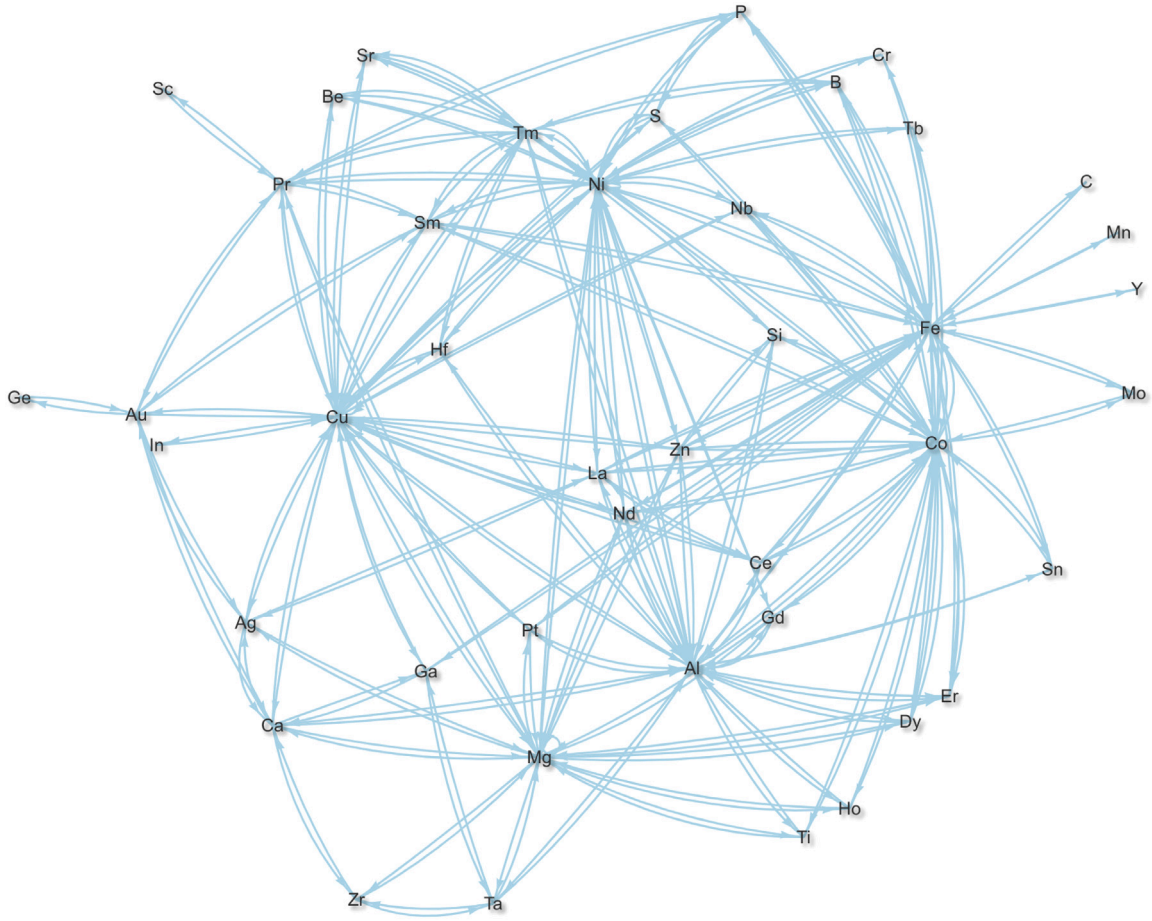


Fig. 5. Materials directed network. Nodes represent different elements and edges means that two elements belong to some vitreous metal.

Table 5
Elements with more difference BT-C^1 in rankings.

Node	BT	C^1	C^2	C^3	C^4
Si	11	23	20	20	8
Pr	21	24	28	24	23
Fe	1	3	2	4	6
Ca	12	10	7	8	5
Al	4	1	4	1	3
Co	5	2	1	3	4
Ni	3	4	3	2	2
Mg	6	6	6	6	12
Cu	2	5	5	5	7
Ag	16	18	16	17	18
Er	22	25	27	28	27
Pd	7	9	11	11	13
Nd	19	21	23	27	28
Sr	17	19	10	9	1
Y	14	11	17	15	20
Pt	20	15	18	19	21
Gd	27	22	25	25	25
La	15	12	15	14	17
Ce	25	17	19	18	19
Zr	9	7	9	7	11
Hf	33	22	21	23	26
Be	34	27	22	21	14
Sc	31	26	24	22	22

in order to connect with other less common elements than the typical ones with which it is usually composed: Be y Sc.

With respect to the elements that make up the sparse networks, those that have weak connections and are relevant to reinforce the compound are the following, from highest to lowest in order of relevance: Pr, Er, Pd and Nd. On the other hand, the elements that have strong connections are, according to relevance: Sr and Ni. Finally, in the analysis of sparse networks, it is worth noting the results of important elements in order to connect with other less common elements than the typical ones with which it usually composes: Fe and Ag.

Discussion. Despite the great potential of metallic glasses, the researchers who make them have been hampered by a shortage of basic-science knowledge of these materials. Currently, combinations of elements that make up vitreous metals with interesting characteristics are still being discovered, such as: (Cu–Zr–Ti) [35], (Ca–Mg–Cu) [36], (Ca–Mg–Zn) [37], (Ni–Nb–Sn) [38], (Ni–Nb–Ta) [39], (Ti–Cu–Ni) [40] and (Ti–Cu–Ni–Mo–Fe) [41].

The search for a parameter or criterion that indicates the high Atom Formation Capacity (CFA) continues to be of great interest and reason for scientific efforts nowadays. Some simple parameters have been suggested, based on physical properties and characteristic temperatures. The best known is the reduced glass transition temperature Trg proposed by Turnbull [42], which is defined as Trg (Tg/Tl), where Tg and Tl are the glass transition temperature and the liquid temperature, respectively. Based on the kinetics of crystal nucleation and the viscosity of the liquid, when a liquid alloy is cooled to Tg the viscosity increases rapidly and glass is formed. The tendency to glass formation will be greater the higher the Trg values are [42].

From the analysed results, the top-5 ranking in BT are the main elements that make up most vitreous metals, more specifically the so-called Bulk Metallic Glass (BMGs). The main elements extracted from the measurement are those that currently connect the most glass-forming elements [43]. In other words, they are not only present in most of the glass metals analysed but also show a great connection

with the rest of the elements analysed. Besides, we could show the importance of five elements, Fe, Ni, Cu, Al and Co. Furthermore, it is noteworthy that in the ranking of C^1, these five elements also are the 5-top positions in the ranking, but not in the same order of importance. These elements, which form both dense and sparse clusters, are the most important elements within their materials. Moreover, these elements are connected with other less common than the typical ones with which vitreous metals are usually formed (they could form new materials). Elements such as Al and Co, by forming dense networks, and Ni, by forming sparse networks, will be very important in the cohesion of the compounds in which they participate (they will have intra-cluster relevance). On the other hand, elements such as Fe form sparse networks indicating their relevance in the connections between families, at the inter-cluster level.

If we apply Eq. (12), we obtain that the connectivity value of the network is 0.3305. If we remove the first element of the ranking obtained based on C^1 (Al) we got a value of 0.352. Likewise, if we remove the second one (Co) the connectivity was 0.3755. Taking into account that a higher value means worse connectivity, we conclude that the sum of all the shortest paths is higher than the original network. The impact on the vitreous-metal network is two: on the one hand, the intra-cluster elements, Al and Co, can be extracted from the components and the remaining elements would remain connected; on the other hand, the average of Trg will be increased because of the shortest path is higher than the original network. On the contrary, if we remove the elements Fe and Cu, the connectivity values are reduced (0.3259 and 0.3291 respectively) but the number of disconnected walks is too high (252 removing Fe and 86 removing Cu). This fact reinforces the idea that these elements have an important role as inter-cluster nodes in the network, as it has been shown in previous results of our centrality model.

4.3. Human mobility dataset

The third network to test our approach has been generated from the nation-wide human mobility study released by the Spanish Ministry of Transportation (SMT) in December 2020.[2] This study relies on a spatial tessellation of the Spanish territory comprising 3216 regions (hereby *Mobility Areas, MA*). Each region usually corresponds to a particular urban area. The study indicates the number of human displacements between each pair of MAs on an hourly basis. To collect such trips, the study made use of the Call Detail Records (CDRs) from 13 million users of an unspecified Spanish mobile-phone carrier during a 3-month period (from April 1, 2020 to June 20, 2020). In its raw form, the dataset comprises 830,450,300 trips among MAs. Fig. 6a shows the distribution of trips in the dataset.

Given this dataset, we composed a network as a directed graph $G = (V, E, W)$, where node set V comprises the 3216 MAs, E is the set of links that represent the trip connections among each pair of areas and the set of edge weights W labels each edge with the total number of human trips from the origin to the destination MA for the whole period of study.

From the aforementioned graph, we used the D2RWBT algorithm to analyse the relevance of each Spanish MA regarding its connectivity to other regions. This connectivity is based on the number of human trips among all the areas. This way, we compared the Classical Betweenness (BT) with the six indexed extracted by our algorithm.

Since we detected a high variability between both metrics, we have chosen 6 particular nodes (out of 3216) to better explain the actual differences between BT and our approach. These nodes are included in Table 6.

Table 6
Some nodes with highlighted properties.

Node	BT	C^1	C^2	C^3	C^4
Málaga (MLG)	341	9	139	25	183
Balearic Island towns (BIT)	777	17	6	272	463
Palma de Mallorca (PLM)	753	3	2	3	105
Cáceres (CAC)	290	706	371	99	4
Soria (SOR)	431	621	26	613	613
Argamasilla and Puerto Llano (APL)	94	664	614	207	21

Nodes MLG, BIT and PLM: These 3 MAs belong to a dense cluster or community (as their high values in the C^1 index shows in contrast to BT). Concerning the Relationship Importance index (C^2), we observe that the node MLG (comprising the city of Malaga in the south of Spain) has weak connections to its closer MAs (C^1 is higher than C^2 rank). In other words, a low number of human trips traverse this node. This means that this node plays the role of reinforcing an intra-cluster (useful to connect different parts within a specific area). This makes sense as the city of Malaga, one of the largest urban areas in the south of Spain (yellow region in Fig. 6b). For that reason, it is reasonable to think that the inner transportation flows within this area is quite large.

Regarding the BIT node, results show that it reinforces the cluster or community too but its intra-connectivity is higher than node MLG (stronger links mean a higher number of trips involving this node as an intermediate stop). In that sense, the BIT node comprises different towns in the Balearic Islands which are one of the most important touristic landmarks of Spain (black region in Fig. 6b). Hence, these areas had a very intense human mobility during the period of study, from April to June, as it covered part of the summer season. This explains the high intra-mobility detected by the proposed metrics.

Last, the PLM area plays the same role than previous nodes but with a high C^4 rank which means a higher relevance in its cluster/community. This node represents the city of Palma de Mallorca, the capital of the Balearic Islands (grey region in Fig. 6b). This justifies the aforementioned relevance of the node.

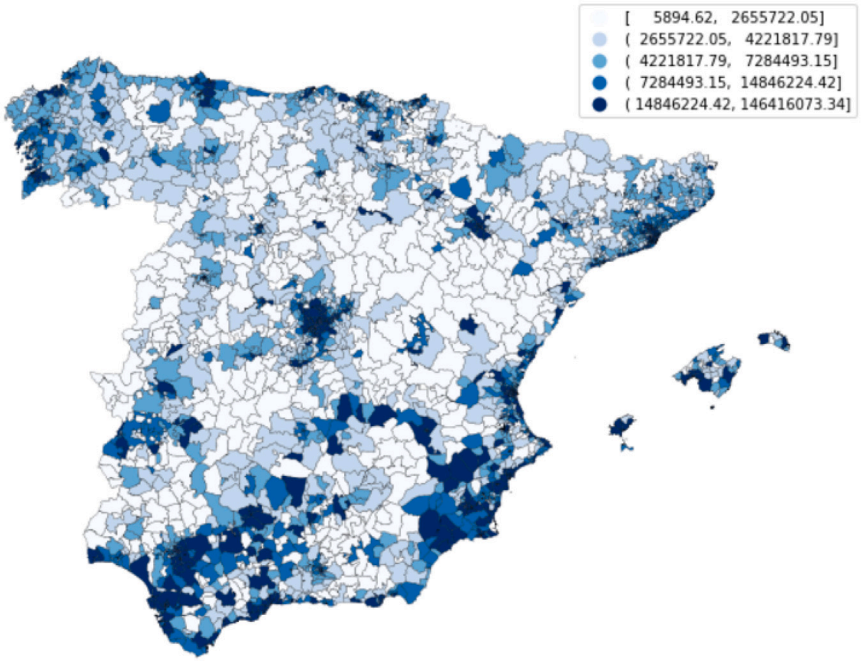
Nodes CAC, SOR and APL: These nodes belong to a sparse cluster or community as the fact that their BT is higher than their C^1 shows. This is justified by the fact that these three nodes cover Spanish areas with a very low population density. More in detail, if we observe the Relationship Importance index (C^2), we see that SOR is linking to its neighbouring with weak links (C^1 is higher than C^2 rank). This means that the human mobility around this node is low. Again, this can be explained by the fact that this node represents the city Soria (green region in Fig. 6b), a small city in the centre of Spain which is suffering the effects of severe depopulation.[3]

Moreover, the node CAC is classified as a relevant intra-cluster node because it reinforces the cluster or community and its intra-connectivity has stronger links (a higher number of trips use this node as an intermediate stop). This reflects the fact that Caceres city (the urban area covered by the node) is an important intermediate stop of the trips moving between Spain and Portugal (red region in Fig. 6b).

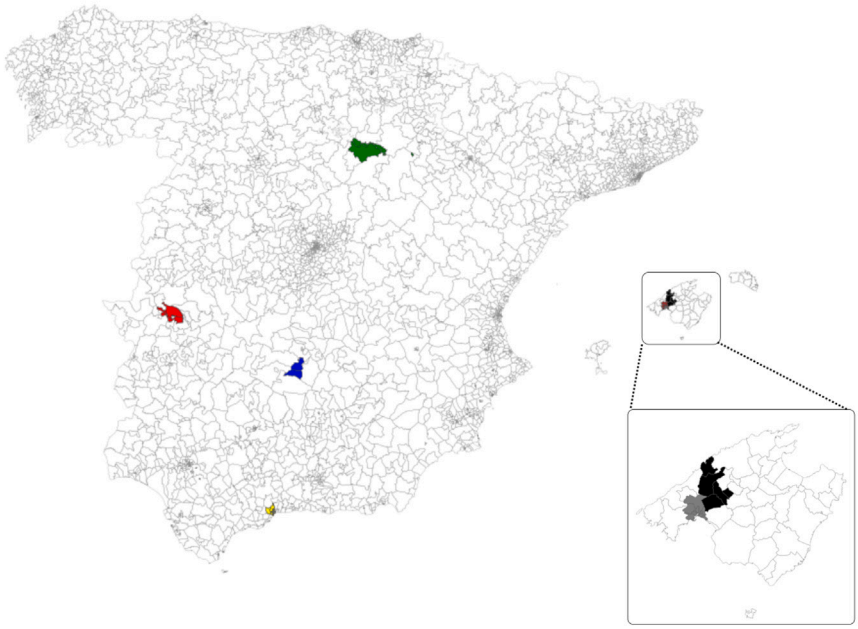
Finally, the node APL is a relevant inter-cluster node (high rankings in C^3 and C^4, individual and global relevance). As result, this node is very relevant to connect a sparse area of mobility with other areas, and plays the same role than previous nodes but a high C^4 rank means an increasing relevance in the cluster/community. This node spatially covers two nearby Spanish towns, Argamasilla and Puerto Llano, both in the centre of Spain (blue region in Fig. 6b). This region is not crossed by any important transportation infrastructure so it justifies its sparse mobility. Its relevance comes from the fact that both towns are two of the most important ones in terms of population in the region.

[2] https://www.mitma.es/ministerio/covid-19/evolucion-movilidad-big-data/opendata-movilidad

[3] https://www.espon.eu/soria



(a) Geographical distribution of the sheer number of outgoing trips of the Spanish MAs during the entire period of study according to the SMT study.



(b) MAs that were identified by the algorithm. Node CAC is shown in red, SOR in green, APL in blue, MLG in yellow, BIT in black and PLM in grey.

Fig. 6. Maps showing the details of the human mobility dataset.

Table 7
Correlations between centrality indices in ranking in the human mobility dataset.

	C^1	C^2	C^3	C^4
C^1	–	88.12%	88.25%	77.1%
C^2	88.12%	–	89.11%	83.02%
C^3	88.25%	89.11%	–	96.82%
C^4	77.1%	83.02%	96.82%	–

As in Sections 4.1 and 4.2, we validated the model by means of testing the connectivity (see Eq. (12)). Following the same strategy of removing the nodes (Málaga, Balearic Island towns, Palma de Mallorca, Cáceres, Soria, Argamasilla and Puerto Llano), the resulting values are 4813.8, 4813.8, 4816.2, 4814.9, 4811.2, and 4816.3 at the same order of the Table 6. The original network has $Q = 4816.1$. The variability of the connectivity between different nodes is low because removing a unique node in a high dense network does not affect excessively. However, we can observe a pattern that verifies our model, as the best Q found is Palma de Mallorca (with 0.1% of unreachable paths less) and Soria is the worse (with 0.1% of unreachable paths more).

To conclude, according to the correlations between our metrics in this network (see Table 7), we can see that the C^3 and C^4 indices have

Table 8
Comparison of the top-10 nodes in the 3 real-world networks. C^1 is our interaction index, BT is the classical betweenness centrality, DP is dominant path effective distance gravity model (DP_{EffG}) and RW is random-walk effective distance gravity model (RW_{EffG}).

#Rank	Residence hall				Vitreous metals				Human mobility			
	C^1	BT	DP	RW	C^1	BT	DP	RW	C^1	BT	DP	RW
1	184	70	70	7	2	14	14	14	367	355	367	491
2	70	184	169	70	9	11	22	22	115	367	391	236
3	169	169	184	169	14	27	11	11	33	491	267	33
4	144	17	120	184	27	2	45	27	267	731	731	739
5	41	113	128	120	11	9	27	45	324	183	393	367
6	73	36	65	128	22	22	41	41	183	773	33	115
7	3	71	131	65	45	30	7	9	731	236	491	267
8	85	27	166	131	41	3	21	7	739	115	283	717
9	120	120	135	166	30	45	26	21	717	391	637	331
10	2	3	113	135	7	41	9	26	773	739	115	355

a very high correlation factor, 96.82%. This reflects that almost all the nodes in this mobility study have a similar individual and global relevance within the network (there are not very relevant intra-cluster nodes with respect to the other nodes into their clusters).

