

Gestion de Connaissances en Temps Réel depuis des Flux Massifs de Données et Apprentissage Automatique

Badre Belabbess*, Jérémy Lhez**
Olivier Curé***

*badre.belabbess@atos.net

**jeremy.lhez@u-pem.fr

***olivier.cure@u-pem.fr

Résumé. L'analyse en temps-réel de données massives envoyées par des capteurs a connu ces dernières années un essor important. Du fait de l'hétérogénéité de ces données, l'application de modèles de machine learning spécialement calibrés pour des cas d'usages précis a permis d'extraire et d'inférer des informations de très grandes valeurs. Néanmoins, peu de systèmes proposent une implémentation distribuée sur un vrai cluster industriel permettant de tirer profit de capacités de calcul décuplées. Nous présentons ici une démonstration de détection d'anomalie sur réseau souterrain d'eau potable en île-de-France réalisé avec notre plateforme, dénotée WAVES.

1 Introduction

Les avancées technologiques en termes de communication sans fil et de microélectronique ont mené au développement de capteurs intelligents toujours plus efficaces et déployables à large échelle. Les domaines d'application se sont alors rapidement diversifiés avec, entre autres, la surveillance d'habitat (A. Rozyyev et F.Subhan (2011)), la géolocalisation d'objets communicants (S.Chauhdary (2009)) et la gestion d'environnement (L.Lee et C.Chen (2008)). Le recours intensif à ces capteurs a conduit à la génération d'un large volume de mesures dynamiques, hétérogènes et géographiquement distribuées. Si ces informations sont analysées efficacement, cela pourrait aider à inférer automatiquement de nouvelles connaissances à haute valeur ajoutée.

Le système décrit ici s'inscrit dans un projet de recherche pour le déploiement d'une solution industrielle nommée WAVES¹. Il s'agit d'une plateforme de traitement en temps-réel de flux massifs provenant de capteurs installés sur un large réseau souterrain d'eau potable. Un de ces traitements correspond à la détection d'anomalies dans la consommation d'eau correspondant à une fuite sur le réseau. Ce projet est né de la nécessité de trouver des solutions innovantes pour réduire les déperditions d'eau qui sont estimées en moyenne à 20% du volume d'eau introduit dans le réseau². Dans cette démonstration, nous présenterons l'architecture globale du

1. Détails sur le projet WAVES disponible à l'adresse : <http://waves-rsp.org/>

2. Rapport de l'Observatoire des services publics d'eau et d'assainissement : www.services.eaufrance.fr/docs/synthese/rapport/Rapport_SISPEA\%202011_resume_DEF.pdf/