EURASIP Journal on
Advances in Signal Processing
a SpringerOpen Journal

RESEARCH   **Open Access**

CrossMark

A novel resource scheduling method of netted radars based on Markov decision process during target tracking in clutter

Zhenkai Zhang[1,2*] and Yubo Tian[1]

Abstract

In order to improve the radio frequency stealth ability of phased array radars, a novel resource scheduling method of the radar network for target tracking in clutter is presented. Firstly, the relationship model between radar resource and tracking accuracy is built, and the sampling interval, power, and waveform will influence predicted error covariance matrix through transition matrix and measurement noise. Then, radar resource scheduling algorithm based on Markov decision process which is converted to be a binary optimization problem is proposed, and an improved binary wind-driven optimization method is presented to solve that problem. The radar and its radiation parameters will be selected for better radio frequency stealth performance and tracking accuracy. Simulation results show that the proposed algorithm not only has excellent tracking accuracy in clutter but also has better radio frequency stealth ability comparing with other methods.

Keywords: Power control, Radio frequency stealth, Target tracking, Sampling interval, Waveform selection

1 Introduction

Low probability intercept (LPI) is one of the important features of modern radars. In order to improve the LPI performance, we need to reduce not only the radar cross section (RCS) but also the radiation of the radars which is also called radio frequency stealth (RFS) [1]. In order to achieve the important tactical requirement of RFS, it is necessary to dynamically manage the radar resource, such as sampling interval, power, and waveform. As we know, larger sampling interval will lead to less radiation times, and lower power and waveform agility will bring better performance of RFS. The research of [2] considers an advanced pulse compression noise radar waveform possessing salient features and noise waveforms and demonstrates the RFS characteristic of the waveform with different κ values. The work in [3] evaluated the mutual interference and low probability of interception capabilities of noise waveforms. The research in [2, 3] focuses on the waveform feature without the consideration of radar efficiency. A novel adaptive sampling interval

algorithm is presented in [4], which optimizes the sampling interval based on particle swarm optimization in order to obtain excellent performance for phased array radar. For better tracking performance instead of RFS ability in [4], an adaptive maneuvering target tracking algorithm for Doppler radar is proposed in [5], where both the sampling interval and transmitting waveform can be adjusted according to the predicted error covariance output adaptively. Both the works in [4, 5] design the radiation parameters only for the single radar. For multiple networked radars, the efficient dynamic power control and management algorithms are presented in [6], and optimal sensor management algorithms are developed for controlling the active sensor emission to minimize the threat posed to all the platforms. Another LPI optimization strategy is proposed for target tracking in radar network architectures in [7, 8], where Schleher intercept factor is minimized by optimizing transmission power allocation among netted phased array radars in the network. However, almost all of those works do not consider the scheduling and control method of combined parameters, such as radar power, sampling interval, and waveform parameters. Moreover, it is impossible

* Correspondence: zzk@nuaa.edu.cn
[1]College of Electronic Information, Jiangsu University of Science and Technology, Zhenjiang, People's Republic of China
[2]Centre for Communication Systems, Durham University, Durham, UK

for phased array radars to schedule such parameters in an infinite scope.

For the resource allocation of phased array radars in the network, this paper will present a novel scheduling method of sampling interval, power, and waveform parameters in a limited library during tracking in clutter. The remainder of this paper is organized as follows. Section 2 introduces the interacting multiple model probability data association algorithm. Section 3 presents the relationship model between radar resource and tracking accuracy. The proposed radar resource scheduling algorithm based on Markov decision process is explained in Section 4. Simulations of the proposed algorithms and comparison results with other methods are provided in Section 5. The conclusions are presented in Section 6.

2 Interacting multiple model probability data association

The proposed resource scheduling method will be used during the tracking process in clutter. As we know, the interacting multiple model (IMM) algorithm has demonstrated its ability to catch up the unknown maneuver, while the probability data association (PDA) approach has shown its relative easy-implementation and good performance properties in the target tracking community [9, 10]. The paper prefers to combine them together and use IMMPDA for the target tracking in clutter.

Let x_k and z_k represent the state vector and the observation vector, respectively; the state equation and transfer equation at time k are:

$$x_{k+1} = Fx_k + w_k \qquad (1)$$

$$z_{k+1} = Hx_k + v_k \qquad (2)$$

The process noise w_k and the measurement noise v_k are mutually uncorrelated zero-mean white Gaussian processes with covariance matrices Q and B_k, F is the transition matrix, and H is the observation matrix. The IMMPDA algorithm includes the following steps.

