Harnessing Social Media Streams for Local Information Needs

<u>Dyaa Albakour</u>

@dyaaa

University of Glasgow

UCREL CRS, Lancaster University, 12 March 2015

Social media

As of December 2012 [1]***:***

- #users on Facebook **1.2 billion**
- #tweets per day **190 million**
- #pictures to Flickr **3,000/min**

- #people accessing the Web from mobiles **818.4 million**
- ***26%*** of mobile app usage is social networking [2]

The "2013 Q1 report" of the Global Web Index:[3]

- ***A rise*** in ***active engagement*** across all social platforms with Twitter the fastest growing (access from mobile phones)

1 http://www.statisticbrain.com/social-networking-statistics/
2 http://www.pswebsitedesign.com/social-media-and-mobile-phones/
3 http://www.globalwebindex.net/Stream-Social

Local Information Needs and Social Media

- **Local Search is attracting more demand**
  - Local Search constitutes 43% of Google Queries [1]
  - *What is happening near me now?*

  "near me", "in Lancaster", "on campus"
  - Activities I can do now or later today
- **People are using social media to reflect on real-world events in real-time [1]**
  - Communicating to their social circle (what is happening? what are they doing? where are they? ..)
  - *Sporting events, earthquakes, protests, riots..*

1 http://chitika.com/insights/2012/local-search-study/

[1] Yardi, S., and boyd, D. Tweeting from the town square: Measuring geographic local networks. In ICWSM'10.

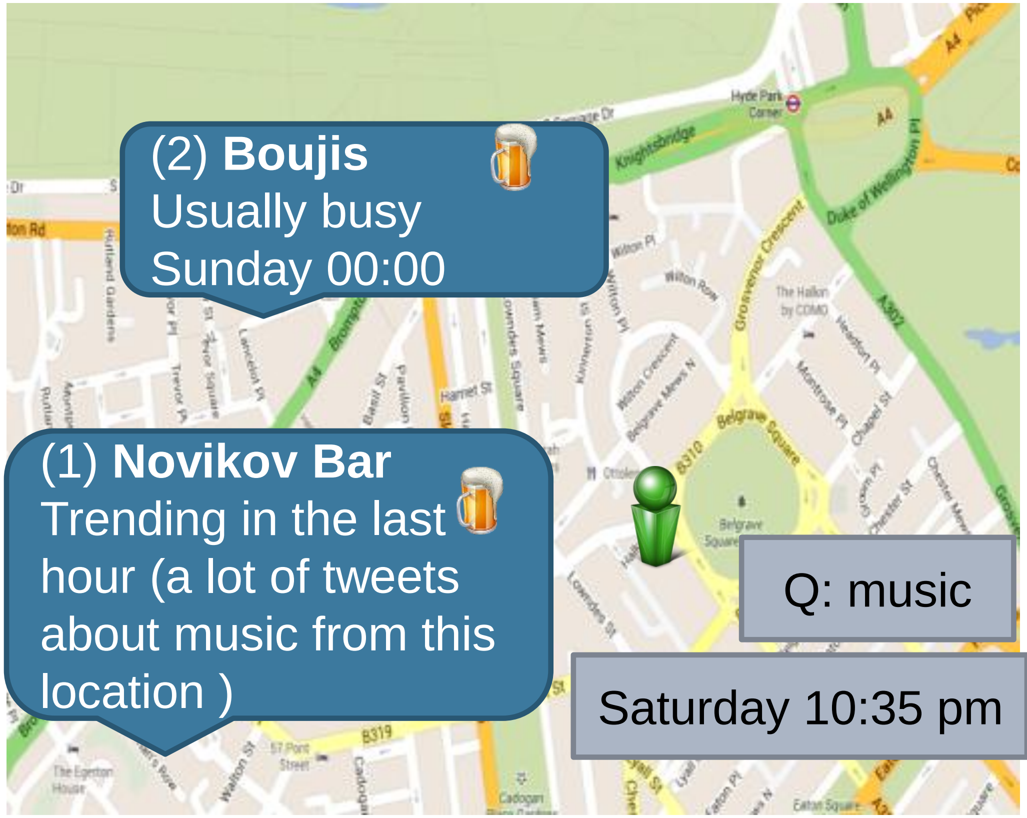
Interaction Scenarios

Input

- Keyword queries or zero-queries;
- Context (time, location and/or user profile)

Output

- *Retrieve* and rank events that has currently started from social media posts
- *Filter* social media content about the event
- *Anticipate and recommend* locations that may have interesting activities for the user



In this talk

Local Event Retrieval using Twitter as a Social Sensor

Twitter Real-time Filtering

Anticipation and Personalised Venue Recommendation using Location-bases Social Networks (LBSNs)

Using Twitter as a Social Sensor

LOCAL EVENT RETRIEVAL FROM SOCIAL MEDIA

Local Event Retrieval from Social media

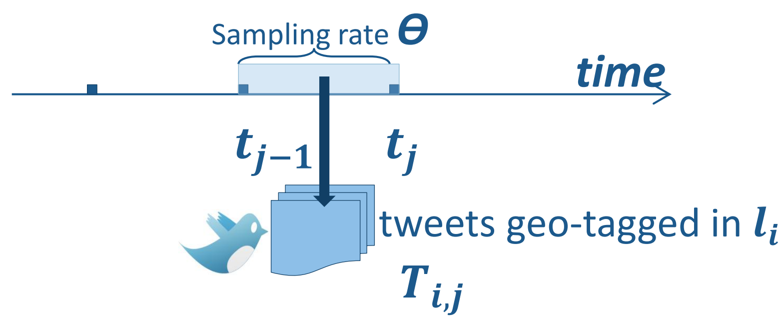
Contributions

- The new task of Local Event Retrieval from Twitter (Twitter as a social sensor)
- A framework for Local Event Retrieval
- Evaluation with a newly created dataset using crowdsourcing and local news feeds

M-D. Albakour, C. Macdonald and I. Ounis. Identifying Local Events by Using Microblogs as Social Sensors. In proceedings of OAIR 2013.

Local Event Retrieval using Twitter

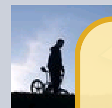
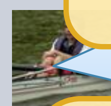
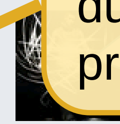
- **Given a user query (q):**
  - Retrieve a ranked list of **local events** that are **relevant** to the user query (q)
- **We model a location as a time series**



- **What people tweet reflect what is happening in a location at a certain time**
- **A local event has (1) a starting time and ; (2) a location**

Ranking function $R(q, < l_i, t_j >)$:
Rank tuples $<l_i, t_j>$ according to how likely t_j represents a starting time of a matching event that occurred in l_i using the tweets

Example of responses for query (concert)

Rank	Start Time	Location	Description (Tweets)
1	Today 19:15	Wembley	 @TheBeachBoys so enjoying your **concert**, singing and dancing away, thank you xx  1st song already has whole of Wembley Arena on their feet #BeachBoys #DoItAgain
2	Yesterday 20:00	London O2	 With Anna at the #CNBLUEinLondon **concert**! http://t.co/RVAVKrvv
3		..	..

