

# Análisis estructural y exploración evolutiva a partir de la red temática de investigación de un campo: caso de la negociación de alta frecuencia

Structural analysis and evolutionary exploration based on the research topic network of a field: a case in high-frequency trading

Mengran Xia; Han Huang; Hongyu Wang; Jing Lin

**Note:** This article can be read in its English original version on:  
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/86789>

Cómo citar este artículo.

Este artículo es una traducción. Por favor cite el original inglés:

**Xia, Mengran; Huang, Han; Wang, Hongyu; Lin, Jing** (2022). "Structural analysis and evolutionary exploration based on the research topic network of a field: a case in high-frequency trading". *Profesional de la información*, v. 31, n. 3, e310314.

<https://doi.org/10.3145/epi.2022.may.14>

Artículo recibido el 29-11-2021  
Aceptación definitiva: 19-04-2022



**Mengran Xia**

<https://orcid.org/0000-0003-1592-847X>

Zhongnan University of Economics and Law  
School of Statistics and Mathematics  
182# Nanhu Avenue, East Lake High-tech  
Development Zone  
430073 Wuhan, R. P. China  
[xiamengran@stu.zuel.edu.cn](mailto:xiamengran@stu.zuel.edu.cn)



**Han Huang**

<https://orcid.org/0000-0002-1517-9731>

Wuhan University  
School of Information Management  
299# Bayi Road, Wuchang District  
430072 Wuhan, R. P. China  
[huanghan@whu.edu.cn](mailto:huanghan@whu.edu.cn)



**Hongyu Wang**

<https://orcid.org/0000-0002-5063-9166>

Wuhan University  
School of Information Management  
299# Bayi Road, Wuchang District  
430072 Wuhan, R. P. China  
[wanghongyu@whu.edu.cn](mailto:wanghongyu@whu.edu.cn)



**Jing Lin** ✉ (Autor de correspondencia)

<https://orcid.org/0000-0001-8885-6708>

Zhongnan University of Economics and Law  
Library  
182# Nanhu Avenue, East Lake High-tech  
Development Zone  
430073 Wuhan, R. P. China  
[linjing@zuel.edu.cn](mailto:linjing@zuel.edu.cn)

## Resumen

Este estudio tiene como objetivo analizar sistemáticamente la dinámica de distribución de los temas de investigación y descubrir el estado de desarrollo de la investigación en el campo específico, lo que proporcionará una referencia práctica para desarrollar servicios profesionales de conocimiento de la materia en la era de los grandes datos. La red de temas de investigación se construye y analiza utilizando métodos y herramientas de la cienciometría. Se realizan estadísticas básicas sobre las características de la red para revelar el estado de la investigación. Se llevan a cabo la detección de comunidades, la ordenación de nodos y otros pasos para generar el diagrama aluvial evolutivo. Luego, se analizan los resultados relevantes para explorar la estructura del conocimiento del campo específico y el contexto evolutivo de los temas de investigación. Se ejecuta un análisis de visualización de la estructura de red del último período para distinguir conceptos relacionados y predecir las tendencias de investigación. Tomando como caso el comercio de alta frecuencia (HFT), este estudio logra un análisis cienciométrico diversificado de la red de temas de investigación y una exploración de la evolución multidimensional de los temas de investigación relevantes en el campo específico, lo que permite obtener

algunos conocimientos. (1) Seis temas principales en HFT: liquidez y microestructura del mercado, eficiencia del mercado, mercado financiero, mercado incompleto, cointegración y descubrimiento de precios, y estudio de eventos. (2) El enfoque de investigación sobre los mercados se transfirió gradualmente de internacional a emergente, mientras que la atención continua a los problemas relacionados con la volatilidad/riesgo. (3) El énfasis cambiará de la teoría a la práctica, las tecnologías (big data, etc.) y las teorías (behavioral finance, etc.) tendrán más interacción con HFT. Se propone una idea de investigación efectiva para revelar la estructura del conocimiento de campo y analizar el contexto evolutivo de los temas de investigación, lo que demuestra las percepciones del conocimiento.

### Palabras clave

Red de temas de investigación; Análisis evolutivo; Cienciometría; *NEViewer*; *Gephi*; Negociación de alta frecuencia; Tendencias; Gráficos.

### Abstract

This study aims to systematically analyze the distribution dynamics of research topics and uncover the development state of the research in the specific field, which will provide a practical reference for developing professional subject knowledge services in the era of big data. The research topic network is constructed and analyzed using methods and tools of scientometrics. Basic statistics on network characteristics are performed to reveal the research status. Community detection, node ordering, and other steps are conducted to generate the evolutionary alluvial diagram. Then, relevant results are analyzed to explore the knowledge structure of the specific field and evolutionary context of research topics. Visualization analysis on the network structure of the latest period is executed to distinguish related concepts and predict the research trends. Taking high-frequency trading (HFT) as a case, this study achieves diversified scientometrics analysis of the research topic network and multi-dimensional evolution exploration of the relevant research topics in the specific field, which obtaining some knowledge insights. (1) Six major topics in HFT: liquidity & market microstructure, market efficiency, financial market, incomplete market, cointegration & price discovery, and event study. (2) The research focus about markets gradually transferred from international to emerging, meanwhile continuous attention to volatility/risk related issues. (3) The emphasis will change from theory to practice, technologies (big data, etc.) and theories (behavioral finance, etc.) will have more interaction with HFT. An effective research idea is proposed to reveal the knowledge structure of field and analyze the evolutionary context of research topics, which demonstrating the knowledge insights.

### Keywords

Research topic network; Evolutionary analysis; Scientometrics; *NEViewer*; *Gephi*; High-frequency trading; HFT; Emerging trends; Graphs.

#### Financiación

Este trabajo es un resultado del proyecto de investigación "Detección de tendencias emergentes basada en un gráfico de conocimiento de ciencia abierta a gran escala" (Nº 71874129), financiado por el *National Natural Science Fund of China* (2019).

## 1. Introducción

Los artículos científicos son canales fundamentales para difundir tecnología y conocimiento a través del tiempo y el espacio (Li; Gong; Zhang, 2011). Con la creciente inversión en investigación y desarrollo, la bibliografía científica se ha expandido rápidamente a gran escala y ha aumentado enormemente. Frente a los datos masivos en el campo científico (Guo *et al.*, 2014), la mayoría de los investigadores necesitan métodos más rápidos e inteligentes para rastrear los puntos críticos de investigación y comprender las tendencias emergentes (Howe *et al.*, 2008). Las redes de conocimiento formadas por la autoorganización en la comunicación científica, como las redes de citas y las redes de temas de investigación, reflejan la estructura y los procesos del conocimiento de cada tema o campo. Los kits de herramientas como *Citespace*, *VOSviewer* y *NEViewer* (Wang; Cheng, 2013) forman parte sistemáticamente de los métodos y tecnologías para explorar conjuntos de datos científicos y permiten obtener diversos tipos de análisis de visualización de estas redes de conocimiento (Gu *et al.*, 2019; Moral-Muñoz *et al.*, 2020; Wang; Wang; Huang, 2021). Con el análisis de la evolución de los temas de investigación en un tema/campo específico, se pueden descubrir sus puntos críticos y su estado de desarrollo, lo que facilita comprender la situación. Ello se ha convertido gradualmente en una forma fundamental de obtener formas de conocimiento sobre cualquier tema en el campo de la biblioteconomía y las ciencias de la información en la era de los grandes datos (*big data*) (Huang; Wang; Wang, 2020).

Para afrontar la gestión de los datos masivos en el campo científico, los investigadores necesitan métodos más rápidos e inteligentes para rastrear los puntos críticos de investigación y comprender las tendencias emergentes

Con el desarrollo constante y la adopción generalizada de tecnologías de comunicación de nueva generación, computación de alto rendimiento, inteligencia artificial y otras tecnologías se crean nuevas condiciones para la investigación en diversos temas/campos. Los avances tecnológicos han repercutido también en las finanzas y han dado lugar a un nuevo campo de investigación: el comercio de alta frecuencia (*high-frequency trading*, HFT). Nacido del comercio asistido por computadora y el comercio algorítmico (*algorithmic trading*, AT) (*Staff of the Division of Trading and Markets*, 2014), HFT es un método que utiliza sistemas informáticos para tomar decisiones comerciales en un corto tiempo después de procesar los datos adecuados y realizar análisis cuantitativos (**Aldridge**, 2009) con alta velocidad y baja latencia (**Currie; Seddon**, 2016). Comenzó a utilizarse en 1998 después de que la *Comisión de Bolsa y Valores de los Estados Unidos* (*U.S. Securities and Exchange Commission*, SEC) autorizara oficialmente el comercio electrónico (*IBISWorld*, 2012). Posteriormente, gracias al desarrollo de las redes de comunicación, informática, inteligencia artificial y otras tecnologías, y la revolución de las reglas de negociación, como la *Regulación de Sistemas Alternativos de Negociación* (*Alternative Trading Systems Regulation*), la *Regulación del Sistema Nacional de Mercados* (*Regulation National Market System*) y el régimen de *tick size* (el precio mínimo que un valor puede mover en un intercambio), la proporción de HFT en el mercado financiero global continúa aumentando (**Hoffmann**, 2014; **Biais; Foucault; Moinas**, 2015). Eso significa una transformación radical en la forma en que opera el mercado de valores. Mientras tanto, el “pool oscuro” (*dark pool*) desarrollado por HFT ha aumentado la opacidad de las operaciones del mercado (**Zupko**, 2021); la implementación de estrategias HFT ha suscitado públicamente preocupación sobre la equidad del mercado; y los choques ocasionales y los eventos disruptivos han planteado desafíos regulatorios (**Goldstein; Kumar; Tumbas**, 2014). Los investigadores deben estudiar, resumir y abordar las implicaciones de estos desarrollos.