4.4. Validation with other approaches

In order to validate our method with different networks, we have compared our centrality measure C^1 with the classical betweenness centrality and an adaptation of the effective distance gravity model [15]. This second model is based on the effective distance which takes into account the local topology of the network and the flows across edges [44].

There are different approaches of computing this distance, focusing on two specific measures that allow us to validate our proposed method, the dominant path effective distance (DPED) [45] and the random-walk effective distance (RWED) [46].

The Effective distance Gravity model (EffG) computes a centrality score (Eq. (14)) for each node on the basis of the interaction scores between all pair of nodes in terms of their degrees and effective distances,

$$C_{EffG} = \sum_{j=1, j\neq i}^{N} \frac{D^+(i)D^-(j)}{ED^2_{j|i}}, \tag{14}$$

where $ED_{j|i}$ is the effective distance from node i to node j.

We have compared our method with the classical betweenness centrality and the EffG centrality scores given two different effective distances, namely DPED (DP_{EffG}) and RWED (RW_{EffG}).

As Table 8 shows, the number of consistent nodes between our method and the other methods is high. Given this high number of coincidences with other approaches, it would be possible to validate our proposal as a consistent centrality measure. Moreover, this evaluation complements the main goal of this paper, the development of a directed centrality model based on a random-path betweenness centrality (C^1) and a palette of indices that measure the importance of a node as an intermediate link and describe some behavioural properties.

5. Conclusions

In this paper we propose the *Directed Two-way Random Path Betweenness Centrality Algorithm (D2RWBT)* a new centrality model for directed networks. By means of this model we obtain a ranking of the nodes in directed networks to describe their relevance within the network as transition nodes. More in detail, the model describes a node by means of four indices that provide information about the density of its cluster/community (dense or sparse), the strength of its connections, the relative and absolute importance in the network, or the relevance as intra or inter-cluster node.

We can briefly describe of the four indices as follows, (i) the first centrality index (C^1) measures the interaction of a node (as a transition node) in random walks of cycles of length 4 in directed networks, (ii) the second one, the Relationship Importance Index, measures the importance of a node in its community in terms of connectivity, (iii) the Individual Relevance Index states whether a node is relevant as intra or inter-clusters in dense or sparse communities, and finally, (iv) the Global Relevance Index measures the importance of the previous assigned role into its cluster. In that sense, we have compared these indices with other measures, obtaining quite consistent results.

This model lets us understand the behaviour of a node into a cluster or community in a dense network. For instance, we studied and analysed its practical implementation in 3 real dense networks, (i) a friendship network where we could explain the relevance of a student in the network (introvert or extrovert student, the force of his friendship links with other students, etc.), (ii) a new vitreous-metal network that lets us detect the most relevant elements in a materials dataset that may help to find new materials, and (iii) a human-mobility network that could explain the mobility of the people between different cities in a specific time period. The experimental results help us to justify our model: the first network, related to a residence hall, was used as a sample to explain the extraction of conclusions using the algorithm flowchart. Given the second network, that represents the glass transition temperature of known vitreous metals, the most relevant elements to the cohesion of compounds in dense (Be and Sc) and sparse (Ag and Fe) combinations, and the most important element as inter-cluster node (Fe) were obtained. This measure can be interesting to create new vitreous metals. Finally, we analysed the mobility of a medium-size country like Spain. In this setting, the proposed algorithm was able to identify important transition nodes in very different spatial areas covering both touristic landmarks and depopulated regions. This knowledge could be instrumental to better define nation-wide transportation policies.

As future work, we would consider the recent availability of very large networks, there is a clear need for scalable measurements. The proposed algorithm includes operations amenable to be parallelized since it has been designed by keeping the scalability requirement in mind. Thus, one clear future course of action will be the parallelization of some key components of the algorithm.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Financial support for this research has been provided under grant PID2020-112827GB-I00 funded by MCIN/AEI/10.13039/501100011033.

References

[1] Linyuan Lü, Tao Zhou, Link prediction in complex networks: A survey, Physica A 390 (6) (2011) 1150–1170.
[2] Maciej Kurant, Patrick Thiran, Layered complex networks, Phys. Rev. Lett. 96 (13) (2006) 138701.
[3] Réka Albert, Hawoong Jeong, Albert-László Barabási, Diameter of the world-wide web, Nature 401 (6749) (1999) 130–131.
[4] M.E.J. Newman, The structure and function of complex networks, SIAM Rev. 45 (2) (2003) 167–256.
[5] Petter Holme, Beom Jun Kim, Chang No Yoon, Seung Kee Han, Attack vulnerability of complex networks, Phys. Rev. E 65 (5) (2002) 056109.
[6] Young Ae Kim, Hee Seok Song, Strategies for predicting local trust based on trust propagation in social networks, Knowl.-Based Syst. 24 (8) (2011) 1360–1371.
[7] Zhiwei Lv, Nan Zhao, Fei Xiong, Nan Chen, A new measure of identifying influential nodes: Efficiency centrality, Commun. Nonlinear Sci. Numer. Simul. 47 (2017) 151–163.
[8] S.H. Strogatz, Exploring complex networks, Nature 410 (6825) (2001) 268—276.
[9] Jurek Kolasa, Complexity, system integration, and susceptibility to change: Biodiversity connection, Ecol. Complex. 2 (4) (2005) 431–442.
[10] Ernesto Estrada, Örjan Bodin, Using network centrality measures to manage landscape connectivity, Ecol. Appl. 18 (7) (2008) 1810–1825.
[11] Linton C. Freeman, Centrality in social networks conceptual clarification, Social Networks 1 (1979) 215–239.
[12] Javier Borge-Holthoefer, Yamir Moreno, Absence of influential spreaders in rumor dynamics, Phys. Rev. E 85 (2012) 026116.
[13] Cai Gao, Xin Lan, Xiaoge Zhang, Yong Deng, A bioinspired methodology of identifying influential nodes in complex networks, PLoS One 8 (2013) e66732.
[14] Haotian Zhang, Shen Zhong, Yong Deng, Kang Hao Cheong, LFIC: Identifying influential nodes in complex networks by local fuzzy information centrality, IEEE Trans. Fuzzy Syst. (2021).
[15] Qiuyan Shang, Yong Deng, Kang Hao Cheong, Identifying influential nodes in complex networks: Effective distance gravity model, Inform. Sci. 577 (2021) 162–179.
[16] Kyle Robert Harrison, Mario Ventresca, Beatrice M. Ombuki-Berman, A meta-analysis of centrality measures for comparing and generating complex network models, J. Comput. Sci. 17 (2016) 205–215.
[17] Lei Meng, Guiqiong Xu, Pingle Yang, Dengqin Tu, A novel potential edge weight method for identifying influential nodes in complex networks based on neighborhood and position, J. Comput. Sci. 60 (2022) 101591.
[18] Linton C. Freeman, A set of measures of centrality based on betweenness, Sociometry 40 (1).
[19] Ilkka Kivimäki, Bertrand Lebichot, Jari Saramäki, Marco Saerens, Two betweenness centrality measures based on randomized shortest paths, Sci. Rep. 6 (1) (2016) 1–15.
[20] Mark E.J. Newman, A measure of betweenness centrality based on random walks, Social Networks 27 (1) (2005) 39–54.
[21] Petter Holme, Jari Saramäki, Temporal networks, Phys. Rep. 519 (3) (2012) 97–125.
[22] Ulrik Brandes, Daniel Fleischer, Centrality measures based on current flow, in: STACS 2005. Annual Symposium on Theoretical Aspects of Computer Science, Springer Berlin Heidelberg, 2005, pp. 533–544.
[23] Manuel Curado, Rocio Rodriguez, Leandro Tortosa, Jose F. Vicent, A new centrality measure in dense networks based on two-way random walk betweenness, Appl. Math. Comput. 412 (2022) 126560.
[24] Ulrik Brandes, Thomas Erlebach, Network Analysis: Methodological Foundations, Vol. 3418.
[25] Enrico Bozzo, Massimo Franceschet, Resistance distance, closeness, and betweenness, Social Networks 35 (3) (2013) 460–469.
[26] Manuel Curado, Return random walks for link prediction, Inform. Sci. 510 (2020) 99–107.
[27] Linton C. Freeman, Cynthia M. Webster, Deirdre M. Kirke, Exploring social structure using dynamic three-dimensional color images, Social Networks 20 (2) (1998) 109–118.
[28] P. Kindlmann, F. Burel, Connectivity measures: A review, Landscape Ecol. 23 (2008) 879–890.
[29] Willens Klement, R.H. Willens, POL Duwez, Non-crystalline structure in solidified gold–silicon alloys, Nature 187 (4740) (1960) 869–870.
[30] Ming-Xing Li, Yi-Tao Sun, Chao Wang, Li-Wei Hu, Sungwoo Sohn, Jan Schroers, Wei-Hua Wang, Yan-Hui Liu, Data-driven discovery of a universal indicator for metallic glass forming ability, Nature Mater. 21 (2) (2022) 165–172.
[31] Benjamin T Afflerbach, Carter Francis, Lane E Schultz, Janine Spethson, Vanessa Meschke, Elliot Strand, Logan Ward, John H Perepezko, Dan Thoma, Paul M Voyles, et al., Machine learning prediction of the critical cooling rate for metallic glasses from expanded datasets and elemental features, Chem. Mater. 34 (7) (2022) 2750–2761.
[32] Yanglin Li, Shaofan Zhao, Yanhui Liu, Pan Gong, Jan Schroers, How many bulk metallic glasses are there? ACS Comb. Sci. 19 (11) (2017) 687–693.
[33] Da-wei Ding, Jing Tan, An-hui Cai, Liu Yong, Wu Hong, An Qi, Peng-wei Li, Yan Zhang, Yang Qing, Fe–C micro-alloying effect on properties of Zr53Al11. 6Ni11. 7Cu23. 7 bulk metallic glass, Trans. Nonferr. Met. Soc. China 31 (9) (2021) 2945–2954.
[34] Haishun Liu, Qi Jiang, Juntao Huo, Yue Zhang, Weiming Yang, Xiaopeng Li, Crystallization in additive manufacturing of metallic glasses: A review, Addit. Manuf. 36 (9) (2020) 101568.
[35] Akihisa Inoue, Wei Zhang, Tao Zhang, Kei Kurosaka, Thermal and mechanical properties of Cu-based Cu-Zr-Ti bulk glassy alloys, Mater. Trans. 42 (6) (2001) 1149–1151.
[36] Kenji Amiya, Akihisa Inoue, Formation, thermal stability and mechanical properties of Ca-based bulk glassy alloys, Mater. Trans. 43 (1) (2002) 81–84.
[37] E.S. Park, D.H. Kim, Formation of Ca–Mg–Zn bulk glassy alloy by casting into cone-shaped copper mold, J. Mater. Res. 19 (3) (2004) 685–688.
[38] S. Han, X. Feng, Z.H. Lu, D. Johnson, R. Wood, Erratum:"transparent-cathode for top-emission organic light-emitting diodes", Appl. Phys. Lett. 83 (13) (2003) 2719.
[39] Minha Lee, Donghyun Bae, Wontae Kim, Dohyang Kim, Ni-based refractory bulk amorphous alloys with high thermal stability, Mater. Trans. 44 (10) (2003) 2084–2087.
[40] Y.C. Kim, W.T. Kim, D.H. Kim, A development of Ti-based bulk metallic glass, Mater. Sci. Eng. A 375 (2004) 127–135.
[41] Nobuyuki Nishiyama, Kenji Amiya, Akihisa Inoue, Bulk metallic glasses for industrial products, Mater. Trans. 45 (4) (2004) 1245–1250.
[42] David Turnbull, Under what conditions can a glass be formed? Contemp. Phys. 10 (5) (1969) 473–488.
[43] Akira Takeuchi, Akihisa Inoue, Classification of bulk metallic glasses by atomic size difference, heat of mixing and period of constituent elements and its application to characterization of the main alloying element, Mater. Trans. 46 (12) (2005) 2817–2829.
[44] Dirk Brockmann, Dirk Helbing, The hidden geometry of complex, network-driven contagion phenomena, Science 342 (6164) (2013) 1337–1342.
[45] Aurélien Gautreau, Alain Barrat, Marc Barthelemy, Global disease spread: Statistics and estimation of arrival times, J. Theoret. Biol. 251 (3) (2008) 509–522.
[46] Flavio Iannelli, Andreas Koher, Dirk Brockmann, Philipp Hövel, Igor M Sokolov, Effective distances for epidemics spreading on complex networks, Phys. Rev. E 95 (1) (2017) 012313.

Manuel Curado He received his degrees in technical engineering and computer science from University of Cordoba, Spain, in 2009 and 2013, respectively. He received his Ph.D. degree in computer science from the University of Alicante in 2018. Since 2020 he has been a lecturer at the Department of Polytechnic School of the Catholic University of Murcia (UCAM). He has conducted research stays at the Computer Vision and Pattern Recognition Lab of the University of York, as well as the European Centre for Living Technology lab at the University Ca'Foscari of Venice. His research interests include pattern recognition, machine learning, computer vision and bioinformatics.

Rocío Rodríguez Gómez She Holds a Degree in Geological Engineer from the University of Salamanca, Master in Industrial Research Technologies and, since 2016, Ph.D. in the mechanical and materials engineering programme. D. contract professor in the department of mechanical engineering specializing in numerical modelling and currently a researcher in the Fracture and Structural Integrity group at the University of Salamanca.

Fernando Terroso-Sáenz obtained his B.S. and Ph.D. in Computer Science from the University of Murcia in 2009 and 2013. Since 2017 he is associate professor at Catholic University of Murcia (UCAM). He has published more than 30 articles in international journals and congresses. His research areas include smart mobility, human-generated data analysis and mobile sensing.

Leandro Tortosa He holds the Degree in Mathematical Sciences at the University of Valencia in the speciality of Theoretical Mechanics and Astronomy. Since 2000 I am a Ph.D. in Computer Engineering. Currently, I am Associate Professor in the Department of Computer Science and Artificial Intelligence of the Polytechnic School of the University of Alicante (Spain). My researcher tasks are within University Institute for Computing Research in the research group of Análisis y Visualización de Datos en Redes.

José F. Vicent He holds a Degree in Physical Sciences at the University Complutense of Madrid in the speciality of Automatic Calculation. Since 2007 I am a Ph.D. in Computer Engineering. Currently, I am Associate Professor in the Department of Computer Science and Artificial Intelligence of the Polytechnic School of the University of Alicante (Spain). My researcher tasks are within University Institute for Computing Research in the research group of Análisis y Visualización de Datos en Redes.

4.4. ARTÍCULO 4

Este artículo ha sido publicado en la revista *Computational Materials Science*, en la editorial *ElSevier*.

Se enmarca dentro de un proyecto nacional competitivo de Ministerio de Ciencia e Innovación, titulado *ALLEGRO: Smart multi-modal crowdsensing-based system as a service oriented to the prediction of social problems*.

A continuación, se muestra algunos indicios de calidad de la revista, y otros indicios adicionales de otras plataformas.

Tabla 5. Otros indicios de calidad del artículo 4

Fecha de publicación	Citescore	5- Year Impact Factor	Lecturas en ResearchGate
Diciembre 2022	6.1	3.571	21

Computational Materials Science 218 (2023) 111986

Contents lists available at ScienceDirect

Computational Materials Science

journal homepage: www.elsevier.com/locate/commatsci

Full length article

Understanding the metallic glasses formation by applying a centrality measure based on betweenness

Rocío Rodríguez [b,1], Manuel Curado [a,*,1], Leandro Tortosa [c,1], Jose F. Vicent [c,1]

[a] Polytechnic School, Catholic University of Murcia, Campus Los Jeronimos, s/n, E-30107 Murcia, Spain
[b] Department of Mechanical Engineering, Catholic University of Ávila, Calle Canteros, s/n, E-05005, Ávila, Spain
[c] Department of Computer Science and Artificial Intelligence, University of Alicante, Campus de San Vicente, Ap. Correos 99, E-03080, Alicante, Spain

ARTICLE INFO

Dataset link: https://github.com/manucurado/D2RWBT.git

Keywords:
Metallic glasses
Centrality measure
Networks
Random walks

ABSTRACT

Key nodes have a significant impact, both structural and functional, on complex networks. Evaluating the importance of a node, from an information flow perspective, has applications in many areas but, to a great extent, they are focused on the shortest path perspective. Then, there exist measures based on the representation of the frequency in which each node plays the role of a transition node by random trajectories between two other nodes. Development of a specific kind of material, metallic glasses, has been mainly based on trial and error methods and, this approach is inefficient in terms of time and cost. Therefore, the development of new approaches for metallic glasses design remains an open topic for researchers in this field. The understanding of the role of each element could be useful for the generation of new metallic glasses and the improvement of existing ones. This work aims to study the formation capacity of metallic glass from the point of view of complex networks, using random walk betweenness centrality approaches, as well as the proposal of a new hypothesis that can serve as a roadmap for future research in this field of generation of new metallic glass. The results and discussions indicate that the proposed approach may be of great interest, for the development of metallic glasses, given the great correlation between the results of this study and the actual obtaining of metallic glasses in recent research.

1. Introduction

Biological, social or communication systems, among others, are adequately described through complex networks whose nodes represent individuals or organizations, and their links represent the relationships or interactions between them. Many of these networks belong to social fields, or represent problems as the mobile network, urban mobility or electricity networks [1,2]. In recent decades, this interdisciplinary field has provided different models that have achieved a significant impact, as layered [3], multilayer [4] and multiplex networks [5], and methods focused on the nodes [6] and predicting links to densify the network [7–9]. Other works are focusing on the study of the structure [10] and to find vulnerabilities [11].

Some problems represented as interconnected networks have a special difficulty because of they are dense networks and their study is very complex, as trust propagation in social networks [12], or the early detection of Alzheimer's disease [13]. To solve these tasks, several approaches seek to identify the influential nodes in a network [14–16].

