ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2023, Vol. 84, No. 4, pp. 460–473.
Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, 2023.

CONTROL IN TECHNICAL SYSTEMS

Optimization of the Cluster-Variant Method of Constructing a Multi-Position Direction Finding System for Conditions of a Priori Uncertainty

Yu. G. Bulychev[*,a] **and E. N. Chepel'**[*,b]

[*] *AO "VNII "Gradient", Rostov-on-Don, Russia*
e-mail: [a] *profbulychev@yandex.ru,* [b] *evgeny_c@bk.ru*

Received November 19, 2020
Revised September 13, 2022
Accepted October 26, 2022

Abstract—The possibility of constructing a multi-position direction finding system for the case of a priori uncertainty, based on the application of the principles of multiplication of single marks of the location of the emitting target (multistructure principle) and their subsequent partition into classes (clustering principle) is considered. The criteria and algorithms for detecting the resulting cluster and for constructing the optimal estimation of target location stable to anomalous measurement errors are presented, taking into account the time costs of their computer realization. Practical recommendations and results of comparative analysis of different algorithms are given.

Keywords: radiating target, multi-position direction finding system, target location, bearing, cluster, selection, multistructure principle, clustering principle, optimality criteria

DOI: 10.25728/arcRAS.2023.42.90.002

1. INTRODUCTION

In [1, 2] the cluster-variant method (CVM) of solving the triangulation problem (which is alternative to the known methods of passive location, for example in [3–20]) is developed, allowing under some limitation on the number of unreliable azimuthal and angular measurements to form stable estimates of target location in conditions of significant a priori uncertainty (such as lack of reliable knowledge about distribution laws of measurement errors, gaps in measurements, presence of tool and methodical errors, "degradation" of the structure, etc.). Under such conditions, which often arise in practice, the results of direction finding may allow the presence of unknown anomalous measurement errors (AMEs). These errors can be of very different nature of origin (deterministic unknowns or random ones) and related uncertainty in their formalized description. Often even minimally necessary information for effective application of known adaptive methods of measurement processing, which guarantee the consistency of formed estimations (see, for example, [21–23]), is absent. In addition, these methods do not take into account the specifics of construction and functioning of a multi-position direction finding system, in particular the geometric factor due to the location in space of direction finders and targets. For a large class of such systems, not only some averaged estimate of efficiency is important, but also a guaranteed for a specific sample of measurements (i.e. "here and now") estimate of the target location. The only factor taken into account is often the assumption of the maximum possible number of unreliable azimuth and angle measurements, at which it is still possible to obtain a reliable estimate, taking into account a given number of direction finders.

In [1, 2], the multistructure principle (which consists of multiplying single triangulation marks based on all possible measurement sets that provide the correct mark without bearing errors) and

the clustering principle (which consists of dividing all generated marks into classes based on the chosen optimality criterion) are used to deal with AMEs. In [1], the principle itself of multistructural formation of a family of single marks is outlined, but there is no effective rule for combining them to form the resulting estimate of the target location. This shortcoming is to some extent eliminated in [2], where an algorithm for selecting the resulting cluster is proposed, which allows detecting unreliable measurements, as well as forming a resultant estimate of target motion parameters that is stable to AMEs. We are talking about the construction of a multi-position direction finding system, which is an alternative to classical systems created within the framework of traditional statistical methods (maximum a posteriori probability density, maximum likelihood, least squares, and minimum of various misalignments [3–13]), which have proved themselves well for the case of correct measurements (in the absence of AMEs).

However, [2] introduces a serious restriction: the number of partitioning clusters of unit marks must be specified a priori. Furthermore, [2] does not take into account the possibility of choosing the optimal resulting cluster and the optimal resulting estimate of the target location. The possibilities of CVM decomposition and the issues of operability of resulting estimate formation are not investigated as well.

This paper further develops CVM [2] in terms of its optimization, including the introduction of better optimality criteria and assumptions regarding the parameters of the clusters used. We demonstrate the possibility to choose the most efficient (in terms of accuracy) algorithm for processing bearing in abnormal measurement conditions, taking into account its possible decomposition (two-step approach) and time consumption for obtaining the resulting estimate of the target location by each algorithm of this family.

2. PROBLEM FORMULATION

In order to compactly describe the proposed method, we will mainly consider the case of a single target, which allows us to omit cumbersome notations and calculations. The generalization of the method to the case of an unknown number of many targets is considered in Section 6. In the XYZ Cartesian coordinate system, we consider a multi-target bearing system (hereafter simply the system), consisting of a set of direction finders ($\Pi_m,\ m = \overline{1, M}$), where the position of each Π_m is given by the vector $\boldsymbol{\rho}_m = [\rho_{xm}, \rho_{ym}, \rho_{zm}]^{\mathrm{T}}$. The true position of the target is characterized by the vector $\boldsymbol{\lambda}_{\mathrm{tr}} = [x_{\mathrm{tr}}, y_{\mathrm{tr}}, z_{\mathrm{tr}}]^{\mathrm{T}}$ with the constraint $\boldsymbol{\lambda}_{\mathrm{tr}} \in \boldsymbol{\Lambda} = \Lambda_x \times \Lambda_y \times \Lambda_z$, the vector $\boldsymbol{\lambda} = [x, y, z]^{\mathrm{T}}$ is used for an arbitrary (model) point of the XYZ system.

Vectors $\mathbf{Y}_\alpha = \left[\tilde{\alpha}_m, m = \overline{1, M}\right]^{\mathrm{T}}$ and $\mathbf{Y}_\beta = \left[\tilde{\beta}_m, m = \overline{1, M}\right]^{\mathrm{T}}$ of primary measurements of the azimuth $\tilde{\alpha}_m = \alpha_m + \Delta\alpha_m$ and the spot angle $\tilde{\beta}_m = \beta_m + \Delta\beta_m$ respectively can be assigned to the system's bearing, where $\Delta\alpha_m$ and $\Delta\beta_m$ are measurement errors with an unknown law distribution.

Let us denote by ς_α the number of azimuthal measurements that do not contain AMEs, and by ς_β the angular ones. We suppose that

$$\gamma_\alpha < \varsigma_\alpha \le M, \quad \gamma_\beta < \varsigma_\beta \le M, \tag{1}$$

where γ_α and γ_β are natural numbers that set the maximum permissible number of unreliable measurements by azimuth and elevation, respectively, and the choice of γ_α and γ_β (along with the number M) should take into account the possibility of forming the required number of sets (they were discussed in the introduction), ensuring the effective application of the previously mentioned principles of multistructure and clustering.

