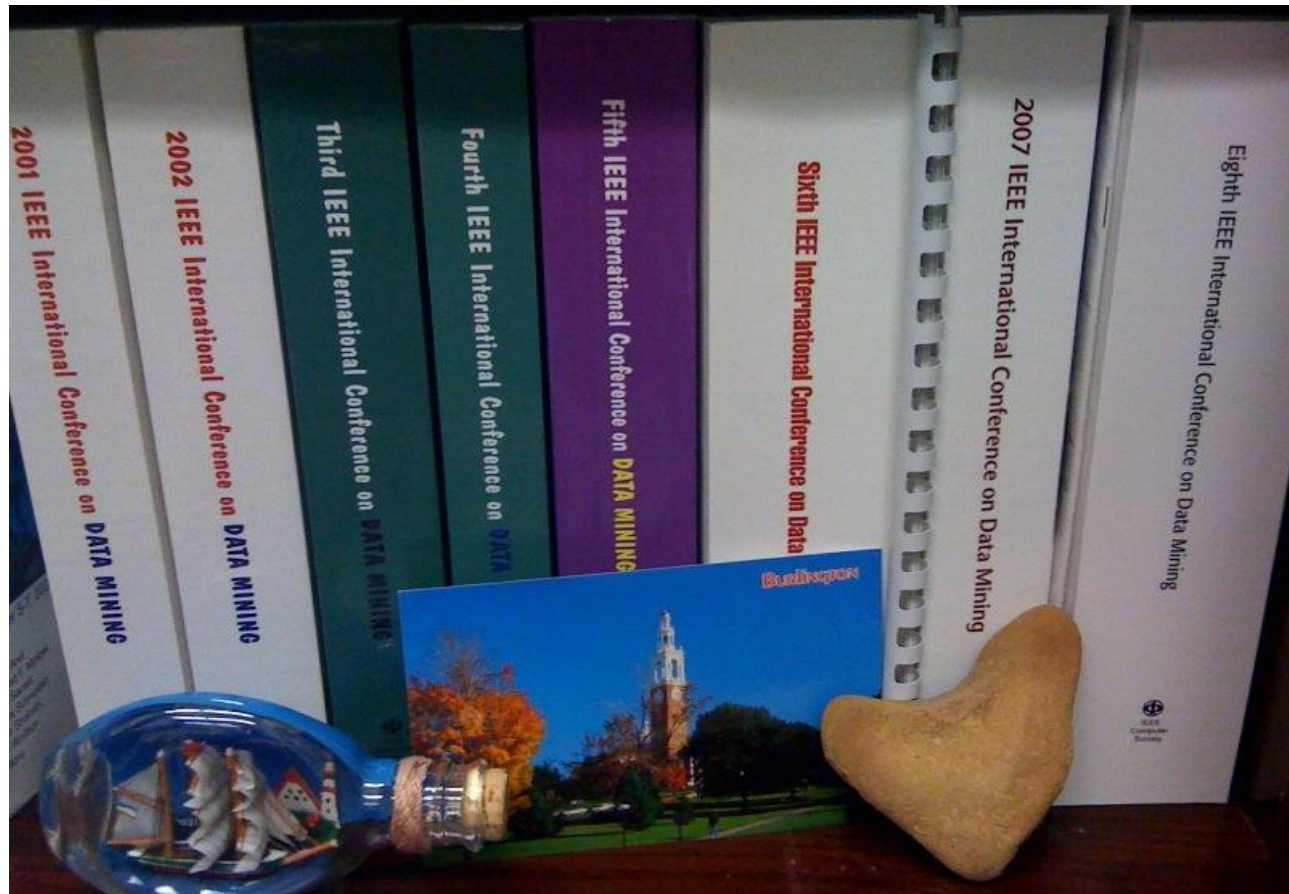




Welcome to 2013 ICDM Community Meeting



ICDM 2013

IEEE International Conference on Data Mining

Dallas, Texas, USA



ICDM 2013: PC Chair Review Report

Professor George Karypis

University of Minnesota – Twin Cities

Professor Hui Xiong

Rutgers, the State University of New Jersey



Paper Statistics

- Submitted: 809
 - 55 different countries
 - 475 (58.7%) student first authors
- Accepted: 159 (19.65%)
 - 94 regular papers (20 minute presentation)
 - 65 short papers (10 minute presentation)
- 22 Technical Sessions



Review Process

- 33 Vice Chairs and 234 PC members
- Triple blind reviewing process
- **Phase 0: Violation of Submission Requirements (76 papers)**
- **Phase 1:** 3 reviewers assigned to each paper (based on bids)
- **Phase 2:** Each paper assigned to a vice chair who guided the review discussion and made the commendation.
- **Phase 3:** Program Chairs made the final decision based on all input (seeking additional input when necessary)



