

## **Proposition de Sujet de Thèse pour Contrat Doctoral UCA**

Adresse e-mail à utiliser pour toute correspondance :

Andre.Ferrari@unice.fr

### **Titre de la thèse**

Apprentissage collaboratif et décision en ligne sur graphes dynamiques

### **Thesis Title**

Online learning over dynamic networks for novelty detection

### **Directeur de Thèse (HDR ou assimilé)**

Nom : Ferrari

Prénom : André

Téléphone : 04 92 07 63 49

Courriel : Andre.Ferrari@unice.fr

### **Laboratoire d'accueil**

LAGRANGE

### **Co-directeur**

Nom : Richard

Prénom : Cédric

HDR : Oui

Unité de recherche : Lab. J.-L. Lagrange

Téléphone : 04 92 07 63 49

Courriel : cedric.richard@unice.fr

### **Domaine Scientifique**

DS8 - Sciences pour l'Ingénieur

## Description du sujet

Les 5 dernières années ont été marquées par un engouement majeur pour le traitement des méga-données, soutenu par le renouveau de l'Intelligence Artificielle et une très grande variété d'applications associées. Un large éventail de ces problèmes est structuré en graphes et nécessite de s'adapter à des dynamiques variant dans le temps. Les exemples abondent et concernent de nombreux systèmes de surveillance, de gestion et de prévention des risques, des instruments d'observation de la Terre ou de l'Univers, ou encore de gestion de réseaux de transport, de télécommunication et de distribution d'énergie.

Ces systèmes interconnectés sont constitués d'un très grand nombre d'agents (capteurs, processeurs, actionneurs) liés entre eux par une topologie de connexion. Ces agents sont éventuellement autonomes. Ils peuvent potentiellement interagir et collaborer, de façon dynamique, afin d'accomplir leur mission plus efficacement. La topologie des graphes considérés est susceptible d'évoluer au cours du temps. Une spécificité supplémentaire réside dans le fait que le volume des données acquises en chaque nœud, et l'agilité requise pour les analyser, incitent à privilégier des traitements locaux, sur graphe, par rapport à toute solution centralisée. De même, un traitement en ligne des données est nécessaire afin de s'adapter à des signaux dont les caractéristiques peuvent varier au cours du temps.

L'objectif du travail de thèse proposé est d'explorer de nouvelles méthodes d'apprentissage adaptatives, distribuées et collaboratives sur des graphes. Il a pour ambition de proposer des contributions méthodologiques pour l'extraction automatique d'informations des flux de données aux différents nœuds de ces graphes en levant les verrous suivants : volume des données acquises en chaque nœud, taille du graphe, suivi des variations temporelles et spatiales, charge calculatoire induite sur chaque nœud, volume du transfert de données entre les nœuds au regard du bénéfice induit par un traitement collaboratif. L'accent sera mis sur des problèmes d'apprentissage pour la décision (détection, classification). Un intérêt particulier sera porté sur la détection de changements ou d'anomalies (novelty detection). Celle-ci reposera sur la collaboration des nœuds du graphe en vue de détecter un régime anormal apparu à un instant donné et dans une zone donnée. De façon connexe, le travail de thèse portera également sur le problème de clustering visant à partitionner le graphe de façon dynamique suivant des caractéristiques communes aux données mesurées par les différents nœuds. Un point essentiel de cette recherche est qu'elle visera à répondre à ces différents problèmes, d'une part en s'affranchissant d'une modélisation paramétrique des signaux mesurés en chaque nœud, d'autre part en développant des algorithmes non supervisés afin de limiter l'intervention d'une expertise extérieure.

Soulignons que les travaux les plus récents en traitement du signal sur graphe se sont jusqu'ici essentiellement concentrés sur des signaux et graphes statiques, sans tenir compte de la dimension temporelle. Seuls quelques rares travaux ont commencé à s'intéresser à des signaux sur graphe dynamiques, mais sans s'inscrire dans des problèmes d'apprentissage automatique, distribué et collaboratif comme proposé dans ce travail de thèse.

1. D. Shuman, S. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging

field of signal processing on graph: Extending high-dimensional data analysis to networks and other irregular domains," IEEE Signal Processing Magazine, vol. 30, pp. 83–98, 2013.

2. R. Nassif, C. Richard, A. Ferrari, and A. H. Sayed, "Diffusion LMS for multitask problems with local linear equality constraints," vol. 65, no. 19, pp. 4979–4993, 2017.
3. R. Nassif, C. Richard, A. Ferrari, and A. H. Sayed, "Proximal multitask learning over networks with sparsity-inducing coregularization," IEEE Transactions on Signal Processing, vol. 64, no. 23, pp. 6329–6344, 2016.

## Description of the thesis

The last 5 years have witnessed an explosion of interest in Big Data processing, supported by the renewal of Artificial Intelligence and a wide variety of applications. A wide range of these problems are network-structured and need to adapt to time-varying dynamics. Examples abound and concern, e.g., systems for risk management and prevention, instruments for observation and study of Earth and the universe, or monitoring systems for transport, telecommunications and energy distribution networks.

These interconnected systems usually consist of a collection of networked agents (sensors, processors, actuators) linked together by a connection topology. These agents are possibly autonomous. They can potentially interact and collaborate dynamically to accomplish their mission more effectively. The topology of the graphs considered is likely to evolve over time. An additional specificity lies in the fact that the volume of data acquired at each node, and the agility required to analyze these data, encourage the use of local in-network processing over any centralized solution. Similarly, online data processing is required to accommodate signals whose characteristics may vary over time.

The objective of the proposed thesis is to explore new methods for adaptive, distributed and collaborative learning on graphs. It aims to propose methodological contributions for the automatic extraction of information from data flows at the nodes of these graphs by removing the following barriers: volume of the data acquired at each node, size of the graph, monitoring data subject to temporal and spatial variations, computational load induced at each node, volume of the data transfer between the nodes regarding the benefit induced by a collaborative treatment. The focus will be on machine learning problems for the decision making (detection, classification), on novelty detection in particular. The learning algorithms will build on the collaboration of the nodes to detect an abnormal regime at a given time instant and in a given area. In a related way, the PhD work will also deal with the clustering problem aiming at partitioning the graph dynamically given characteristics common to the data acquired at the different nodes. An essential point of this research is that it will address these questions, on the one hand by avoiding a parametric modeling of the signals measured at each node, and on the other hand by deriving unsupervised algorithms to limit the intervention of external expertise.

It should be noted that most of recent works on graph signal processing have been mainly focused on static signals and graphs, regardless of the time dimension. Only a few works have begun to be interested in dynamic graph signals, but without considering machine learning problems in multi-agent settings as proposed in this thesis work.

1. D. Shuman, S. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graph: Extending high-dimensional data analysis to networks

- 
- and other irregular domains," *IEEE Signal Processing Magazine*, vol. 30, pp. 83–98, 2013.
2. R. Nassif, C. Richard, A. Ferrari, and A. H. Sayed, "Diffusion LMS for multitask problems with local linear equality constraints," vol. 65, no. 19, pp. 4979–4993, 2017.
  3. R. Nassif, C. Richard, A. Ferrari, and A. H. Sayed, "Proximal multitask learning over networks with sparsity-inducing coregularization," *IEEE Transactions on Signal Processing*, vol. 64, no. 23, pp. 6329–6344, 2016.

### **Informations complémentaires**