# The Percolation Theory Based Analysis of Data Transmission Reliability via Data Communication Networks with Random Structure and Kinetics of Nodes Blocking by Viruses

Dmitry Zhukov
Institute of Information Technologies
Moscow State Technical University, MIREA
Moscow, Russia
e-mail: zhukovdm@yandex.ru

Sergey Lesko
Institute of Information Technologies
Moscow State Technical University, MIREA
Moscow, Russia
e-mail: sergey@testor.ru

*Abstract*—**In the paper, open global computer networks and data communication networks are considered as structures with a random topology. Processes of epidemics spreads are described by percolation theory. "The percolation thresholds", fraction of blocked nodes at which the whole network loses its working capacity, are calculated for different numbers of communications per node. For the real data communication networks with the average number of communications per node in the range of 2.5 to 3.5, the share of the used equitype equipment and the software types should not exceed the margin from 48% to 63%.**

***Keywords: data communication network; blocking nodes; network topology, the percolation threshold; virus distribution dynamics***

## I. Introduction

An important task in ensuring reliable functioning of data communication networks, as well as protection of the transferred information, is the study of the formation of groups of network node physically connected by communication channels but blocked (excluded from operation) for some reason. For example, blockage is possible during computer viruses epidemics. Under certain conditions, such groups of blocked nodes can increase in size and form clusters, which can lead to an overall loss of functionality of the data communication network. For instance, a cluster can form when there is some blockage of a backbone node of data network at the regional or city level. Alternatively, a cluster can originate in a base station of a mobile network as a result of peak load or overload, or when there is a computer virus epidemic in computer networks, which blocks the operation of different network equipment. Our objective is to develop a model describing the processes of nodes clustering based on the percolation theory, the main assumptions of which will be stated further on.

Historically, any data communication network starting from the city region level has an irregular random structure. The brightest example of such a network is the Internet. This is caused by many factors among which we can single out the following: providers having different network and communication equipment, a fluctuating number of subscribers with constantly changing connection topology and many others.

At present time, spreading of epidemics is often described as a process with structure similar a Kailey tree with random number of connections [1]-[2].

One can pay closer attention to a number of works by R. Pastor – Satorras and A. Vespignani, where the authors study the problem of defining the probability of infection depending on the node distance from the source of threat in networks of different scale and with varying number of nodes [3]–[7]. The authors used the scale and number of nodes as topological parameters; however, there were no special insights into the diversity of networking structures and the blocked nodes clustering.

In common case a scale free graph can have any number of nodes. Figure 1a shows such a graph with the total number of nodes equaling 100.

The description of virus epidemic topology using a scale free graph model produces interesting results. However, at some stage, infected network nodes can start sending copies of viruses to already infected nodes, and the process topology will look as shown in Figures 1b and 1c [1].With the help of a scale free graph model, we can consider the data transfer traffic dynamics [8], [9], as well as the processes of network structure hierarchical growth [10].

Obviously, if the amount of blocked nodes is not too large, there will be an "open" route (a way formed by unblocked nodes) between two randomly selected nodes located at a distance. We will refer to the amount of blocked nodes at which the network becomes nonfunctional as a *percolation threshold*–the network will be functional below this value despite the fact that it contains some nodes or their groups (clusters) blocked by viruses. Above the percolation threshold, the whole network turns off and loses its data transfer functionality. There is no "open" way between two randomly selected nodes.

Studying processes of blocked nodes clusters formation and data percolation in networks with different (including random) topology has a lot of scientific and practical importance for the development of topology of data communication networks having high fault-tolerant features.

It would greatly help improve their technical and economic, as well as operational characteristics, and create new methods and methodologies of computer network and applications protection.

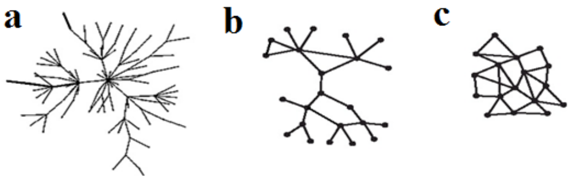

Figure 1. Different types of a scale free graph [2]: (a) – general case, (b)- the beginning of mutual DDoS attacks,(c) – the process of mutual DDoS attacks becomes considerable.

Besides, we should distinguish between two notions:

- Physical connections between the nodes. Two nodes are considered neighbors if they have a direct (without an intermediary) communication channel.
- Address linkage between the nodes. A virus can send its copy to a randomly selected node with a random IP-address instead of sending it to its physical neighbor.