We assume that the kth and lth sets must not coincide for all $k, l \in \{1, \ldots, \overline{N}\}$, $k \ne l$ where $\overline{N}$ corresponds to the minimum required number of $\gamma_{\alpha\beta} = \gamma_\alpha + \gamma_\beta$ azimuthal and angular measurements, sufficient to construct the set. We will also assume that the number of measurements that

do not contain AMEs should be at least $D_{\alpha\beta} = D_{\alpha s} + D_{\beta s}$, $s = \overline{1,S}$, $D_{\alpha s}, D_{\beta s} \in \{1,\dots,M\}$, where $D_{\alpha s}$, $D_{\beta s}$ are the components for the sth representation variant of the number $D_{\alpha\beta}$ which provides the construction of the set (S is the total number of such variants). It is clear that in order to build a set it is necessary that $D_{\alpha s} \geq \gamma_\alpha$ and $D_{\beta s} \geq \gamma_\beta$. The total amount of possible sets for the $(\gamma_\alpha, \gamma_\beta)$ pair is equal to

$$\overline{N} = \left(\sum_{n=\gamma_\alpha}^{M} C_M^n\right)\left(\sum_{n=\gamma_\beta}^{M} C_M^n\right), \tag{2}$$

respectively, the number of sets formed by the measurements that do not contain AEMs (for the sth variant) is

$$\overline{\overline{N}}_s = \left(\sum_{n=\gamma_\alpha}^{D_{\alpha s}} C_{D_{\alpha s}}^n\right)\left(\sum_{n=\gamma_\beta}^{D_{\beta s}} C_{D_{\beta s}}^n\right), \tag{3}$$

and their minimal amount (for all variants)

$$\overline{\overline{N}}_{\min} = \min_s \overline{\overline{N}}_s. \tag{4}$$

In (2) and (3), the corresponding numbers of combinations appear in parentheses under the sum sign.

According to the results of measurements for all sets of triangulation marks are plotted and $\alpha_{m[k]} = \alpha_m + \Delta\alpha_{m[k]}$ and $\beta_{m[k]} = \beta_m + \Delta\beta_{m[k]}$ secondary bearings are calculated, and it is assumed that for the sets containing no measurements with AMEs, the following conditions are fulfilled:

$$\begin{aligned} \left|\Delta\alpha_{m[k]}\right| &= \left|\alpha_{m[k]} - \tilde{\alpha}_m\right| < \varepsilon_{\alpha m}, \\ \left|\Delta\beta_{m[k]}\right| &= \left|\beta_{m[k]} - \tilde{\beta}_m\right| < \varepsilon_{\beta m}, \end{aligned} \qquad m = \overline{1,M}, \tag{5}$$

where $\varepsilon_{\alpha m} > 0$ and $\varepsilon_{\beta m} > 0$ are set thresholds.

To perform clustering (selection) of marks, the number of conditions fulfilled is checked (separately by azimuth and elevation angle): if for the kth set this number does not satisfy the constraints in (5) or the mark does not belong to the $\mathbf{\Lambda}$ area, the set in question is discarded.

The remaining secondary sets after selection (their amount is $N \leq \overline{N}$) correspond to the marks $\boldsymbol{\lambda}_{[n]}$, where $n = \overline{1,N}$, $N \leq \overline{N}$. The formed set $\mathbf{X} = \left\{\boldsymbol{\lambda}_{[1]},\dots,\boldsymbol{\lambda}_{[N]}\right\}$ is divided into clusters K_q, $q = \overline{1,Q}$ (hierarchical agglomerative algorithm with Euclidean norm is used for clustering by analogy with [24–28]). According to this algorithm, the marks $\boldsymbol{\lambda}_{[1]},\dots,\boldsymbol{\lambda}_{[N]}$ are grouped in succession (in steps): first the closest, and then more and more distant from each other. In the first step, each mark is treated as a separate cluster.

The clustering algorithm sets the following mapping $f : \mathbf{X} \to \{\mathrm{K}_q, q = \overline{1,Q}\}$, $\mathrm{K}_q = \{\lambda \in \mathbf{X} | f(\lambda) = q\}$, with $\mathbf{X} = \bigcup \mathrm{K}_q$ (where $\bigcup$ is the cluster association symbol by index $q = \overline{1,Q}$) $\mathrm{K}_k \bigcap \mathrm{K}_r = \varnothing$, $k, r \in \{1,\dots,Q\}$, $k \neq r$ and $\mathrm{K}_q \neq \varnothing$ $\forall q = \overline{1,Q}$.

To solve the problem of finding the best cluster and constructing the optimal resulting estimate corresponding to it (the unified detection-estimation problem), we use the following solver function:

$$F(\boldsymbol{\lambda}, q) = \sum_{m=1}^{M}\left[\left(\frac{\alpha_m(\boldsymbol{\lambda}) - \tilde{\alpha}_m}{2\pi}\right)^2 W_{\alpha m}(q) + \left(\frac{\beta_m(\boldsymbol{\lambda}) - \tilde{\beta}_m}{2\pi}\right)^2 W_{\beta m}(q)\right], \tag{6}$$

where $\alpha_m(\boldsymbol{\lambda})$ and $\beta_m(\boldsymbol{\lambda})$ are model bearing, $W_{\alpha m}(q)$ and $W_{\beta m}(q)$ are dimensionless normalized weight coefficients.

In (6), the number 2π is used for normalization and to obtain the dimensionless function $F(\boldsymbol{\lambda}, q)$, and the partial weights are given as follows:

$$\begin{aligned} W_{\alpha m}(q) &= L_q^{-1} \sum_{l=1}^{L_q} \varphi\left(\frac{|\alpha_{mql} - \tilde{\alpha}_m|^2}{\varepsilon_{\alpha m}^2}\right), \\ W_{\beta m}(q) &= L_q^{-1} \sum_{l=1}^{L_q} \varphi\left(\frac{\left|\beta_{mql} - \tilde{\beta}_m\right|^2}{\varepsilon_{\beta m}^2}\right), \end{aligned} \tag{7}$$

where L_q is the number of marks in the cluster K_q, α_{mql} and β_{mql} are secondary bearings of the mark $\boldsymbol{\lambda}_{ql}$ of the cluster K_q, $l \in \{1, \ldots, L_q\}$, $\varphi(p) = 1 - p$ for $p \le 1$, $\varphi(p) = 0$ for $p > 1$. It is obvious that $0 \le W_{\alpha m}(q) \le 1$ and $0 \le W_{\beta m}(q) \le 1$.

The formula (7) allows taking into account the number of marks (L_q), the thresholds of acceptable errors of secondary direction finding ($\varepsilon_{\alpha m}, \varepsilon_{\beta m}$) and the presence of AEMs (by introducing the indicator function $\varphi(p)$). The coefficients $W_{\alpha m}(q)$ and $W_{\beta m}(q)$ show the specific contribution of bearings $\tilde{\alpha}_m$ and $\tilde{\beta}_m$ to the formation of all K_q cluster marks. They can be called the coefficients of correspondence of the bearings $\tilde{\alpha}_m$ and $\tilde{\beta}_m$ to the marks of the cluster K_q. The greater the value of the coefficient, the greater the confidence in the primary bearing to which this coefficient is related. In classical triangulation estimation such role is played by coefficients inversely proportional to the square of error variances $\Delta\alpha_m$ and $\Delta\beta_m$.

