A Walk-based Model on Entity Graphs for Relation Extraction

Fenia Christopoulou\textsuperscript{1,3}, Makoto Miwa\textsuperscript{2,3}, Sophia Ananiadou\textsuperscript{1,3}

\textsuperscript{1}School of Computer Science, The University of Manchester, UK
\textsuperscript{2}Toyo Technological Institute, Nagoya, Japan
\textsuperscript{3}Artificial Intelligence Research Center (AIRC), National Institute of Advance Industrial Science and Technology (AIST)

\textbf{Motivating Example:} Tofting and capital are related through the word "capital" and with the word "wages".

New Walk Generation

1. Pair & Context Representations
   - Target pair: named entity pair of interest
   - Target pair context: sentence words excluding pair
   - Relative positions: entity \(e_i\) word \(w_j\) to entity \(e_j\)
   - Target entities \(e_i, e_j\) representations: \(v_{c,i} = [e_i, t_i, p_{c,i}]\)
   - Pair \((e_i, e_j)\) context representation:
     \[ C_{c,i,j} = W_c [v_{c,i}, v_{c,j}, p_{c,i}, p_{c,j}] \]

2. Edge Representation
   - Initial representation: \(\theta = (q, \alpha, \omega)\)
   - Classification layer
   - Edge layer
   - Walk layer
   - Final representation: \(\theta_{final} = \theta \times \theta\)

\textbf{Settings}

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Training Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACE 2005</td>
<td>Adam Optimizer</td>
</tr>
<tr>
<td>7 entity types</td>
<td>L2 Regularization</td>
</tr>
<tr>
<td>6 relations types</td>
<td>Early Stopping</td>
</tr>
</tbody>
</table>

**SoA Models:**
- SPTree \[2\]
- CNN \[3\]

**One pair/sentence:**
- \(\checkmark\)
- \(\checkmark\)
- \(\checkmark\)
- \(\checkmark\)
- \(\checkmark\)
- \(\checkmark\)

**Data split:**
- Train/test 5-fold

**Results**

<table>
<thead>
<tr>
<th>Model</th>
<th>P</th>
<th>R</th>
<th>F1 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPTree</td>
<td>70.1</td>
<td>61.2</td>
<td>65.3</td>
</tr>
<tr>
<td>Baseline</td>
<td>72.5</td>
<td>53.3</td>
<td>61.4</td>
</tr>
<tr>
<td>No walks (l = 1)</td>
<td>71.9</td>
<td>53.6</td>
<td>62.7</td>
</tr>
<tr>
<td>Walks (l = 2)</td>
<td>69.9</td>
<td>58.4</td>
<td>63.6</td>
</tr>
<tr>
<td>Walks (l = 4)</td>
<td>69.7</td>
<td>59.5</td>
<td>64.2</td>
</tr>
<tr>
<td>Walks (l = 8)</td>
<td>71.5</td>
<td>53.3</td>
<td>62.4</td>
</tr>
</tbody>
</table>

**Table 1:** Performance on ACE 2005 test set.
- \(\checkmark\): Walks model \(l = 4\) approximates the state-of-the-art.
- \(\checkmark\): Longer walks improve recall.
- Too long walks degrade performance.

**References**


**Contact:** efenia.christopoulous@manchester.ac.uk