Thematic Timeline Generation

NGUYEN Kiem Hieu

School of Information & Communication Technology,
Hanoi Univ. of Science & Technology
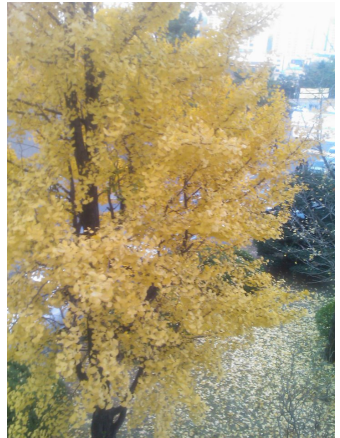1 rue Dai Co Viet, Bach Khoa, Hai Ba Trung, Hanoi
`http://is.hust.edu.vn/~hieunk`
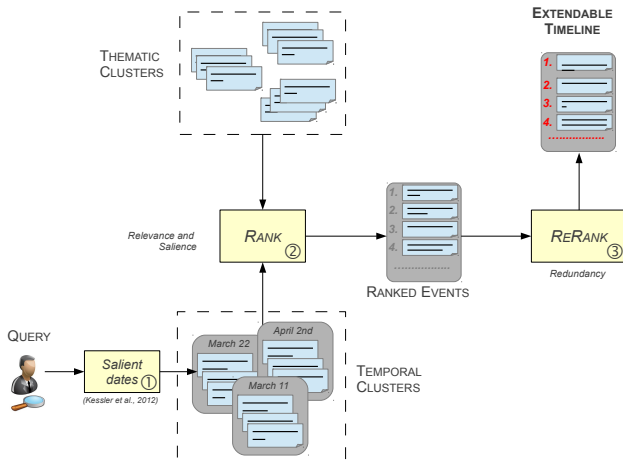hieunk@soict.hust.edu.vn

HAUI, Nov-2015

Outline

A pipeline framework for thematic timeline generation

KEYWORDS: discovery us space shuttle nasa

- **Jul 26**: Discovery blasts off on a mission to supply the International Space Station, the first space shuttle flight since Columbia broke up in February 2003, killing all seven aboard.
- **Jul 26**: NASA announces that a piece of insulating foam broke free of Discovery's external fuel tank on lift-off, possibly damaging tile on the orbiter's heat shield – the damage that triggered Columbia's break-up on re-entry.
- **Jul 27**: NASA grounds the remaining shuttle fleet: Endevour and Atlantis. The Challenger was lost in January 1986.
- **Jul 28**: NASA says that at least one shard of protective foam "might have struck" a wing of the Discovery during launch, but expressed confidence the craft would return safely to Earth. Discovery does a slow back flip before docking with the International Space Station, exposing its thermal shield to space station astronauts, who take pictures with digital cameras.
- **Jul 29**: Discovery Commander Eileen Collins says the crew is "surprised" and "disappointed" to find that NASA had not solved the foam debris problem.

...

Riyadh (AFP) - Saudi Arabia's King Abdullah bin Abdulaziz died on Friday and was replaced by Crown Prince Salman, the OPEC-kingpin's royal court said in a statement. The late monarch's half brother Moqren was named crown prince, according to the statement. King Abdullah, believed to be around 90 years old, was hospitalised in December suffering from pneumonia and had been breathing with the aid of a tube. He died on Friday "at 1:00 am (2200 GMT)" and would be buried later in the day following afternoon prayers, said the statement. In recent years, his advanced age and poor health had raised concerns about the future leadership of one of the world's key oil producers. Abdullah's half-brother Salman, 79, was named crown prince in June 2012 following the death of Prince Nayef bin Abdulaziz. Salman had been representing the king at most recent public events because of the monarch's poor health.

In March 2014, King Abdullah named his half-brother Prince Moqren as a second crown prince, in an unprecedented move aimed at smoothing succession hurdles. Moqren, who was born in 1945, is the youngest of Abdulaziz's sons. Since the death in 1952 of King Abdulaziz al-Saud, the founder of Saudi Arabia, the throne has systematically passed from one of his sons to another. But many of them are old or have died. The former crown princes Sultan and Nayef died in 2011 and 2012 respectively.

