Radar Surveillance using Learned Beam Placement Strategies in Non-homogeneous Environments

Student: Jameson D. Morgan

Student Email: morgan.219@wright.edu

Faculty: Dr. Brian D. Rigling

Faculty Email: brigling1@udayton.edu

AFRL Sponsor: Dr. Andrew Duly

AFRL Directorate: AFRL/RY

PA #: AFRL-2024-5798

Motivation

- **Phased array radars** (PARs) are increasingly popular and are excellent for surveillance because beamforming enables radar beam agility [1]
- **Surveillance** is the detection of previously undetected targets (search) and the maintenance of previously detected targets (tracking) [2]
- The **state of practice** is either the track-while-scan or the search-and-track protocol [1]
    - Both use a **raster scan** for the search task [1]
- **Target activity** in the environment may be **non-homogeneous**
    - Neither state-of-practice protocol is able to use knowledge of target activity to improve surveillance performance
- Contribution: An **activity based, data driven PAR surveillance algorithm** that learns target activity patterns to improve surveillance beam placement

Activity-Based Beam Placement Algorithm

- **Goal:** Learn target activity patterns to improve surveillance beam placement
- **Approach:** A neural network (NN) based **reinforcement learning (RL) agent**
  - Takes a state as input
  - Predicts beam to illuminate using NN
  - The NN is trained by taking actions and backpropagating received reward
- **Novel Contribution**:
  - **Reward** function that permits **online learning** in deployed radar
  - An **activity-based** illumination perspective that only illuminates a sector when needed to follow changes

Do Nothing Action

Radar Observes Difference in Beam

Radar Observes No Difference in Beam

Figure 3.1. The proposed PAR surveillance beam placement algorithm. The goal of the algorithm is to improve the beam placement strategy of PARs engaged in surveillance by incorporating learned target activity patterns. Central to our solution is a NN-based RL agent for deciding the beam to illuminate given the state. **To Train:** The RL state is fed to the RL agent, time progresses by a timestep, the radar takes the specified action, and the reward that results is stored. Three different rewards are possible based on the action chosen, the radar's current perception of the environment, and the true environment condition. The state, action, and reward are recorded in the current trajectory, and the RL state is updated based on the action taken. The previous state-action-reward cycle is repeated for a set number of timesteps after which the trajectory is passed to the REINFORCE policy gradient algorithm to update the agent NN to favor actions with higher rewards given the associated states. The agent is trained for a set number of epochs. **To Deploy:** The RL state is fed to the RL agent to predict where the radar should look in the upcoming timestep. Time advances by one timestep and the radar takes the action chosen by the RL agent (illuminate a particular sector or do nothing). The RL state is updated to reflect the action taken and the process is repeated.

Algorithm Details

Surveillance Environment
- 1-D with maximum of one active target

$$\begin{bmatrix} g_{k+1} \\ v_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & \Delta \\ 0 & 1 \end{bmatrix} \begin{bmatrix} g_k \\ v_k \end{bmatrix} + \begin{bmatrix} \frac{1}{2}\Delta^2 \\ \Delta \end{bmatrix} \omega_k$$

- Target motion using constant velocity model
  - Motion model is unknown to radar
  - g_k denotes current position (m)
  - v_k denotes current velocity (m/s)
  - Δ denotes resolution between timesteps (s)
  - k denotes the current timestep
  - $\omega_k \sim N(0, \sigma_\omega^2)$ is white Gaussian acceleration noise
- Environment is discretized into N sectors
- Each sector i has a probability of appearance $p_{ap}^{(i)}$
- Each sector i has a probability of disappearance $p_{dp}^{(i)}$

Radar Model

$$p_t(i \mid B_j) = Pr\{\hat{c}_i^{(k)} = c_i^{(k)} \mid B_j\} = \begin{cases} p_d, & i = j \\ 0, & i \neq j \end{cases}$$

- Probability of observing true sector condition $c_i^{(k)}$
  - Only beam B_i illuminates sector i
  - $\hat{c}_i^{(k)}$ denotes the estimated sector condition
  - p_d denotes the probability of detection

