A Distributed Expectation-Maximization Algorithm for OTHR Multipath Target Tracking

Hua Lan, Yan Liang, Zengfu Wang, Feng Yang, and Quan Pan
School of Automation, Northwestern Polytechnical University
Key Laboratory of Information Fusion Technology, Ministry of Education
Xi'an, Shaanxi, 710072, P.R.China

***Abstract*—Consider the problem of joint state estimation and mode identification for over-the-horizon radar (OTHR) multipath target tracking, where multiple resolvable detections corresponding to the different propagation modes come from the same target, a consensus based distributed expectation-maximization (EM) framework is proposed in this paper. By regarding the OTHR as a sensor network where each propagation mode corresponds to a sensor node, the proposed distributed EM deals with the multiple data association and state estimation jointly based on the network consensus theory. In the *E-step*, each mode independently calculates local state estimate by using its associated measurement. A consensus filter is used to exchange its localized estimate with its neighbors and then fuse them. In the *M-step*, each mode uses the estimated global state to find the local optimal measurement in the nearest neighbor sense. Based on the above iterative and consensus mechanism, the global optimal state estimate is obtained in the *E-step* while the local optimal measurement for each mode is chosen in the *M-step*. The proposed distributed EM algorithm is more accurate and computational effective than the conventional EM algorithm via a numerical example of OTHR target tracking.**

***Index Terms*—multipath, expectation maximization, consensus filter, OTHR**

I. INTRODUCTION

Skywave over-the-horizon radar (OTHR) systems operating in the high-frequency (HF) band (3-30 MHz) exploit signal refection from the ionosphere to detect and track airborne and surface targets at ranges an order of magnitude greater than conventional line-of-sight radars. A chief advantage of OTHR systems is their ability to persistently monitor remote geographical regions where microwave radar coverage is either not feasible or convenient [1]. The capability of OTHR systems to provide cost-effective early-warning surveillance over wide areas at ground distances of up to 3000km makes them uniquely important in a number of applications [2].

However, a major complicating factor in OTHR is that HF radiowave propagation through the ionosphere often gives risk to multiple propagation paths, or ray modes, between the radar and a target, which results in multiple resolved detections at each radar dwell and in multiple slant tracks from a single target [3]. Furthermore, the multipath propagation is unavoidable since it is often impossible to select a radar operating frequency that results in single-mode propagation to the region of interest [4]. Consequently, the difficulty in tracking a target for OTHR arises from the uncertain origin of the measurements and the uncertainty in the presence of multipath propagation. Much attention has been paid on the multipath state estimation for OTHR. In our opinions, these methods belong to one of the following two categories.

The first category is track-based where multi-path tracks are obtained in slant or radar coordinates based on probabilistic data association (PDA) [5-7] or Viterbi data association (VDA) [8-11], and further fused in ground coordinates via coordinate registration transformation, which requires knowledge of the ionospheric propagation conditions. The most well-known fusion solution is the multi-hypothesis multipath track fusion [12-14] through establishing propagation path-dependent track-to-target association hypotheses, calculating the hypothesis probabilities, and further obtaining the fused track in the minimum mean squared error sense.

An alternative category is measurement-based where track prediction and measurement-target-mode association are implemented in slant coordinates and the track updating fulfilled in ground coordinates. One well-known scheme is multipath probabilistic data association (MPDA) [4, 15], which combines multi-path measurements for state updating. Another scheme is multipath Viterbi data association (MVDA) [16], which chooses the optimal pass (measurement and propagation associations) to maximize the likelihood function. Recently, there have been some developments on OTHR multitarget tracking such as multiple detection MHT (MD-MHT) [17] and multiple detection JPDA (MD-JPDA) [18], which also belong to the measurement-based category.

In fact, the OTHR multipath data association and state estimation problem involves both identification and estimation. Identification in this sense includes data association and propagation mode, while estimation includes path-conditional state estimation and multipath track fusion. Essentially, the identification and estimation are highly coupled and affect each other. Namely, the identification mistake of multipath data association deteriorates state estimation and track fusion while the estimate error triggers the identification risk. However, both track-based methods and measurement-based methods combine identification with estimation. It is highly demanded to develop the joint optimization of identification and estimation. Considering the case of multiple observations of a single target in clutter, Pulford and Logothetis [19] presented the expectation maximization data association (EMDA) for fixed-interval Kalman smoothing conditioned on the MAP estimation of measurement-to-mode association sequence in the expectation-maximization (EM) framework. However, as mentioned in [19], the EMDA is suitable for off-line, batch computation. It is suggested that an approximate method is needed to decrease the computational cost. In our previous works, we proposed the joint multipath data association and state estimation (JMAE) algorithm [20] to obtain the approximate solutions, which is more computationally effective than the EMDA and is superior to the well-known MPDA. However, both of the EMDA and the JMAE are based on the traditional expectation-maximization (EM) with a centralized architecture whereby all the valid measurements from each mode are combined to estimate the target's trajectory. The centralized EM algorithm often converges very slowly and has expensive computational burden, especially when applied to the high-dimension missing data (association hypothesis), such as OTHR multipath target tracking. Consequently, it is necessary to find a distributed EM algorithm to decrease the computational burden. However, to the authors' knowledge, this topic is still open.

In this paper, we regard the OTHR multipath target tracking as a sensor network where each propagation mode corresponds to a sensor node in the network, and a distributed EM algorithm based on consensus filter (DCEM) is proposed to perform the multiple data association and state estimation jointly. In the *E-step*, each propagation mode calculates the local, state estimate independently, and receives the local estimates of its neighboring nodes to update its local state estimate by using a consensus filter. In the *M-step*, the local optimal measurement for each mode is searched by using the updated state estimate via the nearest neighbor algorithm, which is desirable to decrease the computational cost.

The rest of this paper is organized as follows: Section II formulates the problem of OTHR multipath target tracking. In Section III, the joint classification and estimation scheme based on DCEM is proposed. Simulation results are presented in Section IV followed by concluding remarks in Section V.

