



Brands in NewsStand: Spatio-Temporal Browsing of Business News

Ahmed Abdelkader, Emily Hand and Hanan Samet

Center for Automation Research, Institute for Advanced Computer Studies
Department of Computer Science, University of Maryland at College Park

Motivation

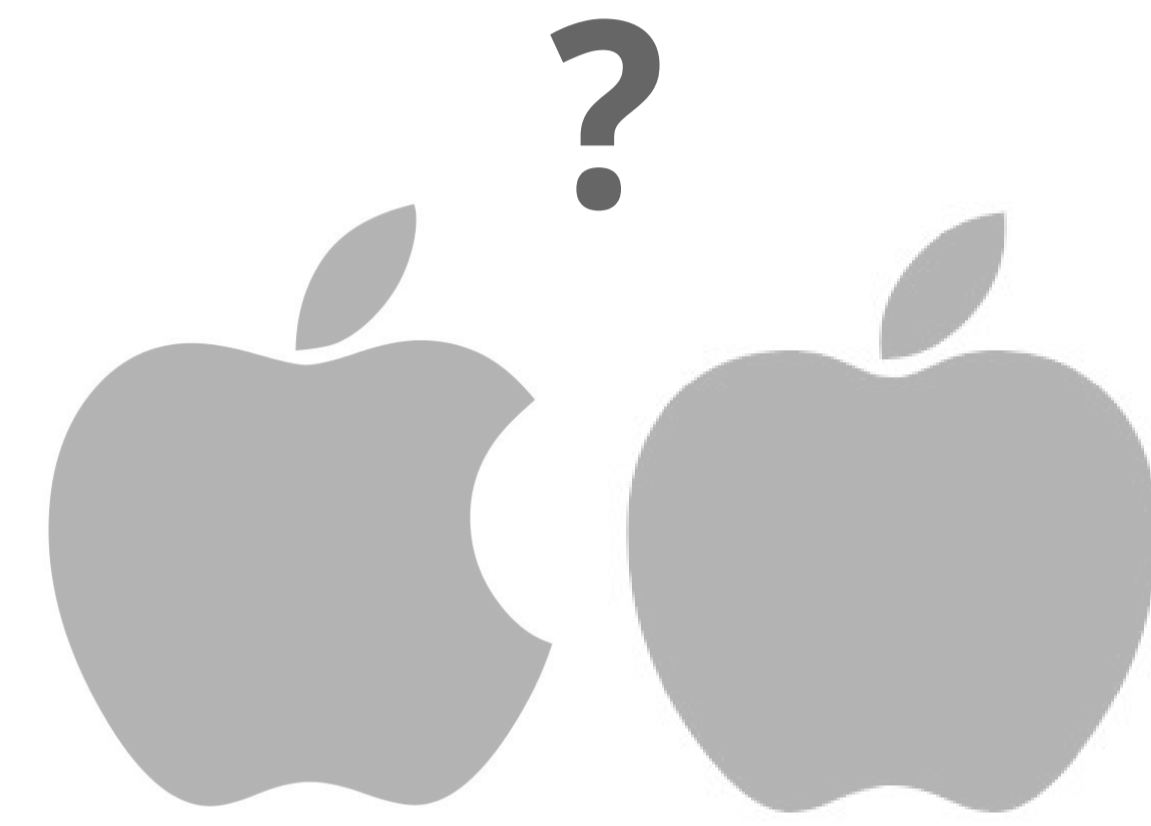
300M Unique Pageviews / Day
Top 10 News Websites Only

2X Volatility of Stock Prices
On Days with Relevant News

Challenges



How to learn new brand names?



How to disambiguate brand names from normal words?



How to recognize relevant news articles?

