Entity-Level Event Impact Analytics
Govind Govind

To cite this version:
Govind Govind. Entity-Level Event Impact Analytics. WSTNET Web Science Summer School 2018 (WWSSS 2018), Jul 2018, Hannover, Germany. hal-02102815

HAL Id: hal-02102815
https://hal.archives-ouvertes.fr/hal-02102815
Submitted on 17 Apr 2019
Objective: Automatically Predict the Event Diffusion into Foreign Language Communities

**Conceptual Approach**

“Discovery of semantic connections to languages from named entities in the article”

**ELEVATE Framework**

- Recursive exploration of `yago` relations
- **Country centric**: `<isCitizenOf>`, `<diedIn>`, `<isLocatedIn>`, `<isLeaderOf>`, `<isPoliticianOf>`, `<wasBornIn>`, `<livesIn>
- **Organization centric**: `<owns>`, `<created>`, `<worksAt>`
- Linking of entities to languages

**ELEVATE Pipeline**

1. **Event Data Collection**
2. **Named Entity Disambiguation**
3. **Entity-level Analytics**
4. **Semantic Aggregation**
5. **Spread Prediction**

**Spread Prediction**

- **Task**: Pick the best candidates from all the scored languages
- **Adjusted Thresholding**
  - Threshold(θ) = average spread in the ground truth
  - k-fold cross-validation
  - Risk of picking the irrelevant languages
- **Multi-label Classification**
  - Output labels as the languages in event spread
  - Candidate language scores as feature vectors
  - Classifiers decide the spread

**Other Research Works**

- **ELEVATE-Live**
  - Fine-grained entity-level Web content classification
  - Concise semantic representation of documents based on their entities

**Publications**


**Experimental Results**

<table>
<thead>
<tr>
<th>Method</th>
<th>Precision</th>
<th>Recall</th>
<th>F1</th>
<th>4PC</th>
<th>5days</th>
<th>20days</th>
<th>40days</th>
<th>5days</th>
<th>20days</th>
<th>40days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random</td>
<td>0.00769</td>
<td>0.00769</td>
<td>0.00769</td>
<td>0.00769</td>
<td>0.00769</td>
<td>0.00769</td>
<td>0.00769</td>
<td>0.00769</td>
<td>0.00769</td>
<td>0.00769</td>
</tr>
<tr>
<td>ELEVATE</td>
<td>0.23473</td>
<td>0.23473</td>
<td>0.23473</td>
<td>0.23473</td>
<td>0.23473</td>
<td>0.23473</td>
<td>0.23473</td>
<td>0.23473</td>
<td>0.23473</td>
<td>0.23473</td>
</tr>
<tr>
<td>Infokit</td>
<td>0.20584</td>
<td>0.20584</td>
<td>0.20584</td>
<td>0.20584</td>
<td>0.20584</td>
<td>0.20584</td>
<td>0.20584</td>
<td>0.20584</td>
<td>0.20584</td>
<td>0.20584</td>
</tr>
<tr>
<td>OntoLink</td>
<td>0.33103</td>
<td>0.33103</td>
<td>0.33103</td>
<td>0.33103</td>
<td>0.33103</td>
<td>0.33103</td>
<td>0.33103</td>
<td>0.33103</td>
<td>0.33103</td>
<td>0.33103</td>
</tr>
<tr>
<td>BroaderLink</td>
<td>0.23054</td>
<td>0.23054</td>
<td>0.23054</td>
<td>0.23054</td>
<td>0.23054</td>
<td>0.23054</td>
<td>0.23054</td>
<td>0.23054</td>
<td>0.23054</td>
<td>0.23054</td>
</tr>
<tr>
<td>Entropy</td>
<td>0.49803</td>
<td>0.49803</td>
<td>0.49803</td>
<td>0.49803</td>
<td>0.49803</td>
<td>0.49803</td>
<td>0.49803</td>
<td>0.49803</td>
<td>0.49803</td>
<td>0.49803</td>
</tr>
<tr>
<td>Entropy2</td>
<td>0.63398</td>
<td>0.63398</td>
<td>0.63398</td>
<td>0.63398</td>
<td>0.63398</td>
<td>0.63398</td>
<td>0.63398</td>
<td>0.63398</td>
<td>0.63398</td>
<td>0.63398</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>Precision</th>
<th>Recall</th>
<th>F1</th>
<th>4PC</th>
<th>5days</th>
<th>20days</th>
<th>40days</th>
<th>5days</th>
<th>20days</th>
<th>40days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dataloc</td>
<td>0.13684</td>
<td>0.13684</td>
<td>0.13684</td>
<td>0.13684</td>
<td>0.13684</td>
<td>0.13684</td>
<td>0.13684</td>
<td>0.13684</td>
<td>0.13684</td>
<td>0.13684</td>
</tr>
<tr>
<td>BroaderLink</td>
<td>0.13684</td>
<td>0.13684</td>
<td>0.13684</td>
<td>0.13684</td>
<td>0.13684</td>
<td>0.13684</td>
<td>0.13684</td>
<td>0.13684</td>
<td>0.13684</td>
<td>0.13684</td>
</tr>
<tr>
<td>Entropy</td>
<td>0.49803</td>
<td>0.49803</td>
<td>0.49803</td>
<td>0.49803</td>
<td>0.49803</td>
<td>0.49803</td>
<td>0.49803</td>
<td>0.49803</td>
<td>0.49803</td>
<td>0.49803</td>
</tr>
<tr>
<td>Entropy2</td>
<td>0.63398</td>
<td>0.63398</td>
<td>0.63398</td>
<td>0.63398</td>
<td>0.63398</td>
<td>0.63398</td>
<td>0.63398</td>
<td>0.63398</td>
<td>0.63398</td>
<td>0.63398</td>
</tr>
</tbody>
</table>

**Short Biography**

3rd year PhD student University of Caen, France Advisor: Prof. Marc Spaniol

**Research Interests**

- Entity-level analytics
- Data aggregation via LOD
- Deep learning for noisy data analytics

**Education**

- Master degree in Maths and Computing from Indian Institute of Technology (IIT) Patna (India)
- Bachelor degree in computer science from Guru Jambeshwar University of Science and Technology Hisar (India)