More relevant tweets

Increased activity of tweeting in those locations during those times (than previously observed)

A Framework for Local Event Retrieval

Terrier Team | University of Glasgow

Two Components:

- Topically related tweets to $\boldsymbol{q}$ in location $\boldsymbol{l_i}$ at around $\boldsymbol{t_i}$
- Increasing tweeting activity

$$R(q, < l_i, t_j >) \sim (1-\lambda).\boxed{S(q, T_{i,j})} + \lambda.\boxed{\mathbf{E}(\mathbf{q}, < l_i, t_j >)}$$

Q … ets T_i … ne qu

(1

The voting model to aggregate ranking of individual tweets

Quantifies the **change** in the **tweeting activity** at time t_j in location l_i

(2) The change component

The Change Component

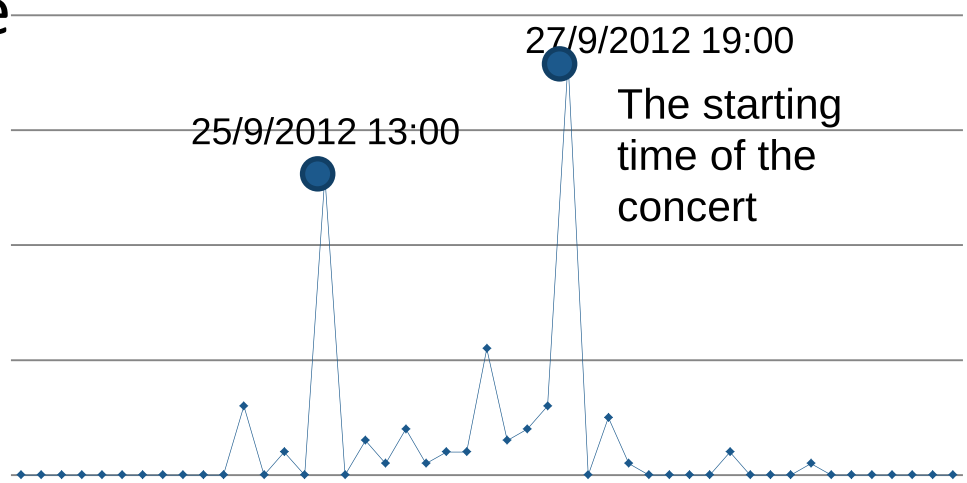
How do we estimate the change in the tweeting activity?

Change point Analysis

- Quantify how likely is the tweeting activity (about a topic) is an outlier with respect to previous observations.

- Apply the Grubb Test [2]
  Normalised score (0..1)

#Tweets about "beach boys" in London



The tweeting activity: is measured by the topical component score $S(q, T_{ij})$

[2] F. Grubb. Procedures for detecting outlying observations.*Technometrics, 11, 1969*

Experiments

Research Question:

- What is the ***impact*** of the different components, in our framework, on the **ranking effectiveness**?

Datasets

Code	Tweets	Events	Queries
Entire London	1.28m geo… within Lon… 12 days (2… → 3/10/12…	• ***Young believers choir concert***	
4 boroughs	864k geotagged within 4 c… boroughs… London 12 days (… → 3/10/1…	• From local news… • ***Hospital volunteers honoured***	The **title** of the…

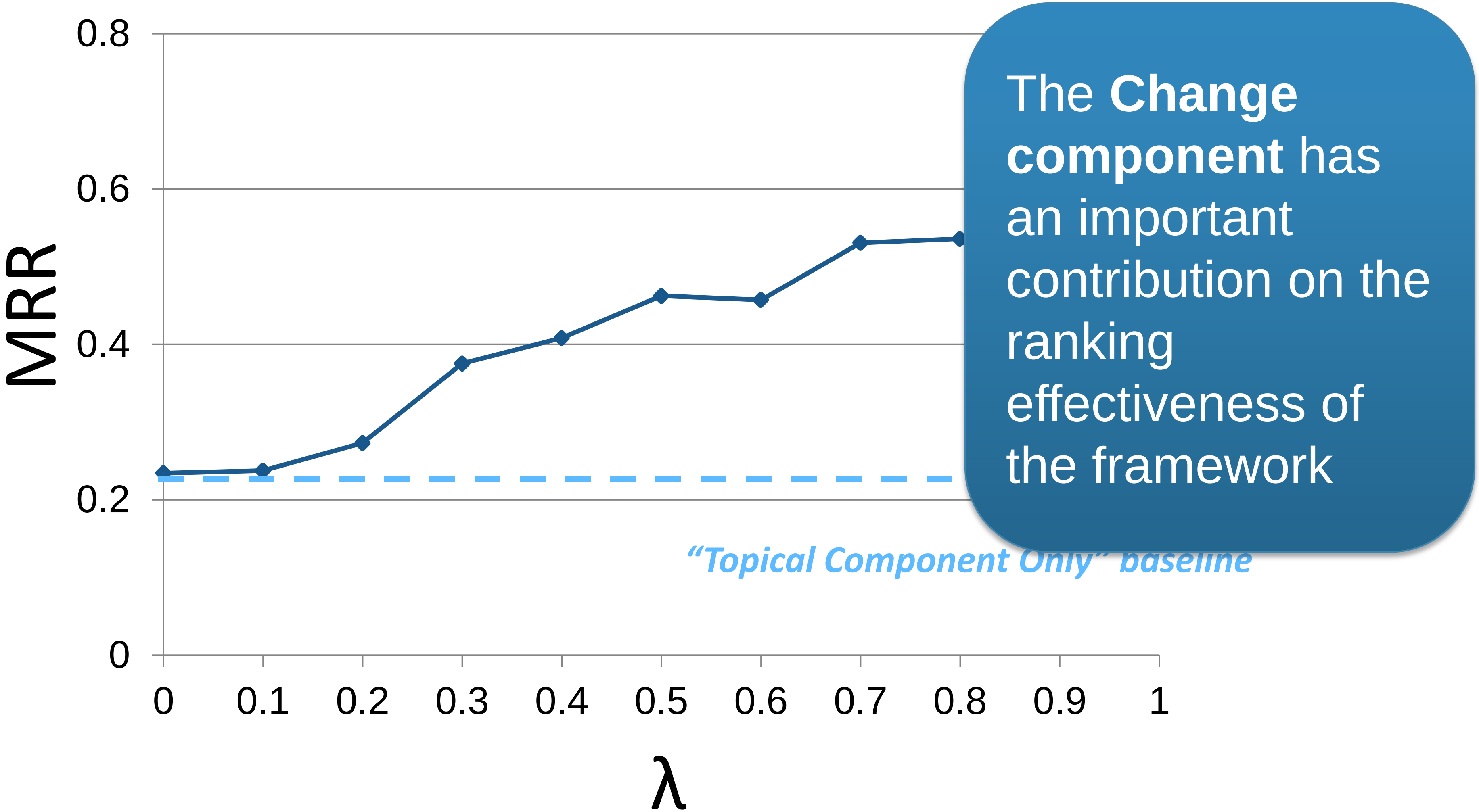
- Coarse-grained
- Single event for each query
- Popular events