(1) Input interaction

The initial state for prediction of each model is a mixture of the states from the last cycle of all models with the mixing probabilities:

$$\mu_{k-1|k-1}^{s|t} = \frac{\theta^{s|t}\mu_{k-1}^s}{\sum\limits_{s=1}^{N} \theta^{s|t}\mu_{k-1}^s} \qquad (3)$$

where μ_{k-1}^s is the model probability of the model s in the last cycle, $\theta^{s|t}$ is the probability for the transition from

model s to model t, and $\sum\limits_{s=1}^{N} \theta^{s|t} = 1$, N is the model numbers. The initial state $\hat{x}_{k-1|k-1}^{0t}$ and covariance $P_{k-1|k-1}^{0t}$ of the model t are:

$$\hat{x}_{k-1|k-1}^{0t} = \sum\limits_{s=1}^{N} \hat{x}_{k-1|k-1}^s \mu_{k-1|k-1}^{s|t} \qquad (4)$$

$$P_{k-1|k-1}^{0t} = \sum\limits_{s=1}^{N} \mu_{k-1|k-1}^{s|t} \left[\hat{x}_{k-1|k-1}^s - \hat{x}_{k-1|k-1}^{ot}\right]\left[\hat{x}_{k-1|k-1}^s - \hat{x}_{k-1|k-1}^{ot}\right]^T$$
$$+ \sum\limits_{s=1}^{N} \mu_{k-1|k-1}^{s|t} P_{k-1|k-1}^s \qquad (5)$$

where $\hat{x}_{k-1|k-1}^s$ and $P_{k-1|k-1}^s$ are the state and covariance of the model s of the last cycle, respectively.

(2) Measurement validation

The validated measurement z_k at time should meet the condition as:

$$\left[z_k - \hat{z}_{k|k-1}\right](S_k)^{-1}\left[z_k - \hat{z}_{k|k-1}\right]^T < \gamma \qquad (6)$$

where $\hat{z}_{k|k-1} = \sum\limits_{t=1}^{N}\sum\limits_{s=1}^{N} \theta^{s|t}\mu_{k-1}^s HF^t\hat{x}_{k-1|k-1}^s$, S_k is the associated covariance matrix, and γ is the validation gate which represents a threshold that is associated with the acceptability of the measurements.

(3) Model filtering

At the prediction step, the state and covariance are predicted using the estimate of the previous step. The predicted state and covariance of model t at time k can be represented as:

$$\hat{x}_{k|k-1}^t = F^t\hat{x}_{k-1|k-1}^{0t} \qquad (7)$$

$$P_{k|k-1}^t = F^t P_{k-1|k-1}^{0t}(F^t)^T + Q \qquad (8)$$

The updated state and covariance with the ith measurement can be represented as:

$$\hat{x}_{i,k|k}^t = \hat{x}_{k|k-1}^t + G_k^t r_{i,k}^t \qquad (9)$$

$$P_{i,k|k}^t = P_{i,k|k-1}^t + G_k^t HP_{i,k|k-1}^t \qquad (10)$$

The innovation covariance S_k^t and Kalman gain G_k^t of model t at time k is represented as:

$$S_k^t = HP_{k|k-1}^t H^T + B_k \qquad (11)$$

$$G_k^t = P_{k|k-1}^t H^T \left(S_k^t\right)^{-1} \tag{12}$$

where $r_{i,k}^t = z_{i,k} - \hat{z}_{k|k-1}^t$ and $z_{i,k}$ is the ith measurement.

(4) Measurement probability computing

$\beta_{i,k}^t$ is the probability of event during which $z_{i,k}$ is the correct measurement from the target. $\beta_{0,k}^t$ is the probability of event during which none of the validated measurements is corrected. The association probabilities of $\beta_{i,k}^t$ and $\beta_{0,k}^t$ are

$$\beta_{i,k}^t = \frac{\exp\left(-r_{i,k}^t \left(S_k^t\right)^{-1} r_{i,k}^t / 2\right)}{b_k + \sum\limits_{i=1}^{W_k} \exp\left(-r_{i,k}^t \left(S_k^t\right)^{-1} r_{i,k}^t / 2\right)} \tag{13}$$

$$\beta_{0,k}^t = \frac{b_k}{b_k + \sum\limits_{i=1}^{W_k} \exp\left(-r_{i,k}^t \left(S_k^t\right)^{-1} r_{i,k}^t / 2\right)} \tag{14}$$

where W_k is the number of validated measurements at time k, $b_k = W_k (1 - P_{D,k|k-1})(P_{D,k|k-1} V_k)^{-1}$, $P_{D,k|k-1}$ is the detection probability, and V_k represents the validation region.

(5) Update of the model probability

The likelihood function Λ_k^t corresponding to the model t is

$$\Lambda_k^t = \frac{P_{D,k|k-1}^t}{W_k P_{D,k|k-1}^t + \left(1 - P_{D,k|k-1}^t\right) V_k} V_k^{-W_k+1} \left[\frac{b_k}{\left|2\pi S_k^t\right|^{1/2}} + \sum_{i=1}^{W_k} \rho\left[r_{i,k}^t; 0, S_k^t\right]\right] \tag{15}$$

where $\rho\left[r_{i,k}^t; 0, S_k^t\right] = \frac{1}{\left|2\pi S_k^t\right|^{1/2}} \exp\left(-\frac{1}{2}\left(r_{i,k}^t - 0\right)' S_k^t \left(r_{i,k}^t - 0\right)\right)$.

The updated model probability is

$$\mu_k^t = \frac{1}{\bar{c}} \Lambda_k^t \sum_{s=1}^{N} \theta^{s|t} \mu_{k-1}^s \tag{16}$$

where $\bar{c} = \sum\limits_{t=1}^{N} \mu_k^t$ is the normalizing constant.

