# Experimental Validation of Cognitive Radar Anticipation using Stochastic Control

Colin Horne, Matthew Ritchie, Hugh Griffiths
Department of Electronic and Electrical Engineering
University College London
London, UK
Email: colin.horne.14@ucl.ac.uk

Folker Hoffmann, Alex Charlish
Sensor Data and Information Fusion Department
Fraunhofer Institute for Communications,
Information Processing and Ergonomics FKIE
53343, Wachtberg, Germany

***Abstract* — Cognition applied to radar systems is a growing area of research. The majority of cognitive radar research is focused on theory and simulation with little experimental validation. Prior research proposed the application of anticipation within a cognitive radar, demonstrating by simulation that this can provide significant improvements in tracking performance, when compared to non-cognitive radar tracking methods. The approach applied a POMDP algorithm to control the timing of track updates for a target while anticipating the loss of measurements within a known time period/region. This work aims to expand on this concept by using data from a real radar, NetRAD, in order to validate the application of anticipation when tracking a human target.**

***Keywords— Cognitive Radar; Tracking; Anticipation.***

## I. Introduction

Cognitive radar is a relatively new field of research, which has received significant attention in recent years. This interest is due to the potential benefits offered by the addition of cognitive capabilities to radar systems, enabling those systems to dynamically exploit the advanced radar hardware and signal processing now available. The potential benefits include, but are not limited to, resilience to RF interference, interoperability of systems via intelligent spectrum sharing and more efficient use of available resources, and are a direct consequence of the increased adaptability of the overall system given the addition of flexibility in the transmitter subsystem.

The fundamental concept underlying cognitive radar is the capability to sense an environment, learn salient features of the environment, and to intelligently adapt the behavior of the sensor and/or the supporting platform, possibly in terms of intra-pulse characteristics of the transmitted waveform, transmit antenna beamshape, the target revisit or dwell time, or the location of the sensor, and then to re-sense the environment with the aim of improving performance cycle by cycle. As part of this processing both short-term online memory and longer term potentially offline prior knowledge is utilized. This simple perception-action cycle which was first applied to radar systems by Haykin [1], is at the heart of all cognitive radar research, providing a closed loop feedback system round the transmitter, environment and receiver.

Very little research has been published which experimentally validates the benefits of cognitive radar. One reason for this, at least if a real-time demonstration of cognition is the goal, are the requirements for very capable hardware platforms offering flexible control of radar operating parameters, and the computing power to exploit the capabilities. Developments over the last several years provide for both requirements with systems such as the Ohio State University (OSU) radar testbed CREW [2] becoming available. The work reported in [2] demonstrates encouraging results in applying highly flexible RF and processing capabilities to the real-time cognitive radar problem.

The purpose of the work described in this paper is to experimentally validate the simulation results reported in [3] concerning the improvement in tracking performance achieved when a known obstruction between sensor and target is anticipated such that extra resources can be expended on improving track accuracy prior to the loss of returns. The experimental layout is illustrated in Fig 1. The improved track accuracy, prior to entering the obscuration, allows for better track prediction during the obscuration with the aim of reacquiring the target, without track loss, on the termination of the obstruction. Fundemental metrics of track performance are successful track association and reduction of loss of track. This cognitive processing technique allows for improvement of both of these.

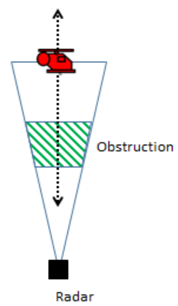


**Figure 1:** Experimental setup

## II. Anticipation

A number of characteristics/properties have been identified [1] as being required for a radar to be considered as cognitive. These include memory, intelligence, attention and the perception-action cycle. Learning is also central to the concept of cognition. The capability of anticipation has recently been proposed as a cognitive ability as argued in [3]. Anticipation can be regarded as the use of experience, or previously gained knowledge, plus a perception of current conditions, to predict a future condition or situation. The work reported in [3] demonstrates the theoretical benefit gained from using anticipation within a cognitive radar tracking application, and this research paper aims to provide experimental validation of this concept.

