Special Session on
Clustering and Co-clustering

Title: Clustering and Co-clustering
Duration: Half-day (to one day)

Description

Everyday, huge amounts of data are generated by users via the web, social networks, etc. Clustering/Coclustering techniques are a tool of choice to help organize the huge collections of data that increasingly beset us. Clustering/Coclustering is an unsupervised learning approach that allows one to discover global structures in the data (i.e. clusters). Given a dataset, it identifies different data subsets which are hopefully meaningful. The discovered clusters are deemed interesting if they are while instances within each (co)cluster share similar features. This (co)clustering problem has motivated a huge body of work and has resulted in a large number of algorithms. (Co)Clustering has thus been used in numerous real-life application domains such as marketing, city planning, and so forth.

Clustering algorithms are a tool of choice to explore these high-dimensional data sets. However numerous questions remain open as:

- What are the last advances in
    - supervised clustering that combines the main characteristics of both traditional clustering and supervised classification tasks?
    - quality criteria?
    - clustering for big data?
    - evolving clustering?
    - clustering events or time series?
    - …

This special session offers a meeting opportunity for academics and industry researchers belonging to the communities of Computational Intelligence, Machine Learning, Experimental Design, and Data Mining to discuss new areas of (co)clustering, One goal of this special session will be two-fold: On the one hand, to look for new algorithms and techniques proposals based on (co)clustering. On the other hand, to look for new application domains, real problems, where the application of (co)clustering have demonstrated an outstanding performance or interpretation abilities against other traditional approaches.

Publication opportunities: Papers should be submitted to IJCNN. We encourage papers that describe new algorithm and applications of (co)clustering in real-world. In the industrial context, the main difficulties met and the original solutions developed, have to be described.

Paper acceptance and publication will be judged on the basis of their quality and relevance to the special session themes, clarity of presentation, originality and accuracy of results and proposed solutions.

The set of proposed topics includes, but is not limited to::
- Clustering, Coclustering
- Supervised Clustering (Coclustering)
- Semi-Supervised Clustering (Coclustering)
- Quality Criteria for Clustering (Coclustering)
- Measure of Variable Importance for a Clustering (Coclustering)
- Automatic tuning of Cluster Number (Cocluster Number)
- Clustering for Big Data
- Method to assess the evolution of a Clustering (Coclustering)
- Constrainted Clustering (Coclustering)

The list of application domain is includes, but it is not limited to:
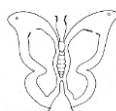
- Evolving textual information analysis
- Evolving social network analysis
- Dynamic process control and tracking
- Intrusion and anomaly detection
- Genomics and DNA micro-array data analysis
- Adaptive recommender and filtering systems
- …

A list of Applicative domains could be found in
P. Berkhin « Survey of clustering data mining techniques », Accrue Software, San Jose, CA, 2002.

Organizers:

Vincent Lemaire
ORANGE-LABS, Lannion, France
2 avenue Pierre Marzin
22300 Lannion
vincent.lemaire@orange-ftgroup.com
http://www.vincentlemaire-labs.fr

Pascal Cuxac
INIST-CNRS,
Recherche Développement
2 allée du Parc de Brabois
CS 10310
54519 Vandoeuvre les Nancy Cedex
pascal.cuxac@inist.fr,
https://sites.google.com/site/pascalcuxac/

Jean-Charles Lamirel
TALARIS (ex-TALARIS) -LORIA,
Campus Scientifique
BP 239
54506 Vandoeuvre-lès-Nancy
jean-charles.lamirel@loria.fr

Important dates:
- Paper submission: February 5, 2015
- Paper decision notification: March 25, 2015
- Camera-ready submission: April 25, 2015

Organizers:

Vincent Lemaire (Orange Labs, France, vincent.lemaire@orange.com) was born in 1968 and he obtained his undergraduate degree from the University of Paris 12 in signal processing and was in the same period an Electronic Teacher. He obtained a PhD in Computer Science from the University of Paris 6 in 1999. He thereafter joined the R&D Division of France Telecom where he became a senior expert in data-mining. His research interests are the application of machine learning in various areas for telecommunication companies with an actual main application in data mining for business intelligence. He developed exploratory data analysis and classification interpretation tools. Active learning and data-space exploration are now his main research interests. He obtained his Research Accreditation (HDR) in Computer Science from the University of Paris-Sud 11 (Orsay) in 2008.

