Uma imagem com Tipo de letra, logótipo, símbolo, Gráficos

Descrição gerada automaticamente

Uma imagem com Saturação de cores, Lilás

Descrição gerada automaticamente

**Comparing similarity between two territory partitions in districting problems**

**Indices and practical issues**

**Sene Conté**

Thesis to obtain the Master of Science Degree in

**Computer Science and Engineering**

Supervisors: Prof. Luís Manuel Silveira Russo Prof. José Rui De Matos Figueira

**Examination Committee**

Chairperson: Prof. Pedro Tiago Gonçalves Monteiro

Supervisor: Prof. Luís Manuel Silveira Russo

Member of the Committee: Prof. Vasco Miguel Gomes Nunes Manquinho

**(Month Year)**

**Declaration**

I declare that this document is an original work of my own authorship and that it fulfills all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

**Abstract**

This study addresses the complex problem of assessing similarity between two partitions of a district map, where each partition is modeled as a connected, undirected, and planar graph. Comparing similarity between such partitions is crucial in applications like urban planning, political districting, and geographic analysis, where variations in partitioning can have significant impacts on representation and resource distribution. Traditional similarity indices, such as the Rand Index, Fowlkes-Mallows Index, and Jaccard Index, are used as foundational measures. However, these indices are extended to incorporate weighting factors for districts, allowing for a more accurate comparison by accounting for variations in district characteristics, such as population density or geographic importance. By using weighted similarity indices, this research aims to achieve a refined measurement of similarity that captures the nuanced impact of each zone. This approach offers an enhanced framework for partition analysis, improving upon existing techniques that often overlook regional disparities. Ultimately, the study presents a comprehensive solution for similarity assessment in partitioned maps, applicable to various fields that rely on accurate and equitable territorial analysis.

**Keywords**

Districting; Weighted Pair Counting; Similarity Indices; Urban Planning; Territorial Maps; Perturbations

**Resumo**

Este estudo aborda o problema complexo de avaliar a semelhança entre duas partições de um mapa distrital, onde cada partição é modelada como um grafo conectado, não-direcionado e planar. Comparar a semelhança entre estas partições é crucial em áreas como o planeamento urbano, a demarcação de distritos eleitorais e a análise geográfica, onde diferentes partições podem ter impactos significativos na representação e na distribuição de recursos. Utilizam-se índices de semelhança tradicionais, como o Índice de Rand, o Índice Fowlkes-Mallows e o Índice de Jaccard, como medidas fundamentais. No entanto, estes índices foram estendidos para incorporar fatores de peso dos distritos, permitindo uma comparação mais precisa ao considerar variações nas características dos distritos, como densidade populacional ou importância geográfica. Através do uso de índices de semelhança pesadas, esta investigação visa obter uma medição refinada da semelhança, capturando o impacto específico de cada zona. Esta abordagem oferece uma estrutura melhorada para a análise de partições, superando técnicas existentes que frequentemente ignoram as disparidades regionais. Por fim, o estudo apresenta uma solução abrangente para a avaliação de semelhança em mapas particionados, aplicável a diversas áreas que dependem de uma análise territorial precisa e equitativa.

**Palavras Chave**

Distritos; Contagem de Pares Pesadas; Índices de Similaridade; Planeamento Urbano; Mapas Territoriais; Perturbações.

# 

# Introduction

Political districting

Problem modeling

Motivation

Organization of the document

Pair Counting Indices

Rand Index

Fowlkes Mallows Index

Jaccard Index

Properties of the indices.

Weighted Pair Counting

Mathematical formulation

Validation of formulas

Comparing with traditional pair counting (optional)

Implementation

Algorithm

Data structure

Perturbation algorithm

Complexity analysis

Solution

Experimental results

Data Set Generation and Preprocessing

Results

Conclusion

Conclusions

System Limitation and Future Work

Bibliography

Appendix