In order to understand the role of nodes in a network, node centrality analysis can provide an efficient framework. There exist different kinds of centrality measures which classify nodes in a network based on how they influence or are affected by other nodes through their connection topology, attending to different purposes [17,18]. These measures allows us to order the nodes by a specific measure of node importance that accounts for global structural properties of complex networks [19], which can be used to order the nodes by a specific node importance ranking [20].

Evaluating the importance of a node from an information flow perspective between pairs of them, exists several methods classified as betweenness centrality measures [6]. This approach has applications in many areas such as transport, communications or infrastructure networks, and they are focused on the functioning of the network from a shortest path perspective [21]. That is to say, the betweenness centrality measures the percentage of topological shortest paths that must go through the specific node.

Nevertheless, there are problems that are not appropriate using shortest paths. Thus, exists measures based on the representation of the frequency in which each node plays the role of a transition node

* Corresponding author.
E-mail addresses: rocio.rodriguez@ucavila.es (R. Rodríguez), mcurado@ucam.edu (M. Curado), tortosa@ua.es (L. Tortosa), jvicent@ua.es (J.F. Vicent).
[1] All authors contributed equally to the work.

https://doi.org/10.1016/j.commatsci.2022.111986
Received 6 September 2022; Received in revised form 23 November 2022; Accepted 12 December 2022
Available online 23 December 2022

by random trajectories between two other nodes [22]. These kinds of measures are useful to understand the connectivity properties of static and dynamic networks, but influential nodes can work as a hub in dense networks. For a better understanding of these networks, in [23] is presented a new measure that increases the ranking of nodes belonging to dense clusters to detect the weakness of a network by removing important nodes from the network. However, this method is focused on undirected networks. In [24], a study of the characteristics of each node in directed dense network is shown, which can be used in different problems as social networks, human mobility and materials.

Development of metallic glasses (MG), non-crystalline materials made up of pure metals or combinations of metals and metalloids [25]. They were first reported 60 years ago, they are non-crystalline solid alloys and have an atomic arrangement directly inherited from the liquid state. Their atomic structure has been basically understood by using a dense random packing scheme (as it can be seen in [26]). The advent of alloys that can be made glassy in sections up to a few cm thick has stimulated further interest in this category of materials. Due to their stability and the mechanical properties [27], these new metallic glasses open up new areas of application and facilitate fundamental studies. They stand out for their particular mechanical properties since they can be up to three times harder than steel, more elastic than ceramic materials and are less brittle than transparent oxide-based glass. Despite the great potential of these materials, researchers making them have been hampered by a lack of basic scientific understanding of how these materials are shaped. Currently, combinations of elements that make up metallic glasses with characteristics of great interest are still being discovered, such as Cu–Zr–Ti [28], Ca–Mg–Cu [29], Ca–Mg–Zn [30], Ni–Nb–Sn [31], Ni–Nb–Ta [32], Ti–Cu–Ni [33] or Ti–Cu–Ni–Mo–Fe [34]. This could be prevented by combining them with other materials or by inducing a second phase to form a composite. These composites have enhanced thermo-physical properties [35].

These materials have several properties which have been deeply studied by researchers, in which exists an empirical criteria based on characteristic thermal temperature (glass transition temperature, onset crystallization temperature and liquidus temperature). In this sense, several works have been proposed to evaluate and predict the glass-forming ability of MG systems [36–39]. However, in some ways, the proposed empirical parameters are not universally applicable to the entire MG family, although some are reasonably effective in evaluating GFA [40,41]. In addition, possible new empirical parameters are currently being studied as corroborated in [42,43]. The main problem of these materials is that they have been mainly based on trial and error methods guided by rules of thumb [44]. This approach is well known to be inefficient in terms of time and cost. Therefore, the development of new approaches for MG design, particularly chemically complex metallic glasses with good glass-forming ability (GFA) properties, remains an open topic for the MG community.

The last decade has seen significant developments in data science and machine learning models based on supervised learning to address various problems related to MG design, such as the GFA of alloys, GFA-related characteristic temperatures, and various other mechanical and magnetic properties. Thus, a classifier algorithms such as support vector classifier (SVC) [45] or random forest classifier (RFC) [46,47] or back-propagation neural network (BPNN) [48] or Graph Neural Networks [49], among others, have been utilized to label MGs into different categories. It should be noted that in [48] the authors trained a back-propagation neural network algorithm by building a dataset consisting of 3277 ternary alloys compositions, through which metallic glass and non-metallic glass classes can be accurately identified. In the same way, regression algorithms such as support vector regression (SVR) [50], random forest regression (RFG) [51,52], correlation-based neural network (CBNN) [53], have been used to construct the compositions-properties predictive models.

Nevertheless, the researchers are facing challenges regarding the development of reliable and adaptable Machine learning models due to the lack of sufficient high-fidelity GFA data for the different types of metallic glasses [54].

With this in mind, a study and analysis of the formation of known metallic glasses by the application of a centrality measure based on betweenness (on a directed networks) is proposed in the adapted D2RWBT algorithm. The main objective is to know the characteristics and the importance of each element in the material composition, as well as the relationships between all elements that make it up. Thus, this knowledge not only lets us understand the role of each element in the formation of metallic glasses, but also is useful as a future road map for future research for the generation of new MG and the improvement of existing ones.

The traditional centrality measures, applied to directed networks, serve to find out which elements within the metallic glasses are mostly used in their formation. However, the choice to study the formation of metallic glasses through the D2RWBT centrality measure is due to the exhaustive analysis that this measure makes (through its four indices) of the behaviour of each element that is part of the complex networks of metallic glasses.

To achieve the goals, this paper is organized as follows. Section 2 proposes the research scheme to put forward the basic hypotheses: construction of the metallic glasses network using the proposed algorithm (D2RWBT), assignment of weights to the edges of the network and ranking the nodes. Section 3 focuses on measuring the impact of the behaviour of each element (node) of the metallic glasses network. Afterwards, a Discussion Section 4 with the common characteristics that the metallic glasses elements present is proposed. Finally, we draw some conclusions about our study.

2. Methodology

In this section, we explain the adapted algorithm to obtain a set of centrality indices that let us to describe the process of construction of the metallic glasses network from a reduced glass transition temperature approach as well as extracting additional element information by means of the D2RWBT model.

2.1. Reduced glass transition temperature (Trg)

The application of the D2RWBT algorithm requires the use of a directed network with weights. At this point the Trg plays a fundamental role since it represents the weights associated with each edge that relates the elements on the materials involved. It is necessary to point out that the Trg parameter is one of the most analysed empirical parameters for predicting glassy metal formation despite its limitations [55,56], being a current parameter in terms of its study [57–59].

For most metallic glasses, the glass formation range coincides with the eutectic region where the molten metal is stable at a lower temperature than in the other regions of the phase diagram. Therefore the reduced glass transition temperature is calculated by the equation

$$Trg = \frac{Tg}{Tl}, \tag{1}$$

where Tg is the glass transition temperature (the temperature at which a metallic glassy material lost its glassy state) and Tl is the liquidus temperature (the temperature at which a material is fully liquid, and the maximum temperature at which crystals can coexist with melt in thermodynamic equilibrium).

This parameter is used to assess the glass forming ability (GFA) of an alloy [60]. The great values of Trg in metallic glasses are to the stronger GFA [61]. Current research confirms the role of Trg as a fundamental parameter in GFA, using the Trg as one of the fundamental parameters to determine. The influence of Zn on CFA on Al–Cu–Mg (Zn) alloys [62], the influence of Sn on the plastic deformation ability of Fe–Si–B–P–Sn bulk metallic glasses, [63] or the structure and mechanical properties of Cu45Zr48Al7 bulk metallic glass and metallic glass matrix composites [64], which is why it is considered in this study as the most relevant relationship factor between the different elements that make up vitreous metals. These studies and notation let us to establish the base of the network construction.

2.2. Directed two-way random path betweenness centrality algorithm (D2RWBT)

It is well known that the importance of a node in a network can be measured by means of centrality measures. Among a large number of existing centrality measures, the betweenness centrality quantifies the number of times a node acts as a bridge along the shortest path between two other nodes. In [23], a new centrality measure (based on betweenness) for undirected networks with the characteristics of reinforcing dense communities was proposed. A related model for directed networks has been presented in [24]. This measure is composed of four indices whose main objective is to characterize the centrality in the context of cycles of length 4.

This approach measures the centrality as follows:

Being $G(V, E, W)$ a directed weighted graph, where V is the set of n vertices/nodes, E is the set of links/edges and $W \in \mathbb{R}^{n\times n}$ is the weighted matrix. We represent an edge from a node i (origin) to a destination node j as e_{ij} and the weight of link e_{ij} as w_{ij}, with $e_{ij} \in E$.

The degrees of each node are denote as D^+ (in-degree) and D^- (out-degree), that are column vectors containing the sum of the weights of the in and out-edges for each node.

As the main objective of the method is looking for paths or relationships with 4 nodes (cycles of length 4), we evaluate by means of Eq. (2) all possible 2-step paths between two nodes i and j through a transition node t ($i \to t \to j$) and all possible 2-step paths between each pair of nodes but now from origin node j to destination node i, with a transition node k. ($j \to k \to i$). From these calculations, we define the matrix P by means of a 2-length random path between three different nodes (origin, transition and destination nodes).

$$P_{itj} = \frac{w_{it}w_{tj}}{D^+(i)D^-(j)}, \tag{2}$$

where $D^+(i)$ and $D^-(j)$ is the ith component of the corresponding vectors. Notes that w_{it} and w_{tj} are the weights of the edges e_{it} and e_{tj}, respectively.

Finally, the probability P_{ij} of a 4-step path between an origin node i to another destination node j, crossing two different transition nodes t and k is evaluate in Eq. (3):

$$P_{ij} = P_{itj}P_{jki} \tag{3}$$

In general, in complex networks it is possible to distinguish two types of edges: those who connect nodes belonging to the same community or cluster named intra-cluster links and others that bridge the communities named inter-cluster links. Taking this into account, the D2RWBT approach is very useful and effective to analyse the behaviour of the nodes from the point of view to the intermediate nodes. This model is based on betweenness centrality and it calculates four indices characterizing different properties of the nodes.

Then, given the graph G, four indices characterizing different properties or behaviour of each node in the network are calculated. The calculation process of the four indices is summarized in the following Fig. 1:

These four indices, summarized in Fig. 1, can be defined as follows:

1. The Interaction Index (C^1) represents the interaction of a node in the network, that is, it indicates the number of times that a specific node acts as a transition node.
2. The Relationship Importance Index (C^2) is defined in order to analyse the importance of a node in the network and help us to differentiate when a node is important in dense or sparse communities.
3. The Individual Relevance Index (C^3) allows us to know whether a transition node is relevant in terms of intra- or inter-cluster.
4. Finally the Global Relevance Index (C^4) enables to know if a node is relevant as an inter-cluster or if it belongs to a dense cluster.

Remark that, in this model the isolated nodes and self-loops have zero centrality (see [24] for more information).

2.3. Constructing metallic glasses network

In this subsection, we describe the process of construction of the metallic glasses network by the extraction of additional element information with the D2RWBT model. In Fig. 2 a flow chart of the proposed methodology are presented.

Using a dataset with 586 different materials, a single directed weighted network $G = (V, E, W)$ is constructed, representing the composite set of information about all these materials. Such materials are made up of 45 different chemical elements representing the nodes of the network V. The set of links, E, corresponds to the relationship between the forming materials (1561 edges). Finally, the weight of each edge between nodes i and j (w_{ij}) is the average Trg values of the elements on the materials involved.

Currently, they continue to discover combinations of elements that form metallic glasses [42] and, thanks to this, the most updated databases used in current research on the formation of metallic glasses reach almost 500 compositions materials [65,66]. Nevertheless, in this paper, we use a database of metallic glasses with (586), the maximum number of these materials discovered so far; these are made up of 45 different chemical elements. Through the application of the proposed centrality model, we study the contribution of single elements to a mixture's ability to form MGAs. Moreover, the formal notation of these material compositions implied the elements and their percentage of each element in the composition (i.e. in Fig. 2 the material Fe50Co30Cu20 means that the main element is the Iron with 50%, and two secondary elements: Cobalt with 30% and Copper with 20%).

The goal is to extract the contribution of single elements to a mixture's ability to form MGAs from the perspective of the connections that each chemical element that constitutes them between the rest of the known metallic glass. For this task, we attend to the reduced glass transition temperature parameter (Trg) as a relationship factor. Using this metallic glass dataset, we can construct a network that represents the role of all elements that involved in the formation of these materials. This dataset contains all the elements that make up the current metallic glasses as well as their relevant chemical characteristics, but we focus on Trg.

It should be noted that it is a directed network in which a material is given by the directed links between the main element and the rest (secondary or complementary) that form the material composition. In addition, since it is a weighted network, each link has a weight that is given by the average relationship given by Eq. (4).

$$W_{ij} = \frac{\sum_{k=1}^{N} Trg_k R_k^i}{N}, \tag{4}$$

being N the number of composition materials where elements i (the main component) and j (secondary element) are involved, Trg_k is the Trg of the composition material k and R_k^i is the percentage of the element i in the composition k. Note that, in the construction of the network, the relationships between the secondary elements are ignored because we focus on the major element in the composition in this study.

In Fig. 2 we show how the graph has been formed, first the data of each metallic glass has been extracted (elements, percentage of the elements present and their associated Trg). In each element of metallic glass, the relationship is established from the element with the highest percentage in the material to the other elements that make it up. The edges between the majority element and the other elements that form the material, being a directed graph, are made in both directions, being the weight of each edge the atomic percentage of the element of exit of the edge by the Trg of the metallic glass to which they belong. Once this operation of all the metallic glass of the database is carried out, all the information generated is joined by building the network. To do this, as shown in the toy sample construction of the network (Fig. 3), the sum of all the weights obtained in the different edges of these same elements in all the metallic glass analysed among the number of these edges is established as the weight of each edge of one element to another. Once

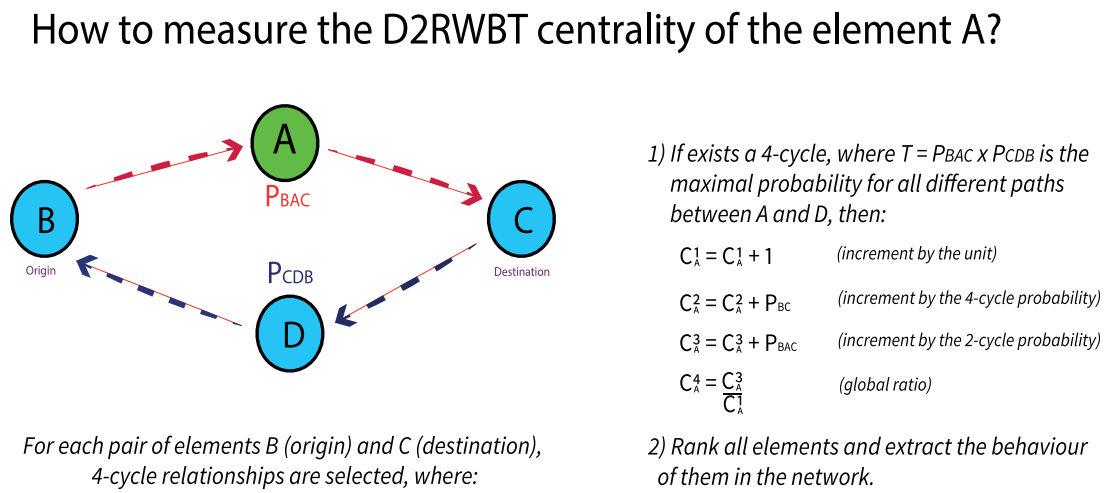


Fig. 1. Measuring the betweenness centrality of an element (node) of a directed network using D2RWBT.

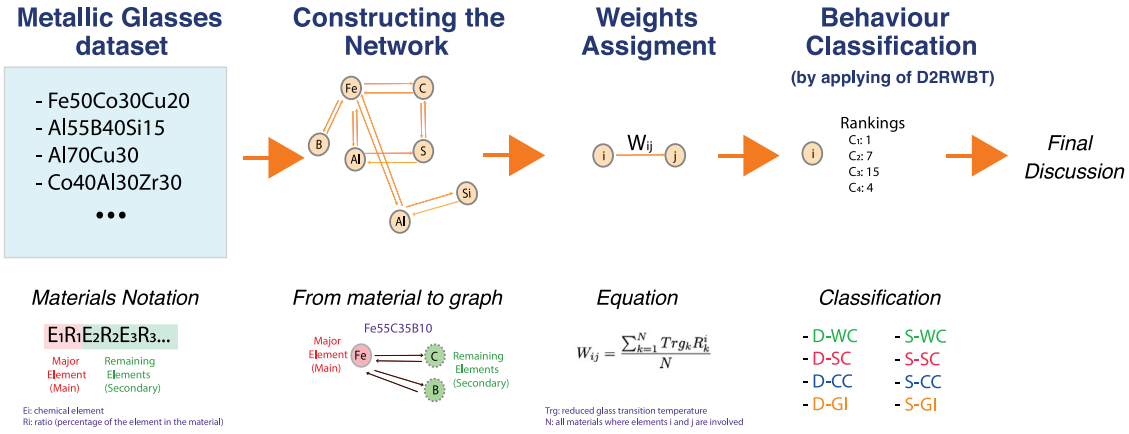


Fig. 2. Global scheme of the proposed methodology.

the operation of each edge of one element towards another has been carried out, the final network is made.

The resulting network studied illustrates the connections of the elements that form most of the known metallic glasses by the relationship between the elements through the reduced transition temperature (Trg) characteristic parameter. The resulting directed network is shown in Fig. 4.

2.4. Extracting additional element information by means of the D2RWBT model

The method presented in Section 2.2 lets us extract additional information about each node of a network. In this work, we apply this model to a real and new metallic glasses network, aiming to understand the behaviour of a specific node.

It should be noted that this model is composed by four ranking indices that classified each node of the network

Table 1 summarizes the behaviour classification of a node of the network, taking into account two groups of elements: Dense and Sparse. Each behaviour classification has an associated role, which can be relevant as inter or intra-cluster, or be useful to reinforce the cluster that belonging it. For instance, an element classified as D-WC means that belongs to a dense cluster, where its links with its neighbouring elements are weak and its role in the network is relevant to reinforce the cluster.

The following subsections show the physical meaning of the relationship of the four indices in terms of the role of an element in the metallic glasses network. First, we show the mathematical equalities of their indices that let us characterize each element in the network and their associated behaviour classification in dense networks. Then, we show the counterpart clusters (sparse).