Throughout this paper, the superscripts "-1" and "T" represent the inverse and transpose operations of a matrix, respectively; $\mathcal{N}\{x;\mu,P\}$ represents the Gaussian probability function of x with mean μ and covariance P; E is mathematical expectation. The notation $\otimes$ refers to the Kronecker product. I and 0 denote the identity matrix and the zero matrix with proper dimensions, respectively. $I\{\bullet\}$ denotes the indicator function, which equals one if the event $\{\bullet\}$ is true, or zero otherwise. For a vector x, define $\mathcal{D}(x,P) = x^T P^{-1} x$, where P is a positive-definite matrix.

II. Problem formulation

Consider a single target in the clutter environment. The target state in ground coordinates at time instant k is defined by $x_k = [R_k, \dot{R}_k, \vartheta_k, \dot{\vartheta}_k]^T$, which consists of ground range R, range rate $\dot{R}$, bearing ϑ and bearing rate $\dot{\vartheta}$. The discrete-time state equation is

$$x_{k+1} = f(x_k) + v_k \tag{1}$$

where the state transition function $f(\bullet)$ is given; v_k is a zero-mean white Gaussian noise vector with the known covariance Q_k; k is time instant.

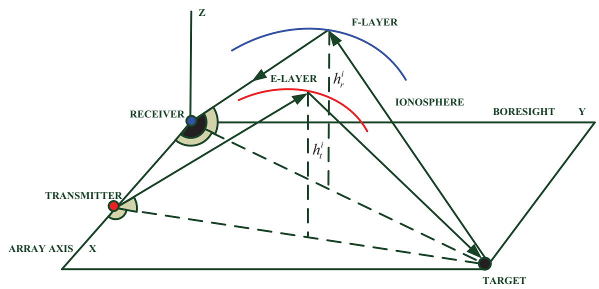


Fig. 1: Radar signal propagation via the ith mode [4]

The multipath propagation is shown in Fig.1. The virtual ionospheric heights on transmission and reception via the ith propagation mode are denoted by h_t^i and h_r^i. The target measurement via the ith propagation mode is detected with the detection probability P_d^i, and the corresponding measurement of the detected target measurement at time instant k is

$$y_k^i = h^i(x_k) + w_k^i, \qquad i = 1,2,...,t \tag{2}$$

where $y_k = [r_k, \dot{r}_k, \theta_k]^T$ consists of slant range r_k, range rate $\dot{r}_k$ and azimuth θ_k; t is the number of possible ionospheric propagation modes; the measurement function $h^i(\bullet)$ is assumed known in contrast to the approach in [15]; the measurement noise w_{k+1}^i is a zero-mean white Gaussian noise vector with known covariances $R_{k+1}^i > 0$. The initial state x_0 is Gaussian distributed with known mean $\bar{x}_0$ and associated covariance Σ_0. Here v_k, w_k^i and x_0 are mutually independent.

The standard uniform and Poisson models are used to represent the clutter. The probability density function (pdf) of clutter measurements in a region $G(k)$ with the corresponding volume $V_G(k)$ is assumed to be uniformly distributed [4]:

$$p_c(y_k) = \begin{cases} V_G(k)^{-1} & \text{if } y_k \in G(k) \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

The number of clutter points in a region $G(k)$ with the corresponding volume $V_G(k)$ is assumed to be Poisson distributed with point mass function $\mu_c(\bullet)$:

$$\mu_c(n) = \frac{(\lambda V_G(k))^n e^{-\lambda V_G(k)}}{n!}, \quad n = 0,1,2,... \tag{4}$$

where λ is the spatial density of the clutter, and the corresponding clutters set is denoted by C_k .

The *validation region* for each filter is established in order to reduce the association hypothesis events that described in [4]. Denote the volumes of the gates and its corresponding number of measurements at time instant k via ith propagation mode by V_k^i and m_k^i, respectively.

The measurement set of OTHR system at time instant k is defined by

$$Y_k = \{y_k(j)\}_{j=1}^{M_k} \tag{5}$$

$$y_k(j) \in \{y_k^i\}_{i=1}^{t} \bigcup \{C_k\} \tag{6}$$

where M_k is the number of measurements at time instant k.

Definition 2.1 Define the measurement-to-target association and the target measurement-to-mode association at time instant k by $a_k \in \{-1, 0, ..., M_k\}$ and $b_k \in \{1, 2, ..., t\}$, respectively:

- "$a_k = -1$": the target measurement does not exist and all measurements are due to clutter;
- "$a_k = 0$": the target exists, but all measurements are due to clutter;
- "$a_k = n, n > 0$": the target exists, and the nth measurement is originated from the interested target;
- "$b_k = i$": the detected target measurement is via the ith propagation.

Notation 2.1 Denote the sequence of measurements $Y_{k-l}^k = \{Y_{k-l}, ..., Y_k\}$, the sequence of states $X_{k-l}^k = \{X_{k-l}, ..., X_k\}$, the sequence of measurement-to-target associations $A_{k-l}^k = \{a_{k-l}, ..., a_k\}$ and the sequence of measurement-to-mode associations $B_{k-l}^k = \{b_{k-l}, ..., b_k\}$, respectively.

The model parameters Θ are defined as follows:

$$\Theta = \{f, h^i, h_t^i, h_r^i, R^i, Q, P_d^i, \bar{x}_0, \Sigma_0, \lambda, A, B, t\} \tag{7}$$

Assuming the model parameters Θ are exactly known, and given the measurement sequence Y_{k-l}^k, the optimal (in a MAP sense) estimates of X_{k-l}^k are obtained by maximizing the probability density function of X_{k-l}^k conditioned on Y_{k-l}^k, i.e.,

$$X^{MAP} = \arg\max_{X_{k-l}^k} f(X_{k-l}^k | Y_{k-l}^k, \Theta) \tag{8}$$

However, the model parameters Θ are partly known, that is, both the measurement-to-target associations A_{k-l}^k and the measurement-to-mode associations B_{k-l}^k are unknown. As a consequence, the OTHR multipath target tracking can be treated as the problem of estimation in the presence of incomplete data. The EM algorithm has been widely used in the engineering and statistical literature as an iterative optimization procedure for computing maximum likelihood (or MAP) parameter estimates of *incomplete data problem* [21]-[24], which is helpful in improving the performance of estimation and identification under multiple uncertainties, and avoids an exhaustive search through a larger solution space.

