MONITORING MOVING OBJECTS USING UNCERTAIN WEB DATA

Mouhamadou Lamine Ba, Sebastien Montenez, Talel Abdessalem, Pierre Senellart
Institut Mines–Télécom ; Télécom ParisTech ; CNRS LTCI (Paris, France)
first.last@telecom-paristech.fr

Moving objects and the World Wide Web

Consider the problem of tracking real-world *moving objects* such as


cars


ships


celebrities


group of humans


cyclones

...

Main characteristics of moving objects

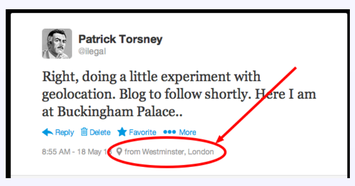
- Timestamped location data, e.g., countries visited by Barack Obama in 2014 and the time periods
- Other meta-information such as name, size, maximum reachable speed, acceleration patterns, etc.

Purpose of analysis, mining, and tracking tasks

- Pattern discovery or prediction of trajectories, and locations
- Better understanding of certain natural phenomena such as epidemics propagation
- Improvement of city services, regulation of route traffic, etc.


Currently used methods for tracking moving objects are often complex, mostly rely on application-specific resources and costly equipment (e.g., satellite or radar tracking of ships and aircrafts)

The Web is a wide source of public geo-tagged items and general information about various real-world moving objects



- Geo-tagged items (e.g., posts, tweets, or pictures) about various moving objects pollute social media like Twitter, Facebook, Flickr, ect.
- General Information about moving objects like celebrities, flights and ships can often be found online, e.g., on Wikipedia or Yellow Pages services

This paper aims at demonstrating

- The inference of object location and trajectories based on social Web data with many imprecision and inconsistencies
- An estimation of users' truthworthiness with respect to the precision of the location information they share on social media
- The gathering of more complete and accurate general data about various moving objects by the integration of different Web sources
- **MARITIME TRAFFIC APPLICATION** which is a system enabling to track various kinds of ships based on uncertain Web information

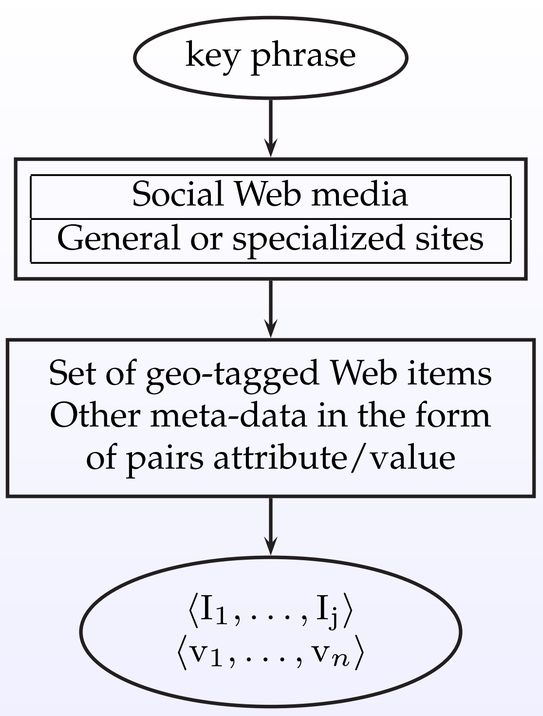
Data Extraction Process

INPUT
(e.g., ship name)

SEARCH & EXTRACTION
(APIs, Crawlers, Parsers)

POST-PROCESSING
(Filtering of Web items,
Mapping of attributes)

OUTPUT
(Spatio-temporal data,
Meta-information)



Uncertainty Estimation

Precision of location data. Evaluate the precision of a location against three criteria, and then mark it as less precise (LP) or more precise (MP)

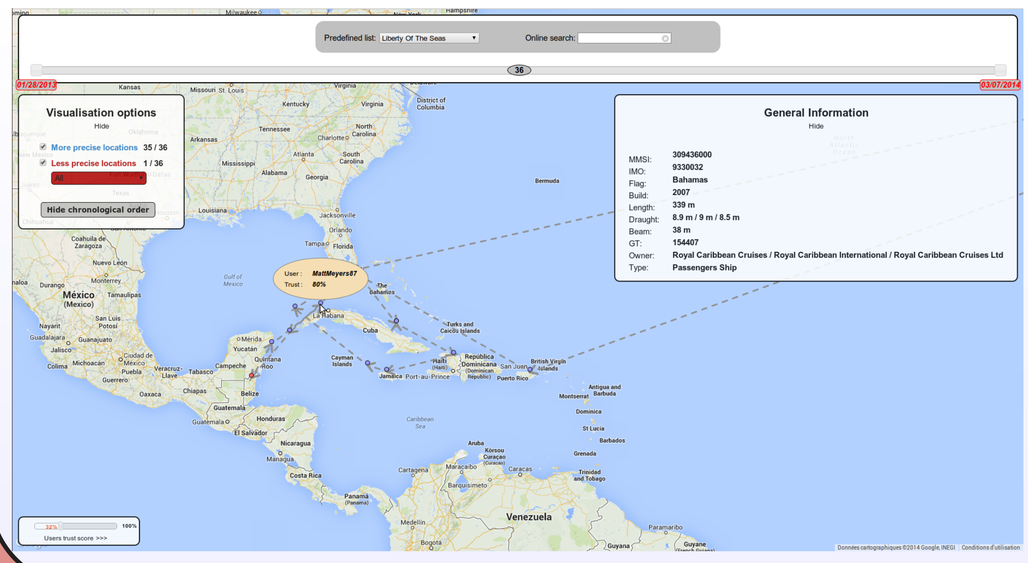
- **Outliers** (Out) w.r.t. a time window T and a maximum distance K
- **On-land locations** (Land) w.r.t. all the water areas on the Earth
- **Far-fetched trajectories** (Far) w.r.t. a maximum reachable speed V

User truthworthiness in sharing media. Compute as the proportion of items with good locations regarding the entire set of shared items about the same moving object weighted by some priori correctness scores

$$Trust(U) = \frac{\alpha \times |MP(I)| + \beta \times |Land(I)| + \gamma \times |Out(I)| + \theta \times |Far(I)|}{|MP(I)| + |Land(I)| + |Out(I)| + |Far(I)|}$$

Correctness of attribute values. Integration of values from different Web sources and estimation of their correctness by majority voting

MARITIME TRAFFIC APPLICATION



Used Web sources. Geo-located pictures from Flickr for ship locations – GrossTonnage, MarineTraffic, ShipSpotting, ShippingExplorer, and Wikipedia for ship features

Main features of the application

- Inference of ship locations and hypothetical itineraries for given time intervals
- Detection and filtering of more vs. less precise locations of given ships, and trusted users
- Integration of multiple sources for a more complete and accurate insight about ship characteristics