In the latter case, the virus epidemic development topology looks like the Kailey tree (network) with a random number of communications, while in the former, the structure of physically connected infected nodes will be more complex and it has almost never been studied.

In Section II, we prove a choice of object and research methodology.

In Section III, we provide the description and discussion of results for data transmission modeling received in framework percolation theory for networks with random topology.

In Section IV we state the main conclusions drawn on the basis of the results received in the work.

## II. Subject and Study Methodology

We base our choice of network structure for a complex study of topology influence on its reliability on the fact that assessment of *real networks similarity* and different theoretical types of topologies on the basis of modeling can help single out a network (or types of networks) with the features closest to those of real networks (for example, the Internet), which is important for analysis of processes happening in the existing networks and ensuring their reliability.

Figure 2 shows the map of a mobile network operating in one of the Russian Federation regions.

The given map shows that real networks of data communication have a random structure similar to the one shown in Figure 1c; therefore, this article scrutinizes the random network with a set of communications per node.

As mathematical apparatus of the conducted research, we used the percolation theory, its basics being represented in [11]-[15].

During the modeling, we made the following assumptions: all nodes ($10^6$) of a computer network create a single network with a specific topology. Blocking of nodes occurs when infected with a computer virus. The virus can send its copies ($10^2$) from any node to any other arbitrary node (with probability of infection of $5 \cdot 10^{-3}$) by selecting its address from the entire set of address space (not necessarily physically connected nearby sites). At the next steps of the epidemics, the infected nodes are sending copies of the virus to other nodes in the network etc.

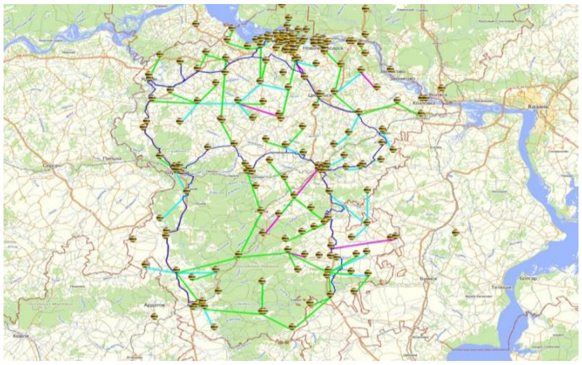

Figure 2. The location of base stations in the north of Chuvashia round the capital – the city of Cheboksary.

The average cluster size of blocked sites was determined at each step by numerical modeling methods; the infection process was carried out until the network reached the percolation threshold.

From a mathematician’s point of view, the percolation theory should be attributed to the probability theory in graphs. The most widespread problems of the percolation theory are the *lattice problems*, viz. the node problem and the connection problem. Let us consider a continued square grid. We shall name the points of line crossing *nodes (vertexes* of the graph); the lines themselves will bear the name of *communications (graph edges)*.

In the connection problem one tries to find an answer to the following: which share of communications should be eliminated (cut off) for the net to fall into two equal parts? In the node problem the nodes are blocked (removed, all the connections with the node being cut off) and one searches for the share of blocked nodes leading to network falling apart. In the percolation theory, a chain of connected items is called a *cluster*. A cluster connecting two opposite sides of the system is dubbed *percolating*, *infinite*, *spanning* or *connecting*. Below the percolation threshold, there are only clusters of a finite size.

The staff members of IBM R&D Centre (Scott Kirkpatrick, Winfried Wilcke, Robert Garner, and Harald Huels) studied the possibility of applying the percolation theory to the data storage systems [16]. They proposed the following model. A data cube of 1000 base elements connected by a cubic-cell type contained two types of cells (nodes), viz. the ones containing the immediate data and the cells (nodes) ensuring fault-tolerance – data replication. Since each node of such a system should not only provide the data output but also ensure data passage through a storage array (the access to other data), it was reasonable to employ the percolation theory. The use of the percolation theory allowed proving that it is enough to have just one copy of replicated data to ensure continuous fault-tolerant operation of the network. In case of excessive replication (two or more copies of data), one could observe a trespass over the percolation threshold, formation of non-conducting cluster in a cubic lattice, which led to the system operation failure. This model was put to good use in the system of data storage ‘Ice cube’ supplied by IBM Company and allowing the creation of a 32-terabyte array of data.

