A Machine Learning Framework for Performance Prediction of an Air Surveillance System

A Machine Learning Framework for Performance Prediction of an Air Surveillance System

Juha Jylhä#1, Marja Ruotsalainen#2, Ville Väisänen#3, Kai Virtanen*4, Mikko Harju*5, Minna Väilä#6

#*Laboratory of Signal Processing, Tampere University of Technology, Tampere, Finland*

[1]juha.jylha@tut.fi, [2]marja.ruotsalainen@tut.fi, [3]ville.j.vaisanen@gmail.com, [6]minna.vaila@tut.fi

* *Department of Mathematics and System Analysis, Aalto University, Espoo, Finland*

[4]kai.virtanen@aalto.fi, [5]mikko.harju@aalto.fi

***Abstract*—The optimal use of a surveillance radar system requires proper understanding about the system behavior in different configurations, modes, and operating conditions. This paper proposes a machine learning framework for producing and validating the performance model of the surveillance radar system. The framework consists of an optimization method for the parameterization of a radar model and a machine learning method for the modeling of a tracker. Optimization and machine learning is based on the satellite navigation data of cooperative aircraft and corresponding track data from the surveillance system. The aim is to learn the system performance in a wide range of operating conditions using the extensive measurement history and then to predict the present performance with high accuracy at specified locations in the airspace. The feasibility of the proposed framework is assessed using real data.**

I. Introduction

Modern surveillance systems comprise networked, remote-controllable multifunction sensors and information processing capabilities to provide a picture of the air situation. The surveillance system is controlled by operators aiming at the best possible situational awareness. The actions and performance of every, possibly self-adaptable, radar may be hard to predict for every circumstance. It is even harder to predict the overall performance of the networked system with high accuracy taking into account all the relevant factors, such as the weather and the radar operation mode. However, understanding the system performance in all conditions plays a key role when making decisions related to the optimal use of the system. Typical solutions for system performance assessment are based on physical modeling and simulation studies that are verified with measurements [1], [2], [5], [4], [8], [9], [10].

In this paper, we focus on the learning and validation of the performance model based on measurement data. We divide the research problem into two subproblems: the learning of a radar model and the learning of a tracker model. The radar model follows the classic radar equation, and the challenge is to determine its parameter values so that the model is able to predict the performance accurately in dynamically varying conditions taking the radar mode setup into account. The tracker modeling, instead, is defined as a machine learning problem where the aim is to learn the relationship between the radar performance and the tracker performance, i.e. the tracker model uses the learned radar model as its input. Concatenated, the radar model and the tracker model predict the tracker performance based on the radar model input. The resulting model is fast to evaluate (without Monte-Carlo simulations) but still aims to produce accurate predictions for all the learned conditions.

We propose a machine learning framework, illustrated in Fig. 1, where the accuracy of the performance model is optimized by automatically learning the parameter values of the radar model and the structure of the tracker model. The measured data provide information about the real-world behavior of the system for the creation and validation of the radar and tracker models. The measurements include the satellite navigation data from aircraft and the surveillance system data, i.e. detections and tracks.

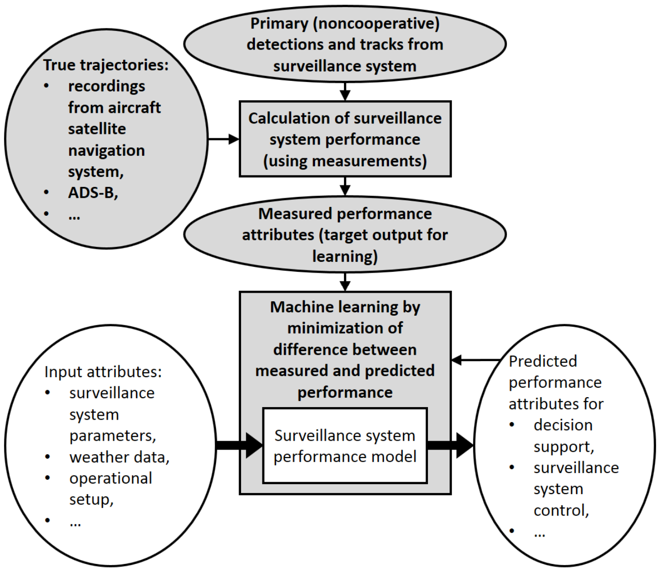

Fig. 1 The block diagram illustrates our machine learning framework. The gray blocks correspond to processing performed offline and represent the utilization of the measurement data for updating the surveillance system performance model for the condition in question. All different conditions are aimed to be learned. The white blocks correspond to processing performed online and represent the real-time performance prediction including all the learned operating conditions.

