Data Analytics and Machine Learning based on Trajectories

Felix Opitz, Kaeye Dästner, Bastian von Haßler zu Roseneckh-Köhler, Elke Schmid
Airbus
Wörthstr.85, 89077 Ulm
GERMANY

felix.opitz@airbus.com,
kaeye.daestner@airbus.com,
bastian.vonhasslerzuroseneckh-koehler@airbus.com,
elke.schmid@airbus.com

ABSTRACT

Modern surveillance networks are able to provide trajectories of all kind of vessels and aircraft worldwide or at least within extended environment. Best known are Automatic Dependent Surveillance – Broadcast (ADS-B) and (Satellite-) Automatic Identification System (AIS) used in air and maritime surveillance. Both of them are cooperative systems. It is foreseeable that ongoing trends like Internet of Things (IoT), digitalisation, automotive, smart cities and decentralisation (blockchain) will enable additional systems in the near future – e.g. in ground surveillance. Besides these systems sensor networks based on ground installations or mounted on airborne and space-based platforms deliver object trajectories independent of any cooperation. Examples include GMTI radar-based systems operating on UAV platforms or optical systems based on high altitude pseudo satellites (HAPS). These surveillance systems enable the extraction of mid- and long-term trajectories of any kind of objects. The real challenge will be the related situational awareness and the estimation of the intent of the tracked objects. Activity-based intelligence and the determination of patterns of life are a significant challenge for new systems. Otherwise, even modern surveillance systems are not able to take a real advantage of the gathered data. Data Analytics and Machine Learning are the most disruptive technologies to handle such massive data problems. To address trajectories various techniques are useful: Trajectories are partitioned into specific segments of interest using cluster algorithms. This help to decode their pattern of life based on unsupervised machine learning. There are dedicated metrics, which are used to compare trajectories as geometric objects. Or topological data analytics is applied to classify trajectories embedded in their specific environments. Trajectories can be sorted in different classes with dedicated representatives which ends up in statistical graphs and Markov transition models. This allows predictive analytics and the identification of anomalous behaviour. On the other hand, transponder and broadcast systems provide additional attributes of the tracked trajectories. This opens the whole area of supervised machine learning. The derived predictors realise the determination of object types and activities based on their trajectories. Finally, these new data analytic techniques have to be integrated in existing near real time surveillance systems. This requires specific system architectures as well as a completely new software and hardware landscape. So, trajectory-based Machine Learning is embedded on local or global clouds and uses dedicated mechanisms for distributed and parallel processing.

1 TRAJECTORIES AND THEIR SOURCES

A trajectory of an object like an aircraft, maritime vessel or a car is defined as a sequence

$$P = (p_1, \dots, p_n)$$

of positions p_i. According to the object's environment these are latitude, longitude and altitude. It is assumed that these positions are in chronological order. Trajectories may be gathered by different sensors or receivers. We distinguish between trajectories coming from cooperative and uncooperative sources.

1.1 Sources of Trajectories

Cooperative sources include Automatic Dependent Surveillance-Broadcast (ADS-B) and Automatic Identification Systems (AIS). ADS-B has become indispensable in aviation in the last ten years [1], [2], [3], [4]. It is nowadays essential for air traffic control (ATC) and even for air safety since ADS-B can support Traffic Alert and Collision Avoiding systems (TCAS) as well [5], [6], [7]. Dense distribution of ADS-B receivers worldwide ensures a continuous coverage over land but results also in huge data traffic. Several ADS-B data providers, e.g. Flightradar24 [8], ADS-B Exchange [9], FlightAware [10], etc., have started to store this data and offer them for commercial use. The data contains the aircraft identity by the ID of the International Civil Aviation Organization (ICAO) [11], the radio call sign, the flight number but also the aircraft location and kinematics for a given time, e.g. course and speed over ground, flight level and climb or sink rate. The location is based on GPS, while the dynamics are calculated from platform sensors.

AIS is an automatic tracking system used on ships and by vessel traffic services (VTS) [12], [13]. AIS is required by the International Maritime Organization's International Convention for the Safety of Life at Sea for international voyaging ships with 300 or more gross tonnage and all passenger ships regardless of size. The most important use case is collision avoidance for water transport. Vessels fitted with AIS transceivers can be tracked by AIS base stations located along coast lines or satellite based AIS receivers (S-AIS). AIS equipped ships have an assigned mobile service identity (MMSI) to guarantee their unique identification. The position information is also based on GPS receiver complemented by additional electronic navigation sensors, such as a gyrocompass or rate of turn indicator.

Both, ADS-B and AIS can deliver long term trajectories because of the inherent association information based on ICAO or MMSI codes. This makes them very valuable for statistics including data science and machine learning.

However, data received from cooperative tracking has gaps, since ADS-B and AIS transmitters are not always mandatory. Additionally, military aircraft have been observed to switch off ADS-B transmission during certain missions. Small boots are often not equipped with AIS. Finally, vessels switch off the transponders during illegal operations or perform spoofing.

