

Maestría en Ingeniería Matemática Propuesta de Tesis

Título de la propuesta

Graph Signal Processing: algoritmos, teoría y aplicaciones para el estudio de procesos en redes

Identificación del proponente

Nombre: Gonzalo Mateos

Lugar de trabajo: Dept. of Electrical and Computer Engineering,
Goergen Institute for Data Science,
University of Rochester, Rochester, NY, USA

Información de contacto: gmateosb@ece.rochester.edu

<http://www.ece.rochester.edu/~gmateosb>

Area Tematica

Teoría (algebraica) de grafos, Procesamiento de Señales, Optimización, Network Science

Perfil esperado del estudiante

Se requieren conocimientos de Álgebra Lineal y se recomiendan conocimientos básicos de Teoría de Grafos, Optimización, o Estadística, e interés en programación enfocada al análisis estadístico de datos.

Resumen

A la par del crecimiento en tamaño e importancia de los sistemas interconectados modernos, y en cuanto se tornan más complejos y heterogéneos, existe una necesidad urgente de desarrollar una

teoría holística de redes. Al mismo tiempo, el diluvio de información impulsado por los avances en la Web y las redes sociales, el e-commerce, así como en las tecnologías de secuenciación del genoma humano, han hecho que las técnicas de aprendizaje estadístico, inferencia y predicción a partir de datos masivos sean cada vez más relevantes. Esta tendencia se ha visto impulsada por una combinación única del interés generalizado en el análisis científico *a nivel de sistema* (a diferencia del reduccionismo), y la capacidad de *recolección de datos de forma masiva* en diversas áreas. Extraer y sumarizar información relevante a partir de grandes cantidades de datos con una perspectiva enfocada en redes, constituirá un paso crucial para obtener nuevos conocimientos en diversos campos de la ciencia y la ingeniería – y las técnicas de procesamiento (y tratamiento estadístico) de señales tendrán un rol importante para cumplir dicho objetivo.

En este contexto, hacer frente a los desafíos actuales en la intersección de Network Science y Big Data requiere ampliar el alcance del procesamiento de señales espacio/temporales clásico, e incorporar también señales definidas sobre los nodos de un grafo. Dichas señales son fundamentales en una multitud de áreas de la ciencia y la ingeniería, por ejemplo los niveles de expresión de genes en redes de regulación genética, la propagación de epidemias en una red social, o los niveles de congestión en los buffers de una red de telecomunicaciones, por mencionar solo unos pocos. Bajo el supuesto de que las señales están relacionadas con la topología del grafo subyacente, el objetivo de *Graph Signal Processing* (GSP) es desarrollar algoritmos que aprovechen esta estructura relacional de manera fructífera, y que permitan hacer inferencias acerca de dichas relaciones aún cuando éstas sólo se observan parcialmente. Existe un desbalance evidente entre nuestra comprensión científica de señales definidas sobre dominios regulares (tiempo o espacio) y las señales definidas sobre grafos arbitrarios. El conocimiento humano sobre sistemas y señales temporales se desarrolló a lo largo de décadas, impulsado por necesidades reales en áreas tales como las comunicaciones, el procesamiento de audio, o el control de procesos. Por el contrario, la fascinación sobre la conectividad incesante de las sociedades modernas, y el acceso a datos de redes de alta calidad son acontecimientos recientes. Esto ha generado una necesidad urgente de comprender mejor las señales o procesos dinámicos definidos en grafos, lo que naturalmente fomentará el desarrollo de GSP en los próximos años.

Esta propuesta busca atacar esta problemática poniendo especial énfasis en el modelado, identificación, y control de procesos de red distribuidos – a menudo conceptualizados como señales definidas en los vértices de un grafo. Proyectos concretos incluyen el desarrollo de algoritmos y teoría para el muestro y reconstrucción de señales (suaves, o de banda limitada) en grafos, análisis espectral (transformadas de Fourier en grafos), inferencia de topologías no observadas, detección de anomalías en redes, modelado y predicción de procesos de difusión y percolación, e implementación de filtros lineales e invariantes en grafos. La idea es trabajar y validar los resultados con datos en el dominio de interés del estudiante, por ejemplo redes estructurales o funcionales en el cerebro, cascadas de información en Twitter o la WWW, tráfico en Internet, redes de interacción de proteínas, perfiles de opinión en redes sociales, o flujos de energía en la red eléctrica.

Referencias

1. D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, “The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains,” *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 83–98, Mar. 2013.
2. A. Sandryhaila and J. M. F. Moura, “Big Data analysis with signal processing on graphs,” *IEEE Signal Process. Mag.*, vol. 21, no. 5, pp. 80–90, 2014.
3. E. D. Kolaczyk, *Statistical Analysis of Network Data: Methods and Models*, Springer, New York, NY, 2009.