FACTA UNIVERSITATIS (NIŠ)
Ser. Math. Inform. Vol. 30, No 5 (2015), 567–581

CLUSTERING METHODS FOR CLASSIFICATION OF ELECTRONIC DEVICES BY PRODUCTION BATCHES AND QUALITY CLASSES

Lev A. Kazakovtsev, Aljona A. Stupina, Victor I. Orlov, Margarita V. Karaseva and Igor S. Masich

Abstract. Authors consider the problem of electronic units packaging with highly reliable electronic components and preventing ingress of low-grade counterfeit products that does not meet the requirements for reliability. When making any electronic circuits, it is desirable to use electronic and radio components with the same characteristics which is most likely achieved using components (chips, transistors, diodes, capacitors, relays, crystals, resistors, etc.) manufactured as a single production batch. If the production method is not exactly known, only affordable way to improve the quality is the comprehensive testing of the delivered production batches. The paper discusses the problem of identifying a production batch of electronic and radio components delivered from a provider based on the test results. The problem is reduced to a series of problems of cluster analysis a special genetic algorithm is applied for. In addition, the testing problem of electronic and radio products is presented as pattern recognition without a teacher. A new algorithm for data classification in the multidimensional feature space is given. It was proposed to group objects on the basis of the distances analysis, i.e., the algorithm does not require knowledge about a number of classes in contrast to the majority of well-known algorithms for taxonomy.
Keywords: Clustering, production batches, quality classes, cluster analysis, genetic algorithm.

1. Introduction

The quality of an electronic component base applied while assembling various electronic nodes leaves much to be desired. The electronic component base manufactured by the placecountry-regionUSA differs with respect to quality classes (Commercial/Industry, Military, Space) [2, 3]. From class to class, the cost of the product increases. Regardless the class of products, as an ideal, the whole production batch is manufactured in the same conditions from the same batch of raw materials.

Received May 06, 2015.; Accepted August 03, 2015
2010 *Mathematics Subject Classification.* Primary 90B85; Secondary 90C27, 90B80

At the same time, suppliers of semiconductors and other devices do not guarantee the homogeneity of devices batch. It is especially true when they use products of foreign manufacture. In this case, the only available method of the elements quality improving is complex testing of the delivered batches for homogeneity check of production batches and division of component groups with identical characteristics from the combined batches (hypothetically).

The result of testing conducted by a manufacturer or a testing center is a set of parameters of each element. The result of each type of testing is a numeric characteristic (more often it is voltage or current in this or that circuit in this or that conditions). One should make a conclusion about homogeneity or heterogeneousness of production batch according to the manufacturing conditions on the basis of these characteristics analysis. The insignificant (at first glance) fluctuations of some characteristics at once let us make a conclusion that a part of the production batch is manufactured in different conditions though strict conditions are applied to the range of these characteristics.

Nowadays, the k-means problem is the most widespread model of cluster analysis [4]. In its turn, cluster analysis is an universal tool of classification and statistical data processing [5, 6]. The problem of k-means can be referred to the problems of continuous location theory [7]. In reality, this problem is brought to k-point finding (centers, centroids) in d-dimensional space of characteristics (here d is a number of characteristics measurements) such that a sum of distances from each of the data vectors to the nearest of the k centers have achieved the minimum. In the space of characteristics, as a rule, squared Euclidean metric is used as in this case finding the center of each cluster is an elementary task to be performed in one step and the computational complexity of the algorithm decreases in comparison to the algorithm with Euclidean metric when the calculation of the center (median) is an iterative process [8, 9]. The data vectors are sets of characteristics; in our case, they are the data of each product test results in the form of a numerical characteristics. The given problem is a problem of global search [10].

For these problems, we use ALA-algorithm (Alternating Location-Allocation) [10]. The given algorithm is a procedure of local search. It starts with specifying some initial solution, the initial set of centers when a subset of k points-vectors of data and means the alternate sets of data vectors finding for each center, for which this center is the closest one, and it also means a new center finding for each of these sets. In the case with the problem of k-means a median is selected as a new center, in the case with a p-median problem a new center is a solution of Weber problem that uses an iterative procedure of Weisfeld [8] or its modifications which are more advanced [9].

