

2018-00358 - PhD Position / Scientific computing / Optimization, machine learning and statistical methods

Type de contrat : CDD de la fonction publique
Niveau de diplôme exigé : Bac + 5 ou équivalent
Fonction : Doctorant
Niveau d'expérience souhaité : Jeune diplômé

A propos du centre ou de la direction fonctionnelle

Our aim is to develop tight formulations for combinatorial problems by combining the latest reformulation techniques, such as Lagrangian and polyhedral approach, non-linear programming tools and graph theoretics tools. Through industrial partnerships, the team targets large scale problems such as those arising in logistics (routing problems), in planning and scheduling, in network design and control, and in placement problems (cutting stock problems).

Research themes

Our project brings together complementary expertise in combinatorial optimization : Mixed Integer Programming (Polyhedral, Lagrangian and decomposition approaches, Branch-and-Price-and-Cut Algorithms), Quadratic programming (semi-definite-programming), and Graph Theory (for induced properties and implicit representation of solutions). We develop approximate solutions for large scale problems through mathematical programming based primal heuristics.

International and industrial relations

We have an associated team in Brazil through which we collaborate with Artur Pessoa and Eduardo Uchoa (Universidade Federal Fluminense) and Marcus Poggi (PUC-Rio)
Our Industrial partners are Pascale Bendotti and Marc Porcheron (EDF, R&D Dpt OSIRIS), and Fabien Rodes (société Exeo Solutions).

Contexte et atouts du poste

Within the framework of a partnership

- collaboration between 4 Inria teams: RealOpt, Zenith, Storm and Tau (IPL HPC-BigData)

Mission confiée

Recently, several frameworks such as TensorFlow [1] and PyTorch [2] emerged and represent the DL network as a directed graph whose nodes represent convolution operations and edges represent data dependences between them. The goal of this PhD thesis is to work on how to allocate the convolution operations and how to schedule them to achieve a better efficiency, typically in the context of platforms consisting of heterogeneous resources such as GPUs and multicore nodes.

The goal of this PhD Thesis is to improve the scheduling and resource allocation strategies along several directions. First, the resource allocation algorithm does not take into account the specificities of the application. Indeed, it is for instance close to the default StarPU scheduling algorithm [3] used for general task graphs.

Second, it has been proved that for specific applications such as linear algebra kernels, injecting some static knowledge based on a more sophisticated scheduling algorithm can strongly improve the performance of greedy algorithm [4]. Third, in the context of DL, the same graph of convolution layers is used many times on different input data along the execution of the DL algorithm, what is close to the context of steady state scheduling [5], that has been proved to be more tractable than general scheduling. At last, another opportunity is to develop high level simulation techniques, that could be used in particular to detect bottlenecks with respect to a DL network and to a parallel architecture. This possibility could more speculatively be especially interesting in the context of DL, since it may help to redesign the network itself to cope with bottlenecks. We will first concentrate on classical layers (Fully Connected Layers, Convolutional Layers, Recurrent Layers) before considering PL@ntNet [6] as a target network.

Informations générales

- **Thème/Domaine** : Optimisation, apprentissage et méthodes statistiques
- **Ville** : Talence
- **Centre Inria** : CRI Bordeaux - Sud-Ouest
- **Date de prise de fonction souhaitée** : 03-09-2018
- **Durée de contrat** : 3 ans
- **Date limite pour postuler** : 30-06-2018

Contacts

- **Equipe Inria** : REALOPT
- **Recruteur** :
Beaumont Olivier /
olivier.beaumont@inria.fr

L'essentiel pour réussir

Technical skills and level required: The candidate will be required to have a solid background in Combinatorial Optimization (scheduling, resource allocation, online algorithms) and/or in Deep Learning (TensorFlow, PyTorch) and a taste for both domains.

Conditions pour postuler

Thank you to send :

- updated CV
- cover letter
- letters of recommendation eventually
- Master Degree Transcripts

Sécurité défense :

Ce poste est susceptible d'être affecté dans une zone à régime restrictif (ZRR), telle que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L'autorisation d'accès à une zone est délivrée par le chef d'établissement, après avis ministériel favorable, tel que défini dans l'arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un poste affecté dans une ZRR aurait pour conséquence l'annulation du recrutement.

Politique de recrutement :

Dans le cadre de sa politique diversité, tous les postes Inria sont accessibles aux personnes en situation de handicap.

Attention: Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.

[1]. M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467, 2016.

[2] Pytorch, <http://pytorch.org>

[3] C. Augonnet, S. Thibault, R. Namyst, and P.-A. Wacrenier. Starpu: a unified platform for task scheduling on heterogeneous multicore architectures. *Concurrency and Computation: Practice and Experience*, 23(2):187–198, 2011.

[4] E. Agullo, O. Beaumont, L. Eyraud-Dubois, and S. Kumar. Are static schedules so bad? a case study on cholesky factorization. In *Parallel and Distributed Processing Symposium, 2016 IEEE International*, pages 1021–1030. IEEE, 2016.

[5] O. Beaumont, A. Legrand, L. Marchal, and Y. Robert. Steady-state scheduling on heterogeneous clusters. *International Journal of Foundations of Computer Science*, 16(02):163–194, 2005.

[6] D. Barthélémy, N. Boujemaa, D. Mathieu, M. Jean-François, A. Joly, and E. Mouysset. The pl@ntnet project: plant computational identification and collaborative information system. 2011.

Principales activités

These research directions require the joint knowledge of experts in deep learning algorithms, dynamic runtime scheduling and scheduling theory and will benefit in particular to Pl@ntNet application.

The PhD student will be localized in Bordeaux and will be co-supervised by Olivier Beaumont (RealOpt) and Alexis Joly (Zenith), in close collaboration with Guillaume Charpiat (Tau) and Samuel Thibault (Storm). Several stays (1 week) in Saclay and Montpellier will be scheduled during the PhD Thesis.

Avantages sociaux

- Subsidised catering service
- Partially-reimbursed public transport

Rémunération

1982€ / month (before taxes) during the first 2 years, 2085€ / month (before taxes) during the third year.