RL State

$$s_k = (\hat{\boldsymbol{c}}^{(k)}, \hat{C}_h^{(k)}, A_h^{(k)}, \delta_k)$$

- $\hat{\boldsymbol{c}}^{(k)}$ denotes vector of N estimated sector conditions
- $\hat{C}_h^{(k)}$ denotes set of previous h environment conditions
- $A_h^{(k)}$ denotes set of previous h actions
- δ_k denotes delay since a detected event with respect to the upcoming timestep

Actions

$$a_k \in \{0, 1, \ldots, N-1, N\}$$

- Actions 0 through $N-1$ represent picking beams B_0 through B_{N-1}
- Action N represents the do-nothing action where no radar illumination occurs

Reward

$$r(s_k, a_k) = \begin{cases} k - q, & \text{if do} - \text{nothing} \\ d_{max} - \delta_k, & \text{if sector different} \\ \mu, & \text{if sector not different} \end{cases}$$

- q denotes time of last do-nothing action
- d_{max} denotes maximum allowed event detection delay
- μ denotes a penalty for visiting a sector with no change

Results | Experiment 1

Per Sector Setup	Sector									
	0	1	2	3	4	5	6	7	8	9
P_ap	0.02	0	0	0.05	0	0	0.02	0	0	0.02
P_dp	1	0	0	0	0	0	0	0	0	1
Velocity (m/s)	{240}	N/A	N/A	{-240, 240}	N/A	N/A	{-240, 240}	N/A	N/A	{-240}
Width	Approximately 50 timesteps									

- Performance of the proposed algorithm on an environment with **non-homogeneous** probability
- Setup
  - $\Delta = 0.005$ seconds
  - $d_{max} = 500$
  - $p_D = 1$
  - $\omega_k \sim N\left(0, 15^2\right)$
  - $h = 2$
  - Per sector details given in table
- Results:
  - 5,000 MC sims. each 5,000 timesteps
  - NN weights from 5,000 training epoch
- Takeaways (w.r.t. raster)
  - Proposed **lowers average detection delay**
  - Proposed scan pattern **follows target activity**
  - Proposed scan pattern provides **more free time**
  - Like raster, proposed **completely detects all targets** with no missed activity

Figure 5.1. 2-D histogram of detection delays for the experiment one evaluation MC trials. **TOP LEFT:** The appearance delay quantity in log10 that results when a raster scan is used in the environment outlined in experiment one. **TOP RIGHT:** The disappearance delay quantity in log10 that results when a raster scan is used in the environment outlined in experiment one. **BOTTOM LEFT:** The appearance delay quantity in log10 that results when the proposed approach is used in the environment outlined in experiment one. **BOTTOM RIGHT:** The disappearance delay quantity in log10 that results when the proposed approach is used in the environment outlined in experiment one. Note that the colorbar range is different between top figures and bottom figures.

Figure 5.2. Barcode plots showing a subset (first 500 timesteps) of a particular MC trial's action history for the raster and proposed approaches. The black dashed line provides a trace for each target in the environment as it moves. A green rectangle is used to denote the first time the target was detected in the sector and a red rectangle is used to denote the time when the target disappearance was detected in the sector. **LEFT:** Action history for a raster scan approach in the environment outlined in experiment one. The “DN” label on the y-axis refers to the do-nothing action which is unused in a raster scan. **RIGHT:** Action history for the proposed approach in the environment outlined in experiment one.

Results | Experiment 2

Per Sector Setup	Sector				
	0	1	...	8	9
P_ap	0.02	0.02	...	0.02	0.02
P_dp	1	0	...	0	1
Velocity (m/s)	{240}	{-240, 240}	...	{-240, 240}	{-240}
Width	Approximately 50 timesteps				

- Performance of the proposed algorithm on an environment with **uniform** appearance probability
- Setup
  - $\Delta = 0.005$ seconds
  - $d_{max} = 500$
  - $p_D = 1$
  - $\omega_k \sim N\left(0, 15^2\right)$
  - $h = 2$
  - Per sector details given in table
- Results:
  - 5,000 MC sims. each 5,000 timesteps
  - NN weights from 5,000 training epoch
- Takeaways (w.r.t. raster)
  - Proposed **lowers average detection delay**
  - Proposed scan pattern **follows target activity**
  - Proposed scan pattern provides **more free time**
  - Like raster, proposed **completely detects all targets** with no missed activity