In the general case, a continuous problem of k-means can be formulated as follows:

$$\arg\min_{X_1,\ldots,X_k\in R^d} \sum_{i=1}^{N} \min \min_{j\in\{\overline{1,k}\}} L\left(X_j, A_i\right). \tag{1.1}$$

Here, $\{A_1,\ldots,A_N\}$ is a set of the data vectors in a d-dimensional space $X_1 \ldots X_N$ are

initial points (cluster centers), $L(\,)$ is some distance function (a metric).

In the case of Euclidean metric $L\left(X_j, A_i\right) = \sqrt{\sum_{k=1}^{d}\left(x_{j,k} - a_{i,k}\right)^2}$ we have a p-median problem. Here, $X_j = \left(x_{j,1}, ..., x_{j,k}\right) \forall j = \overline{1,p}, A_i = (a_{i,1}, ..., a_{i,k}) \forall i = \overline{1,N}$. In the case of squared Euclidean metric $L\left(X_j, A_i\right) = \sum_{k=1}^{d}\left(x_{j,k} - a_{i,k}\right)^2$, $w_i = 1 \forall i = \overline{1,N}$, we have a k-means problem.

Algorithm 1. ALA-algorithm.

Require: initial centers $X_1, ..., X_k \in \{A_1, ..., A_N\}$.

Step1: For each data vector $A_i \in A_1, ..., A_N$ find the nearest center $C_i = \arg\min_{j=\overline{1,k}} L\left(A_i, X_j\right)$. Form k sets (clusters) of data vectors for which each of k centers is the nearest one: $C_j^{clust} = \left\{i \in \left\{\overline{1,N}\right\} | C_i = j\right\}$;

Step2: For each cluster C_j^{clust}, $j = \overline{1,k}$ calculate its center X_j;

Step3: If at least one center has been changed at step 2, go to step 1;

Step4: Otherwise STOP.

In the case of k-means problem, such an algorithm is called a standard k-means procedure (p-means).

For the k-means problem, new cluster center finding is a simple problem:

$x_{j,k} = \sum_{i \in C_j^{clust}} a_{i,k} / \left|C_j^{clust}\right| \forall k = \overline{1,d}$. Here, $\left|C_j^{clust}\right|$ is cardinal number.

To reduce the computational cost at Step 2, the centers of only those clusters whose structures were changed at Step 1 are recalculated.

The result of the described algorithm depends on the choice of the initial center's values. The known k-means++ procedure [11] has the advantage over the random choice of initial centers ensuring the accuracy of the result $O(log(p))$. However, this accuracy may be unacceptable for many practically important problems. In this case various techniques of sets recombination of the initial centers are used.

The dependence of results of the ALA-algorithm as any procedure of local search on the set of initial values is a serious problem from the point of view of the results of the classification algorithm reproducibility: at different runs of the algorithm depending on the choice of initial values of the cluster centers, the same data vectors can belong to different clusters (in terms of the element base classification, it means to belong to various batches) or to the same cluster (to a single production batch). Thus, it is necessary to develop an algorithm for the cluster analysis which gives a stable result.

Excellent results can be obtained for the cluster analysis problems and the classification using methods Information Bottleneck Clustering method [12]. At the initial step of this algorithm, each data vector forms a separate cluster. Then "extra" clusters are removed from the system one by one until the desired number is reached. Each time, one cluster is removed; this removal gives the smallest increment of the total quadratic distance between the data vectors and the nearest cluster

centers. Such algorithms are extremely slow [12]. Some compromise demonstrates genetic algorithms with greedy heuristics [13] originally developed for solving p-median problem on a network. In the form proposed in [14], it can be applied for continuous problems solving. The idea of the approach can be summarized as follows [7].

Algorithm 2. Genetic algorithm with greedy heuristics for p-median problem.

We have: Population size N_p.

Step1: Form (randomly or using the procedure k-means++) N_p of different initial values $\chi_1, \ldots \chi_{N_p} \subset \{\overline{1,N}\}$, $|\chi_i| = k \forall i = \overline{1,N_p}$ – sets of indices of vector data of the power p, used as initial solutions for the ALA-algorithm. Estimate the value of the objective function for each of the initial solutions $F_{fitness}(\chi)$, calculated here and further with use of Algorithm 3, store the values of the given function to variables $f_1, \ldots, f_{N_k}$.

