🏆 *reviewed paper*

Moving Objects Tracking in Distributed Maritime Observation Systems

Vasily V. Popovich, Victor I. Ermolaev, Andrei V. Makshanov, Stanislav A. Vlasov

(Prof. V. Popovich, SPIIRAS, #39, 14 Line, St. Petersburg, 199178, e-mail: popovich@oogis.ru)
(Dr. V. Ermolaev, SPIIRAS, #39, 14 Line, St. Petersburg, 199178, e-mail: ermolaev@oogis.ru)
(Prof. A. Makshanov, SPIIRAS, #39, 14 Line, St. Petersburg, 199178, e-mail: makshanov@oogis.ru)
(Dr. S. Vlasov, SPIIRAS, #39, 14 Line, St. Petersburg, 199178, e-mail: svlasov@oogis.ru)

1 ABSTRACT

This paper considers the processes of target tracking complicated by big data gaps in complex media such as Distributed Maritime Observation System (DMOS). The main purpose of DMOS is to support the favorable navigation conditions, monitoring, life-saving on the sea for different ships in harbors, maritime roads and open sea. DMOS can be considered as a heterogeneous distributed computer system, it includes different layers of services at different levels of abstraction: ship, harbor; regional and global levels. Such a framework is based on several satellite and maritime information systems that nowadays favor the integration of maritime data (e.g., AIS, ECDIS, OPTIMARE, GMDSS). Despite of the input data big volume the situations exist when gaps (e.g., time delays from minutes to hours) between target observations' points are different. Well known algorithms of target tracking do not work properly in such situations. The proposed approach describes synthesis of analytical and simulation methods at tactical hypothesis development for the cases when suitable direct analytics is not applicable. Also, a joined artificial techniques' scenario approach to tactical situation hypothesis development is proposed.

2 INTRODUCTION

An experience so far accumulated by this paper's authors at developing grand distributed systems of maritime monitoring shows that disregarding a great number of various interacting systems (e.g., AIS, ECDIS, OPTIMARE, GMDSS, etc) the problem of objects' tracking in the sea remains actual and is somewhat far from its solution. Even in the coastal zone the occurrences arise when the information systems and radars discontinue the release of information as caused by different reasons.

The case when many targets are clumping within one space's resolution gives a rise to other problems. The above problem is known under a special notation "Cloud Targets". In spite of their heterogeneous nature the given problems in their mathematical statement are homothetic. The currently available literature does not present any appropriate mathematical solutions of the above problems, and in this regard this paper proposes and discusses an originally developed approach.

2.1 Statement of the information integration problem in DMOS

In some area functions a distributed maritime observation system (DMOS) that incorporates N information sources and controls. The following elements are used as sources of information about situations:

- Space-distributed monitoring facilities (mobile and immobile), that possess different capabilities in registration of the objects under monitoring, in measuring coordinates and parameters of the objects' motion, and in intensity of the information release to the controls;
- Other distributed observation systems that release their data discretely in an integrated form.

Upon the processing the data from the i – th monitoring facility or from the - th distributed system ( ) are send to the DMOS controls in a form f discrete messages, at that, each message is related to one, j-th object (j = 1,…, M, M – is a number of objects operating within the system function zone) Each message could be represented as vector :

$$C_{ij} = [\ i\ \ j\ K_{ij}\ \varphi_{ij}\ \lambda_{ij}\ h_{ij}\ k_{ij}\ v_{ij}\ P_{ij}\ t_{ij}\ ],$$

Here i – is the information source number;

j – is the object number assigned by the *i-th* information source;

K_{ij} - is the monitoring object class;

φ_{ij} - is the object's locality latitude;

λ_{ij} - is the object's locality longitude;

h_{ij} - is the object's locality height (depth);

k_{ij} – is the object's source;

v_{ij} – is the object's velocity (speed);

P_{ij} – other object's features that can be generated by the monitoring facility, like feature of novelty, feature of object's maneuver attribute, feature of the object's loss, etc.

t_{ij} – the time of an information receipt (may differ from the message receipt time).

