Trajectory Detection and Summarization over Surveillance Data Streams

Kostas Patroumpas, Athena Research Center, e-mail: kpatro@athenarc.gr
Eva Chondrodima, University of Piraeus, e-mail: evachon@unipi.gr
Nikos Pelekis, University of Piraeus, e-mail: npelekis@unipi.gr
Yannis Theodoridis, University of Piraeus, e-mail: ytheod@unipi.gr

Motivating points

Data sources available in maritime and aviation surveillance provide multiple, heterogeneous, voluminous, fluctuating, and noisy *data streams* from a large number of vessels and aircraft moving over large geographical areas. This kind of *big mobility data* is important in mission-critical applications not only for *real-time monitoring* of the whereabouts of these moving objects at sea and in the sky, but also for tracking their *evolving trajectories*. In such scenarios, trajectories must be constructed from online surveillance data: terrestrial or satellite AIS messages from vessels, and ADS-B messages from aircraft or tracklogs from ATC radars. Keeping in pace with the continuously accumulating data (huge *Volume*), the frequent updates (*high Velocity*), and the various types of collected messages (large *Variety*) is a very challenging task.

Towards this goal, in this chapter we present **Synopses Generator**, an online data processing framework that provides single-pass techniques for succinct, lightweight representation of trajectories without harming the quality of the resulting approximation. Instead of retaining every incoming position for each object, we propose to drop any predictable positions along trajectory segments of "normal" motion characteristics, while also applying additional noise filters. Except for adverse weather conditions, traffic regulations, local manoeuvres close to ports and airports, congestion situations, accidents, etc., most vessels and aircraft normally follow almost straight, predictable routes at open sea and in the air, respectively. It turns out that a large amount of raw positional updates can be suppressed with minimal losses in accuracy, as they hardly contribute additional knowledge about their actual motion patterns. Instead of resorting to a costly trajectory simplification algorithm, we opt to reconstruct their traces approximately from judiciously chosen *critical points* along their trajectories.

Summary of our approach

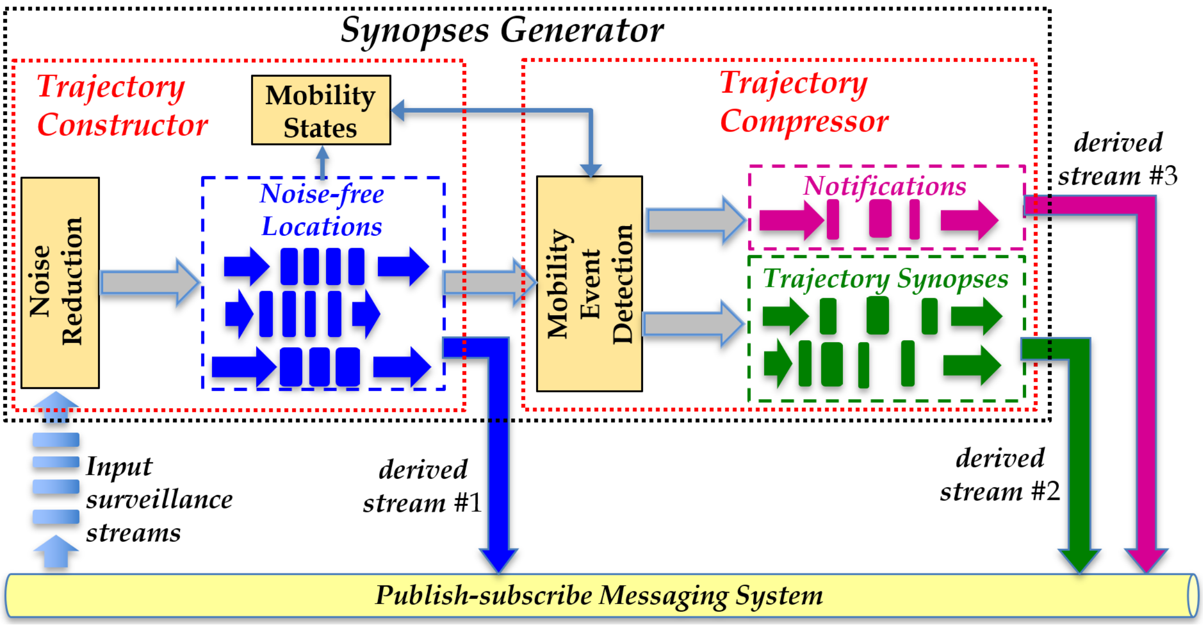

Figure 1: Online processing flow of the Synopses Generator framework

The proposed Synopses Generator consists of two main modules, as illustrated in Figure 1. The first one, called ***Trajectory Constructor,*** maintains distinct sequences of timestamped positions per moving object. This process also involves discarding any inherent *noise* detected in the streaming positions due to e.g., delayed arrival of input messages, duplicate positions, crosswind or sea drift, discrepancies in GPS measurements, etc. The second module, called ***Trajectory Compressor***, drops any predictable positions along "normal" trajectory segments. As exemplified with the examples in Figure 2 (maritime use case) and Figure 3 (aviation use case), effectively this framework keeps only those critical points conveying salient *mobility events* per monitored object (vessel or aircraft):