Remark 2.1. In reality, the behavior of the ionosphere and the resulting propagation modes is highly complex. The virtual heights of ionosphere including h_t^i and h_r^i are not known exactly. Meanwhile, the number of propagation modes t is also time-varying. Both of them are provided by the ionosphere detection device subsystem. In this paper, a constant ionospheric model is used analogously as in [4].

Remark 2.2. The OTHR multipath target tracking involves both identification of the multiple data association and state estimation where the performance of identification and estimation affects each other, the EM based algorithms including EMDA and JAEM for joint decision and estimation are interesting exploring. The EMDA algorithm regards the multipath data association as a triple of measurement-target-mode association, and uses the Kalman smoother to estimate the state in the *E-step*, while searches the optimal measurement-target-mode based on the Viterbi algorithm in the *M-step*. The EMDA is optimal, but has a large computational complexity. As stated in [19], a fast and approximated solution is thus desirable. The JMAE algorithm is one of the computational-effective solutions, which converts the triple association into two binary associations. Instead of finding the optimal association, the JMAE purses *a posterior* probability of each mode and its corresponding pseudo-measurement to estimate the local state, and further obtain the fused estimate. However, both of them consider the OTHR multipath target tracking as a centralized processing, the computational complexity increases rapidly as the increase of measurements and modes. Moreover, as an iterative optimization method, the efficiency of the EM algorithm is affected by the rate of convergence. In the case of OTHR, the number of measurements per scan is in the tens of thousands, and the number of ionospheric models is always more than four. Thus, the missing data is of high dimension hypothesis in the EM framework consisting of dense measurements and many possible mode association, which implies that the EM algorithm converges slowly. The key problem in the OTHR is to find the fast implement of the EM algorithm, the distributed EM algorithm uses the consensus filter to diffuse local information of each mode to the others, and thus has the advantage of parallel computing.

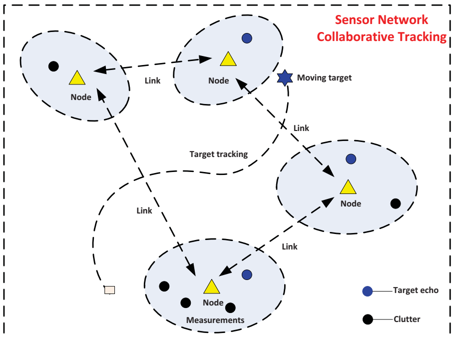
The objective of this paper is to propose a distributed EM algorithm to perform the state estimation and identification of multipath data association jointly, and decrease the computational cost to meet the needs of practical application.

III. Consensus based distributed EM algorithm

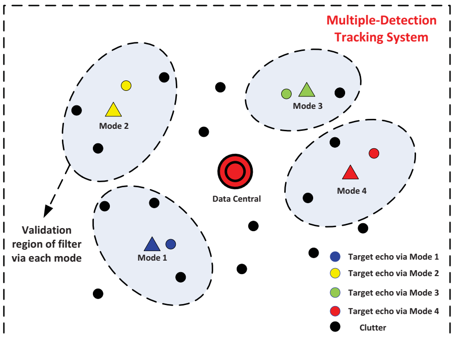
A. Correspondence Between OTHR and Sensor Network

The OTHR multipath target tracking consists of multiple simultaneous measurements from the same target along with unwanted clutters. Each propagation mode has its independent target measurement with detection probability and localized possible candidate association hypothesis. The OTHR can be regarded as a special sensor network, and each node corresponds to a sensor node. Motivated by the above similarities, the OTHR target tracking is transformed into the sensor network collaborate tracking problem as shown in Fig.2.

Distributed estimation and tracking is one of the most fundamental collaborative information processing problems in sensor networks. Decentralized Kalman filtering [25,26] involves state estimation using a set of local Kalman filters that communicate with all other nodes. Control-theoretic consensus algorithms have proven to be effective tools for performing network-wide distributed computation tasks such as computing aggregate quantities and functions over networks [27,28]. Distributed EM is widely used for density estimation and clustering in sensor networks [29,30]. Different from the centralized architecture, the distributed architecture is interested in performing collaborative tracking, but without the need for a central fusion node. While each node collaborate by exchanging appropriate messages between neighboring nodes to achieve the same effect as they would by communicating with



(a) Target tracking in sensor network



(b) Target tracking in OTHR

Fig. 2: Target tracking in sensor network and OTHR

a central fusion node. The distributed collaborative has the advantage of more robust and less computational complexity than that of the centralized fusion architecture.

B. Distributed EM for Joint Identification and Estimation

Consider the sensor network $\mathcal{G} = (\varepsilon, \mathcal{V})$ denote an undirected graph with vertex set $\mathcal{V} = \{1, 2, \cdots, s\}$ and edge set $\varepsilon \subset \{\{i, j\} | i, j \in \mathcal{V}\}$, where each edge $\{i, j\}$ is an unordered pair of distinct nodes. A graph is *connected* if for any two vertices i and j, there exists a sequence of edges $\{i, k_1\}$, $\{k_1, k_2\}$, $\cdots$, $\{k_s, j\}$ in ε. Let $\mathcal{N}_i = \{j \in \mathcal{V} | \{i, j\} \in \varepsilon\}$ denote the set of neighbors of node i. The degree of vertex i is defined as $d_i = |\mathcal{N}_i|$ and the maximum degree is $d_{max} = \max_i d_i$. The OTHR multipath target tracking can be regarded as special sensor networks, where each mode in the OTHR system corresponds to a sensor node in the sensor network. The special sensor network is full-linked connected and all the sensor node is synchronized to a common clock, since the OTHR system itself collects all the measurements for each mode at the same time just like a fusion center and thus each mode can diffuse its information to the others.