For all *i*, *j* the covariance matrices that describe an uncertainty of the *j-th* object coordinates' assessment by the *i-th* source are known;

$$\psi_{ij} = \begin{bmatrix} \sigma_x^2 & K_{xy} \\ K_{xy} & \sigma_y^2 \end{bmatrix},$$

where σ_x^2, σ_y^2 - are variances of the object's locality assessments in the x and y directions correspondingly, K_{xy} – is the above assessments covariance.

The monitoring system controls have available the additional information that specifies a situation within a region and affects the information processing, like information about the:

- Localities of the monitoring system's elements;
- monitoring facilities characteristics (detection range of the objects from various classes, resolution, silent spaces, etc.);
- Region infrastructure (fairwaters, navigable waterways, navigational aids and other);
- Hydrometeorological situation;
- Ice conditions;
- Fishing conditions.

This additional information allow for putting forward certain tactical situational hypotheses that are formalized in a form of exogenous Bayes distributions being sequentially shaped as messages continue to arrive.

DMOS is intended to solve problems of integrating the information arriving from various monitoring facilities and other distributed systems; these problems are as follows:

(a) calculation and construction of the zones of monitoring objects' possible localities.

(b) extrapolation of the discriminated object locality at a definite point of time.

(c) identification and mapping of the tracks for objects followed by the observation system elements.

3 BAYES RECURSIVE ASSESSMENT UNDER RESTRICTIVE STATEMENT

Task of the track analysis [5] represents one of the versions in the problem of recursive non-linear assessment [2], and an extended Kalman filter is the most commonly used approach for the case. The task is exposed to linearization within a neighborhood of the predicted value and the sought density is approximated by the Gaussian density that may not match the real data structure and, in particular, lead to divergence. Other analytic approaches are based upon the approximation of the first two density moments [4, 7]. The better direct numerical approach uses the density assessment at the grid nodes in the state space [3], at that, the case of multi dimensional space requires calculations for an excessively great number of points (nodes), thus, involving considerable computational power for each point (node).

It is assumed that a state vector $x_k \in R^n$ keeps evolving in accordance with the dynamics model

$$x_{k+1} = f_k(x_k, w_k), \qquad (1m)$$

where $f_k : R_n \times R_m \rightarrow R_n$ - is the system transient function, $w_k \in R^m$ - is a sequence of errors with a zero average (mean) and independent of the state vector past and future values. The probability density w_k is assumed known. In terms of a discrete time there exist measurements $y_k \in R^p$ linked to the state vector values by the observation equation:

$$y_k = h_k(x_k, v_k), \qquad (2m)$$

where $h_k : R^n \times R^r \to R^p$ - is the measurements' transient function and $v_k \in R^r$ - is the other sequence with a zero average (mean) and known probability density independent of the past and future system states and noises. The initial density is assumed known for the state vector and function f_i, h_i, i =1,…, k.

Exogenuos information based upon the current tactical situational hypothesis in most cases is formalized as an even distribution in some polygon U_k, independent of the past and future system states and noises. Let us denote the appropriate density as $u_k(x_k)$.

So, at the step k the available information is represented by the sequence of measurements' $\{y_i\}$ and the exogenous density $u_k(x_k)$: $D_k = \{y_i, i = 1,...,k, u_k(x_k)\}$. The task consists in arriving at the values of state vector $p(x_k|D_k)$ based on complete information available by the moment k. This process is performed recurrently in two stages: prognosis and update.

Suppose that the sough density $p(x_{k-1}|D_{k-1})$ at step k-1 is already received. Then using the system dynamics model a priory density can be received as follows:

$$p(x_k|D_{k-1}) = \int p(x_k|x_{k-1})\, p(x_{k-1}|D_{k-1})\, dx_{k-1}. \qquad (3m)$$

The probability model for the state vector $p(x_k|x_{k-1})$ is the Markov one and is determined by the system of equations:

$$p(x_k|x_{k-1}) = \int p(x_k|x_{k-1}, w_{k-1})\, p(w_{k-1}|x_{k-1})\, dw_{k-1}.$$

Under an assumption $p(w_{k-1}|x_{k-1}) = p(w_{k-1})$ we have

$$p(x_k|x_{k-1}) = \int \delta(x_{k-1} - f_{k-1}(x_{k-1}, w_{k-1}))\, p(w_{k-1})\, dw_{k-1},$$

