

Structural-parametric optimization of the experimental data decomposition into approximated fragments

Rudolf Neydorf^{1,*}, Orkhan Iarakhmedov¹

¹Don State Technical University, Rostov-on-Don, Russia.

Abstract. In the process of modeling non-linear objects are often used piecewise approximation, spline methods or other methods based on their experimental data fragmentation. To effectively use fragmentation, it is necessary to develop and examine all the described stages of the approach. In this paper, a new approach to the selection of the structure of the decomposition of matrix data has been developed and proposed. It is based on the method of the ant colony, specially adapted for solving similar problems. Along with the characteristic properties of ant algorithms, special elements of interaction between agents of different colonies were introduced into this modification, for the efficient operation of the algorithm. The approach is considered as a tool for implementing the initial stage of any methods using data fragmentation. In the process of work, a software tool was designed and implemented in C#. The main data structures used in the software are described. Adjustment of the tuning parameters and the graphic output interface of the work process are implemented in the software tool. The graphical output was implemented by built-in tools. Test cases in detail describe the stages of data fragmentation under different initial parameters.

1 Introduction

In the process of modeling various objects and phenomena, many difficulties arise. They often occur in connection with a significant non-linearity of the test object [1-2]. In this case, approximative approaches are used, oriented to fragmentation of an array of experimental data (ED), for constructing a mathematical model (MM), because the fragments can be described by exactly considerably less complex MM. The most popular in solving these problems is the use of piecewise approximation [3], spline approximation [4-5], and methods of radial basis functions [6]. A comparatively recently developed of "Cut-Glue" approximation method [7-9], which has proved to be a very effective and promising method of mathematical description of nonlinear dependencies, is oriented towards fragmentation too. An obvious advantage of the "Cut-Glue" method over the others is the description of data by a single function, which is also analytical.

The initial stage of any of the listed methods is the splitting of the ED into fragments with the least pronounced nonlinearity. This dramatically increases the possibility of an accurate description of each ED. Therefore, the task of efficient fragmentation of data is extremely important, because is an important preparatory stage to any method focused on fragmentary approximation.

In connection with the fact that combinatorial methods for the formation and comparison of variants of ED fragmentation clearly refer to NP-completeness, in this paper, one of the possible approaches based on the biological phenomenon of rivalry of ant colonies was

proposed and investigated [12-13], since preliminary studies have shown promising results.

1.1 Formulation of the problem

It is necessary to develop an search suboptimization algorithm of dividing the area of ED into fragments with a phased use of the scenario of seizure of territories by developing ant colonies.

2 Algorithmic model

As a domain of disputed territories, an ED matrix is considered. To divide it into fragments, the rival ant colonies method (RACM) is used, which simulates the behavior of ant colonies in the division of their habitats [10-16], and adapted for the task. It is based on the mechanisms of interaction of ants with each other, with the environment, as well as colonies as societies among themselves.

In the process of partitioning, we can distinguish the following stages: initialization, agent migration, pheromone laying and checking the break condition, which are also in the classical ant algorithm.

All investigated space is traditionally divided by a grid, but not conditional, but rigidly tied to segments of step-by-step variation of factors when obtaining ED [14-16]. The number of colonies is varied by the "Cut-Glue" approximation algorithm, because it affects the number of fragments obtained after the partition. At their casual arrangement there are a number of restrictions. The colony can not be located on the boundaries of ED. This condition is established for a more even distribution of

* Corresponding author: ran_pro@mail.ru

territory between colonies. Also, the standard normal distribution is used in a random arrangement. Each colony generates a certain initial number of ants in the range that the operator specifies. This is done to ensure that the colonies differ in numbers, and therefore in capture capabilities.

There are several features of territory development by ants, as well as their interactions inside and outside the colony. Since the splitting as a result must consist of a set of rectangular-shaped fragments, a rule for preserving the rectangular shape of the territory under study is provided.

Moving virtual ants is carried out on a probabilistic basis [14-16]. At each iteration, each agent considers 8 possible directions to move as shown in Figure 1. Since the ants can expand the territory of the colony either by mastering unoccupied territories, or by capturing strangers, the corresponding stages are envisaged.