- **Common Issues in the review comments**
 - 1) Picking the wrong fights
 - 2) No contribution or unclear contributions
 - 3) No evidence of research challenges
 - 4) Poor organization (no logic or wrong logic)
 - 5) Weak validation or no validation
 - 6) Unaware of relevant literature
 - 7) Showing ignorance in the paper
 - 8) Poor writing skills
 - 9) Do not show professionalism in the paper



1. Picking the wrong fights

- **A problem has been well studied and there is no new perspective.**
 - ❖ **Difficult to produce novel findings**
- **A problem has no application value.**
 - ❖ **Starting from business pains**
- **A problem has too many competitors.**
 - ❖ **High risk for redundant findings**
- **A problem is difficult to be validated.**
 - ❖ **Lack of data, domain knowledge, benchmark**



2. No contribution or unclear contributions

- **A combined use of existing techniques and no novel application.**
 - ❖ **Difficult to justify the contributions**
- **No evidence to show the proposed solution is better than existing solutions**
- **The contributions are there, but not clearly described.**
 - ❖ **Emphasize on wrong subjects or poor writing**
- **Over claimed contributions**



3. No evidence of research challenges

- **Research challenges should be clearly highlighted and repeated in the paper**
 - ❖ **A good way is to use examples**
- **The challenges are not properly highlighted**
 - ❖ **Simply describing what you have done, but not showing how difficult it could be done**
- **The problem has no research challenge.**
 - ❖ **Is there an application challenge?**



4. Poor organization

- **One appealing and consistent story**
 - ❖ **Don't put many topics**
- **No logic or poor logic**
 - ❖ **Be clear with every step**
- **Abstract should tell everything in a concise and exciting way**
 - ❖ **The first impression is very important**
- **Introduce should tell the story in a detailed and challenging way.**



5. Weak Validation or No Validation

- **There are four validation approaches**
 - ❖ **Pure Theoretical validation**
 - ❖ **Hybrid Validation:**
 - Partial Theoretical and Partial Empirical Validation**
 - ❖ **Pure Empirical Validation**
 - ❖ **Case Studies**
- **The work without validation is not complete**



6. Relevant Literature

- Find good sources and the most relevant papers. Also, predict who will be the reviewers?
- Make note of page numbers, URLs, and quotable passages so that you can properly cite your sources.
- Refer the related work in a positive way. Always acknowledge the merits of these work. However, do point out the difference!



7. Things to show your ignorance

- **A formal definition of common sense**
- **A poor description of the literature**
- **Papers do not properly cited**
- **Wrong use of validation metrics**
- **Wrong use of validation methods**
- **Wrong use of experimental data**
- **Treat some common sense as new findings**
- **Write the paper in an unprofessional way**
- **Simply copy math equations from the textbook**



8. Poor writing skills

- **Do not use complex and long sentences.**
 - ❖ **The purpose is to convey the ideas**
- **Always use examples to illustrate the methods**
 - ❖ **Be clear with every step and show challenges**
- **A paper should be lively and active.**
 - ❖ **Develop a balanced portfolio of all the elements, such as figures, tables, examples, proofs**
- **Avoid widow and orphan sentences**
- **Avoid grammar errors and typos**



9. Show professionalism

- **Clearly show the problem motivation**
- **Clearly point out the research gap**
- **Clearly state the research challenges**
- **All the required items should be in the paper**
- **Use Latex and professional tools to draw figures**
- **Put a really high standard on the paper presentation**
- **Anything in the paper should be demonstrated in a professional way**



Thank You !



ICDM '14 Call for Papers

- In ICDM '13 registration bag
- Important Dates:
 - Workshop proposals: April 4
 - Conference submissions: June 24
 - Acceptance notifications: September 24
 - Conference dates: December 14-17