Where $\delta(\cdot)$ - the Dirac delta function. Delta function appears since under the unknown x_{k-1} and w_{k-1} the assessment x_k is received from the determinate relation (1m). At the step k become available the measurement y_k and exogenous distribution $u_k(x_k)$ that are used to update a priori distribution of the state vector based on the Bayes rule

$$p(x_k|D_k) = \frac{p(y_k|x_k)\, p(x_k|D_{k-1})\, u_k(x_k)}{p(y_k|D_k)}, \qquad (4m)$$

where a normalizing multiplier is determined through the relation

$$p(y_k|D_k) = \int p(y_k|x_k)\, p(x_k|D_{k-1})\, u_k(x_k)\, dx_k. \qquad (5m)$$

The conditional density y_k under the given x_k, $p(y_k|x_k)$ is determined by the measurements' model and by the known statistics distribution v_k:

$$p(y_k|x_k) = \int \delta(y_k - h_k(x_k, v_k))\, p(v_k)\, dv_k. \qquad (6m)$$

In the equation (4m) the measurement y_k and density u_k are used for a correction of the prognosis starting from the preceding step and the receiving of a posteriori distribution for the state vector.

The equations (4m) and (6m) give a formal solution for the problem of the Bayes recurrent assessment. Here an analytical solution is only possible within a restricted class of evolution models, measurements and exogenous information. The case when f_k and h_k are linear, w_k and v_k are the additive Gaussian noises with the known covariance matrices, and u_k are the Gaussian densities is considered to be the most important, though for many applications this problem statement turns out the inadequate one.

4 RESTRICTIVE BAYES BOOTSTRAP-FILTER

Suppose that there exists a set of independent realizations for a random vector

$$\{ x_{k-1}(i),\ i = 1,\dots, N \},$$

conforming to the distribution $p(x_{k-1}|D_{k-1})$. Bootstrap-filter [2,3] is an algorithm allowing based on the above to receive a set of independent realizations { $x_k(i)$, i=1,…, N }, approximately conforming to the distribution $p(x_k|D_k)$ Such a filter presents a mechanism approximately emulating relations (3m) and (4m).

Prognosis: each sampling realization is transformed in accordance with the formula

$$x_k{*}(i) = f_{k-1}(x_{k-1}(i), w_{k-1}(i)),\ i = 1,\dots,N\ ,$$

Where $w_{k-1}(i)$ – are sampling realizations of the random vector with distribution $p(w_{k-1})$. So, in accordance with (3m), $x_k{*}(i)$ - is a vector sample matching distribution $p(x_k|D_{k-1})$.

Update: Based on y_k measurement the likelihood function is derived and weight coefficients are received for each realization:

$$q_i = \frac{p(y_k|x_k{*}(i))\cdot u_k(x_k{*}(i))}{\sum_{j=1}^{N} p(y_k|x_k{*}(j))\cdot u_k(x_k{*}(j))}\ . \qquad (7m)$$

In this way is defined the discrete distribution over the ensemble of points $\{x_k{*}(i),\ i=1,\dots,N\}$ with a probabilistic weight q_i in the i-th point. Let us construct out of this discrete distribution N random realizations and generate a sample $\{x_k(i),\ i=1,\dots, N\}$ so that for any i, j

$$P\{x_k(i) = x_k{*}(i)\} = q_i.$$

These prognosis and update procedures compose one the *k-th* iteration step. At the first step the process is being initialized by N realizations of the random vector with a known distribution $p(x_1)$.

The update procedure is based on the results received by Smith and Gelfand in [14]. Their work proves that Bayes formula could be interpreted as a weighted bootstrap [5]. Suppose that set $\{x_k{*}(i)\}$ is received as a random sample out of a continuous distribution with density $G(x)$, and it is necessary to receive a sample out of distribution with a density proportional to $L(x)\ G(x)$, where $L(x)$ – is a known nonnegative function. The Smith-Gelfand theorem claims that an empirical distribution of a random sample out of discrete distribution concentrated in points $\{x_k{*}(i)\}$ with probabilistic weights

$$q_i = \frac{L(x_k{*}(i))}{\sum_{j=1}^{N} L(x_k{*}(j))}$$

is a distribution converging one assumed that $N \to \infty$ to a distribution with sought density. In the case under consideration $G(x)$ replaced by $p(x_k|D_{k-1})$, and $L(x)$ – by $p(y_k|x_k)\cdot u_k(x_k)$.

