



Offre n°2024-07479

## Doctorant F/H Traitement statistique de « low data » issus de capteurs passifs : application au suivi d'ouvrages d'art

Type de contrat : CDD

Niveau de diplôme exigé : Bac + 5 ou équivalent

Fonction : Doctorant

### A propos du centre ou de la direction fonctionnelle

Le centre Inria de l'Université de Lille, créé en 2008, emploie 360 personnes dont 305 scientifiques répartis dans 15 équipes de recherche. Reconnu pour sa forte implication dans le développement socio-économique de la région des Hauts-De-France, le centre Inria de l'Université de Lille entretient des relations étroites avec les grandes entreprises et les PME. En favorisant les synergies entre chercheurs et industriels, Inria participe au transfert de compétences et d'expertise dans le domaine des technologies numériques et donne accès au meilleur de la recherche européenne et internationale au bénéfice de l'innovation et des entreprises, notamment dans la région.

Depuis plus de 10 ans, le centre Inria de l'Université de Lille est situé au cœur de l'écosystème universitaire et scientifique lillois, ainsi qu'au cœur de la Frenchtech, avec un showroom technologique basé avenue de Bretagne à Lille, sur le site d'excellence économique EuraTechnologies dédié aux technologies de l'information et de la communication (TIC).

### Contexte et atouts du poste

Dans le cadre d'un projet financé par Bpifrance, cette thèse s'inscrit dans une collaboration entre l'équipe Inria MODAL, le Cerema (Centre d'études et d'expertise sur les risques, l'environnement, la mobilité et l'aménagement) et la société SilMach (société en micromécanique MEMS).

Ces dernières années ont vu une accélération de la dégradation de nos chaussées et nos ponts en raison du vieillissement des structures, de l'évolution climatique et de l'augmentation de la charge autorisée des poids lourds. Afin de mieux mesurer ces dégradations et anticiper la maintenance des infrastructures routières, plusieurs initiatives se mettent en place.

A travers leur projet commun ROAD-AL<sup>[1]</sup>, Inria et le Cerema étudient conjointement les outils numériques permettant de modéliser ces phénomènes grâce à une instrumentation des structures. Cette initiative est complétée et renforcée par le projet SIRCAPASS coordonné par la société SilMach et qui vise à utiliser une nouvelle technologie de capteurs passifs MEMS (*Micro Electro-Mechanical Systems*) pour cette instrumentation.

Ces capteurs passifs ont l'avantage d'être résilients dans des conditions hostiles tout en ne nécessitant aucun apport énergétique (« zéro énergie »). Il s'agit typiquement de composants mécaniques passifs de deux types : soit des roues dentées dont le signal de sortie correspond à un comptage fourni par les crans d'une roue en rotation, soit des systèmes de fusible mécanique dont le signal de sortie correspond à une information binaire correspondant à la casse d'un fusible (ou binaire multivarié en cas d'un train de fusibles multi-calibrés). Ces capteurs sont actionnés par des mouvements de structures (ouvrage d'art typiquement) ne nécessitant ainsi aucun branchement électrique. Leur relative « simplicité » leur permet aussi de réaliser leur temps de mission avec très peu d'entretien ou de surveillance.

En contrepartie, ces capteurs passifs produisent des données particulièrement parcimonieuses en fréquence et en type (comptage, binaire) au regard de capteurs plus traditionnels fournissant typiquement des données continues quasiment en temps réel. On se retrouve alors avec des données que l'on pourrait qualifier de « low data ». Les problématiques de SilMach vis-à-vis de ces low data sont alors les suivantes : (1) capacité à produire une estimation effective de la durée de vie résiduelle d'un ouvrage, (2) comparaison de la qualité de ces estimateurs avec ceux de capteurs actifs traditionnels, (3) calibration des capteurs passifs (niveau conception) pour minimiser cet écart de qualité et donc aider à proposer des capteurs passifs optimaux.

Informations complémentaires:

- co-encadrement de la thèse par André Orcési (chercheur au Cerema)
- le lieu principal de la thèse sera à déterminer, entre Champs-sur-Marne (Cerema) et Villeneuve d'Ascq (centre Inria de l'Université de Lille), avec des déplacements réguliers, et pris en charge financièrement, entre les deux sites

