

# Clusterization in networks based on the principles of translational partitioning of space

*Kirill Gorshkov,  
Ph.D, National Research University Higher School of Economics  
Moscow, Russia, 101000, 20 Myasnitskaya Street  
[godograf@list.ru](mailto:godograf@list.ru).*

*Valery Rau,  
Prof. Dr., Vladimir Branch of the Russian Presidential Academy of National Economy and Public  
Administration, Vladimir, Russia, 600017, 59a Gorky Street  
[vgrau@mail.ru](mailto:vgrau@mail.ru)*

*Oleg Nikitin  
Prof. Dr., Vladimir State University named after Alexander and Nikolay Stoletovs  
Vladimir, Russia 600000, 87 Gorky Street  
[olnikitin@mail.ru](mailto:olnikitin@mail.ru)*

*Saleh Hadi  
PhD, Associate Professor of National Research University Higher School of Economics,  
Associate Professor of Vladimir State University named after Alexander and Nikolay Stoletov,  
Moscow, Russia 125319  
[hsalekh@hse.ru](mailto:hsalekh@hse.ru)*

*Ekaterina Kuznetsova  
Vladimir State University named after Alexander and Nikolay Stoletovs  
Vladimir, Russia, 600000, 87 Gorky Street  
[indikatrina@rambler.ru](mailto:indikatrina@rambler.ru)*

**Abstract:** This paper discusses ways of structuring networks using clusterization. In the work, the probabilities of reaching fixed network nodes are calculated for various characteristics of a data transmission system. It will be shown here the clusterization in the network allows us to specify digraph, the numbering of vertices in which makes it possible to calculate the stages of transmission using multiplication table of substitution.

**Keywords:** broken symmetry group, clusterization in networks, digraph, data transfer.

## Introduction

Currently, the systems based on the network organization are of great interest. Moreover, equally determined challenges are faced both by the sphere of development of the material components of data transmission tools, in particular, in the field of quantum communication [1-3] and nanophotonics [4], and by the sphere concerning the structural organization of networks without regard to the specific features of their main components. The network organization is a specification not only for information and communication systems of data transmission, although they are affected primarily with a complex, changing topology and adaptive algorithms of functioning, but also for the systems in the financial sector (blockchain) or the network retail. The intensive introduction of systems based on the conception of the Internet of Things [5] in various

spheres of human activities presents a number of problems in regard of finding the best options for topologies or clustering, in the case of mobile node properties, for example, in USN (Ubiquitous Sensor Network)[6]. New interest in peer-to-peer networks arose due to the great commercial success of Uber, which spawned the term “uberization” in relation to various market sectors, as well as emerging and development of the so-called cryptocurrency redistribution systems (Bitcoin [7]). Moreover, taking in account that network nodes are mostly devices, rather than personalities, whose number is limited within the population, the number of nodes can reach several trillions [5], which leads to the necessity of creation of all sorts of ways of structuring: continuous transmission of transit data can become the cause for the failure of the power supply, and a large volume of traffic leads to the overflow of the receiving buffers. For example, clustering organization offers some solutions, when implemented in various variants in USN wireless sensor networks and MSN (Mobile Sensor Network) as well as in the way of dividing the network address' space into smaller sub-nets [8].

## 1 Self-organization of sensor networks based on symmetric organization

The connectivity criterion for such networks is described in detail in the articles [9, 10]. Connectivity is considered as a measure for a possible intertwist of sensory nodes with each other (Figure 1). If we consider MSN, which is characterized by a sufficiently large number of sensor nodes distributed in the so-called sensory field, then the connectivity criterion is:

$$S = \{s_1, \dots, s_n\}.$$

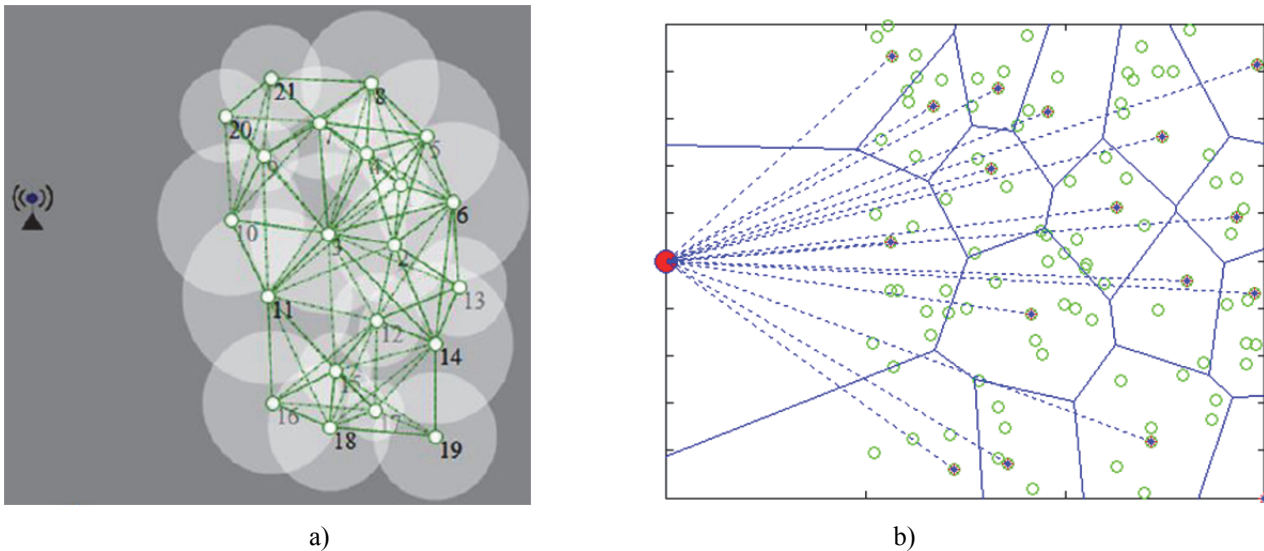


Figure 1 – An example of connectivity in the sensor network (a) and clustering (b) [9]

The sensory node interconnection area  $s$  is the area in which one sensor node can interact with other sensor nodes. The limit of interconnection ( $R_i$ ) of an arbitrary sensor node  $s_j$  is the maximum distance between nodes  $s_i$  and  $s_j$ , where  $s_j$  is within the interconnection range  $s_i$ .

The number of sensors closest to the node is calculated as  $N(s_i) = \{s_j: d(s_i, s_j) < R_i\}$  where  $d$  is the Euclidean distance between  $s_i$  and  $s_j$ . With this in mind, the connectivity criterion can be defined as following:  $ConC(s_i) = |N(s_i)|$ .

From the mathematical point of view, this condition presents the Delaunay condition for the  $(R, r)$ -system: the distance from any point of the set to the point of the same set nearest to it is greater than or equal to some fixed segment of length  $r$ , and the distance from any point of space to the closest point from the one of the system points are less than or equal to some fixed segment of length  $R$  (Figure 2a).

As an example of such an  $(r, R)$ -system can serve a system containing symmetry, which Delaunay first used to describe the periodic structures of the material. A similar approach can also be applied to clusterization networks based on the principles of symmetric partitioning of space (with the topology of the dual symmetric graph selected as a network model (Fig. 2b) [11]).

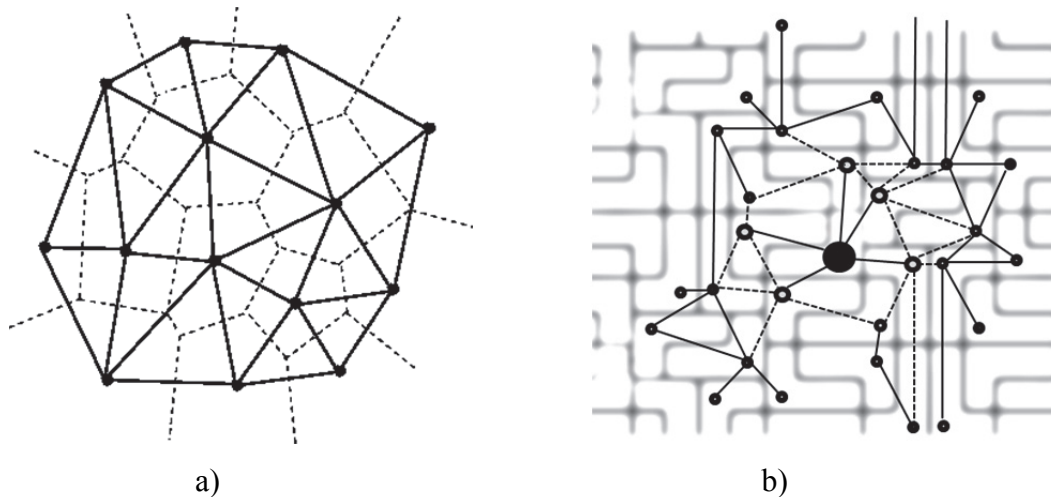


Figure 2 – An example  $(R, r)$ -Delaunay systems – (a); symmetric partition into polyominoes and its corresponding neighbourhood graph – (b)

