Statistical modeling of a scale network of nanosatellites

To cite this article: Y A Mostovoy and V A Berdnikov 2018 J. Phys.: Conf. Ser. 1096 012184

View the article online for updates and enhancements.
Statistical modeling of a scale network of nanosatellites

Y A Mostovoy$^1$ and V A Berdnikov$^1$

$^1$Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

Abstract. Statistical modeling of nanosatellite networks having no motion control of the center of mass and therefore randomly distributed in an intended service area of the interturn ground track interval is considered. This modeling is conducted on square structured matrices as service area models by methods of percolation theory. The concept of programmable service area percolation has been introduced implementing in two phases. Numerically, using the results of statistical modeling of two-phase operations, the concentration value of NS of the stochastic base has been obtained ensuring the minimum of aggregate expenditures.

1. Introduction
In the course of evolutionary development, the methods for solving the problem of global observation from Earth's remote sensing (ERS) satellites have reached a high level of excellence. In order to increase the observation range to a significant interturn interval fraction, these devices are equipped with complex navigation and angular orientation systems allowing, as per appropriate control by angular movement, to observe predetermined targets out of the way of the ground trace. Such satellites can control the parameters of their trajectory and position on it by virtue of the inboard propulsion system independently or by commands from the Earth. Within the ERS concept, classes of heavy satellites and so-called small ERS satellites are developed by single or several similar complex satellites forming the system.

The article considers another concept for data accessing by ERS [1, 2, 3, 4]. There is simultaneously a large number of very simple ultra-small satellites – nanosatellites (NS) on the trajectory which, being united in a scale network, will solve the global ERS problem in a distributed way.

Obviously, the simplest NS, because of their small mass, may not have their own propulsion system for the center of mass control system and, therefore, the organized placement of separate NS in the network (general cluster) during NS flight cannot be supported.

The task for determining necessary number of NS, randomly distributed in space, is proposed to be solved by statistical modeling. The task for obtaining and flowing of ERS information through a network of satellites randomly placed at the interturn interval (in a random environment) complies well with the percolation theory formulations. In this case, appearance at a certain concentration – probability for location of the object in the intended space area of a stochastic percolation cluster which overlaps the interturn interval – makes it possible to determine the number of NS in the cluster [1, 2, 3].

The task statement of percolation theory is as follows. There is a matrix the random cell part of which is "black", conducting the flow, and the rest is "white", not conducting the flow. The "black" cells contacting with the edges (not vertices) form random conducting clusters that form and increase together with the increase in concentration of "black" cells. It is necessary to find minimum concentration of "black" cells at which a through path is formed on the black cells across the entire
matrix in the intended direction, in other words, concentration ($p_n$) at which the whole matrix starts to
draw [1, 5, 6, 7, 8, 9, 10, 11, 12].

In this case, concentration ($p$ – relative proportion of black nodes with randomly homogeneous
filling of the matrix) is probability of the black object in the matrix cell [1, 5]. In this case, the
developed modeling algorithm allows us to use not only uniform but also modal laws for distribution
of objects for matrix cells [2, 3, 4, 13, 14].

The article defines overall appearance of such NS, necessary composition of systems, system
formation of NS that possess the required properties. An adequate apparatus of statistical modeling is
proposed to study issues for forming NS system with the coverage of the operational area (interturn
interval). As a result, the key issue of the distributed ERS task is considered: the required number of
such basic NS.

2. Statistical phenomena of randomly distributed objects
Statistical modeling on square matrices allows to discover and investigate three statistical phenomena.
The first statistical phenomenon – the presence of the threshold for stochastic percolation as a
"breakdown" of the matrix by a conducting percolation cluster is described in detail under [1, 5]. The
second statistical phenomenon considered under [2, 3, 4, 13, 14, 15] – the presence of the
concentration value at which the average number of clusters formed has the maximum.

With regard to the problem of creating scale networks of NS, the use of the threshold for stochastic
percolation to determine the number of satellites required in a network leads to an excessive number of
satellites as shown under [2]. In this case, the percolation cluster turns out to be "loose and branched"
with a large number of unnecessary "dead" branches.

