

Rapport de projet : Rectangle Tree

Hugo Charels 000544051 Mickael Kovel 000396950

28 Avril 2023

Contents

1	Introduction	3
2	Structure	3
3	Création	4
3.1	Split quadratique	4
3.1.1	Pick Seed	4
3.1.2	Pick Next	4
3.2	Split linéaire	4
3.2.1	Pick Seed	5
3.2.2	Pick Next	5
4	Recherche	5
5	Expériences sur donnees réelles	5
5.1	Belgique - Secteurs statistiques	5
5.2	France - Communes	5
5.3	Monde - Pays	5
5.4	Monde - Villes	5
5.5	Analyse	5
6	Conclusion	5
7	Références bibliographiques	5

1 Introduction

Dans de nombreuses applications, la manipulation de polygones est essentielle. Ces polygones peuvent représenter des régions administratives, des informations géologiques, des zones de dessins vectoriels ou encore des images médicales. Lorsqu'il s'agit de déterminer à quel polygone appartient un point donné, une question se pose : comment résoudre efficacement le problème du "Point in Polygon" (PIP) ?

L'algorithme PIP, connu depuis 1962, consiste à compter le nombre de fois qu'une demi-droite partant du point traverse une arête du polygone. Si ce nombre est pair, le point est à l'extérieur du polygone, sinon, il est contenu à l'intérieur. Cependant, si l'on a des milliers de polygones complexes à tester avec un grand nombre de points, une méthode naïve devient inefficace.

Une première optimisation consiste à utiliser l'enveloppe ou "minimum bounding rectangle" (MBR), qui est le plus petit rectangle horizontal englobant totalement le polygone. Avant d'appliquer l'algorithme PIP, on vérifie d'abord l'inclusion du point dans le MBR, ce qui permet de réduire le nombre d'appels à PIP.

Cependant, lorsque le nombre de polygones et de MBR est élevé, une approche hiérarchique basée sur l'algorithme R-Tree devient pertinente. En regroupant les MBR de manière optimale, on peut éviter de considérer tous les MBR pour chaque point à tester.

Dans ce projet, nous allons implémenter et évaluer deux variantes de l'algorithme R-Tree (quadratique et linéaire) pour résoudre efficacement le problème du PIP. Nous testerons ces variantes avec différents paramètres pour évaluer leurs performances.

Le rapport détaillera la méthodologie utilisée, les résultats obtenus et une analyse approfondie des performances des variantes de l'algorithme R-Tree. Nous discuterons également des défis rencontrés et des perspectives pour de futures recherches.

En conclusion, ce projet vise à résoudre efficacement le problème du PIP en utilisant l'algorithme R-Tree. Les résultats obtenus nous permettront de déterminer la meilleure variante en fonction des paramètres et de formuler des recommandations pour d'éventuelles améliorations.

2 Structure

Dans cette section, nous allons explorer la structure d'un R-Tree, une arborescence utilisée pour organiser efficacement les données géospatiales et résoudre le problème du "Point in Polygon" (PIP). L'objectif est de comprendre comment cette structure hiérarchique permet de regrouper les MBR (minimum bounding rectangles) de manière optimale.

Un R-Tree est une structure de données arborescente utilisée pour l'indexation spatiale des objets géométriques. Il est conçu pour stocker et organiser efficacement des données multidimensionnelles, telles que des polygones, des points ou

des rectangles.

L'idée fondamentale d'un R-Tree est de regrouper les MBR de manière hiérarchique. Chaque nœud de l'arbre représente un MBR, qui peut être un polygone ou un rectangle englobant un groupe de MBR plus petits.

- Nœuds internes : Les nœuds internes de l'arbre contiennent des MBR qui englobent plusieurs autres MBR. Ils servent de guides pour naviguer dans la structure et réduire la recherche de MBR pertinents lors de la résolution du problème du PIP.
- Feuilles : Les nœuds feuilles de l'arbre contiennent les MBR individuels et sont généralement associés à des objets géométriques spécifiques. Ils sont utilisés pour effectuer les tests d'inclusion du point lors de la résolution du PIP.

L'algorithme R-Tree utilise des critères spécifiques pour regrouper les MBR de manière optimale. Ces critères visent à maximiser la compacité des groupes de MBR et minimiser leur superposition avec d'autres groupes.

- Critère de regroupement spatial : Les MBR ayant une proximité spatiale élevée sont regroupés ensemble pour réduire la recherche dans les parties de l'arbre non pertinentes pour le point en cours d'évaluation.
- Critère de compacité : Lorsque plusieurs MBR sont candidats pour un regroupement, celui qui crée le MBR englobant le plus compact est préféré. Cela garantit une meilleure utilisation de l'espace de stockage et une réduction des opérations de recherche.

Nous détaillerons la recherche d'un point dans un R-Tree dans la section 4.

3 Création

ici on va écrire la création

3.1 Split quadratique

ici on va écrire le split quadratique

3.1.1 Pick Seed

ici on va écrire le pick seed

3.1.2 Pick Next

ici on va écrire le pick next

3.2 Split linéaire

ici on va écrire le split linéaire

3.2.1 Pick Seed

ici on va ecrire le pick seed

3.2.2 Pick Next

ici on va ecrire le pick next

4 Recherche

ici on va ecrire la recherche

5 Expériences sur donnees réelles

ici on va ecrire les experiences sur donnees réelles

5.1 Belgique - Secteurs statistiques

ici on va ecrire les experiences sur les secteurs statistiques

5.2 France - Communes

ici on va ecrire les experiences sur les communes

5.3 Monde - Pays

ici on va ecrire les experiences sur les pays

5.4 Monde - Villes

ici on va ecrire les experiences sur les villes

5.5 Analyse

ici on va ecrire l'analyse

6 Conclusion

ici on va ecrire la conclusion

7 Références bibliographiques

ici on va ecrire les references bibliographiques