Figure 6.1. 2-D histogram of detection delays for the experiment two evaluation MC trials. **TOP LEFT:** The appearance delay quantity in log10 that results when a raster scan is used in the environment outlined in experiment two. **TOP RIGHT:** The disappearance delay quantity in log10 that results when a raster scan is used in the environment outlined in experiment two. **BOTTOM LEFT:** The appearance delay quantity in log10 that results when the proposed approach is used in the environment outlined in experiment two. **BOTTOM RIGHT:** The disappearance delay quantity in log10 that results when the proposed approach is used in the environment outlined in experiment two. Note that the colorbar range is different between top figures and bottom figures.

Figure 6.2. A barcode plot showing a subset (first 500 timesteps) of the action history for the proposed approach in the environment outlined in experiment two. The black dashed line provides a trace for each target in the environment as it moves. A green rectangle is used to denote the first time the target was detected in the sector and a red rectangle is used to denote the time when the target disappearance was detected in the sector. “DN” is the do-nothing action.

Results | Experiment 3

Per Sector Setup	Sector				
	0	**1**	**...**	**34**	**35**
P_ap	0.02	0	...	0	0.02
P_dp	1	0	...	0	1
Velocity (m/s)	{240}	N/A	...	N/A	{-240}
Width	Approximately 15 timesteps				

- **Vary** p_D and evaluate effect on completion % and number of unused timesteps
- Setup
  - $\Delta = 0.025$ seconds
  - $\omega_k \sim N(0, 15^2)$
  - $d_{max} = 1000$
  - $h = 5$
  - Per sector details given in table
- Results:
  - 5,000 MC sims. each 5,000 timesteps
  - NN weights from 5,000 training epoch
  - Comp. %: Targets with no missed activity
  - Unusued: # of timesteps radar not used
- Takeaways (w.r.t. raster)
  - Proposed **increases completion %**
  - Proposed **increases number of unused**

Figure 7.1. Experiment three results. **TOP:** The completion percentages for each pairing considered. For reference, no target was detected at 100% completion percentage in the raster scan cases, and the proposed approach detected 40,703, 34,609, and 37,658 complete targets (i.e., completion percentage of 100%) on the P_D = 1, P_D = 0.9, and P_D = 0.8 scenarios, respectively. In experiment three, a total of 40,710 targets were generated across all the trials in each pairing considered. **BOTTOM:** The number of unused timesteps for each pairing considered. In both plots, the bold horizontal line within the box is the mean value, the lower edge of the box is the value one standard deviation lower than the mean, the upper edge of the box is the value one standard deviation higher than the mean, the lower whisker is the minimum value, and the upper whisker is the maximum value.

Related Works

- Works [3] – [8] are the closest related works as they explicitly consider non-homogeneous target environments
  - Per-sector target arrival and motion modeled as a Poisson point process
  - **Requires *a priori* dynamics knowledge**
- Work in [9] uses genetic algorithm to build target likelihood map of untracked targets to improve search
  - **Focuses solely on search**
- Bayesian filtering-based sensor management is common
  - Examples: [10] – [12]
  - **Requires knowledge of probability density functions**
  - **Solutions often resort to approximate algorithms due to complexity**
- Limited number of machine learning-based beam placement works
  - Examples: [13] – [15]
  - **Few machine learning works exclusively focus on beam placement**

Conclusion

- **Demonstrated** that **PAR surveillance in non-homogeneous environments** can be improved by incorporating environment knowledge
- The algorithm is **online-capable** as the reward function does not require access to quantities that are not available to a deployed radar
- Algorithm incorporates an **activity-based, data-driven** approach
- Future work may include:
  - Making the scenario more realistic by allowing for more than one target in the environment
  - Adjusting the reward function used to explore operator-set surveillance adjustments for other task(s) of interest
  - Extending the algorithm to allow for the dynamics to change over time

References

[1] C. D. Bailey, “Radar antennas,” in *Principles of Modern Radar: Basic Principles*, M. A. Richards, J. A. Scheer, and W. A. Holm, Eds., Raleigh, NC, USA: SciTech, 2010, ch. 9, sec. 9.7.2, p. 329

[2] J. A. Scheer, “Radar search and overview of detection in interference,” in *Principles of Modern Radar: Basic Principles*, M. A. Richards, J. A. Scheer, and W. A. Holm, Eds., Raleigh, NC, USA: SciTech, 2010, ch. 3, pp. 87–113.

