Science mapping and its applications in multidisciplinartiy analysis

Zhesi Shen

Center of Scientometrics, National Science Library, CAS

shenzhs@mail.las.ac.cn

Nov. 2, 2018

Background
Project-driven

Science Map and
Interdisciplinary

Zhesi Shen

2 Background

Methods

Results
Science Mapping
Diversity

Conclusions and
Discussion

▶ Refining CAS Journal Classification system
according to Chinese Ministry of Education, the
Academy Degrees Committee and NSFC
=> Journal Relatedness

Background
Project-driven

- Refining CAS Journal Classification system according to Chinese Ministry of Education, the Academy Degrees Committee and NSFC
  => Journal Relatedness

- Project-oriented journal classification/selection/recommendation/semi-supervised
  => Journal features

Background
Research-driven

Science Map and
Interdisciplinary

Zhesi Shen

3 Background

Methods

Results
Science Mapping
Diversity

Conclusions and
Discussion

Measuring multi/inter-disciplinarity

Diversity measure

- Variety: how many elements involved
- Balance: how even are they distributed
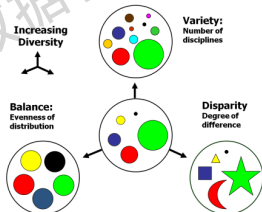- *Disparity*: how different are they
  => Journal Relatedness

Fig 1. Schematic representation of the attributes of diversity, based on Stirling [53].

Background

Science Map and
Interdisciplinary

Zhesi Shen

4 Background

Methods

Results
Science Mapping
Diversity

Conclusions and
Discussion

Relatedness + Features

- local view: e.g., direct citation
- neighborhood view: e.g., co-citations
- global view: e.g., science mapping, network distance

Methods

1. JCR2017 Journal citation graph

Methods

Science Map and
Interdisciplinary

Zhesi Shen

Background

5 Methods

Results

Science Mapping
Diversity

Conclusions and
Discussion

1. JCR2017 Journal citation graph
2. Graph embedding: Node2Vec to learn high
   dimension features

Methods

1. JCR2017 Journal citation graph
2. Graph embedding: Node2Vec to learn high dimension features
3. relatedness/clustering/classification

Methods

Science Map and
Interdisciplinary

Zhesi Shen

Background

5 Methods

Results
Science Mapping
Diversity

Conclusions and
Discussion

1. JCR2017 Journal citation graph
2. Graph embedding: Node2Vec to learn high dimension features
3. relatedness/clustering/classification
4. visualization: dimension reduction, e.g., t-sne

Methods
Node2Vec

Science Map and
Interdisciplinary

Zhesi Shen

Background

6 Methods
Science Mapping
Diversity

Conclusions and
Discussion

1. generate random walks:
   considering
   multiple-order
   neighbors

2. skip-gram model : to
   learn features$\vec{x}$
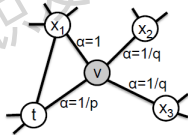
Figure 2: Illustration of the random walk procedure in *node2vec*.
The walk just transitioned from t to v and is now evaluating its next
step out of node v. Edge labels indicate search biases α.

Results
Science Mapping

Science Map and Interdisciplinary

Zhesi Shen

Background

Methods

Results
7   Science Mapping
   Diversity

Conclusions and Discussion
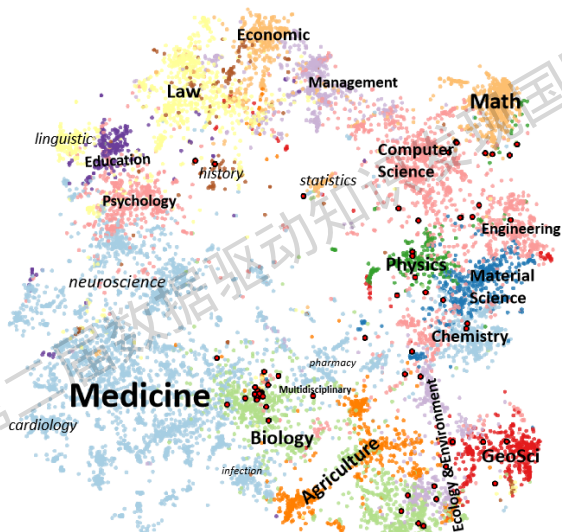
18   CoS

Results
Science Mapping: Journals of Multidisciplinary Sciences
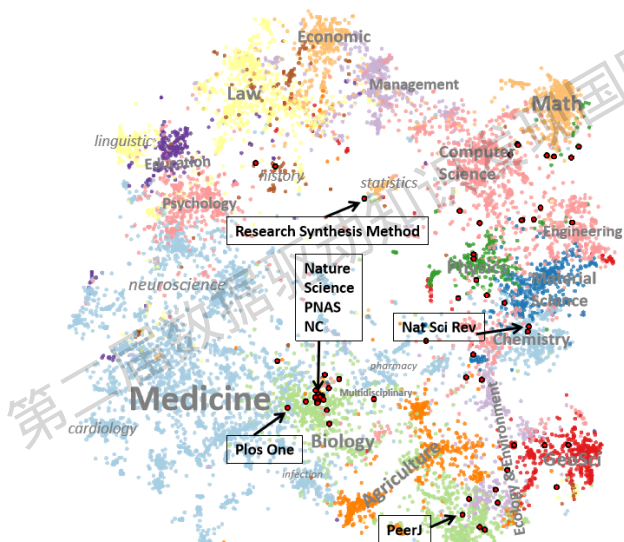
Results
Vector Norm

