Traveling Trends
Social butterflies or frequent fliers?

with O. Varol, F. Menczer, A. Flammini

Emilio Ferrara - @jabawack
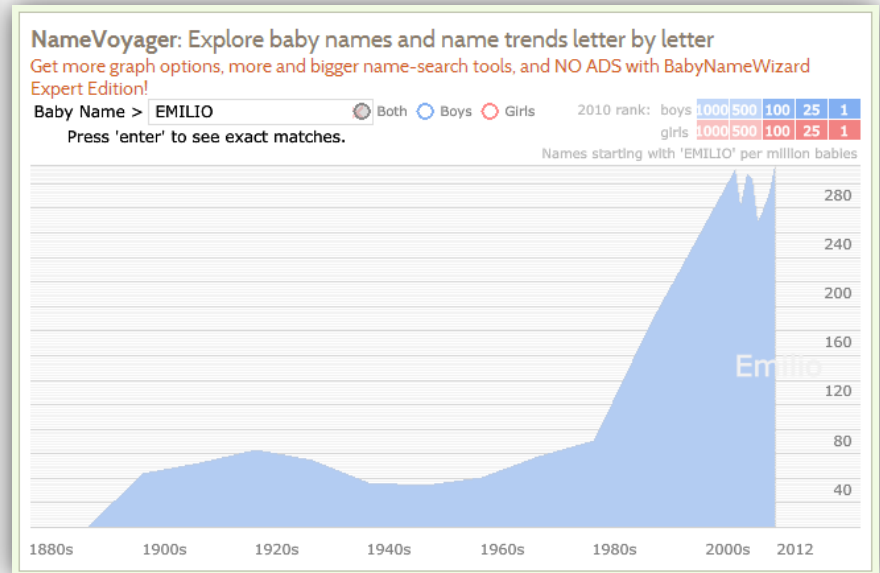
Center for Complex Networks and Systems Research

School of Informatics and Computing
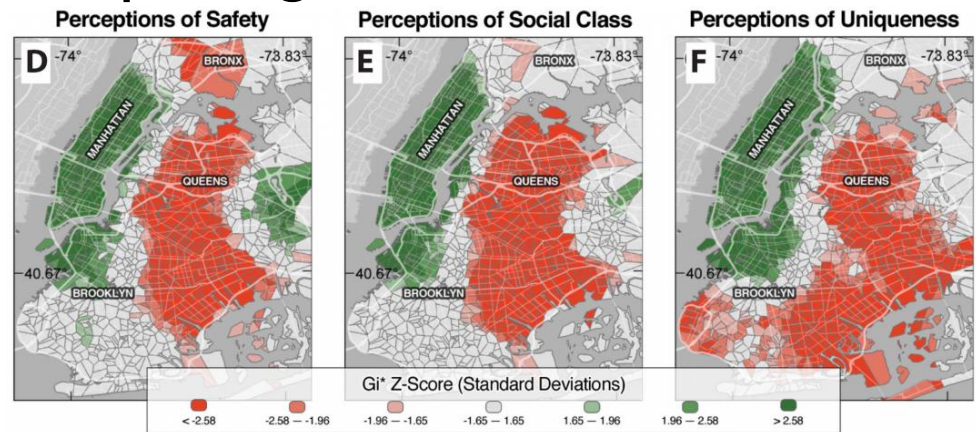
Indiana University Bloomington (USA)

Trends and collective attention

▶ What is a trend? – Baby names trends

Trends and collective attention

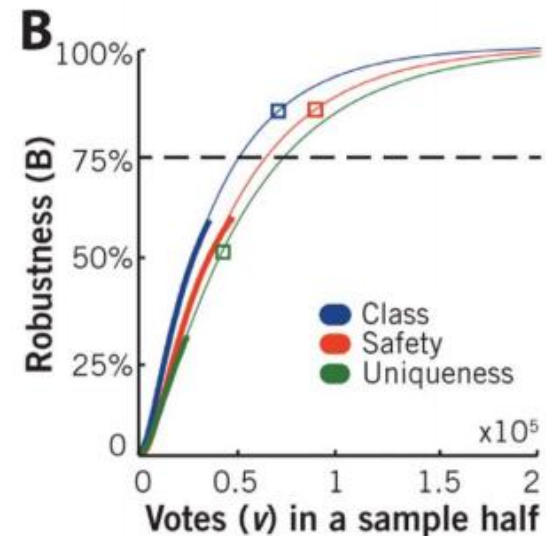▶ ## What is a trend? – House pricing trends

Perceptions of Safety | Perceptions of Social Class | Perceptions of Uniqueness

Gi* Z-Score (Standard Deviations)

A Which place looks safer?

