«Компьютерные науки и информатика»

Научная статья

УДК 004.94
DOI: 10.17072/1993-0550-2024-2-54-60

Моделирование распространения сетевого вируса в локальной компьютерной сети методами теории перколяции

Мария Михайловна Бузмакова[1], Егор Александрович Воробьёв[2]
[1,2] Пермский государственный национальный исследовательский университет, г. Пермь, Россия
[1]mbuzmakova@psu.ru
[2]teddehhh.study@gmail.com

Аннотация. В рамках работы исследовано распространение сетевого вируса в локальной компьютерной сети. Были предложены две перколяционные модели, описывающие два вида сетей: проводные и беспроводные. Порог перколяции соответствует доле зараженных компьютеров в сети, при которой сеть теряет работоспособность. Для моделей были разработаны и реализованы алгоритмы заполнения решетки занятыми узлами, распределения занятых узлов по кластерам, поиска перколяционного кластера, определения порога перколяции. Был проведен численный эксперимент по оценке порога перколяции и его зависимость от различных характеристик вируса.

Ключевые слова: *теория перколяции; компьютерный вирус; сетевой вирус; компьютерное моделирование*

Для цитирования: *Бузмакова М.М., Воробьёв Е.А.* Перколяционная модель распространения сетевого вируса // Вестник Пермского университета. Математика. Механика. Информатика. 2024. Вып. 2(65). С. 54–60. DOI: 10.17072/1993-0550-2024-2-54-60.

Статья поступила в редакцию 25.04.2024; одобрена после рецензирования 13.05.2024; принята к публикации 15.05.2024.

«Computer Science»

Research article

A Net Virus Spreading in a Local Computer Network Modeling With Using Percolation Theory Methods

Mariya M. Buzmakova[1], Egor A. Vorobiev[2]
[1,2] Perm State University, Perm, Russia
[1]mbuzmakova@psu.ru
[2]teddehhh.study@gmail.com

Abstract. A net virus spreading in a local computer network is investigated in this paper. Two percolation models were proposed to describe two types of networks: wired and wireless. The percolation threshold corresponds to the fraction of infected computers in the network at which the network loses its operability. Algorithms for lattice filling by occupied nodes, for distributing occupied nodes into clusters, for searching a percolation cluster and for the percolation threshold determining were developed and implemented for the models. A numerical experiment was conducted to estimate the percolation threshold and its dependence on various virus characteristics.

Keywords: *percolation theory; computer virus; net virus; computer modeling*

For citation: Buzmakova, M. M. and Vorobiev, E. A. (2024), "A Net Virus Spreading in a Local Computer Network Modeling With Using Percolation Theory Methods", *Bulletin of Perm University. Mathematics. Mechanics. Computer Science*, no. 2(65), pp. 54-60. DOI: 10.17072/1993-0550-2024-2-54-60.

The article was submitted 25.04.2024; approved after reviewing 13.05.2024; accepted for publication 15.05.2024.

Введение

Стремительное развитие информационных технологий сопровождается постоянным обменом информацией, что, в свою очередь, сопровождается и распространением компьютерных вирусов. Несмотря на богатство антивирусных программ, каждый день возникают все новые и новые угрозы. Актуальным является изучение распространения вирусов в компьютерных сетях с целью его своевременного обнаружения, предотвращения и дальнейшей защиты.

Начало всех исследований в области компьютерной вирусологии было положено в 1940-х гг., когда Джон фон Нейман поставил перед собой задачу построения модели машины, сложность которой могла бы возрастать подобно биологическим организмам в условиях естественного отбора. Предположительно, такая задача была поставлена для того, чтобы вычислительные машины смогли развиваться самостоятельно с течением времени. Но на тот момент не было реализации подобной модели, поэтому в данном направлении проводились дополнительные исследования.

Первая попытка реализации была совершена Лайонелом Пенроузом в 1957 г. Английский психиатр, медицинский генетик и математик впервые показал пример самовоспроизводящейся механической структуры [1]. Логическая часть его идеи была основана теории фон Неймана, а физическая часть заключалась в том, чтобы создать простые блоки или кирпичики с такими свойствами, чтобы из них можно было построить самовоспроизводящийся механизм.

Спустя несколько лет в 1966 г. Артур Беркс опубликовал книгу "Теория самовоспроизводящихся автоматов" на основе лекций фон Неймана [2], в которой были изложены основные идеи для реализации такого механизма в машинной среде.