There are no analytical models elaborated to describe the percolation processes and to study random networks with multiple communications. Their research is possible only by numerical methods of modeling. For this purpose, at first it is necessary to construct a structural model of a network (see Figure 3), then, to choose a couple of any arbitrary nodes and using numerical modeling methods to define at what part of unblocked nodes in the considered network there is a freeway between A and B nodes. Then, this procedure is likewise performed for any other couples of nodes (in our case for the couples of C and D, E and F nodes in Figure 3 etc.). After that, with statistical averaging the results of separate experiments, we determine the average value of percolation threshold for all considered couples of nodes.

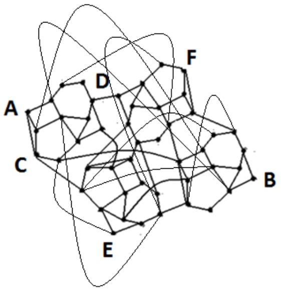

Figure 3. Data transmission random network structure.

## III. RESULTS AND DISCUSION

### A. *Percolation (flow) of information in a random network of data communication*

Table 1 presents the results of numerical modeling to find percolation threshold for random networks with the set of ways between nodes (see Figure 3) and various averages of communications per node.

With an increasing average number of communications per network node, the time and computing resources consumption significantly increases as well. For this reason, we had to choose the number of communications per node ranging from 2.5 to 15 in our numerical modeling.

In Figure 4 the dependence of the results given in Table 1 is shown. The percolation threshold decreases monotonically to 0.115 with the growth of the communications average per one network node. There is really no need to carry out numerical modeling at great values of average of communications per node, and it is possible to extrapolate the results onto the area of great values.

The graphical type of dependence in Figure 4 is similar to exponential law, therefore it can be described by function: $P(x) = P_0 e^{-z}$ , where P(x) is the percolation threshold value at the average of communications per node equaling some value x, z=1/x; $P_0$ is the percolation threshold valueat an infinitely large number of communications per node.

TABLE I. PERCOLATION THRESHOLDS FOR RANDOM NETWORKS

| **Network type** | **Average number of communications per one network node** | **Fraction of the activated nodes at which there is conductivity in the network ($n_c$ - percolation threshold)** |
|---|---|---|
| ***A random network with a set of paths between nodes*** | 2.36 | 0.515 |
| | 2.82 | 0.425 |
| | 3.29 | 0.365 |
| | 4.70 | 0.270 |
| | 4.75 | 0.250 |
| | 6.15 | 0.150 |
| | 6.17 | 0.185 |
| | 6.75 | 0.175 |
| | 9.41 | 0.170 |
| | 10.02 | 0.150 |
| | 10.31 | 0.130 |
| | 10.69 | 0.135 |
| | 11.07 | 0.115 |
| | 13.10 | 0.115 |

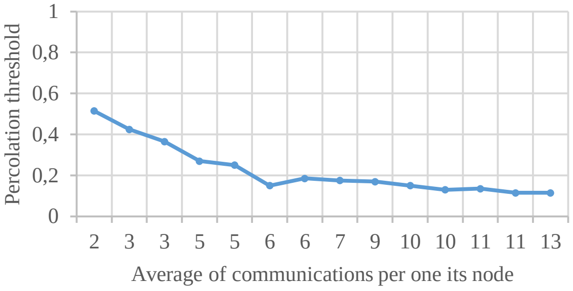

Figure 4. Dependence of percolation threshold size in random network on the average of communications per its one node.

As Figure 5 reveals, the data presented in Table 1 are well linearized in coordinates: lnP(x) depending on z=1/x (a natural logarithm of the percolation threshold is an inverse value to the average of communications x per node) that confirms the possibility of using the function: $P(x) = P_0 e^{-z}$.

Points in Figure 5 mark the experimental data, and the solid line corresponds to the linear dependence:
y = 4.39z-2.41, with a big value of correlation coefficient that equals 0.95.

At z=1/x=0 (corresponds to the case x = ∞) we receive: y=ln$P_0$ =-2.41, and the value of the percolation threshold at infinitely large number of communications per $P_0$ node will be equal 0.09.