Methodology

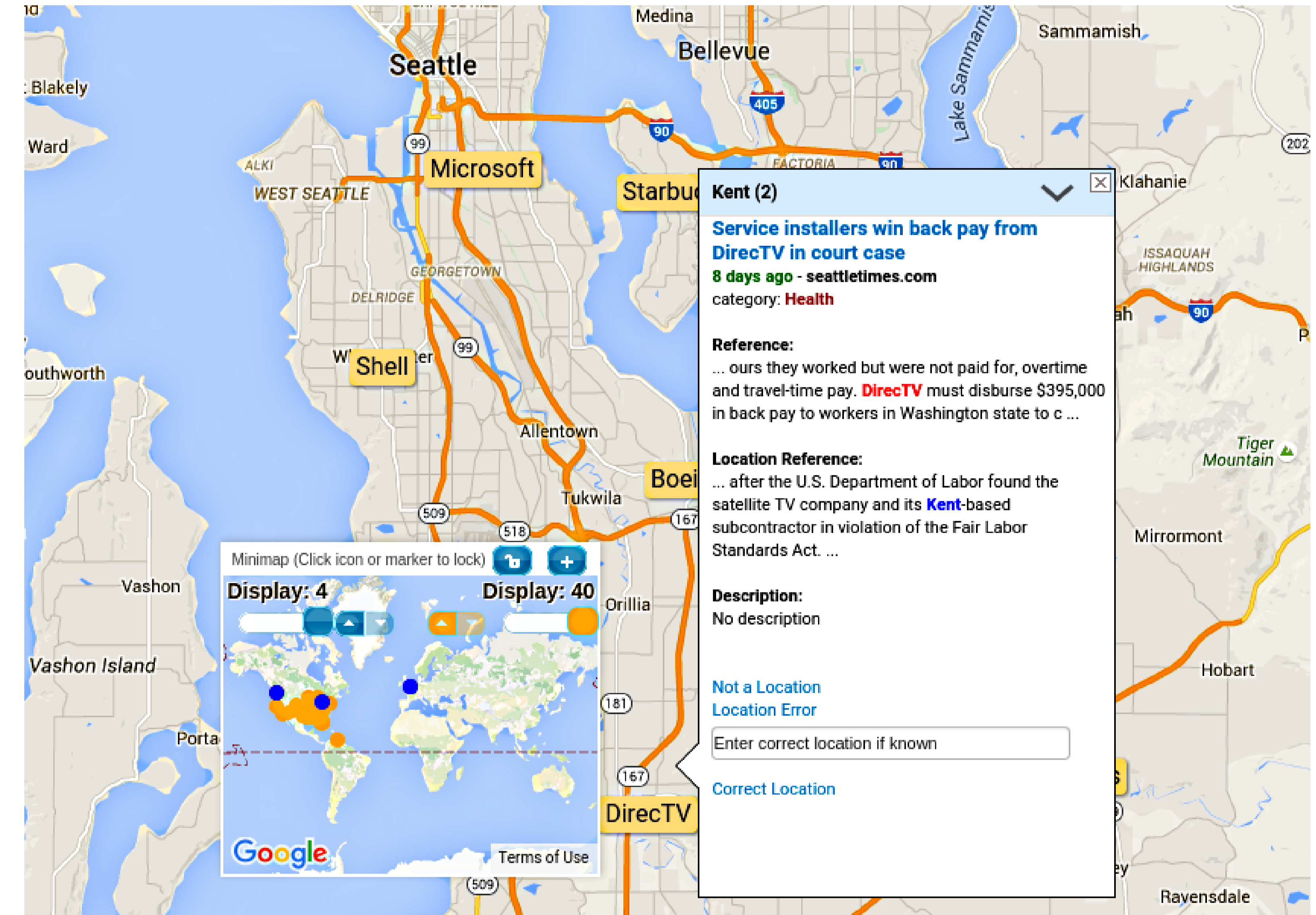
- Use a fixed list of 1000 brand names obtained from the web.
- For each sentence containing brand name words we produce two flags:
 - ▷ Brand flag: true if the brand name word references a real business entity.
 - ▷ Context flag: true if the treatment of the business is relevant.
- Evaluate using a dataset of 1190 instances (365 +ve and 825 -ve).