The collaborative image of the city: mapping the inequality of urban perception
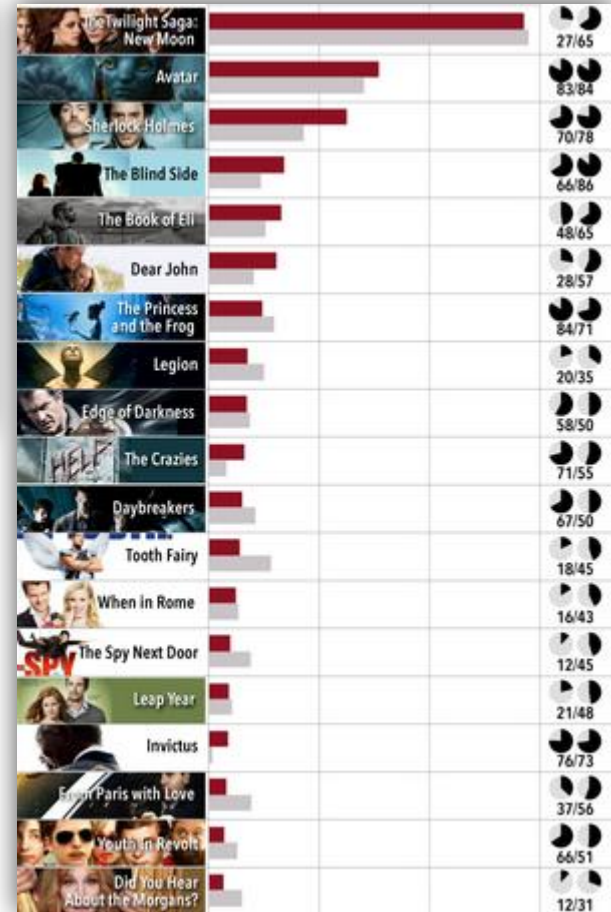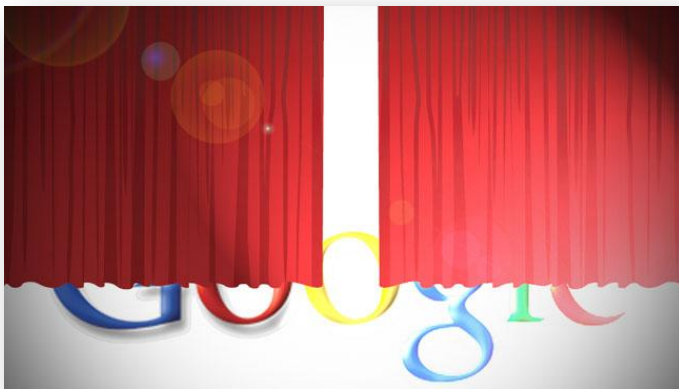P Salesses, K Schechtner, CA Hidalgo. PloS one 8 (7), e68400, 2013

Trends and collective attention

▸ ## What is a trend? – Box office predictions

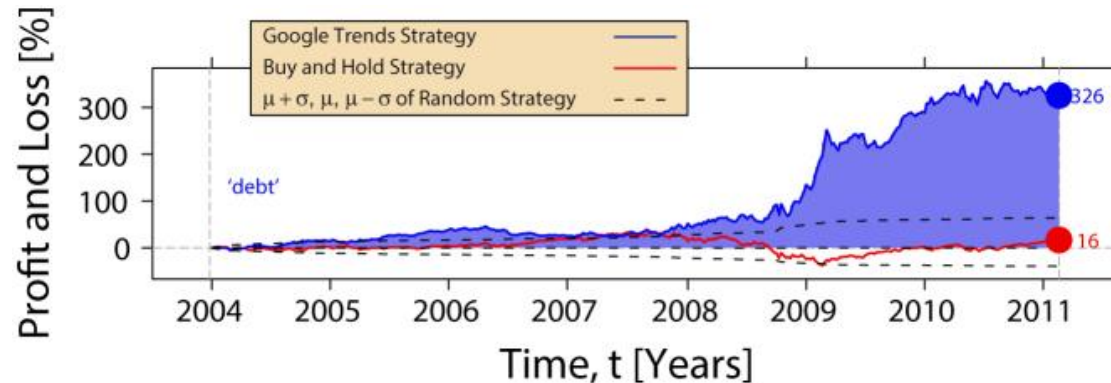Predicting the future with social media. S Asur, BA Huberman. WI-IAT, 2010

Early prediction of movie box office success based on Wikipedia activity big data
M Mestyán, T Yasseri, J Kertész. PLoS ONE 8(8): e71226, 2013

