datAcron

Big Data Analytics for Time Critical Mobility Forecasting

Elias Alevizos, Alexander Artikis and George Paliouras

January 20, 2017

Project Concept

datAcron

Vision

- ▶ Management and integration of
    - ▶ voluminous data
    - ▶ heterogeneous data-at-rest (archival data)
    - ▶ data-in-motion (streaming data) sources
- ▶ Goal
    - ▶ safety and effectiveness of critical operations for large numbers of moving entities in large geographical areas
- ▶ use cases
    - ▶ maritime
    - ▶ aviation

Partners

- University of Piraeus (*Coordinator*)
- Fraunhofer Gesellschaft zur Forderung der Angewandten Forschung (Germany)
- National Centre for Scientific Research "Demokritos"
- Ecole Navale Groupement Interet Public (France)
- NATO Science and Technology Organisation (Belgium)
- Boeing Research & Technology Europe S.L.U. (Spain)
- Centro de Referencia Investigacion Desarrolo e Innovacion ATM, A.I.E. (Spain)
- IMIS Global Limited (United Kingdom)

Role of NCSR

- Adaptive Complex Event Recognition
  - Automatically generate and refine event definitions from data
  - Supervised machine learning
  - In-stream learning, aiming at creating adaptive models over spatial complex events
- Robust Complex Event Recognition
  - Integrate state-of-the-art recognition systems with probabilistic approaches
  - Distributed probabilistic reasoning
- Complex Event Forecasting
  - Support "forward" recognition

Current Status

- Project in its second year
- Domain understanding and data preparation
- Challenges
    - Multiple data sources
    - ... that need to be combined
    - ... which means patterns are context-context-context dependent
    - vague descriptions from use case partners
    - (they also need to better understand their domain)
    - ... which means tons of requirements analysis
    - lack of ground truth
- Work on forecasting (with partial matches)

Research Directions

- Complex Event Forecasting
    - Both use case partners (maritime, aviation) very interested in forecasting
    - But, not much research in relational forecasting (time-series forecasting does not count!)
    - We want to (semi-)automatically extract the conditions that predict the occurrence of an event

Research Directions

- Robust Complex Event Recognition
  - Data sources very noisy (e.g., in the maritime case, up to 35% of messages from vessels contain errors)
  - We need a probabilistic framework
  - Messages may arrive with long delays (need to be able to revise)
  - Same information may arrive from multiple sources (e.g., radars with overlapping coverage)
  - All the above mean: high complexity
  - We need a distributed probabilistic system (very few systems have done this)

Research Directions

- Adaptive Complex Event Recognition
    - Even use case partners might not be able to provide pattern definitions
    - More important: lack of ground truth (and sometimes even lack of "raw" messages)
    - Maybe unsupervised learning after all?
    - Patterns may evolve (e.g., trafficking vessels changing routes)
    - Need to adapt online (we are ambitious indeed!)