(6) PDA filtering of every model

The state and covariance estimate of model t can be written as:

$$\hat{x}_{k|k}^t = \sum_{i=1}^{W_k} \beta_{i,k}^t \hat{x}_{i,k|k}^t + \beta_{0,k}^t \hat{x}_{k|k-1}^t \tag{17}$$

$$P_{k|k}^t = P_k^t - \left(1 - \beta_{0,k}^t\right)\left(I - P_{D,k|k-1}^t G_k^t H\right) P_{k|k-1}^t \tag{18}$$
$$+ \beta_{0,k}^t P_{k|k-1}^t$$

where $P_k^t = G_k^t \left(\sum\limits_{i=1}^{W_k} \beta_{i,k}^t r_{i,k}^t \left(r_{i,k}^t\right)^T - r_k^t (r_k^t)^T\right)(G_k^t)^T$, $r_k^t = \sum\limits_{i=1}^{W_k} r_{i,k}^t$.

(7) Estimate and covariance combination

The final estimate $\hat{x}_{k|k}$ of the state and covariance $P_{k|k}$ can be represented as:

$$\hat{x}_{k|k} = \sum_{t=1}^{N} \hat{x}_{k|k}^t \mu_k^t \tag{19}$$

$$P_{k|k} = \sum_{t=1}^{N} \mu_k^t \left\{ P_{k|k}^t + \left[\hat{x}_{k|k}^t - \hat{x}_{k|k}\right]\left[\hat{x}_{k|k}^t - \hat{x}_{k|k}\right]^T \right\} \tag{20}$$

The radar and its radiation parameters will be selected during target tracking. So IMMPDA in this section is used to tracking the target in clutter. And the proposed radar resource scheduling method is proposed based on IMMPDA.

3 Relationship model between radar resource and tracking accuracy

In this paper, the radar resource includes the sampling interval, radiated power, the pulse width, and carrier frequency of the transmitted waveform. All of these parameters have an impact on the tracking accuracy, which can be seen in Fig. 1. The sampling interval has an effect on the computing of the transition matrix. The detection
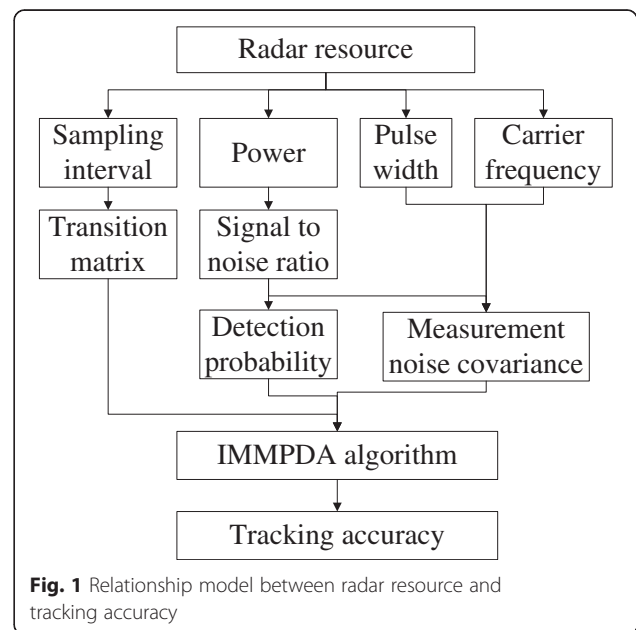
Fig. 1 Relationship model between radar resource and tracking accuracy

probability and measurement noise are under the influence of power, pulse width, and carrier frequency of the transmitted signal. In addition, the transition matrix and measurement noise covariance will decide the tracking accuracy of IMMPDA algorithm in section 2.

3.1 Transition matrix
In the IMMPDA algorithm, the motion model includes constant velocity model and coordinated turn rate model, and transmission matrix can be represented as Eqs. (21) and (22), respectively.

$$F_{CV} = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{21}$$

$$F_{CT} = \begin{bmatrix} 1 & \dfrac{\sin\omega T}{\omega} & 0 & \dfrac{1-\cos\omega T}{\omega} \\ 0 & \cos\omega T & 0 & \dfrac{\sin\omega T}{\omega} \\ 0 & \dfrac{1-\cos\omega T}{\omega} & 1 & \dfrac{\sin\omega T}{\omega} \\ 0 & \sin(\omega T) & 0 & \cos\omega T \end{bmatrix} \tag{22}$$

where T is the sampling interval and ω is the turn factor.