It should be noted that, logically, it has to tend to 0; however, the obtained result can be explained as follows. At a very large number of communications, there can be a change in the law of dependence of percolation threshold on the number of communications. Nevertheless, it is possible to claim that the received dependence remains fair for random networks with significantly large number of communications per node. Thus, for a random network with an infinitely large number of communications per node it is enough to have the fraction of the activated nodes equal to 0.09 from the total number so that there could appear a carrying-out chain of nodes and the network could solve the set information task. At the average of communications equaling 100, the threshold of a percolation will amount to 0.094, and at 10 to 0.139 respectively.

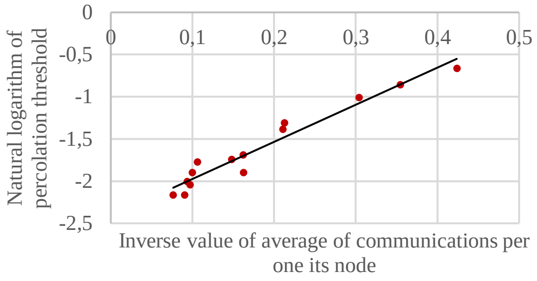

Figure 5. The logarithmic dependence of percolation threshold value in a random network on the inverse value of average of communications per its one node.

In practice, the number of communications of a random network per node will from 2.5 to 3.5, which yields the percolation threshold values ranging from 0.52 to 0.37. We calculated it using the data given in Figure 5 and the equation = 4.39z-2.41

### B. *Clustering of a random network*

If we consider the transition of any node of a random network from the efficient (not infected) state in the blocked state as a random process (with some probability of transition), this probability has to influence the average size of a cluster (a group of the nodes directly interconnected) of the blocked nodes.

By numerical experiments, we studied the influence of blocking probability on the average size of blocked nodes cluster of random networks with various average numbers of communications per node. Network with 10000 nodes were investigated. Two limit cases were considered: one for a small average of communications per node of a random network (see Table 2 and Figure 6) and the other limit case is for a big average of communications.

In Figure 6, curve 1 corresponds to the average of communications per node of a random network equaling 2.13; curve 2 depicts *idem* equaling 2.53; curve 3 is for 2.80 and curve 4 outlines for value 3.27 respectively. As the data in Table 2 and in Figure 6 demonstrate, the size of the cluster of the blocked nodes depends on the average of communications and probability of infection. With the growth of the average of communications, at the fixed probability of blocking, the size of the cluster increases. We can observe a similar situation with a great average of communications per node of a random network.

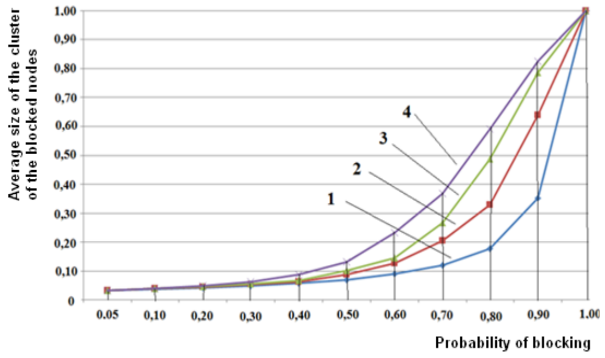

Figure 6. Dependence of the average size of the cluster of the blocked nodes on probability of blocking (for example, infections with a virus).

TABLE II. DEPENDENCE OF THE AVERAGE SIZE OF THE CLUSTER OF THE BLOCKED NODES ON PROBABILITY OF BLOCKING

| **Probability of blocking** | **The average size of the blocked nodes cluster (in fractions relative to the total number)** | | | |
|---|---|---|---|---|
| | ***For the network with an average of communications per node equaling 2.13*** | ***For the network with an average of communications per node equaling 2.53*** | ***For the network with an average of communications per node equaling 2.80*** | ***For the network with an average of communications per node equaling 3.27*** |
| 0.05 | 0.033 | 0.033 | 0.033 | 0.033 |
| 0.1 | 0.037 | 0.039 | 0.039 | 0.039 |
| 0.2 | 0.041 | 0.044 | 0.047 | 0.049 |
| 0.3 | 0.049 | 0.054 | 0.056 | 0.062 |
| 0.4 | 0.059 | 0.062 | 0.067 | 0.087 |
| 0.5 | 0.070 | 0.087 | 0.101 | 0.132 |
| 0.6 | 0.090 | 0.126 | 0.146 | 0.233 |
| 0.7 | 0.120 | 0.206 | 0.268 | 0.368 |
| 0.8 | 0.178 | 0.330 | 0.487 | 0.591 |
| 0.9 | 0.353 | 0.638 | 0.785 | 0.825 |
| 1 | 1 | 1 | 1 | 1 |

### C. *Kinetics of blocking nodes in the address space of computer networks and achieving the percolation threshold*

Currently, researchers recourse to empirical susceptible–infectious (SI) and susceptible–infectious–recovered (SIR) models originating from biology [17] to describe the kinetics of infection of data communication networks. However, instead of the empirical ones, some more reasonable mathematical and information models are required for the adequate description of the blocking processes.