**Gomber et al.** (2011) creían que la pregunta “¿El HFT es beneficioso o perjudicial para la economía?” era la más destacada en la investigación relacionada con HFT. Algunos investigadores han concretado puntos de vista importantes sobre el impacto de HFT en la calidad de los mercados financieros (**Chung; Lee**, 2016; **Linton; Mahmoodzadeh**, 2018; **Virgilio**, 2019), y otros han discutido a fondo las tecnologías y las reglas regulatorias involucradas en HFT y las han complementado con datos obtenidos de entrevistas (**Currie; Seddon**, 2016). Muchos investigadores han intentado categorizar la bibliografía académica sobre HFT, como **Gomber et al.** (2011), que consideran tres cuestiones principales: calidad del mercado, equidad y colocación, y penetración y rentabilidad del mercado. **Goldstein, Kumar y Graves** (2014) clasificaron la investigación sobre el comercio computarizado y de alta frecuencia en la última década en seis categorías: 1) desempeño del mercado, 2) estrategias y prácticas, 3) evolución, 4) velocidad, 5) equidad y 6) implicaciones regulatorias. Estos estudios han explorado la lógica de HFT para algunos problemas específicos basándose en bibliografía y en datos empíricos. Sin embargo, dado que estos estudios generalmente se enfocan a aspectos que no pueden analizar los conceptos relacionados de HFT de manera objetiva, y los datos utilizados son limitados en cantidad y variedad, los estudios existentes no brindan una imagen completa de la estructura del conocimiento en el campo de HFT. Con el desarrollo de la *cienciometría* se aplican más perspectivas de investigación y métodos analíticos para analizar los nuevos fenómenos y los nuevos problemas (**Serenko**, 2021) en *bibliotecología* (**Garcia; Lueck; Yakel**, 2019), *contabilidad* (**Lei; Deng; Liu**, 2020) y otras materias. Con la *cienciometría* se ha analizado la estructura del conocimiento y las tendencias en algunos campos de investigación emergentes, como *Bitcoin* (**Yu; Sheng**, 2020), *ciudades inteligentes* (**Moradi**, 2020), *cuidado de la salud en la nube* (**Gu et al.**, 2020), etc. Pero hasta donde sabemos, actualmente no hay ninguna revisión sistemática o análisis *cienciométrico* de HFT publicado en revistas<sup>1</sup>, ya sea en el campo de las finanzas o de la *biblioteconomía* y ciencias de la información, lo que ha traído ciertos obstáculos a la investigación de los investigadores y a la práctica de los profesionales.

Tomando HFT como un caso, este estudio se enfoca en revelar y “leer a distancia” su estructura del conocimiento y la evolución de los temas de investigación relacionados. En este estudio se utilizarán *NEViewer*, *Gephi* (**Bastian; Heymann; Jacomy**, 2009) y otros softwares de análisis visual de redes para representar cuantitativamente y analizar los artículos recopilados, y en definitiva el *statu quo* de HFT. Este estudio descubrirá de manera integral los micro detalles de investigación del tema, distinguirá la connotación de los conceptos relevantes y hará una inferencia sobre el estado de desarrollo de la investigación, lo cual será una referencia importante para investigadores y profesionales de la HFT. Se mostrará una nueva forma de obtener conocimientos usando una red de temas de investigación y analizando su dinámica y su estado de desarrollo. Específicamente, este estudio tomará HFT como un caso para explorar las siguientes preguntas:

Este estudio propone un nuevo método para obtener información sobre el conocimiento usando una red de temas de investigación, analizando la dinámica de distribución de la investigación por temas y el estado de desarrollo de la investigación

- ¿Cuál es la estructura de distribución y la dinámica de evolución de los temas de investigación en este campo específico?
- ¿Cómo explorar el desarrollo de la investigación del campo reflejado por los resultados del análisis anterior?

## 2. Metodología

### 2.1. Fundamento de esta investigación

Básicamente, este estudio propone un análisis de la evolución de los temas de investigación a partir de su red temática, como se muestra en la figura 1.

En el esquema de la figura 1 se incluyen las siguientes tres etapas:

#### Etapas 1: Recopilación y preprocesamiento de datos

Web of Science (WoS) es una de las bases de datos más utilizadas en el entorno académico, en la que Science Citation Index (SCI) y Social Sciences Citation Index (SSCI) son índices reconocidos internacionalmente para la evaluación y para obtener estadísticas científicas. Por lo tanto, utilizar artículos publicados en revistas indexadas por SCI/SSCI para analizar la evolución de los temas de investigación puede garantizar la validez y confiabilidad de los resultados. Específicamente, buscamos en la base de datos WoS con los términos seleccionados en el campo "Topic", delimitamos el año, el idioma, el tipo de documento y otras reglas de filtrado, y luego exportamos este conjunto básico de documentos. Cada documento incluye el número de acceso, el año de publicación, la fuente, el autor, la institución, las palabras clave del autor y el resumen, etc. Las palabras clave del autor suelen tener algunos problemas formales, como la morfología, ser abreviaturas y a veces contener caracteres especiales. Por lo tanto, es necesario completar el preprocesamiento de datos sobre las palabras clave del autor, como la lematización, la sustitución de siglas y la eliminación de stop words (palabras vacías) mediante un programa.

#### Etapas 2: Construcción de la red temática de investigación.

La red de temas de investigación se compone de redes continuas de co-palabras de múltiples períodos, lo que refleja la correlación y el contexto evolutivo de los temas de investigación. En esta etapa, en primer lugar, el conjunto de datos básico se divide en varios períodos mediante la combinación de las características del desarrollo del campo y el método de longitud constante con la misma distancia. Luego, las palabras clave únicas y las co-ocurrencias entre ellas en cada período se cuentan como nodos y bordes de la red de co-palabras, respectivamente. Finalmente, de acuerdo con los requisitos de NEViewer y Gephi, los datos de la red de palabras compartidas de cada período se guardan en formatos NWB y CSV, que se utilizan como entrada de la siguiente etapa.

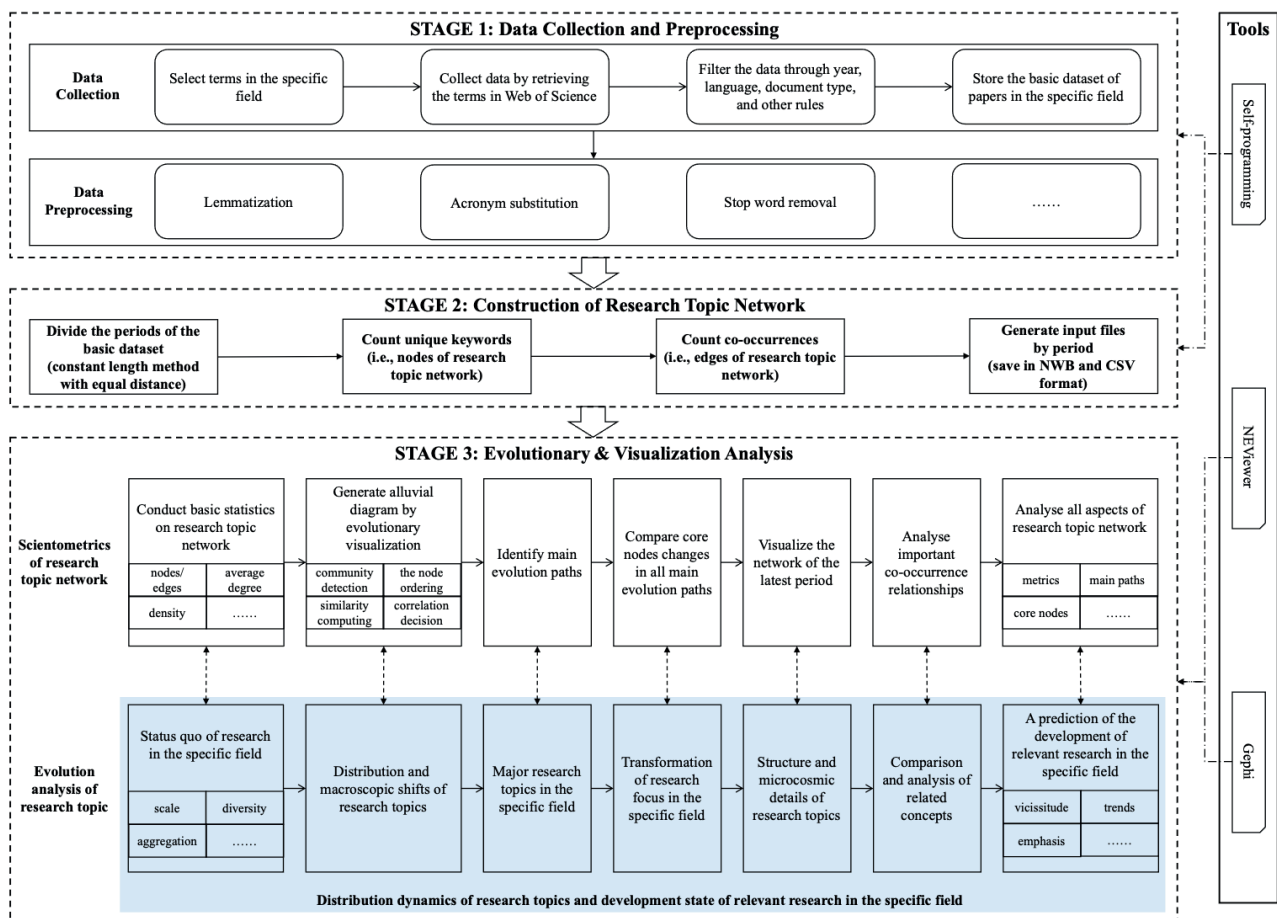


Figura 1. Esquema de la investigación



### Etapa 3: Análisis evolutivo y de visualización

Análisis cuantitativo diversificado de la red de temas de investigación y exploración de la evolución multidimensional del tema de investigación en el campo que se corresponden entre sí. Desde la perspectiva de las características básicas de la estructura de la red, la agregación y asociación de la comunidad, los nodos/bordes y otros elementos constituyentes, la red del tema de investigación se mide y analiza para explorar científicamente el *status quo* y las tendencias de desarrollo de la investigación en el campo, y para revelar visualmente el proceso macro y los detalles micro de la evolución de los temas de investigación. En esta etapa, la red de temas de investigación se mide científicamente a través de 7 pasos. Finalmente, se revela la dinámica de distribución del tema de investigación y se descubre el estado de desarrollo de la investigación.