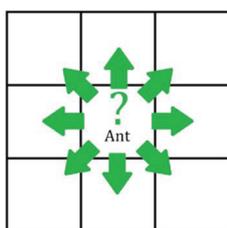


Fig. 1. Available options for moving

Expanding one of the four boundaries of the colony, the ants test it for the presence of someone else's pheromone. If the new border area is completely free, it is endowed with the pheromone of the colony, and the extension continues in a similar way. In the event that an someone else's pheromone is found on the territory under development, the ants return to their assigned territory and are looking for new ways of expansion.

The pheromone is used to highlight areas of the colonies. Its updating occurs at each moving of ants. Finished fragments after the splitting pass through the optimization and gluing stages.

In conditions when each colony has 4 boundaries that form a rectangle or a square, there can be a free space between the colonies, which is not occupied by any of the colonies (wasteland). To solve such situations, the algorithm includes the mechanism of interspecific competition of colonies bordering on the wasteland. The essence of this mechanism is that the disputed territory begins to retreat to that colony that occupies a smaller area. The introduced priority is due to the fact that when it is expanded the changes in the fragmentation structure will be minimal. Too small fragments during the partitioning are often inefficient, both from the position of the quality of the approximation, and from the position of the number of fragments. If the stage of "reclaiming the wasteland" leads to the formation of a new wasteland, the procedure for assessing neighboring fragments is repeated and a new applicant is appointed to capture it.

The break condition at the stage of partitioning is the satisfaction of the integrity condition and rectangularity of all the fragments obtained, as well as the absence of wasteland in the fragmented space.

Since fragments should have adjacent boundaries, it is necessary to work out a rule in which these faces will be determined. In this case, each fragment will include one more vector of values to the right and the bottom of the fragment, unless from these sides it is not limited by the value of the matrix.

Therefore, the result of a single execution of RACM is the variant of data fragmentation. Multiple launch of RACM allows you to generate a sample of options to search for the optimal or suboptimal by the total accuracy of the approximation data under study. These variants go through further stages of approximation, "gluing" and estimating the error of describing the entire dependence in accordance with the "Cut-Glue" approximation paradigm.

2 Conceptual-mathematical model

In connection with the task, it is necessary to build a RACM MM, radically different from the classical route model [10-11]. This is due to the fact that ant colony algorithm was first applied to the fragmentation problem solution. Nevertheless, some principles for the movement of agents remain unchanged; are the basis of the very principle of the "sequential-kinematic" solution of the problem.

Consider ED as a set of D numbers that are the values of the responses (output variables) of the described object and are associated with the points of the r -dimensional factor space (FP) of the input (independent) variables. In this paper, we consider a two-dimensional version of a phase transition.

For any fragmentation methods and algorithms (FA), the following two basic conditions must always be fulfilled, ensuring the correctness of the procedure.

1) The combination of all the obtained experimental data fragments (EDF) forms the entire investigated set of ED:

$$D_1 \cup D_2 \cup \dots \cup D_n = D \quad (1)$$

where $D_i, i = \overline{1, n}$ – fragments of a partition; n – number of fragments; D - set of ED.

2) The intersection of all the obtained EDF forms a subset of ED, which is the union of the boundaries of all fragments:

$$D_1 \cap D_2 \cap \dots \cap D_n = G_1 \cup G_2 \cup \dots \cup G_n \quad (2)$$

where $G_i, i = \overline{1, n}$ – the edges of fragments, which in the general are the partition.

The first condition is obvious. The second condition excludes the overlapping of fragments in which the boundary data of some fragments coincide with the internal data of other fragments. When the EDF merges ("gluing together") with the fulfillment of these conditions, the distortions in the transition between them are the least significant [7-9].

As already noted, the early agents master the space of experimental data presented in the form of a matrix, while moving, they are guided by a certain behavior and direction, which has both random and meaningful components.

Meaningful components are determined by the overall goal of FA and depend on many factors: the state of the boundaries of their colony, the state of the nearest border areas and the territory's fullness with a pheromone and etc.

Random components of displacement are used to expand the range of search for various solutions of the problem, greater uniformity of development of the territory in all directions, and so on.