To create such a model, we have considered the network consisting of L computers, for example, in which viruses can reproduce with coefficient of reproduction equal to ξ. Viruses start spreading before they are detected and before an efficient antivirus appears, which can efficiently eliminate the viruses. An antivirus program appears only at a certain step of virus distribution, lagging behind the start of virus distribution by $h_0$ steps, i.e., at step k, $k = h - h_0$ (there is a delay).

The number of the antiviruses appearing at step (k + 1) (at step (h + 1) for viruses) is designated as $N_{k+1}$, and the number of viruses appearing at step k (step h for viruses) is denoted as $N_k$. In other words, $N_k$ will be equal to the number of computers, at which at k step an antivirus will be available, and $N_{k+1}$is equal to the number of computers, at which antivirus will be available at step (k + 1).

The number of computers infected at step (h + 1) can be defined as $P_{h+1}$, and the number of computers infected at step h can be indicated as $P_h$. The change in the number of infected computers is equal to the difference between the number of infections and the number of viruses destroyed at step (h + 1).

There are the following random events that form the complete system:

1. A computer is infected with a virus with probability of $\frac{P_h}{L}$.
2. There is an antivirus at the computer with probability of $\frac{N_k}{L}$.
3. There is neither a virus, nor an antivirus at the computer with probability of $\left\{1-\frac{P_h}{L}-\frac{N_k}{L}\right\}$.

The number of infections at step (h + 1) will be equal to $\xi P_h\left\{1-\frac{P_h}{L}-\frac{N_k}{L}\right\}$ as the infection of the already infected computer is not considered, and the computer where an antivirus is installed cannot be infected.

The number of viruses eliminated at step (h + 1) has to make up $P_h\frac{N_k}{L}$, where $\frac{N_k}{L}$ is the probability that at step (h+1) any of $P_h$ viruses existing at step h can encounter an antivirus. Thus

$$P_{h+1}-P_h = \xi P_h\left\{1-\frac{P_h}{L}-\frac{N_k}{L}\right\}-P_h\frac{N_k}{L} \tag{1}$$

The change in the number of computers where the antivirus is installed at step (k +1) is defined by $N_{k+1} - N_k$ difference:

$$N_{k+1}-N_k = \xi P_h\left\{1-\frac{N_k}{L}\right\}, \tag{2}$$

where $\xi P_h$ implies that the antivirus is installed at step (h + 1) at those computers where a virus has been detected at step h, and $\left\{1-\frac{N_k}{L}\right\}$ means that the antivirus is installed only where it has not been present.

Since the duration of each step is equal to τ, the duration of the whole process t and number of steps h are interconnected by the following ratio of $t=h\tau$ and $t_0 = h_0\tau$ ($k = h - h_0$), where $t_0$ is the time when the antivirus springs into acting (its action lags behind the onset of viruses by the interval time value $t_0$).

Proceeding from the number of steps h and k to the process duration, we will receive:

$$P(t+\tau)-P(t) = \xi P(t)\left\{1-\frac{P(t)}{L}-\frac{N(t-t_0)}{L}\right\}- \\ -P(t)\frac{N(t-t_0)}{L} \tag{3}$$

$$N(t-t_0+\tau)-N(t-t_0) = \xi P(t)\left\{1-\frac{N(t-t_0)}{L}\right\} \tag{4}$$

We will denote $t\text{-}t_0 = y$ and, having decomposed (3) and (4) into a Taylor row, we will receive:

$$\tau\frac{dN(y)}{dy}+\frac{\tau^2}{2}\frac{d^2N(y)}{dy^2}+\cdots = \xi P(t)\left(1-\frac{N(y)}{L}\right) \tag{5}$$

$$\tau\frac{dP(t)}{dt}+\frac{\tau^2}{2}\frac{d^2P(t)}{dt^2}+\cdots = \\ = \xi P(t)\left\{1-\frac{P(t)}{L}-\frac{N(t-t_0)}{L}\right\}- \\ -P(t)\frac{N(t-t_0)}{L} \tag{6}$$

with an entry condition of N (y=0) = P ($t_0$), where $y = t - t_0$.