Therefore, under [3, 4] two-phase operations in scale networks are considered. In the first phase, a
stochastic base is created at concentration of objects much lower than the threshold for stochastic
percolation, and in the second phase, artificial programmable percolation is provided by introducing
additional objects in the optimal way at the available intercluster intervals.

As a result, total concentration of objects in programmed percolation becomes more than half the
threshold of stochastic percolation and is in the neighborhood of concentration specific for the second
statistical phenomenon.

In this article, new results have been obtained with respect to the analysis of the "path width" of
programmed percolation, determined by the specific "tortuousness tract of the path" of programmed
percolation which increase up to the threshold of stochastic percolation, after which they begin to
decrease. The developed algorithm for completing the aggregate of stochastically formed clusters of
the base with the goal to obtain the ways for controlled percolation of minimum length is called as
"lightning" due to the specific tortuous appearance of the way obtained [3, 4, 13, 14].

This feature of the path width and length of programmable percolation is the third statistical
phenomenon for formation of random clusters on scale networks.

The analysis of programmable percolation path width allows to reduce the required number of
objects for the selected type of coverage with the operating environment, while maintaining the
selected optimal concentration of objects only within the bandwidth equal to the width of the
programmable percolation path.

3. The network of NS, its tasks and features
According to the current classification, NS include the machineries of weight up to 10 kg [1]. Each
working NS used in the swarm, due to weight restrictions, shall have a simple orientation system, may
also not have a propulsion system to control the movement of the center of mass and direct connection
to the Earth but it shall necessarily have connections with the similar ones. In the structure of clusters,
NS can solve the tasks of distributed data collection, monitoring of migration of objects on the Earth's
surface, communication tasks and remote sensing.

The purpose of the considered ERS concept is not that instead of a heavy and expensive satellite, a
simple and light one and a small rocket is launched, but that the efficiency of targeted operations, for
example, by ERS, is increased due to the use of a large number of NS covering the intended service area at the same time. At the same time, a unique remote sensing satellite is replaced by a large number of simple NS produced in series. At the same time, the position of each NS in space shall be known for controlling the ERS process and binding the results. With this end in view, each satellite shall have a receiver for the satellite navigation system and a receiver and a transmitter for network inter-satellite communication through which, in particular, its own ephemerides can be transmitted. There must be several server satellites (SSR) in the NS network which may not solve the target task and maintain communication with the working NS of the nearest clusters on the one hand and with the Earth on the other hand. These satellite-servers shall be brought into the necessary intercluster intervals of NS network after it is formed. Apparently, the most advantageous option may be the solution of the problem for installing objects at intercluster intervals, when the satellite servers are transported and installed in the right place of the formed NS network by some special maneuvering satellite – the "bus". Alternatively, SSR has a propulsion system. The share of SSR (MS) relative to the amount of NS is small (within 20%) and depends on the chosen concentration of NS in the operating environment band. On the other hand, SSR in relation to each other shall be along the line of sight (without shading by the Earth). Obviously, management for the "system" of such NS to cover the intended service area is impossible and shall not be required. These issues shall be solved by the number of working satellites in the cluster and above-mentioned structure of the general cluster (availability of satellite servers). If there is a possibility to "build" NS in such a way that its structure covers the interturn interval \( L_{MB} \) at the equator with observation sectors, then the target global observation criterion would be fulfilled with the number of NS equal to \( N \). Here \( N \) – the number of NS to be determined from the expression:

\[
L_{NS} = \sum_{i=1}^{n} l_i = Nl_i
\]

Being \( l_i \) is the observation sector of the 1st NC or service area for each NS which is assumed as constant for each NS.

However, the lack of means for controlling the orbit after separation from the carrier rocket, the presence of random perturbations of the orbit of each NS makes it impossible to build a general cluster like a deterministic "system" and it is necessary to ascertain the position of NS relative to each other randomly over time within the framework of the NS aggregate.