An area is counterposed to a node or a set of nodes combined into one cluster. Space consisting of equal squares - polyominoes, the connection between nodes (or Clusters Head) is depicted for the case, when the polyominoes has joint sides. Thus, if a neighbourhood relation is given on the set of polyominoes, one can always proceed to the geometrical description of the network in the form of a graph. To implement a clustering procedure based on symmetric partitioning, layer-by-layer growth algorithms for partitioning and packaging, elementary polyomino clusters [12] or a periodic graph growth algorithm [13] can be used as the basis for automatization of the process of data transmission in networks. For example, the algorithms for layer-by-layer growth of the partitioning or growth of a periodic graph can also be used in the formation of wireless sensor networks as a ground for the self-organization. A sensor network is a set of sensors distributed in a chaotic manner in a certain area, called the sensor field, where with time the network should organize itself without the participation of an external administrator. Any node of the sensor network can perform the functions of both a terminal and a transit node, and data transmission is carried out in a step-by-step manner (multi-hop network). At least one of the other sensors (Fig. 5a, b) must be located within the coverage area of the signal generated by the sensor.

In this way, the connection is being formed between the nodes located in such an area, the connected nodes are usually called adjacent. Network reliability is determined by the number of "neighbours": the more connected nodes there are, the more reliable is the network: the failure of one of the sensors should not affect the ability of data transfer, which can be implemented through any alternative routes [9,10].

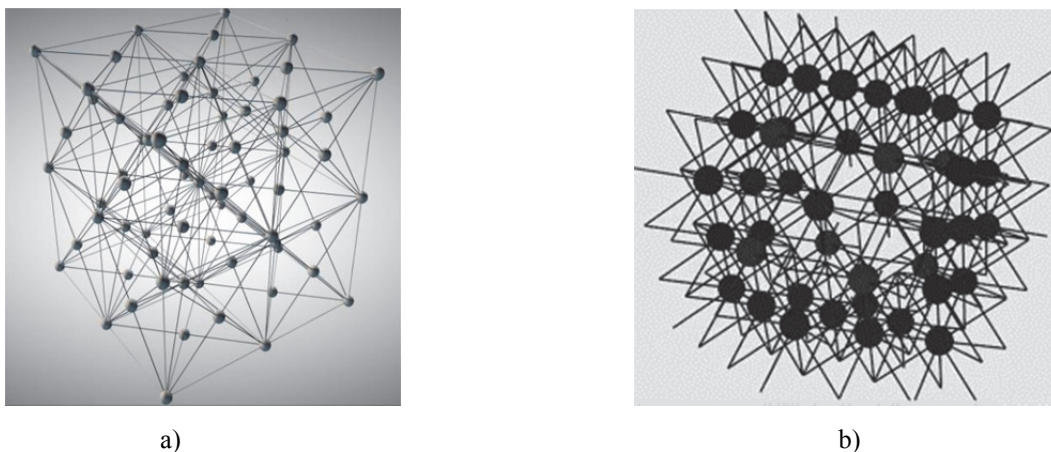


Figure 3 – Examples of a sensor network in three-dimensional (a) space [6] and the forming of a solid body structure in a discrete space, based on the partition (b)

By analogy with crystals (Figure 5b), whose growth can be modelled in a discrete periodic space, the formation of links can occur in a discrete periodic space, where to each network node an elementary descriptor of such a space is assigned — i.e. polyomino. The neighbourhood relation of the set of polyominoes also makes it possible to form a connection between the nodes corresponding to the neighbouring (possessing a common side) polyominoes (Fig. 6).

