

ING1: PROJET DE PARCOURS SIE EXTRACTION AUTOMATIQUE DE DONNEES

Rédigé par : Yannick Le Nir et Rachid Chelouah	Ref : ING1-PRO-PAR-SIE
A l'intention de : Etudiants des ING1	Créé le : 07/06/2013

1. Introduction

OpenStreetMap est une carte du monde librement modifiable. La communauté des utilisateurs intervient donc dans la création/modification des éléments cartographiés. Il est de plus, extrêmement facile d'extraire des parties de la carte et de les exporter sous différents formats, notamment en XML.

Dès lors, il s'agit d'un objet d'étude très intéressant pour manipuler de l'information réelle (il ne s'agit plus d'un jeu de données virtuel et souvent limité, mais des données cartographiques complètes et régulièrement mises à jour).

L'objectif de ce projet de parcours est double. Il s'agit dans un premier temps d'analyser le format XML associé à OpenStreetMap, afin d'isoler certaines composantes qui vous semblent utiles pour la réalisation d'un outil d'extraction, d'analyse et de reporting cartographique. Un tel outil se présente sous la forme d'une liste d'actions effectuant des analyses plus ou moins simples sur les données de la carte. Vous pouvez par exemple vouloir récupérer l'ensemble des magasins répertoriés sur une zone géographique, ou encore déterminer le nombre moyen de feux de croisement par type de route, ou encore effectuer des comptages par zone géographique ,...

L'ensemble de ces analyses peuvent être alors représentées (ce que nous appelons reporting), via des diagrammes adaptés, par exemples camembert ou barres.

Le second objectif de ce projet est donc de réaliser un outil informatique permettant ces analyses, constitué de deux parties distinctes : une base de données contenant l'ensemble des informations extraites de la carte organisées afin d'être adaptées aux requêtes que vous souhaitez réaliser ; une application Java permettant d'interroger la base de donnée afin de produire les résultats de l'analyse sous forme de diagrammes.

2. Modalités

Afin de réaliser ce projet par binômes, nous allons vous fournir deux cartes extraites d'OpenStreetMap, l'une assez petite (une centaine de kilo-octets) représentant votre campus (Cergy ou Pau) et ses proches alentours, et l'autre représentant une grande partie de votre ville (d'une centaine de méga-octets).

Vous allez devoir utiliser ces données afin de réaliser votre application d'extraction d'analyse et de reporting.

Vous devrez fournir une archive contenant l'ensemble des programmes réalisés ainsi qu'un rapport détaillant la réalisation et l'utilisation de votre projet, en incluant les aspects gestions de projets.

Une soutenance complétera le rapport et aura lieu le vendredi 14 juin, avec possibilité d'effectuer une démonstration.

3. Description du logiciel à développer

Le projet consiste à modéliser une base de données qui va servir de sauvegarde des données définies dans le fichier XML d'implémenter cette base de données de l'alimenter et de faire des traitements statistique sur les données ainsi remplies. Le détail de ce logiciel est défini ci-dessous.

3.1 Base de données

- 1. Analyse du format XML issu du logiciel OpenStreetMap
- 2. Modélisation sous Merise d'un modèle de données pour la sauvegarde des données issues du fichier XML
- 3. Implémentation du script de création de la base de données

EXTRACTION AUTOMATIQUE DE DONNEES

4. Implémentation du script d'alimentation de cette base de données (à tirer du fichier XML fourni). Ce script sera généré automatiquement à l'aide d'un moteur XSLT.

3.2 Statistiques descriptives

Vous devez proposer une analyse de ces données. Cette analyse doit être suffisamment complexe afin d'utiliser des analyses univariées et multivariées (bivariées) des différents composants cartographiques que vous avez choisi d'extraire. Le rendu des résultats seront numériques et graphiques (se conférer à l'application Java ci-dessous).

3.3 Application java

La troisième partie du projet consiste à réaliser un client Java permettant d'interroger la base de données afin d'en extraire les informations nécessaires au traitement et au reporting (visualiser sous forme de diagrammes camembert ou barre en utilisant l'API JFlowChart).

Les différentes étapes sont les suivantes :

- Découverte de JDBC pour Oracle via le tutoriel présent sur AREL, permettant la communication entre Java et votre base de données. Vous n'avez pas à implémenter cette fonctionnalité que vous verrez en l'an prochain dans le parcours SIE. Vous devez simplement suivre le tutoriel afin de l'appliquer pour votre projet.
- 2. Écrire les requêtes SQL nécessaires à votre projet pour remplir des les tableaux nécessaire à vos traitements statistiques.
- 3. Utilisation de JflowChart pour le reporting.
- 4. Écrire une IHM Java permettant de manipuler vos requêtes de reporting aisément et de visualiser les résultats sous forme de diagrammes en utilisant l'API JFlowChart que vous devez télécharger.

4. Classes Java, Packages Java et jeux de tests fournis

- Des briques logiciels de JDBC : pour communiquer entre une application Java et base de données SQL
- JFlowChart : pour des représentations graphiques
- Deux Datasets : pour les jeux de données tirés de l'application OpenStreetMap.