Notation 3.1 For each mode i, denote the sequence of measurements in validation region $Y^i_{k-l:k} = \{Y^i_{k-l}, ..., Y^i_k\}$, the sequence of local states $X^i_{k-l:k} = \{X^i_{k-l}, ..., X^i_k\}$, the sequence of local measurement-to-target associations $A^i_{k-l:k} = \{a^i_{k-l}, ..., a^i_k\}$, respectively. Here $a^i_j = n$ with $n \in \{1, ..., m^i_j\}$ means the measurement n origins from target via ith mode, otherwise $a^i_j = 0$ from clutter. Let $X^k_{k-l} = \{X_{k-l}, ..., X_k\}$ be the global states.

Definition 3.3 Define the complete-data log-likelihood function of each mode i as $L^i_{k-l:k}$ and its corresponding conditional expectation, also called Q-function $\mathcal{Q}^k_{k-l:k}(r)$ at time interval $[k-l, k]$ by

$$L^i_{k-l:k} = \log p\left(X^i_{k-l:k}, Y^i_{k-l:k}, A^i_{k-l:k} | Y^i_{1:k-l-1}\right) \quad (9)$$

$$\mathcal{Q}^i_{k-l:k}(r) = E_{X^i_{k-l:k}}\left(L^i_{k-l:k} | Y^i_{k-l:k}, \hat{A}^i_{k-l:k}(r)\right) \quad (10)$$

where $\hat{A}^i_{k-l:k}(r)$ is the association estimate of $A^i_{k-l:k}$ at the rth iteration.

The multipath target tracking of OTHR is regarded as target tracking in the sensor networks, and each mode corresponds to a sensor agent. The topology of network is modeled by an undirected full-linked communication graph, since the OTHR system receives measurements from all the mode just likes a fusion center, and each mode diffuses its information to other modes. The functional flow of the proposed DCEM algorithm is shown as Fig.3.

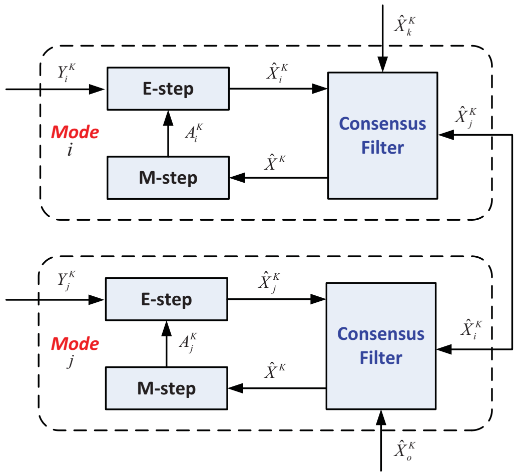


Fig. 3: The functional flow of the DCEM framework

Each node consists of a local EM algorithm and a consensus filter. The local EM algorithm using the local measurements to estimate the local state in the *E-step*, and the consensus filter updates the local state by using the information from other nodes. All the local estimates converge to the global state estimate after the consensus filter, and further the optimal association of each node is obtained by using the global state estimate. The proposed DCEM has the following advantages.

1) The proposed DCEM is more computationally effective than the EMDA, since it converts the triple association of measurement-target-mode into measurement-target association of each mode, and each mode communicates with each other by using the consensus filter. In a sense, the proposed DCEM is a parallel processing just as JMAE, while the EMDA is serial for deal the multipath data association.
2) The proposed DCEM has more estimation performance than the JMAE. The JMAE has two aspects of approximation, that is, it is an approximate rather than the

optimal solution for maximizing Q-function in the *M-step*, and the approximate solution is based on the approximate fusion of all the mode state estimates that are correlative. However, the DCEM avoids the approximate problems posed by the JMAE.

3) The proposed DCEM has more robust than others. The consensus filter is proved very effective in the distributed estimation as its convergence and robustness. Particularly, in the case of OTHR target tracking where the detection probability of each is very low, and the local state estimate is not stable. The DCEM diffuses each mode information over the networks and converges to the global state.

(1) Localized Mode-Based EM

The local EM algorithm just uses the local measurements that fill in the validation region of each mode filter to perform the state estimation, and ignores the influence from other modes. For each mode i, we treat local measurement set $Y^i_{k-l:k}$, the target state sequence $X^i_{k-l:k}$ and the sequence of association events $A^i_{k-l:k}$ as incomplete data, missing data and unknown parameters, respectively. The corresponding local EM consists of the following two iterative steps:

$$\textit{E-Step}: \mathcal{Q}^i_{k-l:k}(r) = E\left(L^i_{k-l:k}|Y^i_{k-l:k}, \hat{A}^i_{k-l:k}(r)\right) \quad (11)$$

$$\textit{M-Step}: \hat{A}^i_{k-l:k}(r+1) = \underset{A^i_{k-l:k}}{\arg\max}\, \mathcal{Q}^i_{k-l:k}(r) \quad (12)$$

The complete data log-likelihood function $L^i_{k-l:k}$ defined by (9) can be expressed as follows:

$$\begin{aligned} L^i_{k-l:k} =& \log p\left(X^i_{k-l:k}, Y^i_{k-l:k}, A^i_{k-l:k}|Y^i_{1:k-l-1}\right) \\ =& \sum_{j=k-l}^{k} \log p\left(y^i_j|x^i_j, a^i_j\right) \\ &+ \sum_{j=k-l}^{k} \log p\left(x^i_j|x^i_{j-1}\right) \\ &+ \log p\left(x^i_{k-l-1}|Y^i_{1:k-l-1}\right) \end{aligned} \quad (13)$$