- Finer-grained
- Single event for each query
- More localised events

Experimental Setup

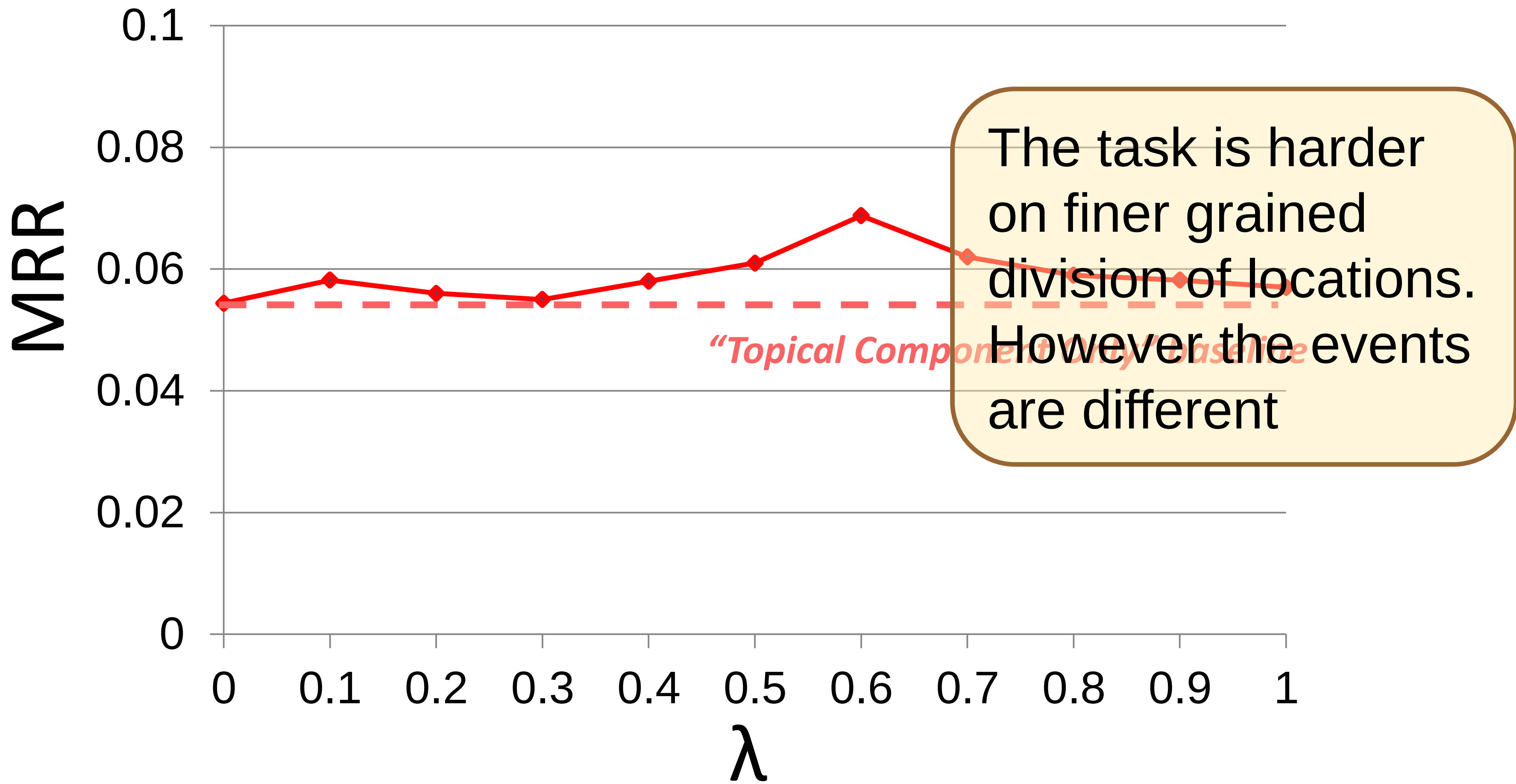
- A sampling rate of 15 minutes
- *DFReeKLIM* for ranking tweets in the voting model
- **Baseline**: using the topical component only (λ=0)
- Evaluation methodology inspired by the video segmentation evaluation for assessing the accuracy of *correctly identifying* the *starting time* of an event

Results (Entire London)



Scores obtained for the MRR measure

Results (4 boroughs)



Scores obtained for the MRR measure

REAL-TIME TWITTER FILTERING

Real-Time Tweet Filtering

- **Producers:** Huge activity around the globe (on average around 5700 published tweets per second)[1]
- **Consumers:** want to stay up-to-date with **relevant** content (not everything!)

Interested in a Topic(s): a query, and a starting time

1 https://blog.twitter.com/2013/new-tweets-per-second-record-and-how

Challenges

Terrier Team | University of Glasgow

Tweets are very short documents

→ vocabulary mismatch

US Unemployment

The tweet does NOT contain any of the terms in the query or the user profile

1- Sparsity (Brevity)

Google News #RonPaul Chairman Ron Paul to Tackle the Fed and Jobs - The New American http://goo.gl/fb/CLjEp

Thu Feb 03 2011 16:26:30

Challenges

Twitter is a **highly dynamic** social medium (it reflects a highly dynamic world!)

Long-term interest

Classic

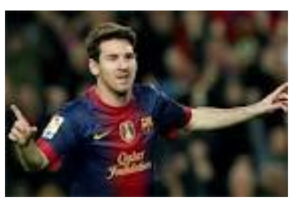
Who is going to play/miss?