## Mission confiée

Cette thèse vise précisément à répondre aux trois problématiques précédentes dans le cadre générique de l'estimation statistique. La clé de résolution consiste à constater que les low data observées peuvent être considérées comme des censures des données continues qui seraient issues de capteurs actifs traditionnels. Le problème peut ainsi être réinterprété comme une question de modélisation à deux niveaux : un modèle sur les données continues (non observées) et un modèle de censure (conduisant aux low data observées). Il s'agit donc d'effectuer un travail de recherche proposant des modélisations réalistes, puis de les estimer, et enfin de les valider (comparaison à des données tests où les valeurs continues seraient exceptionnellement observées, ou alternativement méthodes de sélection de modèles probabilistes). Ces modélisations devraient être aussi suffisamment génériques pour produire des estimateurs des données continues non observées (avec leurs incertitudes associées) tout en ne disposant que des low data initiales, permettant ainsi d'effectuer des traitements industriels classiques déjà disponibles pour les données continues, ce qui aurait l'avantage de permettre aux low data d'être utilisées « sur le terrain » sans changer profondément les protocoles existant de traitement.

## Principales activités

**Le programme de travail de thèse s'échelonne comme suit sur les trois années.**

M1-M6 : Identification des types de low data disponibles chez SilMach. Identification des différents objectifs métiers visés par ces données. Mise en regard de la littérature scientifique sur le traitement habituel de ces types de données. Mise en évidence de la perte informationnelle (au sens théorie de l'information) entre les low data et les données continues collectées en temps réel (de type séries chronologiques, typiquement multivariées).

M6-M12 : Modélisation probabiliste graduelle, des cas les plus simples aux plus complexes, aussi bien pour les low data que les données continues non observées. Obtention d'une mesure de qualité d'estimation et simulation des données continues non observées, depuis les low data observées. Implémentation et expérimentation.

*À la fin de la première année, un ou plusieurs modèles probabilistes seront implémentés et testés, permettant de répondre à des questions liées à des capteurs (passifs, actifs) « moyennement » complexes. L'étudiant(e) sera familiarisé(e) avec les contraintes métier et techniques.*

M12-M18 : Dans le cadre de ces capteurs (passifs, actifs) moyennement complexes, utilisation de la modélisation pour donner des préconisations sur la conception de capteurs passifs optimaux.

M18-M32 : Reprise des étapes précédentes mais cette fois avec des capteurs passifs et actifs complexes (cas multivariés, capteurs en corrélations, séries chronologiques multiples, données non manquantes au hasard). De nouveau on ira jusqu'à l'implémentation et l'expérimentation sur données réelles et sur le terrain.

M32-M36 : finalisation et rédaction du rapport de thèse

## Compétences

Les compétences attendues sont

- Connaissance de niveau M2 en statistique et machine learning
- Connaissances en probabilités et en estimation/simulation pour données manquantes
- Compétences en R et Python
- Bonne compétence en Anglais

## Avantages

- Restauration subventionnée
- Transports publics remboursés partiellement
- Congés: 7 semaines de congés annuels + 10 jours de RTT (base temps plein) + possibilité d'autorisations d'absence exceptionnelle (ex : enfants malades, déménagement)
- Possibilité de télétravail et aménagement du temps de travail
- Équipements professionnels à disposition (visioconférence, prêts de matériels informatiques, etc.)
- Prestations sociales, culturelles et sportives (Association de gestion des œuvres sociales d'Inria)
- Accès à la formation professionnelle
- Sécurité sociale

## Rémunération

1ère et 2ème année : 2100 € (salaire mensuel brut)

3ème année : 2190 € (salaire mensuel brut)

le salaire est donné à titre indicatif car le financement principal se fera via le Cerema

## Informations générales

- **Thème/Domaine** : Optimisation, apprentissage et méthodes statistiques Statistiques (Big data) (BAP E)
- **Ville** : Villeneuve d'Ascq
- **Centre Inria** : [Centre Inria de l'Université de Lille](#)
- **Date de prise de fonction souhaitée** :2024-10-01
- **Durée de contrat** :3 ans
- **Date limite pour postuler** :2024-06-30

## Contacts

- **Équipe Inria** : [MODAL](#)
- **Directeur de thèse** :  
Biernacki Christophe / [Christophe.Biernacki@inria.fr](mailto:Christophe.Biernacki@inria.fr)

## A propos d'Inria

Inria est l'institut national de recherche dédié aux sciences et technologies du numérique. Il emploie 2600 personnes. Ses 215 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3900 scientifiques pour relever les défis du numérique, souvent à l'interface d'autres disciplines. L'institut fait appel à de nombreux talents dans plus d'une quarantaine de métiers différents. 900 personnels d'appui à la recherche et à l'innovation contribuent à faire émerger et grandir des projets scientifiques ou entrepreneuriaux qui impactent le monde. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 200 start-up. L'institut s'efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l'économie.

**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.

## Consignes pour postuler

### 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.