By the fact that the initial state, process noises and measurement noises are all Gaussian, the items in (13) can be rewritten as

$$p\left(x^i_{k-l-1}|Y^i_{1:k-l-1}\right) = \mathcal{N}\left\{x^i_{k-l-1}; \bar{x}^i_{k-l-1}, \Sigma^i_{k-l-1}\right\} \quad (14)$$

$$p\left(x^i_j|x^i_{j-1}\right) = \mathcal{N}\left\{x^i_j; f(x^i_{j-1}), Q_j\right\} \quad (15)$$

$$p\left(y^i_j|x^i_j, a^i_j = n\right) = \begin{cases} p\left(y^i_j(n)|x_j\right), & a^i_j \in \{1, ..., m^i_j\} \\ \{V^i_j\}^{-1}, & a^i_j = 0 \end{cases} \quad (16)$$

$$p\left(y^i_j(n)|x^i_j\right) = \mathcal{N}\left\{y^i_j(n); h^i(x^i_j), R^i_j\right\} \quad (17)$$

where $\bar{x}^i_{k-l-1}$ and Σ^i_{k-l-1} are the mean and covariance of system state at time $k-l-1$ of ith propagation mode, respectively.

The Q-function $\mathcal{Q}^i_{k-l:k}$ is obtained through taking that takes conditional expectation of (13) with respect to the state sequence $X^i_{k-l:k}$:

$$\begin{aligned} \mathcal{Q}^i_{k-l:k}(r) =& E\left\{\sum_{j=k-l}^{k} \log p\left(y^i_j|x^i_j, a^i_j\right) p\left(x^i_j|x^i_{j-1}\right)\right\} \\ =& \sum_{j=k-l}^{k}\left\{\mathcal{D}\left(y^i_j(n) - H^i\hat{x}^i_{j|k-l:k}, R^i_j\right)\right\} \\ &+ \sum_{j=k-l}^{k}\left\{tr\left\{(H^i)^T(R^i_j)^{-1}H^iP^i_{j|k-l:k}\right\}\right\} \\ &+ \sum_{j=k-l}^{k}\left\{\mathcal{D}\left(\hat{x}^i_{j|k-l:k} - F\hat{x}^i_{j-1|k-l:k}, Q^i_j\right)\right\} \\ &+ \sum_{j=k-l}^{k}\left\{\mathcal{D}\left(P^i_{j|k-l:k} - FP_{j-1|k-l:k}, Q^i_j\right)\right\} \end{aligned} \quad (18)$$

where $x^i_{j|k-l:k} = E\left\{x^i_j|Y^i_{k-l:k}, \hat{A}^i_{k-l:k}\right\}$ and $P^i_{j|k-l:k} = Cov\left\{x^i_j|Y^i_{k-l:k}, \hat{A}^i_{k-l:k}\right\}$ are the mean and covariance of x^i_j, respectively. These can be obtained from the fixed-interval Kalman smoother consisted of the forward and backward filtered outputs as follows:

$$\begin{aligned} \hat{x}^i_{j|k-l:k} =& P^i_{j|k-l:k}(P^i_{j|k-l:j})^{-1}\hat{x}^i_{j|k-l:j} \\ &+ P^i_{j|k-l:k}(P^i_{j|j+1:k})^{-1}\hat{x}^i_{j|j+1:k} \end{aligned} \quad (19)$$

$$P^i_{j|k-l:k} = \left\{(P^i_{j|k-l:j})^{-1} + (P^i_{j|j+1:k})^{-1}\right\}^{-1} \quad (20)$$

After obtaining the Q-function $\mathcal{Q}^i_{k-l:k}$ in the *E-step*, now we need to maximize it in the *M-step*. Since the maximization in (10) is over association events $A^i_{k-l:k}$ with Q-function, we ignore the terms that independent with $A^i_{k-l:k}$ and get the following expression.

$$\mathcal{Q}^i_{k-l:k}(r) = \sum_{j=k-l}^{k}\left\{\mathcal{D}\left(y^i_j - H^i\hat{x}^i_{j|k-l:k}, R^i_j\right)\right\} \quad (21)$$

The association events sequence $A^i_{k-l:k}$ are independent over the time sequence, then the optimal association event is obtained by the nearest neighbor algorithm, otherwise a Viterbi algorithm is adopted to find the optimal association sequence in the case of dependence. The optimal association event a^i_j at time instant j for mode i is

$$a^i_j = \underset{n\in\{0,...,m^i_j\}}{\arg\min}\left\{\mathcal{D}\left(y^i_j(n) - H^i\hat{x}^i_{j|k-l:k}, R^i_j\right)\right\} \quad (22)$$

(2) Consensus filter over the network

The consensus filter always guarantees the local state reach to stationary point. They are many ways to perform the consensus filter, here we present an iterative consensus filter. The consensus filter uses its neighboring local state estimates to update its local state estimate and has the formation as follows:

$$\mathcal{X}^i_{k-l:k}(r+1) = \mathcal{X}^i_{k-l:k}(r) + \eta \sum_{j\in\mathcal{N}_i} \left(\mathcal{X}^j_{k-l:k}(r) - \mathcal{X}^i_{k-l:k}(r)\right) + \eta \sum_{j\in\mathcal{N}_i} \left(\hat{X}^i_{k-l:k} - \mathcal{X}^i_{k-l:k}(r)\right) \quad (23)$$

$$\mathcal{P}^i_{k-l:k}(r+1) = \mathcal{P}^i_{k-l:k}(r) + \eta \sum_{j\in\mathcal{N}_i} \left(\mathcal{P}^j_{k-l:k}(r) - \mathcal{P}^i_{k-l:k}(r)\right) + \eta \sum_{j\in\mathcal{N}_i} \left(P^i_{k-l:k} - \mathcal{P}^i_{k-l:k}(r)\right) \quad (24)$$

where $\hat{X}^i_{k-l:k}$, $P^i_{k-l:k}$ are the local state estimate and corresponding local estimate error covariance of mode i that obtained by the local KS by Eqs.(19)-(20), respectively; $\mathcal{X}^i_{k-l:k}$, $\mathcal{P}^i_{k-l:k}$ are the updated state estimate and its corresponding estimate error covariance, respectively. η is the weight, which satisfies $\eta \leq \frac{1}{d_{max}}$. The DCEM algorithm is summarized in Table.1.