The criterion for optimal detection-evaluation comes down to the following:

$$\boldsymbol{\lambda}^* = \arg\min_{\boldsymbol{\lambda}} F(\boldsymbol{\lambda}, q^*), \tag{8}$$

$$q^* = \arg\max_q W(q) = \arg\max_q \left\{ M^{-1} \sum_{m=1}^{M} [W_{\alpha m}(q) + W_{\beta m}(q)] \right\}, \tag{9}$$

where $W(q)$ is the integral dimensionless normalized weight coefficient, $0 \le W(q) \le 1$.

Taking into account the (1)–(9), it is required to: construct an algorithm for solving the triangulation problem in the optimal formulation; consider a quasi-optimal computationally economical two-step CVM (first by azimuth and then by elevation); based on the analysis of known clustering algorithms, substantiate an effective algorithm for combining partial triangulation marks into chained clusters; conduct a numerical comparative analysis of algorithms A_1 (for known CVM), A_2 and A_3 (respectively for the developed optimal and two-stage quasi-optimal) in accuracy and efficiency; give a numerical comparative analysis with the maximum likelihood method (for conventional and extended options).

3. OPTIMAL ALGORITHM FOR SOLVING THE TRIANGULATION PROBLEM

To calculate the model bearings, we will use the well-known formulas for the Cartesian and Cartesian-rectangular coordinates relation:

$$\begin{cases} \alpha_m(\boldsymbol{\lambda}) = \arccos\left\{(x - \rho_{xm})\left[(x - \rho_{xm})^2 + (y - \rho_{ym})^2\right]^{-1/2}\right\}, \\ \beta_m(\boldsymbol{\lambda}) = \arcsin\left\{(z - \rho_{zm})\left[(x - \rho_{xm})^2 + (y - \rho_{ym})^2 + (z - z_{ym})^2\right]^{-1/2}\right\}. \end{cases} \tag{10}$$

Minimization of the solver function $F(\boldsymbol{\lambda}, q)$ on the vector argument $\boldsymbol{\lambda}$ leads to the equation

$$\left[\frac{\partial F(\boldsymbol{\lambda}, q)}{\partial \boldsymbol{\lambda}}\right]^{\mathrm{T}} = \mathbf{0}, \tag{11}$$

where $\mathbf{0} = [0, 0, 0]^{\mathrm{T}}$.

After expanding the partial derivatives in (11), we get a system of scalar equations

$$\begin{cases} \sum\limits_{m=1}^{M} \left[\left(\dfrac{\alpha_m(\boldsymbol{\lambda}) - \tilde{\alpha}_m}{2\pi} \right) W_{\alpha mq} \dfrac{\partial \alpha_m(\boldsymbol{\lambda})}{\partial x} + \left(\dfrac{\beta_m(\boldsymbol{\lambda}) - \tilde{\beta}_m}{2\pi} \right) W_{\beta mq} \dfrac{\partial \beta_m(\boldsymbol{\lambda})}{\partial x} \right] = 0, \\ \sum\limits_{m=1}^{M} \left[\left(\dfrac{\alpha_m(\lambda) - \tilde{\boldsymbol{\alpha}}_m}{2\pi} \right) W_{\alpha mq} \dfrac{\partial \alpha_m(\boldsymbol{\lambda})}{\partial y} + \left(\dfrac{\beta_m(\boldsymbol{\lambda}) - \tilde{\beta}_m}{2\pi} \right) W_{\beta mq} \dfrac{\partial \beta_m(\boldsymbol{\lambda})}{\partial y} \right] = 0, \\ \sum\limits_{m=1}^{M} \left[\left(\dfrac{\alpha_m(\boldsymbol{\lambda}) - \tilde{\alpha}_m}{2\pi} \right) W_{\alpha mq} \dfrac{\partial \alpha_m(\boldsymbol{\lambda})}{\partial z} + \left(\dfrac{\beta_m(\boldsymbol{\lambda}) - \tilde{\beta}_m}{2\pi} \right) W_{\beta mq} \dfrac{\partial \beta_m(\boldsymbol{\lambda})}{\partial z} \right] = 0, \end{cases} \tag{12}$$

where the partial derivatives are disclosed considering (10).

The system (12) corresponds to partial estimates $\boldsymbol{\lambda}^*(q)$, $q = \overline{1, Q}$. As the resultant of $\boldsymbol{\lambda}^* \in \{\boldsymbol{\lambda}^*(1), \ldots, \boldsymbol{\lambda}^*(Q)\}$ that partial estimate $\boldsymbol{\lambda}^*(q^*)$ (where $q^* \in \{1, \ldots, Q\}$) is chosen that, according to (9), meets the criterion

$$W(q^*) > W(q), \quad q \neq q^*. \tag{13}$$

The choice of parameter values appearing in conditions (1)–(5) can always ensure the uniqueness and accuracy of the solution of the triangulation problem, given (10)–(13). Thus, with increasing values of parameters M, ς_α and ς_β, reliable detection of the resulting cluster K_{q^*} containing the largest number of marks formed by valid structures is always ensured. In turn, the choice of parameters $\varepsilon_{\alpha m}$ and $\varepsilon_{\beta m}$ affects the accuracy of triangulation estimation.

If we restrict ourselves to a single cluster, remove the restrictions in (5), and assign the weight coefficients $W_{\alpha m}(q)$ and $W_{\beta m}(q)$ to the inverse values of the variances of the direction finding errors (distributed according to the normal law), then the above solution in (10)–(13) will correspond to the known maximum likelihood triangulation estimate for the case of no AMEs (see, for example, [5, 10, 11, 13]).

The algorithm for solving the triangulation problem in the cluster optimal formulation is reduced to the following:

1. Clustering of triangulation marks is performed by constructing clusters K_q, $q = \overline{1, Q}$ (recommendations for clustering are given in the fourth section).
2. For each cluster K_q the number of marks (L_q), partial and integral weight coefficients ($W_{\alpha m}(q)$, $W_{\beta m}(q)$ and $W(q)$) are calculated.
3. Using criterion (13) solve the detection problem, i.e. find the number q^* of the resulting cluster K_{q^*}.
4. For the cluster K_{q^*} solve the system of equations (12), choosing the Cartesian coordinates of the center of this cluster as the initial condition. As a result, given (13), we obtain the resulting (optimal) estimate of the target location $\boldsymbol{\lambda}^* = \boldsymbol{\lambda}^*(q^*) = [x(q^*), y(q^*), z(q^*)]^{\mathrm{T}}$.

