



<b>Evento</b>	Salão UFRGS 2020: SIC - XXXII SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS
<b>Ano</b>	2020
<b>Local</b>	Virtual
<b>Título</b>	A CLUSTERIZAÇÃO ESPECTRAL E SUAS APLICAÇÕES
<b>Autor</b>	LUCAS SIVIERO SIBEMBERG
<b>Orientador</b>	LUIZ EMILIO ALLEM

# A CLUSTERIZAÇÃO ESPECTRAL E SUAS APLICAÇÕES

Lucas Siviero Sibemberg<sup>1</sup>

Prof. Dr. Luiz Emilio Allem<sup>2</sup> (orientador)

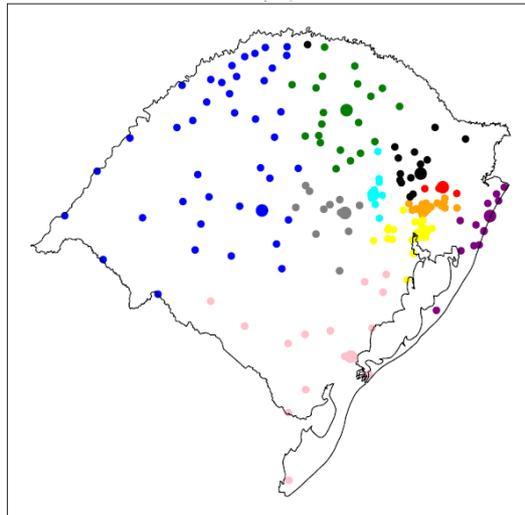
Universidade Federal do Rio Grande do Sul

## Resumo

A clusterização consiste em particionar um conjunto de dados (geralmente grande) em um número menor de grupos, onde os dados de um mesmo grupo possuem mais afinidade que dados de grupos diferentes. Essa técnica é amplamente utilizada para a análise exploratória de dados, com aplicações em diversos campos de pesquisa. Em particular, uma técnica de clusterização tem se destacado: a Clusterização Espectral. O objetivo dela é olhar para um conjunto de dados como um grafo com pesos  $G = (V, E, \omega)$ , onde os dados representam os vértices de  $G$ , o conjunto de arestas  $E$  contém as arestas conectando os elementos de  $V$  e  $\omega$  é uma função que atribui um peso positivo  $\omega_{ij}$  para cada aresta  $ij$  de  $E$ . O método utiliza os  $k$  autovetores associados aos  $k$  menores autovalores da matriz laplaciana normalizada  $\mathcal{L}$  associada ao grafo  $G$ . Para enfim, obter um particionamento do conjunto de dados. O algoritmo completo pode ser conferido em [1].

Uma aplicação prática da clusterização espectral é considerar o conjunto de  $n = 167$  municípios maiores municípios gaúchos como vértices de um grafo, utilizando uma medida baseada na migração pendular, isto é a quantidade de pessoas que viajam diariamente para outro município. A clusterização está ilustrada na Figura 1.

Figura 1 - Clusterização obtida por clusterização espectral com respeito a medida de afinidade baseada na migração pendular para  $k = 10$  clusters. Os vértices maiores correspondem a cidade mais populosa do cluster.



Fonte: Os autores (2020)

## Referências

- [1] NG, A. Y.; JORDAN, M. I.; WEISS, Y. On spectral clustering: Analysis and algorithm. **Proceedings of the 14th NIPS: Natural and Synthetic**, p. 849-856, MIT Press, 2001.

<sup>1</sup> lucas.siviero@ufrgs.br

<sup>2</sup> emilio.allem@ufrgs.br