Trends and collective attention

▶ ## What is a trend? – Financial market

[Quantifying trading behavior in financial markets using Google Trends](). T Preis, HS Moat, HE Stanley. Scientific reports 3, 2013
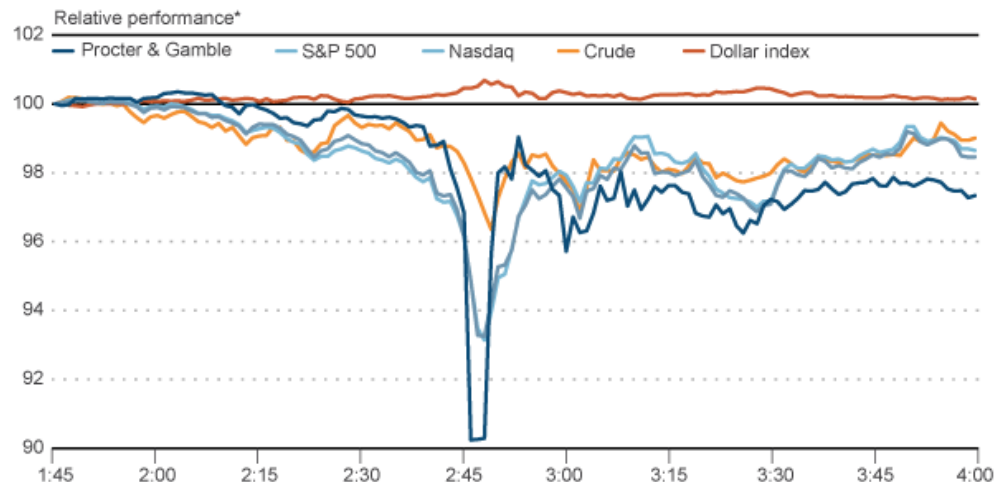
Gaming collective attention

▸ ## How to game the system to drive collective attention

  ▸ Persuasion: orchestrated manipulation, rumors, advertisement…

Toward detecting persuasion campaigns in social media. E Ferrara, O Varol, S Malinchik, F Menczer, A Flammini. Arxiv, 2013

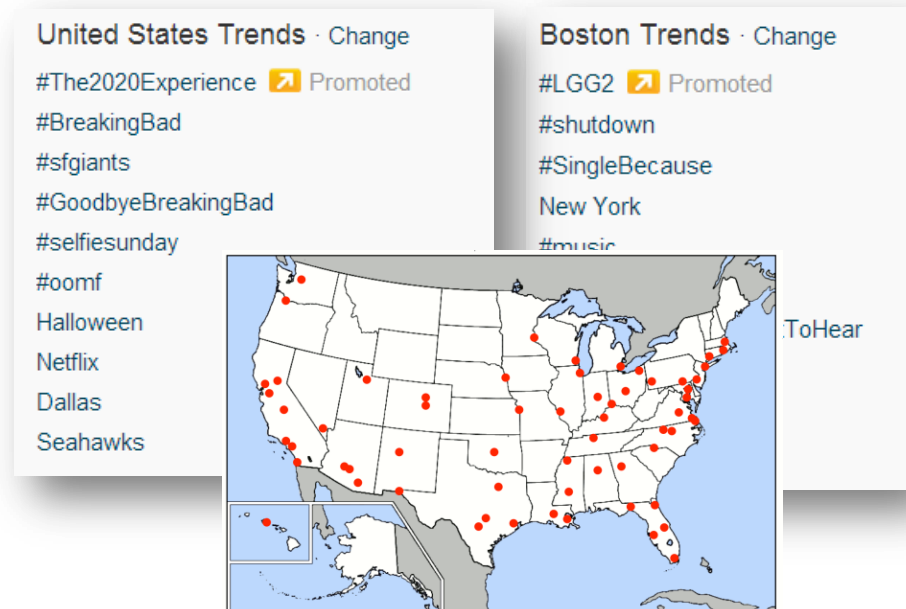# Twitter trends dataset

- Datasource: **twitter**
    - We collect 63 **US cities trends** and also **US national trends**
    - Real-time monitoring (10 min intervals, 100% uptime)

- Period: 50 days, from April, 12th 2013 to May, 31st 2013
    - We remove promoted hashtags

- Total trends: 11,402
    - 4,513 hashtags – 6,889 phrases

United States Trends · Change
#The2020Experience ↗ Promoted
#BreakingBad
#sfgiants
#GoodbyeBreakingBad
#selfiesunday
#oomf
Halloween
Netflix
Dallas
Seahawks

Boston Trends · Change
#LGG2 ↗ Promoted
#shutdown
#SingleBecause
New York
#music

ToHear

Results outline

▸ Geo-temporal trend analysis

▸ Trendsetters and trend-followers

▸ A conjecture: Social butterflies or frequent fliers?

Spatio-temporal trend analysis

▸ **Trends spatial bimodal distribution:**

    ▸ Most trends are popular only in one/few cities

    ▸ Many trends spread in all country

    ▸ The remainder fail to achieve global popularity

▸ **Trends temporal distribution (lifetime):**

    ▸ Lifetime broadly distributed:

        ▸ 68% < 20m – 95% < 6h

        ▸ 0.3% > 1d

    ▸ Entropy defined as

$$\mathcal{S}^j = -\sum_i P_i^j \log P_i^j, \ \text{with} \ P_i^j = \frac{t_i^j}{\sum_k t_k^j},$$

t_i^j the time topic j trended in location i

▸ **Trends reaching more places live longer**
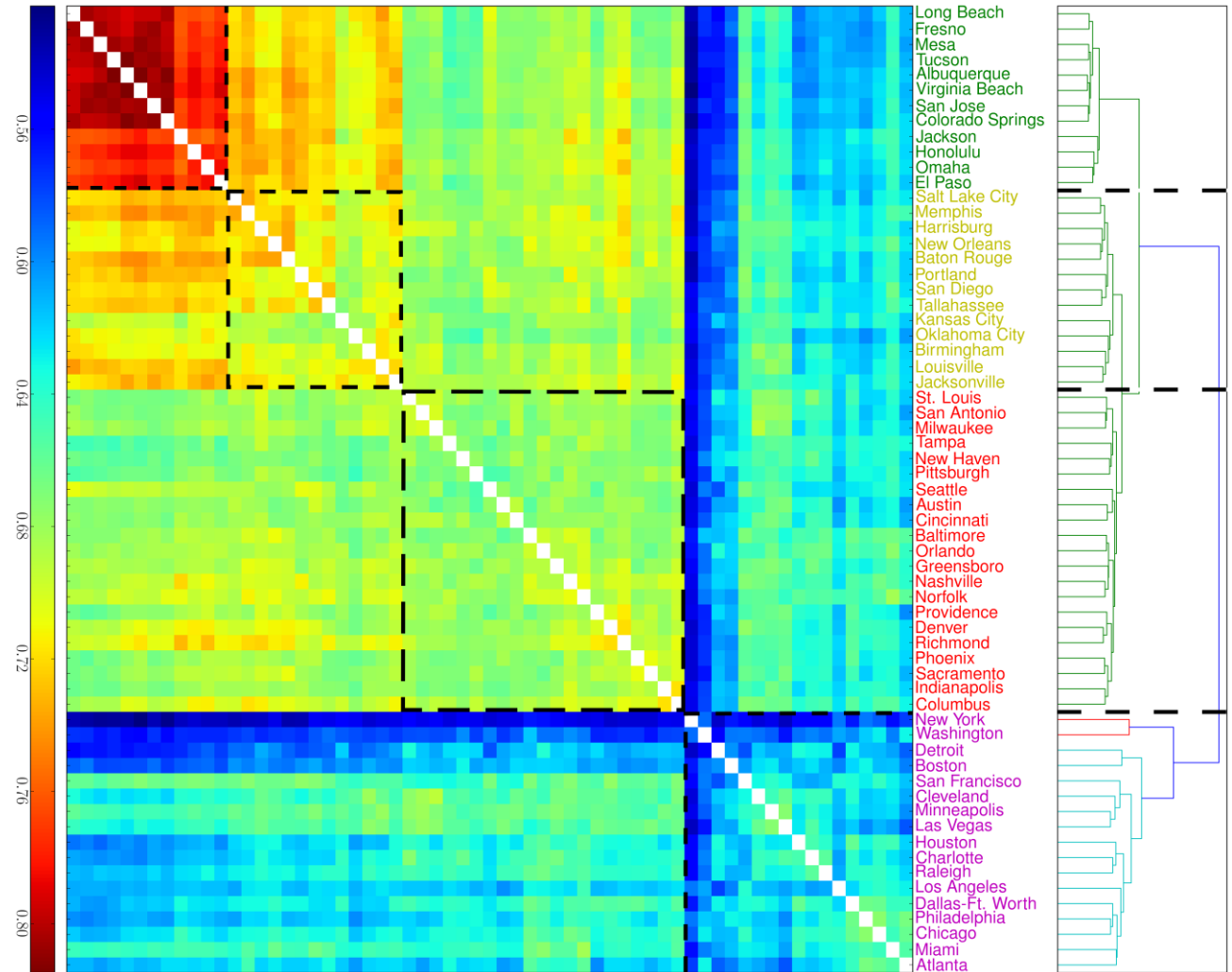
    ▸ Low entropy: low expected lifetime
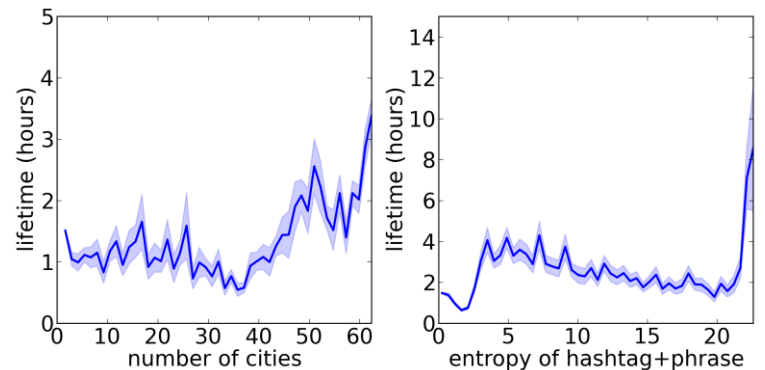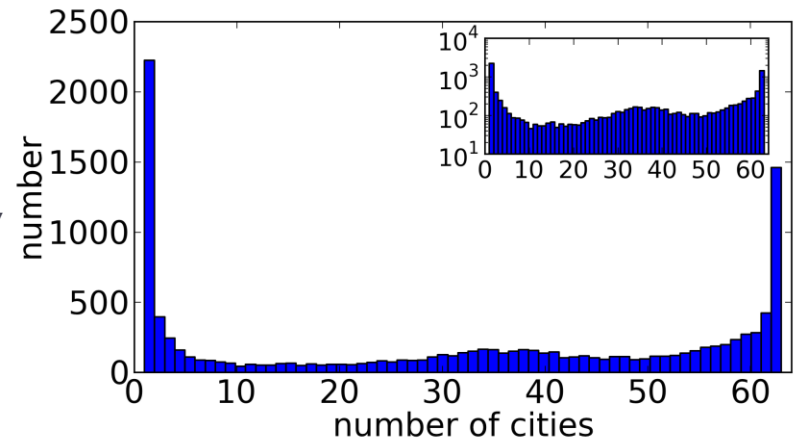
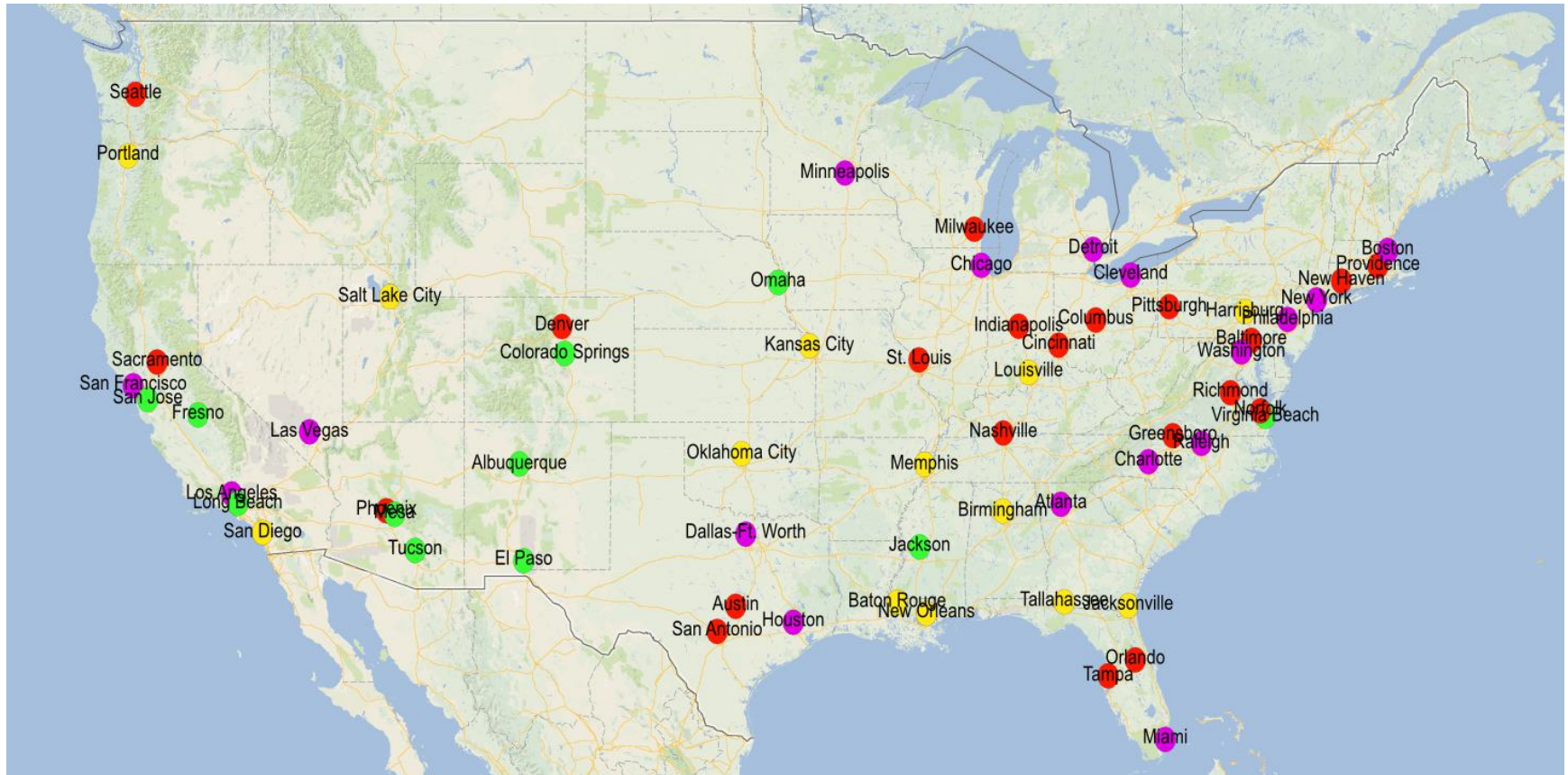# Spatial trend similarity analysis