Remark. If we use the known approximate approach to construct a maximum-likelihood triangulation estimate (see, for example, [13]), based on approximation of the residuals $\alpha_m(\boldsymbol{\lambda}) - \tilde{\alpha}_m$ and $\beta_m(\boldsymbol{\lambda}) - \tilde{\beta}_m$ using approximate data on the range to the target (in this case, the range to the cluster center), then instead of (12) we can obtain an appropriate linear system of algebraic

equations. In this case there is no need to select an initial condition, which is typical for nonlinear system of equations in (12).

In the classical statistical formulation, the triangulation problem is solved on the basis of all measurements (regardless of whether they are good or bad) subjected to joint optimal processing, taking into account the pre-assigned a priori weights. Algorithm A_2 also operates with all measurements, but with consideration of a posteriori weights ($W_{\alpha m}(q)$, $W_{\beta m}(q)$ and $W(q)$) which are generated directly from the direction finding results using a family of sets. Combined result estimation, which implies simultaneous consideration of the a priori and a posteriori weights, is also possible.

The implementation of the proposed A_2 algorithm (taking into account the preparation of input data) is associated with the following computational costs $\Gamma = \Gamma_1 + \Gamma_2 + \Gamma_3$, where Γ_1 is the cost of constructing a family of marks, Γ_2 is the cost of clustering the marks and selecting the resulting cluster, and Γ_3 is the cost of constructing the resulting estimate. When solving the triangulation problem in a linear (approximate) version, the main costs Γ_1 are associated with solving a set (volume $\overline{N}$) of systems of linear equations with a square matrix of size 3×3. In this case, a parallel process of obtaining desired solutions is organized, which is most important for real-time processing of measurements in a multi-position direction finding system. The Γ_2 costs are mainly related to the simplest operations of finding a set (volume $N(N-1)/2$) of Euclidean distances between marks (that have undergone selection) and sorting them into different clusters, Γ_3 is related to solving a set (volume L_q) of systems of linear equations with a square matrix of size 3×3. Implementation of the above operations in specialized computing environments is not particularly difficult and allows us to provide a real-time mode. Comparative results of realization of CVM in a universal computer environment are shown in Section 7.

4. RECOMMENDATIONS FOR MARK CLUSTERING

In case of the A_1 algorithm, detection of invalid measurements and construction of the resulting estimate are related to the cluster core (this is its densest part plus marks of close neighboring clusters), which is effective when there are only spherical clusters. When working with clusters of chain-shaped form, the accuracy of the estimation based on the A_1 algorithm significantly deteriorates. This is confirmed by a real example, the results of which are shown in Figs. 1 and 2. Here the options for grouping the marks for a target observed at an angle of 45 degrees from the horizontal axis of the plane coordinate system are shown. These figures correspond to experimental results for the flat case and the system consisting of five azimuthal direction finders evenly spaced on a circle with a radius of 10 km and centered at the origin (Fig. 1 corresponds to a range to the target of approximately 20 km relative to the circle center, Fig. 2 to 50 km), bearing measurements were accompanied only by acceptable random errors (without AMEs), the solid lines correspond to the direction of direction finder-target, the points correspond to the marks that have passed selection. The scales of Figs. 1 and 2 are different, whereas in Fig. 1 the division value in abscissa and ordinates axis are the same and are 50 m, in Fig. 2—200 m in abscissa and 50 m in ordinates.

In the well-known CVM, the introduction of kernels was designed to minimize the errors in triangulation estimation when Q was not chosen well. Let Q be one more than the real number of clusters (e.g, due to an error of the system operator), then some of them will be artificially split into two clusters. If the priority cluster corresponding to the true position of the target is subjected to such a partitioning, it can lead to a significant increase in the estimation error. Figure 3 shows such a case, where black dots correspond to marks assigned to the first cluster (the target is in this cluster), gray dots correspond to the second cluster, $\square$ and $\times$ marks show centers of clusters and true target mark respectively, ellipse shows some neighborhood of the first cluster center (marks of first and second clusters, which fall in this neighborhood, make the first cluster core), shaded square

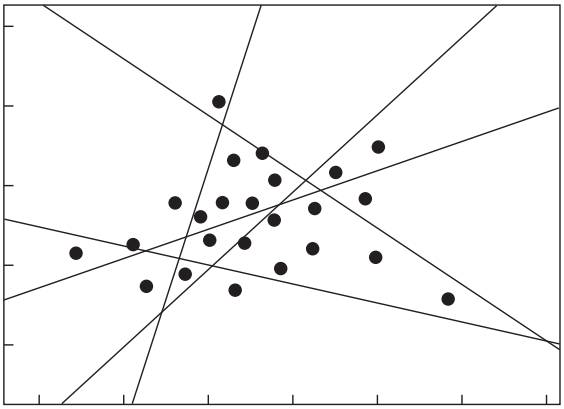

Fig. 1. Formation of a spherical cluster.

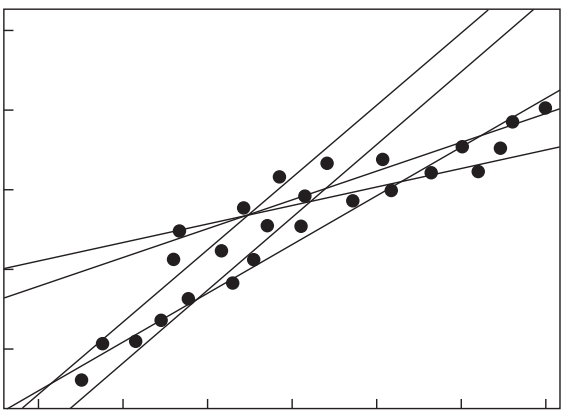

Fig. 2. Formation of a chained cluster.

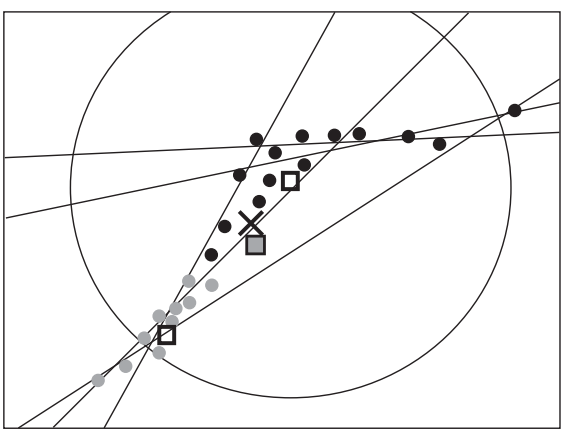

Fig. 3. Formation of the cluster kernel.

marks center of the first cluster core. Due to the mistake made (splitting the original chain cluster into two smaller clusters) the following negative consequences are possible: an error in selecting the priority cluster and, as a consequence, a significant bias in the resulting estimate; the processing completely ignores the marks of the competing (second) cluster; the number of useful marks from the original chained cluster, which could participate in the processing, decreases sharply. The situation can be somewhat improved by introducing kernels (one of them, corresponding to the