Step2: If stop conditions are achieved, then STOP. The solution is the initial solution χ_{i^*}, the least value f_i. corresponds to. For finding the final solution, the ALA-algorithm runs.

Step3: Choose randomly two indexes $k_1, k_2 \in \{\overline{1,N}\}$, $k_1 \neq k_2$.

Step4: Obtain the intermediate solution $\chi_c = \chi_{k_1} \bigcup \chi_{k_2}$.

Step5: If $|\chi_c| > k$, then go to Step 7:

Step6: Calculate $j^* = \arg \min_{j \in \chi_c} F_{fitness}(\chi_c \backslash \{j\})$. Eliminate j^* from χ_c: $\chi_c = \chi_c \backslash \{j^*\}$. Go to Step 5.

Step7: If $\exists i \in \{\overline{1,N_p}\} : \chi_i = \chi_c$, then go to Step 2.

Step8: Choose index $k_3 \in \{\overline{1,N_p}\}$. Two indexes $k_4, k_5 \in \{\overline{1,N}\}$ are chosen randomly, if $f_{k_4} > f_{k_5}$ then $k_3 = k_4$, otherwise $k_3 = k_5$.

Step9: Substitute χ_{k_3} and the corresponding value of the objective function: $\chi_{k_3} = \chi_c$, $f_{k_3} = F_{fitness}(\chi_c)$. Go to Step 2.

The determination of the objective function is realized with the algorithm:

Algorithm 3. Calculation of the objective function $F_{fitness}(\chi)$. We have: the initial solution χ.

Step1: Run Algorithm 1 with the initial set of centers $\{A_i | i \in \chi\}$, obtain a set of centers $\{X_1, \ldots, X_p\}$.

Step2: Retrieve value $F_{fitness}(\chi) = \sum_{i=1}^{N} w_i \min_{j \in \{\overline{1,p}\}} L\left(X_j, A_i\right)$.

So, step 4 of Algorithm 2 generates an intermediate solution, a set of power up to $2p$ where elements consistently are excluded (Step 6) to achieve the power p. Thus, at each iteration, a number of computations of the function $F_{fintess}$ corresponding to the current power of set of is required. We proposes to reduce a number of centers in the intermediate solution generated at Step 4. One of the ideas applied for the local search made ALA-algorithm improvement is to replace a part of the centers in the solution randomly selected data vectors [14]. In our case, we add a number

of elements to the initial solution of the another initial solution, and then exclude sequentially centers from the initial solution until p centers leave. A number of added centers is chosen randomly. Thus, Step 4 of Algorithm 2 takes the following form:

Step 4.1. Choose an integer $r \in \{\overline{1,p}\}$.with a generator of random numbers.

Step 4.2. Choose randomly a subset $\chi^*_{k_2}$ of the power r from the setχ_{k_2} .

Step 4.3. Obtain intermediate solution $\chi_c = \chi_{k_1} \cup \chi^*_{k_2}$.

We apply the version above to the real datasets.

2. Results of Method Based on k-Means Clustering

Various numbers of tests will be conducted depending on the class of electronic component and the destination node which element will be used in installation. For example, 55 measurements distinguished by the dimension and the measured physical value were taken for a circuit STK403-090.Splitting of the combined microcircuit batch into 2 and 3 clusters representing the estimated production batches

The developed clustering algorithms were applied to classify batches of electronic products according to production batches different in production conditions and, consequently, characteristics measured at conducting destructive tests. Moreover these algorithms were applied to determine a number of batches in the tested number of products.

The space of measurement results is normalized [15].

The MDS (Multi-Dimensional Scaling) method, visualization tool ELKI [16] and GNUPlot were applied for visualization of measurement results and the classification of algorithms tested electronic products.

We applied 10 combined batches of various components (diodes, voltage-reference diodes, field-controlled transistors, integrated circuits), each of them contained from one to seven production batches which differ in manufacturing conditions. The volume of combined batches is from 60 to 1250 units.

The algorithms of classification ran for each batch 10 times with different values of parameter k (estimated number of productionnbatches) from 1 to 10.

