

# Propuesta de Investigación

Convocatoria para Plaza de Investigadora Visitante  
Departamento de Matemáticas Aplicadas y Sistemas,  
Universidad Autónoma Metropolitana, Unidad Cuajimalpa

Dra. Daniela Aguirre Guerrero

---

El presente documento describe la propuesta de proyecto de investigación que llevaría acabo como Profesora Visitante del Depto. de Matemáticas Aplicadas y Sistemas de la Unidad Cuajimalpa. El proyecto involucra *ciencia de redes, ciencia de datos y aprendizaje automático*, temas de vanguardia en las Ciencias de la Computación.

## I. CARACTERIZACIÓN DE MODELOS DE EVOLUCIÓN DE REDES DE COLABORACIONES CIENTÍFICAS, APLICANDO APRENDIZAJE AUTOMÁTICO.

### A. PALABRAS CLAVE

Ciencia de redes, ciencia de datos, aprendizaje automático, análisis de redes sociales.

### B. INTRODUCCIÓN

Una de las aplicaciones más populares de la ciencia de redes consiste en el estudio de la evolución de redes complejas [1, Capítulo 6], tales como las redes de colaboraciones científicas, específicamente las redes de citaciones y de coautorías. Una particularidad de estas redes es que están sujetas a eventos externos que modifican su dinámica de evolución, es decir eventos disruptivos. Por ejemplo, un programa de estímulo a la investigación interdisciplinaria podría provocar la desaparición de comunidades especializadas en una disciplina [2]. Descubrir los mecanismos de evolución que subyacen en la dinámica de estas redes, permitiría predecir la formación y desaparición de grupos (comunidades) de colaboradores, el surgimiento de autores y tópicos *populares*, etc. De esta manera, la caracterización de modelos de evolución de dichas redes, permitiría estimar la respuesta de comunidades científicas ante nuevas políticas de estímulo a la investigación, o ante situaciones emergentes, como la pandemia de COVID-19.

El enfoque básico en el desarrollo de modelos de evolución de redes complejas, considera el uso de modelos de redes con enlaces sin pesos (redes no ponderadas) [3]. Sin embargo, el peso de los enlaces es una característica fundamental de las redes de coautorías, debido a que indica el nivel de colaboración entre dos autores. Más aun, recientes investigaciones muestran que, a diferencia de la redes no ponderadas, las redes multi-capa basadas en hipergrafos describen a mayor detalle los niveles de colaboración en redes científicas [4].

Sin embargo, aun es necesario probar la exactitud de estos modelos en redes susceptibles a eventos disruptivos. Lo anterior presenta un tema en la frontera del conocimiento de la ciencia de redes y ciencia de datos, el cual se refiere al desarrollo de modelos de redes de colaboraciones científicas basados en hipergrafos.

En este proyecto de investigación, se propone el desarrollo de modelos de la evolución de la redes de colaboraciones de científicas susceptibles a eventos disruptivos. Las colaboraciones científicas se modelarán mediante hipergrafos y posteriormente se aplicarán técnicas de aprendizaje automático para obtener funciones matemáticas de evolución de parámetros topológicos de la red, tales como centralidad de nodo, diámetro y densidad de la red, etc. Es importante mencionar que aunque es común aplicar aprendizaje automático a grandes volúmenes de datos relacionales, recientes estudios indican los beneficios de aplicar aprendizaje automático a redes complejas de gran escala [5, 6, 7, 8].

### C. METODOLOGÍA DE INVESTIGACIÓN Y CRONOGRAMA DE ACTIVIDADES

La siguiente tabla presenta las etapas de la metodología de investigación a seguir, junto con su duración y actividades asociadas.