II. Performance Model of Air Surveillance System

A model of system performance is typically meant to predict certain application-specific attributes that describe the fundamental capability of the system to perform its tasks.

Next Subsections II.A and II.B introduce the radar and tracker performance models that we use in our framework.

A. Radar Model

The core of the radar performance model is the calculation of signal-to-noise-ratio (S/N). It is estimated with the so-called radar equation which combines the properties of the radar with the reflectivity of the target, the effects of the environment, and the geometry. The equation for monostatic radar is formed as in [3]:

$$\frac{S}{N} = \frac{P G_{\mathrm{T}} G_{\mathrm{R}} \lambda^2 \sigma G_{\mathrm{SP}} G_{\mathrm{I}}}{(4\pi)^3 R^4 (F k T_0 B + S_{\mathrm{W}}) L}, \tag{1}$$

in which

P = Peak transmit power (W)
G_{T} = Transmit antenna gain
G_{R} = Receive antenna gain
λ = Wavelength (m)
σ = Radar cross section (RCS) of the target (m^2)
G_{SP} = Radar signal processing gain
G_{I} = Radar pulse integration gain
R = Distance between the radar and the target (m)
F = Receiver noise figure
k = Boltzmann's constant, $1.38064852 \times 10^{-23}$ J/K
T_0 = Receiver noise temperature (K)
B = Receiver noise bandwidth (Hz)
S_{W} = Radar signal power produced by backscattering from precipitation (W)
L = Radar-related losses.

Our aim is to incorporate real-world phenomena and their variation into the model using measurement data. The effects of these phenomena on the radar performance are typically hard to estimate otherwise. For example, the term S_{w} is important to consider in real-world conditions where the atmosphere is not ideal. However, determining a proper value for S_{w} is not entirely straight-forward, as it depends on the type of the precipitation, the distance between the radar and the target, and terrain shadowing besides the properties of the radar.

The model could be extended to consider propagation effects in more detail [5]. The inclusion of such a propagation factor into (1) could involve the modeling of the shadowing, the Fresnel, or the diffraction effects caused by the terrain, as well as the attenuation caused by the atmospheric conditions.

The probability of detection P_{d} is determined by the integration of Ricean probability distribution based on the S/N and the desired value for the probability of false alarm P_{fa} [3]. It is worth noting that, in addition to considering a target with a constant RCS, (1) can be used to calculate the S/N to obtain the P_{d} for a RCS distribution representing the target [10].

B. Tracker Model

A typical performance model of a multi-sensor tracking system involves a Monte Carlo simulation where the tracking algorithms [1], [5] are combined with a sophisticated simulation software that produces flight trajectories and radar detections and tracks. Simplified methods for predicting the performance of the system are needed [2], e.g. to reduce the computational load of the prediction for real-time applications or to produce the prediction in diversified conditions for which the simulation system setup is not fast and easy to implement.

We formulate the performance model in such a way that it predicts the tracker performance as a predefined level. In this paper, the levels in the experiments are referred to as "excellent", "good", "moderate", "weak", and "poor". The attributes we consider for the definition of these levels are the probability and the accuracy of the track. In practice, a system operator is assumed to define these performance levels by setting thresholds for the tracker performance attributes. However, the proposed framework does not restrict the use of a more sophisticated definition of the tracker quality levels. The definition of the tracker performance as levels enables us to use machine learning, i.e. classification, for the tracker modeling.

III. MEASUREMENT-BASED PERFORMANCE ASSESSMENT

The underlying challenges in the performance prediction relate to determining the validity and accuracy of the performance models when the system is highly complex and even self-adaptive towards the prevailing conditions. How can, for example, the type and behavior of the targets as well as the weather and the atmosphere be taken into account? We use the recorded flight trajectories of aircraft and the corresponding detections and tracks from a surveillance system to assess the true performance of the radar and the tracker; see Fig. 1.

