

Suite de graphes et détection d'anomalies

Projet

Les techniques de détection d'anomalies sont à présent très utilisées dans le monde académique ainsi qu'en entreprise. Elles permettent de détecter (un ou) des instants où intervient un changement dans le signal observé en temps réel. Ce changement est caractéristique d'un phénomène d'intérêt tel que la panne d'un moteur, la défaillance d'un capteur, la hausse du cours d'une action, une cyber-attaque...

En faisant intervenir la notion de noyau reproduisant, on peut détecter de tels changements remarquables à partir de signaux non numériques et/ou de très grande dimension constitués à partir de réalisations de graphes aléatoires ou de séquences d'ADN par exemple. Cependant, de telles approches reposent sur le choix d'un noyau adapté à chaque type de données considérées.

Les objectifs du stage sont :

1. identifier les techniques classiques de détection d'anomalies *online*,
2. utiliser les noyaux semi-définis positifs pour réaliser la détection dans une suite de graphes,
3. analyser théoriquement la procédure résultant pour en garantir les performances,
4. mettre en oeuvre la stratégie sur de vraies données en l'implémentant de façon efficace.

Prérequis

Outre de bonnes compétences en probabilités et statistique, le candidat devra savoir programmer en R afin de pouvoir évaluer les performances des approches envisagées sur simulations ainsi que sur de vraies données.

Encadrant : Alain Celisse, maître de conférences.

Length: 4-6 mois.

Opportunity: Le stage pourrait donner lieu à une thèse.

Laboratory: MODAL équipe-projet Inria, Lille.

Contact: Alain Celisse (celisse@math.univ-lille1.fr).

Bibliography

1. Sylvain Arlot, Alain Celisse. Segmentation in the mean of heteroscedastic data by cross-validation. *Statistics and Copmputing*, 2011, Volume 21, 4, 613-632.
2. Sylvain Arlot, Alain Celisse, and Zaïd Harchaoui. Kernel change-point detection.