<b>Etapas / Duración</b>	<b>Actividades</b>
<b>Minería de datos / 2 meses</b>	<ul style="list-style-type: none"> <li>- Definir el universo de colaboraciones científicas a estudiar, así como el periodo de años para la recolección de datos.</li> <li>- Diseñar una aplicación que automatice la obtención de información anual de colaboraciones reportadas en la <i>Web of Science</i> [9].</li> <li>- Identificar los eventos externos (disruptivos), por ejemplo estímulos a la investigación, que serán considerados en la definición del modelo.</li> </ul>
<b>Análisis de datos / 5 meses</b>	<ul style="list-style-type: none"> <li>- Desarrollar una biblioteca de métodos para automatizar la creación de redes de coautorías basadas en hipergrafos.</li> <li>- Calcular métricas topológicas de las redes, tales como centralidad de nodos, diámetro, distribución de grado, numero de comunidades, tamaño del componente mayor, coeficiente de agrupamiento, etc.</li> <li>- Estudiar la evolución anual de las métricas calculadas y su correlación con los eventos disruptivos.</li> <li>- Aplicar técnicas de aprendizaje automático en la definición de funciones de evolución para cada métrica.</li> </ul>
<b>Descripción y prueba del modelo / 2 meses</b>	<ul style="list-style-type: none"> <li>- Definir formalmente el modelo de evolución de la red, con base en las funciones de evolución obtenidas.</li> <li>- Probar la efectividad del modelo en los primeros años de la red.</li> </ul>
<b>Diseminación de resultados / 2 meses</b>	<ul style="list-style-type: none"> <li>- Escritura de un artículo de investigación sobre los modelos de evolución obtenidos, el cual se sometería para su publicación en la revista <i>IEEE Transactions on Network Science and Engineering</i> incluida en el JCR.</li> </ul>

### D. RECURSOS REQUERIDOS PARA EL DESARROLLO DEL PROYECTO

- Bibliografía básica correspondiente a los temas de ciencia de datos, aprendizaje automático, ciencia de redes y análisis de redes sociales. Disponible en el acervo bibliográfico

de la Unidad Cuajimalapa.

- Acceso a las bases de datos y buscadores como la *Web of Science* y *Scopus*. Disponible en la BIDIUAM.
- Acceso a Interfaces de Programación de Aplicaciones (API, por sus siglas en inglés) de la *Web of Science* y *Scopus* para ser utilizadas en las etapas de Minería y Análisis de Datos. Debido a que me encuentro registrada como desarrolladora de la *Web of Science* y *Scopus*, cuento con acceso a sus APIs.
- Computadora personal con acceso a Internet.

### E. METAS

- Informe con los principales modelos de evaluación de Redes Complejas susceptibles a eventos disruptivos.
- Organización de seminario de investigación, con el fin de difundir los resultados obtenidos y de generar un espacio de dialogo entre especialistas del Análisis de Redes Sociales.
- Publicación de un artículo de investigación en una revista indexada en el *Journal Citation Report*.
- Dirección de un proyecto terminal de la Licenciatura en Ingeniería en Computación, o bien, de la Licenciatura en Matemáticas Aplicadas.

## REFERENCIAS

- [1] A.-L. Barabási and M. Pósfai, *Network science*. Cambridge: Cambridge University Press, 2016.
- [2] M. E. J. Newman, "Coauthorship networks and patterns of scientific collaboration," *Proceedings of the National Academy of Sciences*, vol. 101, no. suppl 1, pp. 5200–5205, 2004.
- [3] R. Albert and A.-L. Barabási, "Topology of evolving networks: Local events and universality," *Phys. Rev. Lett.*, vol. 85, pp. 5234–5237, Dec 2000.
- [4] E. Vasilyeva, A. Kozlov, K. Alfaro-Bittner, D. Musatov, A. M. Raigorodskii, M. Perc, and S. Boccaletti, "Multilayer representation of collaboration networks with higher-order interactions," *Scientific Reports*, vol. 11, p. 11, 3 2021.
- [5] Y. Ding, S. Yan, Y. Zhang, W. Dai, and L. Dong, "Predicting the attributes of social network users using a graph-based machine learning method," *Computer Communications*, vol. 73, pp. 3 – 11, 2016.
- [6] P. Kondeti, L. P. Yerramreddy, A. Pradhan, and G. Swain, "Fake account detection using machine learning," in *Evolutionary Computing and Mobile Sustainable Networks* (V. Suma, N. Bouhmala, and H. Wang, eds.), (Singapore), pp. 791–802, Springer Singapore, 2021.
- [7] H. Shao, L. Wang, and Y. Ji, "Link prediction algorithms for social networks based on machine learning and harp," *IEEE Access*, vol. 7, pp. 122722–122729, 2019.
- [8] R. Sinatra, D. Wang, P. Deville, C. Song, and A.-L. Barabási, "Quantifying the evolution of individual scientific impact," *Science*, vol. 354, no. 6312, 2016.
- [9] "Web of Science: The global citation database." <https://clarivate.com/webofsciencegroup/solutions/web-of-science/>. Accessed: 2021-12-05.