We divided a combined batch of the circuit STK403-090 into 1..9 batches The results for k=2 and k=3 are visualized in Fig.1.

Visually, we can point out 3 clusters for the circuit that corresponds to the 3 actually presented samples of the production batch (a total number is 827 units).

The result of clustering algorithms operation is the correspondence of numbers to the numbers of items in the production batch as well as a number of the estimated production batches and also the total range in the measurement results in the regulatory space that is an objective function of algorithms. The dependence of this total range from a number of the estimated production batches (number of clusters k) is given in Table 1. The table shows that for the product STK403-090,

Table 2.1: Dependence of the objective function on the number of the estimated production batches k of chip STK403-090

k	Batch 1	Batch 1	Batch 2	Batch 2
	$F_{fitness}(\chi)$	% of the previous value	$F_{fitness}(\chi)$	% of the previous value
1	84359	-	32984	-
2	53968	63,9%	29355	89%
3	37985	70,4%	27110	92,3%
4	27329	71,9%	25320	93,4%
5	21098	77,2%	23927	94,5%
6	19865	94,2%	22754	95,1%
7	18952	95,4%	21821	95,9%
8	18091	95,5%	21035	96,4%
9	17357	95,9%	20425	97,1%
10	17021	98,1%	20078	98,3%

the objective function value in the first batch (827 units) sharply decreases with increasing k up to values of k=4 (four production batches). For the second batch (1132units), this behavior of the objective function is not typical due to the fact that this batch includes items of only one production batch. All results are obtained in the time frame not exceeding 5 seconds. For comparison, data processing of the first production batch of 827 units by Information Bottleneck Clustering method needs more than 15 minutes on the machine with Intel Xeon processor with a clock frequency of 2 GHz with comparable by accuracy results. New algorithm gives stable results (minimum range of values) at high speed calculation.

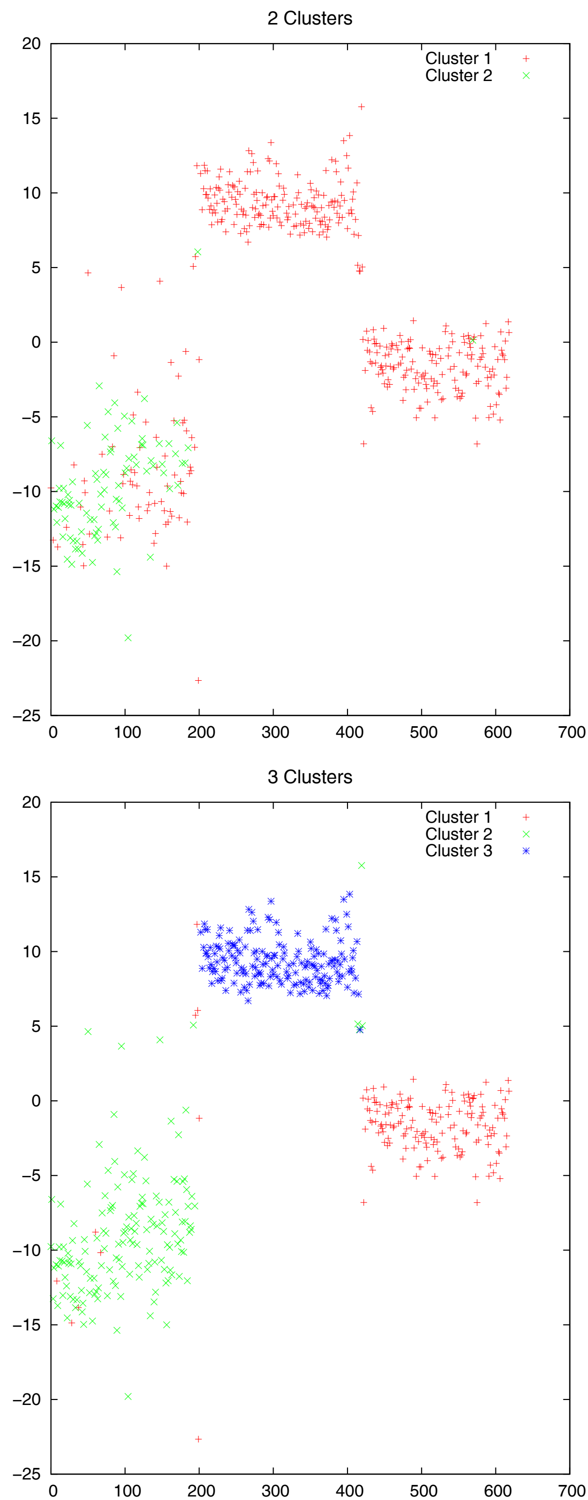

FIG. 2.1: Splitting of the combined microcircuit batch into 2 and 3 clusters representing the estimated production batches