Primary radar in ATC or coastal surveillance or airborne Ground Moving Target Tracking (GMTI) radar do not relay on the cooperation with the aircrafts or maritime vessels. They are also able to deliver extended trajectories as long as the targets are continuously tracked. The partition of all the measured object positions into trajectories requires dedicated tracking systems [14], [15], [16]. In contrast to cooperative systems, a trajectory may be split in different parts if the tracking continuity is interrupted.

1.2 Processing Architectures for Trajectories

The processing of whole trajectories also implies a switch in classical real-time surveillance systems. Their focus was up to now the detection and tracking of the objects which fly or move on land or sea in real time.

Data analysis and Machine learning based on temporary extended trajectories uses a big data architecture – the so called λ-architecture. The data processing is spread out in three different layers [17], [18]:

- Batch Layer
- Serving Layer
- Speed Layer

The batch layer continuously receives and stores the raw trajectories. This may be on a distributed filesystem and/or database, e.g. Hadoop HDFS or Cassandra [19]. Here, the trajectories have to be cleaned, transformed and presented in a way that different applications can access needed information quickly [20].

According to the interest of the user, trajectories have to be processed with methods related to data analytics and supervised or unsupervised machine learning. E.g. trajectories can be used to generate pure statistics – like heat maps. Or their pattern of life can be analysed and visited areas or covered routes can be extracted. If trajectories are attributed with further information related to their vehicle type or activity they can be used to train classifiers or predictors. This is supported by well-established tools like Spark [21] and a broad spectrum of machine learning libraries [22], [23], which enables processing of large data volumes by distributed processing. All this is localised in the serving layer and does not necessarily happen in real time.

The real time aspects are considered in the third layer – the speed layer. Here, the learned patterns and classifiers of the serving layer can be applied to the ongoing data stream to deliver real-time anomaly detection, predictive estimations, classification or even identification. For non-cooperative targets the speed layer possesses a tracking functionality, which can be used to aggregate single measurements to trajectories and to feed the batch layer.

Using this architecture allows a smooth integration of the big data and machine learning environment into already established and proven real time surveillance systems [18], [34]. Dependent on the systems requirements, a real time system can be enriched with either batch- and serving layer or an extraction of the serving layer, containing trained models and batch views only. The latter approach enables theses functionalities also for systems used in the field, which mostly have hardware limitations.

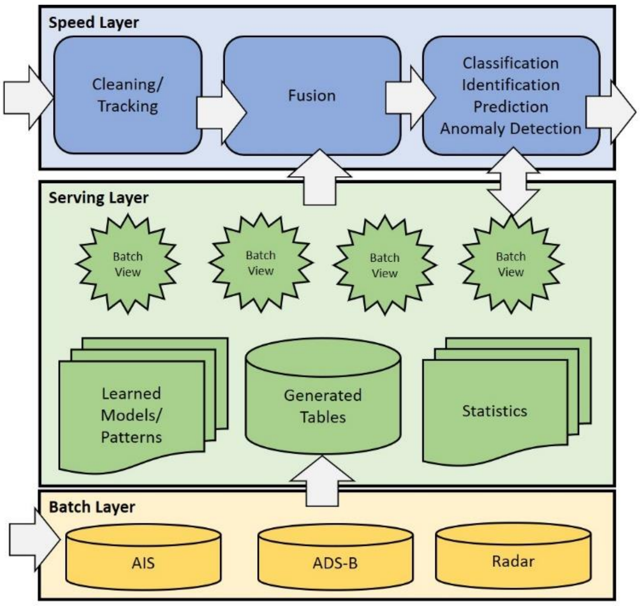

Figure 1: Lambda Architectures.

2 HEAT MAP GENERATION AND VISUAL SITUATION ASSESMENT

Heat maps offer an elementary but nevertheless very intuitive way to improve situation awareness based on trajectories. Basically, this means defining a spatial grid and evaluating statistics for each grid cell. This includes the number of hits with trajectories, the average speed, course, etc. Some statistics will follow mono-modal distributions others are multi-modal. These heat maps can be presented in a graphical layer. Comparing the actual situation with the heat map offers an increase in situation awareness in a very visual way. The generation of such statistics is supported in the serving layer with modern tools of distributed processing like Spark. With these tools it is possible to generate even global heat-maps on long term basis and to adapt these to different time windows continuously. Additionally, these heat maps can be used to generate warnings or alarms.