The equations (5) and (6) essentially differ from the system of equations used in the SIR model. The fundamental differences are:

- In the offered equation of infection (6), the decrease of viruses in the right part is determined not only by a share of nodes susceptible to infection $\left\{1-\frac{P(t)}{L}-\frac{N(t-t_0)}{L}\right\}$, but also by the product of the number of viruses and the probability of their encounter with $\frac{N}{L}$ antivirus, whereas the SIR model implies that the number of viruses decreases with the constant average speed of "immunization" per unit of time γ. Besides, the second derivative reduces the infection speed due to mutually reciprocal attacks (when we transfer it to the right member of the equation, a minus sign appears).
- The SIR model assumes that the speed of antivirus's emergence (the speed of disinfection) linearly depends on the number of available viruses. In the model offered (5), it depends on the probability of $\frac{N}{L}$ antivirus presence in the node and it is indirectly affected via $\frac{d^2N}{dt^2}$ on updating of the antivirus base. When transferring this member of the equation into the right part, a minus sign occurs, and the second derivative implies that the already available antivirus protection requires a base update, and instead of mailing over a network and installation of new antiviruses, they are just being updated.

This approach allows deducing the following differential equation that describes the kinetics of computer viruses epidemic development without protection by an antivirus (N $(t-t_0)$ =0):

$$\frac{\mathrm{dP(t)}}{\mathrm{dt}} = \xi\mathrm{P(t)}\left(1-\frac{\mathrm{P(t)}}{\mathrm{L}}\right) - \mathrm{O}\left(\frac{\mathrm{d^nP(t)}}{\mathrm{dt^n}}\right) \tag{7}$$

The left member of (7) describes in general the speed of emergence of new infected computers or network nodes. The member of the (7) $\xi\mathrm{P(t)}\left(1\text{-}\frac{\mathrm{P(t)}}{\mathrm{L}}\right)$ describes the inception of new infected computers, i.e. the existence in (7) of just this summand implies that all copies of viruses penetrate only the computers that are not infected. Moreover, the member of $O\left(\frac{d^nP(t)}{dt^n}\right)$ view allows considering some part of the dispatched viruses to penetrate the already infected nodes (and thus reduce the infection since there is a minus sign before it).

Figure 7 presents the comparison of results of the empirical SI model and the model we offer, which is based on differential (7) that considers the derived changes of the second and higher order of the number of viruses over time. Curve 1 represents the SI model, and curve 2 depicts the offered model that takes into account the second derivative, curve 3 is *idem* for the third derivative, curve 4 is *idem* for the fourth derivative, and curve 5 is the same model taking into account the fifth derivative. All results are received for identical values of parameters (τ = 25, L = 200000, $P_0$ = 5 and ξ = 2).

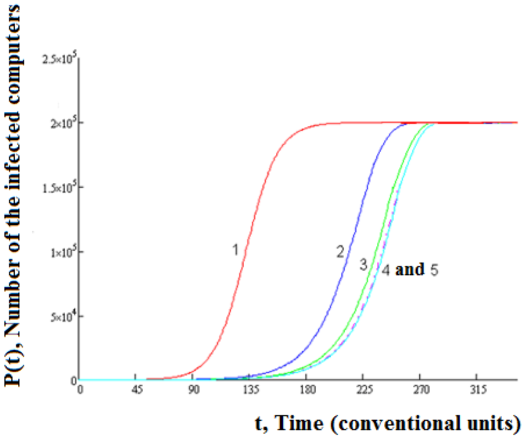

Figure 7. The comparison of SI model (curve 1) and the solution of the equation of the offered model.

Figure 8 presents the comparison of the offered model and the results of observation over development of epidemic of Code Red Worm [18] (the curve describing data is deduced using (7) and taking into account the summands of $O\left(\frac{d^nP(t)}{dt^n}\right)$ form, the dotted line represents experimental data). Figure 8 shows that the data observed and theoretical calculations coincide well with values of $P_0$ =1 (attack begins with one node), ξ =3 (the coefficient of reproduction equals 3, it is chosen to adjust the theoretical curve to the observed data), L=350000 (according to the observations presented in [16], the attacked network consisted of 350000 nodes), τ = 70 (duration of one step of the epidemic development equals 70 conventional units of time or 3.89 hours, it is chosen to adjust the theoretical curve to the observed data). Thus, the number of computers infected per hour is $\beta = \frac{\xi}{\tau} = 0{,}77$ . It coincides with an assessment of 0.7 to 1.8 nodes provided in [18] for random mailing.