3. Method Based on Non-Parametric Classification

While testing electronic and radio products (ERP), the main attention is paid on the types of tests and controls. The control means measuring electrical characteristics under different types of loads on separate elements and nodes of ERP.

Problem of pattern recognition is as follows [17]. Let some product be related to one or another quality category in accordance with the requirements to this type of products. This product is characterized by certain vector parameters values $v = (v_1, \ldots, v_m)$ which are used for diagnosis of the latter (the product may be referred to one of two classes, V_1 or V_2. The typical feature for technical diagnostics problems is the presence of a compact 1-connected domain in the space of characteristics that define a particular class (e.g. satisfactory, medium quality and high quality). The task of diagnosis is considered as a pattern recognition problem and it is brought to designing a solving rule based on the available training samples $\{v = (v_1, \ldots, v_m)\ ; U_i\}$ where U_i is teacher's instructions about the attachment to V_1 or V_2; $i = \overline{1, m}$, m is the sample size.

The number of classes may vary. However, regardless of their number they are always present in the field of mixing representatives of neighboring classes V_3 which leads to the necessity of solving the problem of pattern recognition in a probabilistic setting. This creates 3-alternative task of products diagnosis. The nonparametric pattern recognition algorithms for a 3-alternative case are proposed. The third class defines some border field among classes, i.e. subfield with products of different classes mixing.

The unsupervised pattern recognition (self-learning) is learning without any teacher' instructions about the correctness or invalidity of the system's response to certain shows [18]. Suppose that a set of images X consists of several disjoint subsets X_k, corresponding to various classes of images, characterized by vectors $k = \overline{1, l}$, $x \in X$. Since an image $x \in X$ appears in one or another set X_k where $k = \overline{1, l}$ is accidental, it is natural to consider the probability of emergency of the image x in the class X_k (let us denote it as P_k) and conditional probability density of the vector x inside the appropriate class X_k:$p_k(x) = P(x/k, k = \overline{1, l})$.

In this case, the maxima of the probability densities $P_k(x)$ are above the "centers" of classes corresponded to subsets X_k. When it is not known what class the image x belongs to these conditional probability densities cannot be determined. The joint probability density

$$P(x) = \sum_{k=1}^{l} P_k p_k(x)$$

contains rather complete information about the sets. In particular, it should be expected that the maxima (modes) of the function P(x) will correspond to the "centers" of classes. Therefore the self learning problem is often reduced to the task of restoring the joint probability density $P(x)$ and the "centers", determination and then and class boundaries [19]. So, for the case x=(x_1,x_2) The possible probability

density $P(x)$ is shown in Fig. 2. It shows the probability density $P(x)$ with two maximums, hence there are two classes. The centers of the classes are considered to be the coordinates of the maxima (modes) of the distribution $[x_1(k), x_2(k)]$, k=1,2.

Suppose that we have a set of some objects O_1, O_2,. . . ,O_s defined in the feature space. It is necessary to divide them into groups of objects close to each other in some sense. The information about these objects is defined in the form of a matrix (m is a number of input parameters or features, s is sample size of products). Thus, the classification problem is to divide the feature space into disjoint pieces. Then we consider the algorithm proposed for data grouping when the initial number of classes is unknown.

Algorithm for data grouping. Let we have an observation sample $\{x_i, i = \overline{1, s}\}$ of multidimensional variable $x \in R^m$ where m is the dimension of x. It is necessary to select samples of class observation from the available one, number of classes is unknown. Further we briefly describe a variant of the classification algorithm included the following basic steps.