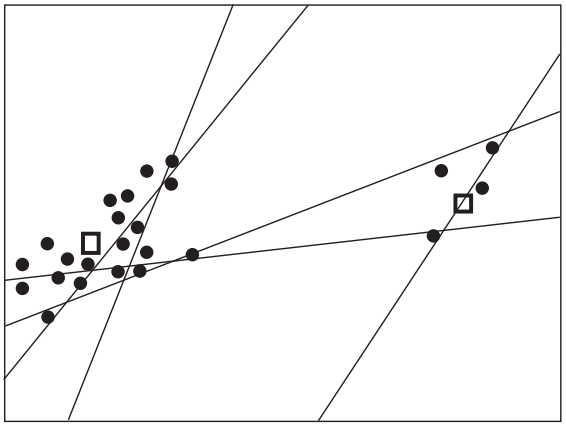

Fig. 4. Location of partial triangulation marks for the flat case.

first cluster, is shown in Fig. 3 in the form of an ellipse). We see that the core of the first cluster "actively" captures the marks of both the first and the second clusters, which increases the number of marks involved in processing, and hence reduces the bias of the resulting score. But even in this case a part of useful marks of the initial chained cluster drops out of the resulting estimation. In the spatial case, when we work with a much larger number of structures, this problem becomes most pronounced.

Figure 4 shows the location of marks for the flat case with allowable random errors in the channels of four azimuthal direction finders $(\Pi_1, \ldots, \Pi_4)$ and the anomalous error (11 degrees) in the channel of the fifth direction finder (Π_5). In Fig. 3, the centers of clusters K_1 (left) and K_2 (right) are marked with a □ (the third cluster K_3 is at a considerable distance and is not shown in the figure). For the flat case and the system of five direction finders we have $\overline{N} = 26$ and $N = 25$ (i.e., twenty-five out of twenty-six markers have passed selection, with twenty markers in K_1, four markers in K_2, and one marker in K_3).

It should also be noted that, in practice, the construction of the kernel requires a certain neighborhood of the cluster center. The complexity of choosing such a neighborhood is caused by the fact that as the distance to the target and/or its appearance in triangulation-incorrect directions grows, the size of clusters also grows. This also forces to modify known CVM and use new approach to choose number of clusters of triangulation marks partitioning.

For the fixed k marks are divided into $Q_k = Q_0 + k$ clusters (where $Q_0 \geq 2$, $k \in \{0, 1, \ldots\}$). The partitioning of marks into clusters should end at step $k = k^*$, when

$$k^* = \arg\max_k S(k), \quad k \in \{0, 1, \ldots\}, \tag{14}$$

where $S_k = S(k)$ is a solving convex function that depends on the cluster parameters and has a maximum point. The value $k = k^*$ corresponds to the optimal number of partitioning clusters $Q_{k^*} = Q^*$.

Analysis of existing cluster theory showed that the most suitable procedure for working with chained clusters as applied to the A_2 algorithm is the automatic non-threshold calculation of the number (Q) of clusters K_q using the "silhouette" coefficient [28]:

$$S(k) = \frac{1}{N} \sum_{q=1}^{Q_k} \sum_{l=1}^{L_q} s(q, l), \quad k \in \{0, 1, \ldots\}, \tag{15}$$

where $s(q,l) = 1 - a(q,l)/b(q,l)$ when $a(q,l) < b(q,l)$; $s(q,l) = 0$ when $a(q,l) = b(q,l)$; $s(q,l) = b(q,l)/a(q,l) - 1$ when $a(q,l) > b(q,l)$; $a(q,l)$ is the average distance from the lth element $(\boldsymbol{\lambda}_{ql})$ of

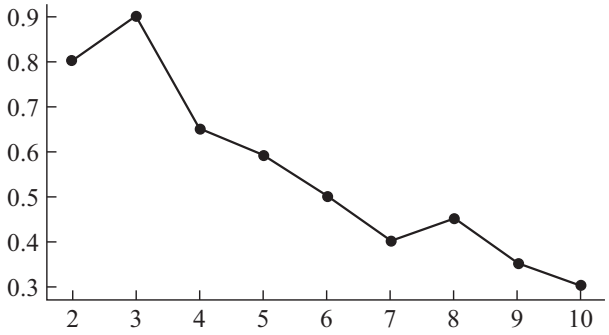

Fig. 5. Dependence of the "silhouette" coefficient on the number of clusters.

the cluster K_q to the other elements of the same cluster; $b(q,l)$ is the average distance from the lth element ($\boldsymbol{\lambda}_{ql}$) of the cluster K_q to the elements ($\boldsymbol{\lambda}_{r1},\ldots,\boldsymbol{\lambda}_{rL_r}$) of the "neighboring" cluster K_r.

The following formulas are used for calculations:

$$a(q,l)=\frac{1}{L_q-1}\sum_{\substack{m=1\\ m\neq l}}^{L_q}|\boldsymbol{\lambda}_{ql}-\boldsymbol{\lambda}_{qm}|\,,$$
$$b(q,l)=\min_{r\neq q}\frac{1}{L_r}\sum_{m=1}^{L_r}|\boldsymbol{\lambda}_{ql}-\boldsymbol{\lambda}_{rm}|\,. \tag{16}$$

Algorithm (14)–(16) splitting the family of triangulation marks into clusters is implemented autonomously (without operator participation), is simple enough for computer implementation and does not require large computational costs. The results of application of cluster algorithms based on the "silhouette" coefficient show their high efficiency and good asymptotic properties in various domains.

Recall that in algorithm A_1 the determination of the number of clusters depends on a threshold value, in the choice of which there is uncertainty, which requires the participation of an experienced operator in solving the cluster partitioning problem. This is due to the fact that the solver function used in algorithm A_1 is monotonically decreasing (with increasing $k\in\{0,1,\ldots\}$) and has no maximum point.

For the example considered in the previous section, the value of the silhouette coefficient for the parameter $Q_k\in\{2,3,\ldots,10\}$ is shown in Fig. 5. It can be seen that if the initial condition $Q_0=2$ is chosen, the maximum is reached when $k=k^*=1$, resulting in $Q_{k^*}=Q^*=3$. Thus, the optimal solution would be to divide all marks into three clusters, which agrees with the data presented in Fig. 4 (i.e., the real number of clusters is equal to the calculated value).

5. TWO-STAGE QUASI-OPTIMAL ALGORITHM

The application of the two-step approach to the implementation of CVM implies an increase in the speed of estimation. For this purpose, let us consider the main provisions of the A_3 algorithm, which is quasi-optimal with respect to the A_2 algorithm.