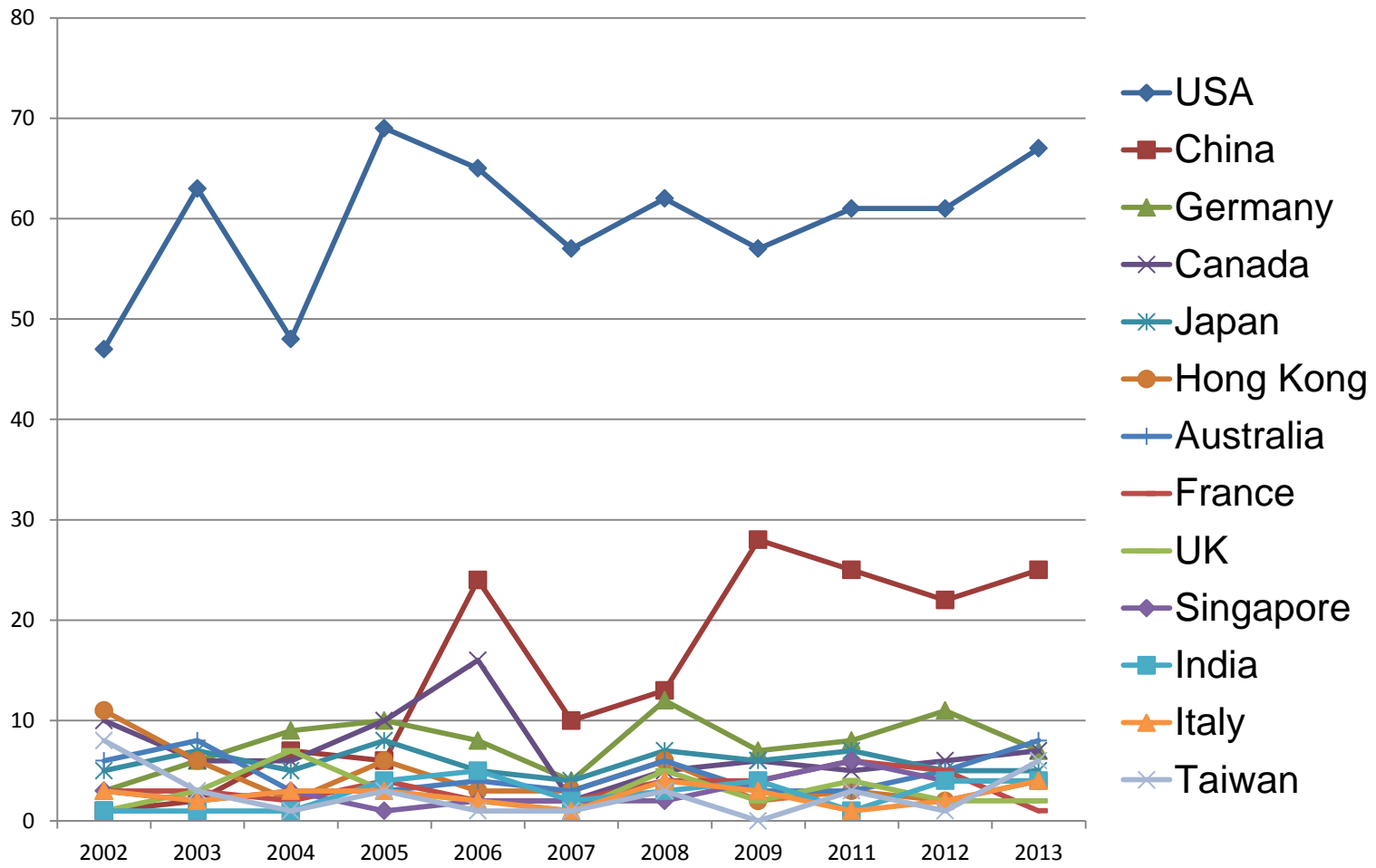


Data Mining on ICDM Submission Data (2013)

Shusaku Tsumoto, Kazuo Misue,
Ning Zhong and Xindong Wu



Trend of Country (2002 to 2013: # of Accepted Papers)





Comparison between 08-12 Countries (Ratio, #>10)

Country	Ratio (2010)	Country	Ratio (2011)	Country	Ratio (2012)	Country	Ratio (2013)
Finland	50%	Singapore	43%	Germany	30%	Taiwan	43%
USA	26%	Korea	40%	Korea	28%	Italy	29%
Italy	25%	Japan	32%	USA	28%	USA	26%
China	25%	Hong_Kong	25%	Singapore	25%	Australia	25%
Germany	22%	Taiwan	25%	Japan	22%	Hong_Kong	25%



Comparison between 09 and 13 Topics (Ratio)

Top5 in 2009	Ratio	Top5 in 2010	Ratio	Top5 in 2011	Ratio	Top5 in 2012	Ratio	Top5 in 2013	Ratio
Statistics and Probability	33%	Statistics and Probability	60%	Pre-processing	26%	Foundation	27%	Visualization	30%
Foundations	25%	Quality Assesment	29%	Security & Privacy	25%	High Performance	24%	Text & Sequence	24%
DM & ML	24%	Pre-processing	26%	Statistics & Probability	25%	Security & Privacy	24%	DM & ML	21%
Text,...	21%	DM & ML	25%	Text & Sequence	23%	DM & ML	23%	Foundation	21%
High Performance	21%	Others	18%	DM & ML	22%	Collaborative Filtering	21%	Stream	21%



Terms in Accepted Abstracts (Top 10)

Emerging

- uncertain graphs
- pairwise similarities
- dynamic patterns
- multi-label learning
- sparse graphs
- conformal prediction framework
- multiple clusters
- information diffusion
- information network

Increasing

- social
- Influence
- graph(s)
- matrix
- learning
- network(s)
- users
- topic
- prediction
- nodes

Decreasing

- frequent patterns
- mining
- techniques
- query
- knowledge
- complex relationships
- prior methods
- patterns
- clustering
- data set