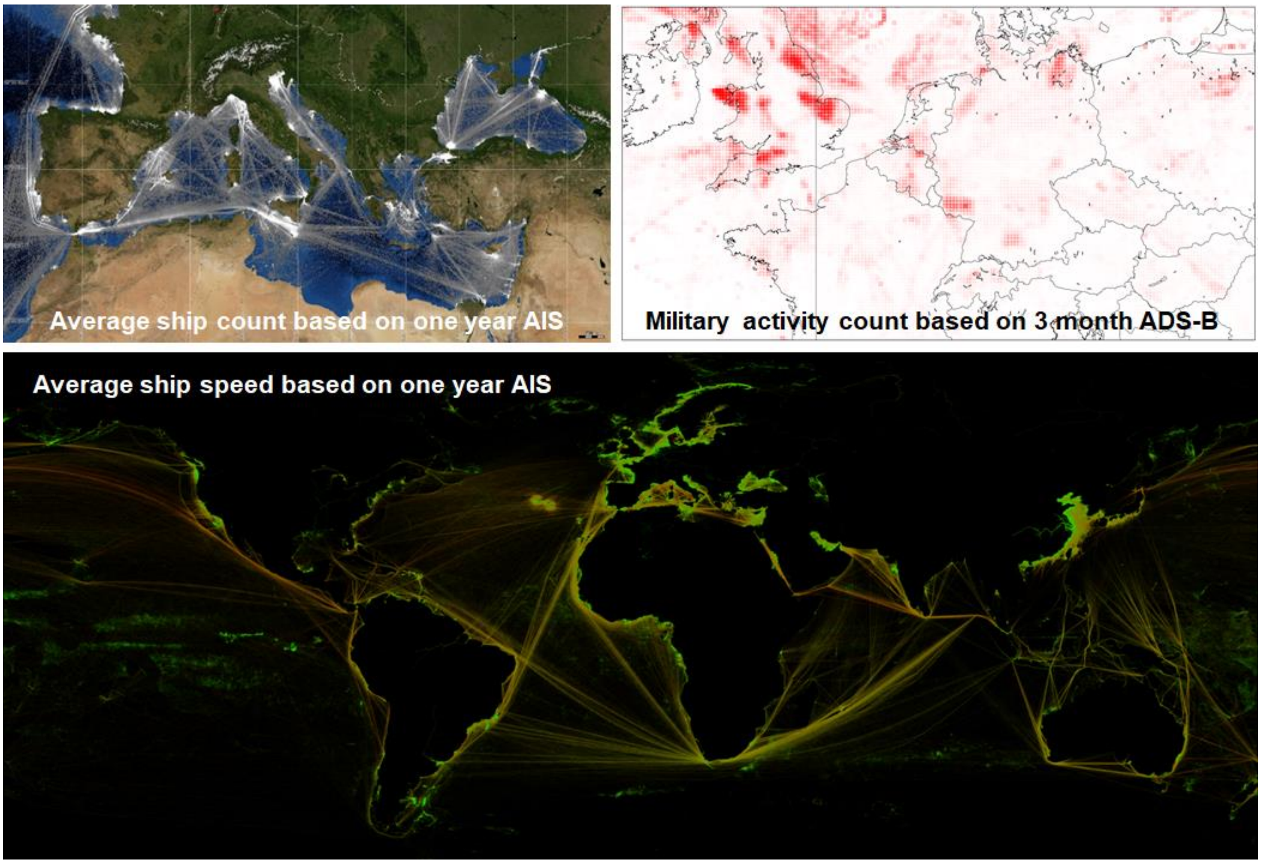
Average ship count based on one year AIS
Military activity count based on 3 month ADS-B
Average ship speed based on one year AIS

Figure 2: Examples for Long term Heat maps.

3 UNSUPERVISED MACHINE LEARNING FOR TRAFFIC ASSESSMENT

3.1 Area of Interest Extraction by Point Clustering

There are several areas of interest a trajectory normally passes through. This could be e.g. airports, heliports, harbours, specific areas close to offshore drilling rigs and wind parks or landing areas of ferries. Plots belonging to these areas are easily to extract from the trajectories. They are characterised e.g. by low speed or climb rate or simply by the fact that they locally start or terminate trajectories. After their extraction these plots can be clustered. The convex hull of these point clusters are candidates for areas of interest [24], [25], [26].

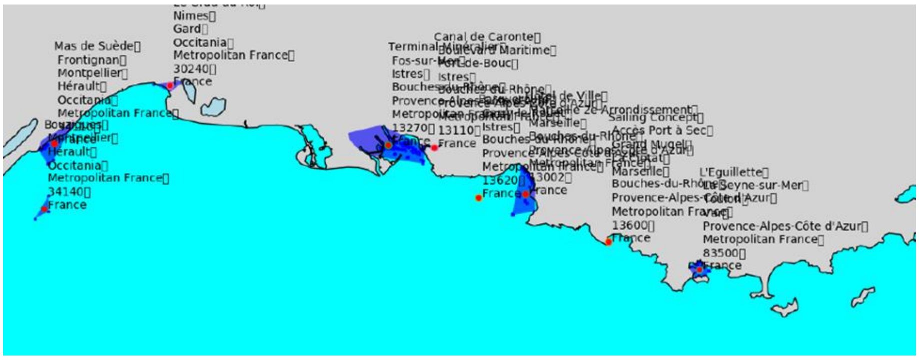

Figure 3: Clustered Harbours and airports.

The most popular cluster algorithms are k-MEAN and density-based spatial clustering (DBSCAN) [22], [23], [26], [27]. The advantage of DBSCAN is the capability to simultaneously determine the number of point clusters, i.e. it applied even if the number of clusters is unknown.