Previous workshop or special session organizations:
- Atelier - Clustering and Co-clustering (CluCo) at EGC 2015 [...]
- Incremental classification, concept drift and novelty detection (IClaNov) - ICDM 2014 [...]
- Workshop of the Discovery Challenge - ECML 2014 [...]
- Incremental learning and novelty detection methods and their applications - ESANN 2014 [...]
- Atelier - Clustering and Co-clustering (CluCo) at EGC 2014 [...]
- Incremental clustering, concept drift and novelty detection (IClaNov) - ICDM 2013 [...]
- Active Learning and Experimental Design (ALED) - IJCNN 2013 [...]
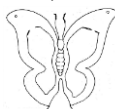
- Incremental classification and novelty detection (CIDN) - EGC 2013 [...]
- Active Learning in Real-World Applications - (ALRA) - ECML 2012 [...]
- Active and Incremental Learning (AIL) - ECAI 2012 [...]
- Incremental classification and novelty detection (CIDN) - EGC 2012 [...]
- Active, Incremental and Autonomous Learning: Algorithms and Applications - IJCNN 2012 [...]
- Workshop on Unsupervised and Transfer Learning (UTL) - ICML 2011 [...]
- Autonomous and Incremental Learning (AIL) - IJCNN 2011 [...]
- Active and Autonomous Learning (AAL) - IJCNN 2010 [...]
- Fast scoring on a Large Databse - KDD 2009 [...]

Jean-Charles Lamirel (LORIA - INRIA, France, lamirel@loria.fr) is Lecturer since 1997. He obtained his PhD in Computer Science in 1995 and his Research Accreditation in the same domain in 2010. He is currently teaching Information Science and Computer Science at the University of Strasbourg and achieving his research at the INRIA laboratory of Nancy. He was a research member of the INRIA-CORTEX project whose scope is Neural Networks and Biological Systems. He has recently integrated the INRIA-TALARIS project whose main concern is automatic language and text processing. Jean-Charles Lamirel main domains of research are textual data mining based on neural networks, multiple viewpoints data analysis paradigms, data mining auto-evaluation methods and evolving data mining. He is board member of the international Webometrics journal: "Collnet Journal of Scientometrics and Information Management" and was taking part in the committee of ICTAI 2011-2012 and WSOM 2012 conference. He is member of the IEE task force on "Evaluation and quality issues in data mining" within the Data Mining Technical Committee and committee member of the corresponding PAKDD-QIMIE 2013 workshop.

Previous workshop or special session organizations:
- Clustering and co-clustering – CluCo - workshop EGC 2014
- Incremental clustering, concept drift and novelty detection (IClaNov), Workshop ICDM 2013
- Incremental classification and novelty detection - CIDN - workshop EGC 2013
- Intelligent analysis of time varying information and concept drift management – IEA/AIE 2012
- Incremental classification and novelty detection - CIDN 2012
- Incremental clustering and novelty detection techniques and their application to intelligent analysis of time varying information - IEA/AIE 2011
- Incremental clustering and novelty detection - CIDN 2011
- International Conference on Webometrics, Informetrics and Scientometrics & 7th COLLNET Meeting in conjunction with the Extra Session on Information Visualization for Webometrics, Informetrics and Scientometrics, Nancy, France, 10-12 May, 2006

Pascal Cuxac (INIST - CNRS, France pascal.cuxac@inist.fr) is Research Engineer at the INIST/CNRS (Institute for Scientific & Technical Information / National Center for Scientific Research) in Nancy, France. He obtained his PhD in Geological and Mining Engineering from the Nancy School of Geology in 1991 and he was working on mechanical behavior of anisotropic rock. In 1993, he joined the CNRS as Research Engineer. Currently, in INIST Research & Development Engineering Service, he takes part in a research program on classification methods for bibliographic corpora, in particular in the development of an incremental unsupervised clustering algorithm.

Previous workshop or special session organizations:
- Clustering and co-clustering – CluCo - workshop EGC 2014
- Incremental clustering, concept drift and novelty detection (IClaNov), Workshop ICDM 2013
- Incremental classification and novelty detection - CIDN - workshop EGC 2013
- Incremental classification and novelty detection - CIDN 2012
- Incremental clustering and novelty detection techniques and their application to intelligent analysis of time varying information - IEA/AIE 2011
- Incremental clustering and novelty detection - CIDN 2011