

NATIONAL UNIVERSITY OF SINGAPORE

School of Computing

PH.D DEFENCE - PUBLIC SEMINAR

**Title: EXPLOITING DECENTRALIZED MULTI-AGENT COORDINATION
FOR LARGE-SCALE MACHINE LEARNING PROBLEMS**

Speaker: Mr Ouyang Ruofei

Date/Time: 7 November 2016, Monday, 09:00 AM to 10:30 AM

Venue: Executive Classroom, COM2-04-02

Supervisor : Dr Low Kian Hsiang, Assistant Professor, School of Computing

Abstract:

Nowadays, the scale of machine learning problems becomes much larger than before. It raises a huge demand in distributed perception and distributed computation. A multi-agent system provides exceptional scalability for problems like active sensing and data fusion. However, many rich characteristics of large-scale machine learning problems have not been addressed yet such as large input domain, nonstationarity, and high dimensionality. This thesis identifies the challenges related to these characteristics from multi-agent perspective. By exploiting the correlation structure of data in large-scale problems, we propose multi-agent coordination schemes that can improve the scalability of the machine learning models while preserving the computation accuracy. To elaborate, the machine learning problems we are solving with multi-agent coordination techniques are:

(a) Gaussian process regression. To perform distributed regression on a large-scale environmental phenomenon, data compression is often required due to the communication costs. Currently, decentralized data fusion methods encapsulate the data into local summaries based on a fixed support set. However in a large-scale field, this fixed support set, acting as a centralized component in the decentralized system, cannot approximate the correlation structure of the entire phenomenon well. It leads to evident losses in data summarization. Consequently, the regression performance will be significantly reduced.

In order to approximate the correlation structure accurately, we propose an agent-centric support set to allow every agent in the data fusion system to choose a possibly different support set and dynamically switch to another one during execution for encapsulating its own data into a local summary which, perhaps surprisingly, can still be assimilated with the other agents' local summaries into a globally consistent summary. Together with an information sharing mechanism we designed, the new decentralized data fusion methods with agent-centric support set can be applied to regression problems on a much larger environmental phenomenon with high performance.

(b) Active learning. In the context of environmental sensing, active learning/active sensing is a process of taking observations to minimize the uncertainty in an environmental field. The uncertainty is quantified based on the correlation structure of the phenomenon which is traditionally assumed to be stationary for computational sake. In a large-scale environmental field, this stationary assumption is often violated. Therefore, existing active sensing algorithms perform sub-optimally for a nonstationary environmental phenomenon.

To the best of our knowledge, our decentralized multi-robot active sensing (DEC-MAS) algorithm is the first work to address nonstationarity issue in the context of active sensing. The uncertainty in the phenomenon is quantified based on the nonstationary correlation structure estimated by Dirichlet process mixture of Gaussian processes (DPM-GPs). Further, our DEC-MAS algorithm can efficiently coordinate the exploration of multiple robots to automatically trade-off between learning the unknown, nonstationary correlation structure and minimizing the uncertainty of the environmental phenomenon. It enables multi-agent active sensing techniques to be applied to a large-scale nonstationary environmental phenomenon.

(c) Bayesian optimization. Optimizing an unknown objective function is challenging for traditional optimization methods. Alternatively, in this situation, people use Bayesian optimization which is a modern optimization technique that can optimize a function by only utilizing the observation information (input and output values) collected through simulations. When the input dimension of the function is low, a few simulated observations can generate good result already. However, for high dimensional function, a huge number of observations are required which is impractical when the simulation consumes lots of time and resources.

Fortunately, many high dimensional problems have sparse correlation structure. Our ANOVA-DCOP work can decompose the correlation structure in the original high-dimensional problem into many correlation structures of subsets of dimensions based on ANOVA kernel function. It significantly reduces the size of input space into a collection of lower-dimensional subspaces. Additionally, we reformulate the Bayesian optimization problem as a decentralized constrained optimization problem (DCOP) that can be efficiently solved by multi-agent coordination techniques so that it can scale up to problems with hundreds of dimensions.