3.2 Signal to noise ratio
The overall netted radar sensitivity can be calculated by summing up the partial signal to noise ratio (SNR) [11] of each transmitter-receiver pair (assuming all signals can be separately distinguished at each receiver), which is given by

$$S_{NR\,netted} = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{P_{t_i} G_{t_i} G_{r_j} \sigma_{ij} \lambda_i^2 T_{dw}}{(4\pi)^3 k_B T_{sij} R_{ti}^2 R_{rj}^2 N_{F_j} L_{ij}} \tag{23}$$

where the thermal noise at each receiver is assumed to be statistically independent, P_{t_i} is the ith peak transmitted power, G_{t_i} and G_{r_j} are the ith transmitter gain and jth receiver gain, σ_{ij} of the target means the ith transmitter and jth receiver which is assumed to be known in our library, λ_i is the ith transmitted wave-length, T_{dw} and k_B are target integration time and Boltzmann's constant, respectively, T_{sij} is the receiving system noise temperature (at a particular receiver), N_{F_j} represents the noise figure at each receiver, L_{ij} is the system loss for ith transmitter and jth receiver, and R_{ti} and R_{rj} are the distance from ith transmitter to the target and distance from target to jth receiver.

Then, the detection probability can be represented as:

$$P_d = \exp\left(\frac{\ln P_{fa}}{1 + S_{NR\,netted}}\right) \tag{24}$$

where P_{fa} is false-alarm probability. The detection probability P_d has an important impact on the tracking performance in the clutter, which is the function of radiated power and target RCS and range.

3.3 Covariance of measurement noise
The covariance matrix B_k of measurement noise is controlled by the radiated power and waveform parameters of the radar during tracking. We assume all the radars in the network transmit the same type of waveform. Range and range-rate measurements are obtained using the type of linear frequency modulated (LFM) Gaussian pulses. The measurement noise covariance is given by:

$$B_k = \begin{bmatrix} \dfrac{c^2 p_u{}^2}{2 S_{NR}{}^k{}_{netted}} & -\dfrac{c^2 b p_u{}^2}{w_0 S_{NR}{}^k{}_{netted}} \\ -\dfrac{c^2 b p_u{}^2}{w_0 S_{NR}{}^k{}_{netted}} & \dfrac{c^2 b p_u{}^2}{w_0 S_{NR}{}^k{}_{netted}} \left(\dfrac{1}{2 p_u{}^2} + 2 b^2 p_u{}^2\right) \end{bmatrix} \tag{25}$$

Here, c denotes the wave speed (m/s), w_0 denotes the carrier frequency (kHz), p_u denotes the pulse width (ms), and b denotes the sweep rate (Hz/s). b can be positive (LFM upsweep), negative (LFM downsweep), or zero. In this paper, all the waveform parameters are assumed to be constant except the signal to noise ratio, carrier frequency, and pulse length at time k.

From Eqs. (23) and (25), we can see that different signal to noise ratio $S_{NRnetted}$ in the radar network can lead to different measurement noise covariance. However, during the tracking process, R_{ti} and R_{rj} are unknown before radar detection in Eq. (23). So R_{ti} and R_{rj} at time k are predicted according to target's velocity, the distance from ith transmitter to the target and distance from target to jth receiver at the last sampling time $k - 1$. So R_{ti} and R_{rj} are replaced by R_{ti}^{pre} and R_{rj}^{pre} which are, respectively, presented as

$$R_{ti}^{pre}(k) = R_{ti}(k-1) + T v_{ei}(k-1) \tag{26}$$

$$R_{rj}^{pre}(k) = R_{rj}(k-1) + T v_{ej}(k-1) \tag{27}$$

where T is the tracking interval, $v_{ei}(k-1)$ and $v_{ej}(k-1)$ are estimated by the ith and jth radar using the tracking algorithm at time $k - 1$.

3.4 Predicted error covariance
Using the state vector x_k and transfer Eq. (2), the observation vector z_{k+1}^{pre} at time $k + 1$ will be predicted. Then, the predicted error covariance matrix P_{k+1}^{pre} will be obtained by Eq. (20) under the clutter environment of the time k. The smaller the trace of P_{k+1}^{pre} is, the better the tracking accuracy will become at time $k + 1$. So the predicted error covariance will be one of the awards in Markov decision process

(MDP) which is employed to manage the radar resource in the next section.

4 Radar resource scheduling model based on Markov decision process

4.1 Markov decision process theory

As its name suggests, MDP processed the Markov property in the sense that system evolution beyond a decision point depends only on the system state and action chosen at that point [12]. The theory of MDP indicates that it is sufficient to locate a stationary policy to achieve optimality, meaning that there is no need to consider the past history when making a decision about which action to perform in a given state.

Under the Markovian and stationary assumptions, a discrete-time MDP [13, 14] is defined by a tuple $\{S, A, T_m, r\}$.

(1) A finite state space $S = \{s_i\}$, $i = 1,...,n$
(2) A finite and non-empty set of available control actions $A(s_i) = \{a_k\}$ $k = 1,...,|A(s_i)|$ associated to each state $s_i \in S$
(3) A real-valued one-step reward function $r: S \times A \rightarrow \mathbb{R}$, where $r = r(s_i, a_k)$ is the reward gained by taking action a_k in state s_i.
(4) $p(s_i, s_j, a_k)$ means the probability that the system will be state $s_j \in S$ when action $a_k \in A(s_i)$ in sate $s_i \in S$ is chosen. The set of these transition probabilities constitute the transition matrix T_m.

In the following, we will denote with $s(t)$, $a(t)$, and $r(t)$ the state, the action, and the reward of the system at time t, respectively. A stationary policy is a function π: $S \rightarrow A$, which maps every state $s_i \in S$ to a unique control action $a_k \in A(s_i)$. When the system operates under policy π, the MDP reduces to a discrete-time Markov chain.