The premise of the work described in [3] is that the knowledge of the future loss of radar performance at a known time (or location) might allow a cognitive radar system to anticipate the consequences of that loss, and for it to compensate in some way in advance of the loss of target detection. One possible scenario is illustrated in Figure 2. In this scenario an electronically steered radar on the sensor platform tracks a target. The loss of target detections is caused by the target passing behind a physical obstruction at a known point in time, for a known period of time. Other scenarios might be envisaged such as the reallocation of the sensor to alternative tasks such as SAR imaging, or weapons control. A further possible scenario is the sensor is a secondary user of the RF spectrum, and must yield the spectrum to the primary function. The challenge is to modify the operation of the radar sensor prior to the loss of detections, such that satisfactory tracking performance is maintained through the period of obstruction at the least possible cost in terms of sensor resourse usage. In this case the revisit interval used by the sensor to track the target is the controlled radar parameter which allows the track quality to be controlled.

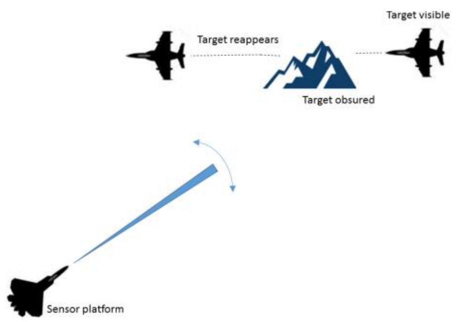


**Figure 2:** Obscuration scenario

The cognitive function of anticipation is realized in [3] by using a rollout based approach [4] to approximate the solution to a partially observable Markov decision process (POMDP), which is a framework for sequential decision making in an uncertain environment. It differs from a Markov decision process (MDP) in that the true state of the system is not fully observable, any measurements of the state being subject to corruption by noise. The POMDP solution attempts to maximise the value of some reward function over the period of assessment by selecting system actions (revisit intervals) which provide the highest probability of high reward balanced against the total resource use.

In this rollout based method the future evolution of a system is approximated by assuming the controller follows a base policy, which enables the future expected reward to be approximated. An approximation of reward is justified as the actual value of the reward is less important than the relative ranking of rewards for the selection of the optimum actions.

This paper continues the use of the POMDP approach to anticipation, but differs from [3] in that the POMDP solution is approximated using the Monte Carlo Tree Search (MCTS) algorithm in place of the policy rollout approach. MCTS has been used with a very high degree of success in applications such as the game Go [5]. MCTS is a Monte Carlo simulation approach in which the simulation is focused on areas where good levels of reward have previously been observed, while still allowing unexplored regions of the search space to be visited in search of more profitable reward paths. In this work the assessment is based on receeding horizon control, in which a limited assessment horizon length is used.

The reward achieved within this POMDP is based on the combination of the predicted positional root mean square error (RMSE), by way of a utility function, and the resource expended to achieve that result. At each decision point, the controller selects the action that provides the largest expected future reward. The POMDP is described further in Section V.

## III. RADAR SYSTEM AND EXPERIMENTS

The nature of the experimental radar system used to capture the data, and therefore the nature of the available data for processing is of fundamental importance, so will be described next.

The radar platform used as part of this research is the University College London (UCL) developed NetRAD multistatic radar system. This is an experimental coherent pulse radar system operating in S-Band at 2.4 GHz. The data collection trials were carried out during July 2015, with the system deployed in an open sports field to the north of London. The data capture was performed using the commonly employed NetRAD radar parameters of a 600 ns pulse width, with a 45 MHz up-ramp chirp. Antennas with 10 degree beam widths in elevation and azimuth were used, these having a gain of 24 dBi, and configured for horizontal polarization. The system is often employed with a PRF of 1 kHz and collecting many range samples, however, due to the restricted range extent available on the experiment site, data was recorded over a limited range, specifically 256 range bins sampled at 100 MHz, and at 5 kHz PRF. These parameters resulted in the ability to record data over an extended period of 2 mins. The 5 kHz PRF provides for the capture of returns from a large number of pulses such that various processing regimes might be implemented in post processing using a subset of pulses, by decimating the original complete dataset to a reduced sample rate equivalent. This offers the ability, for example, for the net PRF and numbers of pulses integrated to be selected based on cognitive algorithms used.