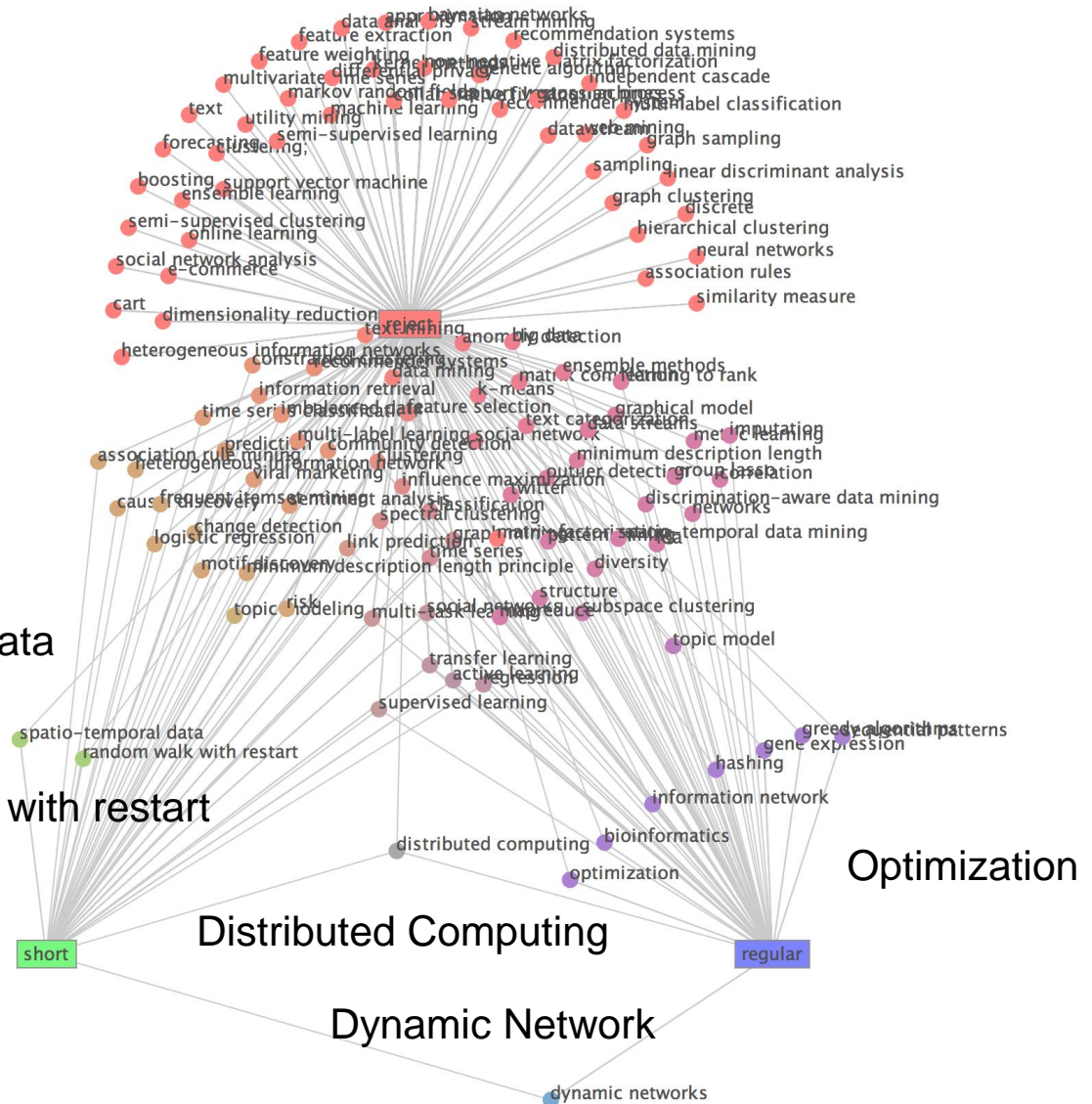


Anchored Map (Keyword vs Decision, 2013)

Spatio-temporal data

Random work with restart

Kazuo Misue,
Anchored Map:
Graph Drawing Technique
to Support Network Mining,
IEICE Trans. Inf. & Syst.,
Vol. E91-D, No. 11,
pp. 2599-2609, 2008.