- *Stop* indicates that an object remains stationary (i.e., not moving) by checking whether its instantaneous speed is lower than a threshold over a period of time.
- *Slow motion* means that an object consistently moves at low speed over a period of time.
- *Change in Heading*: Once there is an angle difference in heading greater than a given threshold with respect to the mean velocity vector (computed over the most recent course of the moving object), the current location should be emitted as critical.
- *Speed change*: Such critical points are issued once the rate of change for speed exceeds a given threshold with respect to the mean speed of the object over a recent time interval.
- *Communication gaps* occur when an object has not emitted a message over a time period, e.g., the past 10 minutes.

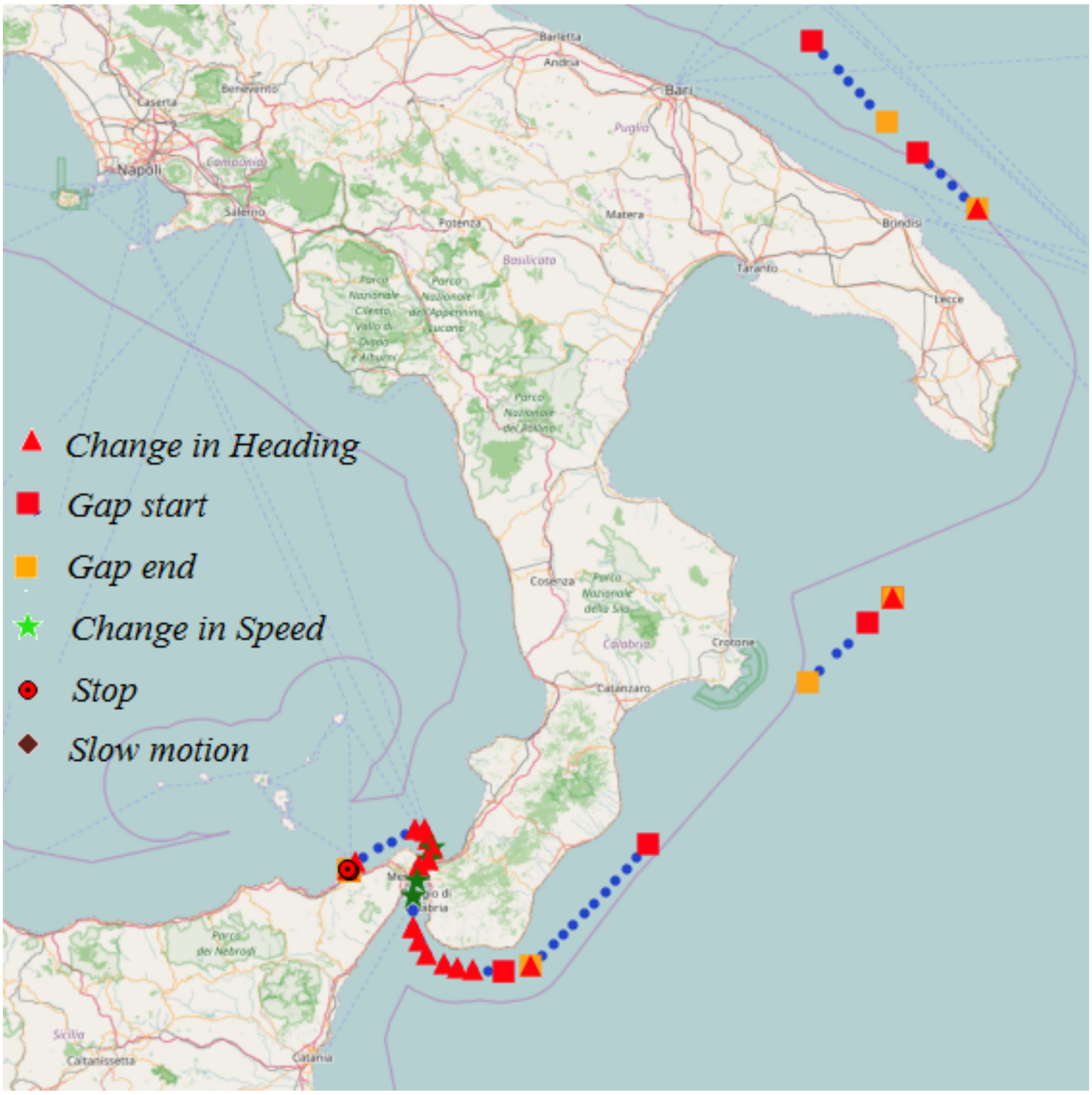

Figure 2: Types of critical points constructing the trajectory synopsis of a vessel