Step1: To determine the distances between all points of the observations sample $\{x_i, i = \overline{1, s}\}$. The distance between two multidimensional points x_i and x_j is calculated as

$$r\left(x_i, x_j\right) = \sum_{q=1}^{m} \left|x_i^q - x_j^q\right|.$$

Step2: To find points located at a relatively great distance from each other (as a radius defining the "long" distance taken from the experimental analysis of the matrix of distances and depended on the average among between all points).

Step3: To choose an array of n elements from the available table of distances: g=$\{r[i], i = \overline{0, n}\}$, so that it contained a pair of points; a distance between them is the largest one, i.e., presumably they are elements of different classes.

Step4: To determine a number of classes and mass centers (x^*) of the selected classes based on the analysis of array g. To divide the array in accordance with the values analysis of its elements distances.

Step5: Find all points satisfying the condition $r(x^*, x) < \delta_1$ where δ_1 is estimated on the base of average distances analysis in the class (it is assumed that these points are of the same class as the initial point x^*).

Step6: To declare the selected points as a class and to exclude them from an observations sample.

Step7: To repeat all steps from Step 3 till the volume of the sample s is not equal to 0.

Let us consider grouping results of the real data conducted using the described algorithm.

Non-parametric data processing algorithms. Let x=(x1,x2,. . . ,xm) be a m-dimensional vector of parameters defining the quality of the certain product. There is a sample $\{x_i\}$, $i = \overline{1, s}$. In this case, non-parametric statistics can be applied as the

density estimation

$$p_s(x) = \frac{1}{sC_s} \sum_{i=1}^{s} \sum_{j=1}^{m} \Phi\left(\frac{x^j - x_i^j}{C_s^j}\right).$$

where $\Phi(\cdot)$ is an integral function with square bell-shaped $\Phi(x - x_i / C_s)$ and a parameter C_s (a fuzziness coefficient) such that is satisfies the convergence conditions [19].

To design a model of the dependencies of the variables, we used nonparametric estimation of Rosenblatt-Parsen regression functions [19]:

$$y_s(x) = \frac{\sum_{i=1}^{s} y_i \prod_{j=1}^{m} \Phi\left(\frac{x^j - x_i^j}{C_s^j}\right)}{\sum_{i=1}^{s} \prod_{j=1}^{m} \Phi\left(\frac{x^j - x_i^j}{C_s^j}\right)}$$

where $x=(x^1, x^2, \ldots, x^m)$ is an m-dimensional vector of input impact of the object; y_i is value taken in this case as the output object.

