

Sujet PLDAC

Analyse de traces utilisateurs dans les réseaux de transport en commun

Vincent Guigue - Nicolas Baskiotis - `prenom.nom@lip6.fr`

Sujet principalement logiciel (peu de bibliographie)

Sujet ouvert à un monôme ou un binôme

Dans le cadre d'un partenariat avec le STIF (Syndicat des Transports en Ile de France), nous travaillons sur les données de validation du métro parisien (liées aux pass Navigo). Le but de ce PLDAC est d'analyser la mobilité des Franciliens dans les transports en commun. Nous nous intéresserons aux différentes motivations des trajets : trajets quotidiens pour le travail, trajets réguliers pour certaines activités, trajets plus rares. L'idée est de croiser une base de données de validation (NAVIGO) avec différentes données disponibles sur le web : données INSEE (revenus, densité de population,...), données google map (cinéma, bar, restaurant, ...), opendata RATP (pour les durées de transport), données tripadvisor, foursquare...

A l'issue du processus d'analyse, nous voulons identifier des comportements typiques en fonction des lieux d'habitat : qui prend le bus, le métro, combien de temps mettent les usagers pour se rendre à leur travail, qui sort le soir, traîne le week-end... Ces comportements seront normalisés puis re-projetés sur la carte parisienne pour proposer une vision originale du territoire Francilien.

Les principales étapes du projets sont les suivantes : 1) prise en main de la base de données des validations 1bis) détermination des éléments de profils visés 2) crowling de données connexes 3) construction des profils, exploitation, projection sur la carte francilienne.

Autour de ce cadre, de nombreuses variantes sont possibles et seront discutées avec les candidats au début du projet puis au cours de son évolution.

Les développements seront principalement effectués en python et l'accès aux données se fait en SQL. D'autres langages sont négociables en concertation avec l'étudiant.