Four Approaches

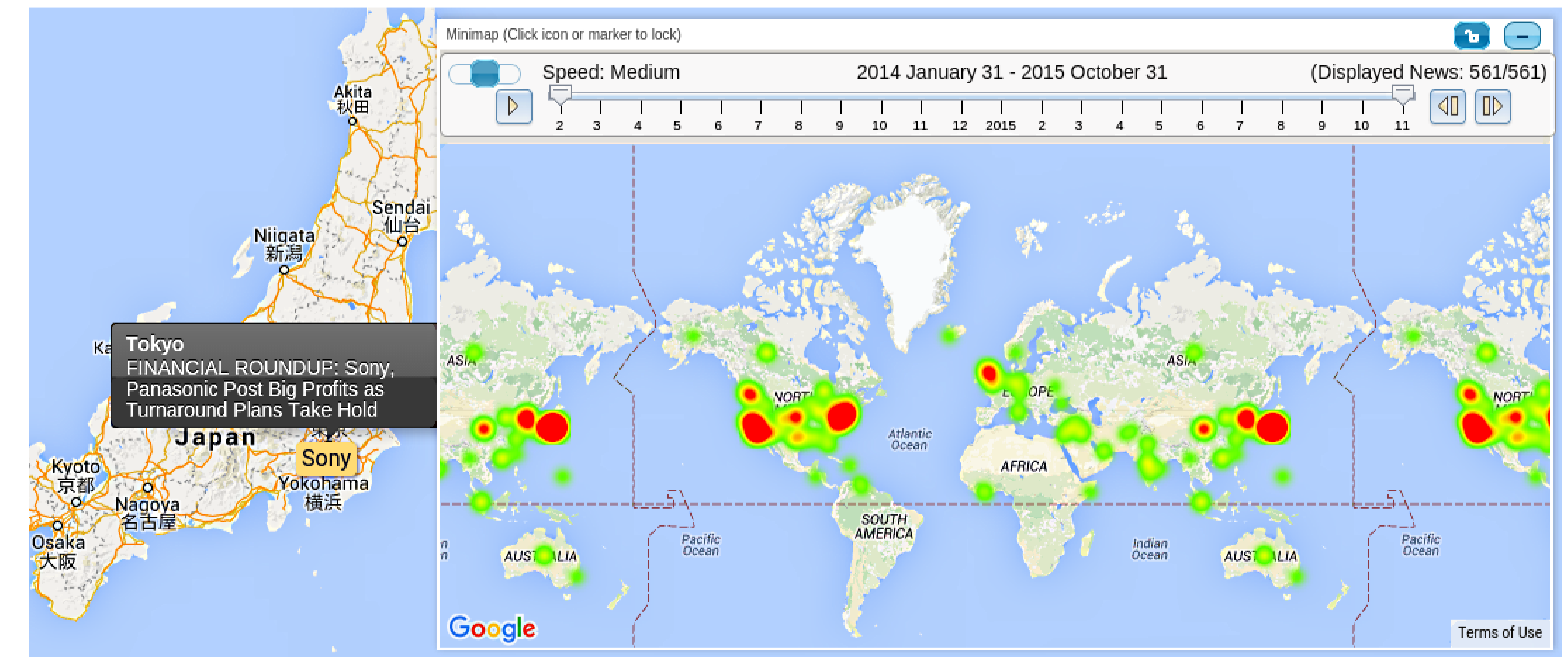
- Plain string matching (baseline). (P=0.31, R=1.00)
- Filter Part-of-Speech tags. (P=0.67, R=0.50)
- Filter Named-Entity Recognition tags. (P=0.47, R=0.61)
- Learn a classifier from a labeled dataset. (P=0.70, R=0.82)

Sample Usage Scenarios

► Map mode: shows brand tags at the associated locations.



► Time mode: gives access to the history of articles referencing a given brand.



Future Work

- Appeal to language models to learn better classifiers.
- Learn new brand names in real-time.
- Design innovative map interfaces for new types of queries.