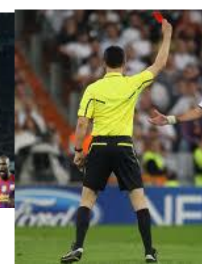
short-term interest

Goals / cards / chances ?

short-term interest

2- Drift of interests





Before the match

During the match

ne

The interests swing between different aspects (subtopics) of the more general topic (due to **events** in the real world)

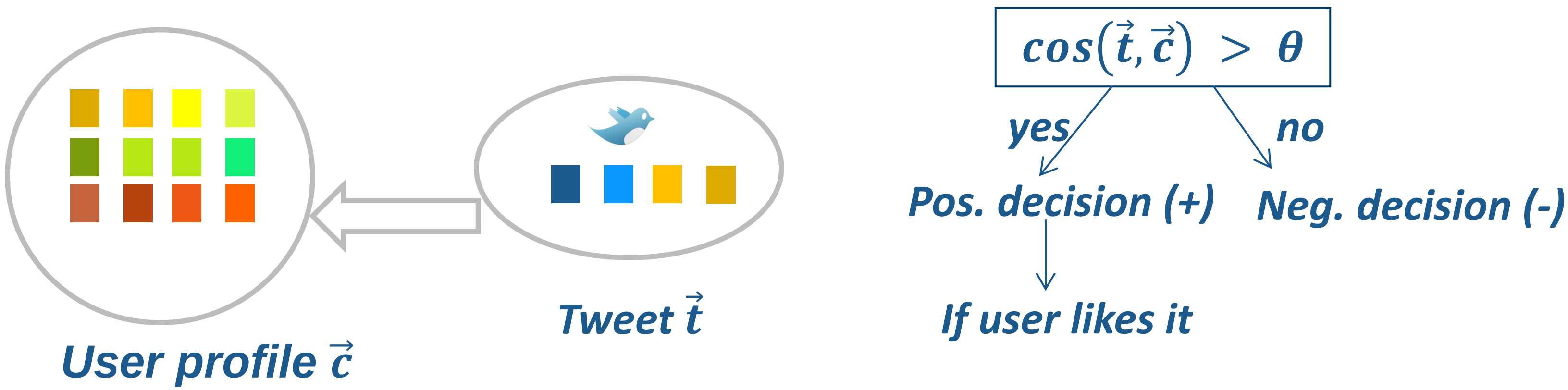
Contributions

- **Build on news filtering approaches to tackle the problem of adaptive tweet filtering**
- **Address the unique challenges in filtering tweets:**
  - Address Sparsity by deriving a richer representation of the user profile
  - Address Drift by balancing between short-term and long-term interests

M.-Dyaa Albakour, Craig Macdonald, Iadh Ounis: On sparsity and drift for effective real-time filtering in microblogs. CIKM 2013: 419-428

Tweet Filtering with Incremental Rocchio

- **We build on a common technique for News Filtering: the popular Incremental Rocchio's classifier (RC) [3]**
  - Build a profile online (vector of terms)



- **We considered another state-of-the-art news filtering approach of Regularised Logistic Regression (LR) [4]**
  - Evaluation suggests that Incremental Rocchio (RC) significantly outperforms LR (full details in the paper).

[3] J. Allan. Incremental relevance feedback for information filtering. In Proc. of SIGIR, 1996.
[4] Y. Zhang. Using bayesian priors to combine classifier for adaptive filtering. In Proc. of SIGIR, 2004

Handling Sparsity

Terrier Team | University of Glasgow

Derive **relevant** and **timely** terms for a richer representation of the centroid using query expansion (QE)

User profile $\vec{c}$

Query

BBC World Service staff cuts

BBC to cut online budget by 25%, cutting 200 websites, and 360 jobs over the next 2 years http://t.co/uD4BDRF

Terms derived with a query expansion (QE) technique

Budget
Grow
Report
Half
Media
Social
..

Top retriev[illegible]weets

(1) BB[illegible]inks online unit to **cut costs** [illegible]d refocus: LONDON (Reuters) - Britain's state-backed public **broadcaster**
(2) Irish Times: BBC World Service confirms cuts: The BBC World Service will **shed** around 650 **jobs**, or more than a qu...
(3) …

Top k

Top k

Index of tweets prior to current tweet

Experiments: Sparsity

- TREC 2012 Microblog Track – Real-time Filtering task
  - Tweets2011 (around 10m tweets over 16 days)
- We have built a real-time filtering infrastructure
  - using Storm and Terrier
- **Experimental Setup**
  - Standard stopword removal and Porter stemming
  - Dirichlet language modelling to weight terms in the vectors
  - Threshold *tuned* on the 10 TREC training topics (38 testing topics)
  - Bo1 DFR for query expansion (as provided by Terrier)

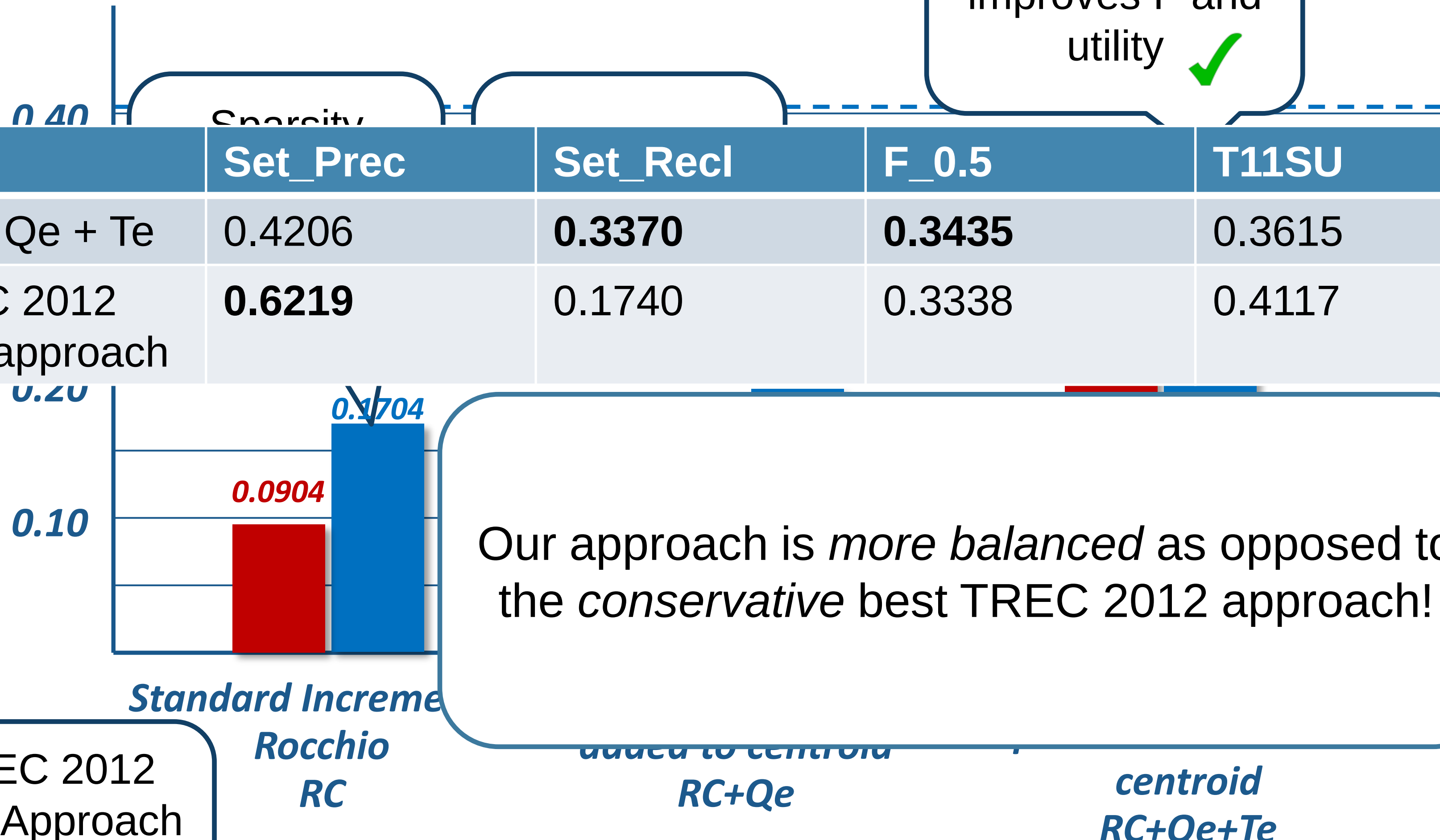
Research Question: Are our adaptations for tackling sparsity, using QE, successful in **improving** filtering **effectiveness?**

