

Automatically identifying periodic social events from Twitter

Florian Kunneman and Antal van den Bosch
Centre for Language Studies, Radboud University
P.O. Box 9103, NL-6500 HD Nijmegen, The Netherlands
{f.kunneman,a.vandenbosch}@let.ru.nl

April 7, 2015

Abstract

Many events that are referred to on Twitter are of a periodic nature, characterized by roughly similar intervals in between. Examples are bi-annual conferences, annual music festivals, weekly television programs and the full moon cycle. We propose a system that can automatically identify such events in an open-domain fashion from a long period of tweets. Based on such a system, any event in a periodic sequence can be enriched by linking it to previous and later editions. Furthermore, periodic event patterns can be leveraged to predict future events before they are mentioned on Twitter.

Our system is applied in three stages. Starting from tweets that go back to december 2010, it firstly extracts explicit event references to build a calendar of events. Secondly, the system finds similar events in time from these calendar entries and scores sequences by their periodicity. In the third stage, future events are predicted based on the most periodic sequences from the past.

We provide an overview of the methods and performance in these three stages, and show how the system is applied in an online setting. This work not only has relevance in a social media context. Many text collections carry periodic patterns, which can be leveraged to disclose knowledge about their domain.