The most important extension to the DBSCAN is known as OPTICS (Ordering Points to Identify the Clustering Structure).

The special challenge of clustering is its implementation for very large datasets in dedicated IT environments like Spark, because these algorithms are per se not parallelisable like k-MEAN. To split the task into several tasks the data has to be separated prior to processing. Data partitioning into geo grid tiles is one way to go, while each tile can then be clustered in parallel [25], [28].

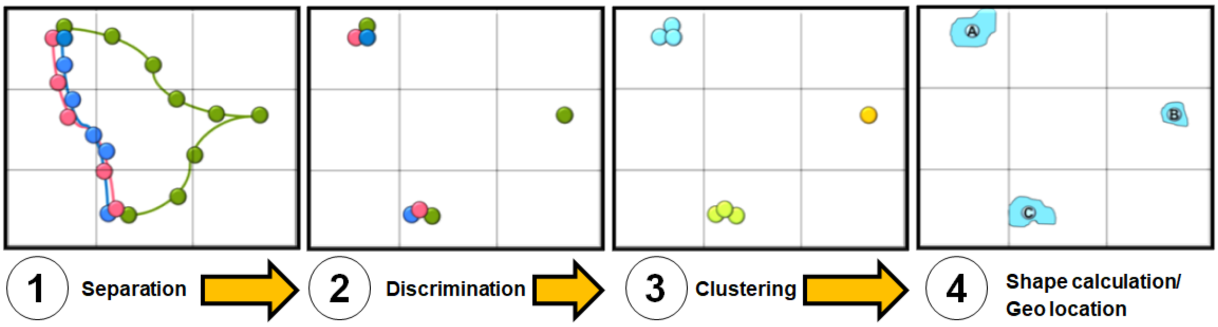

Figure 4: Principle of clustering of aggregated points.

If the data used for clustering contains an identifier, e.g. MMSI or ICAO, relations between objects and harbours/ airports can be extracted, such as the traveling behaviour between harbours/ airports. This finally opens the door to data mining and relational graphs.

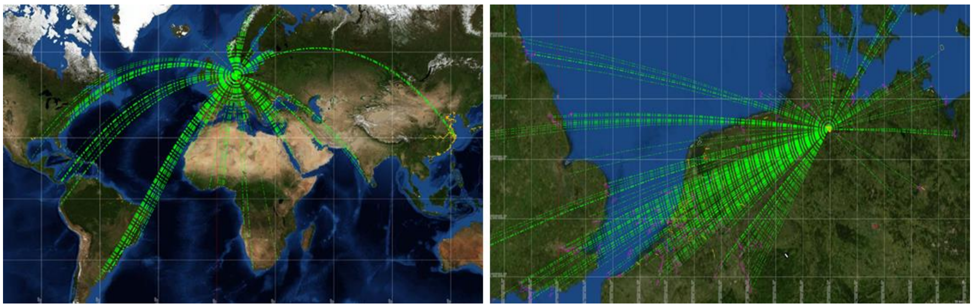

Figure 5: Some Harbour connections of Hamburg based on one-month AIS data.

Besides the extraction of areas of interest, this processing leads to the partition of trajectories into graphs, where the vertices (nodes) correspond to the areas of interest and the edges to the sub-trajectories with known starting and end node. So, this processing ends up in an inherent pattern of life analysis of the trajectory's objects.

2.2 Route Extraction by Trajectory Clustering

If trajectories between two areas or points look similar this may be an indication that they follow the same infrastructure or routes. So, the comparison of trajectories enables an elegant way of traffic flow analysis. This may be used to detect hidden routes e.g. for smuggling or other activities. Or the deviation from highly frequented routes may be a reason to focus attention because of piracy or kidnaping. To judge the similarity between trajectories in a mathematical manner requires the definition of a distance between them.

A widespread method to define a distance between trajectories is inspired by language processing, called dynamic time warping (DTW) [29], [30], [31]. The central idea within this method is the construction of a warping path, i.e. an assignment between the points of two trajectories, see Fig.6. Then one can define a distance between two trajectories as the sum of the warping distances.

Other distances are the Hausdorf [30] or Frechet-Distances [30], [32].

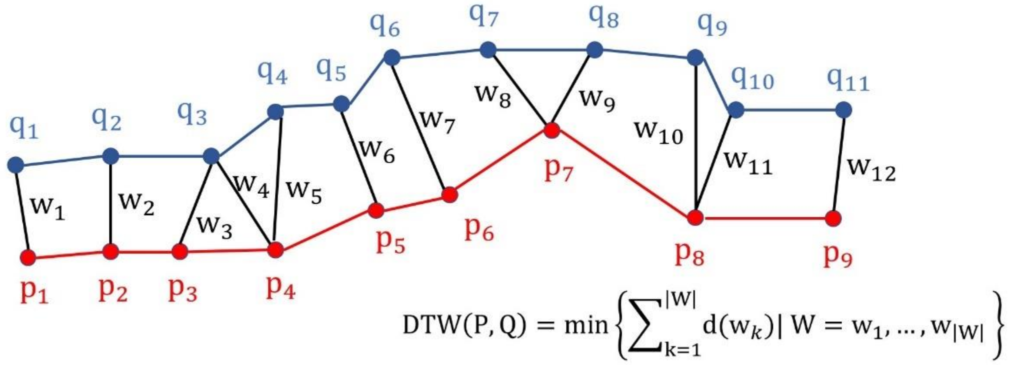
q_1 q_2 q_3 q_4 q_5 q_6 q_7 q_8 q_9 q_{10} q_{11}
w_1 w_2 w_3 w_4 w_5 w_6 w_7 w_8 w_9 w_{10} w_{11} w_{12}
p_1 p_2 p_3 p_4 p_5 p_6 p_7 p_8 p_9
$\mathrm{DTW}(P,Q) = \min\left\{\sum_{k=1}^{|W|} d(w_k) \mid W = w_1, \ldots, w_{|W|}\right\}$

