Weibo-Oriented Chinese News Summarization Task at NLPCC 2015

1. Overview

With the popularity of social media web sites (e.g. Sina Weibo), news stories are usually released and propagated on social media sites in recent years. A short summary for each hot news is posted as a Weibo message, together with a URL link to the full news article. People can quickly understand the news by reading the short summary and then decide whether to read the full news article or whether to comment on the news. Till now, there are many news accounts on Sina Weibo (e.g. 头条新闻, 财经网, etc.) for news release. Currently, the short news summaries given by these accounts are provided by human editors. In this evaluation task, we aim to investigate news summarization techniques for automatically generating short summaries for news release and propagation on Sina Weibo. Traditional news document summarization techniques have been widely explored on the DUC and TAC conferences, and we encourage participants to analyze the differences between traditional news summarization task and this summarization task, and then develop more competitive summarization methods.

2. The Task

The Weibo-Oriented Chinese News Summarization Task is defined as a task of automatically generating a short summary for a given Chinese news article, and the short summary is used for news release and propagation on Sina Weibo. The length of the short summary is less than 140 Chinese characters. The task can be considered a special case of single document summarization. We will provide sample data and test data later. The contents on some specified accounts on Sina Weibo are NOT allowed to be used by the participants.

3. Scoring Metric

The peer summaries will be automatically compared with the reference/model summaries by using the ROUGE toolkit. ROUGE-N F-measure will be used as evaluation metrics.