The behavior of agents depends on many factors: the fulfillment of the rectangle condition, the availability of adjacent faces for expansion, the occupancy of adjacent faces in the expansion of the territory, the state of the integrity of boundaries, and others.

Under different statuses of these conditions, it is possible to distinguish certain features of the colony's behavior.

For example, with a "peaceful" expansion of borders, agents randomly choose the direction and check the whole new face for the possibility of moving and filling it with their pheromone. If the fragments do not contain a foreign pheromone and there are no other restrictions, the territory passes under the possession of the colony and is marked with a certain pheromone (color), as shown in Figure 2 (a) and 2 (b).

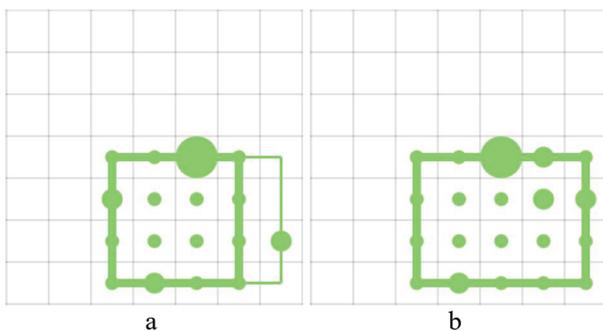


Fig. 2. Process of capturing an unoccupied territory

In a conflict ("aggressive") expansion, agents intentionally shift the boundaries of their colony and mark the entire territory with their pheromone, including those fragments that belonged to another colony. An example of this process is shown in Figure 3 (a) and 3 (b).

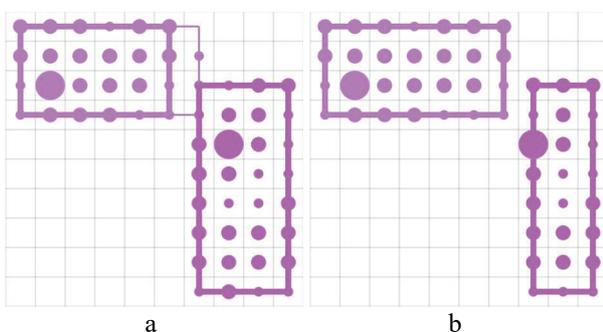


Fig. 3. Process of capturing someone else's territory

Such behavior is caused by the appearance of all kinds of collisions between the basic principles that are mandatory for fragmentation (for example, conditions (1) and (2), the requirement of rectangular fragments, etc.),

and the real configurations, for example, the appearance of the previously described "wastelands". This mechanism is illustrated in detail in the test case later in this work.

In the process of struggle between the colonies, their structure changes because of "wastelands," and the basic rule of rectangularity is violated. In these conditions, agents need to restore their maximum possible territory by retaining the rectangle rule. To do this, all agents return to the colony, and re-expand the boundaries of the territory already marked with their pheromones, as far as possible (see Figure 4 (a) and 4 (b)). At the same time, the pheromone on the side that is partially or completely captured, evaporates, thus a new "wasteland" and the opportunity to change the structure of fragmentation appear.

There are also other collisions, the analysis of which is beyond the scope of this article.

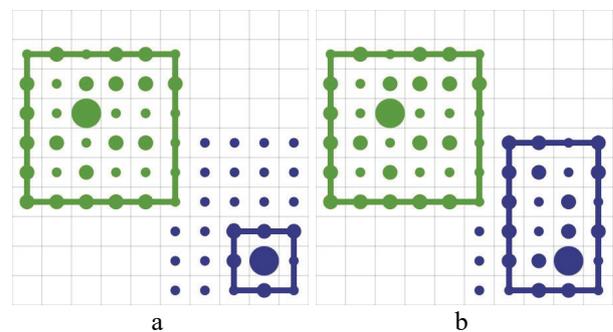


Fig. 4. The process of restoring borders

Based on the developed algorithm, its mathematical and conceptual models (1-2), the "Experimental Data Fragmentation by Rival Ant Colonies Method" (EDF_RACM) software tool (ST) was designed and developed.