TABLE I: The summarized of DCEM algorithm

Step 1: Initialization. ($r=0$) Given measurement set $Y_{k-l:k}$, initialize association sequence of each mode $A^i_{k-l:k}(0)$ to suitable values. Let $\mathcal{X}^i_{k-l:k}(0)$ and $\mathcal{P}^i_{k-l:k}(0)$ to zeros.
Step 2: Iteration.($r=1,2,...$) For each mode i, $i=1,2,...,s$.
Step 3a: *E-Step*. Using the current associated sequence $\hat{A}^i_{k-l:k}(r)$ to estimate the local state $\hat{X}^i_{k-l:k}$ and covariance $P^i_{k-l:k}$ by Eqs.(19)-(20). Step 3b: *Consensus filter*. Update the local state estimate $\mathcal{X}^i_{k-l:k}$ and $\mathcal{P}^i_{k-l:k}$ by other modes information by Eqs.(23)-(24). Step 3c: *M-Step*. Using $\mathcal{X}^i_{k-l:k}$ and $\mathcal{P}^i_{k-l:k}$ to choose the optimal association sequence $A^i_{k-l:k}$ by Eq.(22). Step 3d: *Termination*. If the values of $\mathcal{X}^k_{k-l}(r)$ and $\mathcal{X}^k_{k-l}(r+1)$ are close enough or the number of iteration reaches to the maximum , then iteration terminates, else set $r=r+1$, and go to Step 3a.
Step 4: Recursion. ($r=0, k=k+1$), initialize $\mathcal{X}^i_{k-l:k}$ to the last ultimate value, and go to iterate loop.

IV. Simulation

Considering the same simulation environment for a simple non-maneuvering tracking using a two-layer ionospheric model as [4], the target state equation is assumed to be linear since the target is far away from the receiver, and the detailed scenario parameters refer to Table.2. We compare the proposed DCEM with the JMAE, since simulations were not presented for EMDA in [19] due to its complex computational cost. The stochastic initialization strategies, including emEM [31] and RndEM [32] algorithms, share the common idea of trying different initial values of parameters and choosing the one that yields the largest local maximum. In this paper, we present the simulation of all the methods with different initialization choices, since all of them are based on EM framework.

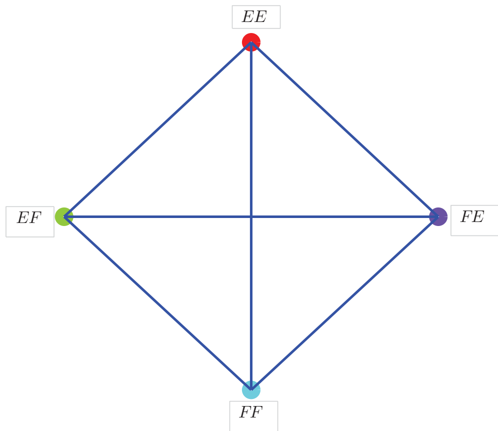
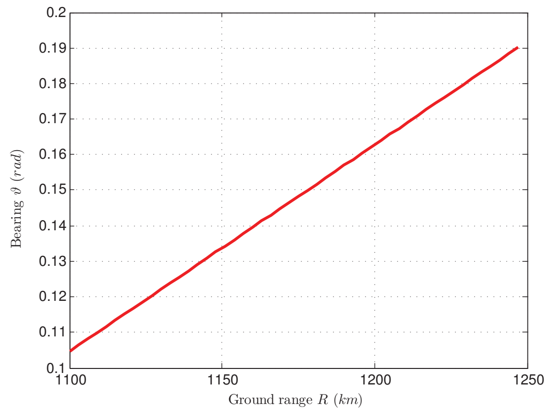
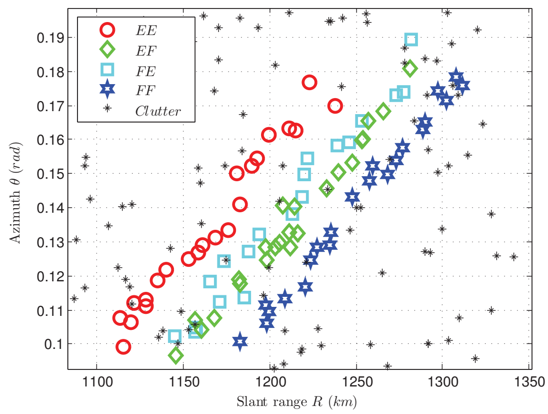


Fig. 4: The network graph for OTHR multipath target tracking



(a) Target trajectory in ground coordinates



(b) OTHR multipath detection

Fig. 5: The scenario of target trajectory and its corresponding OTHR multipath detection

The topological graph of network is shown as Fig.4, each propagation mode i is full-linked with other modes. The target tracking trajectory, OTHR multipath detections and clutters are shown in Fig.5.

The RMSE comparison based on 20 Monte Carlo runs between the proposed DCEM and the JMAE is shown in Fig.6, the estimation performance of the proposed DCEM is better than the JMAE. This is mainly because the JMAE is an approximative solution while the proposed DCEM is an optimal solution. Meanwhile, the DCEM is based on distributed EM and using the consensus filter to diffuse the information over the modes, and thus more robust. Moreover, the DCEM is more computational effective than the JMAE, particularly, the computational cost for DCEM is 855.4s and 1927.5s for JMAE.

