

Cahier des charges de recherche

1. Sujet de recherche

Problème de tournée pollution en temps réel : Une approche basée sur l'analyse des big data avec Apache Spark

2. Contexte et problématique

Dans un contexte de lutte contre le changement climatique et de réduction de l'empreinte écologique, la gestion optimale des tournées de livraison devient cruciale. Le problème de tournée pollution (Pollution Routing Problem, PRP) est une variante du classique problème de tournée de véhicules (Vehicle Routing Problem, VRP), qui intègre les émissions de gaz à effet de serre comme critère d'optimisation.

Dans un environnement urbain dynamique, des paramètres comme le trafic, les conditions météorologiques ou les variations des demandes doivent être pris en compte en temps réel. L'émergence des technologies Big Data et de cadres d'analyse comme Apache Spark offre une opportunité de traiter ces données massives pour concevoir des solutions plus réactives et écologiques.

3. Objectifs de la recherche

1. **Modélisation du PRP en temps réel :**
 - Identifier les paramètres clés affectant les émissions de CO2 en tournée.
 - Construire un modèle adapté à des données dynamiques.
2. **Implémentation avec Apache Spark :**
 - Développer un pipeline de traitement des données en temps réel.
 - Exploiter les bibliothèques Spark pour le traitement distribué et les algorithmes d'optimisation.
3. **Validation des performances :**
 - Comparer la solution à des approches statiques.
 - Évaluer les réductions d'émissions et le gain de réactivité.

4. Méthodologie

4.1. Étude bibliographique

Une revue approfondie des travaux précédents sur :

- Les variantes du VRP et PRP.
- Les approches big data appliquées à l'optimisation.
- Les outils Apache Spark (Spark Streaming, MLlib, GraphX).

4.2. Collecte et traitement des données

- **Sources potentielles :**

- Données de trafic en temps réel (API Google Maps, HERE).
- Données météorologiques.
- Données sur la consommation et les émissions des véhicules.
- **Traitement :**
 - Extraction, transformation et chargement (ETL) via Spark.
 - Construction d'un entrepôt de données dynamique.

4.3. Modélisation et optimisation

- **Modèle :** Formulation mathématique intégrant les émissions de CO2.
- **Algorithmes :** Utilisation d'algorithmes génétiques, recuit simulé ou autres méthodes heuristiques optimisées avec MLlib.

4.4. Expérimentation

- Simulation de scénarios réels.
- Comparaison avec des modèles existants sur des indicateurs clés (temps, émissions, coûts).

5. Livrables

1. Rapport de recherche détaillant la modélisation, les méthodes et les résultats.
2. Un pipeline fonctionnel implémenté avec Apache Spark.
3. Un article scientifique soumis à une revue ou conférence internationale.

6. Planification prévisionnelle

Phase	Durée	Description
Étude bibliographique	2 mois	Revue des travaux existants
Collecte et préparation des données	1,5 mois	Mise en place de l'infrastructure
Modélisation et développement	2 mois	Implémentation avec Spark
Validation et expérimentation	1 mois	Analyse des performances
Rédaction et présentation	1,5 mois	Rapport final et soutenance

7. Références scientifiques

1. R. Kramer, N. Maculan, A. Subramanian and T Vidal, "A speed and departure time optimization algorithm for the pollution-routing problem", *European Journal of Operational Research*, vol. 247, no. 3, pp. 782-787, 2015
2. R. Kramer, A. Subramanian, T. Vidal and F. C Lucídio dos Anjos, "A matheuristic approach for the pollution-routing problem", *European Journal of Operational Research*, vol. 243, no. 2, pp. 523-539, 2015.
3. I. Sbai and S Krichen, "A real-time decision support system for big data analytic: A case of dynamic vehicle routing problems", *Procedia Computer Science*, vol. 176, pp. 938-947, 2020.
4. N. Ouertani, H. Ben-Romdhane, S. Krichen and I Nouaouri, "A vector evaluated evolutionary algorithm with exploitation reinforcement for the dynamic pollution

routing problem", *Journal of Combinatorial Optimization*, vol. 44, no. 2, pp. 1011-1038, 2022.

5. Erdogan, S., & Miller-Hooks, E. (2012). A Green Vehicle Routing Problem. *Transportation Research Part E: Logistics and Transportation Review*.
6. Demir, E., Bektaş, T., & Laporte, G. (2014). A review of recent research on green road freight transportation. *European Journal of Operational Research*.
7. Spark official documentation. Apache Spark. <https://spark.apache.org>.
8. Türkoglu, F., Karaoglan, I., & Atalay, B. (2018). Pollution-routing problem with time windows for dynamic settings. *Computers & Operations Research*.