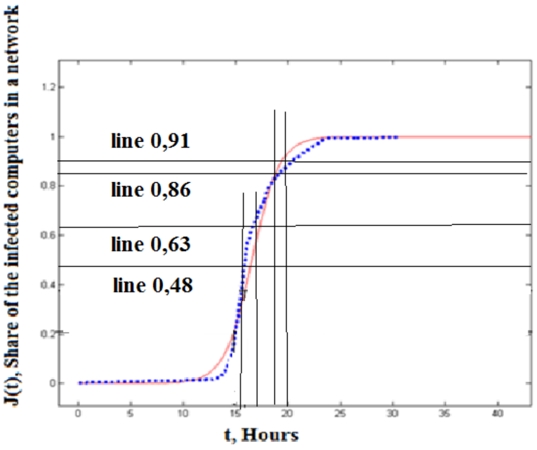

Figure 8. Comparison of observations over epidemic of Code Red worm (dotted line) and calculations according to the offered model (solid line).

In Figure 8, the horizontal lines show the allowance for the values of percolation thresholds ranging from 0.09 to 0.14 for a random network with the set of ways between nodes blocked by viruses (line1-0.09=0.91 and line 1-0.014=0.86). As seen in Figure 8, if the percolation threshold is to be considered as a criterion of reliability and operability of a computer network in general, then the network shut down from the normal operating mode will occur approximately in 19-20 hours after the beginning of infection, which allows taking necessary measures to eliminate epidemic consequences. This begs a question of why infection of computers continued. The answer is that a network consists not only of nodes that can be potentially infected, but also of nodes which cannot be any how infected due to the lack of vulnerabilities since they have another type of software. Yet, these nodes can transmit viruses through the network from infected to not infected nodes, remaining invulnerable. Besides, being infected, network nodes can continue carrying out functions on data transmission.

## IV. CONCLUSIONS

Open global computer networks and networks of data communication can be considered as the structures with random topology, and the processes running in such networks can be described by percolation theory.

During computer viruses epidemics distribution, data communication network nodes blocking can happen, as well as formation of clusters of such nodes. There are a number of blocked nodes at which all network entirely loses operational capacity (hitting the percolation threshold) in spite of the fact that a considerable part of nodes is still in an operational state.

For a random network, in the limit of infinitely large number of communications per node, it is enough to have the fraction of unblocked nodes equaling 0.09 relative to the total number so that there can arise a transferring chain of nodes and the network can solve the assigned information task. At an average of communications equaling 100, the percolation threshold will equal 0.094, and at 10 – 0.139. Thus, it is possible to consider that the fraction from 0.09 to 0.14 nodes keeping operational capacity allows providing overall operability of a network and its reliability. When we reduce the average of communications per node up to 2.5 – 3.5, the network keeps the general operational capacity for number of unblocked nodes from 0.52 to 0.37.

The smaller the average number of connections per node is, the less the blocking time of data communication network as a whole and attaining the percolation threshold turns out to be.

The size of a cluster of the blocked nodes depends on the average of communications and blocking probability. With the growth of an average of communications, at the fixed blocking probability, the size of cluster increases. A similar situation is observed with a large average of communications per node of a random network.

Practical recommendations for protection of any data communication networks against threats of virus attacks are essentially the following. When using the equitype equipment and software to create networks of data communication with a very large number of communications per node, its share should not exceed 86% - 91%. It will allow keeping operability of all network as a whole during epidemics spread of multivector viruses capable of deploying for their penetration not one, but the whole set of vulnerabilities, since it increases the probability that 9% -14% of the employed equipment and types of software will be impregnable. However, in reality the average of communications per node of network varies from 2.5 to 3.5, which yields the percolation threshold ranging from 0.52 to 0.37. Thus, for the real computer information networks the fraction of the employed equitype equipment and software types should be strictly within the limit from 48% to 63% (1-0.52=0.48 and 1-0.37=0.63).