V. Conclusion

Motivated by the OTHR multipath target tracking, a new consensus based EM algorithm is proposed for the joint multipath data association and state estimation, by regarding the OTHR system as a sensor network. The proposed scheme translates the triple association of measurement-target-mode into data association of each mode, and using local EM algorithm to solve it. A consensus filter is adopted to diffuse the local state estimates over the networks. The simulation verifies the proposed DCEM, which is more effective and precise than the JMAE. Along the result of this paper, several further researches could be done. One is how to extend the proposed EM scheme to the case of the uncertainty of ionosphere parameters. Another is to compare the proposed DCEM scheme with EMDA.

Acknowledgments

This work was supported by the National Science Foundation of China under Grants 61135001, 61374023 and 61374159. The first author is thankful for having discussion with Dr Adrian N. Bishop on the consensus based distributed expectation-maximization algorithm.

References

[1] G. Fabrizio, *High Frequency Over-the-Horizon RadarCFundamental Principles, Signal Processing, Practical Applications*, New York : McGraw-Hill Education, 2013.

[2] A. Kolosov, *Over-the-Horizon Radar*. Artech House, Boston, 1987.

[3] R. Anderson, and J. Krolik , "Track association for over-the-horizon radar with a statistical ionospheric model," *IEEE Transactions on Signal Processing*, 2002, 50(11), pp. 2632-2643.

[4] G. Pulford, and R. Evans, "A Multipath Data Association Tracker for Over-the-Horizon Radar," *IEEE Transactions on Aerospace and Electronic Systems*, 1998, 34(4), pp.1165-1182.

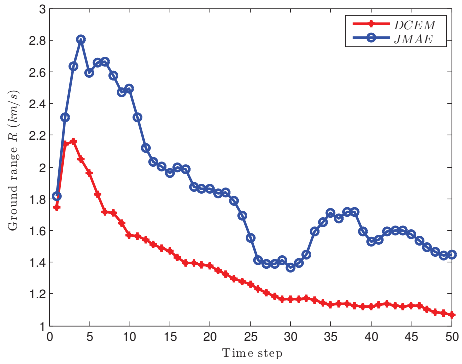
[5] G. Ahmed, F. Majdi, and M. Farooq, "A Comparison of Data Association Techniques for Target Tracking in Clutter," In *Proceedings of the 5th International Conference on Information Fusion*, 2002, 2, pp. 1126-1133.

[6] S. Colegrove, and S. Davey, "PDAF with Multiple Clutter Regions and Target Models," *IEEE Transactions on Aerospace and Electronic Systems*, 2003, 39(1), pp. 110-124.

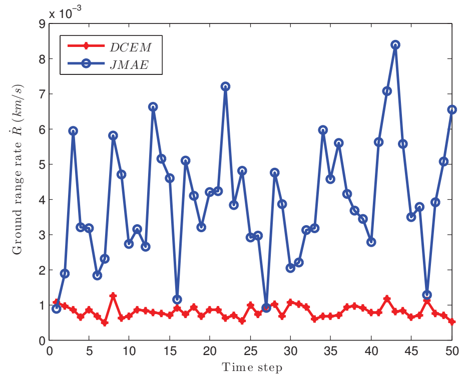
[7] G. Pulford, and R. Evans, "Probabilistic Data Association for Systems With Multiple Simultaneous Measurements," *Automatica*, 1996, 32(9), pp. 1311-1316.

[8] G. Pulford, and B. Scala, "Over-The-Horizon Radar Tracking Using The Viterbi Algorithm," *CSSIP Second Report to DSTO*, 1995, Report No.16 / 95.

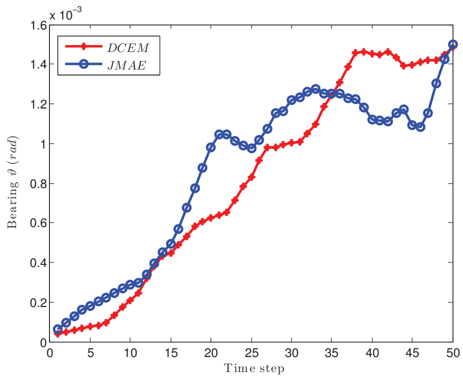
[9] G. Pulford, and B. Scala, "Over-The-Horizon Radar Tracking Using The Viterbi Algorithm," *CSSIP Third Report to DSTO*, 1995, Report No.27 / 95.



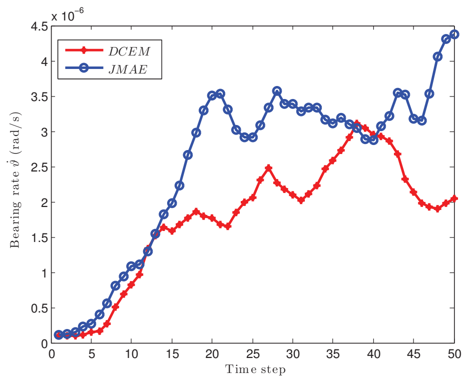
(a) RMSE in ground range



(b) RMSE in ground range rate



(c) RMSE in bearing



(d) RMSE in bearing rate

Fig. 6: The RMSE of DCEM and JMAE with $l = 5$

TABLE II: The parameters setting of the simulation scenario

Category	Parameters	Value
Scenario	number of dwells	20
	time between dwells	20 seconds
	region size (range)	1000-1400km
	region size (azimuth)	0.069813-0.17453rad
	region size (doppler)	0.013889-0.22222km/s
	Target Initial State	(1100km,0.15km/s,0.10rad,8.72e-05rad/s)
	expected number of clutter	400 per dwell
	sensor noise covariances R^i	$diag(25km^2, 1e-6(km/s)^2, 9e-6(rad)^2)$
	detection probability P_d^i	0.4
Ionosphere	number of modes t	4
	ionosphere height(h_t, h_r)	(100km, 260km)
	TX-RX distance	100km
DCEM	state transition matrix F	$I_2 \otimes \begin{bmatrix} 1 & 20 \\ 0 & 1 \end{bmatrix}$
	measurement matrix h^i and its jacobian matrix H^i	see [4] and its correction [33]
	process noise covariance Q	$blockdiag\left(\begin{bmatrix} 7.8e-5 & 4.4e-5 \\ 4.4e-5 & 1.3e-5 \end{bmatrix}, \begin{bmatrix} 1.5e-11 & 1.1e-12 \\ 1.1e-12 & 1.1e-13 \end{bmatrix}\right)$
	initial state covariance $\Sigma_{0\|0}$	$diag(25, 1e-6, 9e-6, 4.5e-8)$
	window length l	5
	Weight of the consensus η	0.3
	gate probability P_g	0.971
	iterative terminated threshold δ_L	$1e-5$
	maximum number of iterations r_{max}	30

[10] G. Pulford, and B. Scala, “Multiregion Viterbi Data Association Tracking for Over-the-Horizon Radar,” *CSSIP Fourth Report to DSTO*, 1997, Report No. 20 / 97.