The main advantage of the proposed approach is that it imposes no constrains on the functions' f_k , h_k and u_k form. The major requirements consist in the following:

- Distribution $p(x_1)$ is known and assumes modeling based on Monte-Carlo techniques;
- Distribution $p(y_k|x_k)$ is known;
- Distributions $p(w_k)$ and $u_k(x_k)$ are known and assume modeling based on Monte-Carlo techniques.

At each step of the filter output appears a vector sample that could be disposed in a variety of ways. Say, a posteriori probability of hitting some zone could be assessed as a rate of the sample values that have hit this zone. In case the reasons exist to suppose that a posteriori distribution is a unimodal one the statistic characteristics for each component of the state vector and their any function.

The proofs of the proposed approach appropriateness are based upon the asymptotic results [8]. To study the bootstrap properties for the finite samples is difficult enough. Moreover, the relation between the samples N volume and the received results accuracy is not clear [2,3], it is influenced upon by three factors:

- The states space dimension;
- The goodness of a priori and a posteriori distributions;

- The number of steps required for recurrence.

It should be expected that N grows rapidly along with the task dimension growth, at that, the rapidity depends on the correlation between the components. Same time under the independent components this rapidity does not depend on the task dimension.

If the state space region where the likelihood function $p(y_k | x_k)$ takes on relatively great values is somewhat restricted in comparison with a similar region for a priori density $p(x_k | D_{k-1})$, many values $x_k*(i)$ will be assigned low weights q_i. At that, the points from the first region will gain the greatest value. In the absence of the system noises all N vectors after several steps can get concentrated in the vicinity of one point. So, modification of the basic algorithm will be possible together with the growth of the points' number emulated by a priory distribution

5 CONCLUSION

Simulation and the developed real time algorithms operation demonstrate their good convergence. Probability of the tracks' recovery for 24 objects within one "cloud" composes over 80%. On the other hand the coordination of data received from heterogeneous sources in one "cloud" is not yet a completely solved problem for the case of large blanks. This problem is not expected to be solved exclusively via mathematic and scenario approach; its solution requires special organizational efforts in information processing and priorities identification among the sources.

6 REFERENCES

1. Gordon N., Salmond D., Smith A. Novel approach to nonlinear/nongaussian state estimation. – Proc. Inst. Elect. Eng., ser.F, 1993, v.40, No. 2, pp.107-113.
2. Doucet A., de Freitas N. (editors). Sequential Monte-Carlo methods in practice. – New York: Springer-Verlag, 2001.
3. Carlin B.P., Polson N.G., Stoffer D.S. A Monte-Carlo approach to nonnormal and nonlinear state space modeling. – JASA, 1992, No. 87, pp. 493-500.
4. West M., Harrison P.J., Migon H.S. Dynamic generalized linear model and Bayesian forecasting (with discussion). – JASA, 1985, 80, pp.73-97.
5. Эфрон Б. Нетрадиционные методы многомерного статистического анализа. – М.: Финансы и статистика, 1988.
6. Chang K.C., Bar-Shalom Y. Joint Probabilistic Data Association for Multitarget Tracking with Possibly Unresolved Measurements and Maneuvers. – IEEE Trans. Auto. Control, 1984, vol. AC-29, pp.585-594.
7. Alspach D.L., Sorensen H.W. Nonlinear Bayesian estimation using Gaussian sum approximation. – IEE Trans. Auto. Control, 1972, vol. AC-17, pp.439-447.
8. Smith A.F.M., Gelfand A.E. Bayesian statistics without tears: a sampling-resampling perspective. – Amer. Stat., 1992, No.46, pp. 84-88.
9. Intelligent GIS for Maritime Observation Systems. Edited by V. Popovich. M.: Nauka, 2013, 283pp
10. Vasily Popovich, Christophe Claramunt. Distributed Maritime Observation Systems for Navigation //Proceedings of the ENC, Vienna, April 23-25, 2013.