- Shared trends similarity:

$$S_{ij} = \frac{|T_i \cap T_j|}{|T_i \cup T_j|}$$

- G: South West

- Y: Midwest

- R: East Coast

- Purple: ?

- Hint: big cities!

- Locality effect

Geography of trends

Geographic representation of the 63 cities and their respective clusters.

Significance of trends spatial clusters

- Clustering significance verification:
  - Compute distribution of similarity values for all pairs of locations belonging to the same cluster (intra-cluster similarity)
  - Compute distribution for all pairs belonging to different clusters (inter-cluster similarity)

- Kernel smoothing:
  - Apply Kernel Density Estimation to estimate the probability density functions of the distributions

- Significance:
  - T-test for any pair of distributions of intra- and inter-cluster similarity to determine if they might origin from the same distribution
  - $P < 0.01$: the four clusters are **significant** at the 99% C.L.
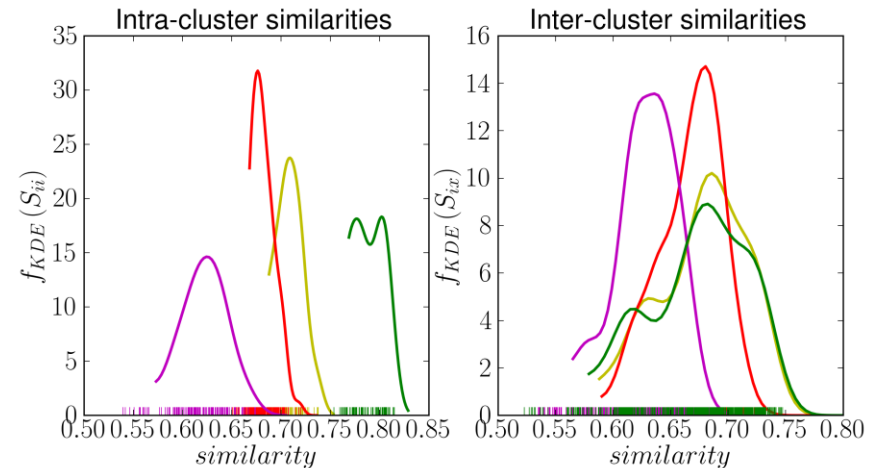
Green	Yellow	Red		Purple	
Long Beach	Memphis	St. Luis	Greensboro	Washington	Dallas-Ft. Worth
Fresno	Salt Lake City	San Antonio	Nashville	New York	Chicago
Mesa	Harrisburg	Milwaukee	Norfolk	Detroit	Philadelphia
Tucson	New Orleans	Tampa	Providence	Boston	Miami
Albuquerque	Baton Rouge	Pittsburgh	Denver	San Francisco	Atlanta
Virginia Beach	Portland	New Haven	Richmond	Cleveland	Los Angeles
San Jose	Tallahassee	Seattle	Phoenix	Minneapolis	Raleigh
Colorado Springs	San Diego	Cincinnati	Sacramento	Las Vegas	Houston
Jackson	Kansas City	Austin	Columbus		
Honolulu	Oklahoma City	Orlando	Indianapolis		
El Paso	Birmingham	Baltimore			
Omaha	Louisville				
	Jacksonville				