Detecting temporal expressions

Riyadh (AFP) - Saudi Arabia's King Abdullah bin Abdulaziz died on Friday and was replaced by Crown Prince Salman, the OPEC-kingpin's royal court said in a statement.

The late monarch's half brother Moqren was named crown prince, according to the statement. King Abdullah, believed to be around 90 years old, was hospitalised in December suffering from pneumonia and had been breathing with the aid of a tube. He died on Friday "at 1:00 am (22:00 GMT)" and would be buried later in the day following afternoon prayers, said the statement.

In recent years, his advanced age and poor health had raised concerns about the future leadership of one of the world's key oil producers. Abdullah's half-brother Salman, 79, was named crown prince in June 2012 following the death of Prince Nayef bin Abdulaziz. Salman had been representing the king at most recent public events because of the monarch's poor health.

In March 2014, King Abdullah named his half-brother Prince Moqren as a second crown prince, in an unprecedented move aimed at smoothing succession hurdles. Moqren, who was born in 1945, is the youngest of Abdulaziz's sons. Since the death in 1952 of King Abdulaziz al-Saud, the founder of Saudi Arabia, the throne has systematically passed from one of his sons to another. But many of them are old or have died. The former crown princes Sultan and Nayef died in 2011 and 2012 respectively.

Temporal normalization

Riyadh (AFP) - Saudi Arabia's King Abdullah bin Abdulaziz died on [mention='Friday' date='23-01-2015' time=''] and was replaced by Crown Prince Salman, the OPEC-kingpin's royal court said in a statement.

The late monarch's half brother Moqren was named crown prince, according to the statement. King Abdullah, believed to be around 90 years old, was hospitalised in [mention='December' date='xx-12-2014' time=''] suffering from pneumonia and had been breathing with the aid of a tube. He died on [mention='Friday "at 1:00 am (22:00 GMT)"' date='23-01-2015' time='22:00:00'] and would be buried later in [mention='the day following afternoon' date='24-01-2015' time='AFTERNOON'] prayers, said the statement.
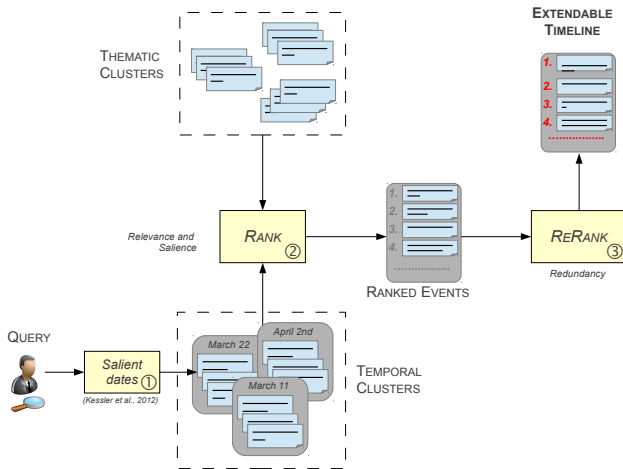
In [mention='recent years' date='' time=''], his advanced age and poor health had raised concerns about the future leadership of one of the world's key oil producers. Abdullah's half-brother Salman, 79, was named crown prince in [mention='June 2012' date='xx-06-2012' time=''] following the death of Prince Nayef bin Abdulaziz. Salman had been representing the king at most recent public events because of the monarch's poor health.

In [mention='March 2014' date='xx-03-2014' time=''], King Abdullah named his half-brother Prince Moqren as a second crown prince, in an unprecedented move aimed at smoothing succession hurdles. Moqren, who was born in [mention='1945' date='xx-xx-1945' time=''], is the youngest of Abdulaziz's sons. Since the death in [mention='1952' date='xx-xx-1952' time=''] of King Abdulaziz al-Saud, the founder of Saudi Arabia, the throne has systematically passed from one of his sons to another. But many of them are old or have died. The former crown princes Sultan and Nayef died in [mention='2011' date='xx-xx-2011' time=''] and [mention='2012' date='xx-xx-2012' time=''] respectively.