[3] K. White, J. Williams, and P. Hoffensetz, “Radar sensor management for detection and tracking,” in *2008 11th Int. Conf. Info. Fusion*, Cologne, Germany, Jun. 30–Jul. 3, 2008, pp. 1–8.

[4] J. L. Williams, “Search theory approaches to radar resource allocation,” in *Proc. 7th US/Aust. Joint Workshop Defense Appl. Signal Process. (DASP)*, Coolum, Australia. 2011.

[5] M. Byrne, K. White, and J. Williams, “Scheduling multifunction radar for search and tracking,” in *2015 18th Int. Conf. Inf. Fusion (Fusion)*, Washington, DC, USA, Jul. 6–9, 2015, pp. 945–952.

[6] M. Byrne, K. White, and J. Williams, “Rolling horizon non-myopic scheduling of multifunction radar for search and track,” in *2016 19th Int. Conf. Inf. Fusion (FUSION)*, Heidelberg, Germany, Jul. 5–8, 2016, pp. 1–8.

[7] A. De Freitas, R. W. Focke, and P. De Villiers, “Response surface modeling for networked radar resource allocation,” in *2018 21st Int. Conf. Inf. Fusion (FUSION)*, Cambridge, UK, Jul. 10–13, 2018, pp. 1–7, doi: 10.23919/ICIF.2018.8455815.

[8] A. De Freitas, R. W. Focke, C. Beyers, and P. De Villiers, “Global optimization for resource allocation in a networked-surveillance radar,” *2019 22th Int. Conf. Inf. Fusion (FUSION)*, Ottawa, ON, Canada, Jul. 2–5, 2019, pp. 1–8, doi: 10.23919/FUSION43075.2019.9011303.

[9] S. Flandermeyer and J. Metcalf, “A particle swarm optimization approach to surveillance resource management,” in *2023 IEEE Int. Radar Conf. (RADAR)*, Sydney, Australia, Nov. 6–10, 2023, pp. 1–6, doi: 10.1109/RADAR54928.2023.10371057.

[10] R. Mahler and T. Zajic, “Probabilistic objective functions for sensor management,” in *Signal Process., Sensor Fusion, Target Recognit. XIII*, vol. 5429, I. Kadar, Ed., Orlando, FL, USA, Aug. 9, 2004, pp. 233–244, doi: 10.1117/12.543530.

[11] A. El-Fallah et al., “Multisensor-multitarget sensor management with target preference,” in *Signal Process., Sensor Fusion, Target Recognit. XIII*, vol. 5429, I. Kadar, Ed., Orlando, FL, USA, Aug. 9, 2004, pp. 222–232, doi: 10.1117/12.543822.

[12] R. A. Romero and N. A. Goodman, “Adaptive beamsteering for search-and-track application with cognitive radar network,” in *2011 IEEE RadarCon (RADAR)*, Kansas City, MO, USA, May 23–27, 2011, pp. 1091–1095, doi: 10.1109/RADAR.2011.5960703.

[13] J. D. Morgan and B. D. Rigling, “Surveillance performance of digital radars in non-uniform target behavior,” in *2023 IEEE Radar Conf. (RadarConf23)*, San Antonio, TX, USA, May 1–5, 2023, pp. 1–6, doi: 10.1109/RadarConf2351548.2023.10149742.

[14] A. Hebb, R. S. Adve, Z. Qu, and Z. Ding, “Belief-rewards to learn radar scheduling policies,” in *2024 IEEE Radar Conf. (RadarConf24)*, Denver, CO, USA, May 6–10, 2024, pp. 1–6, doi: 10.1109/RadarConf2458775.2024.10548136.

[15] E. Krijestorac et al., “Machine learning-assisted computationally efficient target detection and tracking in massive fully digital phased arrays,” *IEEE Trans. Radar Systems*, vol. 1, pp. 353–367, 2023, doi: 10.1109/TRS.2023.3298340.