3 Demo software tool "EDF_RACM"

3.1 Data structures

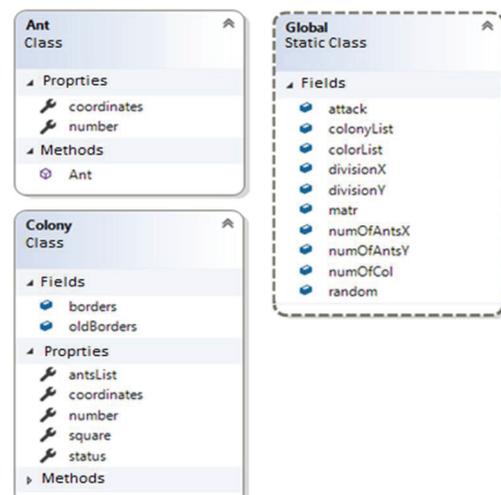


Fig. 5. UML-diagram of EDF_RACM

Figure 5 shows the UML-diagram of the ST classes.

Figure 5 shows that this ST consists of 3 main classes:

- class Ant – class describing an ant as an object;
- class Colony – class describing the colony as an object;
- class Global – class used to store global values and methods.

3.2 GUI

Figure 6 shows the main window of the software tool.

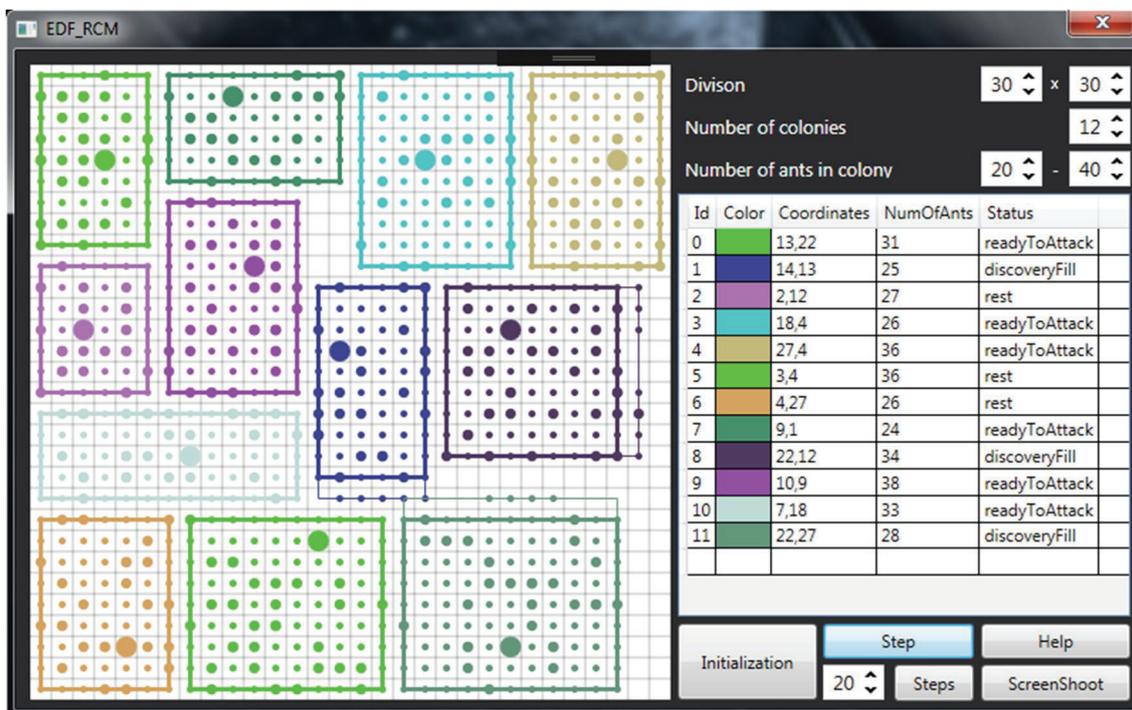


Fig. 6. GUI

The main window of the ST consists of a graphical output window to the right and a panel for setting parameters and outputting information to the left. Thanks to the graphical interface, you can visually follow all the steps of fragmenting experimental data, and using the settings panel you can select a variety of source tasks. Colonies are marked with different colors, and their boundaries are indicated by bold lines.