Text segmentation

- Riyadh (AFP) - Saudi Arabia's King Abdullah bin Abdulaziz died on [mention='Friday' date='23-01-2015' time=''] and was replaced by Crown Prince Salman, the OPEC-kingpin's royal court said in a statement.
- The late monarch's half brother Moqren was named crown prince, according to the statement.
- King Abdullah, believed to be around 90 years old, was hospitalised in [mention='December' date='xx-12-2014' time=''] suffering from pneumonia and had been breathing with the aid of a tube.
- He died on [mention='Friday "at 1:00 am (22:00 GMT)"' date='23-01-2015' time='22:00:00'] and would be buried later in [mention='the day following afternoon' date='24-01-2015' time='AFTERNOON'] prayers, said the statement.
- In [mention='recent years' date='' time=''], his advanced age and poor health had raised concerns about the future leadership of one of the world's key oil producers.
- Abdullah's half-brother Salman, 79, was named crown prince in [mention='June 2012' date='xx-06-2012' time=''] following the death of Prince Nayef bin Abdulaziz. Salman had been representing the king at most recent public events because of the monarch's poor health.
- In [mention='March 2014' date='xx-03-2014' time=''], King Abdullah named his half-brother Prince Moqren as a second crown prince, in an unprecedented move aimed at smoothing succession hurdles.
- Moqren, who was born in [mention='1945' date='xx-xx-1945' time=''], is the youngest of Abdulaziz's sons.
- Since the death in [mention='1952' date='xx-xx-1952' time=''] of King Abdulaziz al-Saud, the founder of Saudi Arabia, the throne has systematically passed from one of his sons to another.
- But many of them are old or have died.
- The former crown princes Sultan and Nayef died in [mention='2011' date='xx-xx-2011' time=''] and [mention='2012' date='xx-xx-2012' time=''] respectively.

- Riyadh (AFP) - Saudi Arabia's King Abdullah bin Abdulaziz died on [mention='Friday' date='23-01-2015' time=''] and was replaced by Crown Prince Salman, the OPEC-kingpin's royal court said in a statement.
- The late monarch's half brother Moqren was named crown prince, according to the statement.
- King Abdullah, believed to be around 90 years old, was hospitalised in [mention='December' date='xx-12-2014' time=''] suffering from pneumonia and had been breathing with the aid of a tube.
- He died on [mention='Friday "at 1:00 am (22:00 GMT)"' date='23-01-2015' time='22:00:00'] and would be buried later in [mention='the day following afternoon' date='24-01-2015' time='AFTERNOON'] prayers, said the statement.
- In [mention='recent years' date='' time=''], his advanced age and poor health had raised concerns about the future leadership of one of the world's key oil producers.
- Abdullah's half-brother Salman, 79, was named crown prince in [mention='June 2012' date='xx-06-2012' time=''] following the death of Prince Nayef bin Abdulaziz.
- Salman had been representing the king at most recent public events because of the monarch's poor health.
- In [mention='March 2014' date='xx-03-2014' time=''], King Abdullah named his half-brother Prince Moqren as a second crown prince, in an unprecedented move aimed at smoothing succession hurdles.
- Moqren, who was born in [mention='1945' date='xx-xx-1945' time=''], is the youngest of Abdulaziz's sons.
- Since the death in [mention='1952' date='xx-xx-1952' time=''] of King Abdulaziz al-Saud, the founder of Saudi Arabia, the throne has systematically passed from one of his sons to another.
- But many of them are old or have died.
- The former crown princes Sultan and Nayef died in [mention='2011' date='xx-xx-2011' time=''] and [mention='2012' date='xx-xx-2012' time=''] respectively.