The *first version* of the A_3 algorithm is based on the fact that it is possible to implement the algorithm discussed in Section 2, first only on azimuthal measurements and sets (the first step), which allows to form a family of plane marks and clusters, and then to determine the estimates x^* and y^* of the two true target coordinates (x_{tr} and y_{tr}). The point (x^*,y^*) in three-dimensional space corresponds to its own target position line. Combining this line and the position cones

(corresponding to the position angles) of all direction finders for different sets, we form new sets, spatial marks and clusters, on the basis of which (based on the algorithm discussed in Section 2) we construct the missing estimate z^* of the true coordinate z_{tr} (second step).

The *second version* of realization of the A_3 algorithm is based on the fact that at the first stage (based on the algorithm considered in Section 2 with regard to azimuthal measurements and sets only), using the coefficients $W_{\alpha m}(q^*)$, unreliable azimuthal measurements are found which are excluded from further processing (step 1). Then, the algorithm from Section 2 (step 2) is implemented on the remaining priority azimuthal and all angular measurements. In this case, the number of the studied sets increases (as compared to the *first version*), which provides an increase in the quality of triangulation estimation in conditions of uncertainty.

Obviously, the two-stage quasi-optimal CVM for the two considered variants does not exhaust the potential of the optimal CVM under uncertainty, but it is more economical from a computational point of view. This is due, first of all, to a significant reduction of the investigated sets, partial triangulation marks, and clusters.

6. GENERALIZING TO THE CASE OF MANY TARGETS

The extension of the developed CVM to this case depends on the purpose, principles of construction and organization of the information-measuring process in a particular system. Two main versions of CVM realization for multipurpose case can be distinguished.

Version 1: When the separation of bearings into classes on the basis of belonging to a given target is separated into a separate stage (the task of bearing identification). Such truncated formulation of a problem is quite widespread in practice at appropriate decomposition of information-measuring process. In this case application of CVM is reduced to the previously considered algorithm with respect to each class. For this purpose the measurement vectors $\mathbf{Y}_j = \left[\mathbf{Y}_{j\alpha}^{\text{T}},\ \mathbf{Y}_{j\beta}^{\text{T}}\right]^{\text{T}}$ are formed, where j is the number of target (class), $j = \overline{1,J}$, $\mathbf{Y}_{j\alpha} = \left[\tilde{\alpha}_{mj}, m = \overline{1,M}\right]^{\text{T}}$ and $\mathbf{Y}_{j\beta} = \left[\tilde{\beta}_{mj}, m = \overline{1,M}\right]^{\text{T}}$. For the jth class (similar to Section 2), the elements K_{jq}, L_{jq}, $W_{j\alpha m}(q)$, $W_{j\beta m}(q)$, $W_j(q)$, q_j^* and the resulting estimations $\boldsymbol{\lambda}_j^* = \boldsymbol{\lambda}_j^*(q_j^*) = \left[x(q_j^*), y(q_j^*), z(q_j^*)\right]^{\text{T}}$ based on these elements are formed, which, taking into account the criteria (8) and (9), provide the solution of the single problem of triangulation estimation of the coordinates of many targets. Version 1 is the most effective from a computational point of view, allowing to organize J parallel channels of measurement processing, but it does not exhaust all potential possibilities of joint processing of available direction finding measurements.

Note 1. If during the solution of the identification problem mixing up of bearings is admitted, a false bearing can be considered as a bearing containing an AME. Developed CVM allows to struggle effectively with such errors irrespective of their nature.

Note 2. It is possible to attract for solving this problem various radio technical parameters (for example, carrier frequency, pulse repetition period, their duration, type of intrapulse modulation, etc.), which are "loaded" with bearings. Such parameters are stored in appropriate formulas, which are widely used in radio reconnaissance to solve the problem of identification of targets.

Version 2. In this case the CVM is applied to all the measurement sets $\mathbf{Y}_\alpha = \left[\tilde{\alpha}_{mj}, m = \overline{1,M}, j = \overline{1,J}\right]$ and $\mathbf{Y}_\beta = \left[\tilde{\beta}_{mj}, m = \overline{1,M}, j = \overline{1,J}\right]$ at once with formation of the clusters K_q, $q = \overline{1,Q}$. Initially, it is necessary to determine the numbers $q_j^* \in \{1,\dots,Q\}$ of the priority clusters $\text{K}_{q_j^*}$, and this is achieved by introducing the criterion $W(q_j^*) \geq \gamma$, where γ is the set target recognition threshold, $\gamma > 0$. Then for each priority cluster the resulting estimates $\boldsymbol{\lambda}_j^* = \boldsymbol{\lambda}_j^*(q_j^*) = \left[x(q_j^*), y(q_j^*), z(q_j^*)\right]^{\text{T}}$, $j = \overline{1,J}$ are constructed.

Version 2 is quite computationally expensive, since it leads to a large number of clusters (Q). The advantage of the version is that the problem of triangulation estimation of coordinates of many targets fulfills the potential capabilities of CVM in the conditions of uncertainty.

7. COMPARATIVE ANALYSIS

Let us consider a system whose direction finders are on a circle with the coordinates $\boldsymbol{\xi}_m = [x_m, y_m, z_m]^{\mathrm{T}} = \left[10^4 \cos(2\pi m/M), 10^4 \sin(2\pi m/M), 0\right]$, $M = 5$, $m = \overline{1,4}$. For the vector $\boldsymbol{\lambda}$, 180 marks were registered, also arranged in the circle:

$$\boldsymbol{\lambda}_k = [x_k, y_k, z_k]^{\mathrm{T}} = \left[5 \times 10^4 \cos(2\pi k/K),\ 5 \times 10^4 \sin(2\pi k/K),\ 3 \times 10^3\right]^{\mathrm{T}}, \quad k = \overline{1,180}.$$

Hereinafter, the coordinates of target and direction finders are given in meters, azimuth, elevation angle, and direction finding errors are given in radians. For each fixed k, it was assumed that in reliable measurement channels the fluctuation errors of measurements are distributed according to the normal law with zero mathematical expectations and correlation matrix $\mathbf{K}_m = \operatorname{diag}\left[\sigma^2_{\alpha m}, \sigma^2_{\beta m}\right] = \operatorname{diag}\left[\sigma^2_{\alpha}, \sigma^2_{\beta}\right]$, where $\sigma_\alpha = \pi/360$ and $\sigma_\beta = \pi/360$, and, in this case, the errors were formed using a random number sensor (errors of different direction finders are independent). The numbers of unreliable measurements (no more than half of all azimuthal measurements and no more than half of all angular measurements) containing AME were chosen randomly. The resulting measurement error (this is the sum of AME and acceptable random error) corresponded to the interval $(3\sigma_\alpha, \pi/6)$ for the azimuth and $(3\sigma_\beta, \pi/6)$ for the location angle. The estimation procedure was performed for each fixed k followed by averaging over a hundred experiments. In A_1 algorithm $Q = 7$, and in A_2 and A_3 algorithms $Q_0 = 2$ was taken as the initial condition. Two numerical characteristics were used for the compared algorithms A_1, A_2 and A_3: $S(\mathrm{A}_i)$ is the integral characteristic of accuracy (expressed in meters), $T(\mathrm{A}_i)$ is the characteristic of computational efficiency (expressed in seconds). For the kth target position and algorithm A_i ($i \in \{1,2,3\}$) we have