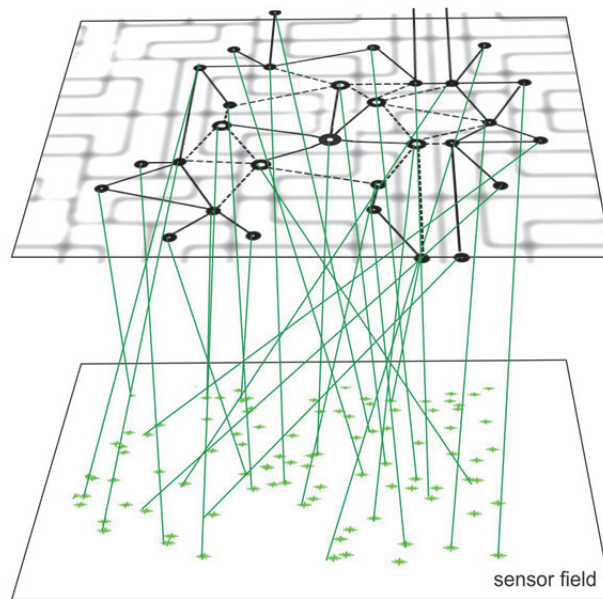


Figure 4 – Sensor field and network self-organization in discrete periodic space

When moving to a “virtual” discrete space, where connections arise in accordance with a given polyomino neighbourhood relation, it becomes possible to structurize the logical topology, including the case of a chaotic distribution of nodes in the sensory field (for example, Poisson ensemble). In such a network, there should not be any problems with the "inclusion" even of such nodes in the network, which doesn't possess radio coverage of any node. The sensor nodes in such a network may have mobile properties, since their logical topology does not depend on the location of the nodes in real space, and the nodes in the network will have a guaranteed number of "neighbours", that in itself ensures the reliability of the network. The presence of symmetry in a discrete network space (built out from the polyominos) allows one to select a fundamental area in it, the transmission of which completely defines the space itself, which can be used as an element of standardization in algorithms [14].

## 2 Routing in a network with clusters based on symmetric partitioning principles

Let's consider, that the clusterization of the network on the principles of symmetric partitioning has been carried out (Fig. 7), and the routing problem has been posed for a finite number of nodes (in order to reduce calculations).

It is necessary to determine the systems properties which can be implemented by various algorithms for finding optimal routes in order to estimate the probability of channeling the data packet from one arbitrary node to another. By analogy with the properties of an intellectual system, we can distinguish the following properties:

1. **“The Faith”** — characterizes the purposefulness of the system (confidence in the ability to reach a certain node). Grades of Faith are:
  - *no faith* — no data packet is being sent;
  - *absolute faith* — passing of a data packet through a node is possible only in one direction and is impossible in the opposite direction;
  - *doubt* — passing of a data packet is possible in the given and in the opposite direction.

2. **“The Memory”** — describes the ability of the system to store information about various routes and, if necessary, to remove the inefficient routes.

3. **“The Knowledge”** — describes the ability of the system to predict and choose the shortest route.

For the network shown in Figure 5, the probability of reaching a data packet from any arbitrary node (for example, 8) to any other arbitrary (for example, 16) can be estimated.

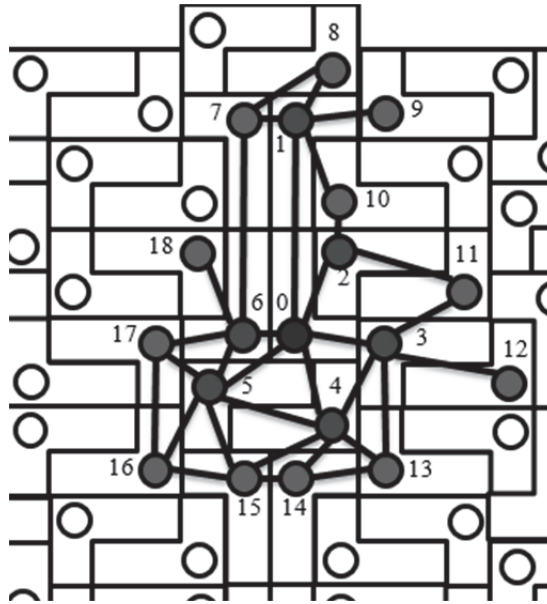


Figure 5 – Clustering network with possible routes [14]

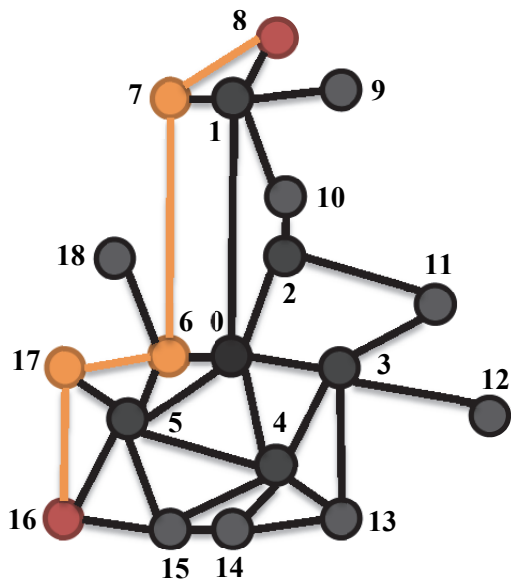
Table 1 – Combinations of system properties and routing efficiency