In the settings panel, the operator can specify the partition, the number of colonies and the range of generated ants. There is also an information area in which the current status of the colonies is reflected. At the

bottom of the tuning panel there are process control buttons.

3.3 Experiments

As a demo, the space of 30x30 dimension was considered. It was generated 8 colonies in a random order. Figure 7 (a) shows the initialization stage, which shows the location of the 8 colonies. Figures 7 (b) and 7 (c) illustrate the intermediate stage and the resulting result in the form of a partition.

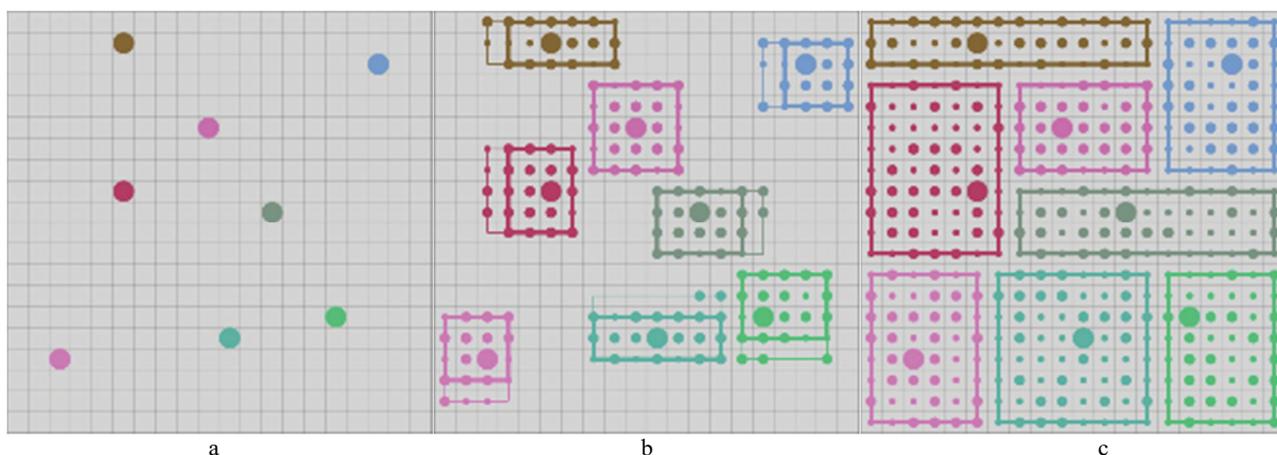


Fig. 7. Stages of the program: a – the initialization stage; b – intermediate stage; c – result

Experiments with matrices of various dimensions were carried out. The results of partitioning the matrices

of dimension 10x10, 30x30 and 40x40 are shown in Figure 8 (a), 8 (b) and 8 (c).