- 1) A través de las estadísticas básicas de los indicadores de estructura, como nodos/bordes, grado promedio (ponderado) y densidad de la red de temas de investigación, se comprende la escala de desarrollo, la diversidad y la agregación de la investigación en el campo específico.
- 2) A través de la detección de la comunidad de temas de investigación, la ordenación de nodos, el cálculo de similitudes y la decisión de correlación en *NEViewer*, un programa de análisis de visualización evolutiva, se genera el diagrama aluvial evolutivo de la comunidad de temas de investigación, para visualizar la distribución de los temas de investigación y sus cambios macroscópicos (nacimiento, crecimiento, fusión, contracción, división y muerte).
- 3) Identificar las principales vías de evolución formadas por comunidades de temas de investigación a lo largo del diagrama aluvial, para conocer los principales temas de investigación en el campo específico.
- 4) Entre los principales caminos de evolución, el orden de los nodos en la comunidad de temas de diferentes períodos (especialmente los cambios de los nodos centrales) refleja la transformación del enfoque de investigación de temas relacionados. Por lo tanto, al comparar los cambios de los nodos centrales en todos los caminos de evolución principales, se revela la transformación del enfoque de investigación de los principales temas de investigación en el campo específico.
- 5) Analizar visualmente la estructura de la red de co-palabras en el último período mediante el software de exploración y visualización de redes *Gephi*, observar intuitivamente la relación de co-ocurrencia entre palabras clave relacionadas y luego explorar la estructura topológica y los detalles microcósmicos de los temas de investigación en el último período.
- 6) En la red de co-palabras, un mayor peso de borde significa una relación de co-ocurrencia más crítica entre las dos palabras clave relacionadas. Analizar estas importantes relaciones, comprender explícitamente la connotación y extensión de conceptos específicos y luego distinguir los conceptos relevantes en el campo.
- 7) Analizar exhaustivamente la situación general de la red de temas de investigación en términos de métricas estructurales, principales caminos de evolución, cambios en los nodos centrales, etc., para estudiar y juzgar el estado de desarrollo de la investigación relevante en el campo específico desde la perspectiva de la vicisitud, tendencias, énfasis, etc.

Después de estas etapas,

- se puede caracterizar cuantitativamente el *status quo* de la investigación en el campo,
- se pueden explorar sistemáticamente los principales temas de investigación y su proceso de cambios macroscópicos, así como la transferencia del enfoque de investigación reflejado por los dos,
- los detalles microscópicos de la investigación en el campo y la connotación de conceptos relacionados pueden analizarse claramente, y
- las tendencias de desarrollo futuras de la investigación relacionada pueden predecirse razonablemente.

En resumen, la idea propuesta en este estudio realizará el análisis dinámico de la distribución de los temas de investigación y el descubrimiento del estado de desarrollo de la investigación relacionada en el campo específico, para obtener algunas ideas de conocimiento.

### 2.2. Preparación de los datos

A través de la función de investigación avanzada en *WoS*, los términos de HFT se seleccionan de la siguiente manera.

- 1) “Comercio de alta frecuencia” y “comercio de alta frecuencia” se utilizan como “Tema” para la recuperación, luego se recopilan las palabras clave del autor incluidas en los artículos buscados.
- 2) Los términos de las materias están incluidos en las revisiones de la bibliografía relevante (**Chung; Lee, 2016; Linton; Mahmoodzadeh, 2018; Virgilio, 2019**).

En cuanto al conjunto de reglas de filtrado, los documentos se restringen a trabajos de investigación y revisiones en inglés de 2000 a 2019 en este caso. Por este medio, los artículos básicos recopilados en el campo de HFT se recuperaron en *SCI/SSCI* y se almacenaron (incluidos 20975 artículos científicos<sup>2</sup>), luego se preprocesaron las palabras clave de los autores.

Combinando la línea de tiempo de HFT (**Budish; Cramton; Shim, 2015; O’Hara, 2015**) y el método de longitud constante con igual distancia, la duración se divide en cuatro periodos de longitud equidistantes de 5 años, a saber:

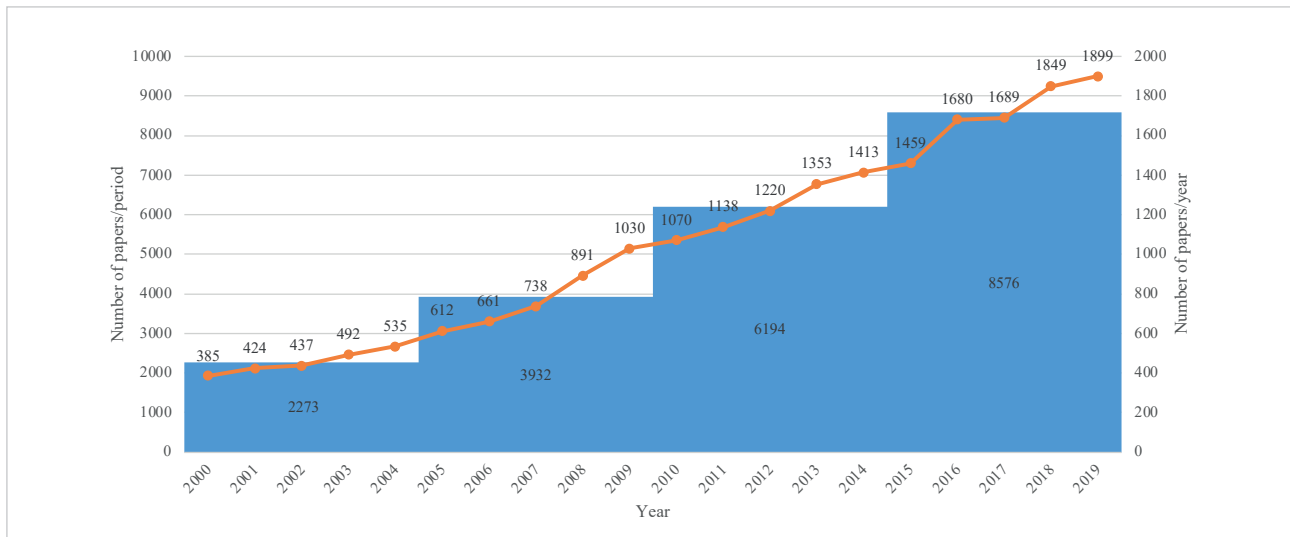


Figura 2. Distribución de artículos científicos en HFT

- 2000-2004 (Período I)
- 2005-2009 (Período II)
- 2010-2014 (Período III), y
- 2015-2019 (Período IV).

Las palabras clave únicas y su relación de co-ocurrencia se cuentan a través de la autoprogramación, y los datos de la red de temas de investigación se guardan en los formatos requeridos por período. El número de artículos que varía con el tiempo se muestra en la figura 2. El número de artículos científicos relevantes en HFT aumenta año tras año. De acuerdo con la ley del crecimiento de la bibliografía científica y el modelo logístico de la bibliografía científica (Wang *et al.*, 2021), HFT aún se está desarrollando. Con el creciente número de artículos, la connotación y la extensión de HFT se enriquecen constantemente. Por lo tanto, la revelación y la “lectura a distancia” de la estructura de conocimiento de HFT y el contexto evolutivo de los temas de investigación son necesidades urgentes para los investigadores y profesionales de HFT, y son de gran importancia para su desarrollo futuro.

### 3. Resultados

Tomando HFT como un caso, en esta sección se presentarán tres resultados principales del análisis evolutivo del tema de investigación, y habrá una breve explicación de la entrada de datos y el conjunto de herramientas.

#### 3.1. Estadísticas básicas de la red del tema investigado

Este estudio construyó una secuencia de red de co-palabras clave no dirigida basada en las palabras clave procesadas del autor de cuatro períodos para formar la red de temas de investigación. La tabla 1 presenta los indicadores estadísticos básicos para la red del tema investigado. Nodes y Edges miden el tamaño de la red e indican el número de nodos y bordes, respectivamente. La densidad utilizada para medir la integridad de la red se calcula mediante la ecuación:

$$2 * \text{Edges} / (\text{Nodes} * (\text{Nodes} - 1))$$

El Average Degree (grado medio) del nodo representa el valor medio de todos los bordes conectados del nodo en la red. El Average Weight Degree (grado de ponderación promedio) adicional considera la ponderación de la ventaja (es decir, el número de coincidencias entre las palabras clave de autor en este caso). El Clustering Coefficient (coeficiente de agrupamiento) muestra un signo general de agrupamiento de nodos en la red, mientras que los Connected Components (componentes conectados) se refieren al número de subgráficos separados en la red.

El análisis de la tabla 1 y la distribución del conjunto de datos básicos revelan lo siguiente:

- Con el creciente número de artículos científicos en HFT desde el Período I hasta el Período IV, los Nodos y los Bordes aumentan en consecuencia, pero la Densidad muestra una tendencia a la baja. Significa que el creciente número de artículos relacionados está acompañado por la ampliación gradual del tamaño de la correspondiente red de co-palabras clave, y los temas de investigación en el campo muestran una mayor diversidad pero más dispersos.
- El grado promedio (ponderado) de cuatro redes de co-palabras clave aumenta con el tiempo, lo que significa conexiones más frecuentes y relaciones de co-ocurrencia entre los temas de investigación en HFT. Sin embargo, el coeficiente de agrupación promedio muestra cambios fluctuantes, lo que indica que su grado general de agregación no conduce a una tendencia clara.
- Los componentes conectados de cuatro redes de palabras clave compartidas cambian en un tipo “V”. El Período II tiene el componente más bajo pero la tasa de crecimiento anual más alta en el tamaño de la red de palabras clave

Tabla 1. Estadísticas básicas de la red del tema de investigación

Índice	2000-2004	2005-2009	2010-2014	2015-2019
Nodos	675	1.583	2.793	4.067
	--	(134,5%)	(76,4%)	(45,6%)
Bordes	2.607	8.224	19.883	32.299
	--	(215,4%)	(141,7%)	(62,4%)
Grado medio	7,724	10,390	14,238	15,883
Grado medio ponderado	9,541	12,423	17,607	19,380
Densidad	0,01146	0,00657	0,00510	0,00391
Componentes conectados	14	8	11	15
Coefficiente de agrupamiento promedio	0,398	0,375	0,444	0,435

Nota: Los números entre paréntesis en la tabla son las tasas de crecimiento interanual de los indicadores correspondientes sobre el período anterior.

compartidas. Se puede inferir que el Período II tuvo una explosión en la cantidad de estudios en comparación con el Período I, y los investigadores tuvieron conocimiento del estudio del período anterior y llevaron a cabo muchas investigaciones de seguimiento. Estos estudios mostraron inicialmente tendencias sistemáticas y connotaciones claras. Los estudiosos difundieron aún más estas connotaciones en los dos períodos siguientes.

### 3.2. Distribución y evolución de los temas de investigación

Las redes de palabras clave compartidas de cuatro períodos se importaron al conjunto de herramientas *NEViewer* para la visualización y el análisis evolutivo. Utilizando el algoritmo de partición de Blondel y el diseño Radical-Tree, se configuraron los nodos que miran hacia atrás. El diagrama aluvial resultante se muestra en la figura 3.

La figura 3 muestra que la evolución de la comunidad de temas de investigación en el campo de HFT tiene un alto grado de continuidad a lo largo de cuatro períodos, con algunas comunidades dividiéndose o fusionándose en general. Específicamente,

- a) el “mercado incompleto” se dividió en “riesgo sistemático” y “optimización de cartera” en el Período II, y este último se dividió aún más en el Período III.
- b) Tanto la “volatilidad” como el “mercado financiero”, que ocuparon un lugar alto y tenían nodos grandes en el Período IV, se separaron del “mercado financiero” en el Período III.
- c) Además de la división en el Período II y el Período III, el “estudio de eventos” también se estaba fusionando en el Período IV, y es la única convergencia en todo el diagrama aluvial.