4.2 Radar resource scheduling model based on MDP

In this paper, the state space is composed of n trace value of tracking error covariance matrix, which represents the tracking accuracy of IMMPDA in Section 2. s_i denotes the ith tracking accuracy. As it is shown in Section 3, the action a_k should denote the kth radiation of radar j with its power p_k, carrier frequency w_{0k}, pulse width p_{uk}, and sampling interval s_{ak}, which have an impact on the tracking performance.

$$a_k = \left\{ p_{kj}^i, w_{0kj}^i, p_{ukj}^i, s_{akj}^i \right\} \tag{28}$$

The reward function in our algorithm is designed as:

$$r(s_i, a_k) = \text{Nor}\big(\text{trace}\big(P_{k|k+1}(a_k)\big)\big) + \text{Nor}(\text{resource}(a_k)) \tag{29}$$

where Nor() is the normalization function, $P_{k|k+1}(a_k)$ is

the predicted tracking error covariance after the action a_k is taken, and trace() is the function of trace operation. resource(a_k) means the power and time cost of the action a_k, resource(a_k) = p_k^i / s_{ak}^i, as we know, larger sampling interval means less radiation times during the tracking.

As the predicted tracking error covariance can be predicted through the relationship between radar resource and tracking performance, the transition probability is supposed to be 1. The expected award in the algorithm can be represented as:

$$V(i) = \min_a \{r(s_i, a_k)\} \tag{30}$$

The MDP problem is the determination of the optimal policy a_k^* minimizing cost V. In this paper, it is difficult for the radar to control the power, carrier frequency, pulse width, and sampling interval freely, so the award design is converted to an optimization problem. Four kinds of radar radiation parameters are designed in this paper, as it is shown in Table 1. During the scheduling, the action a_k is represented as an eight binary sequence $\{a_1 a_2 a_3 a_4 a_5 a_6 a_7 a_8\}$, which implies the radiation parameters of radar resource. The optimization variable in this study is a binary variable that corresponds to the status of the radar resource.

4.3 Improved WDO method for solution of MDP

As it is shown in Section 4.2, finding an optimal solution of an MDP means finding a policy that minimizes V. In order to solve the MDP problem, a binary wind-driven optimization (WDO) method is proposed in this section, which can select the binary sequence for the excellent tracking accuracy and low probability of intercept. This optimization problem is mathematically formulated as (30).

The WDO presented in [15] is not for the optimization of binary solution. In order to solve the binary optimization problem of MDP, the original WDO method is modified with the binary character. The major difference between binary WDO with continuous version is that the amount of position displacement per iteration is rather defined in terms of probabilities that a bit will change to one. In continuous WDO, the coordinates of the air parcel as the parcel positions are represented in continuous values; meanwhile, in binary

Table 1 Binary sequence of the radar resource

$a_1 a_2$		$a_3 a_4$		$a_5 a_6$		$a_7 a_8$	
00	p^1	00	w_0^1	00	p_u^1	00	s_a^1
01	p^2	01	w_0^2	01	p_u^2	01	s_a^2
10	p^3	10	w_0^3	10	p_u^3	10	s_a^3
11	p^4	11	w_0^4	11	p_u^4	11	s_a^4

WDO, the parcel's coordinates are shown in discrete values. Velocity updating in binary WDO is similar to WDO, but it has velocity clamping to balance the exploration and exploitation in the search space. Plus, in binary WDO, the position updating is based on a sigmoid function as shown in Eqs. (30) and (31). The sigmoid function is used to force the real values between 0 and 1.

$$S(u_{new}) = 1/(1 + \exp(-u_{new})) \tag{31}$$

$$x_{new} = \begin{cases} 0 \, if \, S(u_{new}) \leq rand \\ 1 \, if \, S(u_{new}) > rand \end{cases} \tag{32}$$

where $S()$ is a sigmoid function which can scale the positions between 0 and 1, rand is a quasi random number uniformly distributed within the range of (0, 1), and u_{new} and x_{new} are the velocity and position in the next iteration, respectively.

5 Simulations
In this section, Monte Carlo simulations are performed to analyze the performance of the proposed resource scheduling method.

5.1 Design of target trajectory and radar resource
We assume that there are two radars for tracking one target in clutter. In the transition matrix Eqs. (21) and (22), $T = 1$ s and $\omega = 0.1$. Figure 2 shows the target trajectory with its measurement results in 100 s. The clutter for the target is produced randomly during the trajectory. Both the positions of the radars are (0 km, 0 km) and (20 km, 100 km), respectively.