Temporal clusters

- Riyadh (AFP) - Saudi Arabia's King Abdullah bin Abdulaziz died on [mention='Friday' date='23-01-2015' time=''] and was replaced by Crown Prince Salman, the OPEC-kingpin's royal court said in a statement.
- King Abdullah, believed to be around 90 years old, was hospitalised in [mention='December' date='xx-12-2014' time=''] suffering from pneumonia and had been breathing with the aid of a tube.
- He died on [mention='Friday "at 1:00 am (22:00 GMT)"' date='23-01-2015' time='22:00:00'] and would be buried later in [mention='the day following afternoon' date='24-01-2015' time='AFTERNOON'] prayers, said the statement.
- In [mention='recent years' date='' time=''], his advanced age and poor health had raised concerns about the future leadership of one of the world's key oil producers.
- Abdullah's half-brother Salman, 79, was named crown prince in [mention='June 2012' date='xx-06-2012' time=''] following the death of Prince Nayef bin Abdulaziz.
- In [mention='March 2014' date='xx-03-2014' time=''], King Abdullah named his half-brother Prince Moqren as a second crown prince, in an unprecedented move aimed at smoothing succession hurdles.
- Moqren, who was born in [mention='1945' date='xx-xx-1945' time=''], is the youngest of Abdulaziz's sons.
- Since the death in [mention='1952' date='xx-xx-1952' time=''] of King Abdulaziz al-Saud, the founder of Saudi Arabia, the throne has systematically passed from one of his sons to another.
- The former crown princes Sultan and Nayef died in [mention='2011' date='xx-xx-2011' time=''] and [mention='2012' date='xx-xx-2012' time=''] respectively.

Event detection

Riyadh (AFP) - Saudi Arabia's King Abdullah bin Abdulaziz [died]$_{EVENT}$ on [Friday]$_{23\text{-}01\text{-}2015}$ and [was replaced]$_{EVENT}$ by Crown Prince Salman, the OPEC-kingpin's royal court [said]$_{ATTRIBUTION}$ in a statement.

Event selection

Riyadh (AFP) - Saudi Arabia's King Abdullah bin Abdulaziz [died]$_{EVENT}$ **on** [Friday]$_{23\text{-}01\text{-}2015}$ and [was replaced]$_{EVENT}$ by Crown Prince Salman, the OPEC-kingpin's royal court [said]$_{ATTRIBUTION}$ in a statement.

- Riyadh (AFP) - Saudi Arabia's King Abdullah bin Abdulaziz [died]$_{EVENT}$ **on** [Friday]$_{23\text{-}01\text{-}2015}$ and was replaced by Crown Prince Salman, the OPEC-kingpin's royal court said in a statement.
- King Abdullah, believed to be around 90 years old, [was hospitalised]$_{EVENT}$ **in** [December]$_{xx\text{-}12\text{-}2014}$, suffering from pneumonia and had been breathing with the aid of a tube.
- He [died]$_{EVENT}$ **on** [Friday "at 1:00 am (22:00 GMT)"]$_{23\text{-}01\text{-}2015}$ and [would be buried]$_{EVENT}$ later **in** [the day following afternoon]$_{24\text{-}01\text{-}2015}$ prayers, said the statement.
- **In** [recent years]$_{xx\text{-}xx\text{-}xxx}$, his advanced age and poor health [had raised]$_{EVENT}$ concerns about the future leadership of one of the world's key oil producers.
- Abdullah's half-brother Salman, 79, [was named]$_{EVENT}$ crown prince **in** [June 2012]$_{xx\text{-}06\text{-}2012}$ following the death of Prince Nayef bin Abdulaziz.
- **In** [March 2014]$_{xx\text{-}03\text{-}2014}$, King Abdullah [named]$_{EVENT}$ his half-brother Prince Moqren as a second crown prince, in an unprecedented move aimed at smoothing succession hurdles.
- Moqren, who [was born]$_{EVENT}$ **in** [1945]$_{xx\text{-}xx\text{-}1945}$, is the youngest of Abdulaziz's sons.
- Since [the death]$_{EVENT}$ **in** [1952]$_{xx\text{-}xx\text{-}1952}$ of King Abdulaziz al-Saud, the founder of Saudi Arabia, the throne has systematically passed from one of his sons to another.
- The former crown princes Sultan and Nayef [died]$_{EVENT}$ **in** [2011]$_{xx\text{-}xx\text{-}2011}$ **and** [2012]$_{xx\text{-}xx\text{-}2012}$ respectively.

