

ANALYSE DE DONNEES ET APPRENTISSAGE POUR AMELIORER LE REALISME D'UN SIMULATEUR DU SYSTEME FERROVIAIRE

SNCF INNOVATION & RECHERCHE

OFFRE DE STAGE 2019 EN INFORMATIQUE ET DATA SCIENCE

Ce sujet est proposé à un étudiant de M1 ou M2 (école d'ingénieur ou université) spécialisé en informatique et data science. Il peut également convenir pour un stage de césure. Il sera encadré par Mathieu Gagnon (contact : mathieu.gagnon@sncf.fr).

Date du début souhaitée : printemps-été 2019. **Durée** : 4 à 6 mois. **Lieu** : Saint-Denis.

CONTEXTE

Au sein de la Direction Innovation & Recherche de SNCF qui regroupe des dizaines d'experts de toutes spécialités (sciences physiques, confort à bord des trains, data science, nouveaux usages voyageurs, télécoms, capteurs,...), le groupe Modélisation et Optimisation de la Décision aide SNCF à répondre aux défis de la mobilité à haut débit de demain.

ENJEUX

Avec plus de 5 millions de voyageurs et 15000 trains par jour, les défis liés à l'information voyageurs et la gestion du trafic sont cruciaux pour SNCF. Pour essayer de répondre à ces enjeux, un simulateur à événements discrets – qui vise à représenter l'évolution réelle du système ferroviaire – a été développé à la direction Innovation & Recherche.

Les événements considérés ici correspondent aux départs, passages ou arrivées de trains sur une partie du réseau (des gares en particulier). A la fin d'une simulation, on obtient les horaires de passage simulés aux différents points du réseau.

Le déroulement des événements dans la simulation est soumis à des contraintes (espacement minimum à respecter entre les trains, ordre de passage en zone de convergence, temps de parcours théorique) qui visent à représenter les contraintes physiques du système ferroviaire (signalisation, vitesse des trains). **Les contraintes de la simulation ne sont cependant pas complètement fidèles à la réalité, et il est important de calibrer le simulateur pour en améliorer sa performance.**

SNCF dispose d'importants **historiques de données** (horaires observés des trains) qui peuvent à ce titre permettre d'améliorer le réalisme du simulateur.

OBJECTIFS DU STAGE

Le stagiaire devra dans un premier temps se familiariser avec le simulateur, avant d'implémenter de nouveaux modèles pour représenter les contraintes de la simulation – en particulier les **temps de parcours** et la **propagation des retards** – qui peuvent dépendre de différents paramètres (heure de départ du train, type de matériel, zone géographique, etc.). L'étudiant pourra s'appuyer sur des travaux réalisés courant 2018 au sujet de la sélection de retards mineurs et intrinsèques que l'on souhaite injecter dans la simulation.

Pour évaluer les modèles conçus, l'étudiant devra les comparer au modèle de simulation actuel. Pour cela, il disposera d'un projet JAVA existant permettant de lancer des simulations en masse sur des journées passées.

Ce stage requiert les compétences suivantes :

- Appétence pour le traitement de données
- Esprit d'adaptation et rigueur scientifique
- Appétence pour la recherche, autonomie, force de proposition
- Maîtrise du langage JAVA, connaissances en Python ou R

A noter que le stagiaire sera amené à travailler en support d'un doctorant qui travaille sur le même sujet. Un rapport détaillé et soigné sur l'aspect scientifique sera demandé. Le stagiaire devra également rendre compte régulièrement des avancées de son travail.