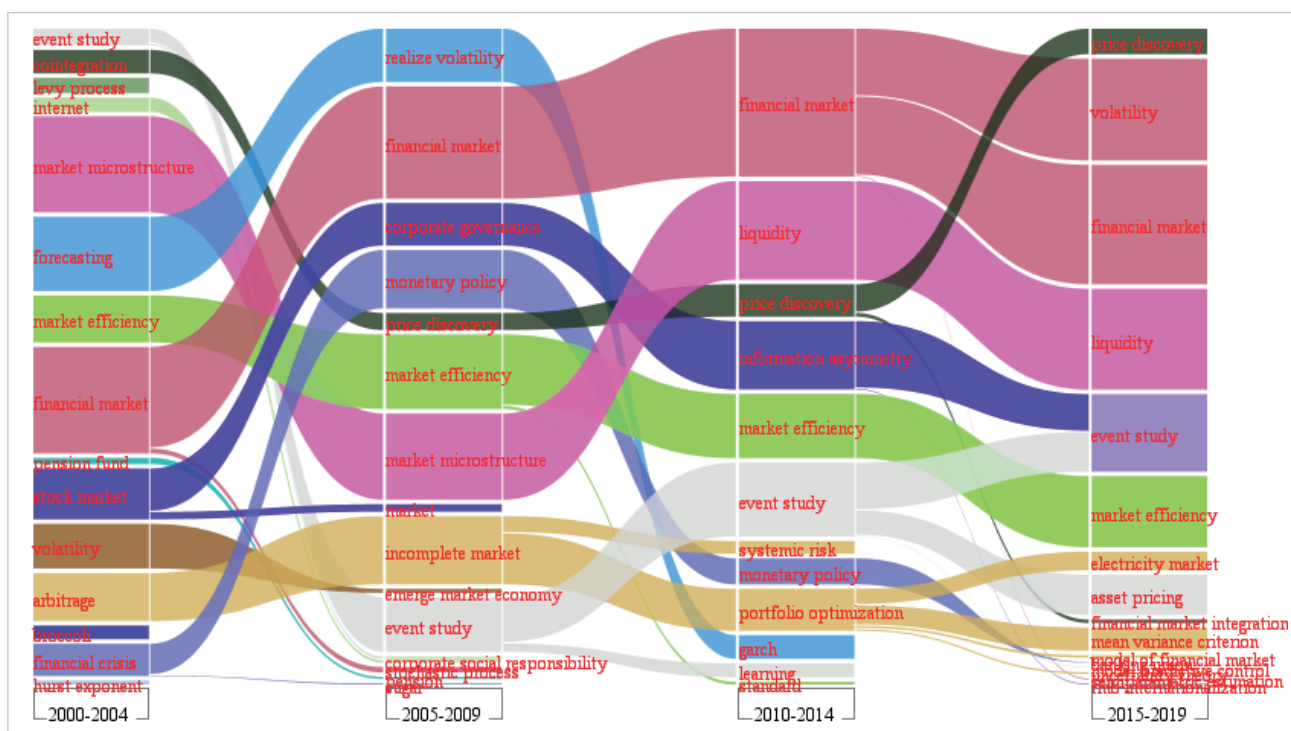


Figura 3. Diagrama aluvial de la evolución de la comunidad en los temas de investigación en HFT (umbral de similitud de la comunidad igual a 0,2)<sup>3</sup>

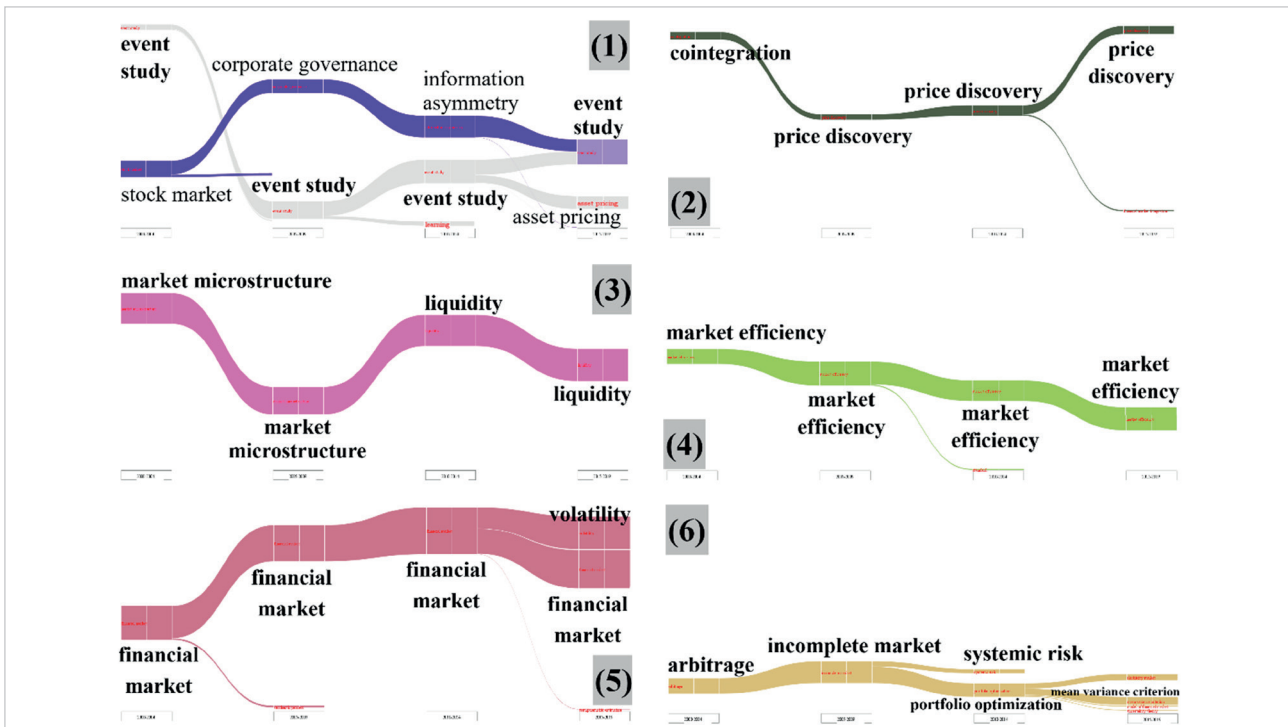


Figura 4. Principales vías de evolución de los temas de investigación en HFT

Hay seis rutas principales en la figura 3, a saber,

- 1) estudio de eventos,
- 2) cointegración y descubrimiento de precios,
- 3) liquidez y microestructura del mercado,
- 4) eficiencia del mercado,
- 5) mercado financiero, y
- 6) mercado incompleto, como se muestra en la figura 4.

En la salida de *NEViewer*, Z-Value mide qué tan cerca se relaciona un nodo con otros en esa comunidad. Es un indicador que revela la importancia de los nodos a nivel local en lugar de global. Para las comunidades relacionadas en las rutas de evolución principales anteriores, sus nodos se clasifican por Z-Value y los nodos centrales correspondientes se muestran en la tabla 2.

### 3.3. Estructura y análisis de la red de co-palabras clave del último período

Para profundizar la información adecuada en la red de temas de investigación, este estudio importó los datos CSV del Período IV a *Gephi*. En *Gephi*, este estudio llevó a cabo las siguientes operaciones:

- 1) seleccionar la función de modularización para agrupar los nodos en la red,
- 2) la reproducción de color se realizó de acuerdo con la agrupación de los nodos,
- 3) filtrar los nodos con valores superiores a 42 (6,5%),
- 4) ajustar el tamaño de los nodos según el grado, y
- 5) seleccionar “Fruchterman Reingold” para ajustar el diseño de los nodos y completar la visualización final.

La estructura de la red de palabras clave compartidas en el Período IV se muestra en la figura 5.

Comparando la figura 5 con la figura 3, se puede observar que el orden del tamaño de los nodos en la figura 5 es el mismo que el tamaño del bloque de color del diagrama aluvial. Las comunidades con más nodos como “mercado financiero”, “volatilidad”, “liquidez”, “estudio de eventos” y “eficiencia del mercado” en la figura 3 se han distribuido específicamente en la figura 5. Este estudio contó los pesos (es decir, el número de co-ocurrencias) de los 10 principales en la red y sus correspondientes nodos de co-ocurrencia, que se muestran en la tabla 3.

Tomando HFT como un caso, este estudio se enfoca en revelar y “leer a distancia” la estructura del conocimiento de los campos temáticos y el contexto evolutivo de los temas de investigación



Tabla 2. Nodos centrales en 6 caminos de evolución principales de temas de investigación en HFT