| No | Conditions                                     | Result   |
|----|--|--|
| 1. | No faith.<br>No memory.<br>No knowledge.       | No data packet is transmitted. The system has no purpose. $P = 0$  |
| 2. | Absolute faith.<br>No memory.<br>No knowledge. | There is a chance to get into the "dead-end" nodes, to which only one communication channel leads (9, 12, 18). The probability of reachability can be calculated through the probability of a reverse event: to get to the "dead end" node. $P = 0.82$ . |
| 3. | Doubt.<br>No memory.<br>No knowledge.          | Data transfer may loop, and the data may not reach the specified node. Possible cycles (1-10-2-0-1), (2-11-3-0-2), (8-1-0-6-7-8), etc. $P = 0.5$ .   |
| 4. | Doubt.<br>Memory.<br>No knowledge.             | The data will be delivered to any node with a probability of $P = 1$ , however, the delivery time without knowing the optimal routes can be long.  |
| 5. | Doubt.<br>Memory.<br>Knowledge.                | The route is selected following the optimal criterion: the shortest in the number of intermediary nodes or the buffering time. $P = 1$ , the gain in time in comparison with the previous condition.   |

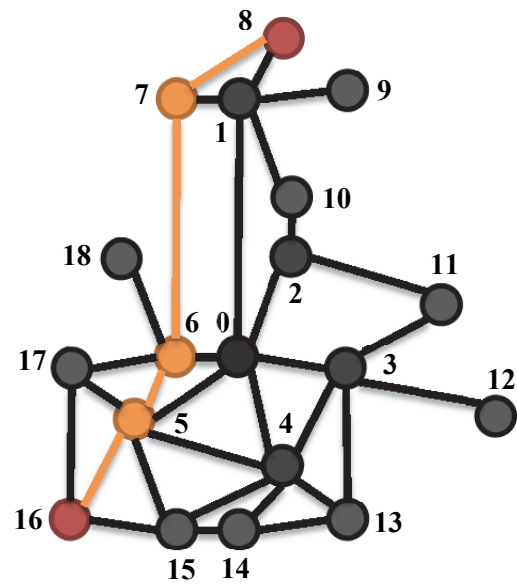
We give a calculation of the probability for the case: “Absolute faith. No memory. No knowledge.” It is necessary to enumerate all possible routes from node 8 to nodes 9, 12 and 18, which are considered to be “dead ends”, then find the probability of passing a packet of data from node 8 to node 16 as well as the probability of a reverse event of falling to a “dead end”.  $P(\text{hit from 8 to 16 nodes}) = 1 - 0.178 = 0.822$ . Determining the shortest paths with the smallest number of intermediary nodes allows one to concentrate on calculating the probabilities for the following 6 routes (Fig. 8).

The pre-condition providing the minimum number of intermediary nodes was chosen as a criterion for obtaining the optimal of all possible routes (therefore, only routes with 3 and 4 mediators were considered in this calculation), the probability of choosing this route, determined by the degree of vertices in the network graph (probability from 1 node to get to 0 –  $1/3$ , and from 0 to 6 node –  $1/4$ ) was taken into consideration. The result of multiplying of the probabilities at all nodes that fall in the route gives the total probability of choosing this route and can signify the measure of its effectiveness (priority in the process of choice when the system possesses “the Knowledge”).