NASA Discovery mission failed

- Jul 26: Discovery blasts off on a mission to supply the International Space Station, the first space shuttle flight since Columbia broke up in February 2003, killing all seven aboard.
- Jul 26: NASA announces that a piece of insulating foam broke free of Discovery's external fuel tank on lift-off, possibly damaging tile on the orbiter's heat shield – the damage that triggered Columbia's break-up on re-entry.
- Jul 27: NASA grounds the remaining shuttle fleet: Endevour and Atlantis. The Challenger was lost in January 1986.
- Jul 28: NASA says that at least one shard of protective foam "might have struck" a wing of the Discovery during launch, but expressed confidence the craft would return safely to Earth. Discovery does a slow back flip before docking with the International Space Station, exposing its thermal shield to space station astronauts, who take pictures with digital cameras.
- Jul 29: Discovery Commander Eileen Collins says the crew is "surprised" and "disappointed" to find that NASA had not solved the foam debris problem.
- Jul 30: Discovery will remain docked at the International Space State a day longer than planned to allow the crew to do more work, NASA announces.
- Jul 31: NASA's head says the agency "goofed" on key safety checks.
- Aug 1: NASA announces a space-walk repair to remove two strands of gap filler dangling from the belly of the orbiter, over concerns that the ceramic fiber could trigger uneven heating of the shield on re-entry.
- Aug 2: Astronauts gather their tools for an unprecedented repair on the shuttle's thermal shield to ensure it can make a safe return to Earth.

26 July, 2005

- NASA said Wednesday it had set a new tentative launch date for the shuttle Discovery of July 26 at 10:44 am (14:44 GMT) from the Kennedy Space Center near Cape Canaveral in Florida.
- NASA closed Discovery 's hatch after the seven crew members were strapped in their seats Tuesday, less than two hours ahead of the space shuttle's planned launch.
- July 26: NASA announces that a piece of insulating foam broke free of Discovery's external fuel tank on lift-off, possibly damaging tile on the orbiter's heat shield–the damage that triggered Columbia's break-up on re-entry.
- NASA set Tuesday as the tentative launch date for the shuttle Discovery, after saying it was confident the technical glitch that delayed the original July 13 launch has been overcome.
- At the soonest, NASA said, Discovery could blast off on July 26.
- Weather conditions for a launch of Space Shuttle Discovery improved significantly, NASA said just hours ahead of Tuesday's scheduled liftoff.
- Discovery blasted off from Florida's Kennedy Space Center on Tuesday on the first shuttle mission since the 2003 Columbia disaster

- NASA said Wednesday it had set a new tenative launch date for the shuttle Discovery of July 26 at 10:44 am (14:44 GMT) from the Kennedy Space Center near Cape Canaveral in Florida.
- NASA closed Discovery 's hatch after the seven crew members were strapped in their seats Tuesday, less than two hours ahead of the space shuttle's planned launch.
- July 26: NASA announces that a piece of insulating foam broke free of Discovery's external fuel tank on lift-off, possibly damaging tile on the orbiter's heat shield–the damage that triggered Columbia's break-up on re-entry.
- NASA set Tuesday as the tentative launch date for the shuttle Discovery, after saying it was confident the technical glitch that delayed the original July 13 launch has been overcome.
- At the soonest, NASA said, Discovery could blast off on July 26.
- Weather conditions for a launch of Space Shuttle Discovery improved significantly, NASA said just hours ahead of Tuesday's scheduled liftoff.
- Discovery blasted off from Florida's Kennedy Space Center on Tuesday on the first shuttle mission since the 2003 Columbia disaster.

Date creation time (DCT) & Title

Title: 'URGENT NASA to launch shuttle Discovery July 26'
DCT: 20, July, 2005$_{BEFORE}$

NASA said$_{ATTRIBUTION}$ Wednesday it had set$_{ROOT}$ a new tenative launch date for the shuttle Discovery of July 26 at 10:44 am (14:44 GMT) from the Kennedy Space Center near Cape Canaveral in Florida.