The simple library of the radar resource is designed as Table 2. The sampling interval, power, pulse width, and carrier frequency have four types of value.
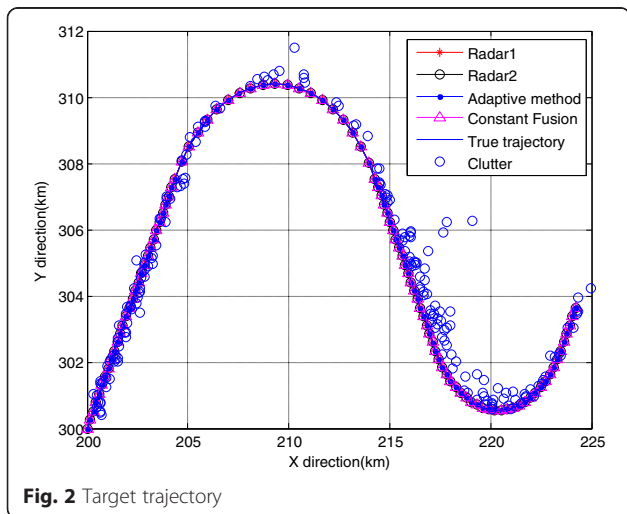
Table 2 Binary sequence of the radar resource

Power (kw)		Carrier frequency (kHz)		Pulse width (ms)		Sampling interval (s)	
p^1	10	w_0^1	10	p_u^1	0.1	s_a^1	1
p^2	20	w_0^2	20	p_u^2	0.2	s_a^2	2
p^3	30	w_0^3	30	p_u^3	0.3	s_a^3	3
p^4	40	w_0^4	40	p_u^4	0.4	s_a^4	4

5.2 Comparison of tracking performance
The proposed adaptive resource scheduling method and the resource scheduling with constant parameters for the tacking fusion in the radar network, which are labeled as "Adaptive fusion" and "Constant fusion," respectively, are realized in the simulation. The simulation also compares the performance with the single working radars which are, respectively, labeled as "Radar1" and "Radar2," both of which use the proposed resource scheduling method-based MDP. The methods of "Constant fusion" is assumed to track the target with maximum power, minimum sampling interval, and constant waveform parameters in order to get the highest detection probability which is nearly equal to 1. In addition, the radar receives the signals reflected from the target when the other radar works. The fusion center tracks the target according to the data received from all the radars in the network.

The root-mean-square error (RMSE) of time k can be formulated as Eq. (32):

$$RMSE(k) = \sqrt{\frac{1}{M_c} \sum_{m=1}^{M_c} \left(x_k - \hat{x}_k^m\right)^2} \tag{33}$$

where M_c is the number of the Monte Carlo simulation, x_k is the true state of the system, and $\hat{x}_k^m$ is the estimated vector at the mth simulation, $M_c = 200$.
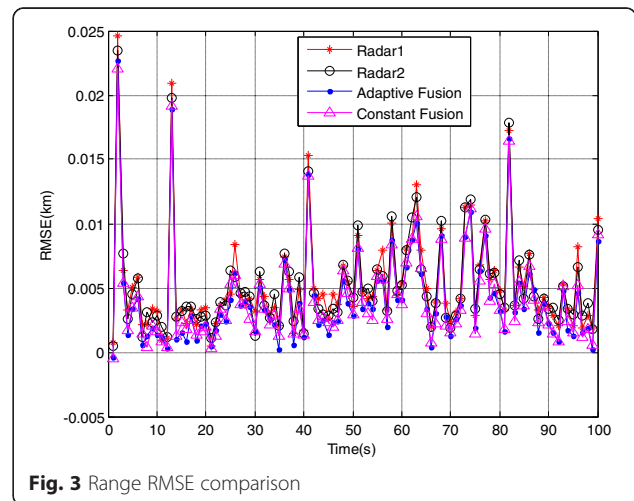