$$S(\mathrm{A}_i) = \sum_{k=1}^{180} S_k(\mathrm{A}_i) = (2\pi/180) \sum_{k=1}^{180} \Delta_k(\mathrm{A}_i),$$

where $\Delta_k(\mathrm{A}_i) = \left\|\bar{\boldsymbol{\lambda}}^*_k(\mathrm{A}_i) - \boldsymbol{\lambda}_k\right\|_2$ is the partial residual, $\boldsymbol{\lambda}^*_{kp}(\mathrm{A}_i)$ and $\bar{\boldsymbol{\lambda}}^*_k(\mathrm{A}_i) = \sum\limits_{p=1}^{100} \boldsymbol{\lambda}^*_{kp}(\mathrm{A}_i)/100$ are unit (for the pth experiment) and averaged (for 100 experiments) estimates of vector $\boldsymbol{\lambda}_k$ respectively.

For a comparative analysis of algorithms A_i (where $i \in \{1,2,3\}$), the relative integral characteristic

$$\delta S(\mathrm{A}_i) = 100 S(\mathrm{A}_i)\, S^{-1}(\mathrm{A}_1)\ \ [\%]$$

and the relative computational efficiency

$$\delta T(\mathrm{A}_i) = 100 T(\mathrm{A}_i)\, T^{-1}(\mathrm{A}_2)\ \ [\%]$$

are further used, taking into account that algorithm A_1 is less accurate and A_2 is more time-consuming than other algorithms. The results of the simulation are presented in the algorithm comparison table.

Table

Algorithms (A_i)	A_1	A_2	A_3
$\delta S(\mathrm{A}_i)$, %	100	50	67
$\delta T(\mathrm{A}_i)$, %	90	100	15

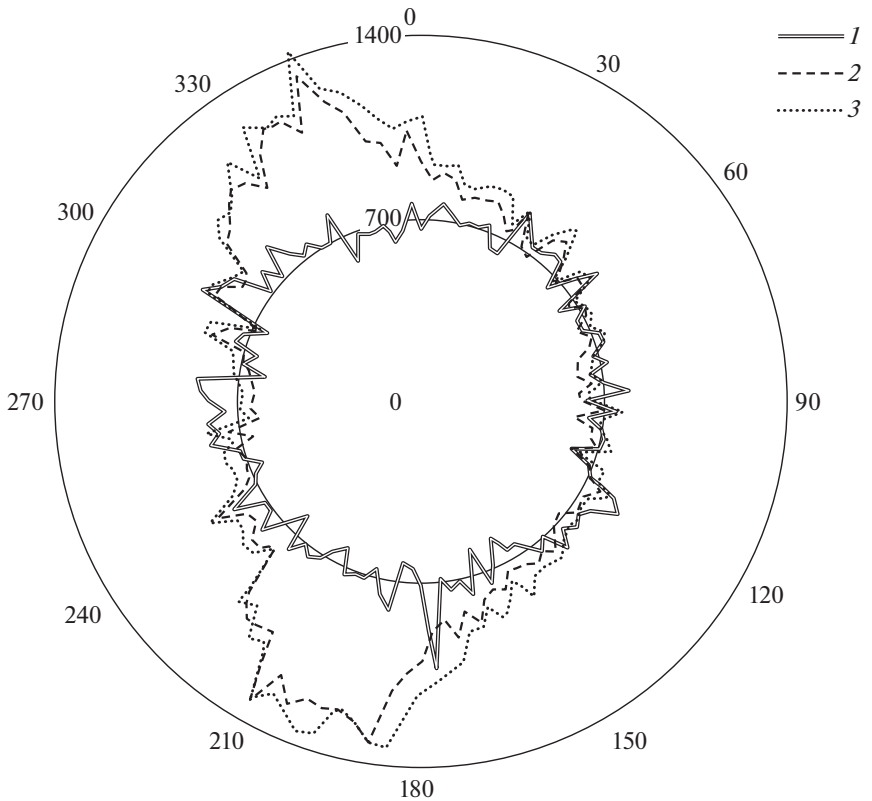

Fig. 6. Comparison with algorithms based on the maximum likelihood method.

We see that in the abnormal conditions of the system the modified algorithm A_2 with adaptive selection of the number of clusters (based on the "silhouette" factor), optimal selection of the resulting cluster (based on the improved criterion) and weighted data processing (using matching coefficients) is significantly better (in terms of accuracy) than algorithms A_1 and A_3. In terms of responsiveness, it is the most time-consuming, but it should be noted that the time cost of implementing procedures (6)–(8) does not exceed 10%. We also see that the A_3 algorithm is inferior to the A_2 algorithm in terms of accuracy, but it requires significantly less time.

We also carried out a comparison of the CVM with the maximum likelihood method (MLM) and the extended MLM (EMLM—for the case when the AMEs are included in the vector of estimated parameters). It was assumed that the measurements with AMEs are fixed—this is the 5th azimuth and the 3rd position angle. For the MLM method, the readings of the 5th and 3rd direction finders were completely excluded from the processing, while for the EMLM method, it was assumed that the numbers of measurements with anomalous errors are known, and only the values of the corresponding errors are unknown.

The resulting error of the compared methods (in meters) is presented in Fig. 6. Here we use the designations: *1*—for CVM, *2*—for MLM, *3*—for EMLM. From the figure, we can see that only the CVM provides reliable estimation for all directions of target sighting.

8. CONCLUSION

The proposed modified CVM allows, based on the A_2 and A_3 algorithms, to construct a resultant estimate of the target location stable to the influence of AMEs under conditions of structural uncertainty. These algorithms both independently and in combination with traditional approaches (e.g., the maximum likelihood method) can be effectively used to improve existing and develop

promising new generation systems. In case of limited computational resources and/or with a large number of direction finders, the two-step approach (based on the A_3 algorithm) can significantly increase the efficiency of triangulation estimation.

Obviously, that essential interest for the theory and practice of multi-position direction finding systems of various types are the following directions of CVM improvement: consideration of systems with single-channel and dual-channel direction finders of various types; modernization of obtained algorithms for the case of non-synchronous and different-precision measurements; building a filtering algorithm taking into account the chain character of clusters, which will be most expressed in consideration of a moving target. Work in these areas is already underway and will be presented to interested specialists in the near future.