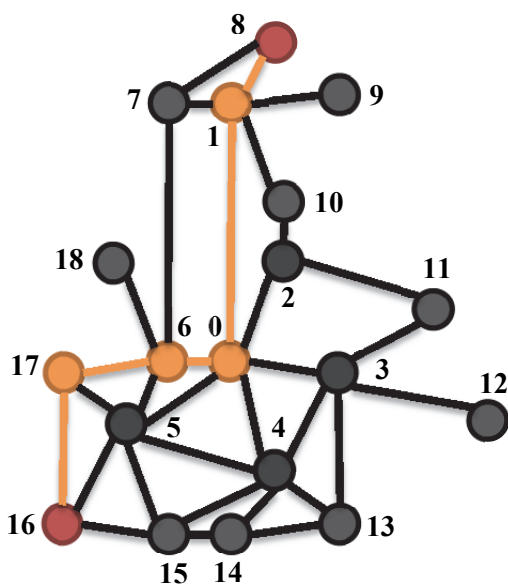




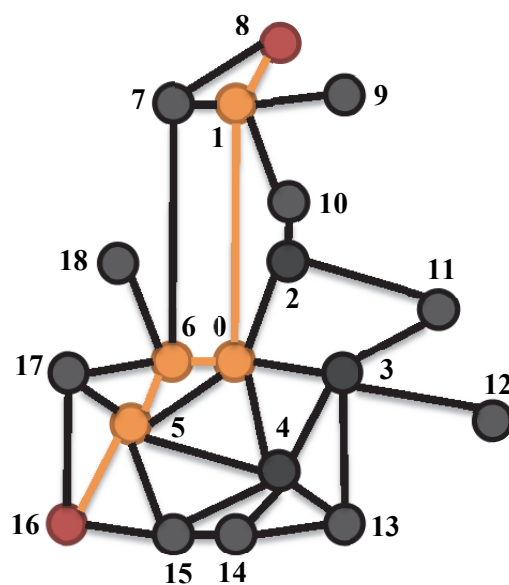
a)



b)



c)



d)

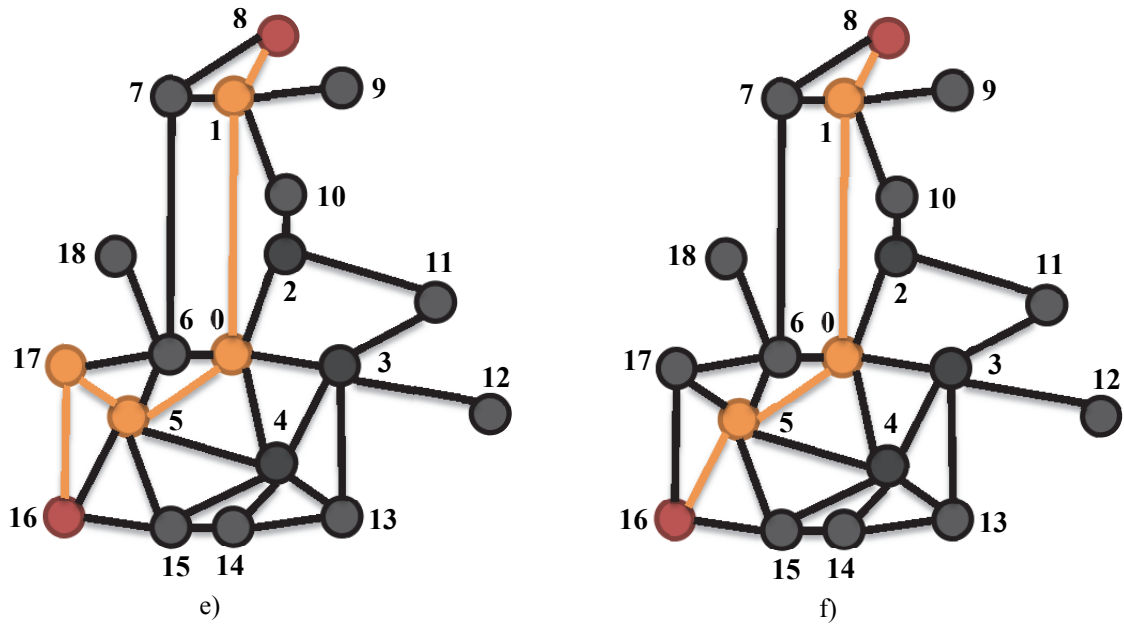


Figure 6 – Routing Options 8-16

Table 2 – Calculation of the probability for each of the routes

| No | Moving around nodes (route) |   |   |    |    |    | Probability of occurrence |
|----|-----------------------------|---|---|----|----|----|---------------------------|
| a  | 8                           | 7 | 6 | 17 | 16 |    | 0,0313                    |
| b  |                             |   |   | 5  | 16 |    | 0,0125                    |
| c  | 8                           | 1 | 0 | 6  | 17 | 16 | 0,0125                    |
| d  |                             |   |   |    | 5  | 16 | 0,0013                    |
| e  |                             |   |   | 5  | 17 | 16 | 0,0025                    |
| f  |                             |   |   |    | 16 |    | 0,0050                    |

It can be seen from the data in the table 2 above, that route (a) contains the minimum number of choices, therefore passing of the data packet along it will be characterized by the highest speed (the time for buffering is reduced), and the probability of reaching a given node along this route is the maximum one. But if the system possesses “the Knowledge”, that is, the complete information about the logical structure of communication nodes, a quick calculation of the shortest route (not occupied and not containing nodes that have failed) allows the system to deliver data to the destination node rapidly. This confirms the increase in network efficiency due to the introduction of “intellectual” properties to the system through appropriate algorithms.