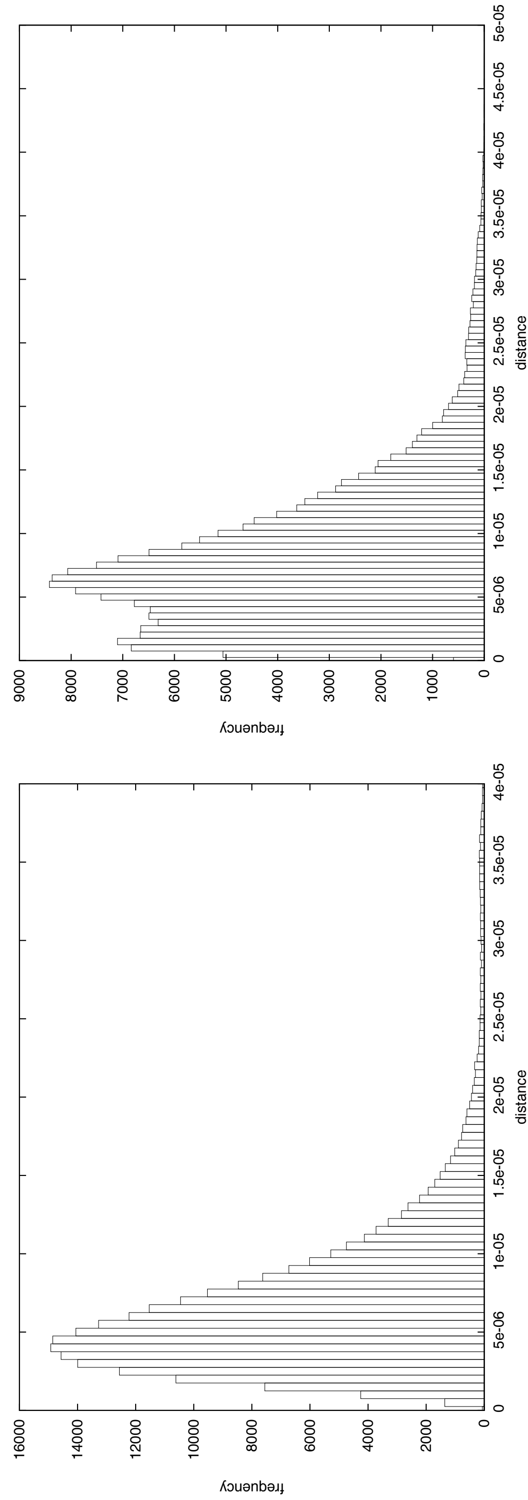

FIG. 3.1: Distance probability density in case of non-homogeneous (above) and homogeneous (below) production batches

4. Processing Test Data for the Non-Papametric Approach

Let us consider the real data classification. We have a sample consisting of circuits of the same production batch. The initial data is represented as a matrix of observations; each of them contains 83 characteristics (m = 83), the sample size of observations s = 198. It is necessary to divide a sample into classes. The experiment is conducted using the described above algorithm with different values of the distance δ. The results are presented below (R_s is the mean distance between two adjacent points).

A preliminary analysis of the data processing characterizing the ERP state leads to the conclusion that there exist quite close linear and nonlinear relations of some variables as in the nonparametric estimation of regression function from observations of some variables shows. However, the density estimation differs from the Gaussian distribution and is of multimodal nature; therefore, there is the necessity to group data that leads to the emergence of clusters in the space of parameters characterizing the product quality. In the class of accepted products in the result of diagnostic test there are two or more groups that can be called "good" and "very good".

Table 4.1: Classification results for the non-parametric approach

Distance δ	Number of classes	Class and number of devices
R_s	3	1:44; 2:73; 3:58; 4:23
0,5 Rs	5	1:33; 2:33; 3:50; 4:64; 5:7
3 R_s	2	1:81; 2:117

5. Conclusion

The problem of classification of the lots of electronic and radio components into production batches with different production conditions on the basis of non-destructive tests can be transformed into the problem of cluster analysis. The application proposed genetic algorithm with a special heuristic allows us solve such problems with a stable result. The algorithm of non-parametric data classification in a multidimensional feature space that does not require a number specification of classes was proposed. However, this method is subject to further research.

REFERENCES

1. D. Allen: *Relations between the local and global structure of fnite semigroups.* Ph. D. Thesis, University of California, Berkeley, 1968.
2. L. Hamiter: *The History of Space Quality EEE Parts in the country-regionplaceUnited States.* In: ESA Electronic Components Conference, ESTEC, ESA SP-313, 12-16 Nov 1990, Noordwijk, The country-regionplaceNetherlands, 1991.
3. C. S. Kirkconnell, T. T. Luong, L. S. Shaw et al.: *High Effi ciency Digital Cooler Electronics for Aerospace Applications.* In: Proc. SPIE 9070, Infrared Technology and Applications XL, 90702Q, 2014, doi:10.1117/12.2053075
4. P.-N. Tan, M. Steinbach and V. Kumar: *Cluster Analysis: Basic Concepts and Algorithms*, Chapter 8, Introduction to Data Mining, Addison-Wesley, 2006, pp.487–567.
5. J. B. MacQueen: *Some Methods of Classification and Analysis of Multivariate Observations.* In: Proceedings of the 5th Berkley Symposium on Mathematical Statistics and Probability, Vol. 1, 1967, pp.281-297.
6. I. S. Masich and E. M. Kraeva: *Oybor zakonomernostey dlya postroyeniya reshayuschego pravila v logicheskih algoritmah raspoznavaniya (Selection of the regularities for logical rule creation in logical recognition algorithms).* Systemi upravleniya i informatsioniye tehnologii **51 no.1.1** (2013), pp. 170-173 *(in Russian).*
7. L. A. Kazakovtsev, V. I. Orlov, A. A. Stupina and V. L. Kazakovtsev: *Modied Genetic Algorithm with Greedy Heuristic for Continuous and Discrete p-Median Problems.* Facta Universitatis (Niš) series Mathematics and Informatics, **30(1)** (2015), pp. 89–106.
8. E. Weiszfeld: *Sur le point sur lequel la somme des distances de n points donnes est minimum.* Tohoku Mathematical Journal, **43(1)** (1937), pp.335–386.
9. Z. Drezner: *The Fortified Weiszfeld Algorithm for Solving the Weber Problem.* IMA Journal of Management Mathematics, published online (2013). DOI: 10.1093/imaman/dpt019
10. L. Cooper: *Location-allocation problem.* Oper. Res. **11** (1963), pp.331–343.
11. N. Mishra, D. Oblinger and L. Pitt: *Sublinear time approximate clustering.* In: 12th SODA, 2011, pp.439–447.
12. Zh. Sun, G. Fox, W. Gu and Zh. Li: *A parallel clustering method combined information bottleneck theory and centroid-based clustering.* The Journal of Supercomputing **69(1)** (2014), pp. 452–467, doi: 10.1007/s11227-014-1174-1