2.4.1. Dense clusters

An element belongs to dense clusters if its C^1 ranking is higher than the betweenness centrality ranking due to the methodology involved in the D2RWBT model. As we study cycles of length four, if several 4-cycles are located in a cluster, the centrality of the nodes of this cluster increases, in contrast to betweenness measure. We can say that, a top-ranking element has relevant inter-cluster connections.

Focusing on the presented network and considering that the elements of dense clusters belongs to a higher number of metallic glasses compounds, we classify each element in dense clusters as follows:

- Unlabelled element: if the four ranking indices have similar values, the method does not describe additional information other than density.
- D-WC: C^1 ranking is higher than C^2 ranking ($C^1 > C^2$). That means that elements play the role of reinforcing the cluster but they have weak connections (low Trg).
- D-SC: C^1 ranking is not higher than C^2 ranking, and C^3 ranking is not higher than C^2 ranking ($C^1 \not> C^2 \not< C^3$). These elements reinforce the cluster with strong connections (high Trg).
- D-CC: C^1 ranking is not higher than C^2 ranking, C^3 ranking is higher than C^2 ranking, and C^4 is low ($C^1 \not> C^2 < C^3 >$

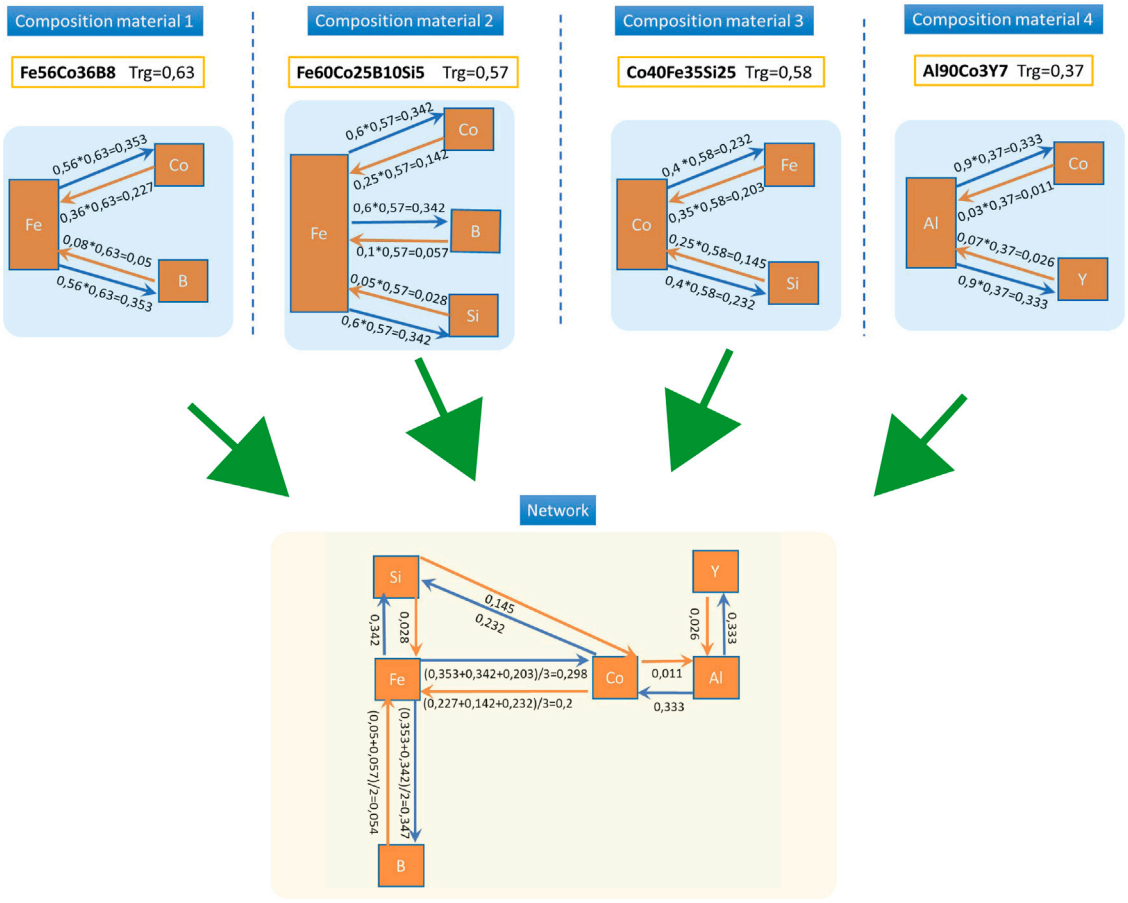


Fig. 3. Toy sample of the construction of the network, where 4 composition materials form a directed network with 6 nodes (elements).

Table 1
Behaviour classification of elements of the network and its associated role.

Behaviour classification	Role
Weak connections in dense clusters (D-WC)	Relevant to Reinforce (RR)
Strong connections in dense clusters (D-SC)	Relevant to Reinforce (RR)
Cohesive compound in dense clusters (D-CC)	Relevant to Reinforce (RR)
Global importance in dense clusters (D-GI)	Very Relevant INTRA-cluster element (VR-INTRA)
Weak connections in sparse clusters (S-WC)	Relevant INTRA-cluster element (R-INTRA)
Strong connections in sparse clusters (S-SC)	Relevant INTRA-cluster element (R-INTRA)
Cohesive compound in sparse clusters (S-CC)	Relevant INTRA-cluster element (R-INTRA)
Global importance in sparse clusters (S-GI)	Relevant INTER-cluster element (R-INTER)

C^4). These elements reinforce the cluster with strong connections (high Trg) and they are relevant in terms of intra-cluster elements in the cluster (an important element to form metallic glasses).

- D-GI: C^1 ranking is not higher than C^2 ranking, C^3 ranking is higher than C^2 ranking, and C^4 is high ($C^1 \not> C^2 < C^3 < C^4$). These elements are very relevant in terms of inter-cluster connections (help to avoid the division of metallic glasses into subgroups).

2.4.2. Sparse clusters

An element belongs to a sparse cluster if its C^1 ranking is lower than betweenness centrality ranking. This fact is marked by the lack of 4-cycles in the cluster where the element is located in contrast to the better ranking as an intermediate element in classical measures (BT). Focusing on the real Metallic Glasses Network we can say that the elements in sparse clusters have limited connections in the network because they participate in fewer metallic glasses compounds. For that, we classify each element in sparse clusters as follows:

- Unlabelled element: if the four ranking indices have similar values, the method does not describe additional information other than density.
- S-WC: C^1 ranking is higher than C^2 ranking ($C^1 > C^2$). That means that elements play the role of reinforcing the cluster but they have weak connections (low Trg).
- S-SC: C^1 ranking is not higher than C^2 ranking, and C^3 ranking is higher than C^2 ranking ($C^1 \not> C^2 < C^3$). These elements reinforce the cluster with strong connections (high Trg).
- S-CC: C^1 ranking is not higher than C^2 ranking, C^3 ranking is not higher than C^2 ranking, and C^4 is low ($C^1 \not> C^2 \not< C^3 > C^4$). These elements reinforce the cluster with strong connections (high Trg) and they are relevant in terms of intra-cluster elements in the cluster (an important element to form metallic glasses).
- S-GI: C^1 ranking is not higher than C^2 ranking, C^3 ranking is not higher than C^2 ranking, and C^4 is high ($C^1 \not> C^2 \not< C^3 < C^4$). These elements are very relevant in terms of inter-cluster connections (help to avoid the division of metallic glasses into subgroups).

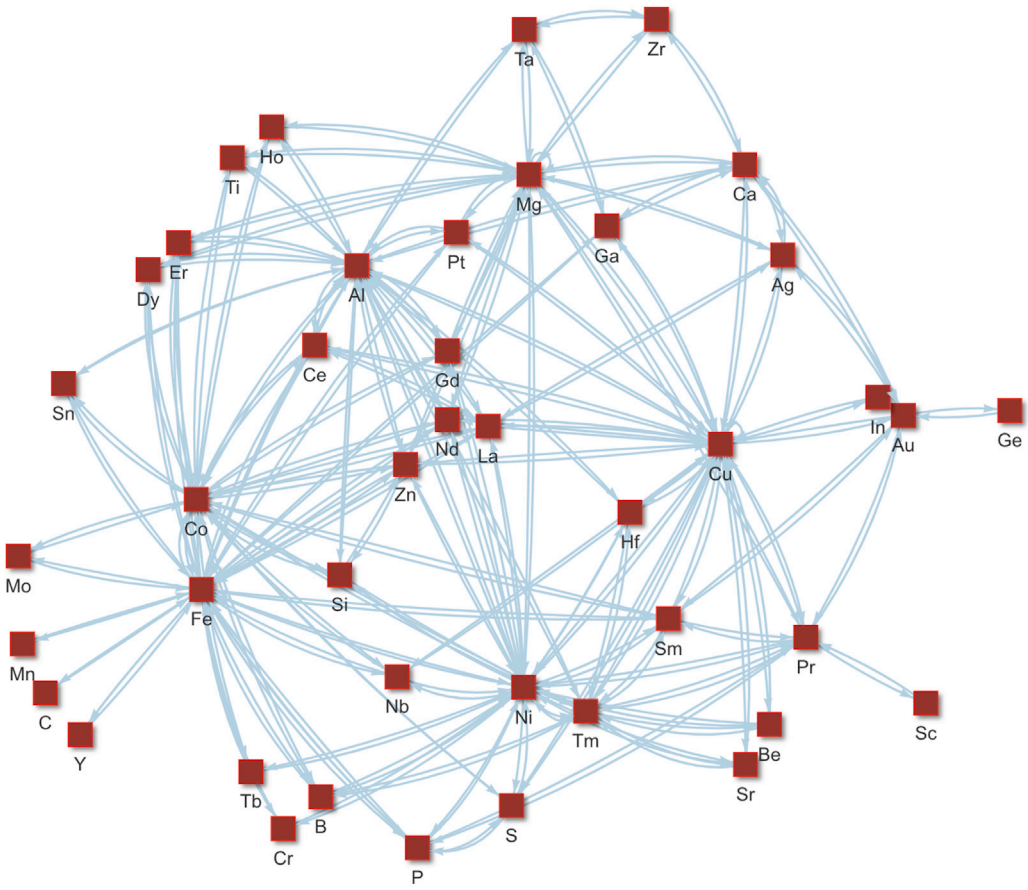


Fig. 4. Materials directed network. Nodes represent different elements and edges mean that two elements belong to some vitreous metal.

3. Results

The results presented in this section have focused on the comparison between the classic centrality measures, Betweenness (BT), and the recent centrality measure (D2RWBT) [24]. This comparison aims to provide more information on the role of each of the elements that make up this complex network of metallic glasses.

Applying the described method (Section 2.2), we analyse the behaviour of each element of the constructed metallic glasses network (Section 2.3). For that, we calculate all indices of this approach and the classical BT, and we establish two main groups: the most important elements that belong to dense and sparse clusters or communities.

After analysing the ranking position of all indices, we can observe how the five elements (Al, Co, Ni, Fe and Cu) that present the highest positions in both main indices rankings (BT and C^1).

Moreover, it should be noted that these five elements are also the 5-top positions in the ranking of C^1, although in a different order of importance. After analysing the clusters or communities of these elements in the network, we can highlight the great importance of them in their clusters, and their relevance with respect to the connection with other elements less common than the involved ones in the actual metallic glasses formation. Elements such as Al and Co, by forming dense networks, and Ni, by forming sparse networks, will represent great importance in the cohesion of the compounds in which they take part (they will have intra-cluster relevance).

Nonetheless, elements such as Fe make up sparse networks indicating its relevance in the connections between families, at the inter-cluster level.

Furthermore, we classify all elements evolved in metallic glasses compositions following the categories explained in Section 2.4 in the next subsections. Moreover, we remove the prefixes D and S of the behaviour classifications to simplify the notation.

3.1. Elements in dense clusters

In Table 2, we show the results of the positions of the different chemical elements analysed, formers of metallic glasses (MG), and the ranking position that they occupy through the calculated centrality indices.

Elements that make up dense networks tend to participate in compounds with very few elements, from highest to lowest ranking difference ($BT - C^1$): Hf, Ce, Be, Sm, Sc, Pt, Gd, P, Co, Al, La, Y, Ca, Ti and Zr

Among those that make up dense networks, we classify the elements according to its behaviour, ordered from highest to lowest relevance (Ti, P and Sm are not classified because we do not obtain additional information from the application of the method):

- WC: Y, Pt, Gd, Al, La, Ce, Zr and Hf.
- SC: Co and Ca.
- CC: Sc.
- GI: **Be**.

In bold, we highlight the Be, because it is a very relevant intra-cluster element in their compounds. Attending their indices, we can see the relevance of its C^4 ranking. This fact marked the importance of all their connections. In other words, this element does not form many compounds (low ranking in almost indices), but its relevance in them is highly (high ranking in C^4). In Section 4, we explain the implications of this role.

The remaining elements play the role of intra-cluster reinforcing the network.

Table 2
The top elements that make up dense networks are ordered by the difference in ranking between BT and C^1 (Rank Diff). The last column represents the coloured behaviour classification of each element. The unlabelled-elements classification means that is relevant in dense clusters but there is not more information about its role in the network.

Element	BT	C^1	Rank Diff.	C^2	C^3	C^4	Classification
Zirconium (Zr)	9	7	2	9	7	11	WC
Titanium (Ti)	10	8	2	8	10	15	–
Calcium (Ca)	12	10	2	7	8	5	SC
Aluminium (Al)	4	1	3	4	1	3	WC
Cobalt (Co)	5	2	3	1	3	4	SC
Yttrium (Y)	14	11	3	17	15	20	WC
Lanthanum (La)	15	12	3	15	14	17	WC
Phosphorus (P)	18	14	4	14	16	16	–
Platinum (Pt)	20	15	5	18	19	21	WC
Gadolinium (Gd)	27	22	5	25	25	25	WC
Scandium (Sc)	31	26	5	24	22	22	CC
Beryllium (Be)	34	27	7	22	21	14	GI
Samarium (Sm)	35	28	7	26	26	24	–
Cerium (Ce)	25	17	8	19	18	19	WC
Hafnium (Hf)	33	20	13	21	23	26	WC

Table 3
Top elements that make up sparse networks are ordered by the difference in ranking between BT and C^1 (Rank Diff). The last column represents the coloured behaviour classification of each element. The unlabelled-elements classification means that is relevant in sparse clusters but there is not more information about its role in the network.

Element	BT	C^1	Rank Diff.	C^2	C^3	C^4	Classification
Silicium (Si)	11	23	−12	20	20	8	–
Gold (Au)	8	13	−5	13	13	10	–
Copper (Cu)	2	5	−3	5	5	7	–
Gallium (Ga)	13	16	−3	12	12	9	–
Praseodymium (Pr)	21	24	−3	28	24	23	WC
Erbium (Er)	22	25	−3	27	28	27	WC
Iron (Fe)	1	3	−2	2	4	6	GI
Palladium (Pd)	7	9	−2	11	11	13	WC
Silver (Ag)	16	18	−2	16	17	18	CC
Strontium (Sr)	17	19	−2	10	9	1	SC
Neodymium (Nd)	19	21	−2	23	27	28	WC
Nickel (Ni)	3	4	−1	3	2	2	SC

3.2. Elements in sparse clusters

As in the previous subsection, we show in Table 3, the resulting chemical elements, formers of metallic glasses (MG), and all the rankings and classification. We can observe how the element Fe in the sparse networks table meets the characteristics of the example presented.

We can see the order, from highest to lowest ranking difference ($BT - C^1$), is: Si, Au, Cu, Ga, Pr, Er, Fe, Pd, Ag, Sr, Nd and Ni.

Similar to the previous subsection, we classify all these elements as follows:

- WC: Pr, Er, Pd and Nd.
- SC: Sr and Ni.
- CC: Ag.
- GI: **Fe**.

In difference to Be in dense clusters, we highlight Fe in sparse clusters, which is classified as the unique relevant inter-cluster element. This role means that this element is key in the network, and any change (shortages, high prices, etc.) will affect almost the whole metallic glasses network.

The remaining elements are relevant as intra-cluster elements of the network.

4. Discussion

Numerous investigations have tried to find out how metallic glasses (MG) are formed, with the great interest of being able to create MG without the need to carry out numerous trial and error tests with little chance of success. Recent research shows that in experiments and simulations for the formation of MG, at the nanoscale, groups of atoms interconnect in these materials to form highly structured *supergroups* [67–69]. Despite these ordered structures, the causes of the order have not yet been discovered. As one of the fundamental objectives of this work, the contributions of the use of complex networks to this knowledge are presented. In the 3 section, the types of connection and the classification according to a measure of centrality based on random walk betweenness have been obtained. Now, the common characteristics that these elements present, according to the classification obtained, will be explained.

4.1. Trend analysis of chemical properties with respect to the elements classification

After the analysis of the rankings og elements, we focusing on the characteristics of the elements that make up the network and due to its importance in the types of chemical bonds established as well as their physical behaviour, it has studied the possible relationship of the role that each element plays in the network and its physico-chemical characteristics: atomic mass, atomic density, valences, oxidation states, covalent, ionic and atomic radius, boiling temperatures, evaporation and fusion, electronegativity and specific heat. From the analysis of the chemical characteristics presented by the different elements and the relationship with the types of networks they form, as well as the role they play in them, trends have been extracted only in terms of atomic density and mass, which are not conclusive results but could serve as an indication in subsequent research.

Trends highlight elements such as Fe in sparse networks and Be in dense networks share the lowest values of atomic masses as well as low densities and low differences between both characteristics. They also have low atomic, ionic and covalent radius together. Both elements play a fundamental role in both types of networks, with Fe being the most relevant element for inter-cluster links, Global Importance in dense or sparse clusters (GI) and Be the most relevant for intra-cluster links (GI). On the other hand, elements such as Co and Ca in sparse networks and Ni and Sr in dense networks are relevant elements that form strong links in intra-cluster (SC), share low atomic mass values, medium-low densities and low differences between both values.

Regarding elements such as Ag in dense networks and Sc in sparse networks, they have not shown similar trends although both have the same function within the networks (CC). This fact is due to the great difference within the rankings, that is, the Sc element is a clear CC and the Ag is by a small margin considered as CC, and the difference between the two is remarkable.