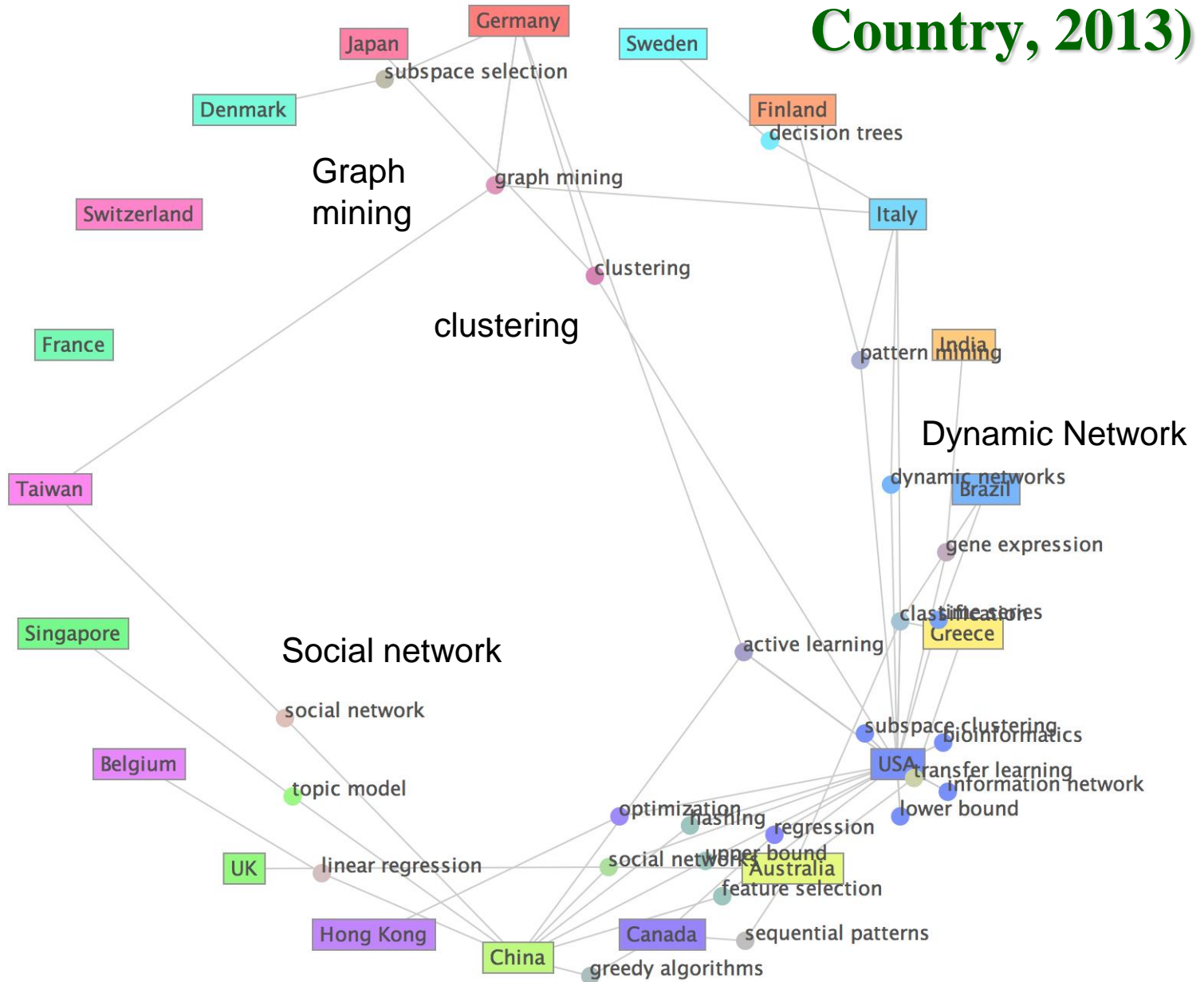


Optimization

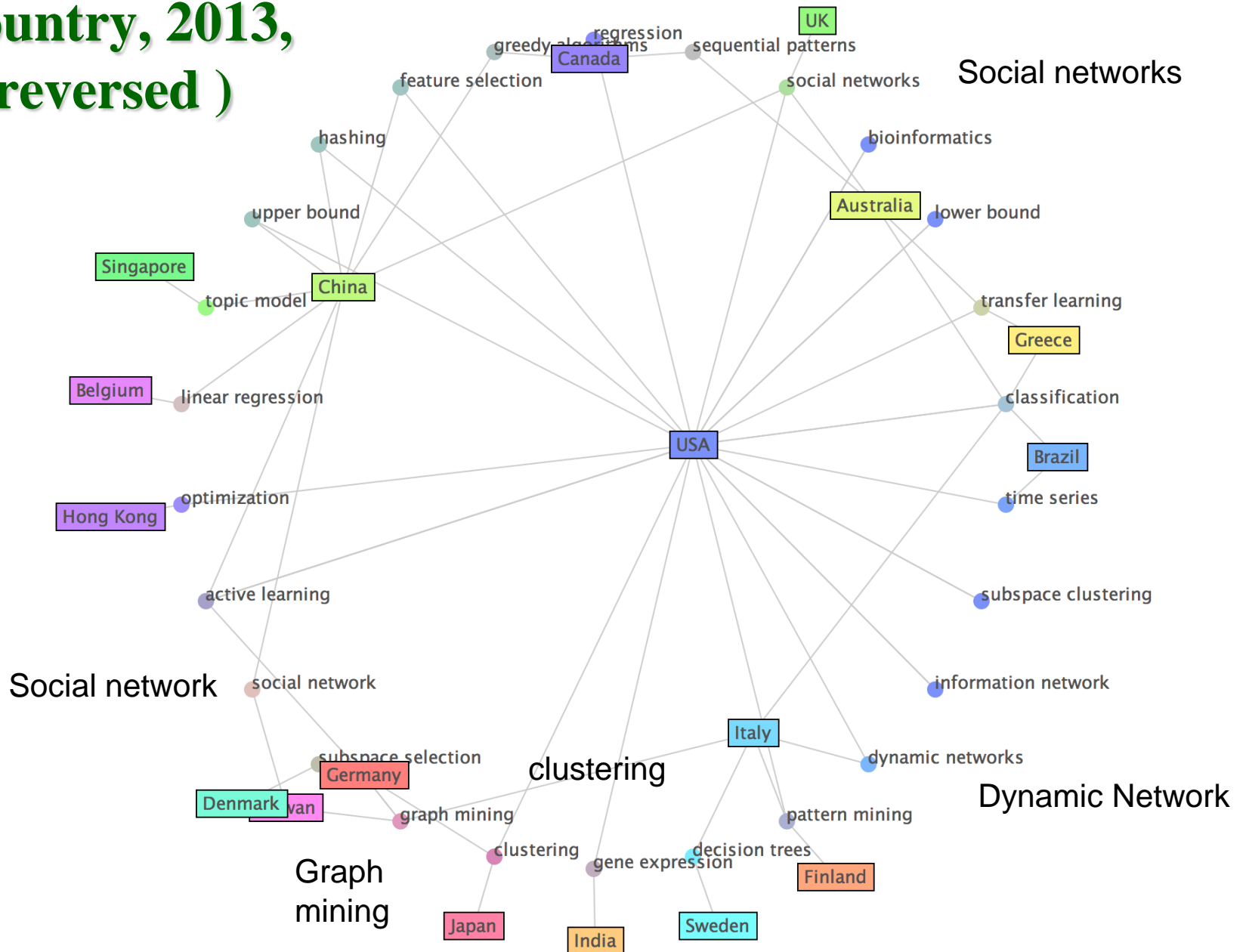
