Chinese Whispers - An Efficient Graph Clustering Algorithm

Nisim Hurst

Tuesday 14 May 2019

Chinese Whispers - An Efficient Graph Clustering Algorithm

The article was written by (Biemann 2006). It was was cited 303 times according to Google Scholar. The task performed was graph-clustering. They used x metric over x.

Hypothesis

A randomized graph-clustering algorithm would perform well of small worlds.

Evidence and Results

Dataset

The dataset used are all from the natural language processing world.

Results

Contribution

The authors propose

Algorithm 1 Chinese Whispers clustering subject to pairwise constraints

Input: An initial graph G of V multidimensional vectors to be clustered, the must-link pairs symmetric relations $ML = \{(V_i, V_j)|i \neq j \cap C(V_i) = C(V_j)\}$ and $CL = \{(V_i, V_j)|i \neq j \cap C(V_i) \neq C(V_j)\}$

Output: An assignment of the clusters for each vertex $\in G.V$

1: **procedure** ConstrainedCW(G,ML,CL)
2: MustLink $\leftarrow$ Connected-Components($G.V$, ML)
3: CannotLink $\leftarrow$ Connected-Components($G.V$, CL)
4: CWWrapper(&G,$MustLink$,$CannotLink$)
5: **return** ClusterAssignments
6: **end procedure**

Small world graphs

Controversial Ideas

Weaknesses

Future Work

References

Biemann, Chris (June 2006). "Chinese Whispers - an Efficient Graph Clustering Algorithm and its Application to Natural Language Processing Problems". In: *Proceedings of TextGraphs: the First Workshop on Graph Based Methods for Natural Language Processing*. New York City: Association for Computational Linguistics, pp. 73–80. URL: https://www.aclweb.org/anthology/W06-3812.