4.2. Classification vs. behaviour of metallic glasses elements in complex networks

The roles of the most relevant elements analysed are discussed in this section. As far as Fe is concerned, it has been obtained as an element that can be of great importance for the formation of MG with which it normally does not conform. The role of Fe can be decisive in being able to form MG in combination with elements that normally have not obtained good results together. This fact and the results are ratified in the formation of numerous MG such as Ti–Fe–Si, Zr–Cu–Fe–Al, Fe–Co–Ni–PB or Fe–Ni–PB [70–73]. The Fe can be used to make strong intercluster links that enable the formation of MG despite its clear tendency to form sparse or dense networks. The element Be is a dense network generator with a very relevant role in the intra-cluster connection. This fact is ratified in numerous metallic glasses as Zr41.25Ti13.75Cu12.5Ni10Be22.5, Zr65-xAl10Ni10Cu15Bex or Ti48Zr29Ni6Ta1Be16 [74–78].

Experimental evidence confirms that the addition of elements such as Sc, which behaves as a Cohesive Compound in dense clusters (CC),

are very effective in improving glass formation ability (GFA) [79]. Recent investigations show that the substitution of some of the elements by Sc improves the formation ability of metallic glasses [79–81]. Just like the element Sc, the element Ag also behaves as a Cohesive Compound (CC) in sparse clusters and has been verified in research as an element whose addition is very relevant to improving the ability to form metal glasses (GFA) [82,83].

It should be noted that the results are also ratified for their use in substituting some elements for others given the role they perform in the network. A clear example of this application is the substitution of Zr by Sc given its higher formation capacity and the common role of intra-cluster links that they play in the network [79]. Another example is the substitution of Be, which is toxic, for Sc in vitreous materials with biomedical applications, given that both play the same role, being Be very relevant and Sc relevant, in intra-cluster links [80]. Based on the results, it is proposed the possibility of using these criteria to analyse the compatibility in the possible substitutions that are made in the metallic glasses formation experiments.

As predictions if it is intended to form a vitreous metal with an element that forms networks of dense type and one that forms scattered networks, an element such as Fe should be added to it, as some possible examples such as Fe–Pr–Pt, Er–Fe–Y, Gd–Fe–Pd, Nd–Fe–La, Sr–Fe–Co or Co–Fe–Ca. The Be given its role in dense networks in intracluster connections could serve to form vitreous metals such as Sc–Be–Hf, Ce–Be–Sm, Pt–Be–Gd, Ca–Be–Al, La–Ce–Be, Be–Co–Ca or Be–Sc–Ca, between others.

On the one hand, to improve the GFA of vitreous metals with dense lattice elements, Sc would be added, as Al–Co–Y–Sc, Al–Ca–Sc, Al–Co–Zr–Sc, Be–Zr–Sc or Gd–Co–Al–Y–Sc. On the other hand, to improve the GFA of vitreous metals with dispersed lattice elements, Ag would be included as Nb–Ni–Ag, Ga–Sr–Ag, Au–Cu–Si–Ag or Au–Si–Ge–Ag.

As example of possible predictions, not confirmed by experiments, we can suggest the replacement of Zr and Be by Sc, in glassy materials such as Sc–Cu–Ni–Al, Sc–Ti–Al–Cu–Ni, Ti–Cu–Ni–Sc or Ti–Ni–Cu–Sn–Sc.

All the predictions exposed have been made applying the roles of each element within the complex network with the aim of improving the GFA but, these predictions will need to be ratified by experiments.

5. Conclusion

The implementation of knowledge about complex networks, especially some types of centrality measures, in the formation of metallic glasses (MG) is of great interest for its development given the great correlation between the results of this study and the actual obtaining of metallic glasses in recent research.

In the case of the experiment proposed to elucidate the formation capacity of the elements that make up MG, the most notable conclusions of this study can be summarized as follows:

- **(i)** As far as the formation of MG is concerned, the elements that make up dense networks gain relevance relative the elements that form sparse networks with the application of the new measure of centrality (D2RWBT) and in comparison with the classic centrality measure, Betweenness (BT). Both D2RWBT and BT rank the same six main elements that form MG.
- **(ii)** Elements such as Sc and Ag, both which behaves as Cohesive Compound in dense and sparse cluster respectively (CC), whose addition is capable of increasing the ability to form metallic glasses.
- **(iii)** Elements such as Fe and Be are of great relevance in the inter and intra-cluster formation respectively. That is to say, edges between elements that belong to different types of networks and elements that usually belong to the same type of network.
- **(iv)** The relevance of elements such as Ca, Co, Ni and Sr for the formation of metallic glasses generating strong intra-cluster links (SC).
- **(v)** The study of the relationship between the physicochemical characteristics and the role played by the elements in the network is not conclusive, only trends have been extracted that can serve as indices for future research.
- **(vi)** Finally, one the main conclusion is to identify, based on this centrality measure (D2RWBT), the compatibility in the substitution of metallic glasses elements for the formation of new metallic glasses according to the role they play in the network.

As future work, it will be interesting to understand the internal properties of the elements of specific materials by combining classical mathematical and statistical methods with other deep learning models as Graph Neural Networks (GNNs).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data and code employed in this paper are available at https://github.com/manucurado/D2RWBT.git.

Acknowledgements

Financial support for this research has been provided under grant PID2020-112827GB-I00 funded by MCIN/AEI/10.13039/501100011033.

References

[1] Wei Tu Qili Gao Rui Cao Jinzhou Cao, Qingquan Li, Chen Zhong, Resolving urban mobility networks from individual travel graphs using massive-scale mobile phone tracking data, Cities 110 (2021) 103077.
[2] Grant McKenzie, Urban mobility in the sharing economy: A spatiotemporal comparison of shared mobility services, Comput. Environ. Urban Syst. 79 (2020) 101418.
[3] Maciej Kurant, Patrick Thiran, Layered complex networks, Phys. Rev. Lett. 96 (13) (2006) 138701.
[4] Mikko Kivelä, Alex Arenas, Marc Barthelemy, James P Gleeson, Yamir Moreno, Mason A Porter, Multilayer networks, J. Complex Netw. 2 (3) (2014) 203–271.
[5] Federico Battiston, Vincenzo Nicosia, Vito Latora, The new challenges of multiplex networks: Measures and models, Eur. Phys. J. Spec. Top. 226 (3) (2017) 401–416.
[6] Linton C. Freeman, A set of measures of centrality based on betweenness, Sociometry (1977) 35–41.
[7] Linyuan Lü, Tao Zhou, Link prediction in complex networks: A survey, Physica A 390 (6) (2011) 1150–1170.
[8] Manuel Curado, Return random walks for link prediction, Inform. Sci. 510 (2020) 99–107.
[9] Adam Trebach, Evan Toth, Woo Hyun Chae, Jatin Patil, Ki-Jana Carter, Thomas Sannicolo, Jeffrey C. Grossman, De-densifying metallic nanowire networks using betweenness centrality, Comput. Mater. Sci. 214 (2022) 111683.
[10] Stefano Boccaletti, Ginestra Bianconi, Regino Criado, Charo I Del Genio, Jesús Gómez-Gardenes, Miguel Romance, Irene Sendina-Nadal, Zhen Wang, Massimiliano Zanin, The structure and dynamics of multilayer networks, Phys. Rep. 544 (1) (2014) 1–122.
[11] Petter Holme, Beom Jun Kim, Chang No Yoon, Seung Kee Han, Attack vulnerability of complex networks, Phys. Rev. E 65 (5) (2002) 056109.
[12] Young Ae Kim, Hee Seok Song, Strategies for predicting local trust based on trust propagation in social networks, Knowl.-Based Syst. 24 (8) (2011) 1360–1371.
[13] Manuel Curado, Francisco Escolano, Miguel A Lozano, Edwin R Hancock, Early detection of Alzheimer's disease: Detecting asymmetries with a return random walk link predictor, Entropy 22 (4) (2020) 465.
[14] Shasha Wang, Yuxian Du, Yong Deng, A new measure of identifying influential nodes: Efficiency centrality, Commun. Nonlinear Sci. Numer. Simul. 47 (2017) 151–163.
[15] Xiaohang Zhang, Ji Zhu, Qi Wang, Han Zhao, Identifying influential nodes in complex networks with community structure, Knowl.-Based Syst. 42 (2013) 74–84.
[16] Javier Borge-Holthoefer, Yamir Moreno, Absence of influential spreaders in rumor dynamics, Phys. Rev. E 85 (2) (2012) 026116.

[17] Ernesto Estrada, Örjan Bodin, Using network centrality measures to manage landscape connectivity, Ecol. Appl. 18 (7) (2008) 1810–1825.
[18] Linton C. Freeman, Centrality in social networks conceptual clarification, Social Networks 1 (3) (1978) 215–239.
[19] Mark E.J. Newman, The structure and function of complex networks, SIAM Rev. 45 (2) (2003) 167–256.
[20] Enrico Bozzo, Massimo Franceschet, Resistance distance, closeness, and betweenness, Social Networks 35 (3) (2013) 460–469.
[21] Ulrik Brandes, Daniel Fleischer, Centrality measures based on current flow, in: Annual Symposium on Theoretical Aspects of Computer Science, Springer, 2005, pp. 533–544.
[22] Mark E.J. Newman, A measure of betweenness centrality based on random walks, Social Networks 27 (1) (2005) 39–54.
[23] Manuel Curado, Rocio Rodriguez, Leandro Tortosa, Jose F Vicent, Anew centrality measure in dense networks based on two-way random walk betweenness, Appl. Math. Comput. 412 (2022) 126560.
[24] Manuel Curado, Rocio Rodriguez, Fernando Terroso-Saenz, Leandro Tortosa, Jose F Vicent, A centrality model for directed graphs based on the two-way-random path and associated indices for characterizing the nodes, J. Comput. Sci. 63 (2022) 101819.
[25] Willens Klement, R.H. Willens, POL Duwez, Non-crystalline structure in solidified gold–silicon alloys, Nature 187 (4740) (1960) 869–870.
[26] Akihiko Hirata, Kaname Matsue, Mingwei Chen, Metallic glasses, 2016, pp. 9–14, http://dx.doi.org/10.1007/978-4-431-56056-2_2.
[27] Greg Parker, Encyclopedia of Materials: Science and Technology, Elsevier, 2001.
[28] Akihisa Inoue, Wei Zhang, Tao Zhang, Kei Kurosaka, Thermal and mechanical properties of Cu-based Cu-Zr-Ti bulk glassy alloys, Mater. Trans. 42 (6) (2001) 1149–1151.
[29] Kenji Amiya, Akihisa Inoue, Formation, thermal stability and mechanical properties of Ca-based bulk glassy alloys, Mater. Trans. 43 (1) (2002) 81–84.
[30] E.S. Park, D.H. Kim, Formation of Ca–Mg–Zn bulk glassy alloy by casting into cone-shaped copper mold, J. Mater. Res. 19 (3) (2004) 685–688.
[31] S. Han, X. Feng, Z.H. Lu, D. Johnson, R. Wood, Erratum:"transparent-cathode for top-emission organic light-emitting diodes"[appl. Phys. Lett. 82, 2715 (2003)], Appl. Phys. Lett. 83 (13) (2003) 2719.
[32] Minha Lee, Donghyun Bae, Wontae Kim, Dohyang Kim, Ni-based refractory bulk amorphous alloys with high thermal stability, Mater. Trans. 44 (10) (2003) 2084–2087.
[33] Y.C. Kim, W.T. Kim, D.H. Kim, A development of Ti-based bulk metallic glass, Mater. Sci. Eng. A 375 (2004) 127–135.
[34] Nobuyuki Nishiyama, Kenji Amiya, Akihisa Inoue, Bulk metallic glasses for industrial products, Mater. Trans. 45 (4) (2004) 1245–1250.
[35] Adit Sharma, Vladislav Zadorozhnyy, Review of the recent development in metallic glass and its composites, Metal 11 (2021).
[36] David Turnbull, Under what conditions can a glass be formed? Contemp. Phys. 10 (5) (1969) 473–488.
[37] Akihisa Inoue, High strength bulk amorphous alloys with low critical cooling rates (overview), Mater. Trans. JIM 36 (7) (1995) 866–875.
[38] Z.P. Lu, C.T. Liu, A new glass-forming ability criterion for bulk metallic glasses, Acta Mater. 50 (13) (2002) 3501–3512.
[39] X.H. Du, J.C. Huang, C.T. Liu, Z.P. Lu, New criterion of glass forming ability for bulk metallic glasses, 2007.
[40] Zhuang Li, Zhilin Long, Shan Lei, Ting Zhang, Xiaowei Liu, Dumin Kuang, Predicting the glass formation of metallic glasses using machine learning approaches, Comput. Mater. Sci. 197 (2021) 110656.
[41] Yogesh Prabhu, Ajeet K Srivastav, DV Gunderov, Jatin Bhatt, Thermodynamic model to predict bulk metallic glass forming composition in Zr-Cu-Fe-Al system and understanding the role of Dy addition, Physica B 624 (2022) 413416.
[42] Ming-Xing Li, Yi-Tao Sun, Chao Wang, Li-Wei Hu, Sungwoo Sohn, Jan Schroers, Wei-Hua Wang, Yan-Hui Liu, Data-driven discovery of a universal indicator for metallic glass forming ability, Nature Mater. 21 (2) (2022) 165–172.
[43] A.V. Maiorova, T.V. Kulikova, A.S. Bykov, New criteria for predicting compositions with high glass-forming ability in Zr–Co–Al alloys, J. Non-Crystal. Solids 594 (2022) 121812.
[44] Akihisa Inoue, Stabilization of metallic supercooled liquid and bulk amorphous alloys, Acta Mater. 48 (1) (2000) 279–306.
[45] M.Z. Li, W.H. Wang, Y.T. Sun, H.Y. Bai, Machine learning approach for prediction and understanding of glass-forming ability, J. Phys. Chem. Lett. 8 (14) (2017) 3434–3439.
[46] Logan Ward, Stephanie C. O'Keeffe, Joseph Stevick, Glenton R. Jelbert, Muratahan Aykol, Chris Wolverton, A machine learning approach for engineering bulk metallic glass alloys, Acta Mater. 159 (2018) 102–111.
[47] Fang Ren, Logan Ward, Travis Williams, Kevin J. Laws, Christopher Wolverton, Jason Hattrick-Simpers, Apurva Mehta, Accelerated discovery of metallic glasses through iteration of machine learning and high-throughput experiments, Sci. Adv. 4 (4) (2018) eaaq1566.
[48] Xiaodi Liu, Xin Li, Quanfeng He, Dandan Liang, Ziqing Zhou, Jiang Ma, Yong Yang, Jun Shen, Machine learning-based glass formation prediction in multicomponent alloys, Acta Mater. 201 (2020) 182–190.
[49] Qi Wang, Longfei Zhang, Inverse design of glass structure with deep graph neural networks, Nature Commun. 12 (1) (2021) 5359.
[50] J. Xiong, T.Y. Zhang, S.Q. Shi, Machine learning prediction of elastic properties and glass-forming ability of bulk metallic glasses, MRS Commun. 9 (2019) 576–585.
[51] Jie Xiong, San-Qiang Shi, Tong-Yi Zhang, A machine-learning approach to predicting and understanding the properties of amorphous metallic alloys, Mater. Des. 187 (2020) 108378.
[52] Binghui Deng, Yali Zhang, Critical feature space for predicting the glass forming ability of metallic alloys revealed by machine learning, Chem. Phys. 538 (2020) 110898.
[53] Majid Samavatian, Reza Gholamipour, Vahid Samavatian, Discovery of novel quaternary bulk metallic glasses using a developed correlation-based neural network approach, Comput. Mater. Sci. 186 (2021) 110025.
[54] Ziqing Zhou, Yinghui Shang, Yong Yang, A critical review of the machine learning guided design of metallic glasses for superior glass-forming ability, J. Mater. Inform. 2 (1) (2022) 2.
[55] A.-H. Cai, Y. Pan, G.-X. Sun, New thermodynamic parameter describing glass forming ability of bulk metallic glasses, Mater. Sci. Technol. 21 (10) (2005) 1222–1226.
[56] D. Wang, Y. Li, B.B. Sun, M.L. Sui, K. Lu, E. Ma, Bulk metallic glass formation in the binary Cu–Zr system, Appl. Phys. Lett. 84 (20) (2004) 4029–4031.
[57] Gayatri Tanuja Guddla, Satyadevi Ambadipudi, Vamsi Krishna Katta, Naresh Kumar Katari, Balaji Rao Ravuri, Influence of titanium content on thermal, mechanical and corrosion behaviour anomalies of nickel-molybdenum-silicate bulk metallic glasses, Silicon 14 (4) (2022) 1571–1581.
[58] Cosmin Codrean, Dragoş Buzdugan, Mircea Vodă, Viorel-Aurel Şerban, Alberto Pertuz-Comas, Effect of partial substitution of Cr with co on glass forming ability, mechanical and magnetic properties in Fe48Mo14Cr15-xCoxB6P8Si7Y2 bulk metallic glasses, Rev. UIS Ing. 21 (4) (2022) 29–38.
[59] Gayatri Tanuja Guddla, Satyadevi Ambadipudi, Shanti Yenduva, Vamsi Krishna Katta, Balaji Rao Ravuri, [Ni-Mo-Si]: Nb bulk metallic glasses: Microstructure, mechanical and corrosion studies, Silicon 14 (6) (2022) 2545–2553.
[60] F.Q. Guo, S.J. Poon, G.J. Shiflet, Investigation of glass formability in al-based multinary alloys, Scr. Mater. 43 (12) (2000) 1089–1095.
[61] Shan Zhang, Chao Wei, Liang Yang, Jingwang Lv, Haoran Zhang, Zhilin Shi, Xinyu Zhang, Mingzhen Ma, Formation ability, thermal stability, and mechanical properties of the Zr50Cu34Al8Ag8 amorphous alloys prepared by different mold materials, Mater. Sci. Eng. A 840 (2022) 142978.
[62] Vanya Dyakova, Georgi Stefanov, Ivan Penkov, Daniela Kovacheva, Nikolay Marinkov, Yana Mourdjeva, Stoyko Gyurov, Influence of Zn on glass forming ability and crystallization behaviour of rapidly solidified Al-Cu-Mg (Zn) alloys, J. Chem. Technol. Metall. 57 (3) (2022).
[63] Xiaoyu Wu, Xining Li, Xue Li, Shengli Li, Effect of Sn on plastic deformation ability of Fe-Si-BP-Sn bulk metallic glasses, Available At SSRN 4111830.
[64] Tomasz Koziel, Grzegorz Cios, Bogdan Rutkowski, Anna Wierzbicka-Miernik, Krzysztof Pajor, Piotr Bała, Structure and mechanical properties of Cu45zr48al7 bulk metallic glass and metallic glass matrix composites, Available At SSRN 4128921.
[65] Benjamin T Afflerbach, Carter Francis, Lane E Schultz, Janine Spethson, Vanessa Meschke, Elliot Strand, Logan Ward, John H Perepezko, Dan Thoma, Paul M Voyles, et al., Machine learning prediction of the critical cooling rate for metallic glasses from expanded datasets and elemental features, Chem. Mater. 34 (7) (2022) 2945–2954.
[66] Yanglin Li, Shaofan Zhao, Yanhui Liu, Pan Gong, Jan Schroers, How many bulk metallic glasses are there? ACS Combin. Sci. 19 (11) (2017) 687–693.
[67] Alain Reza Yavari, A new order for metallic glasses, Nature 439 (7075) (2006) 405–406.
[68] Majid Samavatian, Reza Gholamipour, Dmitry Olegovich Bokov, Wanich Suksatan, Vahid Samavatian, Morteza Mahmoodan, Characterization of nanoscale structural heterogeneity in metallic glasses: A machine learning study, J. Non-Crystal. Solids 578 (2022) 121344.
[69] Gongji Yang, Longfei Liu, Jiaojiao Yi, Lingti Kong, Jinfu Li, Short to medium range order and atomic dynamic property within the surface of nanoscale metallic glasses revealed by molecular dynamics simulations, J. Non-Crystal. Solids 578 (2022) 121348.
[70] Y Guo, I Bataev, K Georgarakis, AM Jorge Jr., RP Nogueira, M Pons, AR Yavari, Ni-and Cu-free Ti-based metallic glasses with potential biomedical application, Intermetallics 63 (2015) 86–96.
[71] Kaifeng Jin, Jörg F. Löffler, Bulk metallic glass formation in Zr–Cu–Fe–Al alloys, Appl. Phys. Lett. 86 (24) (2005) 241909.
[72] Mingcan Li, Huimin Guan, Sen Yang, Xu Ma, Qiang Li, Minor Cr alloyed Fe–Co–Ni–P–B high entropy bulk metallic glass with excellent mechanical properties, Mater. Sci. Eng. A 805 (2021) 140542.
[73] F.E. Luborsky, Crystallization of some Fe Ni metallic glasses, Mater. Sci. Eng. 28 (1) (1977) 139–144.
[74] Hugh A. Bruck, Ares J. Rosakis, William L. Johnson, The dynamic compressive behavior of beryllium bearing bulk metallic glasses, J. Mater. Res. 11 (2) (1996) 503–511.