	2000-2004	2005-2009	2010-2014	2015-2019
1)	<b>estudio de eventos, mercado financiero internacional</b> , comercio electrónico, <b>valoración de activos, información y eficiencia del mercado</b> , rendimiento anormal, fusiones y adquisiciones, seguimiento, valor de mercado, regulación del mercado laboral	<b>gobierno corporativo</b> , regulación, divulgación, fusión, fusión y adquisición, adquisición, costo de capital, bienes raíces, señalización	asimetría de la información, <b>gobierno corporativo</b> , <i>abuso de información privilegiada, adquisición, fusión, fusión y adquisición</i> , reestructuración, negociación informada, oferta pública inicial, adquisición	<b>estudio de eventos, gobierno corporativo</b> , rendimiento anormal, <b>información y eficiencia del mercado</b> , china, uso de <i>información privilegiada, adquisición, fusión y adquisición</i> , mercado emergente, <i>fusión</i>
	mercado de valores, crecimiento económico, <b>gobierno corporativo</b> , globalización, integración financiera, mercado, país en desarrollo, valor para el accionista, intermediación, rendición de cuentas	<b>estudio de eventos, valoración de activos, información y eficiencia del mercado</b> , <i>volumen de negociación, mercado financiero internacional, elección de cartera, tasa de interés de los bonos</i> , creencia heterogénea, capm, <i>general</i>	<b>estudio de eventos, información y eficiencia del mercado</b> , <b> fijación de precios de activos</b> , <i>volumen de operaciones, tasa de interés de los bonos, general, mercado financiero internacional, elección de cartera, decisiones de inversión</i> , mercado financiero general	<b>precios de activos</b> , <i>elección de cartera, volumen de negociación, tasa de interés de los bonos, decisiones de inversión, mercado financiero internacional</i> , política monetaria, tipo de cambio, mercado financiero y macroeconomía, modelo de series de tiempo
2)	<b>cointegración, descubrimiento de precios</b> , autorregresión vectorial, equilibrio, exogeneidad, corrección de errores, modelo de corrección de errores, diseño de mecanismos, estacionalidad, protocolo de negociación	<b>descubrimiento de precios</b> , integración de mercado, <b>cointegración</b> , Japón, <b>causalidad granger</b> , autorregresión vectorial, consumo, precio de las acciones, desbordamiento, liberalización	<b>descubrimiento de precios, cointegración</b> , mercado futuro, china, causalidad, derrame de volatilidad, mercado de materias primas, intercambio de información, materia prima, <b>causalidad granger</b>	<b>descubrimiento de precios, cointegración</b> , intercambio de información, futuro, comercio de pares, etf, futuro de productos básicos, <b>causalidad granger</b> , fondo de comercio de intercambio, proceso de ornstein uhlenbeck
3)	<b>microestructura del mercado, liquidez</b> , asimetría de la información, diferencial entre <b>oferta y demanda</b> , uso de información privilegiada, diferencial, creador de mercado, tamaño de tick, calidad de ejecución, microestructura	<b>microestructura del mercado, liquidez, diferencial entre oferta y demanda</b> , asimetría de información, microestructura, calidad del mercado, pedido límite, libro de pedidos límite, flujo de pedidos, costo de negociación	<b>liquidez, microestructura del mercado</b> , datos de alta frecuencia, <b>diferencial entre oferta y demanda</b> , volatilidad, <b>libro de órdenes límite</b> , volatilidad real, impacto en el precio, ruido de la microestructura del mercado, costo de transacción	<b>liquidez, microestructura del mercado</b> , negociación de alta frecuencia, <b>libro de órdenes con límite</b> , asimetría de información, <b>diferencial entre oferta y demanda</b> , costo de transacción, negociación algorítmica, impacto en el precio, datos de alta frecuencia
4)	<b>eficiencia del mercado</b> , mercado de capitales, <b>finanzas conductuales</b> , gestión de ganancias, divulgación, psicología del inversionista, aprendizaje, acumulación, regulación de cuentas, atención limitada	<b>eficiencia del mercado, finanzas conductuales</b> , caminata aleatoria, economía experimental, mercado de capitales, riesgo, pronóstico del analista, precios erróneos, reacción exagerada, anomalía	<b>eficiencia del mercado, finanzas del comportamiento</b> , hipótesis del mercado eficiente, no linealidad, previsibilidad, anomalía, análisis fundamental, venta al descubierto, hipótesis del mercado adaptativo, riesgo	<b>eficiencia del mercado</b> , análisis técnico, <b>finanzas conductuales</b> , estrategia comercial, previsibilidad de retorno, impulso, hipótesis de mercado eficiente, aprendizaje automático, criptomoneda, bitcoin
5)	<b>mercado financiero</b> , <i>econofísica</i> , proceso estocástico, comportamiento de rebaño, fat tail, economía, gráfico aleatorio, percolación, negocios, heavy tail	<b>mercado financiero</b> , <i>econofísica</i> , volatilidad, mercado de valores, <b>crisis financiera</b> , banco, mercado emergente, exponente de hurst, eficiencia, crecimiento económico	<b>mercado financiero</b> , <i>crisis financiera</i> , <i>mercado de valores, regulación</i> , econofísica, finanzas, crisis, institución, <i>neoliberalismo</i> , <i>globalización</i>	<b>volatilidad</b> , <i>mercado de valores</i> , previsión, garch, realización de volatilidad, tipo de cambio, precio del petróleo, rentabilidad de las acciones, volatilidad desbordante, desbordamiento <b>mercado financiero</b> , <i>financiarización, crisis financiera, regulación, finanzas, unión europea, innovación, neoliberalismo</i> , desarrollo financiero, crecimiento económico
6)	<b>arbitraje, mercado incompleto</b> , costo de transacción, cobertura, maximización de la utilidad, medida de martingala equivalente, medida de martingala, seguro, reclamo contingente, fijación de precios	<b>mercado incompleto</b> , maximización de la utilidad, costo de transacción, cobertura, abuso de información privilegiada, <i>ecuación diferencial estocástica hacia atrás</i> , arbitraje, cadena de markov, <i>ecuación de hamilton jacobi bellman</i> , <i>gravamen proceso</i>	<b>riesgo sistémico</b> , innovación financiera, titularización, estabilidad financiera, riesgo de crédito, derivado, regulación bancaria, educación financiera, desempeño, prima de capital <b>optimización de cartera</b> , valor en riesgo, <b>mercado incompleto</b> , <i>ecuación diferencial estocástica hacia atrás</i> , gestión de riesgos, optimización, comercio algorítmico, <i>proceso de levy, programación dinámica, ecuación de hamilton jacobi bellman</i>	<b>criterio de varianza media</b> , control estocástico, inversión, <i>ecuación de hamilton jacobi bellman</i> , información parcial, inversión óptima, varianza media, <i>programación dinámica, mercado incompleto</i> , consistencia temporal

La letra en rojo en la tabla representa la etiqueta de la comunidad correspondiente a las figuras 3 y 4; la negrita indica el nodo que aparece en los cuatro periodos en la misma ruta; la cursiva significa que el nodo está involucrado en la división o fusión de la comunidad, y el subrayado denota el nodo en la misma ruta que aparece en los primeros tres periodos o en los últimos tres periodos

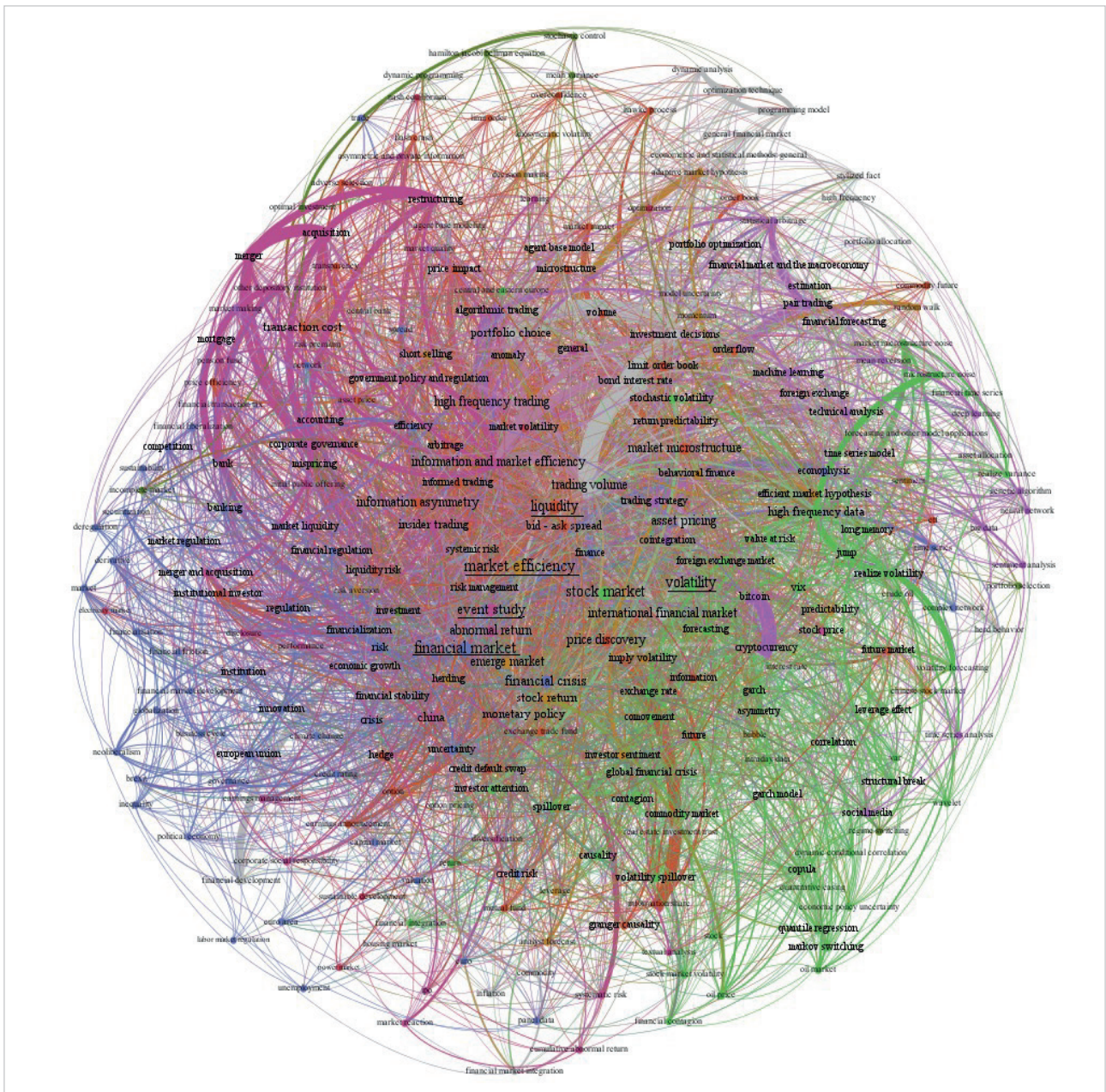


Figura 5. Topología de red de palabras clave compartida de HFT en 2015-2019

Tabla 3. Top 10 de co-ocurrencia en la red de co-palabras del último periodo

Rango	Nodo A	Nodo B	Peso
1	estudio de eventos	información y eficiencia del mercado	69
2	fijación de precios de activos	tasa de interés de los bonos	41
3	estudio de eventos	regreso anormal	40
4	fijación de precios de activos	volumen de comercio	39
5	volumen de comercio	tasa de interés de los bonos	37
6	elección de cartera	decisiones de inversión	32
7	comercio de alta frecuencia	comercio algorítmico	29
8	bitcoin	criptomoneda	27
9	estudio de eventos	volumen de comercio	24
	descubrimiento de precios	compartir información	
10	fijación de precios de activos	información y eficiencia del mercado	23

HFT y AT (*algorithmic trading*) tienen la séptima clasificación más alta de co-ocurrencias en la red de co-palabras clave del Período IV. Como se discutió en la introducción sobre HFT y AT, es necesaria una distinción más profunda entre los dos para

la investigación y la práctica relevantes. Para distinguir los dos conceptos de manera objetiva, los tiempos de co-ocurrencia de ellos con otros nodos se clasificaron en orden descendente, y los 10 principales se muestran en la tabla 4.