It is important to note that NetRAD has fixed antennas which point in a fixed direction, so no bearing information is available. The lack of bearing information means that the scenario described in [3] can be only partially replicated by the captured real radar data. The selection of revisit interval will be made on the basis of range-only measurements.


**Figure 3:** Trials Site Layout

The experimental layout showing the three NetRAD radar nodes used to capture data is illustrated in Figure 3. The green pin shows the nominal minimum range of the target. The target traverses a path between this point and a point short of the tree line. Data was gathered at all three NetRAD nodes allowing for multistatic processing to be considered within the intersection region of the beams, although the work reported in this paper only employs monostatic data. The target must stay within the fixed antenna beam width for it to be visible in the captured data.

Another restriction imposed by the use of NetRAD is that there is no facility to change any radar parameter within a single data capture period. All parameters are pre-set for use in a trials run, and can only be changed after data capture is complete. Cognitive capabilities can therefore only be shown in off-line post processing.

No physical obstruction is included in the captured data. A virtual loss of detections was simulated within the processing software framework, which allows the loss of returns to be configured in terms of duration and time as desired.

An example Range Time Intensity (RTI) plot from a human target moving towards/away from the sensor can be seen in Figure 4. The figure shows 600,000 PRIs of data post Hilbert transform and pulse compression. The target starts at short range (~range bin 38), walks to long range (~range bin 90) and returns. The constant line at range bin 16 is the direct coupling between the transmit and receive antennas, and the signals at ranges 101-105 were produced by returns from the tree line at the end of the field.

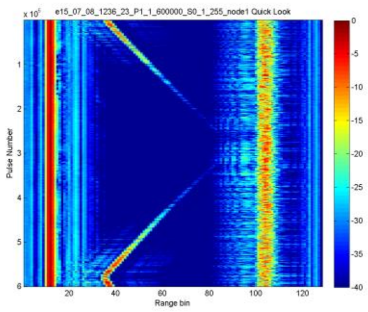

**Figure 4:** RTI plot of walking human target

## IV. TRACKING

The target is tracked in range and range rate using the monostatic measurement information. A Kalman filter is used to perform the tracking function, using a continuous white noise acceleration motion model. Track performance is assessed using the filter calculated position uncertainty, in this one dimensional case, this equates to the square root of the range element of the error covariance matrix.

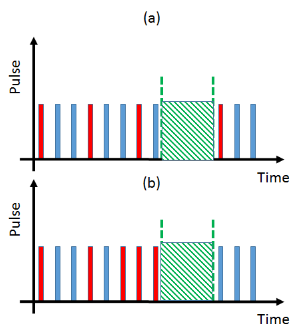


**Figure 5:** Radar pulses

Figure 5(a) illustrates the track revisit behaviour without anticipation where the red pulses represent tracking pulses, and blue the potential pulses not employed, and with anticipation Figure 5(b) where extra pulses are used to improve the track accuracy prior to the obstruction, shown as the green hashed area where no target returns are available.

The effects of these two cases on track RMSE is shown within Figure 6. The graphs illustrate the evolution of the track error over time. During the obscuration, marked by the vertical dashed lines, the track error covariance becomes large and the track is dropped Figure 6(a). However if the track gets sufficiently sharpened before the obscuration, the covariance stays below the limit for dropping the track, allowing the track to be maintained Figure 6(b).

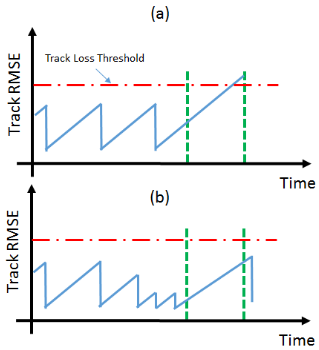


**Figure 6:** Expected track RMSE
(a) Without Anticipation (b) With Anticipation

## V. SIGNAL AND COGNITIVE PROCESSING

### *A.* Target Detection

Target detection is achieved using the radar signal processing chain shown in Figure 7. Hanning windowing is applied to the data prior to a 256 point FFT being carried out on each range cell. Each Doppler filter channel has CA-CFAR processing applied followed by second thresholding and centroid detection to isolate any detection to a single range bin.

