
LUANA RUIZ, Johns Hopkins University
Machine Learning on Large-Scale Graphs

Graph neural networks (GNNs) are successful at learning representations from most types of network data but suffer from limitations in large graphs, which do not have the Euclidean structure that time and image signals have in the limit. Yet, large graphs can often be identified as being similar to each other in the sense that they share structural properties. Indeed, graphs can be grouped in families converging to a common graph limit – the graphon. A graphon is a bounded symmetric kernel which can be interpreted as both a random graph model and a limit object of a convergent sequence of graphs. Graphs sampled from a graphon almost surely share structural properties in the limit, which implies that graphons describe families of similar graphs. We can thus expect that processing data supported on graphs associated with the same graphon should yield similar results. In my research, I formalize this intuition by showing that the error made when transferring a GNN across two graphs in a graphon family is small when the graphs are sufficiently large. This enables large-scale graph machine learning by transference: training GNNs on moderate-scale graphs and executing them on large-scale graphs.