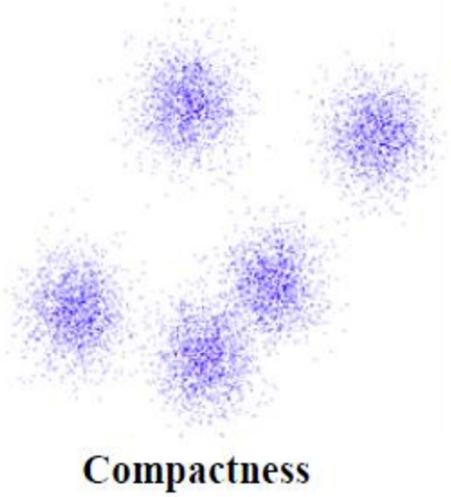
Implementation of Secure Spectral Clustering

Introduction

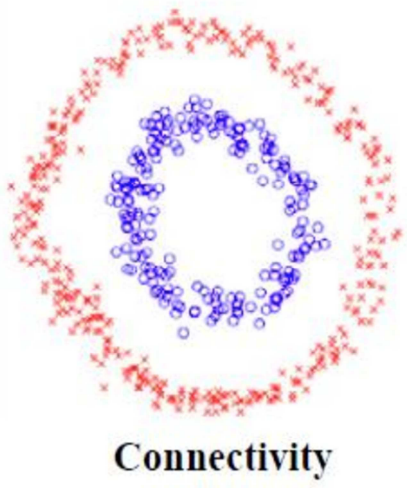
Data clustering is an important problem with many applications in Machine learning, Computer vision, and Signal processing. The object of clustering is to divide a dataset into natural groups of similar items.

Spectral clustering treats the data clustering as a graph partitioning problem, where similar items are connected. This type of clustering typically leads to cluster rings, which are e.g. used for finding bio-markers for health applications. This is different than the common k-means clustering algorithm, which will find compact data clusters.

- Compactness, e.g., k-means, mixture models
- Connectivity, e.g., spectral clustering



Compactness



Connectivity

Security

When the involved data is sensitive, like for most health applications, one would like to find secure solutions, e.g. by encrypting private inputs before performing the clustering algorithm. Since the involved clustering algorithm, which involves computing eigenvalues of matrices, is not easily translated to the encrypted domain, it is important to learn how fast such a 'secure spectral clustering' solution would perform. We are looking for a computer science or mathematics student, who is able and willing to implement our foreseen solution. This could be done during a (paid) TNO internship. For more information contact Thijs Veugen (email: thijs.veugen@tno.nl).