香港浸會大學
HONG KONG BAPTIST UNIVERSITY

DEPARTMENT OF
COMPUTER SCIENCE
計算機科學系

ONLINE SEMINAR
2022 SERIES

Department of Computer Science

Prof. Peter Richtárik

Professor of Computer Science
King Abdullah University of Science and Technology
Saudi Arabia

Register Now

Date: 2 September 2022 (Friday)
Time: 3:00pm – 4:00pm
Registration: http://bit.ly/bucs-ereg

(*Zoom details will only be provided to registrants)

ProxSkip: Yes! Local gradient steps provably lead to communication acceleration! Finally!

ABSTRACT

In this talk I will introduce ProxSkip [1] - a surprisingly simple and provably efficient method for minimizing the sum of a smooth (f) and an expensive nonsmooth proximable (ψ) function. The canonical approach to solving such problems is via the proximal gradient descent (ProxGD) algorithm, which is based on the evaluation of the gradient of f and the prox operator of ψ in each iteration. In this work we are specifically interested in the regime in which the evaluation of prox is costly relative to the evaluation of the gradient, which is the case in many applications. ProxSkip allows for the expensive prox operator to be skipped in most iterations: while its iteration complexity is $O(\kappa log 1/\epsilon)$, where κ is the condition number of f, the number of prox evaluations is $O(\sqrt{\kappa} log 1/\epsilon)$ only. Our main motivation comes from federated learning, where evaluation of the gradient operator corresponds to taking a local GD step independently on all devices, and evaluation of prox corresponds to (expensive) communication in the form of gradient averaging. In this context, ProxSkip offers an effective acceleration of communication complexity. Unlike other local gradient-type methods, such as FedAvg [2], SCAFFOLD [3], S-Local-GD [4] and FedLin [5], whose theoretical communication complexity is worse than, or at best matching, that of vanilla GD in the heterogeneous data regime, we obtain a provable and large improvement without any heterogeneity-bounding assumptions.

Time permitting, I will mention several subsequent extensions, generalizations and improvements [6, 7, 8].

BIOGRAPHY

Peter Richtárik is a professor of Computer Science at KAUST, Saudi Arabia, where he leads the Optimization and Machine Learning Lab. Through his work on randomized and distributed optimization algorithms, he has contributed to the foundations of machine learning and optimization. He is one of the original developers of Federated Learning. Prof Richtárik's works attracted international awards, including a Best Paper Award at aNeurIPS 2020 workshop, Distinguished Speaker Award at the 2019 International Conference on Continuous Optimization, SIAM SIGEST Best Paper Award, and the IMA Leslie Fox Prize. He serves as an Area Chair for leading machine learning conferences, including NeurIPS, ICML and ICLR, and is an Action Editor of Transactions of Machine Learning Research, Area Editor of Journal of Optimization Theory and Applications, and Associate Editor of Optimization Methods and Software.

ENQUIRY

Tel: 3411-2385 Email: comp@comp.hkbu.edu.hk Website: https://www.comp.hkbu.edu.hk/v1/?page=events