Describing the distribution dynamics of computer viruses and using the differential equations of the second and higher order to account for changes in the number of the infected nodes over time, the mathematical model enables allowances for mailing copies to already infected addresses. It will also be significantly better coordinated with the results of observation over computer viruses epidemics in the Internet, than the existing SI and SIR models are.

In the future we will consider probabilistic schemes of transitions between statuses of congestion of data transmission networks. Such formalization is able to receive the differential equation of second order (like Kolmogorov's equation) which modeling stochastic dynamics of change status of congestion a network and to connect them with the results received from the percolation models.

## REFERENCES

[1] J. Nazario, "Defense and Detection Strategies against Internet Worms"; Artech House, 2004, ISBN: 1580535372.

[2] J. Leveille, "Epidemic Spreading in Technological Networks", Information Infrastructure Laboratory HP Laboratories Bristol, 2002, available at http://www.hpl.hp.com/techreports/2002/HPL-2002-287.pdf

[3] R. Pastor-Satorras and A. Vespignani, "Epidemic dynamics and endemic states in complex networks", Physical Review E, vol. 63, 2001, pp. 0661171 – 0661178.

[4] R. Pastor-Satorras and A. Vespignani, "Epidemic spreading in finite scale-free networks", Physical Review E, vol. 65, 2001, pp. 0351081 – 0351084.

[5] R. Pastor-Satorras and A. Vespignani,"Epidemic spreading in scale-free networks", Physical Review Letters, vol. 86, 2001, pp. 3200–3203.

[6] R. Pastor-Satorras and A. Vespignani, "Epidemics and Immunization in Scale-Free Networks", Wiley-VCH, S. Bornholdt and H. G. Schuster (eds.) Handbook of Graphs and Networks: From the Genome to the Internet, 2005, DOI:10.1002/3527602755.ch5.

[7] R. Pastor-Satorras and A. Vespignani,"Immunization of complex networks", Physical Review E, vol. 65, 2002, pp. 036104.

[8] A. Fekete, G. Vattay and L. Kocarev, "Traffic Dynamics in Scale-Free Networks", Complexus, vol. 3, 2006, pp. 97–107, DOI: 10.1159/000094192.

[9] Zhi-Xi Wu, G. Peng, Wing-Ming Wong and Kai-Hau Yeung, "Improved routing strategies for data traffic in scale-free networks", Journal of Statistical Mechanics: Theory and Experiment, vol. 2008, 2008, P11002, DOI:10.1088/1742-5468/2008/11/P11002.

[10] S. Boccaletti, D.-U. Hwang and V. Latora,"Growing hierarchical scale-free networks by means of nonhierarchical processes", International Journal of Bifurcation and Chaos, Vol. 17, No. 7,2007, pp. 2447–2452, DOI:10.1142/S0218127407018518.

[11] G. Grimmett, "Percolation and disordered systems, in Lectures in Probability Theory and Statistics", Ecole d'Eté de Probabilités de Saint-Flour XXVI-1996, Springer Lecture Notes in Math. no. 1665, ed. P. Bernard, 1997, ISBN978-3-540-63190-3.

[12] G. Grimmet, "Percolation", Springer-Verlag, 1999, ISBN978-3-540-64902-1.

[13] M. B. Isichenko, "Percolation, statistical topography, and transport in random media", Rev. Mod. Phys., vol. 64, 1992, pp.961-1043, http://dx.doi.org/10.1103/RevModPhys.64.961.

[14] M. Sahimi, "Applications of Percolation Theory", Taylor & Francis, 1992, ISBN 0748400761.

[15] V.K.S. Shante and S. Kirkpatric, "An Introduction to Percolation Theory", Advances in Physics, vol. 85, 1971, pp. 325-357, DOI:10.1080/00018737100101261.

[16] W.W. Wilcke, R.B. Garner, H. Huels, "Percolation in dense storage arrays", Physica A: Statistical Mechanics and its Applications, Vol. 314(1), 2002, pp. 220-229, DOI:10.1016/S0378-4371(02)01153-6.

[17] C. C. Zou, D. Towsley, W. Gong, "On the performance of Internet worm scanning strategies", Performance Evaluation vol. 63, 2006, pp. 700–723, DOI:10.1016/j.peva.2005.07.032.

[18] C. C. Zou, W. Gong, D. Towsley, "Code Red Worm Propagation Modeling and Analysis", 9th ACM Symposium on Computer and Communication Security, 2002, pp.138 – 147.