Figure 6: DTW and warping path.

Having defined a distance between trajectories, clustering (e.g. DBSCAN or related methods) can be applied to find trajectory clusters, i.e. trajectories which following similar movement patterns [25], [26], [33], [34], [35].

Moreover, averaging methods like the DTW-Barycentre Algorithm [35] can be used to calculate dedicated representatives of these trajectory clusters. These representatives are very convenient to build up a route map. Fig. 7 shows trajectory clusters between airports based on ADS-B and routes based on averaged AIS trajectories.

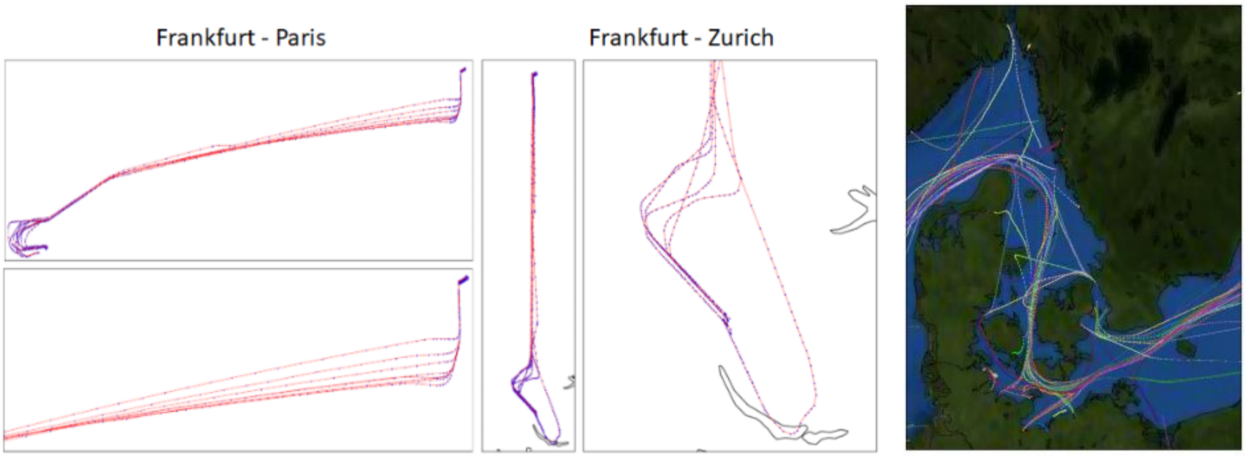

Figure 7: DTW Clusters and Average routes.

2.3 Route Assignment, Prediction and Anomaly Detection

Routes are coarsened by a grid tile and are described by a set of continuous tile indexes. Typically for indexing the OSM standard is used [36], which is given by:

$$x = 2^z * \left\lfloor \frac{longitude + \pi}{2\pi} \right\rfloor$$

$$y = 2^{z-1} * \left\lfloor \frac{1 - \ln\left(\tan(latitude) + \frac{1}{\cos(latitude)}\right)}{\pi} \right\rfloor$$

z is the zoom level, which is e.g. the size of a city for zoom level 11 or small road for zoom level 15 [36].

Using this conversion from geo location to a grid index, each track update can be assigned to a tile. With a query on a route table that is stored in a No-SQL database, every route containing the same index is received. Finally, a comparison between track heading and route direction and including the count of track contributions gives the possible routes the track might travel on. So, for a set of routes *R* the query returned, the following probability estimation applies:

$$p(r_i) = \frac{n_i}{N} * \left(1 - \frac{\lfloor heading - \angle(r_i^j, r_i^{j+1}) \rfloor}{\sum_k \lfloor heading - \angle(r_k^j, r_k^{j+1}) \rfloor}\right)$$

Where

- route $r_i \in R$,
- $n_i$is the count of track contributions for the ith route and $N = \sum_i n_i$,
- $\angle(r_i^j, r_i^{j+1})$ *is* the north oriented angle between the jth and (j+1)th tile centre location..

If the track identifier (MMSI, ICAO) are known they can be compared with the list of identifiers that each route owns. If it is found for a route, it gives the assignment more confidence.