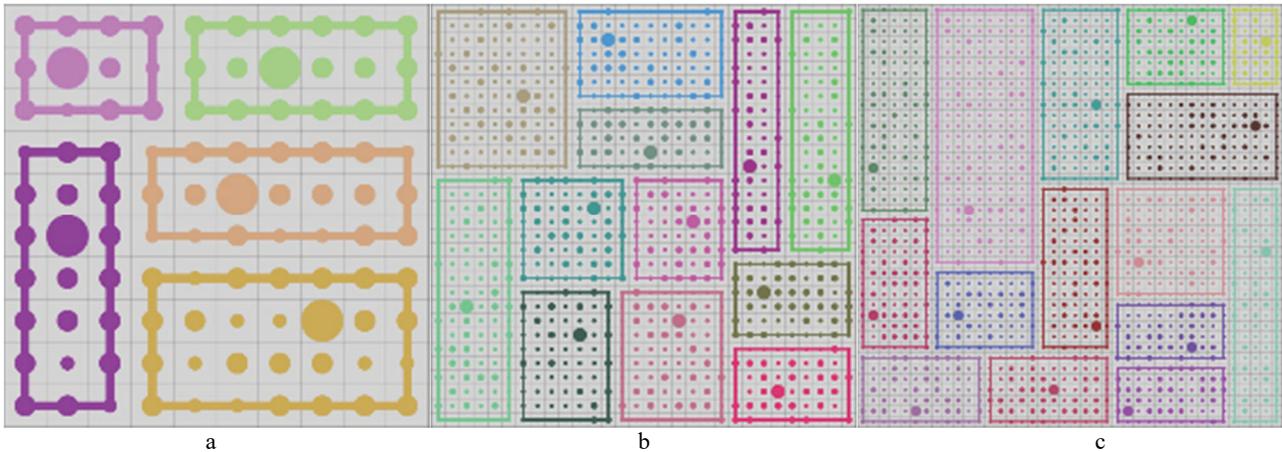


Fig. 8. Examples of partitions: a – partition 10x10; b – partition 30x30; c – partition 40x40.

3.4 Additional experiments

To demonstrate the mechanism of getting rid of "wastelands", an additional experiment was conducted. 6 colonies were generated in a random order. Figure 9 (a) shows the stage in which it is seen that all free space has

been investigated, with the exception of one controversial fragment, and a colony with a minimal territory begins to capture this fragment, advancing its right border. This change entailed a change in the boundaries of another colony and the liberation of a new territory for development (see Figure 9 (b)). The following stages are successively shown in Figure 9 (c), 9 (d), 9 (e) and 9 (f).