1 http://terrier.org/ 2 http://storm-project.net/

Results: Sparsity

Significantly improves F and utility ✓

	Set_Prec	Set_Recl	F_0.5	T11SU
RC + Qe + Te	0.4206	**0.3370**	**0.3435**	0.3615
TREC 2012 Best approach	**0.6219**	0.1740	0.3338	0.4117



Our approach is *more balanced* as opposed to the *conservative* best TREC 2012 approach!

TREC 2012 Best Approach
F_0.5
T11SU

What is Drift?

Terrier Team

University of Glasgow

Illustrative Example

Topic:
BBC World Service Cuts

Jan 24

BBC to cut **online budget** by 25%, cutting 200 websites, and 360 jobs over the next 2 years http://t.co/uD4BDRF

The day when BBC announces that it will **cut its online budget**

Jan 26

BBC World Service axes five **language services** (AFP) - AFP - The BBC World Service has said it will close five o... http://ow.ly/1b23Gf

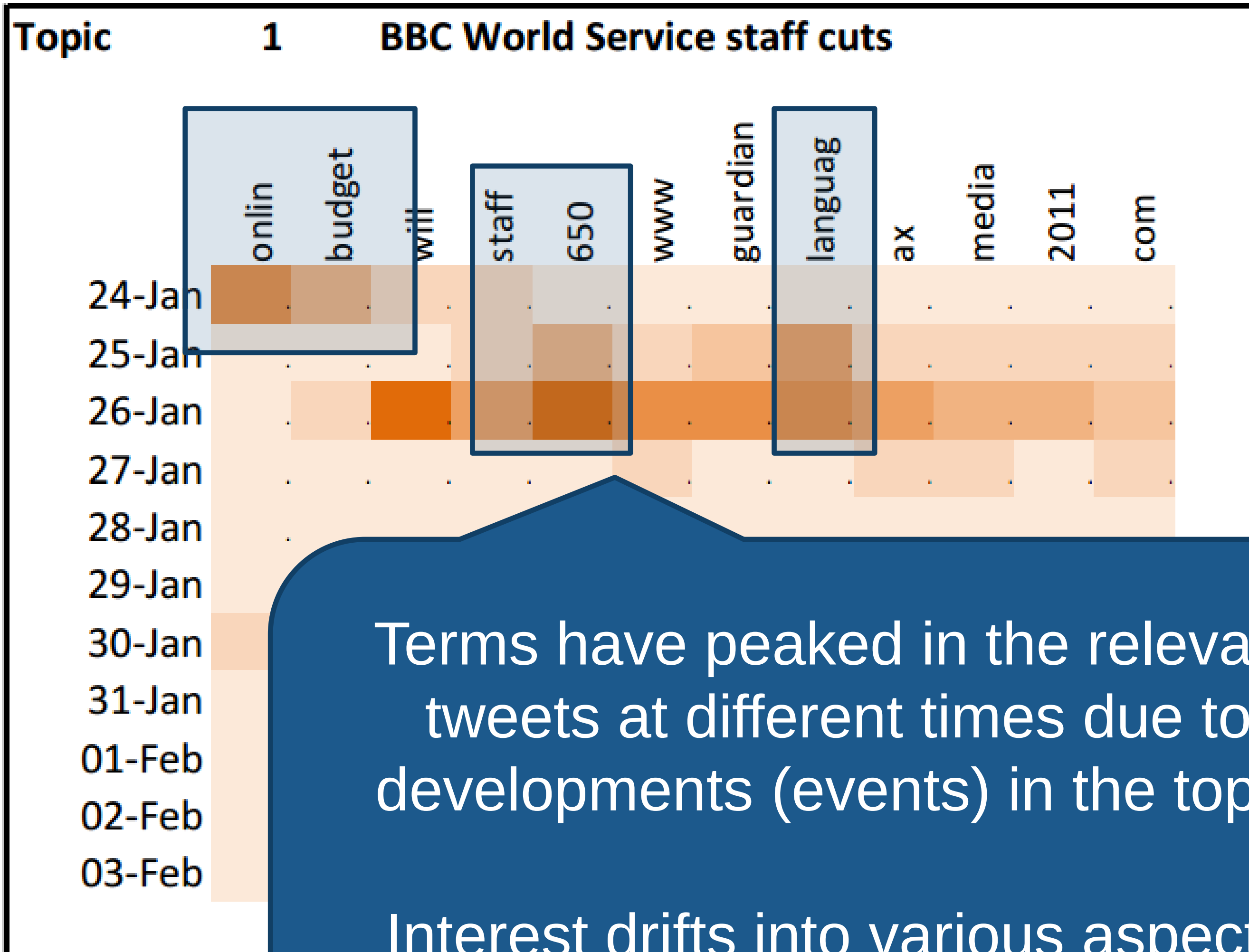
BBC to axe **650 jobs** at World Service after Foreign Office cuts £50million funding: Today's announcement of the c... http://bit.ly/hrC109

On that day, two stories:

1) **Cutting five language services**
2) **Slashing 650 jobs**

Time

Empirical Viewpoint of Drift

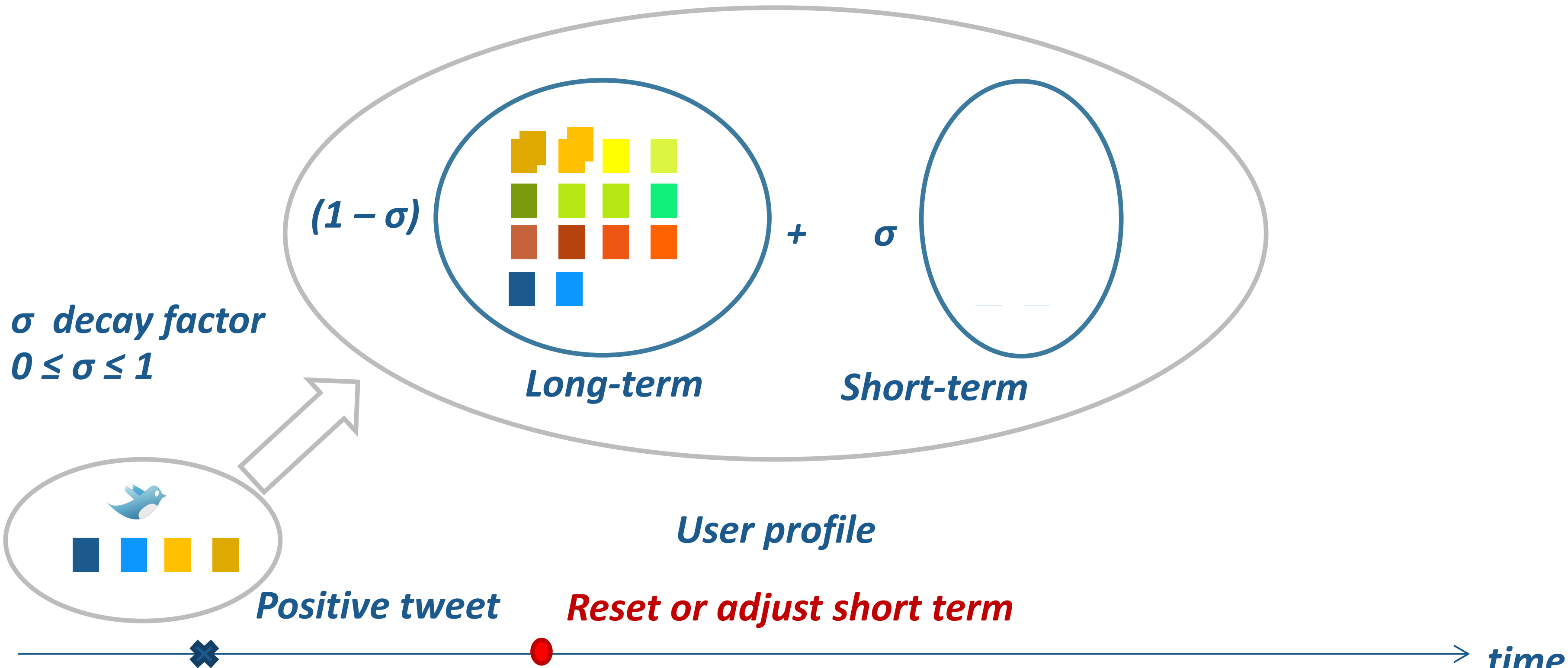


Terms have peaked in the relevant tweets at different times due to developments (events) in the topic

Interest drifts into various aspects (sub-topics) over time

Handling Drift

- Dynamically changing the centroid over time to represent both **short-term** interests and **long-term** interests in the overall topic (combined with the QE approach)



When do we reset/adjust short-term interests?

When does drift occur?

When do we reset/adjust short-term interests?

1. Arbitrary adjustments: The most recent n positive tweets

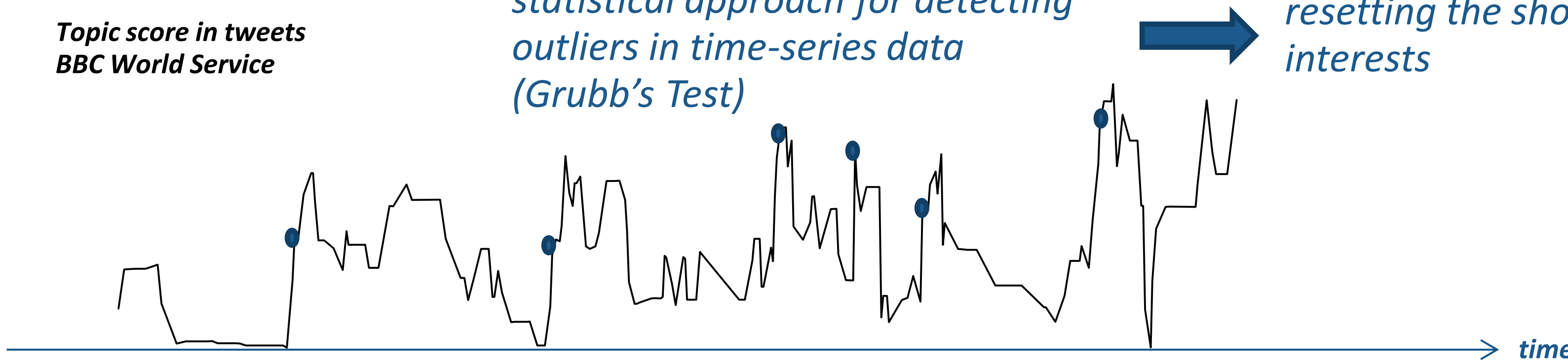
2. Daily adjustments: The tweets in the current calendar day

Adhoc

3. Event detection [5] to automatically identify times when events related to the topic occurred and reset the short-term interests accordingly

Detected events using a statistical approach for detecting outliers in time-series data (Grubb's Test)

Identified times for resetting the short term interests



Event detection can be applied on the Twitter stream itself or external news streams

[5] M. Albakour, C. Macdonald, I. Ounis. Identifying Local Events by Using Microblogs as Social Sensors. In Proc. of OAIR, 2013

Experiments: Drift

Identical setup to the one used before

The QE approach for handling sparsity as a baseline

The newswire stream

- BBC, CNN, Google News, New York Times, Guardian, Reuters, The Register and Wired

Research Questions:
(1) **Adhoc methods** vs. **event detection** for handling drift?

(3) **sensitivity** of the filtering performance to the **decay factor σ**?