В настоящее время компьютерный вирус определяется как программа, целью которой является распространение своих копий. Компьютерные вирусы классифицируются по признакам среды обитания, способу заражения, методам распространения, организации программного кода и деструктивным возможностям [3].

Развитие информационных сетей, объединяющих несколько рядом стоящих компьютеров, способствовало появлению вируса, представляющего угрозу компьютерной инфраструктуре, в которой находятся десятки или даже сотни работающих вычислительных устройств. В отличие от загрузочных и файловых вирусов, сетевой вид обладает свойством самораспространения без использования внешних устройств передачи данных. Такой вид относят к классу вирусов-червей. Проводятся исследования по изучения распространения вирусов [например, 4–9]. Разными авторами предложены оригинальные и модифицированные эпидемиологические модели. При анализе работ в данном направлении наблюдается тенденция к востребованности методов теории перколяции. Использование перколяционных свойств для большого разнообразия существующих архитектур компьютерных сетей и методов атак злоумышленников может способствовать совершенствованию методов антивирусной защиты от нападения, вследствие чего снизится количество успешных атак на информационную инфраструктуру.

Целью настоящей работы является исследование распространения сетевого вируса в локальной сети с использованием подходов теории перколяции. Для достижения цели авторами были предложены и исследованы две перколяционные модели распространения сетевого вируса в локальной компьютерной сети, описывающие два вида сетей: проводные и беспроводные. Порог перколяции соответствует доле зараженных компьютеров в сети, при которой сеть теряет работоспособность.

Постановка задачи

В рамках данной работы предложены и исследованы две решеточные перколяционные модели распространения сетевого вируса в локальной компьютерной сети, описывающие два вида сетей: беспроводные и проводные.

В первой модели беспроводная локальная сеть представлена в виде наиболее распростра-

ненной сети ячеистой топологии, которая описывается простой квадратной решеткой.

Линейный размер решетки – N. Компьютеры в локальной сети – узлы решетки, которые могут быть свободными (не зараженными вирусом) и занятыми (зараженными вирусом). Свободным узлом считается компьютер, сохраняющий полную работоспособность и никак не влияющий на работу соседних компьютеров.

Под занятым узлом подразумевается, что компьютер имеет вредоносный код, способный причинить вред компьютеру, а также передавать свою копию соседним компьютерам. При этом такой компьютер продолжает работать. Один компьютер может заразить другой компьютер, то есть если текущий узел занят, то соседние с ним узлы могут стать занятыми с вероятностью q_1. Кроме того, учитывается способность компьютера к восстановлению, то есть занятый узел может стать свободным с вероятностью q_2. Восстановление возвращает узлу свободное состояние и освобождает компьютер от вредоносного кода.

Соседние занятые узлы образуют кластеры – группы зараженных компьютеров. Концентрация занятых узлов p соответствует степени распространения вируса в локальной сети. Если в системе находится непрерывная группа зараженных компьютеров, проходящих через всю сеть (перколяционный кластер), то можно говорить о наличии перколяции – просачивании сетевого вируса через всю локальную сеть, что означает возможность выхода вируса за пределы локальной сети и блокировку передачи информации между любыми свободными узлами. В модели были применены открытые и периодические граничные условия (ОГУ и ПГУ соответственно). Математически предложенную модель можно описать так:

$$M_1 = < N, p, q_1, q_2, k >,$$

где k – количество испытаний.

Во второй модели применяется локальная сеть, использующая смешанную топологию из трех базовых: "Шина", "Кольцо", "Звезда". Общей сетью является шинная топология, объединяющая в себе множество N подсетей, представленных кольцевой и звездной топологиями. В свою очередь каждая подсеть состоит из z узлов. Граничащие между топологиями узлы являются коммутаторами и обеспечивают связь с соседними локальными сетями. Начало и конец сети определяется первой и последней подсетью соответственно (см. рис. 1).

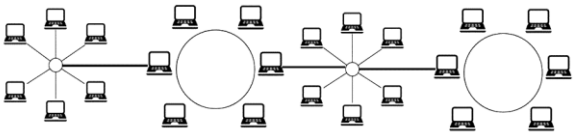

Рис. 1. *Пример смешанной сети второй модели*

Математически модель записывается следующим образом:

$$M_2 = < N, z, p, k >.$$

Основной задачей в рамках каждой из предложенной модели является получение оценки значения порога перколяции при различных параметрах модели. Под порогом перколяции принимается значение концентрации занятых узлов, при которой вероятность возникновения стягивающего кластера равна 0.5. Порог перколяции соответствует критической концентрации зараженных компьютеров локальной сети, при которой сеть теряет свою работоспособность.