Trends pathway analysis

- Backbone extraction: α=0.3 (to keep only significant links)
- Country backbone: East-to-West ⟷ West-to-East
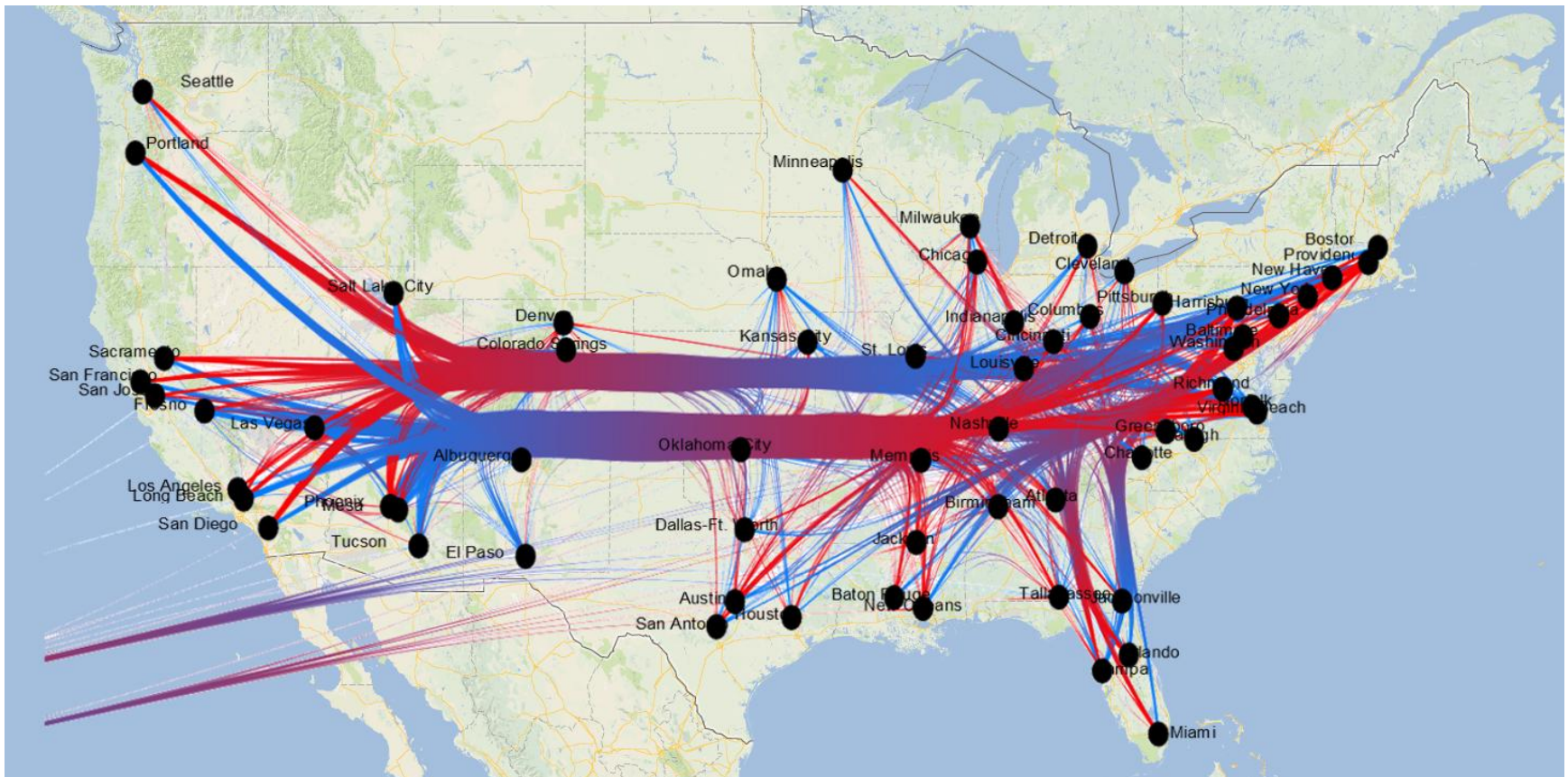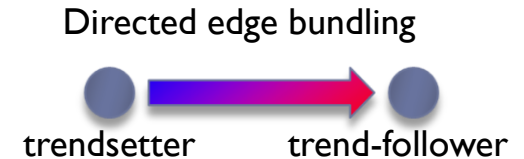
Directed edge bundling

trendsetter   trend-follower

Resutls outline

▸ Geo-temporal trend analysis

▸ **Trendsetters and trend-followers**

▸ A conjecture: Social butterflies or frequent fliers?

Trendsetters and trend-followers

- Q: Are trending topics that become popular at the country level produced uniformly by all cities, or preferentially by some of them?
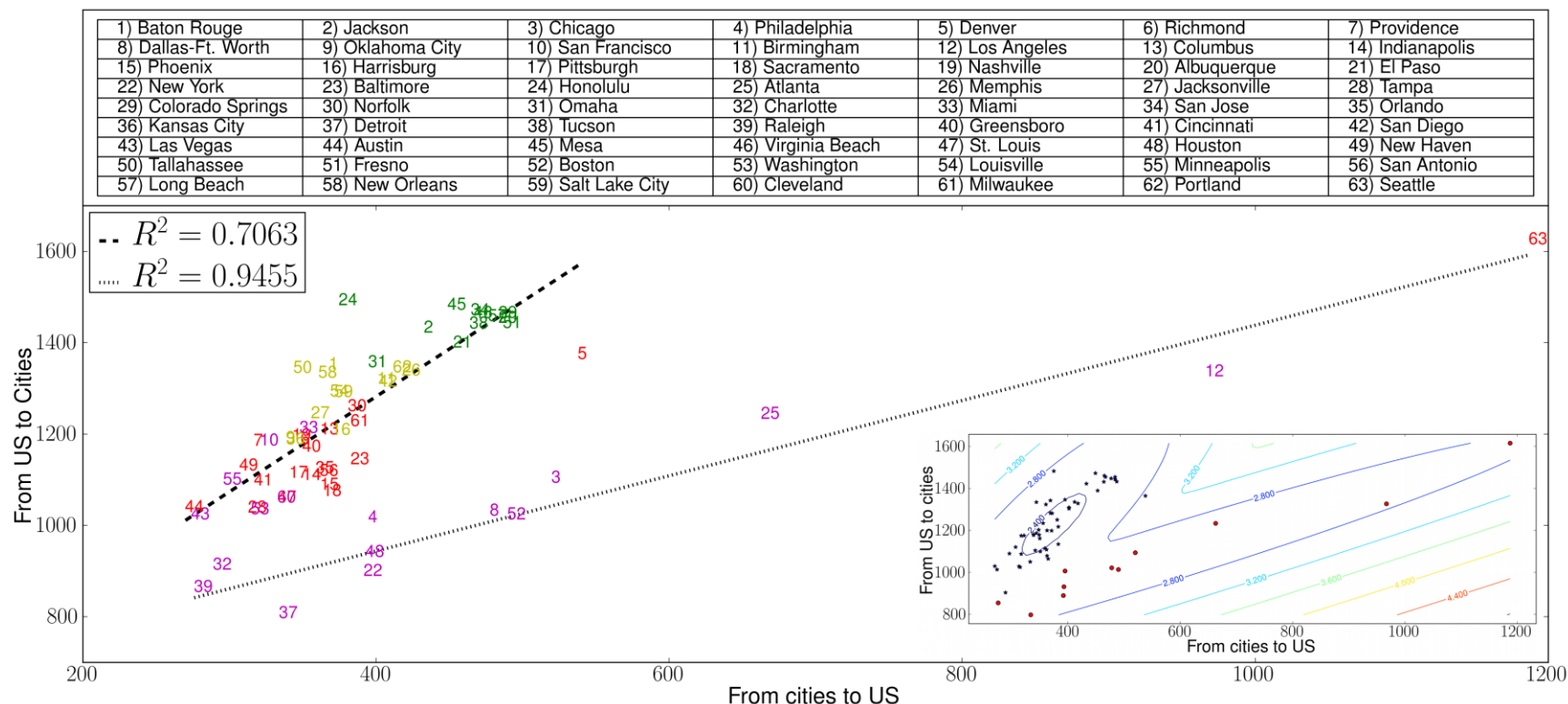
- Sources: **trendsetters**
- Sinks: **trend-followers**