Tabla 4. Nodos centrales en co-ocurrencias con HFT o AT

Nodo A	Nodo B	Peso	Nodo A	Nodo B	Peso
comercio de alta frecuencia	comercio algorítmico	29	comercio algorítmico	comercio de alta frecuencia	29
	<b>liquidez</b>	17		<b>liquidez</b>	5
	microestructura del mercado	13		regulación del mercado	4
	<b>libro de órdenes limitadas</b>	12		red neuronal	4
	<b>calidad del mercado</b>	9		aprendizaje automático	4
	<b>eficiencia del mercado</b>	8		<b>eficiencia del mercado</b>	4
	mercado financiero	8		<b>calidad del mercado</b>	4
	<b>impacto en el precio</b>	6		<b>libro de órdenes limitadas</b>	4
	accidente repentino	6		<b>impacto en el precio</b>	3
	arbitraje estadístico	6		modelo base de agente	3

Nota: La negrita indica que los nodos aparecen en ambas relaciones de co-ocurrencia.

Dado 1) el estado actual de desequilibrio de la investigación y la práctica en HFT entre los mercados financieros emergentes y los internacionales, 2) la discusión social en curso sobre la relación de HFT con los eventos de choque repentino, 3) el hecho de que las finanzas conductuales siempre sirven como nodo central en los caminos principales (algo tratado por varios ganadores del Premio Nobel de Economía), 4) el continuo interés de personas relevantes por la precisión y puntualidad de los pronósticos sobre volatilidad, tendencias, precios, etc., este caso también descubrió los nodos que con frecuencia co-ocurren, como se muestra en la tabla 5.

Tabla 5. Algunas relaciones co-ocurrentes importantes

Nodo A	Nodo B	Nodo A	Nodo B
mercado emergente	la eficiencia del mercado	mercado financiero internacional	fijación de precios de activos
	liquidez		elección de cartera
	mercado financiero internacional		información y eficiencia del mercado
	estudio de eventos		tasa de interés de los bonos
	fijación de precios de activos		estudio de eventos
	microestructura del mercado		volumen de comercio
accidente repentino	comercio de alta frecuencia	aprendizaje automático	comercio algorítmico
	libro de órdenes limitadas		arbitraje estadístico
	modelo base de agente		grandes datos
	riesgo sistémico		comercio de alta frecuencia
	pastoreo		pronóstico
finanzas conductuales	sentimiento de los inversores	aprendizaje profundo	estrategia comercial
	eficiencia del mercado		asignación de cartera
	crisis financiera		aprendizaje automático
	mercado financiero		predicción de volatilidad
	elección de cartera		comercio de alta frecuencia
mercado eficiente	comercio algorítmico		
pronóstico	volatilidad	grandes datos	estrategia comercial
	darse cuenta de la volatilidad		aprendizaje automático
	bolsa de Valores	red neuronal	comercio algorítmico
medios de comunicación social	análisis de sentimientos		pronóstico



## 4. Discusión

### 4.1. Estructura de distribución y dinámica de evolución de los temas de investigación

El análisis de visualización de la red de temas de investigación puede revelar la estructura del conocimiento del campo específico en períodos y el contexto evolutivo de los temas de investigación en este campo. Durante 2000-2019, la investigación en HFT se centró en el estudio de eventos, la cointegración y el descubrimiento de precios, la liquidez y la microestructura del mercado, la eficiencia del mercado, el mercado financiero, el mercado incompleto y otros temas de investigación.

En comparación con otras rutas principales, la ruta de “estudio de eventos” tiene el proceso de evolución más significativo de división y fusión. Los investigadores han realizado un análisis extenso de la cotización de activos y la información y la eficiencia del mercado en los mercados financieros internacionales mediante el método de estudio de eventos. Hay un sub-camino de “gobernanza corporativa”, donde la investigación relacionada se concentra en la situación financiera, las regulaciones, la divulgación, el abuso de información privilegiada y la fusión y adquisición de las empresas que cotizan en bolsa. A través de la dinámica de distribución del nodo de gobierno corporativo y la comparación de los nodos centrales relacionados en los Períodos I y IV, se puede encontrar que el enfoque de la investigación relacionada se ha transferido gradualmente de los mercados financieros internacionales a los mercados emergentes como China.

El tamaño del nodo de las comunidades relacionadas en la ruta de “cointegración y descubrimiento de precios” es más pequeño, y sus nodos centrales no cambian mucho durante cuatro períodos, lo que indica que los temas de investigación relacionados están más concentrados. Además, el “descubrimiento de precios” se convirtió en la principal comunidad del diagrama aluvial en el Período IV, lo que sugiere que fue la más relacionada con otros temas de investigación de HFT en 2015-2019. Este camino involucra algunos métodos de análisis financiero matemático y econométrico de descubrimiento y formación de precios. Los focos han cambiado gradualmente de la autorregresión vectorial, el método de equilibrio y el modelo de corrección de errores a la causalidad, el modelo de intercambio de información y el proceso estocástico. Durante 2000-2014, las externalidades, el diseño de mecanismos, los estudios de efectos indirectos (volatilidad) y otros temas estuvieron en el centro de la investigación relacionada, mientras que en 2015-2019, los estudios sobre la práctica y aplicación del comercio de pares y derivados financieros como futuros y mercado de valores –los fondos negociados estaban calientes. Además, China estuvo a la vanguardia de los nodos centrales durante el Período III, lo que indica que el descubrimiento de precios en los mercados emergentes recibió una atención adicional por parte de la academia en 2010-2019.

El camino de la “microestructura y liquidez del mercado” es el único sin división ni fusión, y sus temas relacionados han formado una gama de investigación más independiente. *Bid ask spread* recibió una atención constante por parte de los estudiosos, pero su entusiasmo había disminuido ligeramente en el Período IV en comparación con otros períodos. Puede ser el resultado de un aumento en la popularidad de la orden limitada, HFT y otros problemas en 2015-2019. En 2000-2014, la investigación más importante se centró en los conceptos de microestructura de mercado, como creador de mercado, *tick size*, orden límite, flujo de órdenes, costo de negociación, datos de alta frecuencia y costo de transacción, lo que puede ser una de las razones del aumento. al calor de conceptos como HFT durante 2015-2019. Además, los investigadores observaron atentamente los conceptos a nivel de transacción, como el uso de información privilegiada, el libro de órdenes limitadas y el costo de transacción, y realizaron investigaciones relacionadas con la asimetría de la información. Aunque algunos nodos, como la asimetría de la información, han surgido en la ruta secundaria del estudio de eventos, subrayaron conceptos a nivel de empresa, como la gobernanza corporativa y fusión-adquisición. Además, la volatilidad y el ruido de la microestructura del mercado también atrajeron la atención de la academia entre 2010 y 2014.

En el camino de la “eficiencia del mercado”, la centralidad de la comunidad se reduce con el tiempo, es decir, la interacción entre la eficiencia del mercado y otros temas de investigación en HFT está disminuyendo. Esto es contrario al aumento de la centralidad de las comunidades relacionadas en el camino del “mercado financiero”, lo que puede indicar que su evolución tiene una correlación moderada. Dado que el comportamiento de los participantes del mercado afecta la eficiencia del mercado, se ha realizado una amplia investigación sobre la psicología del inversor, la atención limitada, el caminar aleatorio, la reacción exagerada, la anomalía, la hipótesis del mercado eficiente, la hipótesis del mercado adaptativo y otros temas desde la perspectiva analítica de las finanzas del comportamiento. En 2000-2004, los investigadores prestaron atención a los conceptos relacionados con la contabilidad financiera, como la gestión de ganancias, la divulgación, la acumulación y la regulación contable. En consecuencia, analizaron el impacto de cambios en factores como el comportamiento financiero y la regulación contable de las empresas cotizadas en la eficiencia del mercado. Durante 2005-2009, los métodos de economía experimental recibieron la atención de los investigadores. Se han realizado varios estudios relacionados con la previsión en 2010-2019, incluida la previsión de analistas, la previsibilidad (de retorno), etc. Además, se pudo encontrar que el enfoque dominante de la previsión está cambiando gradualmente del análisis fundamental al análisis técnico. La investigación general durante 2015-

Seis temas principales en HFT: liquidez y microestructura del mercado, eficiencia del mercado, mercado financiero, mercado incompleto, cointegración y descubrimiento de precios, y estudio de eventos

2019 está sesgada hacia la práctica en el campo de HFT, con conceptos como estrategia comercial, impulso, aprendizaje automático, criptomoneda, *bitcoin* y algoritmo genético ganando popularidad.

El tamaño del nodo de las comunidades relevantes del “mercado financiero” es el más grande de todas las redes de palabras clave compartidas, y su centralidad en la red de temas de investigación también es relativamente mayor. Sin embargo, tal vez por el concepto amplio de mercado financiero, no se advierten en este camino otros conceptos de interés sostenido. De 2000 a 2004, los investigadores se preocuparon por los conceptos de características matemáticas como el proceso estocástico, la cola gruesa y la cola pesada. Considerando que en los siguientes 15 años se prestó más atención a conceptos de entornos de mercado como crisis financiera, mercado emergente, mercado de valores, crecimiento económico, desarrollo financiero y globalización. La econofísica fue el nodo central para las comunidades relevantes en los tres primeros períodos, pero no se mantuvo en el cuarto. La razón puede ser que los investigadores prestaron más atención a la aplicación de los logros de la econofísica en el Período IV (por ejemplo, la predicción de la variación de la volatilidad del mercado debido a las crisis financieras por parte de la dinámica). En general, los nodos directamente relacionados con la volatilidad, como la volatilidad real y el desbordamiento de la volatilidad, solo aparecieron en el Período II y IV. Además, la “volatilidad” se separó de la pre-secuencia “mercado financiero”, se convirtió en la segunda comunidad más grande en el Período IV. En 2015-2019, la investigación relevante sobre la volatilidad se concentró en la medición, el pronóstico y el rendimiento de las acciones y los efectos indirectos asociados. Además, también incluyó el análisis del mercado de valores, tipo de cambio, precio del petróleo y otros objetos de investigación.

La centralidad general del “mercado incompleto” es la más baja entre los seis caminos principales, y sus comunidades asociadas se dividen dos veces en los Períodos II y III, lo que sugiere que los investigadores todavía lo están explorando. Las subdirecciones como el proceso estocástico, el control óptimo y la programación dinámica

El énfasis cambiará de la teoría a la práctica, las tecnologías (*big data*, etc.) y las teorías (*behavioral finance*, etc.) tendrán más interacción con HFT

en la investigación operativa se han utilizado ampliamente en la investigación del mercado incompleto. En este camino, la comunidad de “optimización de cartera”, que se dividió por primera vez, continuó la investigación relevante sobre procesos estocásticos y control óptimo de 2005 a 2009. Sus nodos principales son la ecuación diferencial estocástica hacia atrás, el proceso de exacción y la ecuación de Hamilton Jacobi Bellman, mientras que también se centra en la medición y gestión del riesgo. En 2010-2014, el “criterio de varianza media” de la segunda división siguió la ecuación media de Hamilton Jacobi Bellman y la programación dinámica. Además, la innovación financiera, la titulización, la estabilidad financiera, el riesgo crediticio, la educación financiera y otros temas atrajeron gran atención sobre el “riesgo sistémico” comunitario de la división en el Período III, que rara vez fueron cubiertos por la investigación relevante en otros períodos, probablemente debido a la repentina serie de eventos disruptivos del mercado durante 2010-2014.