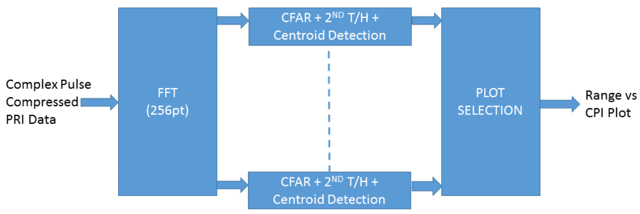


**Figure 7:** Signal Processing Chain

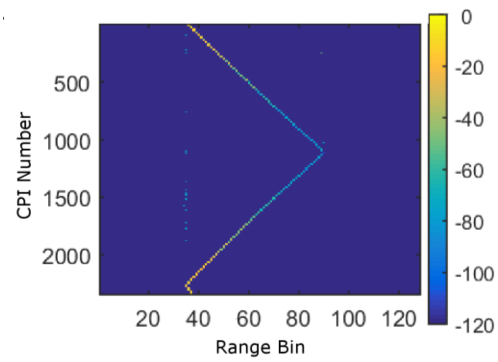


**Figure 8:** RTI - post signal processing chain

Figure 8 shows the RTI corresponding to the detections realised by the signal processing chain when applied to the data from Figure 4.

## B. Partially Observable Markov Decision Process

The use of a POMDP problem formulation as an effective tool for anticipation of the consequences of actions taken by a system is shown in [3] via a simulated tracking example.

The essence of the POMDP is to select an action to be taken which maximizes the reward which could be achieved over some horizon given that optimal actions are selected over the remainder of the horizon. The reward may be defined as:

$$V_H = E\left[\sum_{k=0}^{H-1} R(x_k, a_k)\right] \tag{1}$$

where $V_H$ = reward
$E[\cdot]$ = expectation operator
$H$ = horizon
$R(x_k, a_k)$ = reward function given state $x_k$ and action $a_k$ at time $k$

However, in a POMDP the state isn't known, and instead the reward must be based on a belief state, which is a probability distribution over all possible model states. The reward R(·) based on the state must be modified for use in the belief state:

$$r(b,a) = \sum_{x \in X} b(x) R(x, a) \tag{2}$$

where $r(b,a)$ = reward in belief state $b$ given action $a$
$b(x)$ = the probability belief state $b$ assigns to state $x$

In this work the belief state is estimated by the Kalman filter tracker, and the reward is based on a utility function of the predicted range error calculated by the filter, in a similar arrangement to [3]. The utility function is:

$$u(P_{k+1|k}) = \begin{cases} 0.0 & if\ P_{k+1|k} \geq 1.0 \\ 1.0 & if\ P_{k+1|k} \leq 0.3 \\ \left(\dfrac{1-\sqrt{P_{k+1|k}}/v}{0.7}\right)^{\eta} & otherwise \end{cases} \tag{3}$$

where $P_{k+1|k}$ is the filter predicted mean square error
$v$ is a constant
$\eta$ is a sensitivity parameter

As in [3], the reward available in belief state $b_k$ is then a function of the utility and the amount of resource required to generate this utility:

$$r(b,a) = \frac{u(P_{k+1|k}) \cdot t_r}{r_l} \tag{4}$$

where $t_r$ is the revisit interval
$r_l$ is the resource loading

The formulation of the problem is based on [3], where a more complete derivation can be found.