Results: Spars

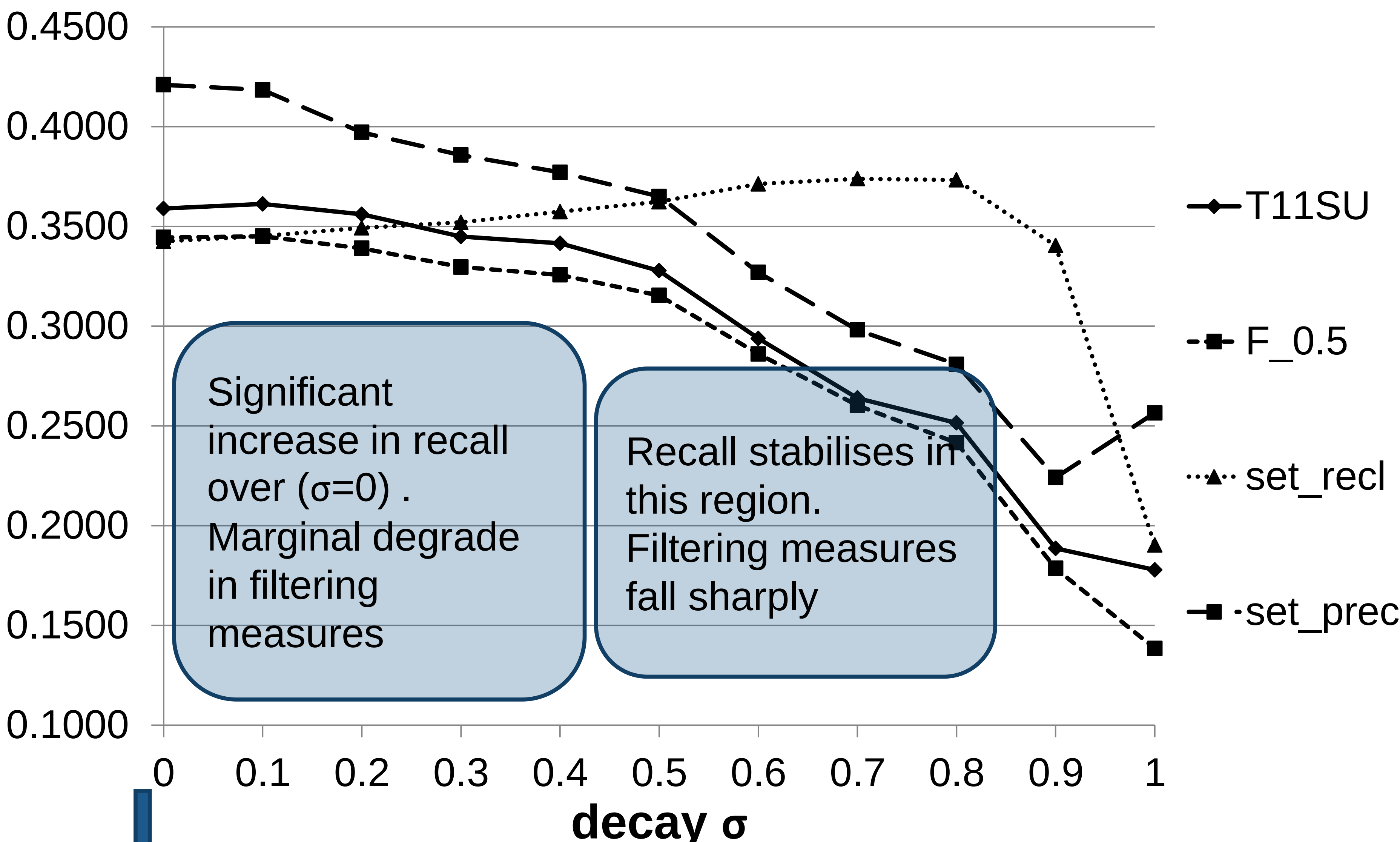
- ✗ Adhoc methods **failed**
- ✓ The recall is slightly improved.

- ✓ The **increase in recall** is **significant**.
- ✓ **Event detection** is helping!
- Differences are marginal when using a different stream for events! (Events overlap in both streams)

Bas RC+				
Arbi (n=1				
Dail (σ =				
Eve Twee (σ = 0.4)				
Event detection using News Streams (σ = 0.4)	0.3724 ▼	**0.3598** ▲▲	0.3198 ▼	0.3351 ▼

A single triangle means the differences are not statistically significant using a paired t-test at p<0.05. Double triangles mean the differences are statistically significant

Sensitivity to decay



Baseline: no difference between short-term and long-term

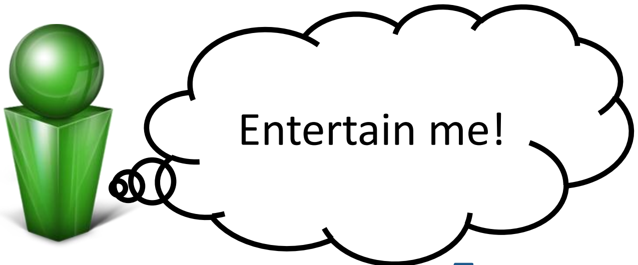
More emphasis on short-term interests

Conclusions

- Tackled sparsity and drift for real-time twitter filtering
- State-of-the-art for real-time twitter filtering
- With an event detection approach to tackle drift, we can significantly improve the filtering recall while only marginally harming the filtering utility

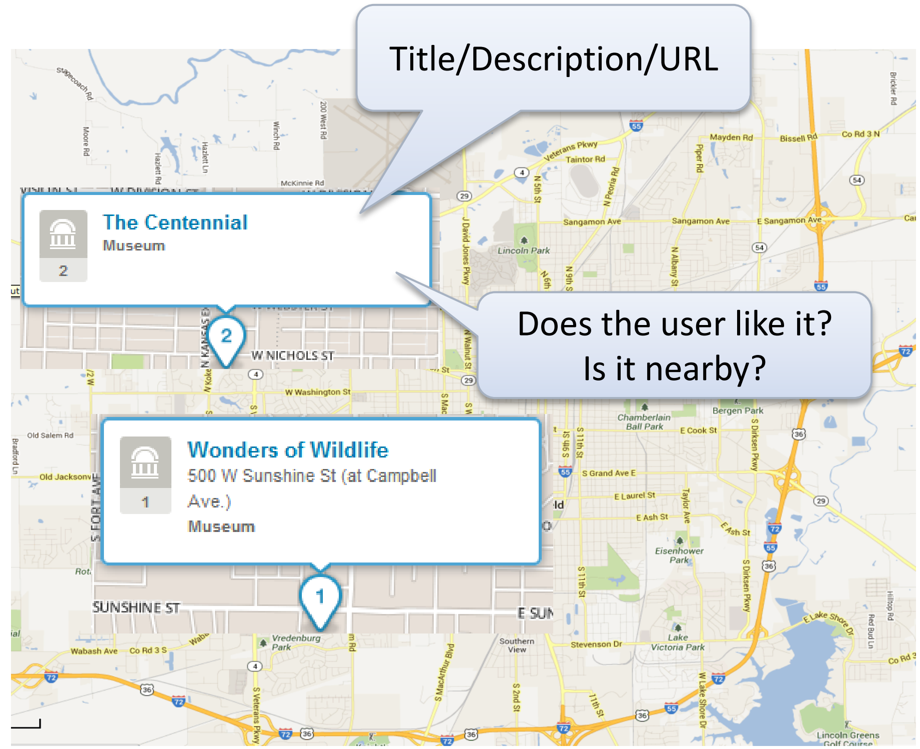
ANTICIPATION AND PERSONALISED VENUE RECOMMENDATION

Venue Recommendation



Zero-query

Location (Springfield) + time (2 pm) + ..



