

Using RAPIDS AI to Accelerate Graph Data Science Workflows

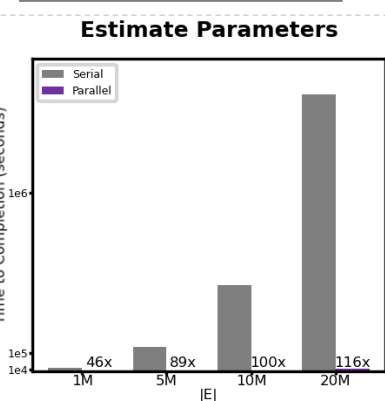
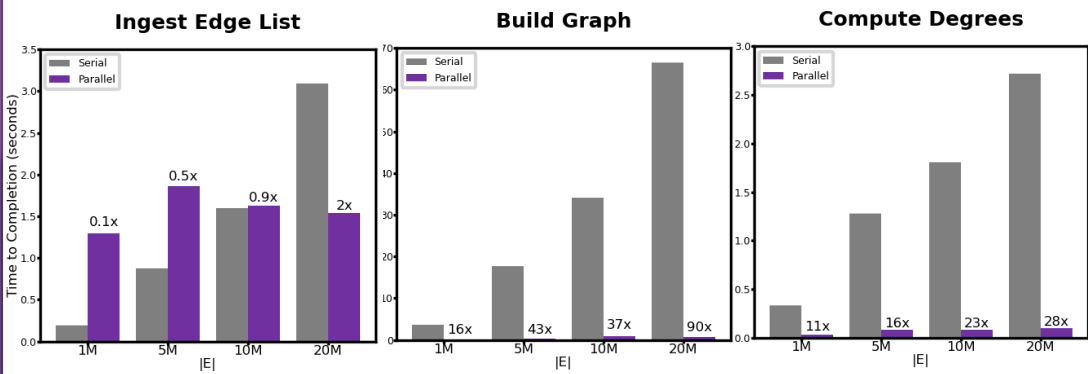
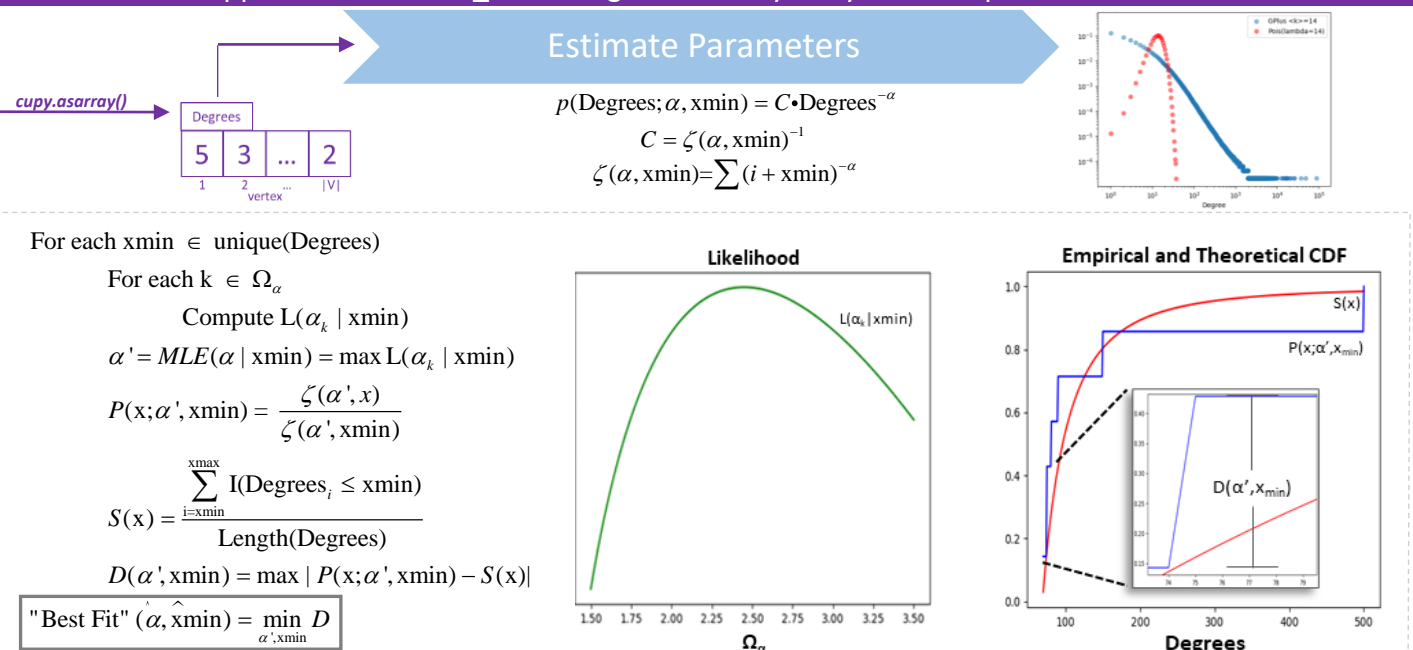
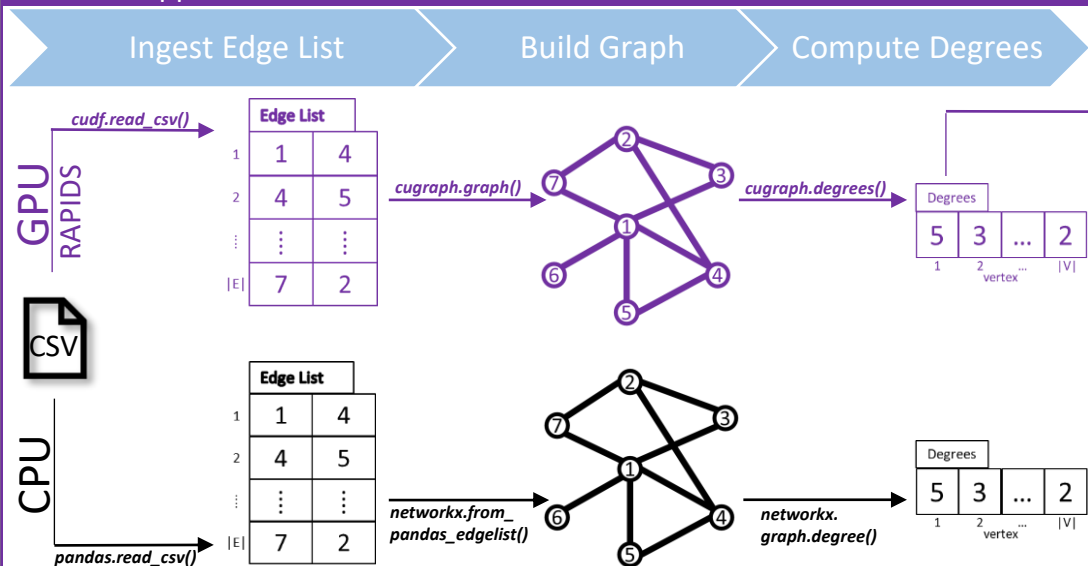