#### 4.2. Estado de desarrollo de la investigación en el campo

Mediante un análisis integral de la escala, las rutas principales y los nodos centrales de la red de temas de investigación, se puede realizar un estudio y una evaluación razonables del estado de desarrollo de la investigación relevante en el campo específico. La investigación relacionada durante 2000-2019 en el campo de HFT se centra en la microestructura del mercado, la eficiencia del mercado, el mercado financiero, el mercado incompleto y otros temas de investigación a nivel de mercado, así como la cointegración y el descubrimiento de precios, la fijación de precios de activos y otros de nivel-formación-precios. El estudio de eventos es un método común en HFT. La volatilidad/riesgo es un concepto comúnmente involucrado en todos los temas de investigación anteriores. En los últimos años, el énfasis de la investigación en HFT ha cambiado gradualmente de la teoría a la práctica y ha prestado más atención a los temas relacionados con la innovación financiera, las finanzas conductuales y la investigación operativa. Además, además de la investigación del entorno general de los mercados financieros internacionales, los mercados emergentes como China han atraído cada vez más el interés de los investigadores en los últimos años, especialmente en los temas relacionados con el estudio de eventos y el descubrimiento de precios.

En este estudio también se observaron detalles adicionales de la red de co-palabras clave en el último período.

- En el campo de HFT, el precio de los activos, la tasa de interés de los bonos y el volumen de negociación están estrechamente relacionados entre sí.
- Entre los temas relacionados con AT (comercio algorítmico) y HFT, la liquidez es una de las principales preocupaciones de los investigadores. Mientras tanto, conceptos como el libro de órdenes límite, la calidad del mercado, la eficiencia del mercado y el impacto en el precio a menudo aparecen junto con AT y HFT. Mientras se discute AT, la investigación relevante se enfoca más en las reglas del mercado y la tecnología. En comparación, la investigación de HFT se concentra más en la microestructura del mercado.
- En los mercados emergentes, la investigación sobre la eficiencia del mercado es más prominente, mientras que en el caso de los mercados financieros internacionales, la investigación sobre la valoración de activos es la más preferida. Así, la investigación sobre mercados emergentes en el campo de HFT se inclina más al nivel de mercado, mientras que la investigación sobre los principales mercados financieros internacionales se inclina más al nivel de transacción.



- Además de HFT, la investigación sobre el accidente repentino implica riesgo sistémico y rebaño (efecto) y otros conceptos.
- La investigación relevante sobre finanzas conductuales en HFT examina principalmente su impacto en la eficiencia del mercado y el mercado eficiente.
- La previsión de volatilidad es más común en HFT que la previsión de precios y tendencias.
- La investigación relacionada con las redes sociales dentro de HFT se centró en el análisis de sentimientos. La idea común es analizar el sentimiento de los inversores y monitorear la información fundamental a través de los modelos de algoritmos basados en los datos de opinión pública en las redes sociales y luego aplicarlos a la estrategia comercial y al modelo de pronóstico para mejorar el retorno de la estrategia y la precisión del modelo.
- El aprendizaje automático y otras tecnologías emergentes se utilizan ampliamente en la investigación y la práctica de HFT, en lo que respecta al arbitraje estadístico, la previsión, la estrategia comercial, la asignación de cartera, etc. En cuanto a la investigación relacionada con la predicción, incluye la aplicación de tecnología de aprendizaje profundo y modelo de red neuronal. Además, el desarrollo y el cálculo de estrategias HFT utilizan técnicas como big data y computación de alto rendimiento, que pueden brindar información útil en una gran cantidad de flujo de información de manera rápida y oportuna.

Combinado con la discusión anterior, este estudio puede descubrir razonablemente el estado de desarrollo de la investigación y la práctica en HFT. La investigación en el campo de HFT aún está en desarrollo, y el recuento de artículos aumentará continuamente en los próximos años. La investigación a nivel de mercado sobre la microestructura del mercado, la eficiencia del mercado, el mercado financiero, el mercado incompleto y la investigación a nivel de formación de precios sobre cointegración y descubrimiento de precios, precio de activos, etc., siguen siendo los focos de futuras investigaciones en el campo de HFT. Estos estudios prestarán más atención a los mercados emergentes en el entorno general de los mercados financieros internacionales. Los logros en los campos de las finanzas conductuales, la econofísica y la investigación de operaciones, así como técnicas como big data, aprendizaje automático y aprendizaje profundo, se aplicarán aún más en la investigación y práctica de HFT, como pronósticos, formulación de estrategias comerciales, asignación de cartera y gestión de riesgos. La investigación relevante y la práctica de la predicción utilizarán métodos de análisis más técnicos y métodos de análisis fundamental combinados con factores macroscópicos, mesoscópicos y microscópicos en el futuro. Con una serie de eventos financieros históricos que ocurrieron en los últimos años, como el colapso de las acciones estadounidenses y la caída de los precios del petróleo, se espera que la investigación futura en HFT se concentre más en la consideración del riesgo sistémico en términos de innovación financiera y estabilidad del mercado.

El enfoque de investigación sobre los mercados se transfirió gradualmente de internacional a emergente, mientras tanto, atención continua a la volatilidad/ relacionada con el riesgo

## 5. Conclusiones

Este estudio propuso una idea general para el análisis de la evolución de los temas de investigación basada en la red de temas de investigación. Tomando el campo emergente de HFT como un caso, la dinámica de distribución del tema de investigación en HFT se descubre sistemáticamente, y luego se estudia y juzga exhaustivamente el estado de desarrollo de la investigación relevante, lo que demostró la efectividad de la idea propuesta.

Concretamente, se utilizan una variedad de métodos y herramientas de la cienciometría para analizar los 20975 artículos de HFT publicados en revistas indexadas por *SCI/SSCI* desde 2000 hasta 2019. Este estudio,

- caracterizó cuantitativamente el status quo de la investigación en el campo,
- exploró sistemáticamente los principales temas de investigación y su proceso de cambios macroscópicos, así como la transferencia del enfoque de investigación reflejado por los dos,
- analizó claramente los detalles microscópicos de la investigación en el campo y la connotación de conceptos relacionados, y
- razonablemente inferido y predijo el estado de desarrollo futuro de la investigación relacionada.

Además, este estudio ha encontrado que

1) la estructura de conocimiento de un campo específico reflejada por la distribución comunitaria de los temas de investigación, es decir, la investigación de HFT se centra en seis temas principales relacionados con el nivel de mercado y la formación de precios (liquidez y microestructura del mercado), eficiencia del mercado, mercado financiero, mercado incompleto, cointegración y descubrimiento de precios, y estudio de eventos).

1) El contexto evolutivo y la transformación del enfoque de la investigación relacionada en el campo se refleja en las principales vías de evolución del tema de investigación y los cambios de sus nodos centrales. El enfoque de la investigación HFT se ha transferido gradualmente de los mercados financieros internacionales a los mercados emergentes como China, pero los estudios relacionados con la volatilidad/riesgo han sido una preocupación constante de la academia y la industria. El énfasis de la investigación en el campo de HFT ha ido cambiando de la teoría a la práctica y ha prestado más atención a los temas relacionados con la innovación financiera, las finanzas conductuales y la investigación de operaciones.

2) La estructura de la red en el último período refleja que el enfoque de la investigación futura en el campo HFT cambiará gradualmente de la teoría a la práctica. Las tecnologías relevantes, como los macrodatos, el aprendizaje automático y el aprendizaje profundo, y las teorías, como las finanzas del comportamiento, la econofísica y la investigación de operaciones, tendrán una mayor interacción con la investigación y la práctica de HFT en el futuro. La investigación relacionada y la práctica de la predicción en HFT utilizarán métodos de análisis más técnicos y métodos de análisis fundamental que combinan factores macroscópicos, mesoscópicos y microscópicos.

3) Los resultados de este caso indican que en los próximos años, los investigadores prestarán más atención a un campo de investigación emergente en finanzas-HFT, y la cantidad de artículos científicos relevantes continuará aumentando en consecuencia. La investigación en HFT continuará enfocándose en los temas relacionados con el nivel de mercado y la formación de precios y se concentrará más en la consideración del riesgo sistémico, mientras que el mayor desarrollo de los mercados emergentes en el entorno general del mercado financiero internacional atraerá más atención de la academia y de la industria.

La idea de investigación propuesta en este estudio, a través del análisis cuantitativo de la red de temas de investigación en el campo, revela efectivamente la dinámica de distribución de los temas de investigación y descubre científicamente el estado de desarrollo de la investigación relacionada en el campo específico. A partir de esto, se pueden obtener conocimientos y puntos de vista profesionales. Este documento se centra en la dimensión de producción en la cuantimetría sin considerar más el impacto, y existen limitaciones en el análisis de casos. Por un lado, la elección de los datos puede ser conservadora. Este caso solo analizó los artículos científicos publicados en las revistas *SCI* y *SSCI* y no consideró otra bibliografía científica, como artículos de conferencias, documentos de trabajo, monografías, etc. Además, hay algunas revistas que no requieren que los autores proporcionen palabras clave., por lo tanto, este caso no incluyó artículos en estas revistas. Por otro lado, aunque este caso ha adoptado una forma diversificada de seleccionar los términos en HFT, todavía no hay garantía de que no se hayan omitido algunos otros temas asociados con HFT, lo que puede limitar la profundidad del análisis y la discusión en este caso. hasta cierto punto. En general, la idea de investigación propuesta en este estudio proporciona una referencia eficaz para el desarrollo de servicios profesionales de conocimiento de la materia, como la guía de selección de temas, la recomendación de puntos críticos y el análisis de tendencias en la era de los grandes datos.

## 6. Notas

1. Observamos que un artículo de una conferencia (**Chou; Lee; Hung, 2016**) presenta los objetivos de un estudio que utiliza análisis bibliométrico de HFT sin contenido ni hallazgos específicos del estudio.