Distributed Computing

Dynamic Network

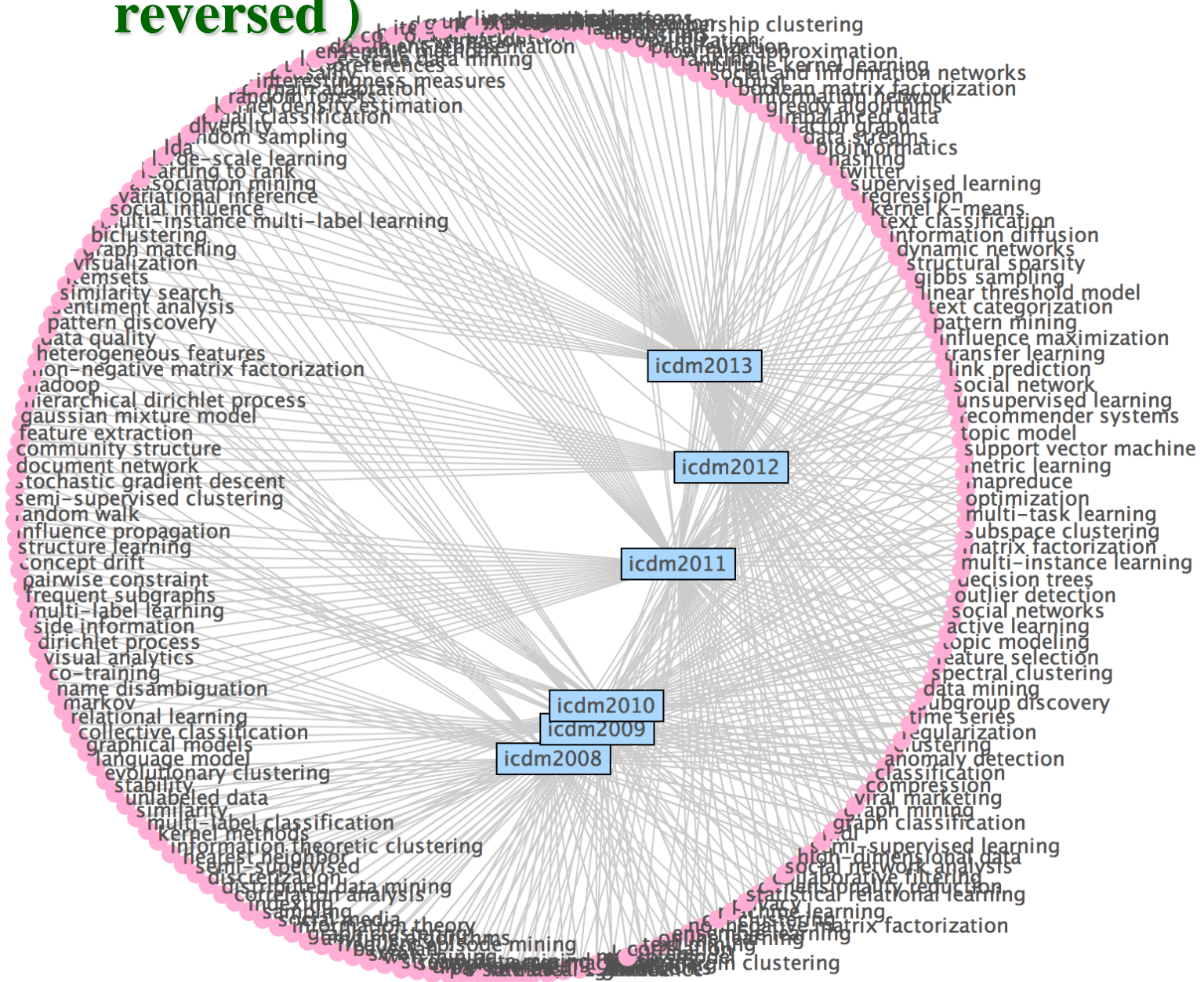
(Keyword vs Country, 2013)



