UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Unitat de Valorització

Multimodal spectral clustering technology

Clustering techniques allow to find natural groups inside very large datasets whose internal structure may be forma Partners to further develop the system and/or to establish commercial agreements along with technical cooperation are sought.

The Challenge

Clustering techniques allow to find natural groups inside very large datasets whose internal structure may be formerly unknown, specially when their size makes a manual analysis impractical.
Most clustering techniques have an important limitation, as they can only deal with one form (known as information mode) of the information being analyzed. Our multimodal clustering technique overcomes this limitation, as it can deal with any number of information modes in a single clustering analysis, extracting common patterns from them.

The Technology

It is a natural extension of the spectral clustering algorithm, a state of the art clustering algorithm that allows to cluster similar data not only by distance, but also by contiguity. This is especially important when analyzing networks of any kind (social, web pages, genes, economic or financial flows, etc.). For example, in an economic study, say company A is related with company B, and B with C. Even if there is no direct connection between A and C, it is important to capture the indirect link between them.

Innovative advantages

The algorithm exhibits the following desirable properties:

- Scalable: it can be applied to any number of data modes.
- Robust: it does not require any adaptation to be applied to new problems.
- Mathematically elegant: directly extends spectral clustering algorithm.

Current stage of development

The algorithm and software workflow is developed.

Applications and Target Market

Applications to all kinds of real-life problems: Bioinformatics, Image processing, Social networks analysis, Econometrics, Machine learning

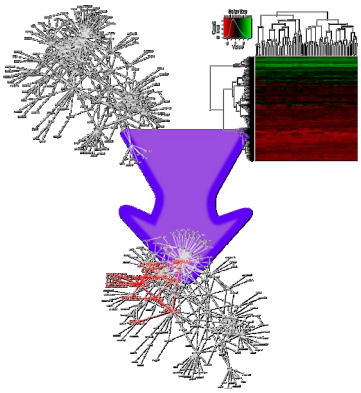
Example of application: social networks: Multimedia data, nowadays we face with everywhere, are a good example of multimodal tasks. The information sources for data analysis come from speech, images, tags, social networks, geo-locations etc. For example, we can think of problem on social event detection by looking at photos on http://www.flickr.com with metadata. The challenge is to find all soccer events taking place in Hamburg (Germany) and Madrid (Spain) in the collection (two clusters). Our flexible approach of multimodal spectral clustering llows to aggregate the information from all the sources (photo image, date and location, tags) to perform the task.

Example of application: bioinformatics: Gene to disease relationship is straightforward in exceptional cases. In most cases this relationship is not evident, and therefore it is useful to include as much information as possible in the studies.
In this example two data modes are used. First, a gene interaction network that shows the interactions between genes. Second, a gene expression matrix, that reflects the changes in gene expression inside cells that undergo different experimental conditions. Our technique allows to use both data modes in a single clustering analysis in order to obtain the most useful information from both sources and obtain a set of genes that are both related in terms of genetic interaction and changes in gene expression.

Reference number

MKT2012/0136_I

Application in Bioinformatics



Application in Social Networks

Useful for Image Processing

Business Opportunity
Technology available for licensing with technical cooperation

Patent Status
Priority application

Contact
Mr. Xavier Estaran Latorre
Licensing Manager
T. + 34 934 134 094
M. +34 626 260 596
f.xavier.estaran@upc.edu

See more technologies at
www.upc.edu/patents/TO
UPC—BarcelonaTech

ICT AND ELECTRONIC TECHNOLOGIES