4. The results of statistical modeling of scale networks on square matrices with random filling

The aggregate of NS geometrically distributed within the boundaries of the interturn interval of the path shall first be assumed as randomly placed on a square matrix with the number of nodes or cells determined \( L \) by the number of rows of the matrix (1). In this case, we assume that the interturn interval is geometrically mapped to the matrix height.

Suppose that in each cell of the matrix (or lattice site), there is a NS with probability, \( p \) or the cell is empty with probability \( 1 - p \). As a first step, we assume that the probable location of NS in the matrix cell is constant throughout the matrix.

The answer to the question 'What shall be the probable location of NS in the cell to create a percolation cluster connecting the upper and lower parts of the matrix?' is given by the percolation theory.

We have examined matrices of size 30 by 30, 50 by 50, 100 by 100, 1000 by 1000, the cells of which have been filled randomly first, taking into account the equiprobable distribution of the objects in the cells. Under [2, 3, 4, 13, 14], modal laws for distribution of objects for the matrix cells have been also considered. Mathematical experiments consisted in constructing a series of several hundred random matrices for each value \( p \) – probability of availability of NS in the matrix cell with further
determination of numerical characteristics for obtained distribution of random clusters of NS with respect to the chosen parameters and calculating their average values.

In the classical percolation theory, it is customary to consider an infinite matrix [1, 5, 12], however, in our case, in accordance with the formulation of the problem, it is necessary to consider the matrices – models of the operating environment of finite dimensions. This raises the problem of estimating the influence of the matrix dimensions on accuracy of the results of statistical modeling. As this probability increases in the range, the 0.1-0.3 matrix is filled with objects, and the number of clusters increases. The maximum value of the number of clusters is achieved with probability of having an object in the cell equal to 0.26.

In this case, a large number of clusters of small sizes is present in the matrix. After this point, when new objects are added with increasing concentration, they begin more actively joining to already formed clusters, the clusters merge and increase in size with decrease in the total number of clusters. Figure 1 also shows dependence of the number of clusters on the matrix dimensions, however, normalization of these results over the area (number of cells) of the matrix (\(L^2\)) allows us to get rid of this dependence.

![Figure 1](image)

**Figure 1.** The dependence of the average number of clusters on probability of availability of the object in the cell for the matrix 50 x 50 (denoted by dashed lines) and 100 x 100 – (a) and the average number of clusters normalized for the matrix area – b).

This is because the number of clusters emergence on the matrix at a certain concentration depends on the area of the matrix \(L^2\), being \(L\) – the size of the square matrix, measured by the number of cells in the line. In contrast to this, the path lengths on the matrix – for example, the length of the percolation path – depend on the size of the matrix \(L\) and can be normalized for it.

Statistical studies on matrices of various sizes have shown the statistical stability of the characteristics of the distribution of clusters and their independence from the matrix size under the above-mentioned normalization for sufficiently large matrices (\(L>15\)).

The above results allow us to estimate the threshold of stochastic percolation by an average value of 0.59, which gives the number of NS in the percolation cluster \(N_{pk} = 0.59L^2\), being, \(L^2\) – the number of cells in the matrix. In turn, the number of the matrix cells depends on the coverage area of communication and target equipment for one NS.

However, such approach for determining the amount of NS gives an obviously excessive amount, since the structure of the stochastic percolation cluster is rather branched and loosened [1, 5].

The maximum point on the curve in Figure 1 with probability of NS location in the matrix cell, equal to 0.26, gives the average maximum number of clusters.

5. **Programmable percolation in two-phase operations. The average length and width of the programmable percolation path**

Further decrease in the required concentration of NS and, consequently, necessary quantity for implementation of the coverage (percolation) of the service area can be achieved if the reference system of randomly distributed NS is first created at a relatively low concentration. Then, at the
second stage of the operation, at intercluster intervals of the "stochastic support", to introduce a number of additional NC minimally controlled in a manner so that together with the existing stochastic clusters they form a continuous percolation path of minimum length in the intended direction.

In this case, we can talk about programmable or controlled percolation as opposed to classical stochastic percolation.