### 3 Assembling information in the server

Let us return to the analysis of the network graph presented in Figure 7. If we consider the network topology, we can see that node number 0 is located in the center and has indirect or direct communication with all nodes of the system, and therefore we'll call it server node and set the task of collecting data from all system nodes into it. Excluding the cycles (for example 1-10-2-0-1)), we define a directed graph of the network. Information may be collected continuously or according to a schedule set by the user.

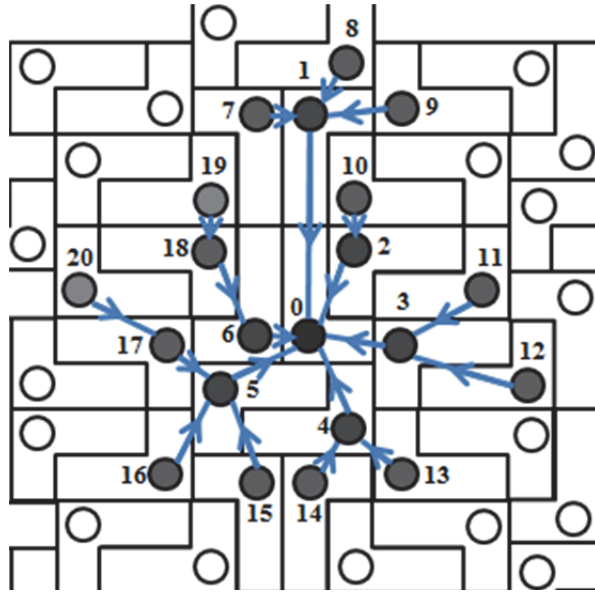


Figure 7 – System of network node connections, oriented neighborhood graph

As it was already shown by the authors in [14], for a more visual representation of such a sensor network, one should consider its graph (Fig. 9). The task was to determine the number of steps necessary for the data located in the side nodes to fall into the server node. For the mathematical formalization of this process, it was proposed to consider permutations along the graph.

Initially, there is a  $g[0]$  permutation (substitution, because the numbers are repeated) describing the relationship of all nodes:  $g[0] = (0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 10\ 11\ 12\ 13\ 14\ 15\ 16\ 17\ 18\ 19\ 20)$ .

By transferring data from node to node, one can get the following substitution  $g[1]$ :

$$g[1] = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 & 20 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 2 & 3 & 3 & 4 & 4 & 5 & 5 & 5 & 6 & 18 & 17 & \end{pmatrix}$$

$$g[1] = (0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 2\ 3\ 3\ 4\ 4\ 5\ 5\ 5\ 6\ 18\ 17)$$

All kinds of substitutions on the oriented graph (figure 9) form a closed set [14]:

$$g[0] = (0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 10\ 11\ 12\ 13\ 14\ 15\ 16\ 17\ 18\ 19\ 20);$$

$$g[1] = (0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 2\ 3\ 3\ 4\ 4\ 5\ 5\ 5\ 6\ 18\ 17)$$

$$g[2] = (0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 6\ 5)$$

$$g[3] = (0\ 0)$$

The substitution multiplication table for the graph shown in Figure 9:

$$\begin{matrix} g[0] & g[1] & g[2] & g[3] \\ g[1] & g[2] & g[3] & g[3] \\ g[2] & g[3] & g[3] & g[3] \\ g[3] & g[3] & g[3] & g[3] \end{matrix}$$

It can be concluded from the above-mentioned version of substitutions and the presented table, that the complete transfer of information from all nodes of this network will take place already in the third step of routing. This can be used to algorithmize the process of collecting data into the server, as well as estimating the number of steps necessary for the system to be able to accumulate data into a common center for further processing and prediction of the systems operational behavior. It should be noted that the substitutions constitute the so-called broken symmetry group [15].

The presented ideas can be used for the automatization of the procedure of collecting biometric data of any given person or to predict the weather in a network of sensor nodes distributed in a certain area.