A. Association of Measured Flight Paths and Measured Tracks

Accurate knowledge on the flight paths of aircraft are produced by onboard satellite navigation systems. The position data can be collected from the cooperative aircraft via automatic dependent surveillance – broadcast (ADS–B) technology or by standard flight recordings. The former provides the position data online and the latter offline.

The acquired trajectory represents the true position of the aircraft during its flight. We have developed a procedure that automatically associates the detection and track data produced by an air surveillance system with these acquired trajectories. The association is performed by calculating similarity metrics, such as mean error and mean absolute error, for all the trajectory-track-pairs and choosing the pairs representing the best match and having similarity values below the predefined threshold. This procedure must be automatic as huge amounts of data have to be processed in order to assess the performance attributes extensively and accurately.

B. Measured Performance

The radar and tracker performance attributes are calculated by comparing the flight trajectories and tracks associated to each other. A similar idea is discussed in [8] from the point of view of radar calibration. In our framework, momentary performance of the surveillance system is calculated for a set of volume cells of the airspace, voxels. The idea is to produce a 3D grid of the attributes that represents the surveillance system performance over the airspace. The measured radar performance attributes include the probability to detect an aircraft of a certain type and the (range, azimuth, altitude, Euclidean) accuracy of the detections. The tracker performance attributes include the probability of the track and the Euclidian accuracy of the tracks. The attributes for each voxel are averaged over the paths, or over the partial paths, within the voxel in question. Fig. 2 illustrates the linkage of the trajectories to the 3D grid. Large amount of trajectories, detections and tracks ensure that the estimation accuracy of the performance attributes is adequate. Further it contributes to the generalization ability of the data-induced models because more measurements involve more different realizations. The good generalization ability is one goal in our model validation and learning framework.

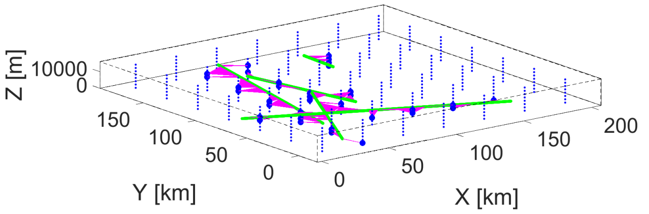

Fig. 2 Flight trajectories are linked with the 3D grid of voxels that represent the airspace. The flight trajectories are represented with green lines and the voxel centers with blue dots. Bold dots highlight the linked voxels and the linkages are visualized with red lines

IV. The Machine Learning Framework

The proposed framework is described generally in Fig. 1. In the learning, we are looking for models which produce radar performance attributes or tracker quality levels that are maximally similar (minimally different) to the measured ones in the voxels of the 3D grid.

A. Optimizing Radar Performance Model

We formulate the optimization of the radar performance model as a search of the optimal values for the model parameters listed in Subsection II.A. Even though an analytic formulation of the optimization problem is possible, we use a brute-force search procedure due to its straightforward capability to optimize basically any radar model parameter or set of parameters. Our method searches the optimal values for the selected model parameters from predefined sets of values.

In our first experiment, the selected parameters (from the list in Subsection II.A) are the peak transmit power and the gain of the radar antenna as a function of elevation. The P_d is simulated using every combination of parameter values, and the model parameterization yielding the minimum mean difference (over all the voxels) between the modeled and measured P_d is determined. The selection of reasonable parameters and their possible values depend on the radar type and the conditions in question. Fig. 3 gives an example of the optimization result. The original prediction (model) clearly differs from the true performance (measured). After the optimization, the prediction follows the measured P_d nicely. Notice that even though Fig. 3 shows only one altitude layer, all the altitudes are taken into account in the optimization. The training data of this first experiment includes only 30 voxels, but the result is promising.