Thus, by increasing concentration of the objects in the stochastic base of the scale network, it is possible to reduce the number of additional controlled NS needed to create the shortest percolation path and vice versa.

Taking into account the different costs of the first-order NS randomly distributed and the second-order NS implemented into certain service areas to achieve an artificial controlled percolation with a minimum path length, one can find the concentration of NS at which the total cost of creating a controlled percolation path will have a minimum.

This minimum cost shall be less than the cost for creating a purely stochastic percolation cluster \( p = p_a \) or the cost for creating a fully managed percolation of the intended service area without the stochastic base \( p=0 \).

Similarly, in other applications of this theory to scale networks, each of the objects stochastically distributed is cheaper than the matrix of the object being implemented at a certain place due to two reasons: availability of the latter means allowing it to be installed in the required intercluster interval, and at the expense of the cost for implementation of the operation itself.

In order to confirm these considerations, a statistical modeling of such two-phase operation has been carried out.

![Graph](image)

**Figure 2.** Dependence of the average number of the objects added to provide controlled percolation on probability of having an object in a cell for matrices in size 50 \( \times \) 50 (indicated by a dashed line) and 100 \( \times \) 100 – a). The same dependence normalized to the matrix size – b).

Figure 2a) shows the average numbers of the objects added during statistical modeling to obtain programmable percolation in the intended direction for matrices of various sizes. Figure 2b) shows dependence data that are normalized for the matrix size. After that, the graphics have been coincided.

It can be seen from this figure that as concentration increases, the number of the objects added to form a minimal path of controlled percolation decreases.

On the other hand, the "tortuosity" of this path and, consequently, its length increases due to use of an increasing number of passing clusters up to the value of the percolation threshold. This is shown in Figure 3.

Figure 4 shows as an example the visualization of several random matrices of different sizes and for different values of probability of availability of an object in the cell. On these figures, the cells, in which the object is contained, are marked with black. In statistical modeling, the matrix is filled with objects randomly using a random number generator. In order to recognize the formed NS clusters and measure their characteristics, the Hoshen-Kopelman algorithm has been used [16, 17].
In the same figures, the shortest paths of controlled percolation passing through the clusters statistically formed are noted, and realized by adding a minimum number of the objects to intercluster intervals. The developed algorithm [14] is a modification of the Dijkstra algorithm.

**Figure 3.** Dependence of the average normalized in size matrix path length of controlled percolation on probability of availability of an object in the cell.

The maximum value of the average shortest path of controlled percolation approximately corresponds to the concentration of the stochastic percolation threshold in the stochastic base. In this case, the path of programmable percolation at this time has the maximum tortuosity (maximum length), and the average number of the added objects is 0. With a further increase in concentration of the objects, the shortest path is rectified and reduced.

**Figure 4.** Visualization of clusters on several of the considered matrices in size 50 x 50 at specified concentrations $p$. The number of the added objects $S^+$ refers to the shortest path of controlled percolation, marked in the figures, via clusters stochastically formed.

$L = 1.12; S = 0.078; S_s = 0.74; p = 0.1; L = 1.35; S = 0.122; S_s = 0.5; p = 0.25; L = 1.57; S = 0.087; S_s = 0.25; p = 0.4$

The figures show: $p$ – concentration, $L$ – length of the lightning path normalized to the matrix size, $S$ – number of the clusters normalized for the matrix area, $S_s$ – number of the objects added to form the shortest controlled percolation path normalized for the matrix size.

The width of the programmable percolation path has been obtained in the course of mathematical experiment on many matrices of various concentrations and sizes. The experiment has been carried out as follows: After obtaining each shortest path of programmable percolation, the matrix has been determined by the value $\Delta j = |j_{max} - j_{min}|$, being $j_{max}$ – rightmost coordinate on the programmed percolation path, $j_{min}$ – respectively the extreme left, $\Delta j$ – size of the band for a particular matrix of size $L$. Further, the obtained results have been averaged (on each matrix of $L$ optimal paths, and total matrices of each size 500) and normalized for the matrix size (the data shown in Figure 5b). The experiment result is shown in Figure 5.
6. Modeling outcomes

Let us denote the cost of each objects randomly distributed \( a \), and the cost of one object installed in a specific place of the scale network (in our case, NS of the second kind) \( \theta(p) \). Then total cost of the two-phase operation \( P \) shall be:

\[
P = a p L^2 + \theta(p) \phi(p) L
\]

(2)

Here, the first component is the cost of the stochastic base of a scale network, and \( p L^2 \) is the number of NS of the first kind in the stochastic base.