Fig. 2 Target trajectory

Fig. 3 Range RMSE comparison

(a) Pulse width

(b) Carrier frequency

Fig. 4 Transmitted waveform feature
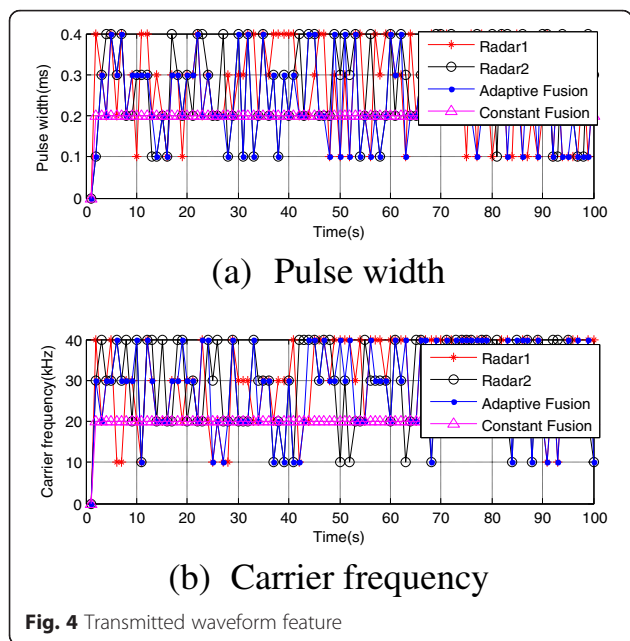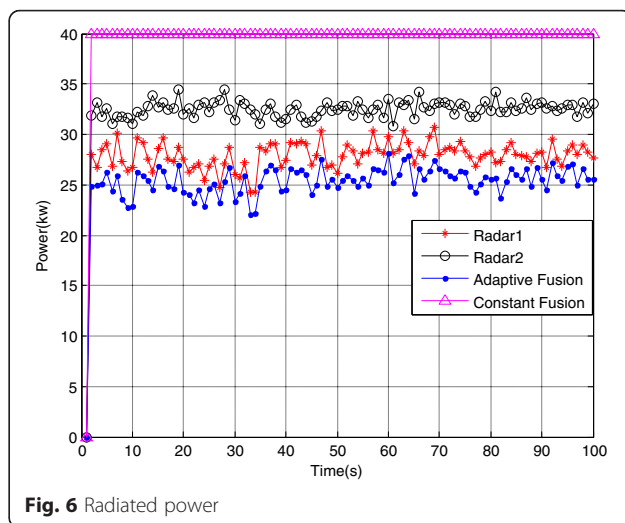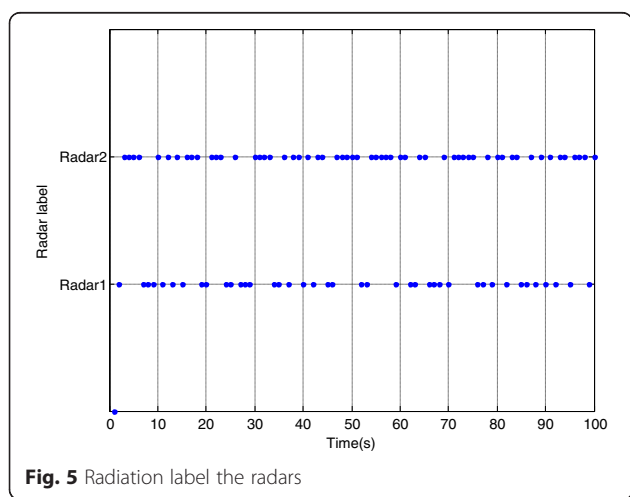

Fig. 6 Radiated power

Figure 3 shows the range RMSE using the four methods. We can see that the proposed method "Adaptive fusion" presents almost the same excellent tracking accuracy with other methods. As the fusion method takes the advantage of the superior decision ability of MDP and optimization performance of binary WDO, it can obtain excellent tracking performance using the more effective tracking data selected from Radar1 and Radar2 in turn.

5.3 Comparison of RFS performance
The transmitted signal feature is shown in Fig. 4, and the pulse width and carrier frequency are shown in Fig. 4a, b, respectively. The results show the waveform agility during the tracking, which will bring low probability of intercept ability of the radars. The radiation

label of the two radars is shown in Fig. 5. We can see that the radars work in turn.

The radiated power and sampling interval are illustrated in Fig. 6 and Fig. 7, respectively. Compared with the other three methods, we can see that the proposed method presents excellent tracking accuracy, but also radiates least power with largest sampling interval. The experimental results indicate that binary wind-driven optimization method has a better global optimization capability, as the method could select a proper set of radar resource according to the decision superiority of MDP and accurate predicted tracking error covariance in IMMPDA, after building the relationship model between the radar resources and tracking performance.

A detailed account of the complexity requirements of the proposed algorithm of "Adaptive fusion," "Constant fusion," "Radar1," and "Radar2" is performed and

Fig. 5 Radiation label the radars

Fig. 7 Sampling interval
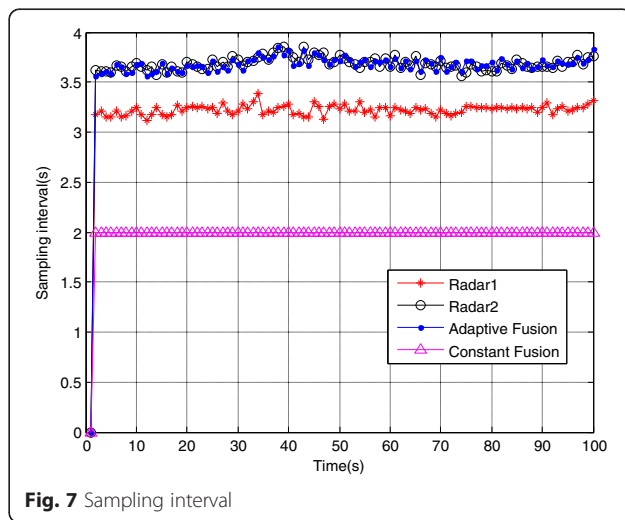
Table 3 Requirements of radar resource scheduling methods

Adaptive fusion	Radar1	Radar2	Constant fusion
$O(N_{ra}N_{p\max}N_{i\max} + W_k + N)$	$O(N_{p\max}N_{i\max} + W_k + N)$	$O(N_{p\max}N_{i\max} + W_k + N)$	$O(W_k + N)$

illustrated in Table 3. $N_{p\max}$ and $N_{i\max}$ are the maximum numbers of air parcels and iterations in WDO process. W_k and N are the numbers of validated measurements and the motion models, respectively, in the IMMPDA algorithm. N_{ra} represents the radar numbers in the network, and N_{ra} is supposed to be 2 in our simulation.

