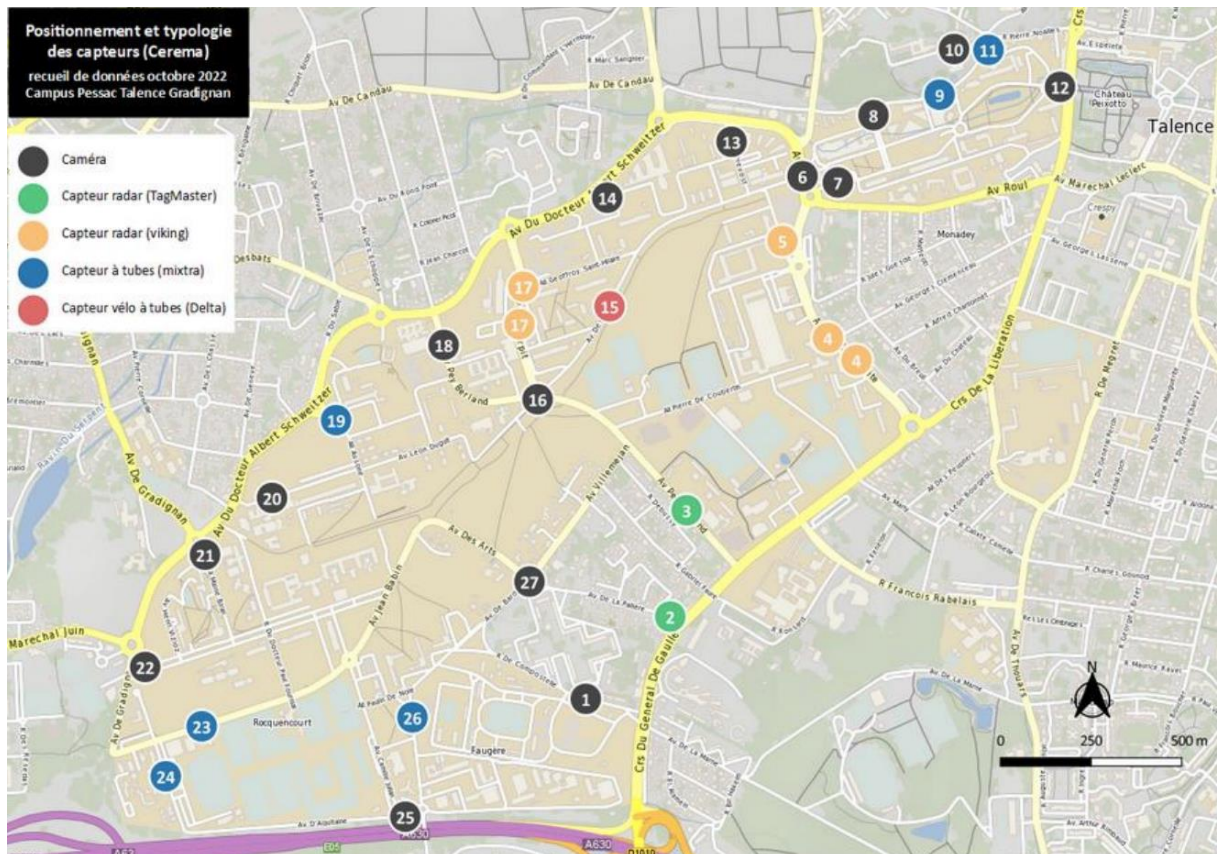


Sujet PLE 2022-2023

Analyse de données de capteurs de circulation

Nous disposons d'un ensemble de capteurs installés sur le domaine universitaire. Ces capteurs permettent de détecter le passage d'un type de véhicule et du sens de son sens de circulation sur les différentes entrées de l'université. Le type de capteur utilisé ainsi que leur localisation sur le campus sont donnés sur la carte ci-dessous.



Les capteurs nous permettent d'obtenir des informations de différent types et dans différent format en fonction des capteurs. Les données sont récupérées dans un fichier unique et distinct pour chaque capteur et stocker au format brut dans dans « Master Dataset HDFS ».

L'objectif du projet est mettre en place un dashboard de visualisation de ses données. Ce dashboard devra permettre prioritairement (1) de suivre l'évolution du trafic, par heure, par jour, par capteur et par type de véhicule, (2) de voir les points de passages les plus fréquentés en fonctions des jours, heures de la journée.

Dans une deuxième temps nous voulons mettre en place des outils d'analyse des données. Par exemple, pour représenter au cours de la journée, le nombre de véhicule se trouvant sur le site. (Différence entre les véhicule entrant et sortant). Un autre exemple est de permettre de mettre en évidence les véhicules traversant l'université uniquement pour raccourcir leur trajet. Vous êtes libre d'intégrer tout type d'analyse sur ce type de données.

