Libraries for Graph to Vector

This comparison of graph2vec libraries aims to select the fastest and scalable for sparse graph. Since its out of the scope to test all libraries, we explored the well-known based on research papers citations and number of users use the library on GitHub repositories.

Metrics used for selection:

- System environment for library to run e.g parallel, distributed, CPUs, GPUs
- Large sparse graph can be handled such as 1 million nodes e.g Yes, No
- Speed how fast compared to overall experience e.g High, Medium, Low

Following some libraries from Google, Google Scholar, and their documents:

Library	System	Implementation	Large Sparse	Speedup
GraphVite	CPUs and GPUs Parallel with CUDA	C++ and Python interface	Yes 1 million and more	High
Deep Graph Library	CPU with CUDA	Python	Yes	High
Pytorch BigGraph	Distribution	Python	Yes	High
NetSMF	Sparse Matrix Factorization	C++	Yes	Medium
PecanPy	Parallel	Python	No	Medium
700x	Sparse Matrix Factorization	Python	Yes	Medium

GraphVite is a general graph embedding engine, dedicated to high-speed and large-scale embedding learning in various applications. By cooperating CPUs and GPUs for learning, it scales to million-scale or even billion-scale graphs. With its Python interface, you can easily practice advanced graph embedding algorithms, and get results in incredibly short time.

Pytorch BigGraph

List of networks embedding techniques

Pytorch BigGraph paper documentation, and github repo