## C. Monte Carlo Tree Search

The solution method selected for this work is the MCTS. MCTS differs from pure Monte Carlo methods, using just a random selection of next action, by attempting to balance the exploitation-exploration choice. The MCTS focuses attention on paths which appear likely to offer high rewards, whilst still allowing alternative paths to be explored which might hold the potential for greater rewards. The mechanism used is called Upper Confidence Bounds on Trees (UCT), which selects which action to investigate based on the maximum value of the following formula:

$$X_i + C\sqrt{\frac{\log(N)}{n_i}} \tag{5}$$

where $X_i$ is the reward currently available from action i
$N$ is the total number of simulations run
$n_i$ is the number of simulations selecting action i
$C$ determines the exploration/exploitation balance

The second element of Eqn. (5) increases for action i as the total number of simulations increases with action i not selected, such that after sufficient simulations the action i will be investigated again.

A significant characteristic of MCTS is that it is an 'anytime' algorithm, meaning that a good solution is achieved even if the search is shortened.

# VI. RESULTS

This section will describe the parameters chosen within the processing and the results obtained.

Revisit times of between 1s and 6s, with 1s intervals
Receeding horizon of 10s
Process noise standard deviation = 0.5m/s
Measurement noise standard deviation = 0.95m
Obscuration duration = 5s
Obscuration points = Case 1 : 19s & Case 2 : 22s

The measurement noise standard deviation can be calculated [7]:

$$\sigma_R = \frac{\Delta R}{k\sqrt{(2 * SNR)}} \tag{6}$$

where $\Delta R$ is the range bin size
$k$ is a constant
$SNR$ is the signal to noise ratio

The value used is based on a 13 dB SNR. In the radar data the SNR varies with range and for the most part is significantly higher than 13 dB. A future development of this work will be to control the SNR by selecting an appropriate integration time.

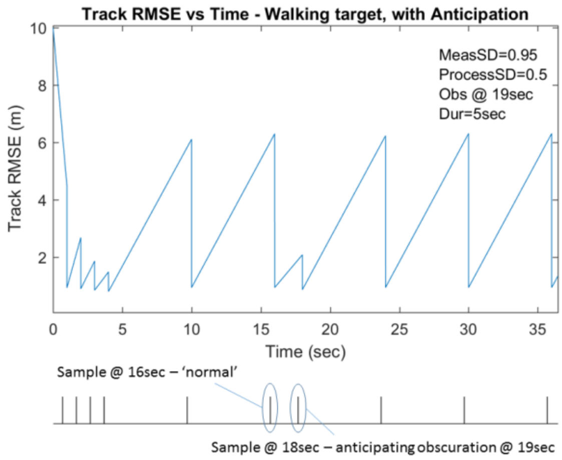


**Figure 9:** Track performance – Anticipation (19sec)

Figures 9-11 show the track errors calculated by the Kalman filter (top plot in each figure) and the sample points selected (lower plot) for various configurations. Figure 9 and Figure 10 use an obscuration point of 19 & 22 seconds respectively, each with a duration of 5 seconds, using the anticipation processing.

The steady state condition is the maximum revisit time of 6 seconds.

The result of anticipating the obscuration is that a sample is taken just prior to the detection loss in each case, at 18s and 21s respectively. The effect seen in [3] whereby a number a samples are taken just prior to the obscuration is not evident as a single sample reduces the tracking error to a suitably small value due to the small measurement noise in this range-only scenario.

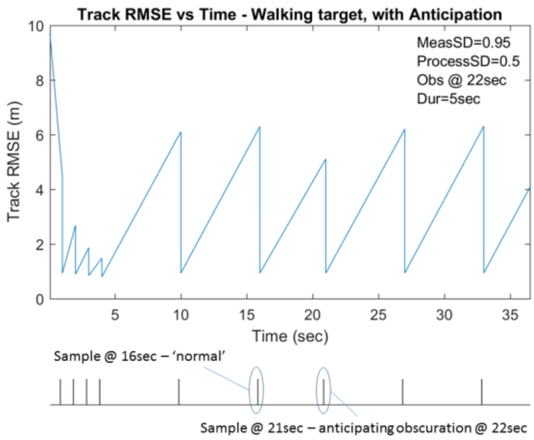


**Figure 10:** Track performance – Anticipation (22sec)

Figure 11 shows the case where no anticipation is employed. The missing sample at 19sec allows the tracking error to increase beyond the usual values, and could result in track loss. The second peak is due to actual missing detections for which the addition of anticipation could not provide any improvement.