Методы исследования

Для моделей были разработаны и реализованы алгоритмы распространения занятых узлов, поиска кластеров и определения наличия перколяции в системе. Были написаны программы на языке программирования C++ с консольным интерфейсом.

Для первой модели алгоритм заполнения квадратной решетки свободными и занятыми узлами описан ниже.

1. Из всех узлов случайно выбирается первый занятый узел.
2. Соседние свободные узлы текущих занятых становятся занятыми с вероятностью q_1.
3. Занятые узлы могут стать свободными с вероятностью q_2.
4. Повторить 2–3 до достижения необходимой концентрации занятых узлов на решетке.

Для маркировки кластеров на решетке используется алгоритм Хошена–Копельмана [10]. Для поиска перколяционного кластера используется алгоритм "поиска в глубину" [11]. Для обеспечения коэффициентов вероятности используется генератор псевдослучайных чисел "Вихрь Мерсена" [12].

Для второй модели разработан следующий алгоритм определения порога перколяции проводилось путем проведения k случайных экспериментов, в каждом из которых:

1) случайным образом конфигурируется общая сеть, состоящая из N подсетей,
2) для каждой подсети находится свой порог перколяции,
3) находится среднее значение порога перколяции для общей сети.

После чего полученные значения по k экспериментам усредняются.

Порог перколяции для подсетей кольцевой или звездной топологии находятся аналитически. Рассмотрим кольцевую подсеть. Подсеть имеет граничные узлы, которые обеспечивают внешнюю связь с соседними подсетями. Однако внутри возникает два пути распространения вируса: верхняя и нижняя цепочка (см. рис. 1). Данные цепочки являются одномерными, порог перколяции для каждого пути p_c = 1, так как кластер из занятых узлов может возникнуть, только если все узлы хотя бы в одной из цепочек заняты [13]. Таким образом, подсеть имеет точное значение порога перколяции p_c = 1.

Звездная топология имеет похожую структуру с деревом Кейли или решеткой Бете. Подсеть представляет первый уровень построения дерева Кейли, когда из одного узла выходит z новых узлов. Для такой структуры вычислена и доказана критическая концентрация p_c=1/(z-1) в [14]. Подставляя значение количества выходящих узлов в данную формулу, можно точно определить критическую концентрацию в подсети.

Далее, чтобы найти порог перколяции для общей сети, необходимо сложить критические концентрации p_{1c}, p_{2c}, …, p_{Nc} и разделить на количество подсетей N.

Численные эксперименты основаны на подходах методов Монте-Карло с применением методов математической статистики и теории вероятностей.

Результаты и их обсуждение

Для первой модели проведен ряд численных экспериментов со следующими параметрами: N = 10; 20; 50; $p = 0; 0{,}1; \ldots; 1$; $q_1 = 0{,}5$; $q_2 = 0; 0{,}1; \ldots; 0{,}6$. Значение q_1 выбрано, исходя из неопределенного количества внешних факторов, а также человеческого фактора. Поэтому вероятность заразить или не заразить компьютер является одинаковой. Параметр q_2 имеет диапазон значений: от низкой эффективности каких-либо противодействующих средств либо их отсутствия до наличия эффективных средств защиты компьютера.

Для каждого набора входных данных было проведено 1000 экспериментов и определено значение порога перколяции по следующей методике: определяется вероятность возникновения перколяционного кластера $P(p)$.

Полученные в ходе компьютерного эксперимента значения вероятности возникновения перколяционного кластера $P(p)$ для каждого набора данных при различных значениях N и q_2 аппроксимируются сигмоидальной функцией:

$$P(p) = (1 + exp\,(-(p - p_c)a)\,)^{-1}.$$

Значение доли заполненных узлов p, при которой вероятность появления перколяционного кластера равна 0.5, является значением порога перколяции (например, рис. 2).

При аппроксимации данных возникает погрешность: при численном эксперименте учитывается погрешность вероятности возникновения перколяционного кластера при каждом значении концентрации занятых узлов с использованием стандартного отклонения среднего $\sigma_{\bar{P}} = \sqrt{\frac{1}{k-1}\sum_{i=1}^{k}(P_i - \bar{P})^2 / k}$.

Погрешность значения порога перколяции – это результат аппроксимации данных в математическом пакете с учетом ошибок исходных данных.