Science Map and
Interdisciplinary

Zhesi Shen

Background

Methods

Results
9 Science Mapping
Diversity

Conclusions and
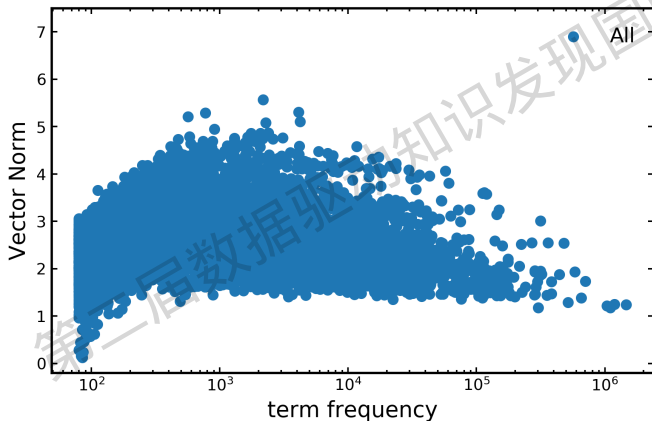Discussion

what the vector can tell
we first define

- term frequency: the number of appearance in the random-walk list <==> network centrality

- vector norm: $|x|_2$

Results
Vector Norm

increase − decrease $p \sim \vec{x} \cdot \vec{y} = |x||y|cos(\theta)$

Science Map and
Interdisciplinary

Zhesi Shen

Background
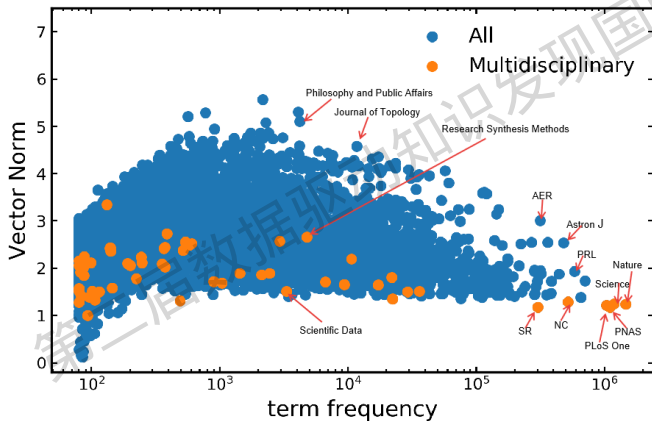
Methods

Results
10 **Science Mapping**
    Diversity

Conclusions and
Discussion

Results
Vector Norm

Vector norm encodes multidisciplinarity.

Science Map and Interdisciplinary

Zhesi Shen

Background

Methods

Results
11 Science Mapping
Diversity

Conclusions and Discussion

Results
Citing/Cited Diversity

Rao-stirling:

$$D = \sum_{i \neq j} p_i p_j (1 - s_{ij}) \qquad (1)$$

- Global similarity: $s_{ij}^G = cos(x_i^G, x_j^G)$
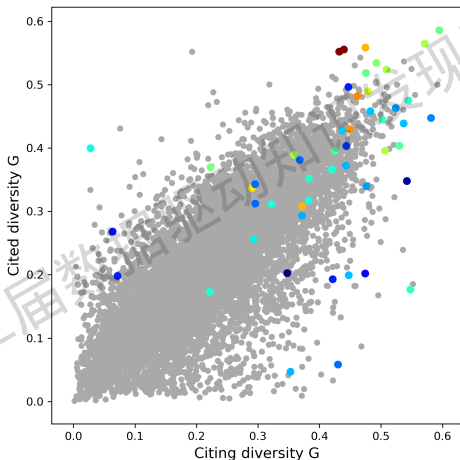- Neighborhood similarity: $s_{ij}^{NB} = cos(c_i^{NB}, c_j^{NB})$

Results
Citing/Cited Diversity

Using global similarity

Science Map and
Interdisciplinary

Zhesi Shen

Background

Methods

Results
Science Mapping
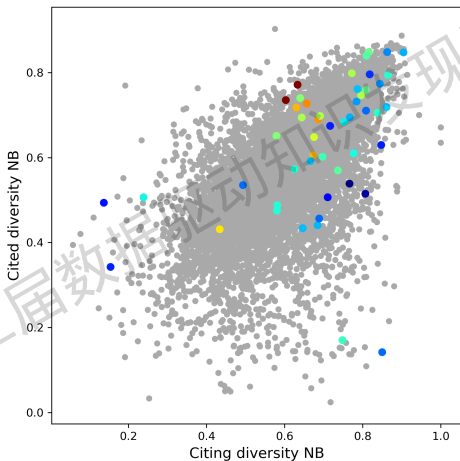13 Diversity

Conclusions and
Discussion

18 CoS

Results
Citing/Cited Diversity

Using neighborhood similarity

Science Map and
Interdisciplinary

Zhesi Shen

Background

Methods

Results
 Science Mapping
14 **Diversity**

Conclusions and
Discussion

18

Conclusions

- node2vec embedded features implying the intrinsic structure of disciplines
- embedded vector encodes both centrality and multidisciplinarity
- Journals in WOS-Multidisciplinary Sciences are highly heterogeneous
- similarity matters in diversity measure
- global similarity seems to be better

CoS

Discussion

- limited
- paper-level diversity measure
- paper-level classification system
- the equivalence between modularity-based community detection and non-negative matrix factorization
- infomap and random-walk based graph embedding

Acknowledgements

Collaborators

- Fuyou Chen, Center of Scientometrics, NSL, CAS
- Liying Yang, Center of Scientometrics, NSL, CAS
- Jinshan Wu, Beijing Normal University

Thank you for your attention!