Travail demandé, Jalon 1 [DataCleaning]:

Dans un premier temps le travail demandé consiste récupérer les données des différents capteurs, à les fusionner dans un seul format que nous pourrons utiliser ultérieurement. Le résultat devra être stocké dans un `sequenceFile` sur HDFS.

- Mesurez le temps de traitement nécessaire pour le datacleaning
- Évaluez la scalabilité horizontale de votre solution.
- Estimez l'infrastructure nécessaire si nous voulons analyser les données pendant 10 ans à l'échelle d'une métropole (nombre de capteurs multipliés par 10.000).

Travail demandé, Jalon 2 [BatchProcessing]:

Dans cette deuxième étape, vous devez mettre en place les pré-calculs nécessaires pour la génération de dashboard par votre « serving layer ». L'ensemble des pré-calculs devront être stockés dans HBASE. Pour chacun des traitements que vous implémentez indiquez les ressources nécessaires, le goulots d'étranglement et mesurez les performances de traitement.

Travail demandé, Jalon 3 [DashBoard]:

En utilisant Leaflet, D3.js, Bootstrap, ou tout autre framework web mettez en place un dashboard permettant de visualiser les informations stockées dans votre « Serving Layer ».

Annexe exemple de donnée des capteurs

Avec les capteurs radar nous collectons les données suivante :

SENS,JOUR,HEURE/MINUTE,SECONDE/CENTIEME,VITESSE,SER,TYPE

1,3,1530,1466,V=042,S=0618,VL

1,3,1530,4347,V=040,S=0613,VL

2,3,1531,3593,V=055,S=0703,VL

1,3,1531,3756,V=041,S=0660,VL

1,3,1531,4736,V=040,S=0631,VL

1,3,1531,3210,V=030,S=0592,VL

Sens 1 : sortie Fac.

sens 2 : entrée Fac

$SER < 205$ 2RM (2 roues); $205 < SER < 1140$ VL (voiture) $SER > 1140$: PL (camion)

Avec les capteurs de tube les données suivantes :

Voie,Horodate,de,passage,Seconde,Centième,Vitesse,(km/h),Type véhicules,EntreFront,Catégorie, Longueur,(dm),T,Presence,(ms),D,EntreFront,(m),Silhouette,Catégorie,STERELA,Poids,total,Distance inter- essieux,Nombre,d'essieux,Poids,E1,(dt),Poids,E2,(dt),Poids,E3,(dt),Poids,E4,(dt),Poids,E5,(dt),Poids,E6,(dt), Poids,E7,(dt),Poids,E8,(dt),Poids,E9,(dt),Distance,F/E1,(cm),Distance,E1/E2,(cm),Distance,E2/E3,(cm), Distance,E3/E4,(cm),Distance,E4/E5,(cm),Distance,E5/E6,(cm),Distance,E6/E7,(cm),Distance,E7/E8,(cm), Distance,E8/E9,(cm),Catégorie,E1,Catégorie,E2,Catégorie,E3,Catégorie,E4,Catégorie,E5,Catégorie,E6, Catégorie,E7,Catégorie,E8,Catégorie,E9,Coef,Piézo,Attaque,Coef,Piézo,Sortie,Véhicule,Valide,Phase, Fonctionnement,Amplitude,Attaque,Amplitude,Sortie

```
0,16/1/20,0:26,32,6,25,0,,VL,0,1,5,,2,,,,,,,,,0,307,,,,,,,,,,,,,1,0,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,
0,16/1/20,0:43,28,19,21,0,,VL,0,1,5,,2,,,,,,,,,0,256,,,,,,,,,,,,,1,0,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,
0,16/1/20,1:24,21,69,27,0,,VL,0,1,5,,2,,,,,,,,,0,231,,,,,,,,,,,,,1,0,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,
0,16/1/20,1:36,4,27,17,0,,PL,0,2,10,,2,,,,,,,,,0,397,,,,,,,,,,,,,1,0,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,
```

Et avec les capteurs caméra les données suivantes :

| ID | Catégorie | date | heure | vers jardin public (nord_sud) | vers gare (sud-nord) | remarques cyclistes |
|----|-----------|------------|----------|-------------------------------|----------------------|---------------------|
| 2 | VL | 11/05/2022 | 11:00:01 | x | | |
| 4 | VL | 11/05/2022 | 11:00:08 | x | | |
| 5 | VL | 11/05/2022 | 11:00:26 | | x | |
| 9 | UT | 11/05/2022 | 11:00:43 | | x | |
| 12 | VL | 11/05/2022 | 11:00:57 | | x | |
| 13 | UT | 11/05/2022 | 11:01:05 | | x | |
| 20 | VL | 11/05/2022 | 11:01:19 | x | | |