Title: 'Discovery docks with space station , as NASA grounds shuttle fleet'

DCT: 28, July, 2005$_{SAME}$

Discovery blasted off$_{ROOT}$ from Florida's Kennedy Space Center on Tuesday on the first shuttle mission since the 2003 Columbia disaster.

Title: 'NASA launches shuttle "foam" investigation'

DCT: 10, August, 2005$_{AFTER}$

A day after Discovery safely returned to Earth following a troubled 14-day mission , the US space agency said$_{ATTRIBUTION}$ senior engineers were examining$_{ROOT}$ what caused pieces of insulation foam to break loose from the external fuel tank during the shuttle ' s launch on July 26.

Date salience

- ▶ Occuring frequently in the query-focused sub-corpus
- ▶ Ocurring not-so-frequently in the whole corpus
- ▶ DCT on the same date: Main events
- ▶ DCT after the date: Secondary events
- ▶ DCT before the date: The events in future
- ▶ Distance between DCT and the date: The closer the more relevant

Event importance

Salient date:

- ▶ Mentioning sentence similar to article title
- ▶ Event as root of sentence
- ▶ Verb tense
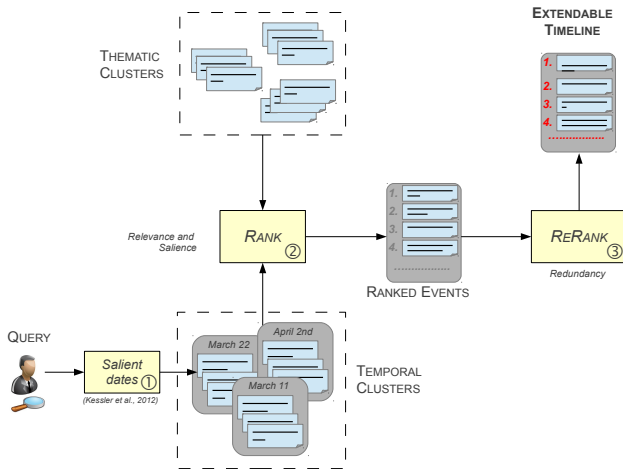- ▶ Indirect speech attribution

Salient Date Detection

Salient date detection:

- ▶ Input: dates
- ▶ Decision: salience
- ▶ Information: meta-data, textual contexts
- ▶ Golden references: timelines created by journalists

Technical approaches:

- ▶ Ad-hoc ranking
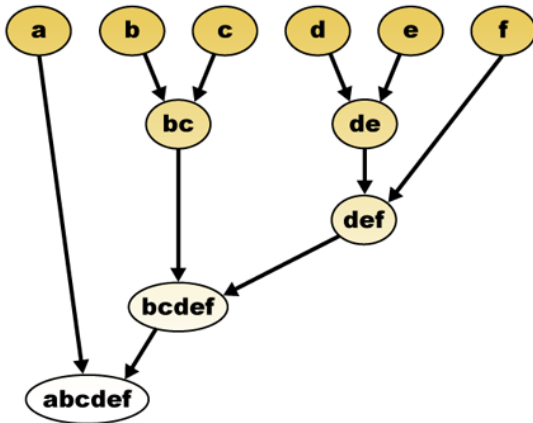- ▶ Classification
- ▶ Learning-to-rank

Text representation

- ▶ Vector space model
- ▶ Graphical structure
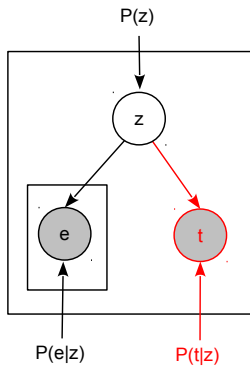- ▶ Graph-based representation

Vector space model

- Term frequency in salient dates
- Global term frequency
- Term normalization
- Linear combination of term relevance and date salience