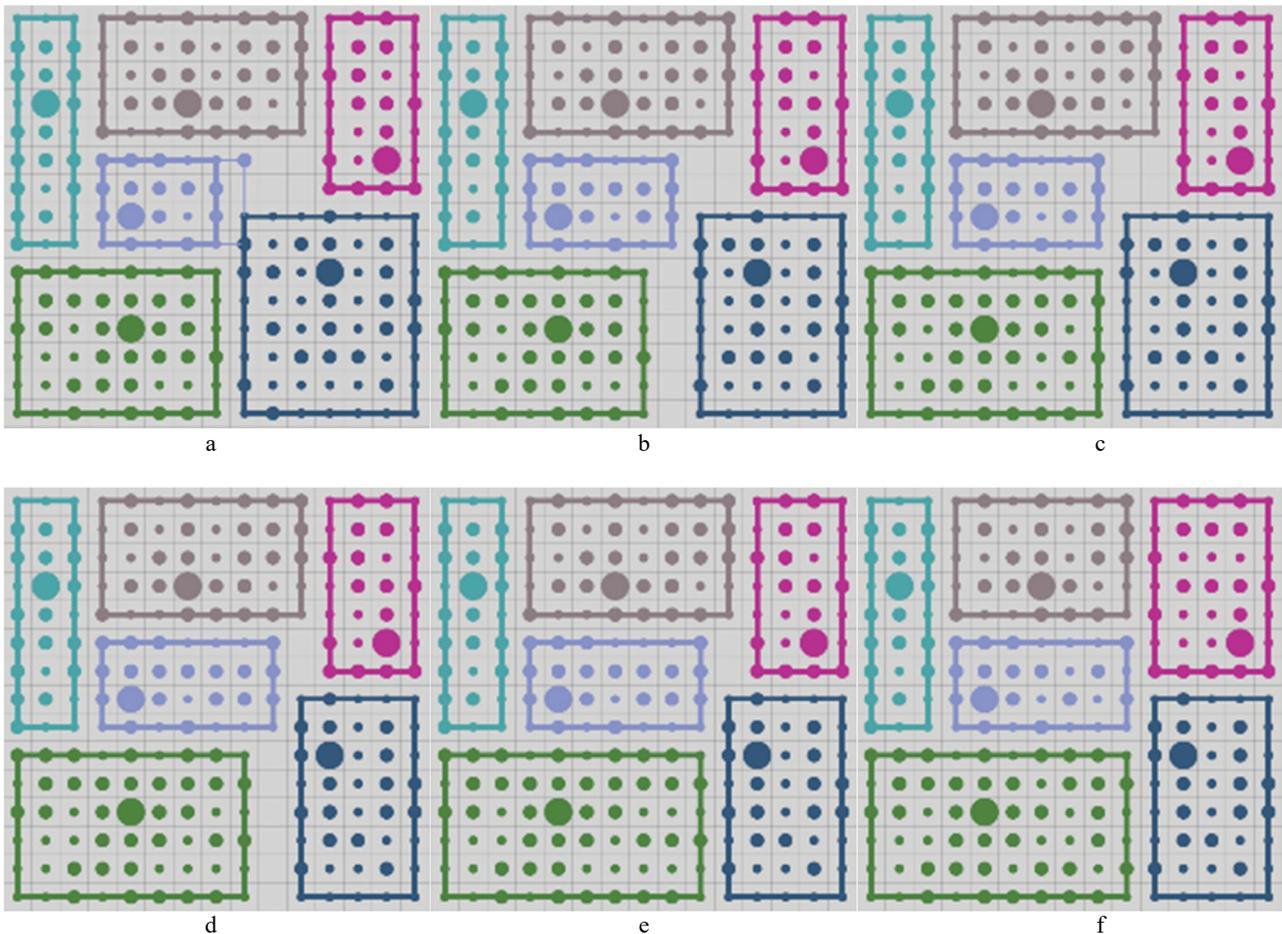


Fig. 9. Stages of the program in the occurrence of "wasteland"

* Corresponding author: ran_pro@mail.ru

3.5 Computational resources and performance

This ST was developed in the Microsoft Visual Studio 2017 environment using the C # programming language. The solution of the tasks was carried out on the PC with the AMD Phenom II P960 process and 6Gb of RAM.

4 Conclusion

The developed modification of the ant algorithm in the RACM variant showed acceptable qualitative and quantitative results in solving problems of varying complexity. The check was carried out on tasks with a partition from 10 to 50. In all tests, the algorithm did not allow errors during fragmentation. The results of its tests showed the expediency of continuing research in the chosen direction. It can be concluded that the developed ST meets all the necessary conditions and allows you to find the required partitioning.

It should be noted that the algorithm requires the involvement of some additional mechanisms to improve performance [14-16].

References

1. G. Boeing, Visual Analysis of Nonlinear Dynamical Systems: Chaos, Fractals, Self-Similarity and the Limits of Prediction. *Systems*. **4** (4): 37. (2016).
2. A.V. Panteleev *Nonlinear control systems: description, analysis and synthesis* (University book, Moscow, 2008)
3. P.-J. Laurent, *Approximation and optimization* (Hermann, Paris, 1972).
4. C. De Boor, *A practical guide to splines* (Springer, 1978)
5. Z. Cieselski, J. Musielak. Spline bases in function spaces. *Approximation theory*, Reidel, 49 (1975).
6. M.D. Buhmann. *Radial Basis Functions: Theory and Implementations*, (Cambridge University Press, 2003).
7. R. Neydorf, *Proceedings of the ASME 2014 International Mechanical Engineering Congress and Exposition*. (Montreal, Quebec, Canada, 2014).
8. R. Neydorf, *SAE Int. J. Aerosp.* **8**(1) 47-54 (2015)
9. R. Neydorf, A. Neydorf, *SAE Technical Paper*, 2016-01-2035 (2016).
10. M. Dorigo, L.M. Gambardella, *IEEE Trans. Evol.Comput.*, **1**(1) 53 (1997)
11. M. Dorigo, G. Di Caro, L. M, *Artificial Life*, **5**(2) 45 (1999)
12. K.P. Fatemeh, F. Fardad, S.N. Reza, *Journal of Academic and Applied Studies*. **3**(2), 15 (2013).
13. M.D. Toksari, *Applied Mathematics and Computation*. **176**, 6 (2006).
14. R. Neydorf, I. Chernogorov, V. Polyakh, O. Yarakhmedov, J. Goncharova, D. Vucinic, *Proceedings of The X International Conference on*

Advanced Engineering Computing and Applications in Sciences ADVCOMP2016, 7 (2016).

15. R.A. Neydorf, O.T. Yarakhmedov, V.V. Polyakh, I.V. Chernogorov, D. Vucinic *ACEX2016-10th International Conference on Advanced Computational Engineering and Experimenting, ACEX280*, (2016).
16. C.Y. Pang, X. Li, *5th IEEE Conf. on industrial Electronics and Applications*, 4 (2009).

* Corresponding author: ran_pro@mail.ru