Venue recommendation has different potential use cases:

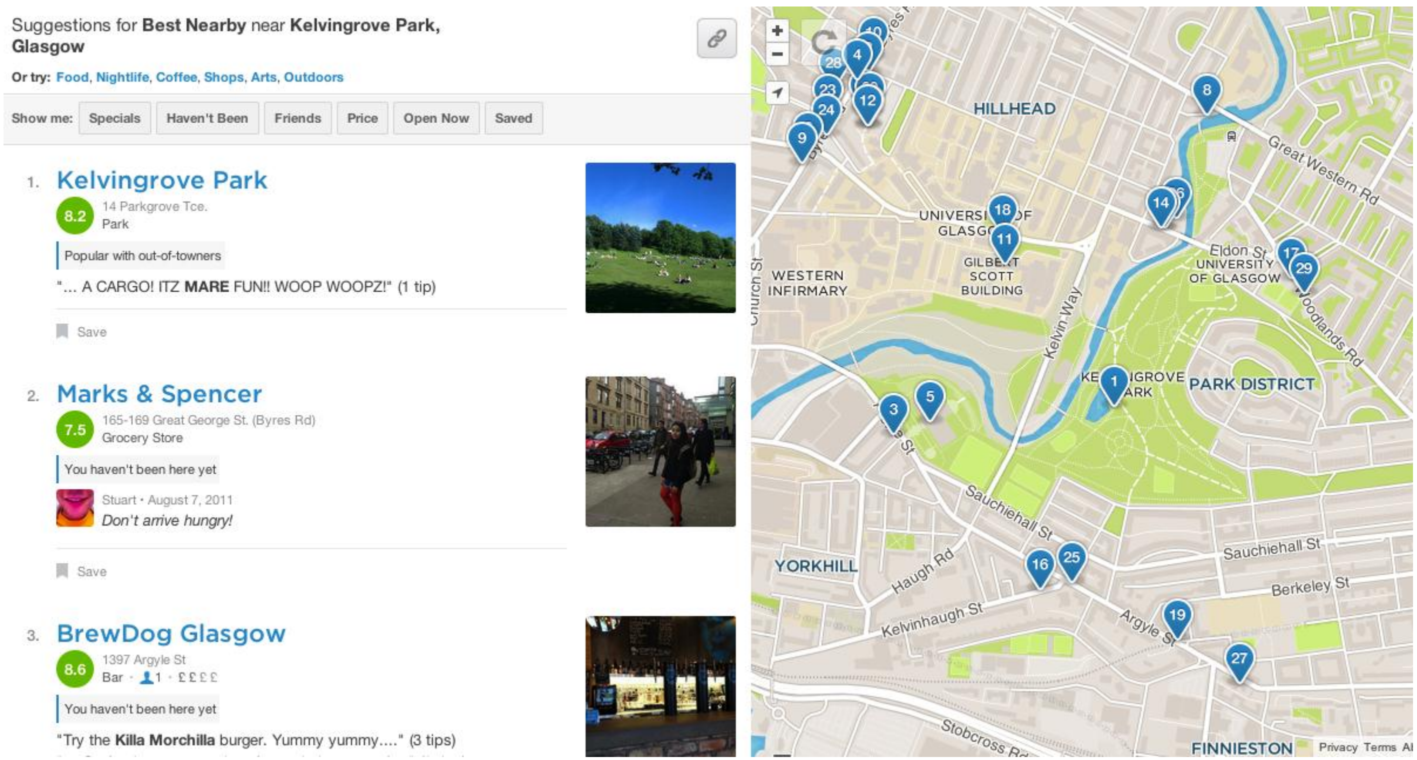
- Tourists use case: *"I have one day in this city, what should I see?"*
- Residents use case: discover/explore new venues, avoid noisy or polluted places, …

Existing Services

What do people currently use?

- A tourist guide, The List, Yelp, FourSquare?

No anticipation of venue popularity...



Recommendations from Foursquare at 10pm, in March

Challenges

Venue recommendation: help users decide where to go

"I'm new to the city. What should I visit?"

We argue that effective venue suggestions should encompass:

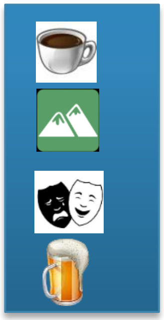
- Cold-start: we don't know where you have been before
- Personalised: recommend venues that I would **like**
- Time-aware: Quality venues will be **popular**

We developed and evaluated a probabilistic model for time-aware and personalised venue recommendation

Ranking Venues

Input

Output



Ranked list of venues

Location (Springfield)

foursquare yelp

P (| ,) ?

Not available!

Venue Popularity

Terrier Team | University of Glasgow

P ( | )

How busy a venue will be later in the near future (in the next few hours)

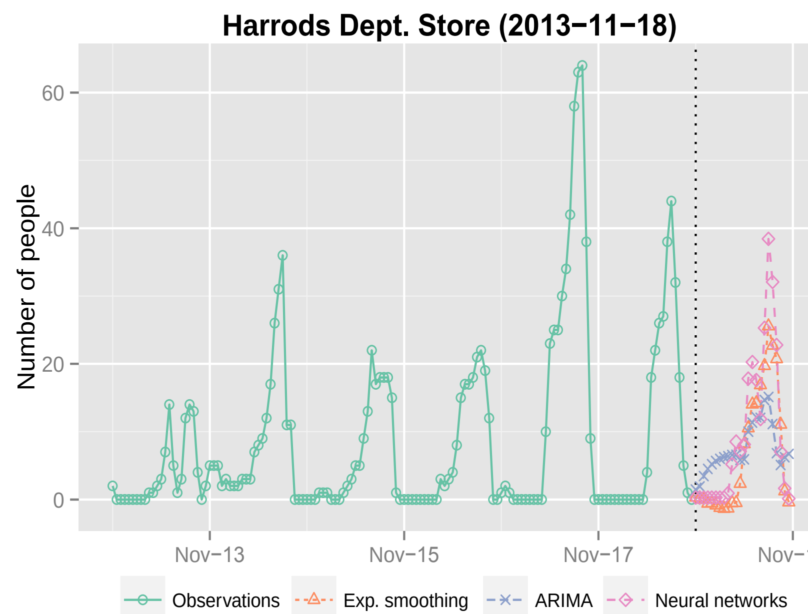
- we anticipate how **popular** the venue will be

Popularity – we forecast the attendance of venues based on past Foursquare checkins

– *Anticipating* the *future attendance*

– Foursquare API as a *social sensor* of the level of venue attendance (“check-ins”)

– time series forecasting models



Personalisation

Not available!

Personalisation

Not available!

Evaluation – venue popularity

Not available!

User Study

Not available!

User Study

Not available!

User Study

Not available!

Results of the User Study

Not available!

Thanks!

Acknowledgments

This work has been carried out in the scope of the EC co-funded project SMART (FP7-287583).

Co-authors:

Romain Deveaud, Craig Macdonald, Iadh Ounis

dyaa.albakour@glasgow.ac.uk