The second component is the cost of the objects of the second kind added for formation of the shortest controlled percolation path through clusters stochastically formed, and \( \phi(p) L \) is the number of these added objects determined by the results of statistical modeling and shown in Figure 2b).

The values \( a \) and \( \theta_0 \) depend on a variety of factors typical to the specific design of nanosatellites, so it is reasonable to evaluate the costs of conducting a two-phase operation as a ratio function for the values of one of the complementary objects, taking into account its installation to the cost of one object of the stochastic base.

Having the results of statistical modeling on a large number of matrices for different concentration values, it is possible to estimate the average number of cells into which additional objects shall be installed in order to form a programmable percolation path. In accordance with Figure 2, the value of the number of complementing objects is \( \phi(p) L \).

Then \( \theta(p) = \frac{\theta_0 \phi(p) L}{L'(p)} \).

With this consideration in mind, equation of total costs (2) takes the form:

\[
P = a p L^2 + \theta_0 \phi^2(p) L^2 \frac{L'(p)}{L^2(p)}
\]

(3)

Let us consider the relative cost of a two-phase operation, reduced to the cost of a single-phase operation with purely stochastic percolation of NS, for which we divide the left and right sides of the resulting equation (3) by \( P_n = a p L^2 \) taking into account that for concentration of the percolation threshold \( p_n \approx 0.6 \) the value is \( \phi(p_n) = 0 \).

Then:

\[
P_{\text{err}} = \frac{P}{P_n} = 1.7p + 1.7 \left( \frac{\theta_0 \phi^2(p)}{\alpha L^2(p)} \right) = 1.7 \left( p + R \frac{\phi^2(p)}{L^2(p)} \right)
\]

(4)

being \( R = \frac{\theta_0}{\alpha} \) we denote the cost ratio of the complementary object to the object cost of the stochastic base, \( L'(p) \) – dependence of the length of the average normalized path of controlled percolation on concentration (Figure 3).
Dependence of the relative cost of a two-phase operation on concentration of the NS is shown in Figure 6 for different values of the $R$ of the ratio cost complementary to the artificial percolation of NS to the cost of NS of the stochastic base.

![Figure 6](image)

**Figure 6.** Dependences of relative costs for conducting a two-phase operation on concentration of NS of the stochastic base with different cost ratios of additional NS and the cost of NS of the stochastic base.

Analysis of the obtained dependencies shows that for two-phase operations in scale networks with the same cost of complementary objects used to create the shortest path of controlled percolation and the cost of the objects of the stochastic base, **optimum value of the concentration – probability of availability of an object in the cell of the stochastic base is 0.26.** This concentration value corresponds to the maximum number of clusters of the stochastic base and is twice as little as the threshold of stochastic percolation. In the event that the $R$ value of the optimal concentration of the stochastic base is not equal, the $1$ value of the optimal concentration increases.

The obtained outcomes not only do not depend on the dimensions of the matrices (after their corresponding normalization) but also statistically do not depend on the location on the random sector matrix through which the programmable percolation path passes. Therefore, the results of statistical processing may be transferred to rectangular submatrices $L$ height of which is greater or equal to the average path width of controlled percolation when the latter condition is fulfilled.

For concentration, the 0.26 relative width of the programmed percolation path is 0.15. Taking the service area sector size for each NS as 50 km, approximately 100 NS of the stochastic base are needed to cover the interturn interval, and the average number of added NS for formation of the shortest path of controlled percolation through clusters of the stochastic base does not exceed 20 per cent of the number of NS of the stochastic base at this concentration. The estimation has been carried out for the circular orbit with an average altitude of about 400 km.