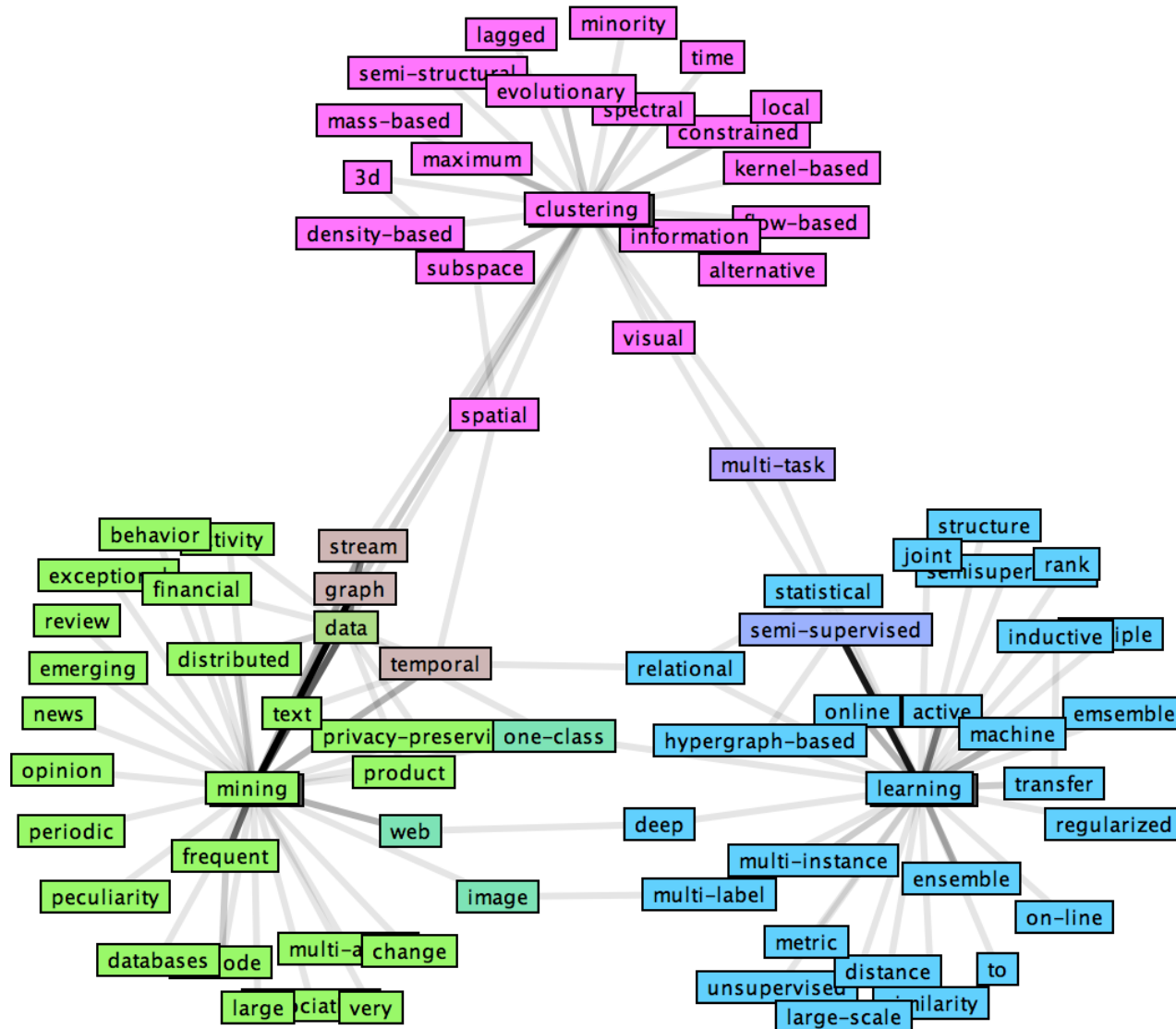
(Keyword vs Country, 2013, reversed)