[75] R.D. Conner, A.J. Rosakis, W.L. Johnson, D.M. Owen, Fracture toughness determination for a be: Ryllium-bearing bulk metallic glass, 1997.
[76] Q. Zheng, J. Du, Low beryllium content Zr-based bulk metallic glass composite with plasticity and work hardenability, J. Appl. Phys. 115 (4) (2014) 043519.
[77] Xueshan Xiao, Fang Shoushi, Wang Guoming, Hua Qin, Dong Yuanda, Influence of beryllium on thermal stability and glass-forming ability of Zr–Al–Ni–Cu bulk amorphous alloys, J. Alloys Compd. 376 (1–2) (2004) 145–148.
[78] RT Zhu, Z Wang, XH Shi, HJ Yang, XT Guo, JW Qiao, Work hardening in Ti48Zr29Ni6Ta1Be16 metallic glass matrix composites at cryogenic temperature, J. Appl. Phys. 131 (13) (2022) 135103.
[79] Jiahua Zhu, Wenjin Gao, Sirui Cheng, Xiaodi Liu, Xiaofeng Yang, Jinsen Tian, Jiang Ma, Jun Shen, Improving the glass forming ability and plasticity of ZrCuNiAlTi metallic glass by substituting Zr with Sc, J. Alloys Compd. 909 (2022) 164679.
[80] Ying Liu, Guan Wang, Haifei Li, Shujie Pang, Kewei Chen, Tao Zhang, TiCuZr-FeSnSiSc bulk metallic glasses with good mechanical properties for biomedical applications, J. Alloys Compd. 679 (2016) 341–349.
[81] SA Uporov, RE Ryltsev, VA Bykov, NS Uporova, S Kh Estemirova, NM Chtchelkatchev, Glass-forming ability, structure and magnetocaloric effect in Gd-Sc-Co-Ni-Al bulk metallic glasses, J. Alloys Compd. 854 (2021) 157170.
[82] E.S. Park, J.Y. Lee, D.H. Kim, Effect of Ag addition on the improvement of glass-forming ability and plasticity of Mg–Cu–Gd bulk metallic glass, J. Mater. Res. 20 (9) (2005) 2379–2385.
[83] Sanjay Saini, A.P. Srivastava, Suman Neogy, The effect of Ag addition on the crystallization kinetics and glass forming ability of Zr-(CuAg)-Al bulk metallic glass, J. Alloys Compd. 772 (2019) 961–967.

V – RESULTADOS

V - RESULTADOS

En la fase preliminar de la tesis, tras el análisis de medidas de centralidad aplicadas a la accesibilidad turística de una ciudad, los resultados más reseñables de la implementación de medidas clásicas de centralidad son los siguientes: se ha podido observar que una cuarta parte (25.42%) de los negocios de la ciudad están ubicados en calles inaccesibles para personas con grandes restricciones de movilidad. Además, el 8.75% de los negocios son inaccesibles para toda persona con algún problema de movilidad (solo son accesibles para personas sanas). Algunas calles podrían hacerse más accesibles a través de diferentes acciones, como mejorar el pavimento, añadiendo una barandilla, y mejorando las aceras. Por ello, tras proponer medidas de mejora leves y poco costosas (por ejemplo, mejorar el pavimento, eliminar obstáculos en aceras o incorporar barandillas en ciertas escaleras/rampas), se incorporan calles/aristas a la red (densificación), y se reduce del 25.42% al 21.32% los negocios alcanzables (mejora de la accesibilidad). Con una fuerte restauración de la ciudad (imposible en este caso de estudio al tratarse de una ciudad patrimonio de la humanidad por la UNESCO), casi todos los negocios serían accesibles por la mayoría de las personas (0.71%).

No obstante, la conclusión más relevante en este primer trabajo es la necesidad de usar una medida de centralidad de intermediación adaptada para aplicarlo en el ámbito de esta tesis (metales vítreos).

Una primera aproximación de la medida de centralidad es la denominada *Two-way random walk betweenness centrality* (2RW). Esta medida no dirigida se ha analizado mediante redes tanto reales como sintéticas.

El primer caso de estudio es una red sintética que simula la propagación de un virus, representando tres grandes comunidades con diferentes densidades locales y globales. Un primer experimento en esta red es un estudio sobre la correlación entre las medidas de centralidad de intermediación clásica y la medida propuesta 2RW. En la comparativa con la medida clásica de intermediación se puede observar cómo 2RW mejora dentro de las comunidades densas. Es decir, la densificación a nivel local mejora la importancia de los nodos que pertenecen a una

misma comunidad o clúster, mientras que a nivel global se asemejan ambas medidas y *PageRank*.

De la segunda red sintética analizada, la cual simula la posible propagación de un virus entre 20 comunidades diferentes, se corroboran los resultados anteriores, y se explica que, a medida que crecen los enlaces entre comunidades (inter-clúster) se pierde la noción de comunidad (pasa a ser un único clúster), siendo *PageRank* y *Betweenness* muy similares.

Complementariamente, se aplica el algoritmo a redes reales para confirmar los resultados observados en redes sintéticas y detectar la aplicabilidad en otros ámbitos. Adicionalmente a los rankings de las medidas indicadas, se establece como información relevante la diferencia de ranking entre ambas medidas (2RW y CBT).

Por ejemplo, se aplica sobre una red de relaciones literaria, basado en las relaciones de los personajes de juegos de tronos y las relaciones entre las tramas que les unen. Se puede concluir que la medida 2RW no sólo destaca los personajes más intermediarios (*betweenness centrality*), sino que señala los nodos influyentes de la historia (*PageRank*). En otras palabras, sirve para detectar que personajes son vitales a nivel de conectividad, que unen las diferentes tramas de la historia (global) y los personajes importantes, que podrían considerarse incluso secundarios, pero que sirven para reforzar la conexión en una trama (local).

Siguiendo esta vía de doble información (local y global, intermediación e importancia), se repite el experimento en una red real de terroristas implicados en un atentado con bomba en trenes en Madrid. Se concluye que la nueva medida de centralidad es capaz no sólo de identificar cual es el terrorista más relevante sino cual es el terrorista clave en los procesos de comunicación de la red y sin el cual esta queda incomunicada entre el mayor número de terroristas posibles. Es decir, se indican que nodos/terroristas deberían desconectarse de la red lo antes posible para aislar la conectividad global de la red.

Esta vía de información dual local-global encaja con el objetivo general de la aplicación en redes de materiales, pero tiene un par de problemas: a) no explica el comportamiento específico o rol que juega cada nodo a nivel local y global, y b) su carácter no dirigido no encaja totalmente con la forma de relacionarse de los elementos químicos en un material.

Por ello, se realiza una adaptación del algoritmo, mediante una segunda propuesta: *Directed Two-way Random Path Betweenness Centrality Algorithm* (D2RWBT). Con esta medida dirigida se propone un conjunto de 4 índices que sirven para caracterizar el comportamiento de cada nodo dentro de la red. Se aplica en tres redes reales: a) una residencia de estudiantes y las relaciones de amistad entre ellos, b) red de movilidad urbana a nivel nacional en España, y c) una red que describe las relaciones de los elementos químicos que conforman los metales vítreos, objetivo general de esta tesis doctoral.

En cuanto a la primera red de amistad de estudiantes evaluada se ha podido explicar a través de la utilización de la nueva medida de centralidad la relevancia de un estudiante dentro de la misma con dependencia del número de amigos que tuviese, conociendo lo fuertes que pueden ser cada uno de los vínculos de amistad que tiene. Es decir, se puede detectar mediante los índices propuestos si un residente tiene más o menos amigos (enlaces), la fuerza o debilidad de esas amistades, la dirección de los mismos (la percepción de amistad puede no ser bidireccional, considerando muy amigo a una persona que no te considera de la misma forma), así como si es relevante dentro del grupo de amigos (nivel local) o juega un papel de conexión con otras amistades de otros grupos (nivel global).

Respecto a la red de movilidad urbana, se ha analizado la movilidad de un país de tamaño medio como España. Con la aplicación de la nueva medida de centralidad en dicha red se ha podido identificar los nodos de transición importantes, en áreas espaciales muy diferentes, que abarcan incluso regiones despobladas. Por ejemplo, se analiza cómo una población es relevante dentro de una red de movilidad desde un punto de vista de su conectividad con otras regiones, la cantidad de trayectos entre los mismos (fuerza de la conexión), la relevancia de esas conexiones (si es con ciudades importantes), así como la importancia de esa ciudad dentro de su ámbito local (provincia/comunidad autónoma) y global (país).

Por último, se presentan los resultados relativos a la red de la formación de los metales vitreos a través de la aplicación de la medida de centralidad propuesta, objetivo principal de la presente Tesis.

Esta medida pretende proporcionar más información sobre el papel que juega cada uno de los elementos que componen esta compleja red de vidrios metálicos.

Para ello, calculamos los cuatro índices extraídos de la medida propuesta de centralidad y la medida *betweenness* clásica, y establecemos dos grupos principales: los elementos que pertenecen a agrupaciones o comunidades densas y dispersas. Tras analizar la posición en el ranking de todos los índices, podemos observar cómo los cinco elementos (aluminio - Al, cobalto - Co, níquel - Ni, hierro - Fe y cobre - Cu) que presentan las posiciones superiores en los rankings de los dos índices principales (BT y C1) son los mismos aunque en un orden de importancia distinto en cada medida.

Tras analizar las agrupaciones o comunidades de estos elementos en la red, podemos destacar la gran importancia que tienen en su clúster, y su relevancia respecto a la conexión con otros elementos menos comunes que los implicados en la propia formación de vidrios metálicos. Elementos como el aluminio y el cobalto, al formar redes densas, y níquel, al formar redes dispersas, representarán gran importancia en la cohesión de los compuestos en los que participan forman parte (tendrán relevancia intra-clúster). No obstante, elementos como el hierro conforman redes dispersas que indican su relevancia en las conexiones entre familias, a nivel inter-clúster.

En las tablas 2 y 3 del artículo 4 se recogen los resultados de las posiciones que ocupan en el ranking de las distintas medidas de centralidad analizadas, BT y C1, así como el valor de los tres índices, C2, C3 y C4. A partir de esta información se han dividido, según la clasificación anterior, en elementos que conforman redes densas y elementos que conforman redes dispersas, respectivamente. Los tres índices calculados proporcionan información sobre la fuerza de sus conexiones, la importancia relativa y absoluta en la red, o la relevancia como nodo intra o inter-clúster C2, C3 y C4 respectivamente.

En la Tabla 8 se muestran los elementos que conforman redes dispersas, hecho que implica que sean elementos que suelen participar en compuestos formados por muchos elementos, y los elementos que conforman redes densas, los cuales tienden a participar en compuestos con pocos elementos. En la tabla se han

posicionado en los dos tipos de redes de mayor a menor importancia como elementos formadores de las mismas.

Atendiendo a la clasificación de la figura 2 del artículo 4, entre los elementos que forman redes densas, aquellos que tienen conexiones débiles y son relevantes para reforzar el compuesto son los siguientes, de mayor a menor en orden de relevancia: yodo, platino, gadolinio, aluminio, lantano, cerio, circonio y hafnio. Por otra parte, los elementos formadores de redes densas que tienen conexiones fuertes son los siguientes por orden de relevancia: cobalto y calcio. Por último, en el análisis de redes densas, cabe destacar los resultados de unos elementos con relevancia en cuanto a la capacidad de conectar con otros elementos menos comunes que con los elementos típicos con los que suelen formar metales vítreos, estos son: berilio *y* escandio.

Tabla 6. Elementos formadores de redes dispersas y redes densas.

Redes dispersas	**Redes densas**
Silicio (Si)	Hafnio (Hf)
Oro (Au)	Cerio (Ce)
Cobre (Cu)	Berilio (Be)
Galio (Ga)	Samario (Sm)
Praseodimio (Pr)	Escandio (Sc)
Erbio (Er)	Platino (Pt)
Hierro (Fe)	Gadolinio (Gd)
Paladio (Pd)	Fósforo (P)
Plata (Ag)	Cobalto (Co)
Estroncio (Sr)	Aluminio (Al)
Neodimio (Nd)	Lantano (La)
Níquel (Ni)	Itrio (Y)
	Calcio (Ca)
	Titanio (Ti)
	Circonio (Zr)

Respecto a los elementos que componen las redes dispersas, los que tienen conexiones débiles y son relevantes para reforzar el compuesto son los siguientes, de mayor a menor en orden de relevancia: praseodimio, erbio, paladio y níquel. Por otro lado, los elementos que tienen conexiones fuertes, según relevancia son: estroncio y níquel. Por último, en el análisis de redes dispersas, cabe destacar los resultados de elementos importantes elementos para conectar con otros menos comunes que los típicos con los que se suele formar metales vítreos estos son: hierro y plata.

VI – DISCUSIÓN Y CONCLUSIONES

VI. DISCUSIÓN Y CONCLUSIONES

En esta tesis doctoral se propone una medida de centralidad adaptada para conocer el comportamiento de los elementos químicos que forman un tipo de materiales específicos, metales vítreos, para entender su formación y proponer teóricamente posibles modificaciones en los mismos o la creación de nuevos mediante el estudio en redes complejas.

En una primera fase de han estudiado, implementado y analizado diferentes medidas de centralidad existentes en un problema de aplicación real (accesibilidad turística de una ciudad). Como conclusión general se extrae la necesidad adaptar una medida de centralidad de intermediación (*betweenness*).

Una primera aproximación de la medida se basa en medir la importancia local y global de un nodo como nodo intermedio mediante caminos aleatorios de doble transición (2RW).

Por tanto, se ha desarrolla una nueva medida de centralidad 2RW que combina la idea de la centralidad de *betweenness* de caminos aleatorios y el algoritmo de predicción de enlaces *Return Random Walk*.

Los resultados experimentales en problemas de diferentes ámbitos (propagación de virus, redes urbanas, literatura, etc.) muestran que la medida funciona correctamente y es una alternativa válida a la medida clásica de intermediación, mejorando su rendimiento en problemas donde las redes sean especialmente densas. Además, incorpora componentes de medidas de *PageRank*, las cuales miden la importancia de los nodos en función a la importancia de aquellos con los que se relacionan/conectan. Adicionalmente, sirven para detectar debilidades dentro de la red.

Sin embargo, se trata de una medida más general que no se adapta completamente a las características de las redes de materiales de esta tesis doctoral, y no permite explicar totalmente el comportamiento y el rol que desempeña cada nodo dentro de la red a nivel local y global.

Por ello, se propone una segunda aproximación, que, aun siendo general y aplicable a diferentes problemas y ámbitos, se adapta específicamente al tipo de

red propuesta. Esta propuesta se denomina Algoritmo de Centralidad de Intermediación Aleatorio Bidireccional Dirigido (D2RWBT), un nuevo modelo de centralidad para redes dirigidas. Mediante este modelo se han obtenido un ranking de los nodos en redes dirigidas para describir su comportamiento y relevancia dentro de la red como nodos de transición. Más en detalle, el modelo describe un nodo mediante cuatro índices que proporcionan información sobre la densidad de su clúster/comunidad (denso o disperso), la fuerza o debilidad de sus conexiones, la importancia relativa y absoluta dentro de la red, o la relevancia como nodo a nivel local (intra-clúster) o a nivel global (inter-clúster).

De la aplicación de la nueva medida de centralidad (D2RWBT) se han podido extraer las siguientes conclusiones interesantes de las redes analizadas, centrando la atención a los resultados obtenidos en las redes de metales vítreos.

En la formación de metales vítreos, se han obtenido muy buenos resultados ratificados por una buena correlación entre estos y la obtención real de vidrios metálicos en recientes investigaciones.

