

REF : DSR_2020_IVA_08_01

Pour postuler : stages@irt-systemx.fr et en copie mostepha.khouadjia@irt-systemx.fr et latifa.oukhellou@ifsttar.fr

Mise en œuvre et évaluation de stratégies de détection d'anomalies au sein d'un réseau de transports en commun

CONTEXTE DU STAGE

Au sein de l'Institut de Recherche Technologique SystemX, situé au cœur du campus scientifique d'excellence mondiale de Paris-Saclay, vous prendrez une part active au développement d'un centre de recherche technologique de niveau international dans le domaine de l'ingénierie numérique des systèmes. Adossé aux meilleurs organismes de recherche français du domaine et constitué par des équipes mixtes d'industriels et d'académiques, ce centre a pour mission de générer de nouvelles connaissances et solutions technologiques en s'appuyant sur les percées de l'ingénierie numérique et de diffuser ses compétences dans tous les secteurs économiques.

Vous serez encadré par un ingénieur chercheur SystemX du domaine Science des données, IA et Interaction.

Vous travaillerez au sein du projet de recherche SystemX IVA (Information Voyageur Augmentée | <https://www.irt-systemx.fr/project/iva/>) dont les partenaires sont SNCF, Île-de-France Mobilités (ex STIF), Kisio digital, SpirOps, et l'IFSTTAR.

Le poste est basé à l'IRT SystemX - Gif sur Yvette [site du Moulon].

DUREE ET DATE DE DEMARRAGE

- Durée du stage : 6 mois avec un démarrage envisagé en mars 2020.
- Stage rémunéré avec participation aux frais de restauration à la cantine entreprise, et aux frais de transport en commun.

PRESENTATION DETAILLEE DU SUJET

Ce stage a pour ambition de mettre en œuvre et d'expérimenter plusieurs modèles de détection d'anomalies (perturbation, incidents, ...) sur le réseau de transport en commun francilien. Cette analyse permettra de mieux comprendre les modalités de propagation d'une perturbation dans le temps et dans l'espace. On cherchera à mettre en évidence les causalités qui peuvent fortement influencer la diffusion et la variabilité des flux de passagers entre les stations.

Ces modèles ont été développés dans le cadre du projet IVA à partir de données télé-billettique et de charges à bord dans les trains. Ils sont basés sur des méthodes d'apprentissage automatique.

L'objectif sera de comparer et confronter plusieurs stratégies de détection sur la base d'un historique d'incidents déclarés par l'opérateur du réseau. Il sera intéressant de quantifier l'impact de ces anomalies à différentes échelles (train, station, ligne, réseau, ...) voire de construire un graphe d'impact sur la base des détections relevées par les différents modèles.

Vos missions seront les suivantes :

- Etat de l'art méthodologique et applicatif.
- Mettre en œuvre et évaluer les stratégies de détection d'anomalies au sein d'un réseau de transport sur la base d'un historique d'incidents enregistrés.
- Analyser et confronter des résultats obtenus par différents modèles de détection d'anomalies et mesurer l'impact des perturbations associées à ces détections sur différents échelles (mission des trains, station, ligne, réseau, ...).

◆ Références :

- [1] Zhang, N., Graham, D.J. and Carbo, J.M. *Using smart card data to analyse the disruption impact on urban metro systems*. Transportation Research Board 98th Annual Meeting Transportation Research Board, (2019).
- [2] Sun, H., Wu, J., Wu, L., Yan, X., & Gao, Z. Estimating the influence of common disruptions on urban rail transit networks. *Transportation Research Part A: Policy and Practice*, 94, 62-75, (2016).
- [3] Kilpi, J., Koskinen, S., & Scholliers, J. (2019). Detection of anomalies in urban traffic from open data. 13th ITS European Congress, Brainport Eindhoven, Netherlands.
- [4] Silva, R., Kang, S. M., & Airoidi, E. M. (2015). Predicting traffic volumes and estimating the effects of shocks in massive transportation systems. *Proceedings of the National Academy of Sciences*, 112(18), 5643-5648.
- [5] A-S. Briand, E. Côme, M. Khouadjia, L. Oukhellou, Detection of atypical events on a public transport network using smart card data, European Transport Conference, Dublin, 2019
- [6] K. Pasini, M. Khouadjia, A. Samé, F. Ganancia, L. Oukhellou, [LSTM encoder-predictor for short-term train load forecasting](#), European Conference on Machine Learning (ECML/PKDD), Würzburg, 2019.

PROFIL ET COMPETENCES

Etudiant BAC+5 en fin de cycle école d'ingénieur ou Master 2, dans le domaine de la science des données, IA, statistiques, mathématiques appliquées ou informatique avec une spécialisation en science des données et le *machine learning*.

Compétences :

- Bon niveau de maîtrise dans l'un des langages de programmation (Python, R).
- Machine Learning, statistiques.
- Base de données.

Aptitudes personnelles :

- Esprit d'analyse, autonomie, travail collaboratif.
- Intérêt pour la recherche appliquée.
- Aptitude à communiquer aussi bien à l'oral qu'à l'écrit (en français et anglais).
- Intérêt pour les transports, la mobilité.