TREC 2013

Temporal Summarization

Jay Aslam
Fernando Diaz
Matthew Ekstrand-Abueg
Virgil Pavlu
Tetsuya Sakai
Example: 2011 Tōhoku Earthquake

Friday
2:46 PM  Magnitude 8.9 earthquake 231 miles northeast of Tokyo, Japan at a depth of 15.2 miles.
Quake is fifth largest in the world (since 1900) and the largest quake ever to hit Japan.

3:00 PM  Pacific Tsunami Warning Center issues tsunami warning for the Pacific Ocean from Japan to the U.S. west coast. Tsunami alerts sound in more than 50 countries and territories.

3:30 PM  Wall of water up to 30 feet high washes over the Japanese coast.

7:39 PM  Casualty reports begin to come in. Kyodo News Service reports at least 32 dead.

8:15 PM  Japanese government declares emergency for nuclear power plant near Sendai, 180 miles from Tokyo. Japan has 54 nuclear power plants.

9:35 PM  4 nuclear power plants closest to the quake are shut down.

10:29 PM  Cooling system at Fukushima nuclear report are reported not working: Authorities say they are “bracing for the worst”.

Saturday
12:39 AM  Fires are reported in at least three Japanese prefectures (Hakodate, Chiba, Miyagi). An oil refinery was ablaze near Tokyo.

12:42 AM  A dam breaks in Fukushima prefecture, washing away scores of homes.

12:59 AM  4 million homes in Tokyo and surrounding areas are without power.

1:07 AM  Kyodo News Service puts number of confirmed deaths at 137.

1:31 AM  Delta cancels 29 flights into and out of Tokyo. American Airlines diverts six flights en route to Tokyo to other airports.

1:57 AM  US Navy announces movement of seven ships toward Japan to assist relief efforts.

2:06 AM  Radiation level in Fukushima No. 1 nuclear power plant reported rising.

3:05 AM  President Obama says Japan's prime minister told him no evidence of radiation leaks from his country's nuclear power plants.

3:24 AM  Japanese trade minister Banri Kaieda says small radiation leak could occur at Fukushima nuclear plant.

3:34 AM  National Police up confirmed death toll to 151.

3:42 AM  Secretary of State Hillary Clinton announces US Air Force planes are headed to Japan carrying coolant for the Fukushima nuclear power plant. The report that the planes carried coolant is later found to be erroneous.
Motivation

• information access is **difficult** during unexpected news events (e.g. earthquakes, hurricanes).
  – sparse (minutes after the event)
  – redundant (hours after the event)
  – noisy (hours after the event)

• information access is **important** during unexpected news events.
  – urgency (especially for those close to the event)
  – concern (for family and friends)
Tasks

• **Sequential Update Summarization:** broadcast *useful, new, and timely* sentence-length updates about a developing event.

• **Value Tracking:** can track the value of important event-related attributes (e.g. number of fatalities, financial impact).
Track Goals

• to develop algorithms which detect sub-events with **low latency**.

• to develop algorithms which **minimize redundant information** in unexpected news events.

• to model information **reliability** in the presence of a dynamic corpus.

• to understand and address the sensitivity of text summarization algorithms in an online, sequential setting.

• to understand and address the sensitivity of information extraction algorithms in dynamic settings.
Task 1:
Sequential Update Summarization

• **corpus**: stream of documents
• **input**: tracking query, event onset time
• **output**: relevant, novel, and timely text updates
• **target**: gold standard, time-stamped updates
Task 1: Sequential Update Summarization

wikipedia revisions

output

input $d_0$ $d_1$ $d_2$ $d_3$ $d_4$ $d_5$ $d_6$

time
Corpus

• desired properties
  – timestamped documents
  – topically relevant
  – diverse

• approach
  – KBA2013
    • July 2012-January 2013
    • web, news, (twitter, facebook)
    • NLP annotations (e.g. segmentation, coref)
    • noisy timestamps (possibly ~1-2 hours late)
    • evaluation on `all sources’ and `twitter only’
Input

• desired properties
  – unexpected/sudden event (e.g. earthquake, hurricane, terrorist attack) with rough onset time.
    • KBA events focus on medium/small, longer term events
  – `easy’ to find subevents

• approach
  – ~10 large events occurring in timespan of corpus
  – <event onset time, keyword query>
  – <event onset time, first wikipedia revision>
2011 Tōhoku earthquake and tsunami

An earthquake occurred on 30 km (80 miles) E of Sendai, Honshu, Japan. The earthquake possible to create regional tsunami on the zone.

- USGSEvent ID usc0001xgp
  

- Integrated Tsunami Watcher Service [http://www.iibc.in/itws/](http://www.iibc.in/itws/)
Output

• desired properties: <time, sentenceid, docid>
  – short, natural language
  – support in corpus

• approach
  – timestamp of the system decision, *not necessarily the the source document*
  – id of sentence detected in the annotated corpus
  – support
    • id of supporting document(s)
Gold Standard Output

• desired properties
  – timestamped text `nugget’
  – standard method for determining importance
  – low latency wrt when nugget was known

• approach
  – nuggets semi-automatically derived from wikipedia revision history.
Evaluation

• desired properties
  – update must be relevant (~precision)
  – system must be comprehensive (~recall)
  – update must be novel
  – update must be timely

• approach
  – recall: fraction of Gold Standard updates that are matches by the system.
  – novelty: fraction of system updates which did not match the same Gold Standard update.
  – timeliness: difference between the system update time and the matched Gold Standard update time.
Research Topics

• Generalizability of previous algorithms
  – temporal summarization [Allan et al. 2001]
  – information filtering […]
  – TDT tracking […]
  – multidocument summarization […]

• Task-specific models
  – what features are important for the task?
  – what optimization objectives are effective for the task?

• Modeling reliability of information
  – is this source of information reliable?

• Algorithms for deferred decision-making
  – what is the tradeoff between timeliness and precision?
Miscellany

• no external data
  – requires time-synchronized external corpora
  – motivation for diverse corpora
Task 2: Value Tracking

• **corpus**: stream of documents
• **input**: tracking query, event onset time, attribute type
• **output**: running estimate of retrospective attribute value
• **target**: gold standard, retrospective attribute value
Task 2: Value Tracking

![Graph showing value tracking over time with blue dots for output and a red line for retrospective target.](image)
Corpus

• same as Task 1
Input

• desired properties
  – attributes with estimates mentioned in the corpus.
  – attributes existing across event types

• approach
  – ~10 large events shared with Task 1
  – attributes
    • fatalities
    • financial impact
  – <event onset time, keyword query, attribute type>
Output

• desired properties
  <time, estimate, docid>

• approach
  – estimate
    • extractive
    • generative
  – support
    • id of supporting document(s)
Gold Standard Output

• desired properties
  – retrospective true value

• approach
  – can be extracted from wikipedia infoboxes
Evaluation

• desired properties
  – update must be accurate
  – update must be timely

• approach
  – cumulative error rate from event onset to the end of the stream.
Research Topics

• Generalizability of extraction algorithms.
• Task-specific models
  – what features are important for the task?
• Modeling reliability of information
  – is this source of information reliable?
• Algorithms for deferred decision-making
  – what is the tradeoff between timeliness and precision?
Miscellany

• no external data
  – requires time-synchronized external corpora
  – motivation for diverse corpora
Schedule

• March 2013:
  – Guidelines and Tasks fully formulated
  – Sample events/values/queries and their gold wikipedia pages released. Participants can get an idea what the track would be like
• June 2013: Test events/values/queries released
• September 2013: runs due
• November 2013: evaluations due
• November 2013: TREC Conference