Text clustering

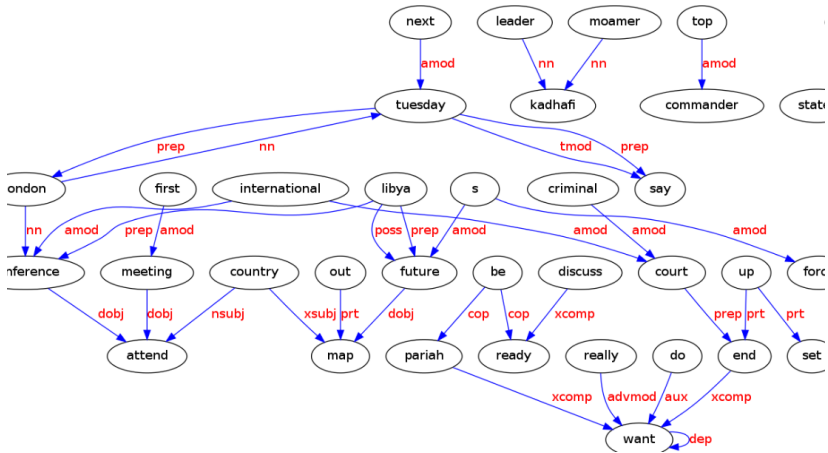# Graphical structure

- ▶ Probability-based presentation
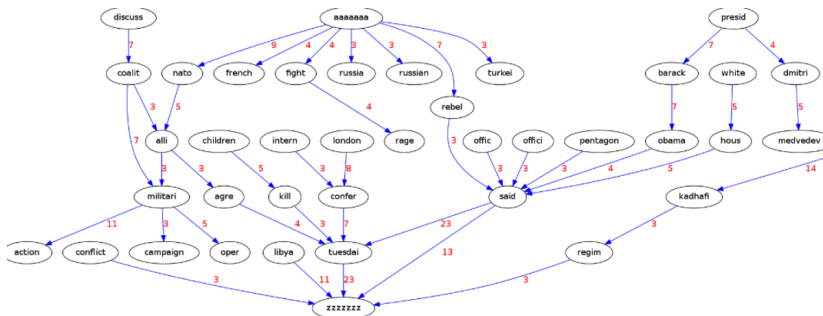- ▶ Dates as 'documents'
- ▶ Term filtering

Graph-based representation

Learning-to-rank

- ▶ Training data generation
- ▶ Meta-data features
- ▶ Temporal features
- ▶ Textual features

Iterative ranking

- Redundancy in texts
- Extendable timelines

Algorithm 1 Reranking algorithm

1: $out \leftarrow \phi$
2: **while** $(!terminate)$ **do**
3:     **for** $e \in S(q) \setminus out$ **do**
4:         $score(e|q)$
5:     **end for**
6:     $e^* = \underset{e \in S(q) \setminus out}{\mathrm{argmax}} \; score(e|q)$
7:     $out \leftarrow out \cup e^*$
8:     $d^* = d(e^*)$
9:     **for** $t \in e^*$ **do**
10:         $used(d^*) \leftarrow used(d^*) \cup t$
11:     **end for**
12: **end while**

Evaluation tasks

- ▶ Salient date detection
- ▶ Text summarization
- ▶ Timeline generation

Salient date detection

Metric

- ▶ Precision
- ▶ Recall
- ▶ F-score

System	MAP
DFIDF	71.46
ML	79.18
Rerank	77.83

Table: Comparison of salient date detection using MAP.

Text summarization

Metric

- ▶ Text summarisation evaluation metric
- ▶ K-length evaluation
- ▶ Precision, recall, F-score

System	P	R
DFIDF	27.24	25.50
ML	29.93	27.54
Rank	28.82	24.47
Rerank	31.23	26.63

Table: Comparison of MDS using ROUGE at 95% confidence interval.