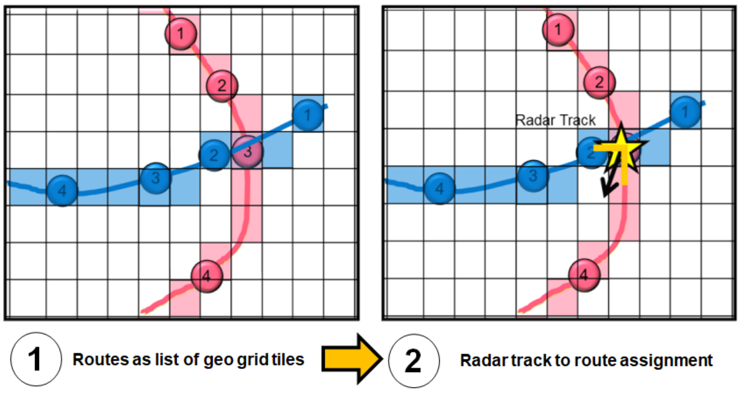
1
2
3
4
Radar Track
1 Routes as list of geo grid tiles
2 Radar track to route assignment

Figure 8: Principle of the track to route assignment.

Finally, route maps and additional route statistics are the basis for numerous anomaly detection approaches. It creates relations between known objects, e.g. vessels or aircraft and new generated data, e.g. harbours, airports, routes. AIS and ADS-B data from data provider enrich this knowledge spectrum with additional data, e.g. nationality, airline, aircraft or vessel type, etc. So, there is much information of different kinds of relations available that offers new opportunities for data mining and anomaly detection.

3 SUPERVISED MACHINE LEARNING FOR CLASSIFICATION

Labelled data, e.g. AIS or ADS-B data, is typically used for supervised machine learning, especially classification topics. Each trajectory point has one or more labels, e.g. the ship type, navigation status for AIS and military flag, aircraft type for ADS-B, which can be learned in association to given so called features of the trajectory. These features are derived from the complete or windowed trajectories and are based on the object's positional and kinematical history. So, the label can be predicted just based on the positional or kinematical features of the trajectory after the training process, which takes place at the serving layer.

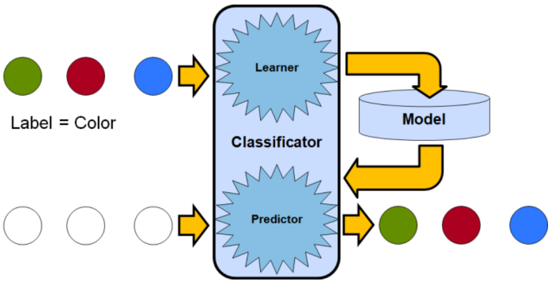

Figure 9: Principle use of a machine learning classifier. Labelled data is used for learning and the model is saved to HDD. The classifier uses e.g. radar data to predict the label.

For classification tasks a lot of classical machine learning algorithms are known, such as: Logistic Regression, Supported Vector Machine, Decision Trees, Ensembles of Decision Trees, e.g. Random Forest, Gradient Boost Tree and Multi-Layer Perceptron, which is the simplest form of a neural network [22], [23], [27].

Which algorithm should be chosen depends on the implementation constraints, e.g. integration environment, learning time, number of labels etc. as well as on the data itself. Normally these algorithms expect an even distribution of all label classes since they learn the association between different feature forming and a label. If the distribution of classes is highly imbalanced, additional strategies have to be considered that consider the occurrence of minority and majority classes. Known strategies are class weighting, under- and oversampling [37], [38], [39], [40], [41], [42], [43]. However, every classifier reacts differently to these strategies but also to the feature data itself. So, it should always be considered, what information is available and how it affects the classifier.

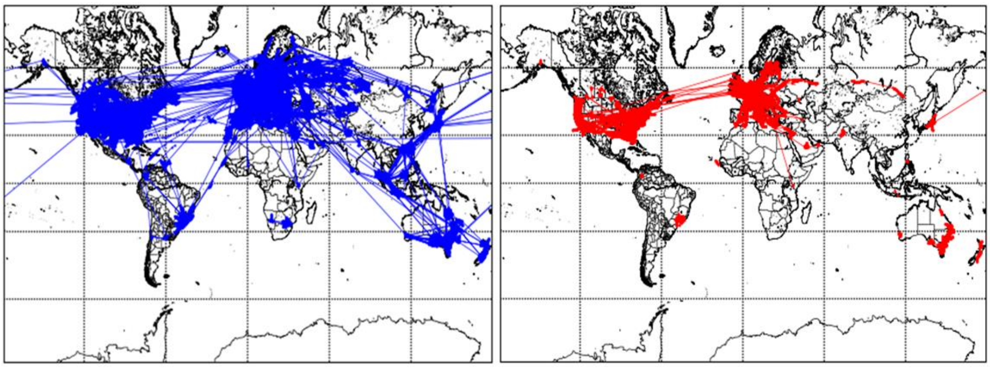

Figure 10: Class Imbalance for military and civil aircraft.