We demonstrated the implementation and the feasibility of the radar model optimization. Our future directions to develop the optimization comprise 1) taking the aircraft heading into account (related to the RCS); 2) a study of optimization parameter selection and how they should be linked together; 3) a study of forming a multi-criteria objective function where the P_d is incorporated with the P_{fa} and the accuracy of the radar (in range, azimuth, and altitude); 4) a study of different optimization procedures (to replace the brute-force search); and 5) utilizing a lot more train and test data that contain a wide range of operating conditions. After those enhancements, the maturity of the proposed framework (concerning the radar modeling) will be much further.

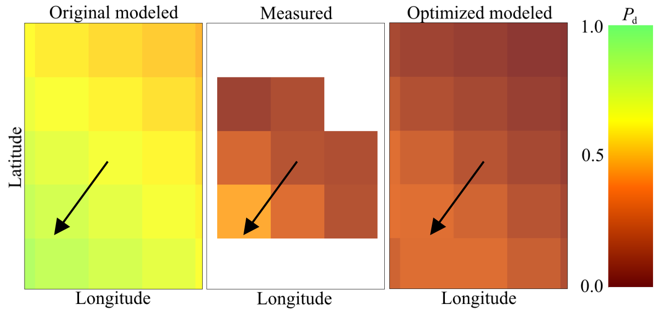

Fig. 3 Measured P_d (middle) compared with original modeled P_d (left) and optimized modeled P_d (right). The direction, where the radar is located, is shown by arrows. White color in the middle image represent area (voxels) with no measurements available in the used data set.

B. Learning Tracker Performance Model

Due to the fact that a tracker as a complex software system is hard to model analytically and physically, we create a tracker model based on machine learning. In the learning, the measured radar performance is the input for the model and the measured performance (quality levels) of the tracker is the target output. However, when using the learned tracker model for real-time predictions, we use the modeled radar performance as the input for the model.

In the experiment considered in this paper, we defined the measured radar performance with the following attributes: 1) the probability of detection as a maximum over all the radars, 2) the Euclidian position error in meters of the radar having the highest probability of detection, 3) the range position error in meters of the radar having the highest probability of detection, 4) the azimuth position error in degrees of the radar having the highest probability of detection, 5) the aggregated scan frequency of radars as a sum where the scan frequency of each radar is weighted using the probability of detection, 6) the weighted range position error in meters where the weights are the probability-of-detection-weighted scan frequencies of radars, and 7) the weighted azimuth position error in meters where the weights are the probability-of-detection-weighted scan frequencies of radars.

We defined the target output using five performance levels as described in Subsection II.B. In learning, we basically look for a simple, easy-to-interpret model that is able to portray the cause and effect relationship between the radar performance and the tracker performance levels. The created model is a classifier that returns the output class (the tracker performance level) for a given input. Classifiers were created using our rule learning method, GPRL (genetic programming based rule learner) [6], [7].

Table I gives an example of the learned tracker performance model. The model consists of four simple rules. According to the model, "excellent" tracker performance, for example, is declared when radar performance attribute *Pd* is larger than threshold t_1 and either radar performance attribute *AccRAvg* (average range accuracy) is smaller than threshold t_2 or *AccAzAvg* (average azimuth accuracy) is smaller than t_3. *AccEuc* refers to Euclidean accuracy and *FScan* to average scanning frequency of all the radars together. Our learning method determines also the numerical values of the thresholds.

TABLE I

AN EXAMPLE OF THE TRACKER PERFORMANCE MODEL CREATED BY OUR GPRL LEARNING METHOD. THE METHOD AUTOMATICALLY DETERMINES THE RULES AND THRESHOLD VALUES BASED ON MEASUREMENT DATA.

```
If (Pd > t1) and ((AccRAvg < t2) or (AccAzAvg
< t3))
    Return "EXCELLENT"
Else if (Pd > t4) and (AccEuc < t5)
    Return "GOOD"
Else if (Pd > t6) and ((Pd > t7) or (AccRAvg
< t8))
    Return "MODERATE"
Else if (FScan > t9) and (AccRAvg < t10)
    Return "WEAK"
Else
   Return "POOR"
```

TABLE II

CONFUSION MATRIX FOR THE MODEL PRESENTED IN TABLE I. THE ROWS OF THE MATRIX REPRESENT THE TRUE PERFORMANCE LEVELS OF THE VOXELS AND THE COLUMNS THE CLASSIFICATION RESULT.