- Weighted sink-source ratio:

$$\omega(n) = \frac{s_{out}(n)}{s_{in}(n) + s_{out}(n)}$$

- Top ranked cities: 4/5 **major metro areas**
- Los Angels & NY: also top worldwide HT producers (Kamath et al. WWW 2013)

- **All sinks** happen to be in the Midwest or Southwest of the country

Location	Rank	$\omega(n)$
Los Angeles	1st	0.806
Cincinnati	2nd	0.736
Washington	3rd	0.718
Seattle	4th	0.711
New York	5th	0.669
	…	
	…	
Kansas City	59th	0.352
Omaha	60th	0.352
El Paso	61st	0.235
Albuquerque	62nd	0.109
Oklahoma City	63rd	0.101

National trendsetters and trend-followers

- ‣ Trendsetters vs. trend-followers
  - ‣ X: no. times a topic trending in a given city later becomes a **national trend**; Y: the reverse effect
  - ‣ Inset: a Gaussian Mixture Model identifying two different trendsetting dynamics

Results outline

▶ Geo-temporal trend analysis

▶ Trendsetters and trend-followers

▶ A conjecture: Social butterflies or frequent fliers?

Social butterflies or frequent fliers?

- Q1: Does mere city size explain the trendsetting dynamics?
    - Larger cities produce more tweets; this yields to **more potential topics competing** for collective attention, but the maximum number of possible trends is fixed to 10 at the same time!
    - As a result, the effect of sheer volume is <u>discounted by construction</u> in the definition of Twitter trends
- Q2: Why the metro areas play such a trendsetting role, then?

Location	Rank	Traffic	Class
New York	1st	54M	‖‖‖‖‖‖
Atlanta	2nd	45M	‖‖‖‖‖‖
Chicago	3rd	41M	‖‖‖‖‖‖
Miami	4th	33M	‖‖‖‖‖‖
Dallas-Ft. Worth	5th	32M	‖‖‖‖‖‖
Washington	6th	31M	‖‖‖‖‖‖
Los Angeles	7th	31M	‖‖‖‖‖‖
Denver	8th	25M	‖‖‖‖‖‖
Charlotte/Raleigh	9th	24M	‖‖‖‖‖‖
Houston	10th	24M	‖‖‖‖‖‖
San Francisco	11th	21M	‖‖‖‖‖‖
Las Vegas	12th	20M	‖‖‖‖‖‖

- 16/17 **purple** cities are also top 20 air traffic hubs!
    - Major travel cities including Atlanta, Chicago, Los Angeles
- Some **purple** cities are not in top 30 most US populated metro areas…
    - Charlotte, Raleigh, and Las Vegas appear among the major traffic hubs!
- Does information travel faster by airplane than over the Internet?

Conclusions and future work

- Trends reflect a **locality effect**: they diffuse locally more than globally in three specific geographical areas
  - East-coast, Midwest, South West
- There is a fourth class of **metropolitan cities**
  - They are spread all over the country
  - They act as trendsetters (they produce most national trends)
  - They correspond to major air traffic hubs
- Open questions:
  - What's the role of **traffic hubs** in trend diffusion?
  - What's the role of **people**?!

Emilio Ferrara   Onur Varol   Filippo Menczer   Alessandro Flammini

Thanks! Questions?
@jabawack

**E Ferrara, O Varol, F Menczer, A Flammini.
Traveling trends: social butterflies or frequent fliers?
ACM COSN 2013**

James S. McDonnell Foundation   DARPA   NSF