Figure 3: Types of critical points constructing the trajectory synopsis of a flying aircraft

- *Change in Altitude* may be detected for aircraft by checking their rate of climb (or descent), i.e., the vertical speed of the aircraft (in feet/sec) when ascending (respectively, descending). Once, this value exceeds a given threshold, a critical point should be issued in the synopsis.
- *Takeoff* is the latest location of an aircraft while still on the ground, as its next location reports an altitude above ground.
- *Landing* for a flying aircraft is the first reported location when it touches the ground.

Such mobility events are identified when the pattern of movement for a given object changes significantly. This derived stream of ***trajectory synopses*** must keep in pace with the incoming raw streaming surveillance data so as to get incrementally *annotated* with semantically important mobility features once they get detected. Thanks to the computation of trajectory synopses, object traces remain lightweight for efficient real-time processing without sacrificing accuracy, whereas they can be actually compared to each other irrespectively of the actual reporting frequency that may differ among objects.

It should be noted that the entire summarization process is working in a *stream-in-stream-out* fashion, i.e., consuming position updates arriving at varying frequencies from numerous objects and producing a derived stream distinct subsequences of (expectedly noise-free) timestamped positions as their evolving trajectory. Most importantly, this process offers in real-time a summarized representation of each trajectory consisting of critical points only, once they get detected from the noise-free positions retained from each object. Each critical point in such trajectory synopses is available in a *streaming fashion* to other applications. Moreover, the results of such trajectory summarization can progressively become *data-at-rest*, e.g., in archived in files, databases, RDF repositories, etc. They can be also semantically linked and contextually enriched with other data (e.g., weather, areas of interest), eventually resulting to *semantic trajectories* available for analytics, querying, forecasting, mining, etc.

Moreover, we implemented a ***software prototype*** application to handle this trajectory detections and summarization using the DataStream API of *Apache Flink* with *Apache Kafka* as broker for the streaming messages. Since these platforms lack native support for spatial and spatiotemporal operations, we introduced our custom data structures for maintaining trajectory segments with support for all mobility operations and functions required in the processing flow of the Synopses Generator. The source code of this prototype has been made publicly available[1].

Last, but not least, another major objective of our work concerns the issue of ***cross-stream processing*** of surveillance data by reconciling and aligning parts of synopses from different sources into a unified streaming summary. In particular, cross-stream processing of ADS-B data can be seen as a means of *"filling-in" gaps* in trajectory representations with information obtained from other sources, despite differences in object identification schemes, reporting frequencies, conflicting coordinates, etc. This real-time integration of data from different sources effectively allows the correlation of data from multiple sources in order to provide a coherent trajectory representation.

Contributions

To the best of our knowledge, this is the first *trajectory-aware* application framework specifically tailored for real-time surveillance over noisy, intermittent, *streaming mobility data* in the aviation and maritime domains. In particular, the proposed framework addresses the following challenges:

- ***Timeliness***. Trajectory detection and summarization can be carried out in real-time. Critical points concerning evolving trajectories of vessels and aircraft are reported at operational latency (ideally within milliseconds, or at most a few seconds) in order to enable immediate action, if necessary.
- ***Compression***. Since trajectory synopses per moving object are extracted from the incoming positions by retaining salient movement features only, this online data reduction can yield huge space savings. Empirical results indicate that at least 70-80% of the raw data may be discarded as redundant, while compression ratio can be up to 99% when frequency of position updates is high.
- ***Quality***. Such compressed representations are also reliable enough in reconstructing trajectories with small deviations (i.e., tolerable approximation error) from original traces, also coping with imperfections (such as network delays, noise, etc.) inherent in real-world surveillance streams.
- ***Scalability***. The Synopses Generator framework can be deployed in both centralized and distributed infrastructures and can manage scalable volumes of frequently updated, streaming positions from large fleets of vessels or aircraft moving over a large area.
- ***Cross-stream processing***. We introduce a methodology to reconcile and align parts of trajectories from different sources into a unified representation and thus “fill in” trajectories with missing points from a given source. A proof-of-concept evaluation offers very promising results in terms of quality.

[1] https://github.com/DataStories-UniPi/Trajectory-Synopses-Generator