DANCINGLINES: An Analytical Scheme to Depict Cross-Platform Event Popularity

Author: Jinning Li, Tianxiang Gao, et al. Tutor: Xiaofeng Gao

Background:
With the development of Internet technology, the primary media for information propagation have been shifting to online media like social networks, search engines, web portals. Popular events are usually disseminated on multiple media. Depicting and analyzing event popularity across different platforms plays a vital role in tracking the public concerns and understanding the event disseminations.

Object:
Quantify the event popularity time series (EPTS) based on the datasets of Baidu and Weibo. Align and analyze EPTSs of different media and visualize the result.

TF-SW:
TF-SW is a semantic-aware popularity quantification model based on an integrated weight coefficient that leverages Word2Vec and TextRank algorithm.

Discard the words unrelated to certain events

\[
\text{sem}(w_i, w_j) = \beta \cdot \frac{w_i^B \cdot w_j^B}{||w_i^B|| \cdot ||w_j^B||} + (1 - \beta) \cdot \frac{w_i^P \cdot w_j^P}{||w_i^P|| \cdot ||w_j^P||}
\]

\[
\text{str}(w_i, w_j) = \frac{\sqrt{\sum_{c_i \in w_i} \text{num}(c_i, w_i)^2 \cdot \sqrt{\sum_{c_j \in w_j} \text{num}(c_j, w_j)^2}}}{\sum_{c_i \in w_i} \text{num}(c_i, w_i) \cdot \sum_{c_j \in w_j} \text{num}(c_j, w_j)}
\]

\[
\text{sim}(w_i, w_j) = \gamma \cdot \text{sem}(w_i, w_j) + (1 - \gamma) \cdot \text{str}(w_i, w_j)
\]

Similarity utilizing semantic and lexical relations

\[
\text{TR}(w_i) = 1 - \frac{\theta}{|G|} + \theta \cdot \sum_{j \to i} \sum_{k \to j} \text{sim}(w_i, w_j) \cdot \text{TR}(w_j)
\]

TextRank gives the importance of each word

\[
\text{pop}(w_i) = \text{fre}(w_i) \cdot \text{weight}(w_i)
\]

\[
\text{weight}(w_i) = \frac{\text{TR}(w_i)}{|C_i|} \cdot \sum_{w_k \in E_i} \text{fre}(w_k)
\]

Experiment:
Our result is proved to be more accuracy and robust than other model on depicting event popularity.

Keywords: Cross Platform, Event Popularity, EPTS, Word2Vec, TextRank, DTW, Visualization