2. En la *WoS*, se obtuvieron 24250 artículos mediante la fórmula de búsqueda TS=(“high frequency trad\*” OR “algorithmic trad\*” OR “alternative trading” OR “alternative trading system” OR “electr\* trad\*” OR “dark pool\*” OR “dark liquidity” OR “market microstructure” OR “market quality” OR “pair\* trad\*” OR “flash\* trad\*” OR “flash crash” OR “insider trad\*” OR “tick size” OR “market maker\*” OR “market making” OR “informed trad\*” OR “bid-ask spread\*” OR “exchange server\*” OR “market volatility” OR “abnormal profit\*” OR “abnormal return\*” OR “order-to-trade ratios” OR “financial transaction tax” OR “Regulation NMS” OR “MiFID II” OR “minimum order resting time\*” OR “floor trading” OR “price discovery” OR “limit order book” OR “market efficiency” OR “flow toxicity” OR “financial market\*” OR “market regulation\*”) AND PY=2000-2019 AND LA= “English” AND DT=(“Article” OR “Review”) NOT DT= “Early Access”, y quedaron 21083 artículos después de la selección. Después de la inspección de datos, se identificaron y excluyeron 108 artículos de 2020, y el conjunto de datos final contenía 20975 artículos.

3. Nota: Los bloques de color rectangulares representan comunidades, la etiqueta roja es el nodo con el valor Z más alto en la comunidad y el tamaño del bloque de color indica la cantidad de nodos que contiene la comunidad. Las curvas aerodinámicas de colores entre los bloques en cada período muestran el proceso de división y fusión de una comunidad en un momento específico, lo que hace que se comprenda más vívidamente la trayectoria evolutiva y la tendencia de la dirección de la investigación. El orden vertical de los bloques de colores en el período indica la centralidad del tema de investigación, y la comunidad correspondiente en la parte superior está más estrechamente relacionada con otras comunidades..

## 7. Referencias

**Aldridge, Irene** (2009). *High-frequency trading: A practical guide to algorithmic strategies and trading systems*. Hoboken: John Wiley & Sons, Inc. ISBN: 978 0 470 56376 2

**Bastian, Mathieu; Heymann, Sebastien; Jacomy, Mathieu** (2009). “Gephi: An open source software for exploring and manipulating networks”. In: *Proceedings of the International AAAI Conference on web and social media*, pp. 361-362. <https://doi.org/10.13140/2.1.1341.1520>

**Biais, Bruno; Foucault, Thierry; Moinas, Sophie** (2015). “Equilibrium fast trading”. *Journal of financial economics*, v. 116, n. 2, pp. 292-313. <https://doi.org/10.1016/j.jfineco.2015.03.004>

- Budish, Eric; Cramton, Peter; Shim, John** (2015). "The high-frequency trading arms race: Frequent batch auctions as a market design response". *Quarterly journal of economics*, v. 130, n. 4, pp. 1547-1621.  
<https://doi.org/10.1093/qje/qjv027>
- Chou, Jerome-Chih-Lung; Lee, Mike Y. J.; Hung, Chia-Liang** (2016). "Bibliometric analysis of emerging trends in high frequency trading research". In: *Frontier computing*, pp. 985-991.  
[https://doi.org/10.1007/978-981-10-0539-8\\_96](https://doi.org/10.1007/978-981-10-0539-8_96)
- Chung, Kee H.; Lee, Albert J.** (2016). "High-frequency trading: Review of the literature and regulatory initiatives around the world". *Asia-Pacific journal of financial studies*, v. 45, n. 1, pp. 7-33.  
<https://doi.org/10.1111/ajfs.12120>
- Currie, Wendy L.; Seddon, Jonathan J. M.** (2016). "The regulatory, technology and market "dark arts trilogy" of high frequency trading: A research agenda". *Journal of information technology*, v. 32, n. 2, pp. 111-126.  
<https://doi.org/10.1057/s41265-016-0025-3>
- Garcia, Patricia; Lueck, Joseph; Yakel, Elizabeth** (2019). "The pedagogical promise of primary sources: Research trends, persistent gaps, and new directions". *Journal of academic librarianship*, v. 45, n. 2, pp. 94-101.  
<https://doi.org/10.1016/j.acalib.2019.01.004>
- Goldstein, Michael A.; Kumar, Pavitra; Graves, Frank C.** (2014). "Computerized and high-frequency trading". *Financial review*, v. 49, n. 2, pp. 177-202.  
<https://doi.org/10.1111/fire.12031>
- Gomber, Peter; Arndt, Björn; Lutat, Marco; Uhle, Tim** (2011). "High-frequency trading". *SSRN Electronic journal*, pp. 1-86.  
<https://doi.org/10.2139/ssrn.1858626>
- Gu, Dongxiao; Li, Tongtong; Wang, Xiaoyu; Yang, Xuejie; Yu, Zhangrui** (2019). "Visualizing the intellectual structure and evolution of electronic health and telemedicine research". *International journal of medical informatics*, v. 130, 103947.  
<https://doi.org/10.1016/j.ijmedinf.2019.08.007>
- Gu, Dongxiao; Yang, Xuejie; Deng, Shuyuan; Liang, Changyong; Wang, Xiaoyu; Wu, Jiao; Guo, Jingjing** (2020). "Tracking knowledge evolution in cloud health care research: Knowledge map and common word analysis". *Journal of medical internet research*, v. 22, n. 2, e15142.  
<https://doi.org/10.2196/15142>
- Guo, Huadong; Wang, Lizhe; Chen, Fang; Liang, Dong** (2014). "Scientific big data and digital Earth". *Chinese science bulletin*, v. 59, n. 12, pp. 1047-1054.  
<https://doi.org/CNKI:SUN:KXTB.0.2014-12-001>
- Hoffmann, Peter** (2014). "A dynamic limit order market with fast and slow traders". *Journal of financial economics*, v. 113, n. 1, pp. 156-169.  
<https://doi.org/10.1016/j.jfineco.2014.04.002>
- Howe, Doug; Costanzo, Maria; Fey, Petra; Gojobori, Takashi; Hannick, Linda; Hide, Winston; Hill, David P; Kania, Renate; Schaeffer, Mary; St Pierre, Susan; Twigger, Simon; White, Owen; Rhee, Seung Yon** (2008). "Big data: The future of biocuration". *Nature*, v. 455, n. 7209, pp. 47-50.  
<https://doi.org/10.1038/455047a>
- Huang, Han; Wang, Hongyu; Wang, Xiaoguang** (2020). "An analysis framework of research frontiers based on the large-scale open academic graph". In: *Proceedings of the Association for Information Science and Technology*, v. 57, n. 1, e307.  
<https://doi.org/10.1002/pr2.307>
- IBISWorld* (2012). *High frequency trading industry in the US - Market research report*.  
<https://www.ibisworld.com/industry-trends/specialized-market-research-reports/advisory-financial-services/intermediaries/high-frequency-trading.html>
- Lei, Lei; Deng, Yaochen; Liu, Dilin** (2020). "Examining research topics with a dependency-based noun phrase extraction method: a case in accounting". *Library hi tech*, ahead-of-print.  
<https://doi.org/10.1108/LHT-12-2019-0247>
- Li, Ping; Gong, Xuhong; Zhang, Qingchang** (2011). "Diffusion of technical knowledge based on international citations: Evidence from China". *Management world*, v. 12, pp. 21-31.  
<https://doi.org/10.19744/j.cnki.11-1235/f.2011.12.003>
- Linton, Oliver; Mahmoodzadeh, Soheil** (2018). "Implications of high-frequency trading for security markets". *Annual review of economics*, v. 10, n. 1, pp. 237-259.  
<https://doi.org/10.1146/annurev-economics-063016-104407>

**Moradi, Shima** (2020). "The scientometrics of literature on smart cities". *Library hi tech*, v. 38, n. 2, SI, pp. 385-398.  
<https://doi.org/10.1108/LHT-12-2018-0203>

**Moral-Muñoz, José A.; Herrera-Viedma, Enrique; Santisteban-Espejo, Antonio; Cobo, Manuel J.** (2020). "Software tools for conducting bibliometric analysis in science: An up-to-date review". *Profesional de la información*, v. 29, n. 1, e290103.  
<https://doi.org/10.3145/epi.2020.ene.03>

**O'Hara, Maureen** (2015). "High frequency market microstructure". *Journal of financial economics*, v. 116, n. 2, pp. 257-270.  
<https://doi.org/10.1016/j.jfineco.2015.01.003>

**Serenko, Alexander** (2021). "A structured literature review of scientometric research of the knowledge management discipline: A 2021 update". *Journal of knowledge management*, v. 25, n. 8, pp. 1889-1925.  
<https://doi.org/10.1108/JKM-09-2020-0730>

**Staff of the Division of Trading and Markets** (2014). "Equity market structure literature review. Part II: High frequency trading".  
[https://www.sec.gov/marketstructure/research/hft\\_lit\\_review\\_march\\_2014.pdf](https://www.sec.gov/marketstructure/research/hft_lit_review_march_2014.pdf)

**Virgilio, Gianluca-Piero-Maria** (2019). "High-frequency trading: A literature review". *Financial markets and portfolio management*, v. 33, n. 2, pp. 183-208.  
<https://doi.org/10.1007/s11408-019-00331-6>

**Wang, Kang; Gao, Jiping; Pan, Yuntao; Chen, Yue** (2021). "Research on the method of multi-position research themes recognition and evolution path". *Library and information service*, v. 65, n. 11, pp. 113-122.  
<https://doi.org/10.13266/j.issn.0252-3116.2021.11.012>

**Wang, Xiaoguang; Cheng, Qikai** (2013). "Analysis on evolution of research topics in a discipline based on NEViewer". *Journal of the China Society for Scientific and Technical Information*, v. 32, n. 9, pp. 900-911.  
<https://doi.org/10.3772/j.issn.1000-0135.2013.09.001>

**Wang, Xiaoguang; Wang, Hongyu; Huang, Han** (2021). "Evolutionary exploration and comparative analysis of the research topic networks in information disciplines". *Scientometrics*, v. 126, n. 6, pp. 4991-5017.  
<https://doi.org/10.1007/s11192-021-03963-6>

**Yu, Dejian; Sheng, Libo** (2020). "Knowledge diffusion paths of blockchain domain: The main path analysis". *Scientometrics*, v. 125, n. 1, pp. 471-497.  
<https://doi.org/10.1007/s11192-020-03650-y>

**Zupko, Robert** (2021). "Global algorithmic capital markets: High frequency trading, dark pools, and regulatory challenges [Book review]". *The social science journal*, Online first.  
<https://doi.org/10.1080/03623319.2021.2015910>

**Dialnet | métricas**  
Nueva edición 2020

**dialnet.unirioja.es/metricas**

Fundación Dialnet