REFERENCES

1. Bulychev, Yu.G. and Golovskoi, V.A., Processing of Measurements of Angle Measuring Systems under a Priori Uncertainty in the Regularized Formulation, *RE*, 2010, tom 55, no. 1, pp. 71–77.
2. Bulychev, Yu.G. and Chepel', Ye.N., Multistructure Method for Triangulation Estimation of Motion Parameters of a Radiating Target under a Priori Uncertainty, *Teor. Sist. Upravlen.*, 2019, no. 6, pp. 26–42.
3. Saibel', A.G., *Osnovy teorii tochnosti radiotekhnicheskikh metodov mestoopredeleniya* (Fundamentals of the Theory of Accuracy of Radio-Technical Methods of Locating), Moscow: Oborongiz, 1958.
4. Kukes, I.S. and Starik, M.E., *Osnovy radiopelengatsii* (Fundamentals of Radio Direction Finding), Moscow: Sovetskoe Radio, 1964.
5. *Teoreticheskie osnovy radiolokatsii* (Theoretical Foundations of Radiolocation), Shirman, Ya.D., Ed., Moscow: Sovetskoe Radio, 1970.
6. Butterly, P.I., Position Finding with Empirical Prior Knowledge, *IEEE Trans.*, 1972, vol. AES-8, no. 3, pp. 142–146.
7. Nunn, W.R., Position Finding with Prior Knowledge of Covariance Parameters, *IEEE Trans.*, 1979, vol. AES-15, no. 3, pp. 204–208.
8. Wax, M., Position Location from Sensors with Position Uncertainty, *IEEE Trans.*, 1983, vol. AES-19, no. 5, pp. 658–662.
9. Kondrat'ev, V.S., Kotov, A.F., and Markov, L.N., *Mnogopozitsionnye radiotekhnicheskie sistemy* (Multi-Position Radio Systems), Moscow: Radio i Svyaz', 1986.
10. Chernyak, V.S., *Mnogopozitsionnaya radiolokatsiya* (Multi-Position Radiolocation), Moscow: Radio i Svyaz', 1993.
11. Lin, X., Kirubarajan, T., Bar-Shalom, Y., and Maskell, S., Comparison of EKF, Pseudomeasurement and Particl Filters for a Bearing-only Target Tracking Problem, *Proc. SPIE-Int. Soc. Optic. placecountry-regionEng.*, 2002, vol. 4728, pp. 240–250.
12. Bulychev, Y.G., Bulychev, V.Yu., Ivakina, S.S., and Nasenkov, I.G., Passiv of Location of moving Targets with prior Information, *Autom. Remote Control*, 2017, vol. 78, no. 1, pp. 125–137.
13. Bulychev, Y.G., Bulychev, V.Yu., Ivakina, S.S., etc. Rationale for Methods of Optimal Estimation of Target Motion Parameters in Triangulation Measuring System, *Teor. Sist. Upravlen.*, 2015, no. 4, pp. 94–110.
14. Gustafsson, F., Particle Filters for Positioning, Navigation and Tracking, *IEEE Transactions on Signal Processing*, 2002, vol. 50, no. 2, pp. 425–437. https://doi.org/10.1109/78.978396
15. Bar-Shalom, Y., Rong, Li X., and Kirubarajan, T., *Estimation with Applications to Tracking and Navigation: Theory, Algorithms and Software*, New York: John Wiley & Sons, 2004. https://doi.org/10.1002/0471221279

16. Valente de Oliveira, J. and Pedrycz, W., *Advances in Fuzzy Clustering and Its Applications*, New York: John Wiley & Sons, 2007. https://doi.org/10.1002/9780470061190

17. Zekavat, S. and Buehrer, R., *Handbook of Position Location: Theory Practice and Advances*, 2nd ed. Hoboken, New Jersey: Wiley-IEEE Press, 2019. https://doi.org/10.1002/9781119434610

18. Zhao, J., Renzhou, G., and Xudong, D., A New Measurement Association Mapping Strategy for DOA Tracking, *Digital Signal Processing*, 2021, vol. 118, pp. 103–228. ISSN 1051-2004. https://doi.org/10.1016/j.dsp.2021.103228 (https://www.sciencedirect.com/science/article/pii/S1051200421002670)

19. Peng, L., Wenhui, W., Junda, Q., Congzhe, Y., and Zhenqiu, S., Robust Generalized Labeled Multi-Bernoulli Filter and Smoother for Multiple Target Tracking using Variational Bayesian, *KSII Transactions on Internet and Information Systems*, 2022, vol. 16, no. 3, pp. 908–928. https://doi.org/10.3837/tiis.2022.03.009

20. Wang, X., Wang, A., Wang, D., Xiong, Y., Liang, B., and Qi, Y., A Modified Sage-Husa Adaptive Kalman Filter for State Estimation of Electric Vehicle Servo Control System, *Energy Reports*, 2022, vol. 8, no. 5, pp. 20–27. ISSN 2352-4847. https://doi.org/10.1016/j.egyr.2022.02.105 (https://www.sciencedirect.com/science/article/pii/S2352484722003523)

21. Widrow, B. and Stearns, S., *Adaptive Signal Processing*, Moscow: Radio i Svyaz', 1989.

22. Granichinin, O.N. and Polyak, B.T., *Randomized Estimation and Optimization Algorithms under Almost Arbitrary Disturbances*, Moscow: Nauka, 2003.

23. Mansur, M.E. and Stepanov, O.A., Algorithms of Complex Processing in the Task of Correction of Navigation System Readings in the Presence of Nonlinear Measurements, *Izv. Tulskogo Gos. Univ., Technical Sciences*, 2016, no. 6, pp. 89–102.

24. Mandel', I.D., *Klasternyi Analiz* (Cluster Analysis), Moscow: Finansy i Statistika, 1988.

25. Williams, U.T. and Lans, D.N., *Methods of Hierarchical Classification*, Malyutov, M.B., Ed., Moscow: Nauka, 1986.

26. Lance, G.N. and Willams, W.T., A General Theory of Classification Sorting Strategies. 1. Hierarchical Systems, *Comput. J.*, 1967, vol. 9, no. 4, pp. 373–380.

27. Granichinin, O.N., Shlymov, D.S., Avros, R., and Volkovich, Z., A Randomized Algorithm for Estimating the Number of Clusters, *Autom. Remote Control*, 2011, vol. 72, no. 4, pp. 754–765.

28. Paklin, N.B. and Oreshkov, V.I., Cluster Silhouettes, in *Sistemnyi analiz v proektirovanii i upravlenii: Sb. tr. XX Mezhdunar. nauchno-prakt. konf.* (Systems Analysis in Design and Management: 20th International Scientific and Practical Conference), St. Petersburg, June 29–July 1, 2016, pp. 314–321.

This paper was recommended for publication by O.A. Stepanov, a member of the Editorial Board