Finally, the classifier which was trained in the serving layer can be delivered to the speed layer and can be applied to the real time data streams of the trajectories.

These classifiers can be used in two ways: First they can be used to detect spoofing within cooperating data sources, like ADS-B and AIS. So, they are used to detect illegal fishing, to discriminate between military and civil aircraft trajectories or other kinds of anomaly detection. Second, they can be used to allow classification capabilities for unlabelled, often uncooperating data sources, like coastal radar networks or GMTI radar applications [18].

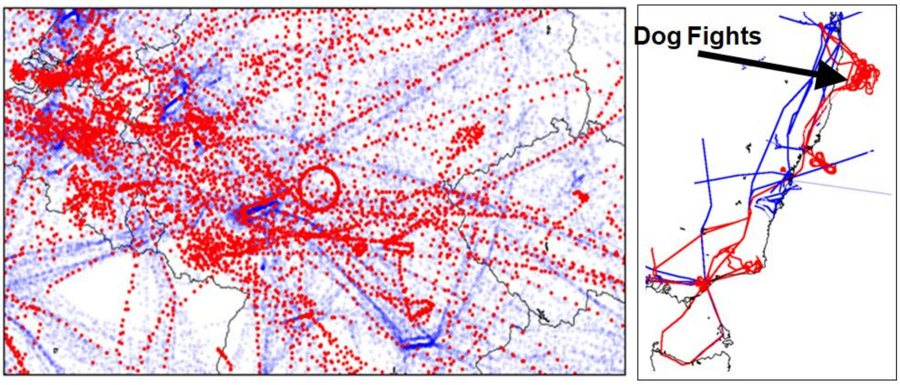

Figure 11: Military surveillance patterns and dog fights.

CONCLUSION

Real time surveillance in military and security becomes a difficult task. Surveillance takes place in an ambiguous and complex environment and has considered asymmetric threats. The trial to address this challenge with an increased amount of data ends up in a big data problem. The user needs support to control the high data rates and to keep decision-making ability.

On the other hand, the new big data world is a chance. Therefore, modern surveillance systems have to integrate big data concepts like lambda architectures. Further, it requires the usage of distributed data processing which is supported by the new Apache open source tool landscape, e.g. Spark, Hadoop, Kafka or Flink. Then worldwide localisation data received by networks like AIS, ADS-B and sensors delivers trajectories which pave the way for advanced data analytics, supervised and unsupervised machine learning within surveillance systems. Finally, this improves decision support.

It is possible to extract areas of interest for airborne, maritime and land scenarios. Further, the traffic between these points can be classified based on cluster algorithms and route maps can be extracted. Labelled trajectories can be used to train classifiers and predictors which can be applied in the real time processing.

REFERENCES

[1] STO Reference Text

[2] Federal Aviation Administration, www.faa.gov/

[3] Eurocontrol, www.eurocontrol.int

[4] Specification for Surveillance Data Exchange ASTERIX Part 12 Category 21 ADS-B Target Reports, EUROCONTROL

[5] John Scardina: "Overview of the FAA ADS-B Link Decision". Federal Aviation Administration, 7 June 2002.

[6] “Report from the ADS-B Aviation Rulemaking Committee to the Federal Aviation Administration”, Federal Aviation Administration, September 26, 2008

[7] Edward A. Lester and R. John Hansman: "Benefits and Incentives for ADS-B Equipage in the National Airspace System", MIT International Center for Air Transportation Department of Aeronautics & Astronautics Report No. ICAT-2007-2 August 2007

[8] Flightradar24: www.flightradar24.com/

[9] ADS-B Exchange: www.adsbexchange.com/

[10] FlightAware: flightaware.com/

[11] International Civil Aviation Organization, www.icao.int

[12] http://www.imo.org

[13] https://www.navcen.uscg.gov/

[14] Samuel Blackman and Robert Popoli: “Design and Analysis of Modern Tracking Systems”, Artech House, Boston, 1999

[15] Wolfgang Koch: “Tracking and Sensor Data Fusion: Methodological Framework and Selected Applications”, Springer Science & Business Media, 2013.

[16] Yaakov Bar-Shalom, X. Rong Li, Thiagalingam Kirubarajan: “Estimation with applications to tracking and navigation: theory algorithms and software”, John Wiley & Sons, 2004.

[17] Marz, Nathan, and James Warren: “Big Data: Principles and best practices of scalable realtime data systems.” Manning Publications Co., 2015.

[18] Dästner,Kaeye, et al.: "Machine Learning Techniques for Enhancing Maritime Surveillance Based on

GMTI Radar and AIS." International Radar Symposium (IRS), 2018 19th International. IEEE, 2018

[19] Apache Cassandra, a no-sql database: cassandra.apache.org

[20] Busyairah Syd Ali, Wolfgang Schuster, Washington Ochieng, Arnab Majumdar: "Analysis of anomalies in ADS-B and its GPS data", GPS Solutions July 2016, Volume 20, Issue 3, pp 429–438, Springer

[21] Apache Spark, a unified analytics engine for large-scale data processing: spark.apache.org

[22] scikit-learn, Machine Learning in Python: scikit-learn.org

[23] Apache Spark MLlib, a scalable machine learning library: spark.apache.org/mllib

[24] Fu, Zhongliang, et al. "A two-step clustering approach to extract locations from individual GPS trajectory data." ISPRS International Journal of Geo-Information 5.10 (2016): 166.

