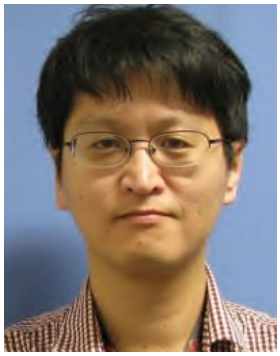
Invited Online Talk



Graph Representation Learning and Scientific Applications

Seung-Hwan Lim

Time: Friday, 08/04/2023, 1:00 PM — 2:00 PM (CST)
Zoom Link: https://unt.zoom.us/j/88308740502

Abstract: Graph is a universal language across science and engineering domains. Machine learning has demonstrated very promising results in 1-D sequential data (e.g., time series) and 2D grid data (e.g., images). A new horizon in machine learning is graph data, which can flexibly represent arbitrary dimensions and arbitrary relationships, through modeling entities as nodes and their relationships as edges. Thus, machine learning on graph data can address a range of non-trivial problems, which cannot be fully addressed by other methods. Performing machine learning on graph data consists of two steps: 1) learning representation of the target graph and; 2) optimizing the learned representation against the learning objective. This talk will focus on the first step, learning graph representation.

A fundamental challenge in earning graph representation is that graphs are not easy to compare with each other as graph isomorphism is an NP with no known polynomial time algorithms. Thus, computational efficiency is a critical issue, while learned representation should accurately capture the relationship in the original graph data. This talk will present one of the most important graph representation learning algorithms, the Weisfeiler-Lehman graph kernel, which is known to be equivalent to graph convolution network, along with relevant operations such as graph similarity search for scientific applications - molecular structure search and neural network model explanation.

Bio: Dr. Seung-Hwan Lim is a research scientist in the computer science and mathematics division at Oak Ridge National Laboratory. His research interests lie at the intersection of machine learning and computing system for various Department of Energy mission applications such as basic sciences, bio-medical domain, and others. He joined Oak Ridge National Laboratory in 2012 as a postdoc and was promoted to staff scientist in 2013. He graduated from the Pennsylvania State University as a Ph.D. in 2012. His PhD topic was parallel and distributed systems with the thesis title of managing performance and energy in large-scale data centers, using probabilistic models. Prior to his Ph.D. degree, he was a software engineer at Samsung Electronics between 2000 and 2005, with a BS/MS degree in computer engineering at Seoul National University, in 1998 and 2000, respectively.