Socio-semantic Networks of Research Publications in the Learning Analytics Community

Soude Fazeli, PhD candidate
Dr. Hendrik Drachsler
Prof. Dr. Peter Sloep
Agenda

1. Introduction
2. Motivation
3. Data processing
4. Network visualization
5. Discussion and conclusion
1. Introduction

- Presenting visualization of the authors and papers network
- Carrying out a deeper analysis of the generated networks
- The main aim
  - To use such a graph of authors and papers to recommend similar items to a target user
2. Motivation

• A list of recommended authors and papers
  • To plan the conference participation more efficiently and effectively

• Awareness support for researchers (Reinhardt et al., 2012; Fisichella et al., 2010; Ochoa et al., 2009; Henry et al., 2009)

• Scientific recommender systems (Huang et al., 2002; Wang & Blei, 2010)

• Such a priority list
  • Support the awareness of the attendees
  • Empower the network of like-minded authors in the attendees’ particular research focus
• **RQ1.** How are the authors connected and which authors share more connections and are more central in terms of sharing commonalities with the others?

• **RQ2.** How are the papers connected to each other in terms of similarity?
1. Finding patterns of similarity between authors and papers
2. Visualizing networks of the LAK authors and papers
3. Data processing

3.1. Tag clouds

• The TF-IDF algorithm
  • Weighted list of the most commonly used terms in research articles
  • Using default algorithm provided by Mahout on the text files extracted from the RDF files
  • Removing the stop words
    • Setting the configuration variables within Mahout to 90%

• Outcome
  • A so-called dictionary of all the terms in the LAK dataset
  • A binary sequence file including the TF-IDF weighted vectors
3. Data processing

3.2. Computing similarity

• Using T-index algorithm (Fazeli et al., 2010)
  • Collaborative filtering recommender algorithm
  • Generates a graph of users
    • The nodes are users and the edges show the relationship between users
    • Originally makes recommendations based on the ratings data of users

• Extending the T-index algorithm
  • Process tags and keywords extracted from the linked data

• Using Jena APIs to
  • Process RDF files
  • Handle Ontology Web Language (OWL) files describing the generated graph of authors and papers
4. Network visualization

4.1. Author network (made by Welkin)
4. Network visualization

4.1. Author network

The degree centrality of the top ten central authors

Then first ten central authors

The degree centrality of the top ten central authors
4. Network visualization

4.2. Paper network
4. Network visualization

4.2. Paper network

The degree centrality of Top ten papers
• RQ1. How are the authors connected and which authors share more connections and are more central in terms of sharing commonalities with the others?

• RQ2. How are the papers connected to each other in terms of similarity?
5. Discussion and conclusion

5.1. RQ1

<table>
<thead>
<tr>
<th>Author</th>
<th>Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hendrik Drachsler</td>
<td>116</td>
</tr>
<tr>
<td>Kon Shing Kenneth Chung</td>
<td>87</td>
</tr>
<tr>
<td>Wolfgang Greller</td>
<td>80</td>
</tr>
<tr>
<td>Javier Melenchon</td>
<td>66</td>
</tr>
<tr>
<td>Brandon White</td>
<td>59</td>
</tr>
<tr>
<td>Vania Dimitrova</td>
<td>50</td>
</tr>
<tr>
<td>Erik Duval</td>
<td>45</td>
</tr>
<tr>
<td>Rebecca Ferguson</td>
<td>44</td>
</tr>
<tr>
<td>Anna Lea Dyckhoff</td>
<td>40</td>
</tr>
<tr>
<td>Simon Buckingham Shum</td>
<td>39</td>
</tr>
</tbody>
</table>

The first ten central authors
• **RQ1.** How are the authors connected and which authors share more connections and are more central in terms of sharing commonalities with the others?

• **RQ2.** How are the papers connected to each other in terms of similarity?
## 5. Discussion and conclusion

### 5.2. RQ2

<table>
<thead>
<tr>
<th>Paper</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning Dispositions and Transferable Competencies: Pedagogy, Modelling and Learning Analytics</td>
<td>Simon Buckingham-Shum, Ruth Deakin Crick</td>
</tr>
<tr>
<td>The Pulse of Learning Analytics Understandings and Expectations from the Stakeholders</td>
<td>Hendrik Drachsler, Wolfgang Greller</td>
</tr>
<tr>
<td>Social Learning Analytics: Five Approaches</td>
<td>Rebecca Ferguson, Simon Buckingham-Shum</td>
</tr>
<tr>
<td>Multi-mediated Community Structure in a Socio-Technical Network</td>
<td>Dan Suthers, Kar Hai Chu</td>
</tr>
<tr>
<td>Modelling Learning &amp; Performance: A Social Networks Perspective</td>
<td>Walter Christian Paredes, Kon Shing Kenneth Chung</td>
</tr>
<tr>
<td>Teaching Analytics: A Clustering and Triangulation Study of Digital Library User Data</td>
<td>Beijie Xu, Mimi M Recker</td>
</tr>
<tr>
<td>Monitoring Student Progress Through Their Written &quot;Point of Originality&quot;</td>
<td>Johann Ari Larusson, Brandon White</td>
</tr>
<tr>
<td>Learning Designs and Learning Analytics</td>
<td>Lori Lockyer, Shane Dawson</td>
</tr>
<tr>
<td>A Multidimensional Analysis Tool for Visualizing Online Interactions</td>
<td>Eunchul Lee, M'hammed Abdous</td>
</tr>
<tr>
<td>Using computational methods to discover student science conceptions in interview data</td>
<td>Bruce Sherin</td>
</tr>
</tbody>
</table>
5. Discussion and conclusion

- The central authors and top ten papers will appear more often in the top recommendation list.
- Although most of the central authors also appear in top ten papers’ list, the order is not the same.
  - Some authors have more than one paper.
  - Not each and every one of the authors’ papers individually has the highest similarity to the other papers.
Soude Fazeli  
PhD candidate  
Open University of the Netherlands  
Centre for Learning Sciences and Technologies (CELSTEC) PO-Box 2960  
6401 DL Heerlen, The Netherlands  
external email: soude.fazeli@ou.nl