13. O. Alp, E. Erkut and Z. Drezner: *An Efficient Genetic Algorithm for the p-Median Problem*. Annals of Operations Research **122(1-4)** (2003), pp.21–42.
14. M. N. Neema, K. M. Maniruzzaman and A. Ohgai: *New Genetic Algorithms Based Approaches to Continuous p-Median Problem*. Netw. Spat. Econ **11** (2011), pp. 83–99, DOI:10.1007/s11067-008-9084-5.
15. F. M. Callier: *Linear System Theory*, New York: Springer-Verlag. ISBN 0-387-97573-X, 1991.
16. H. P. Kriegel, P. Krueger and A. Zimek: *Outlier Detection Techniques (Tutorial)*. In: 13th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2009), Bangkok, Thailand, 2009.
17. N. V. Koplyarova, V. I. Orlov, N. A. Sergeva and V. V. Fedosov: *O neparametricheskih modelyah v zadachae diagnostiki electroizdeliy (On the non-parametric models of problem of electronic components diagnostics)*. Zavodskaya laboratoriya. diagnostica materialov **7** (2014), pp. 73–77 *(in Russian)*.
18. Ya. Z. Tsipkin: *Osnovi teorii obuchayuschikhsya system (Basics of self-learning systems theory)*, Moscow, Nauka, 1970 *(in Russian)*.
19. A. V. Medvedev: *Teoriya neparametricheskikh sistem. Protsessy (Non-parametric systems theory. Processes)*. Vestnik SibGAU **3** (2010), pp. 4–9 *(in Russian)*.

Lev A. Kazakovtsev
Siberian State Aerospace University Named after M.F.Reshetnev
Department of Economical Information Systems
prosp. Krasnoyarskiy Rabochiy 31
660014 Krasnoyarsk, Russian Federation
Siberian Federal University
Department of Economy and Information Technologies of Management
prosp. Svobodniy 79
660041 Krasnoyarsk, Russian Federation
levk@bk.ru

Aljona A. Stupina
Siberian State Aerospace University Named after M.F.Reshetnev
Department of Systems Analysis and Operations Research
prosp. Krasnoyarskiy Rabochiy 31
660014 Krasnoyarsk, Russian Federation
Siberian Federal University
Department of Economy and Information Technologies of Management
prosp. Svobodniy 79
660041 Krasnoyarsk, Russian Federation
saa55@rambler.ru

Victor I. Orlov

TTC NPO PM
ul. Molodezhnaya, 20
662970 Zheleznogorsk, Krasnoyarskii Krai, Russian Federation
ttc@krasmail.ru

Margarita V. Karaseva
Siberian State Aerospace University Named after M.F.Reshetnev
Department of Systems Analysis and Operations Research
prosp. Krasnoyarskiy Rabochiy 31
660014 Krasnoyarsk, Russian Federation
karaseva-margarita@rambler.ru
Igor S. Masich
Siberian State Aerospace University Named after M.F.Reshetnev
Department of Systems Analysis and Operations Research
prosp. Krasnoyarskiy Rabochiy 31
660014 Krasnoyarsk, Russian Federation
i-masich@ya.ru