It is seen that the proposed "Adaptive fusion" method has a comparable computation complexity. As the number of radars increases, the computation time is increased as well. The proposed binary WOD is a suitable search technique for radar resource scheduling in this paper, it has to spend more time on the optimization process than the other methods. It also has been seen that the "constant fusion" method present least complexity with largest waste of radar resource.

As we know, the theory [12] of MDP indicates that it is sufficient to locate a stationary policy to achieve optimality. The proposed method uses MDP to find optimal policy in radar resource scheduling, which can achieve an optimal trade-off between the radar RFS ability and tracking accuracy. In addition, the proposed binary WDO used for solving MDP problem can provide advantages over other optimization methods [15], as it can prevent air parcels from remaining trapped at the boundary for long periods of time and pull them back into the search space. And the Coriolis force in WDO introduces a stochastic effect from other dimensions, providing robustness to the motion of the parcel. In addition, the WDO method proves higher efficiency than the other optimization methods [15], such as particle swarm optimization (PSO) and genetic algorithm (GA).

As a result, the proposed method based MDP and binary WDO can present excellent performance for radar resource scheduling.

6 Conclusions
In this paper, we have presented a new resource scheduling method for the radar network based on MDP. During the target tracking in the clutter, the relation model is built between the radar resource and tracking performance. Then, the selection of radar and its resource can be optimized by an improved WDO method, in order to obtain the maximum reward in the MDP. The simulation results show that the proposed algorithm reduces much more radar resource with excellent tracking performance in clutter. Future research will focus on the reduction of computational complexity for resource scheduling in the radar network.

Authors' information
Zhenkai Zhang was born in Jiangsu, China. He received his doctor degree from Nanjing University of Aeronautics and Astronautics in 2013. His research interests include radar signal processing and optimization methods. Now, he is an academic visitor in Durham University in UK.
Yubo Tian was born in 1971. He is a Ph.D. and a professor in the Department of Electronic Engineering, Jiangsu University of Science and Technology. His research interests include radar signal processing and electromagnetic computing.

References
1. David A. Lynch. Introduction to RF stealth. Raleigh, North Carolina: SciTech Publishing press; 2013.
2. MA Govoni, L Hongbin, JA Kosinski, Low probability of interception of an advanced noise radar waveform with linear-FM. IEEE Trans Aerosp Electron Syst **49**(2), 1351–1356 (2013)
3. T Thayaparan, M Daković, L Stanković, Mutual interference and low probability of interception capabilities of noise radar. IET Radar, Sonar & Navigation **2**(4), 294–305 (2008)
4. Z Zhang, J Zhou, F Wang, W Liu, Multiple-target tracking with adaptive sampling intervals for phased-array radar. J Syst Eng Electron **22**(5), 760–766 (2011)
5. T Cheng, DQ Zou, ZS He, *Adaptive waveform and sampling interval tracking based on estimation accuracy for Doppler radar* (Proceedings of IET International Radar Conference, Xi'an, 2013), pp. 1–4
6. V Krishnamurthy, Emission management for low probability intercept sensors in network centric warfare. IEEE Trans Aerosp Electron Syst **41**(1), 133–151 (2005)
7. S Chenguang, Z Jianjiang, W Fei, *Low probability of intercept optimization for radar network based on mutual information* (2014 IEEE China Summit & International Conference on Signal and Information Processing (ChinaSIP), Xi'an, 2014), pp. 683–687
8. S Chenguang, Z Jianjiang, W Fei, Y Yuxiao, *Minimum mean square error based low probability of intercept optimization for radar network* (IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), Guilin, 2014), pp. 10–13
9. K Jo, J Kim, M Sunwoo, Real-time road-slope estimation based on the intergration of on bradard sensors with GPS using an IMMPDA filter. IEEE Trans Intell Transp Syst **14**(4), 1718–1732 (2013)
10. B Vondra, *Overview of the multisensor IMMPDA filter with an amplitude feature for tracking maneuvering target in cluttered environment* (Proceedings of 21st international conference on applied electromagnetic and communications, Dubrovnik, 2013), pp. 1–6

11. Y Teng, HD Griffiths, CJ Baker, K Woodbridge, Netted radar sensitivity and ambiguity. IET Radar, Sonar & Navigation **11**(6), 479–486 (2007)
12. X Ruofan, M Fumio, T Kishor, *A Markov decision process approach for optimal data backup scheduling* (44th Annual IEEE international conference on dependable systems and networks, Atlanta, 2014), pp. 660–665
13. P-Y Kong, Optimal probabilistic policy for dynamic resource activation using Markov decision process in green wireless networks. IEEE Trans Mob Comput **13**(10), 2357–2368 (2014)
14. G Oddi, M Panfili, A Pietrabissa et al., *A resource allocation algorithm of multi-cloud resource based on Markov decision process*, 1st edn. (Proceedings of IEEE 5th International Conference on Cloud Computing Technology and Science, Bristol, 2013), pp. 130–135
15. Z Bayraktar, M Komurcu, JA Bossard, DH Werner, The wind driven optimization technique and its application in electromagnetics. IEEE Trans Antennas Propag **61**(5), 2745–2757 (2013)