[11] B. Scala, and G. Pulford, “Implementation of a Viterbi Data Association Tracker for Manoeuvring Targets in Over-the-Horizon Radar,” *CSSIP Report to DSTO*, 1997, Report No. 12 / 97.

[12] D. Percival, and K. White, “Multihypothesis Fusion of Multipath Over-the-horizon Radar Tracks,” In *Processings of the SPIE*, Florida, 1998, 3373, pp. 440-451.

[13] D. Percival, and K. White, “Multihypothesis Fusion of Multipath Over-the-horizon Radar,” In *Processings of the SPIE*, 1997, 3163, pp. 363-374.

[14] P. Sarunic, K. White, and M. Rutten, “Over-The-Horizon Radar Multipath and Multisensor Track Fusion Algorithm Development,” *Defence Science and Technology Organisation Salisbury Electronics and Surveillance Research*, 2001, No. DSTO-RR-0223.

[15] G. Pulford, “OTHR Multipath Tracking with Uncertain Coordinate Registration,” *IEEE Transactions on Aerospace and Electronic Systems*, 2004, 40(1), pp.38-56.

[16] H. Liu, Y. Liang, Q. Pan and Y. Cheng, “A Multipath Viterbi Data Association Algorithm for OTHR,” *Proceedings Of 2006 CIE International Conference On Radar*, Shanghai, China, 2006.

[17] T. Sathyan, T. Chin, S. Arulampalam, and D. Suter, “A Multiple Hypothesis Tracker for Multitarget Tracking with Multiple Simultaneous Measurements,” *IEEE Journal of Setected Topics in Signal Porcessing*, 2013, 7(3), pp. 448-460.

[18] B. Habtemariam, R. Tharmarasa, T. Thayaparan, and M. Mallick, “A Multiple-Detection Joint Probabilistic Data Association Filter,” *IEEE Journal of Setected Topics in Signal Porcessing*, 2013, 7(3), pp. 461-471.

[19] G. Pulford, and A. Logothetis, “An Expectation-Maximisation Tracker for Multiple Observations of a Single Target in Clutter,” In *Proceedings of the 36th Conference on Decision and Control*, San Diego, California, 1997, pp. 4997-5003.

[20] H. Lan, Q. Pan, F. Yang and Y. Liang, “Joint multipath data association and fusion for OTHR,” In *Processings 2013 International Conference on Information Fusion*, Istanbul, Turkey, 2013, pp. 1312-1319.

[21] A. Dempster, N. Laird, and D. Rubin, “Maximum Likelihood from Incomplete Data,” *Royal Statistical Society*, 1977, 39(1), pp. 1-38.

[22] G. McLachlan, and T. Krishnan, *The EM Algorithm and Extensions, 2nd Edition*. John Wiley & Sons, 2008.

[23] R. Streit, and T. Luginbuhl, “Maximum likelihood method for probabilistic multi-hypothesis tracking,” In *Processings of the SPIE*, 1994, 2235, pp. 394-405.

[24] H. Lan, Y. Liang, F. Yang, Z. Wang, and Q. Pan, “Joint estimation and identification for stochastic systems with unknown inputs,” *IET Control Theory and Applications*, 2013, 7(10), pp. 1377-1386.

[25] R. Olfati-Saber, ”Distributed Kalman filtering for sensor networks,” *IEEE Conference on 46th Decision and Control*, 2007, pp. 5492-5498.

[26] R. Olfati-Saber, ”Distributed Kalman filtering with embedded consensus filters,” *IEEE Conference on 44th Decision and Control*, 2005, pp. 8179-8184.

[27] R. Olfati-Saber, and J. Shamma, ”Consensus filters for sensor networks and distributed sensor fusion,” *IEEE Conference on 44th European Control COnference*, 2005, pp. 6698-6703.

[28] D. Gu, ”Distributed EM algorithm for Gaussian mixtures in sensor networks,” *IEEE Transactions on Neural Networks*, 2008, 19(7), pp. 1154-1166.

[29] R. Nowak, ”Distributed EM algorithms for density estimation and clustering in sensor networks,” *IEEE Transactions on Signal Processing*, 2003, 51(8), pp. 2245-2253.

[30] L. Xiao, S. Boyd, and S. Lall, ”A scheme for robust distributed sensor fusion based on average consensus,” *IEEE Fourth International Symposium Information Processing in Sensor Networks*, 2005 pp. 63-70.

[31] C. Biernacki, G. Celeux, and G. Govaert, “Choosing starting values for the EM algorithm for getting the highest likelihood in multivariate gaussian mixture models,” *Computational Statistics and Data Analysis*, 2003, 41(3), pp. 561-575.

[32] R. Maitra, “Initializing partition-optimization algorithms,” *IEEE Transcations on Computational Biology and Bioinformatics*, 2009, 6(1), pp. 144-157.

[33] H. Liu, Y. Liang, Q. Pan, and Y. Cheng, “Comments on ‘A Multipath Data Association Tracker for Over-the-Horizon Radar’,” *IEEE Transactions on Aerospace and Electronic Systems*, 2005, 41(3), pp. 1148-1150.