Timeline generation
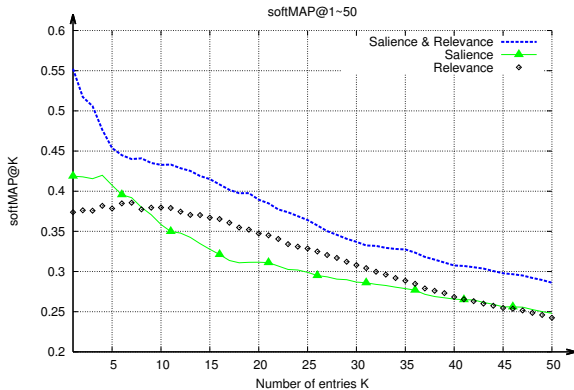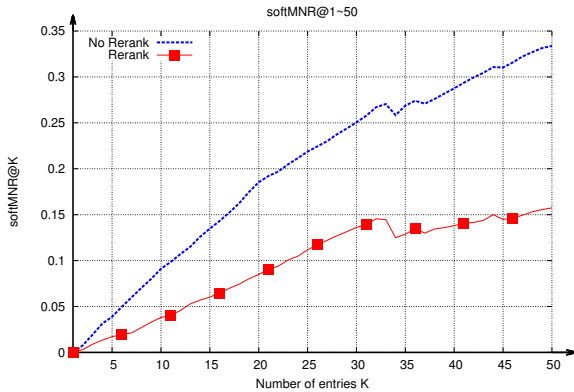
Outline

Introduction
Temporal Analysis
Event Detection
Timeline Generation
Evaluations
Scalability

Temporal-spatial domain

MH370: what we know

By Friday **14 March** data suggests MH370 may have been flown deliberately off course towards the Andaman Islands

18 March Australian government releases map showing two possible flight paths

27 March Estimated route travelling at 450 knots
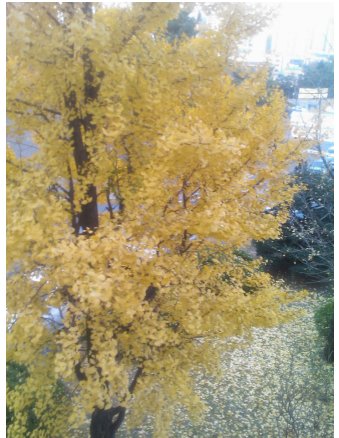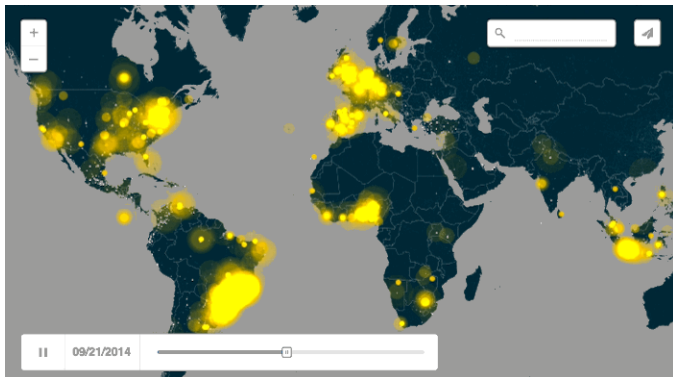
INDIAN OCEAN

Search areas to **23 March**

7 **18 March** Chinese report of debris

10

9

13 **13**

8

4 **3** **2**

10 March Search area, two days after disappearance

Communications turned off. Plane turns west

MALAYSIA

1 **8 March** Plane takes off from Kuala Lumpur airport, heading for Beijing

INDONESIA

5 **16 March** Satellite data shows plane was still in the air seven and a half hours after take off and its last recorded 'ping' was somewhere along a northern or southern arc

AUSTRALIA

27 March Estimated route travelling at 400 knots

21 March French report of debris

24 March Debris sighted

11

12 **27 March** Planned areas to be searched

6 **16 March** Debris sighted

1,224 miles

♦Perth

SOURCES: AUSTRALIAN GOVERNMENT, US NATIONAL TRANSPORTATION SAFETY BOARD, MALAYSIAN GOVERNMENT, NY TIMES

Temporal granularity

- ► Text genres.
- ► Meta-data
- ► Hour, date, month, year...

Multilingualism

- Web corpora
- Morphological analysis
- Temporal analysis