PREDICTED						
Excellent	**Good**	**Moderate**	**Weak**	**Poor**		
2	5	5	0	0	**Excellent**	**TRUE CLASS**
1	6	5	2	1	**Good**	
0	8	14	2	2	**Moderate**	
0	1	5	15	2	**Weak**	
0	0	0	2	7	**Poor**	

The studied airspace comprised 171 voxels. We used half of them (86) as training data and the other half (85) as test data. The learning from the training data produced the model presented in Table I. The generalization capability of the learned tracker model is evaluated using the test data. Table II shows the confusion matrix for the model. One can see, for example, that 15 of the 23 voxels representing the "weak" tracker performance level are correctly classified into the class "weak", five of them into "moderate", two into "poor", and one into "good". Overall, the learned model correctly classified 52 percent of the voxels. It should be noted that incorrect classifications are centered around the correct class. For example, none of the voxels with "excellent" tracker performance is classified as "weak" or "poor". To summarize, our experiments indicate that machine learning clearly has potential for creating the tracker performance model automatically from measurement data.

The preliminary results encourage us to continue developing our learning method further. First, the radar performance attributes used as the tracker model input are simple and more sophisticated attributes could be applied. The current attributes do not consider e.g. the direction from which a radar observes the aircraft. Second, further attention should be paid to the fusion of the performance attributes from multiple radars. Third, the learning method is performed separately for each tracker performance level and the created rules are combined to form an overall model; learning the rules in parallel allowing the learning processes interact with each other could lead to rules that are more coherent (from level to level) and an overall tracker model that is more appropriate.

V. CONCLUSIONS

This paper proposed a framework for learning a performance model for an air surveillance system. It is a concept that can be used for many applications where the accurate prediction of the system performance is needed for diversified operating conditions. We discussed the basic principles of the framework and presented preliminary results.

VI. ACKNOWLEDGEMENT

The authors would like to thank the Finnish Scientific Advisory Board for Defence for funding.

REFERENCES

[1] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*, John Wiley and Sons, New York, NY, 2001.

[2] W. Blair and P. Miceli, "Performance prediction of multisensor tracking systems for single maneuvering targets," *Journal of Advances in Information Fusion*, vol. 7, no. 1, pp. 28–45, June 2012.

[3] P. J. Hannen, *Radar and Electronic Warfare Principles for the Non-Specialist*, 4th ed., Scitech Publishing, Edison, NJ, USA, 2013.

[4] H. Perälä, M. Väilä, J. Jylhä, and A. Visa, "ESPRESS—On efficient bistatic characterization of radar targets," in *Proc. 2015 IEEE Radar Conference*, pp. 259–264, May 2015.

[5] M. A. Richards, J. A. Scheer, and W. A. Holm, *Principles of Modern Radar*, SciTech Pub., 2010.

[6] M. Ruotsalainen, J. Jylhä, and A. Visa, "Reasoning logical rules from multi-sensor data for lifecycle management of aircraft structures," in *Proc. IEEE Congress on Evolutionary Computation 2013*, pp. 3221–3227, June 2013.

[7] M. Ruotsalainen and J. Jylhä, "Learning of a tracker model from multi-radar data for performance prediction of air surveillance system," to be published in *Proc. IEEE Congress on Evolutionary Computation 2017*, June 2017.

[8] S. Symons, D. Harris, L. Hendy, and D. Winter, "SAMT: A toolset for calibration and assessment of radar performance," in *IEE Seminar on Calibration of Active Phased Array Antennas*, pp. 1–5, Mar. 2005.

[9] M. Väilä, J. Jylhä, T. Sailaranta, H. Perälä, V. Väisänen, and A. Visa, "Incorporating a stochastic model of the target orientation into a momentary RCS distribution," in *Proc. 2015 IEEE Radar Conference*, pp. 969–973, May 2015.

[10] M. Väilä, J. Jylhä, V. Väisänen, H. Perälä, A. Visa, M. Harju, and K. Virtanen, "A RCS model of complex targets for radar performance prediction," in *Proc. 2017 IEEE Radar Conference*, May 2017.