[25] Dästner, Kaeye, et al. "Exploratory data analysis for GMTI radar." Radar Symposium (IRS), 2017 18th International. IEEE, 2017.

[26] Nicolas Le Guillarme, Xavier Lerouvreur: Unsupervised Extraction of Knowledge from S-AIS Data for Maritime Situational Awareness

[27] Aurélien Géron: "Hands On Machine Learning with Scikit-Learn & TensorFlow", O'Reilly, 5th Release, 2018

[28] Gui, Zhiming, Haipeng Yu, and Yunlong Tang. "Locating Traffic Hot Routes from Massive Taxi Tracks in Clusters." J. Inf. Sci. Eng. 32.1 (2016): 113-131.

[29] Jin, Xiaoyu, Yang Yang, and Xuesong Qiu. "Framework of Frequently Trajectory Extraction from AIS Data." Proceedings of the 2017 The 7th International Conference on Computer Engineering and Networks. 22-23 July, 2017 Shanghai, China (CENet2017)

[30] Michel Marie Deza, Elena Deza: Encyclopedia of Distances, Springer 2009

[31] Müller, Meinard. "Dynamic time warping." Information retrieval for music and motion (2007): 69-84.

[32] Besse, Philippe, et al. "Review and perspective for distance based trajectory clustering." arXiv preprint arXiv:1508.04904 (2015).

[33] Huanhuan Li, Jingxian Liu, Ryan Wen Liu, Naixue Xiong, Kefeng Wu, Tai-hoon Kim: A Dimensionality Reduction-Based Multi Step Clustering Method fro Robust Vessel Trajectory Analysis. In Sensors 2017

[34] Kaeye Dästner, Susie Brunessaux, Elke Schmid, Bastian von Haßler zu Roseneckh-Köhler, Felix Opitz: "Classification of Military Aircraft in Real-time Radar Systems based on Supervised Machine Learning with Labelled ADS-B Data", 12-th Symposium Sensor Data Fusion, 9-11 October 2018, Bonn, Germany

[35] Petitjean, François, Alain Ketterlin, and Pierre Gançarski. "A global averaging method for dynamic time warping, with applications to clustering." Pattern Recognition 44.3 (2011): 678-693.

[36] https://wiki.openstreetmap.org/wiki/Zoom_levels

[37] Leo Breiman: "Random Forests", Machine Learning October 2001, Volume 45, Issue 1, pp 5–32, Springer

[38] Pedregosa, Fabian, et al. ”Scikit-learn: Machine learning in Python.” Journal of machine learning research 12.Oct (2011): 2825-2830.

[39] Chen, C., Liaw, A & Breiman, L.: "Using Random Forest to Learn Imbalanced Data", Technical Report 666. Statistics Department of University of California at Berkeley, Berkeley, 2004.

[40] Leo Breiman: "Arcing The Edge", Technical Report 486 , Statistics Department University of California, Berkeley, June 1997

[41] Mason, L., Baxter, J., Bartlett, P. L., Frean, Marcus: "Boosting Algorithms as Gradient Descent" In S.A. Solla and T.K. Leen and K. Müller. Advances in Neural Information Processing Systems 12. MIT Press. 1999, pp. 512–518.

[42] L. Mason, J. Baxter, P. Bartlett, M. Frean, "Boosting Algorithms as Gradient Descent in Function Space", Proc. Neural Information Processing Systems Conf., pp. 512-518, 1999.

[43] X. Zeng and T. R. Martinez, “Distribution-balanced stratified cross-validation for accuracy estimation”, Journal of Experimental & Theoretical Artificial Intelligence, vol. 12, no. 1, pp. 1–12, 2000.

[44] Apache Kafka, a distributed steraming platform: kafka.apache.org

[45] S. Hochreiter and J. Schmidhuber, “Long short-term memory”, Neural Computation , vol. 9, no. 8, pp. 1735–1780, 1997.

[46] Guiliana Pallotta, Michele Vespe, Karna Bryan: Vessel Pattern Knowledge Discovery from AIS Data: A Framework for Anamoly Detection and Route Prediction. Entropy 2013, 15, 2218-2245, 2013

[47] Bo Liu, Erico N. de Souza, Stan Matwin, Marcin Sydow:Knowledge-based Clustering of Ship Trajectories Using Density-based Approach. IEEE International Conference on Big Data, 2014.

[48] Liu, Bo. "Parallel Maritime Traffic Clustering Based on Apache Spark."

[49] Andrej Dobrkovic: Using Machine Learning for Unsupervised Maritime Route Discovery. University of Twente.

[50] Salvador, Stan, and Philip Chan. "Toward accurate dynamic time warping in linear time and space." Intelligent Data Analysis 11.5 (2007): 561-580.