T. Hricik¹, D. Bader¹, O. Green²
 1) Ying Wu College of Computing, New Jersey Institute of Technology
 2) NVIDIA Corporation



Challenge 1: CPU based Python tools for graph ETL do not scale
 Approach: RAPIDS AI data structures and methods scale on GPUs

Challenge 2: Power-law parameter estimation requires 2-dimensional grid search
 Approach: Load cudf_series Degrees as CuPy array and compute on GPUs



- [1] B. Huberman, L. Adamic, "Growth dynamics of the world-wide web," Nature, vol. 2, 1999, p.131.
- [2] N. Gong, W. Xu, "Reciprocal versus parasocial relationships in online social networks," Springer Social Network Analysis and Mining (SNAM), 4(1), 2014.
- [3] Clauset A., Shalizi, C.R., Newman, M.E.J., "Power-law distributions in empirical data," SIAM Review, 51, 2009, pp. 661-703.
- [4] Bauke, H., "Parameter estimation for power-law distributions by maximum likelihood methods", European Physical Journal B 58, 67, 2007, pp. 167-173.
- [5] Gong, N.Z., Xu, W., Huang, L., Mittal, P., Stefanov, E., Skar, V. Song, D., "Evolution of social-attribute networks: measurements, modeling, and implications using Google+", ACM Workshop on Social Network Mining and Analysis (SNA-KDD), co-located with KDD, 2012.
- [6] M. Kutner, C. Nachtsheim, J. Neter, Applied Linear Regression Models, 4th ed., McGraw-Hill Irwin: New York, 2004, pp. 100-118.

*This work was partially funded by D. Bader's NVIDIA NVAL Award