Tras el análisis de las clasificaciones de los elementos, se estudiaron las características de los elementos químicos que componen la red debido a la importancia de los tipos de enlaces químicos y del comportamiento físico de los mismos. Por tanto, se ha estudiado la posible correlación del papel que el elemento químico juega en la red y las características físico-químicas características como son; masa atómica, densidad atómica, valencias, estados de oxidación, radio covalente, iónico y atómico, temperaturas de ebullición, evaporación y fusión, electronegatividad y calor específico. De este estudio se han extraído tendencias de relación en cuanto a la densidad atómica y la masa, las cuales son teóricas, teniendo que validarse de forma experimental. Los resultados del estudio de esta relación podrían indicar que ambas características, densidad y masa atómicas, se puedan considerar en investigaciones futuras como indicadoras.

En las tendencias analizadas más relevantes destacan elementos como el hierro en redes dispersas y el berilio en redes densas los cuales comparten los valores más bajos de masas atómicas, así como bajos valores de densidad y escasas diferencias entre ambas características. También ambos elementos poseen radios atómicos, iónicos y covalentes bajos. Tanto el hierro como el berilio tienen papeles fundamentales en ambos tipos de redes siendo el hierro el elemento más relevante para los enlaces entre clúster, de importancia global en clústeres densos o dispersos

y el estroncio el más relevante para los enlaces intra-clúster. Por otra parte, elementos como el cobalto y el calcio en redes dispersas y el níquel y el estroncio en redes densas son los más importantes para los enlaces inter-clúster, comparten valores de masa atómica bajos, densidades medias-bajas y pequeñas diferencias entre ambos valores.

En cuanto a elementos como la plata en redes densas y el escandio en redes dispersas no han mostrado tendencias similares, aunque ambos tienen la misma función dentro de las redes. Este hecho se debe a la gran diferencia dentro de las clasificaciones como, por ejemplo, el escandio y la plata, que son catalogados con este rol por un pequeño margen, y por tanto la diferencia entre ambos es notable.

A continuación, se discute la clasificación de los elementos químicos formadores de metales vítreos y el papel que desempeñan dentro de la red compleja. En cuanto al hierro, se ha obtenido como un elemento que puede ser de gran importancia para la formación de metales vítreos obteniéndose la capacidad más alta para formar metales vítreos con otros elementos que normalmente no conforma este tipo de material. El papel del hierro puede ser decisivo para poder formar vidrios metálicos en combinación con elementos que normalmente no han obtenido buenos resultados juntos. Este hecho queda ratificado por la evidencia de formación de numerosos metales vítreos con estas características como son; Ti-Fe-Si, Zr-Cu-Fe-Al, Fe-Co-Ni-Pb o Fe-Ni-Pb [49-52]. El hierro puede por tanto para formar fuertes enlaces inter-clúster que permiten la formación de vidrios metálicos a pesar de su clara tendencia a formar redes dispersas o densas.

Otros elementos, como el berilio, son generadores de redes densas con un papel muy relevante en las conexiones intra-clúster. Este hecho se ratifica en numerosos vidrios metálicos como: Zr41.25Ti13.75Cu12.5Ni10Be22.5, Zr65-xAl10Ni10Cu15Bexo Ti48Zr29Ni6Ta1Be16 [53-57].

Las pruebas experimentales confirman que la adición de elementos como el escandio, que se comporta como un compuesto cohesivo en redes densas, es muy eficaz para mejorar la capacidad de formación de vidrio [58]. Investigaciones recientes muestran que la sustitución de algunos de los elementos por escandio mejora la capacidad de formación de vidrio de los vidrios metálicos [58-60]. Al igual que el elemento escandio, la plata también se comporta como un elemento cohesivo en redes dispersas y se ha demostrado en numerosas investigaciones que

es un elemento cuya adición es muy relevante para la capacidad de formar vidrios metálicos [61, 62].

Tabla 7. Ejemplos de posibles predicciones de formación de metales vítreos basados en este nuevo entendimiento de su formación.

***Fe* uniones intercluster**	***Be* uniones intracluster**	***Sc en redes densas***	***Ag en redes dispersas***	***Sustitución Zr y Be***
Fe-Pr-Pt	Ce-Be-Sm	Al-Co-Y-Sc	Nb-Ni-Ag	Sc-Cu-Ni-Al
Er-Fe-Y	Pt-Be-Gd	Al-Ca-Sc	Ga-Sr-Ag	Sc-Ti-Al-Cu-Ni
Gd-Fe-Pd	Ca-Be-Al	Al-Co-Zr-Sc	Au-Cu-Si-Ag	Ti-Cu-Ni-Sc
Nd-Fe-La	La-Ce-Be	Be-Zr-Sc	Au-Si-Ge-Ag	Ti-Ni-Cu-Sn-Sc
Sr-Fe-Co	Be-Co-Ca	Gd-Co-Al-Y-Sc	Cu-Ga-Ag	Cu-Sc
Co-Fe-Ca	Be-Sc-Ca	Sc-Be-Gd	Pd-Fe-Ag	Cu-Sc-Ti-Co
Pd-Ni-P-Fe	Sm-Be-Al	Cu-Si-Pd-Sc	Pr-Er-Ag	Sc-Pd
Fe-Si-Y	Be-Y-Gd	Sm-Ca-Sc	Si-Au-Ag-Pd	Al-Co-Sc

Dentro del análisis también se ha estudiado la posibilidad de, en base al papel que cada elemento juega dentro de la red compleja, realizar sustituciones de estos conformando metales vítreos diferentes. Un claro ejemplo de esta aplicación es la sustitución de circonio por escandio, ambos con papeles semejantes dentro de la red, con gran capacidad de formación de vidrio y con alta relevancia en uniones intra-clúster [58]. Otro ejemplo es la sustitución del berilio, el cual presenta elevada toxicidad, por escandio para la formación de materiales vítreos con aplicaciones biomédicas, puesto que ambos juegan papeles semejantes dentro de la red, siendo el berilio muy relevante y el escandio relevante en los enlaces intra-clúster [59]. En base a los resultados, se propone que estos criterios puedan ser utilizados para analizar la compatibilidad en las posibles sustituciones a realizar con vidrios metálicos ya existentes en los experimentos de formación de vidrios metálicos.

En base a todas las consideraciones de las posibles aplicaciones del papel que cada elemento juega en la red se han realizado una serie de predicciones (ver Tabla 9), las cuales tienen como objeto únicamente ilustrar ejemplos de la aplicación de

este nuevo entendimiento de la formación de metales vítreos a través de su estudio como red compleja analizando medidas de centralidad.

Una vez discutidos los resultados obtenidos se pueden extraer, del nuevo entendimiento de la formación de metales vítreos a través de su estudio como redes complejas, las siguientes conclusiones:

a) En cuanto a la formación de metales vítreos, los elementos que componen las redes densas ganan relevancia respecto a los elementos que forman redes dispersas con la aplicación de la nueva medida de centralidad (D2RWBT) y en comparación con la medida clásica de centralidad, *Betweenness* (BT). Los seis elementos con mejor posición en el ranking, es decir los mejores formadores de metales vítreos, son los mismos para las dos medidas de centralidad, BT y D2RWBT, sin embargo, estos elementos no comparten el mismo orden.

b) Elementos como escandio y plata, se comportan como compuestos cohesivos en clústeres denso y disperso respectivamente (CC). La adición de estos elementos es capaz de aumentar la capacidad de formar vidrios metálicos.

c) Elementos como hierro y berilio son de gran relevancia en la formación inter e intra-clúster respectivamente. Es decir, forman uniones entre elementos que pertenecen a diferentes tipos de redes y elementos que, por lo general, pertenecen al mismo tipo de red.

d) La relevancia de elementos como calcio, cobalto, níquel y estroncio para la formación de vidrios metálicos generadores radica en que generan fuertes vínculos intra-clúster (SC).

e) Características de los elementos químicos tales como radio atómico, masa atómica, densidad o diferencia entre estos dos, parecen ser importantes en el papel que cada elemento, de la red, desarrolla en la formación de vidrios metálicos.

f) Por último, una de las principales conclusiones es identificar, sobre la base de esta medida de centralidad (D2RWBT), la compatibilidad en la sustitución de elementos de vidrios metálicos para la formación de nuevos vidrios metálicos según sea el papel que desempeñe el elemento a sustituir.

Como trabajo futuro, se pretende validar estos nuevos materiales de forma experimental, y aplicar este algoritmo en otros tipos de materiales, como aceros.

VII - REFERENCIAS BIBLIOGRÁFICAS

VII- REFERENCIAS BIBLIOGRÁFICAS

[1] Freeman, L. C. (1977). A set of measures of centrality based on betweenness. Sociometry, 35-41.

[2] Kurant, M., & Thiran, P. (2006). Layered complex networks. Physical review letters, 96(13), 138701. https://doi.org/10.1103/PhysRevLett.96.138701.

[3] HolmeP, K., & Yoon, C. N. (2002). Attackvulnerabilityofcomplexnetworks. PhysicalReviewE, 65(5), 056109. https://doi.org/10.1103/PhysRevE.65.056109.

[4] Kivela, M., Arenas, A., Barthelemy, M., Gleeson, J., Moreno, Y., & Porter, M. (2014). Dynamical and spectral properties of complex networks. J Complex Netw, 2, 203-71. https://doi.org/10.1093/comnet/cnu016.

[5] Battiston, F., Nicosia, V., & Latora, V. (2017). The new challenges of multiplex networks: Measures and models. The European Physical Journal Special Topics, 226, 401-416. https://doi.org/10.1140/epjst/e2016-60274-8.

[6] Boccaletti, S., Bianconi, G., Criado, R., Del Genio, C. I., Gómez-Gardenes, J., Romance, M., & Zanin, M. (2014). The structure and dynamics of multilayer networks. Physics reports, 544(1), 1-122., doi:10.1016/j.physrep.2014.07.001.

[7] Moreno, Y., & Perc, M. (2019). Focus on multilayer networks. New Journal of Physics, 22(1), 010201, doi: 10.1088/1367-2630/ab4fcb.

[8] Kim, Y. A., & Song, H. S. (2011). Strategies for predicting local trust based on trust propagation in social networks. Knowledge-Based Systems, 24(8), 1360-1371. https://doi.org/10.1016/j.knosys.2011.06.009.

[9] Wang, S., Du, Y., & Deng, Y. (2017). A new measure of identifying influential nodes: Efficiency centrality. Communications in Nonlinear Science and Numerical Simulation, 47, 151-163. https://doi.org/10.1016/j.cnsns.2016.11.008

[10] Zhang, X., Zhu, J., Wang, Q., & Zhao, H. (2013). Identifying influential nodes in complex networks with community structure. Knowledge-Based Systems, 42, 74-84. doi:10.1016/j.knosys.2013.01.017.

[11] Gosak, M., Markovič, R., Dolenšek, J., Rupnik, M. S., Marhl, M., Stožer, A., & Perc, M. (2018). Network science of biological systems at different scales: A

review. Physics of life reviews, 24, 118-135. https://doi.org/10.1016/j.plrev.2017.11.003.

[12] Yang, R., Wang, B. H., Ren, J., Bai, W. J., Shi, Z. W., Wang, W. X., & Zhou, T. (2007). Epidemic spreading on heterogeneous networks with identical infectivity. Physics Letters A, 364(3-4), 189-193, doi:10.1016/j.physleta.2006.12.021.

[13] Borge-Holthoefer, J., & Moreno, Y. (2012). Absence of influential spreaders in rumor dynamics. Physical Review E, 85(2), 026116, doi:10.1103/PhysRevE.85.026116.

[14] Gao, C., Lan, X., Zhang, X., & Deng, Y. (2013). A bio-inspired methodology of identifying influential nodes in complex networks. PloS one, 8(6), e66732, doi:10.1371/journal.pone.0066732.

[15] Lü, L., & Zhou, T. (2011). Link prediction in complex networks: A survey. Physica A: statistical mechanics and its applications, 390(6), 1150-1170, doi:10.1016/j.physa.2010.11.027.

[16] Albert, R., Jeong, H., & Barabási, A. L. (1999). Diameter of the world-wide web. nature, 401(6749), 130-131., doi:10.1038/43601.

[17] Newman, M. E. (2003). The structure and function of complex networks. SIAM review, 45(2), 167-256, doi:10.1137/s003614450342480.

[18] Klement, W., Willens, R. H., & Duwez, P. O. L. (1960). Non-crystalline structure in solidified gold–silicon alloys. Nature, 187(4740), 869-870, doi:10.1038/187869b0.

[19] Parker, G. (2001). Encyclopedia of materials: science and technology.

[20] Francés, J. F. V., Grau, L. T., & Curado, M. (2020). Las matrices y sus aplicaciones. Publicacions de la Universitat d'Alacant.

[21] Curado, M. (2020). Return random walks for link prediction. Information Sciences, 510, 99-107, https://doi.org/10.1016/j.ins.2019.09.017

[22] Turnbull, D. (1969). Under what conditions can a glass be formed? Contemporary physics, 10(5), 473-488, doi: 10.1080/00107516908204405.

[23] Inoue, A. (1995). High strength bulk amorphous alloys with low critical cooling rates (overview). Materials Transactions, JIM, 36(7), 866-875, doi:10.2320/matertrans1989.36.866

[24] Li, Z., Long, Z., Lei, S., Zhang, T., Liu, X., & Kuang, D. (2021). Predicting the glass formation of metallic glasses using machine learning approaches. Computational Materials Science, 197, 110656, https://doi.org/10.1016/j.commatsci.2021.110656.

[25] Prabhu, Y., Srivastav, A. K., Gunderov, D. V., & Bhatt, J. (2022). Thermodynamic model to predict bulk metallic glass forming composition in Zr-Cu-Fe-Al system and understanding the role of Dy addition. Physica B: Condensed Matter, 624, 413416, https://doi.org/10.1016/j.physb.2021.413416

[26] Li, M. X., Sun, Y. T., Wang, C., Hu, L. W., Sohn, S., Schroers, J., ... & Liu, Y. H. (2022). Data-driven discovery of a universal indicator for metallic glass forming ability. Nature materials, 21(2), 165-172, https://doi.org/10.1038/s41563-021-01129-6.

[27] Maiorova, A. V., Kulikova, T. V., & Bykov, A. S. (2022). New criteria for predicting compositions with high glass-forming ability in Zr–Co–Al alloys. Journal of Non-Crystalline Solids, 594, 121812. https://doi.org/10.1016/j.jnoncrysol.2022.121812

[28] Inoue, A. (2000). Stabilization of metallic supercooled liquid and bulk amorphous alloys. Acta materialia, 48(1), 279-306, https://doi.org/10.1016/S1359-6454(99)00300-6.

[29] Sun, Y. T., Bai, H. Y., Li, M. Z., & Wang, W. H. (2017). Machine learning approach for prediction and understanding of glass-forming ability. The journal of physical chemistry letters, 8(14), 3434-3439, https://doi.org/10.1021/acs.jpclett.7b01046.

[30] Ward, L., O'Keeffe, S. C., Stevick, J., Jelbert, G. R., Aykol, M., & Wolverton, C. (2018). A machine learning approach for engineering bulk metallic glass alloys. Acta Materialia, 159, 102-111. https://doi.org/10.1016/j.actamat.2018.08.002.

[31] Ren, F., Ward, L., Williams, T., Laws, K. J., Wolverton, C., Hattrick-Simpers, J., & Mehta, A. (2018). Accelerated discovery of metallic glasses through iteration of machine learning and high-throughput experiments. Science advances, 4(4), eaaq1566, doi: 10.1126/sciadv.aaq1566.

[32] Liu, X., Li, X., He, Q., Liang, D., Zhou, Z., Ma, J., ... & Shen, J. (2020). Machine learning-based glass formation prediction in multicomponent alloys. Acta Materialia, 201, 182-190. https://doi.org/10.1016/j.actamat.2020.09.081.

[33] Wang, Q., & Zhang, L. (2021). Inverse design of glass structure with deep graph neural networks. Nature communications, 12(1), 5359. https://doi.org/10.1038/s41467-021-25490-x.

[34] Xiong, J., Zhang, T. Y., & Shi, S. Q. (2019). Machine learning prediction of elastic properties and glass-forming ability of bulk metallic glasses. MRS Communications, 9(2), 576-585. https://doi.org/10.1557/mrc.2019.44.

[35] Xiong, J., Shi, S. Q., & Zhang, T. Y. (2020). A machine-learning approach to predicting and understanding the properties of amorphous metallic alloys. Materials & design, 187, 108378. https://doi.org/10.1016/j.matdes.2019.108378.

[36] Deng, B., & Zhang, Y. (2020). Critical feature space for predicting the glass forming ability of metallic alloys revealed by machine learning. Chemical Physics, 538, 110898. https://doi.org/10.1016/j.chemphys.2020.110898

[37] Samavatian, M., Gholamipour, R., & Samavatian, V. (2021). Discovery of novel quaternary bulk metallic glasses using a developed correlation-based neural network approach. Computational Materials Science, 186, 110025. https://doi.org/10.1016/j.commatsci.2020.110025

[38] Zhou, Z., Shang, Y., & Yang, Y. (2022). A critical review of the machine learning guided design of metallic glasses for superior glass-forming ability. Journal of Materials Informatics, 2(1), 2,doi: 10.20517/jmi.2021.12

[39] Cai, A. H., Pan, Y., & Sun, G. X. (2005). New thermodynamic parameter describing glass forming ability of bulk metallic glasses. Material.s science and technology, 21(10), 1222-1226. https://doi.org/10.1179/174328405X58887

[40] Wang, D., Li, Y., Sun, B. B., Sui, M. L., Lu, K., & Ma, E. (2004). Bulk metallic glass formation in the binary Cu–Zr system. Applied Physics Letters, 84(20), 4029-4031, doi: 10.1063/1.1751219

[41] Guddla, G. T., Ambadipudi, S., Katta, V. K., Katari, N. K., & Ravuri, B. R. (2022). Influence of titanium content on thermal, mechanical and corrosion behaviour anomalies of nickel-molybdenum-silicate bulk metallic glasses. Silicon, 14(4), 1571-1581. https://doi.org/10.1007/s12633-020-00841-3.