## Conclusion

This article discusses ways of structuring networks using clusterization, a method of clusterization based on translational partitions of space (flat surface) is proposed, allowing the increase of the efficiency of the network with regard to the failure of nodes. A comparison of networks' efficiency depending on the presence of properties of intellectual systems in them was carried out; the probabilities of data reaching the destination node for various combinations of properties were estimated. For example, the introduction of a route memory algorithm in the considered system allows increasing the probability of



reaching a destination node up to 82%. And if the algorithm for comparing routes in the network is implemented, then choosing the optimal route allows you to save transmission time (the difference between the transmission efficiency of neighboring routes is 2-3%). The possibility of automatization of the process of assembling the data to a server (central) node using the procedure of multiplying substitutions according to a given oriented graph is shown. In the demonstrated example, the number of intermediary nodes in the sensor network is calculated (it is 3) which is necessary for data transfer to the server.

### Acknowledgements

The authors would like to acknowledge financial support from the Russian Foundation for Basic Research under the grant No.18-07-00170.

### References

- [1] C. Elliott Building the quantum network // *N. J. Phys.* 4, 46 (2002).
- [2] Areeya Chantasri, Mollie E. Kimchi-Schwartz, Nicolas Roch, Irfan Siddiqi, and Andrew N. Jordan Quantum Trajectories and Their Statistics for Remotely Entangled Quantum Bits *Phys. Rev. X* 6, 041052 (2016).
- [3] V.I. Emelyanov, Yu.V. Vladimirova *Quantum mechanics. Bits and qubits: Tutorial*// Moscow: Faculty of Physics MSU, 2012. — 176 p. — ISBN 978-5-8279-0108-2.
- [4] A.E. Krasnok, A.I. Denisyuk, P.A. Belov, I.S. Maksymov, A.E. Miroshnichenko, K.R. Simovskii, Yu.S. Kivshar *Optical nanoantennas* // *Uspekhi Fizicheskikh Nauk* 183 (6) Pp.561-589 (2013)
- [5] Rob van Kranenburg. *The Internet of Things: A critique of ambient technology and the all-seeing network of RFID.* — Pijnacker: Telstar Media, 2008
- [6] Chong C.Y., Kumar S.P. *Sensor Networks: Evolution, Opportunities and Challenges.* Proceedings of the IEEE, vol. 91, issue 8, August 2003.
- [7] Christian Catalini. *How blockchain technology will impact the digital economy* // *Oxford Business Law Blog* — 2017.
- [8] Broido V.L., Iliina O.P. *Computing systems, networks and telecommunications: A textbook for universities* // Saint Petersburg: “Piter”, 2011. — 560 p.: ill.— ISBN 978-5-49807-875-5
- [9] Goldstein B.S., Kutcheryavy A.E. *The Communication Networks Post-NGN* // Saint Petersburg, 2014. —160 p.: ill. — ISBN 978-5-9775-0900-8
- [10] Kutcheryavy A.E. *The self-organizing Networks/ Kutcheryavy A.E, Prokop'ev A.V., Kutcheryavy E.A.*// Saint Petersburg, „Lyubavitch“, 2011.
- [11] Gorshkov K.A., Nikitin O.R., Rau V.G., Rau T.F. *Hierarchical networks in the model of a discrete network's space*// *News of the Higher Educational Institutes. Textile Industry.* – Ivanovo – 2015 - №4(385) – Pp. 141-145.
- [12] Maleev A.V., Rau V.G., Potekhin K.A., Parkhomov L.G., Rau T.F. *Et al. Methods of the Discreet Modelling of Packages in Molecular Crystals*// *Reports of the USSR' Academy of Science* — Vol. 315. — 1990. — Pp. 1382-1385.
- [13] Zhuravlev V.G. *The Self-similar Growth of the Periodical Divisions and Graphs*// *Algebra and Analysis.* — №13. — 2001. — Pp. 69-92.
- [14] O. Nikitin, K. Gorshkov, V. Rau, E. Kuznetsova, A. Shubin. *Networks based on periodical fragmentations in broken symmetry group model* // 14th International scientific-technical conference on actual problems of electronic instrument engineering (APEIE) – 44894 Proceedings – Novosibirsk , 2018 – Volume 1, Part 3, Pp.247-250.
- [15] Rau V.G., Lomtev L.A., Rau T.F., Gorshkov K.A., Nikitin O.R. *Computer experiments in groups of permutations with broken symmetry* // *Modern high technologies.* – 2017. – № 3. – Pp. 43-49