

Analyzing Network Data in Python using NetworkX and DyNetworkX

Tanner Hilsabeck, Hung Do, and Kevin S. Xu (University of Toledo)

Description of Topic and Structure

Networks are all around us in many forms, ranging from online social networks to public transportation networks to gene networks in biology. Analyzing network representations of complex systems can reveal many insights about how they are organized. This tutorial introduces methods for analyzing network data using Python. It will be divided into 2 parts.

In part 1, we begin by discussing a variety of measures of network structure, including measures of centrality, transitivity, path lengths, and homophily. We also introduce several common analysis tasks involving networks, including community detection and link prediction. We then demonstrate how to compute these measures and perform these analysis tasks in Python using the well-known NetworkX package¹. We conclude this part with a case study on real social network data.

In part 2, we transition to dynamic or temporal networks, where the structure of the network changes over time. We first introduce different representations of dynamic networks ranging from the discrete-time snapshot or panel data representation to the continuous-time relational event representation. We then discuss several different approaches to analyzing dynamic network data ranging from applying static network measures snapshot by snapshot to identifying temporal motifs. We demonstrate how to conduct such analyses using functionality built into NetworkX and DyNetworkX². We then conclude with a second case study on a real *dynamic* social network.

Expected Audience

We aim to serve a multidisciplinary audience, including researchers and practitioners from the mathematical sciences, the computer and information sciences, the social sciences, and public health. We will assume a relatively low mathematical background for attendees. Attendees with some prior experience in computer programming, but not necessarily in Python, will likely benefit the most from this tutorial given its focus on network analysis using Python packages.

Organizer Bios

Tanner Hilsabeck (DyNetworkX developer, 2019-Present) is a second-year MS student in Computer Science and Engineering at the University of Toledo. During his bachelor's degree, he found a passion for teaching others science and mathematics. His interdisciplinary experience provides unique insights when predicting how applications will interface with end users. His main research interests are in social and information networks and modeling human behavior.

Hung Do (DyNetworkX developer, 2020-Present) is a senior undergraduate student majoring in Data Science with a concentration in Economics at the University of Toledo. He has interest in modeling human behavior, network science, and machine learning.

Kevin S. Xu (PhD, University of Michigan, 2012) is currently an assistant professor in the EECS Department at the University of Toledo and has previously held industry research positions at Technicolor and 3M. He is a recipient of the NSF CAREER award and has previously served as a program co-chair for SBP-BRiMS. His main research interests are in machine learning and statistical signal processing with applications to network science, human dynamics, and health.

¹ <https://networkx.org/>

² <https://github.com/IdeasLabUT/dynetworkx>