7. Conclusions
As a result of the work done, the following may be concluded:

1. The concept of the solution of applied tasks is proposed, in particular, the acquisition of ERS information by a large number of simple ultra-small NS that do not have the means of controlling the motion of the center of mass randomly distributed in the intended service area of the orbit path interturn interval. A device for statistical study of such scale networks of NS based on percolation theory has been developed and proposed.

2. Statistical phenomena for formation of random clusters of NS in the statistical modeling of scale networks of NS on the matrices randomly filled has been considered and studied. It is shown that along with probability of location of the object (NS) in the matrix-concentration cell, which describes the threshold of stochastic percolation, there are two other remarkable points on the probability axis of location of the object in the cell – the point of maximum clustering, in which the average number of
3. The concept of programmable percolation, which is introduced in contrast to classical stochastic percolation, is proposed. Relying on this concept, two-phase operations on scale networks are considered. The obtained outcomes not only do not depend on the dimensions of the matrices but also statistically do not depend on the location on the random sector matrix through which the programmable percolation path passes.

4. An adequate apparatus for statistical modeling is proposed to study the issues for forming a system of NS with the coverage of the operational area (interturn interval). Expressions have been obtained for estimating overall efficiency of the two-phase operation on scale networks and concentration value of NS ensuring the minimum of total costs.

8. References

[1] Moskalev P V and Shitov V V 2007 Mathematical modeling of porous structures (Moscow: Fizmatlit) p 120
[2] Mostovoy Y A 2011 Statistical phenomena for scale distributed clusters of nanosatellites Bulletin of the Samara National Research University 26 80-89
[3] Mostovoy Y A 2013 Two-phase operations in scale networks of nanosatellites Computer Optics 37 120-130
[4] Mostovoy Y A 2013 Controlled percolation and optimal two-phase operations in scale networks of nanosatellites Bulletin of the Samara National Research University 11 53-62
[5] Tarasevich Y Y 2002 Percolation: theory, applications, algorithms (Moscow: URSS) p 109
[6] Golubev A S, Zvyagin M Y and Milovanov D S 2011 The percolation effect in information networks with intermittent communications Bulletin of the Lobachevsky State University of Nizhny Novgorod 2(3) 260-263
[7] Nekrasova A A and Sokolov S S 2010 Investigation of using the percolation theory for controlling data flows in information networks on transport Bulletin of the State University for the Navy and River Fleet (Saint-Petersburg) 32(4) 192-198
[8] Alexandrowicz Z 1980 Critically branched chains and percolation clusters Physics Letters A 80 284-286
[9] Babalievski F 1998 Cluster counting: Hoshen-Kopelman algorithm vs. Spanning three approach International Journal of Modern Physics pp 43-61
[10] Hoshen J and Kopelman R 1976 Percolation and cluster distribution. I. Cluster multiple labeling technique and critical concentration algorithm Phys. Rev. B 14 3438-3445
[11] Stauffer D 1979 Scaling theory of percolation clusters Physics Reports 54 74
[12] Stauffer D and Aharony A 1992 Introduction to Percolation Theory (London: Taylor & Francis) p 180
[13] Mostovoy Y A 2015 Modeling of optimal two-phase operations in random operating environments Optoelectronics, Instrumentation and Data Processing 51 35-41
[14] Mostovoy Y A and Berdnikov V A 2017 Large scale networks security strategy CEUR Workshop Proceedings 1901 187-193
[15] Wilkinson D and Willemsen J F Invasion percolation: A new form of percolation theory 1983 Journal of Physics A: Mathematical and General 16 3365-3376
[16] Galam S and Mauger A 1996 Universal formulas for percolation thresholds Phys. Rev. E 53(3) 2177-2181
[17] Sarshar N, Boykin P O and Roychowdhury V P 2004 Scalable percolation search in power law networks Proceedings of the Fourth International Conference on Peer-to-Peer Computing 2-9