[42] Codrean, C., Buzdugan, D., Vodă, M., Şerban, V. A., & Pertuz-Comas, A. (2022). Effect of partial substitution of Cr with Co on glass forming ability, mechanical and magnetic properties in Fe48Mo14Cr15-xCoxB6P8Si7Y2 bulk metallic glasses. Revista UIS Ingenierías, 21(4), 29-38. https://doi.org/10.18273/revuin.v21n4-2022003.

[43] Guddla, G. T., Ambadipudi, S., Yenduva, S., Katta, V. K., & Ravuri, B. R. (2022). [Ni-Mo-Si]: Nb Bulk Metallic glasses: Microstructure, Mechanical and Corrosion Studies. Silicon, 14(6), 2545-2553. https://doi.org/10.1007/s12633-021-01037-z

[44] Guo, F. Q., Poon, S. J., & Shiflet, G. J. (2000). Investigation of glass formability in Al-based multinary alloys. Scripta materialia, 43(12), 1089-1095. https://doi.org/10.1016/S1359-6462(00)00540-6.

[45] Zhang, S., Wei, C., Yang, L., Lv, J., Zhang, H., Shi, Z. & Ma, M. (2022). Formation ability, thermal stability, and mechanical properties of the Zr50Cu34Al8Ag8 amorphous alloys prepared by different mold materials. Materials Science and Engineering: A, 840, 142978. https://doi.org/10.1016/j.msea.2022.142978.

[46] Dyakova, V., Stefanov, G., Penkov, I., Kovacheva, D., Marinkov, N., Mourdjeva, Y., & Gyurov, S. (2022). Influence of Zn on glass forming ability and crystallization behaviour of rapidly solidified Al-Cu-Mg (Zn) ALLOYS. Journal of Chemical Technology & Metallurgy, 57(3).

[47] Wu, X., Li, X., Li, X., & Li, S. Effect of Sn on plastic deformation ability of Fe-Si-BP-Sn bulk metallic glasses. Available at SSRN 4333716. http://dx.doi.org/10.2139/ssrn.4333716.

[48] Koziel, T., Cios, G., Rutkowski, B., Wierzbicka-Miernik, A., Pajor, K., & Bała, P. Structure and Mechanical Properties of Cu45zr48al7 Bulk Metallic Glass and Metallic Glass Matrix Composites. Available at SSRN 4128921. http://dx.doi.org/10.2139/ssrn.4128921.

[49] Guo, Y., Bataev, I., Georgarakis, K., Jorge Jr, A. M., Nogueira, R. P., Pons, M., & Yavari, A. R. (2015). Ni-and Cu-free Ti-based metallic glasses with potential biomedical application. Intermetallics, 63, 86-96.

[50] Jin, K., & Löffler, J. F. (2005). Bulk metallic glass formation in Zr–Cu–Fe–Al alloys. Applied Physics Letters, 86(24), 241909.

[51] Li, M., Guan, H., Yang, S., Ma, X., & Li, Q. (2021). Minor Cr alloyed Fe–Co–Ni–P–B high entropy bulk metallic glass with excellent mechanical properties. Materials Science and Engineering: A, 805, 140542.

[52] Luborsky, F. E. (1977). Crystallization of some Fe□ Ni metallic glasses. Materials Science and Engineering, 28(1), 139-144.

[53] Bruck, H. A., Rosakis, A. J., & Johnson, W. L. (1996). The dynamic compressive behavior of beryllium bearing bulk metallic glasses. Journal of Materials Research, 11(2), 503-511.

[54] Conner, R. D., Rosakis, A. J., Johnson, W. L., & Owen, D. M. (1997). Fracture toughness determination for a be: Ryllium-bearing bulk metallic glass.

[55] Zheng, Q., & Du, J. (2014). Low beryllium content Zr-based bulk metallic glass composite with plasticity and work hardenability. Journal of Applied Physics, 115(4), 043519.

[56] Xiao, X., Shoushi, F., Guoming, W., Qin, H., & Yuanda, D. (2004). Influence of beryllium on thermal stability and glass-forming ability of Zr–Al–Ni–Cu bulk amorphous alloys. Journal of alloys and compounds, 376(1-2), 145-148.

[57] Zhu, R. T., Wang, Z., Shi, X. H., Yang, H. J., Guo, X. T., & Qiao, J. W. (2022). Work hardening in Ti48Zr29Ni6Ta1Be16 metallic glass matrix composites at cryogenic temperature. Journal of Applied Physics, 131(13), 135103.

[58] Zhu, J., Gao, W., Cheng, S., Liu, X., Yang, X., Tian, J., ... & Shen, J. (2022). Improving the glass forming ability and plasticity of ZrCuNiAlTi metallic glass by substituting Zr with Sc. Journal of Alloys and Compounds, 909, 164679.

[59] Liu, Y., Wang, G., Li, H., Pang, S., Chen, K., & Zhang, T. (2016). TiCuZrFeSnSiSc bulk metallic glasses with good mechanical properties for biomedical applications. Journal of Alloys and Compounds, 679, 341-349.

[60] Uporov, S. A., Ryltsev, R. E., Bykov, V. A., Uporova, N. S., Estemirova, S. K., & Chtchelkatchev, N. M. (2021). Glass-forming ability, structure and magnetocaloric effect in Gd-Sc-Co-Ni-Al bulk metallic glasses. Journal of Alloys and Compounds, 854, 157170.

[61] Park, E. S., Lee, J. Y., & Kim, D. H. (2005). Effect of Ag addition on the improvement of glass-forming ability and plasticity of Mg–Cu–Gd bulk metallic glass. Journal of materials research, 20(9), 2379-2385.

[62] Saini, S., Srivastava, A. P., & Neogy, S. (2019). The effect of Ag addition on the crystallization kinetics and glass forming ability of Zr-(CuAg)-Al bulk metallic glass. Journal of Alloys and Compounds, 772, 961-967.

ANEXO: INDICIOS DE CALIDAD DE LAS REVISTAS

ANEXO I: INDICIOS DE CALIDAD DE LAS REVISTAS

En este anexo se adjuntan tablas e imágenes que resumen e ilustran los indicios de calidad de los cuatro artículos científicos que conforman esta Tesis doctoral por compendio.

Tabla 8. Resumen de los indicios de calidad de los artículos de la tesis por compendio

Nº Art.	Título	Nombre de la revista	Nº de citas
1	*A new methodology to study street accessibility: a case study of Ávila (Spain).*	*International Journal of Geo-Information* (MDPI)	*0*
2	*A new centrality measure in dense networks based on two-way random walk betweenness.*	*Applied Mathematics and Computation* (ELSEVIER)	*9*
3	*A centrality model for directed graphs based on the two-way-random path and associated indices for characterizing the nodes.*	*Journal of Computational Science* (ELSEVIER)	*3*
4	*Understanding the metallic glasses formation by applying a centrality measure based on betweenness.*	*Computational Materials Sciences* (ELSEVIER)	*0*

Tabla 9. Resumen de los indicios de calidad de las revistas

Nombre de la revista	Factor de Impacto (año del F.I.)	Cuartil	Ranking y categoría
International Journal of Geo-Information **(MDPI)**	3.099 (2021)	Q3	88/164 (COMPUTER SCIENCE, INFORMATION SYSTEMS)
Applied Mathematics and Computation **(ELSEVIER)**	4.397 (2021)	Q1	7/267 (MATHEMATICS, APPLIED)
Journal of Computational Science **(ELSEVIER)**	3.817 (2021)	Q1	25/110 (COMPUTER SCIENCE, THEORY & METHODS)
Computational Materials Sciences **(ELSEVIER)**	3.572 (2021)	Q3	184/345 (MATERIALS SCIENCE, MULTIDISCIPLINARY)

Artículo 1.- Curado, M., Rodríguez, R., Jiménez, M., Tortosa, L., & Vicent, J. F. (2021). A New Methodology to Study Street Accessibility: A Case Study of Avila (Spain). *ISPRS International Journal of Geo-Information*, *10*(7), 491.

DOI: https://doi.org/10.3390/ijgi10070491

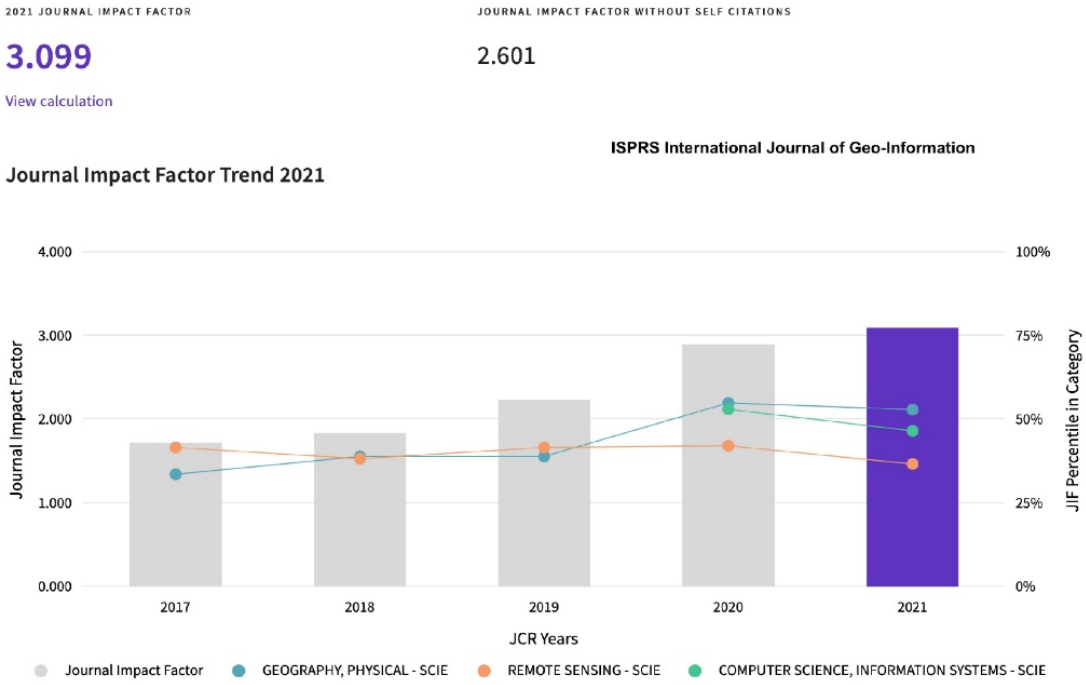


Figura 6.Scores de la revista ISPRS International Journal of Geo-Information, 2021

2.- Curado, M., Rodríguez, R., Tortosa, L., & Vicent, J. F. (2022). Anew centrality measure in dense networks based on two-way random walk betweenness. *Applied Mathematics and Computation*, *412*, 126560.

DOI: https://doi.org/10.1016/j.amc.2021.126560

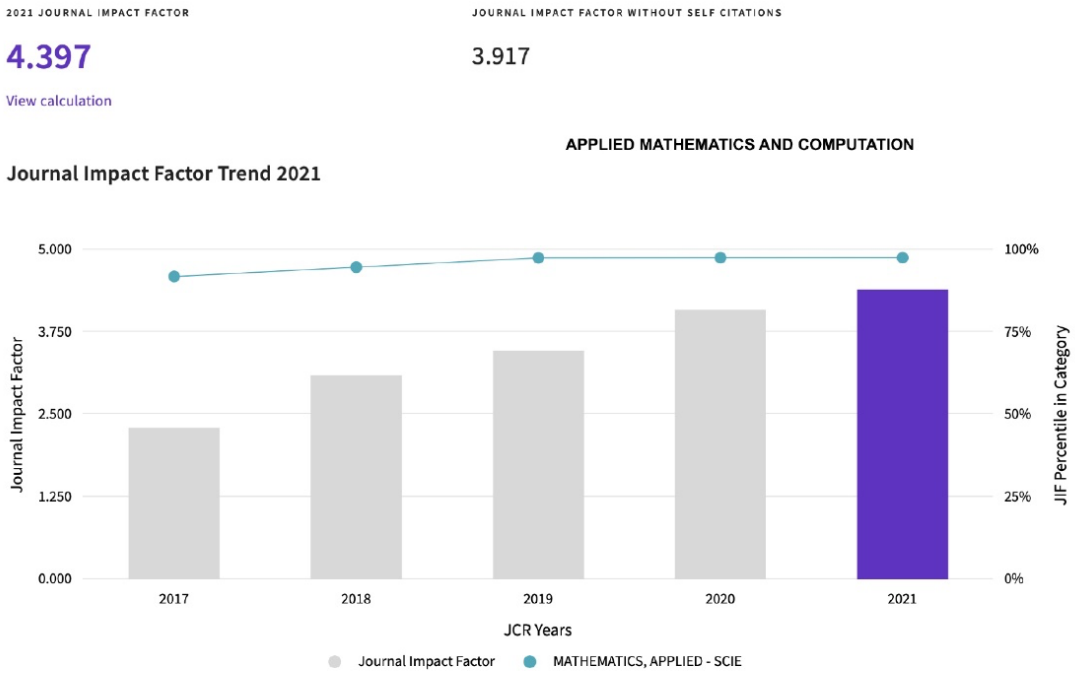


Figura 7. Scores de la revista Applied Mathematics and Computation

3.- Curado, M., Rodríguez, R., Terroso-Sáenz, F., Tortosa, L., & Vicent, J. F. (2022). A centrality model for directed graphs based on the two-way-random path and associated indices for characterizing the nodes. *Journal of Computational Science*, *63*, 101819.

DOI: https://doi.org/10.1016/j.jocs.2022.101819

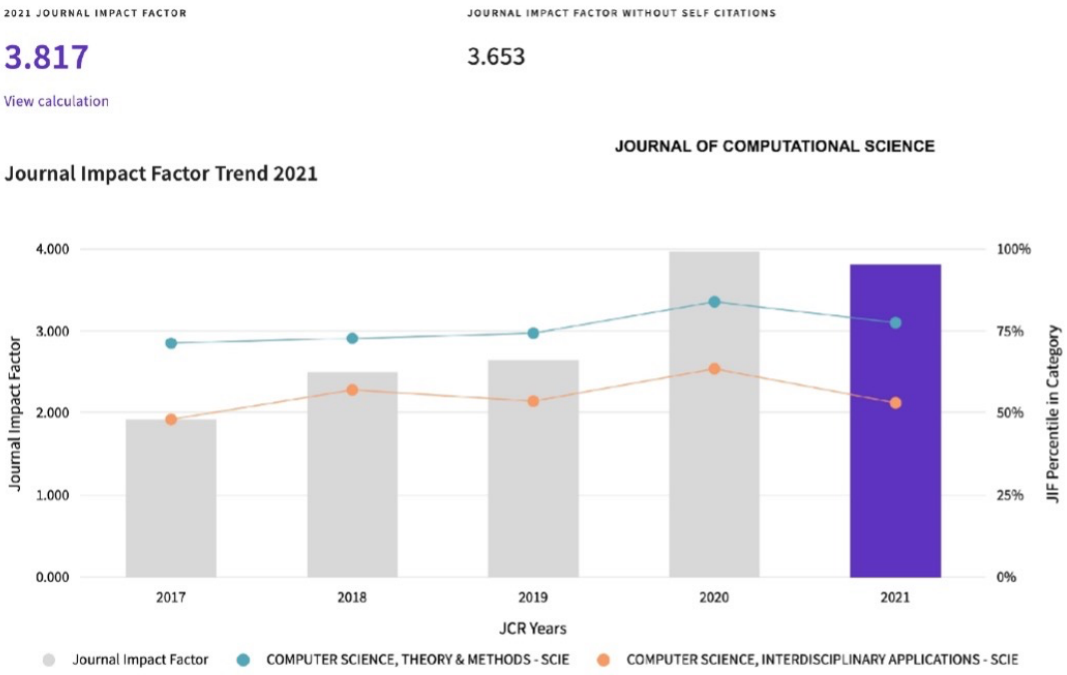


Figura 8. Scores de la revista Journal of Computational Science,

4.- Rodríguez, R., Curado, M., Tortosa, L., & Vicent, J. F. (2023). Understanding the metallic glasses formation by applying a centrality measure based on betweenness. *Computational Materials Science*, *218*, 111986.

DOI: https://doi.org/10.1016/j.commatsci.2022.111986

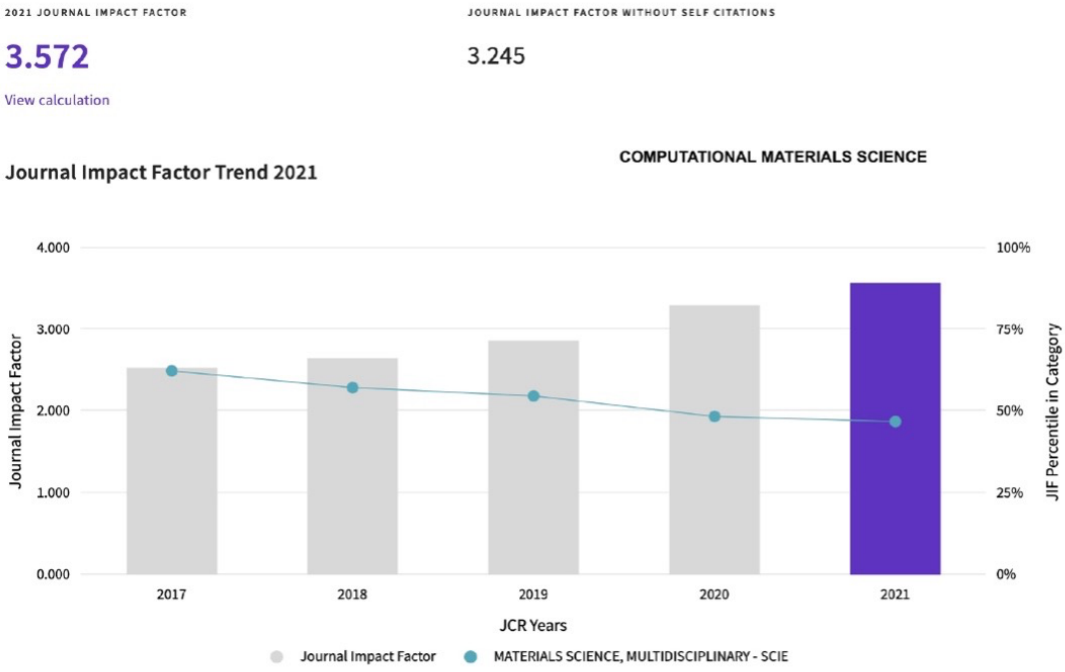


Figura 9. Scores de la revista Computational Materials Science