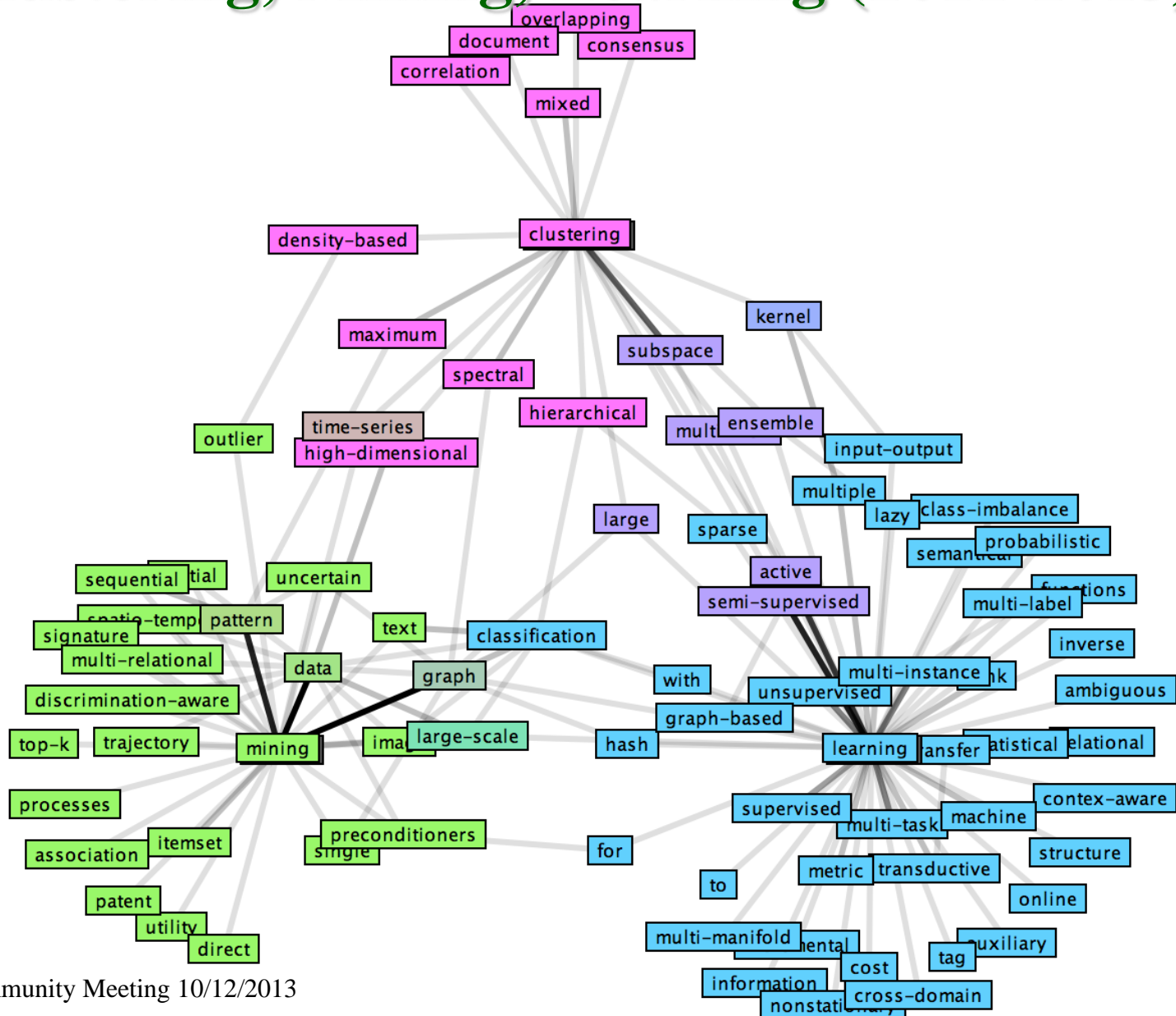
(Keyword vs Year: 2008-2013, reversed)



Adjunct words for Clustering, Mining, Learning (2008-2010)



Adjunct words for Clustering, Mining, Learning (2011-2013)





Data Mining on ICDM Submission Data

- Acknowledgements
 - Many thanks to
 - PC chairs, Vice Chairs and PC members
 - All the authors
 - All the contributors to ICDM2013
 - See you soon in Shenzhen!