## VII. Discussion and Future Work

The use of range-only measurements as provided by the NetRAD data capture restricts the verification of the technique described in [3] as range errors are by the nature of the system very small. The use of a radar system capable of providing range and bearing measurements would afford a more complete investigation.

With the data collected, only off-line processing is possible. The ability to modify radar operating parameters in real time, possibly between successive pulses or pulse bursts will provide the ability to demonstrate action selection 'on-the-fly'. UCL's latest experimental radar system, NeXtRAD will provide such facilities. Initial monostatic trials using the new system will be conducted towards the end of 2016 in South Africa.

Further extensions to this work include control of signal to noise ratio by varying the number of pulses integrated at the revisit point, and modification of the measurement noise standard deviation based on the SNR. As well as data from human walking targets, data was also captured from flights of a small drone and of a car traversing the course. Processing of this data is also planned.

The control achieved using the 'cognitive' approach could equally have been generated by hard coding the revisit time response, or by a rule based approach, however, a balance between perfomance and resource usage has been demonstrated which provides a step towards verifying more complex cognititve systems.

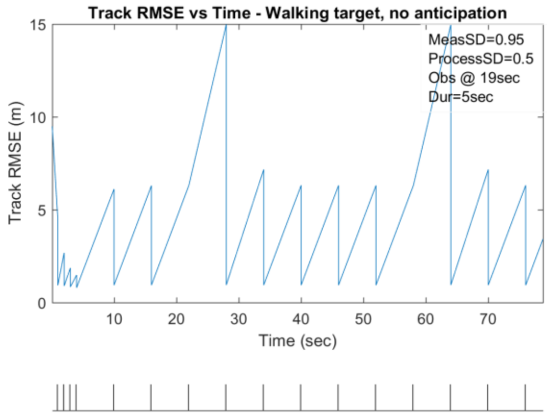


**Figure 11:** Track performance – No Anticipation

The scenario employed in this work relies on the availability of prior information regarding the timing of the obscuration, what might be called the knowledge aided approach. Scenarios could be made more 'cognitive' given learning of where the obscuration occurs over multiple targets or multiple missions.

## VIII. Conclusions

In this paper the cognitive characteristic of anticipation has been investigated using real radar data from the UCL NetRAD system, providing range-only measurements. Monte Carlo Tree Search techniques have been employed to solve a POMDP for the control of the tracking revisit interval in a scenario where anticipating the loss of radar coverage due to an impending obstruction allows the system to compensate such that tracking performance may be maintained despite the signal loss.

## Acknowledgments

This work has been funded by the EPSRC, Dstl and the IET A F Harvey Prize awarded to Prof. Hugh Griffiths in 2013.

## References

[1] S. Haykin, "Cognitive Radar [A way of the future]" IEEE Signal Processing Magazine, vol. 23, January, pp. 30-40, 2006

[2] G. Smith, Z. Cammenga, A. Mitchell, K. Bell, M. Rangaswamy, J. Johnson and C. Baker, "Experiments with cognitive radar, "*2015 IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, 2015.

[3] A. Charlish and F. Hoffmann, "Anticipation in cognitive radar using stochastic control," *2015 IEEE Radar Conference, Arlington, VA, pp. 1692-1697, 2015.*

[4] E. K. P. Chong, C. M. Kreucher, and A. O. Hero, "POMDP Approximation Using Simulation And Heuristics," in *Foundations and Applications of Sensor Management*, 2008, pp. 95 – 119.

[5] S Gelly, L Kocsis, M Schoenauer, M Sebag, D Silver, C Szepesvári and O Teytaud, The grand challenge of computer Go: Monte Carlo tree search and extensions. *Communications of the ACM*, *55*(3), 2012, 106-113.

[6] L. Kocsis and C. Szepesvari. Bandit based Monte-Carlo planning. *15th European Conference on Machine Learning (ECML), pages 282–293, 2006*

*[7] S* Blackman and R Popoli (1999). *Design and Analysis of Modern Tracking Systems.* Norwood: Artech House