Master 2 Internship Proposal: Large-Scale Learning on Graphs

Nicolas Keriven Nicolas Tremblay Romain Couillet
Gipsa-lab, Univ. Grenoble-Alpes, France

Context. In the last decade, machine learning (ML) on *graph data* has known a rapid growth, with the advent of kernel methods on graphs [3], deep neural nets on graphs [9], benchmark datasets, and numerous applications ranging from the analysis of social networks to molecular classification or protein interface prediction. However, existing methods can be expensive in terms of algorithmic complexity and memory storage, and applying them to large graphs with reasonable computing power often remains an issue.

On more classical data such as images or audio, methods such as dimensionality reduction or random kernel features [6] have been successfully used to improve the computational efficiency of algorithms applied to large datasets. In particular, recent advances in *optical computing* open new perspectives [7], by being able to perform specific non-linear random operations in very high dimension at almost no cost, using a so-called *Optical Processing Unit* (OPU).

Goals. The goal of the internship is to explore the possibilities of optical computing and dimensionality reduction for large-scale learning on graphs, from an empirical and theoretical point of view. As a starting point, we will particularly focus on combining these principles with *random subgraph sampling* (and its many variants) [4], which is one of the most efficient machine learning algorithm on graphs and presents many advantages in terms of parallelization and data privacy. Using an OPU developed by a startup company, LightOn, the proposed efficient algorithms will be benchmarked against standard methods.

On the theoretical side, several approaches can be adopted, depending on the interest of the candidate. As noted in previous work [2], a notion related to kernel features and random sampling (here, of subgraphs) is that of *mean kernel map* [1]. A possible objective would be to characterize the discriminative power of this type of maps for models of large graphs. On the other hand, when a particular *structure* is assumed on the probability distribution of a collection of objects, then dimensionality reduction can be effectively performed, for instance by computing a so-called *linear sketch* [2]. The size of the sketch is directly linked to the complexity of the method, and to the structure underlying the object of interest. In light of this, the theory of graphons [5] is, for instance, a primary example of generative models for large random graphs which allows for a complete characterization of the distribution of subgraphs, and therefore, of potential computational gains by dimension reduction. Hence, another possible objective would be to characterize the trade-off between learning power and complexity of the proposed methods, using the theory of graphons and its variants [8].

The internship. This internship will take place at the Gipsa-lab in Grenoble, France, and will be done in close collaboration with LightOn, a startup company based in Paris which develops the OPU. The latter will be used remotely most of the time, but occasional short stays in Paris will be planned to visit LightOn. A "gratification de stage" (small wage) will be provided to the intern.

Application, contact. Please send a CV and short statement of interest to
`firstname.lastname@gipsa-lab.grenoble-inp.fr`
Do not hesitate to contact us if you have any question.

References

[1] Arthur Gretton, Karsten M. Borgwardt, Malte J. Rasch, Bernhard Schölkopf, and Alexander J. Smola. A Kernel Method for the Two-Sample Problem. In *Advances in Neural Information Processing Systems (NIPS)*, pages 513–520, 2007.

[2] Nicolas Keriven, Anthony Bourrier, Rémi Gribonval, and Patrick Pérèz. Sketching for Large-Scale Learning of Mixture Models. *Information and Inference: A Journal of the IMA*, 7(3):447–508, 2018.

[3] Nils M. Kriege, Fredrik D. Johansson, and Christopher Morris. A Survey on Graph Kernels. 2019.

[4] Jure Leskovec and Christos Faloutsos. Sampling from large graphs. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2006:631–636, 2006.

[5] László Lovász and Balázs Szegedy. Limits of dense graph sequences. *Journal of Combinatorial Theory. Series B*, 96(6):933–957, 2006.

[6] Ali Rahimi and Benjamin Recht. Random Features for Large Scale Kernel Machines. In *Advances in Neural Information Processing Systems (NIPS)*, 2007.

[7] Alaa Saade, Francesco Caltagirone, Igor Carron, Laurent Daudet, Angélique Dremeau, Sylvain Gigan, and Florent Krzakala. Random projections through multiple optical scattering: Approximating Kernels at the speed of light. In *IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)*, pages 6215–6219, 2016.

[8] Victor Veitch and Daniel M. Roy. Sampling and Estimation for (Sparse) Exchangeable Graphs. pages 1–25, 2016.

[9] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and Philip S. Yu. A Comprehensive Survey on Graph Neural Networks. pages 1–22, 2019.