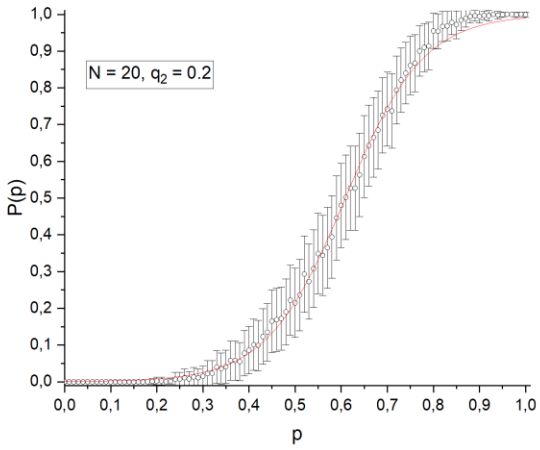

Рис. 2. *Вероятность возникновения стягивающего кластера, при N = 20; q_1 =0,5; $q_2 = 0{,}2$ с открытыми граничными условиями*

Найдены значения порогов перколяции для выбранных входных параметров модели при открытых и периодических граничных условиях для систем конечного размера, данные представлены в табл. 1. Пустые ячейки означают, что для соответствующих входных параметров перколяция не наступила.

Таблица 1. *Значения порога перколяции с погрешностью аппроксимации при различных N и q_2 с открытыми и периодическими граничными условиями*

0,6	0,5	0,4	0,3	0,2	0,1	0	q_2
–	–	–	0,789 (0,001)	0,6332 (0,0009)	0,5356 (0,0009)	0,4735 (0,0008)	N = 10 ОГУ
–	–	0,876 (0,001)	0,679 (0,001)	0,5453 (0,0008)	0,4608 (0,0006)	0,4052 (0,0006)	N = 10 ПГУ
–	–	–	0,7922 (0,0009)	0,6089 (0,0009)	0,5056 (0,0007)	0,4472 (0,0006)	N = 20 ОГУ
–	–	0,9885 (0,0009)	0,747 (0,001)	0,5561 (0,0008)	0,4545 (0,0005)	0,4000 (0,0005)	N = 20 ПГУ
–	–	–	0,8037 (0,0007)	0,587 (0,001)	0,4591 (0,006)	0,4108 (0,0006)	N = 50 ОГУ
–	–	–	0,7935 (0,0005)	0,5602 (0,0007)	0,4350 (0,0006)	0,3910 (0,0004)	N = 50 ПГУ

При увеличении значения параметра q_2 вероятность появления перколяционного кластера значительно снижается. При этом наблюдается разделение значений q_2 на две группы. Первая группа достигает значения порога перколяции и представляет достаточную опасность для компьютерной сети. В нее входят модели при $q_2 = 0; 0{,}1; 0{,}2; 0{,}3$. Вторая группа содержит модели при $q_2 = 0{,}4; 0{,}5; 0{,}6$, которые не так опасны для компьютерной сети. При максимальной концентрации узлов вероятность появления перколяционного кластера находится ниже значения 0,5. Стоит отметить, что модель с параметром $q_2 > q_1$ на всем диапазоне концентрации p имеет вероятность $P(p)$ близкую к нулю, что является адекватным результатом, так как восстановление узлов происходит интенсивнее, чем заражение.

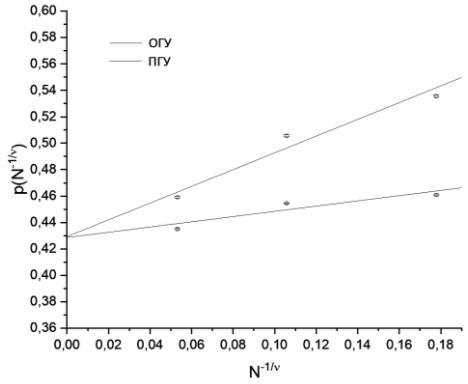

Рис. 3. *Определение порога перколяции для случая бесконечной системы с помощью скейлинга при q_2=0,1*

Для бесконечного случая определены значения порога перколяции с помощью скейлингового соотношения (например, рис. 3), результаты для открытых и периодических граничных условий представлены в табл. 2. Близкие или схожие значения для разных граничных условий говорят о правильности полученных результатов, так как порог перколяции для бесконечной перколяционной системы не должен зависеть от этого параметра.

Таблица 2. *Значения порогов перколяции для случая бесконечной системы*

q_2	Scaling, ОГУ	Scaling, ПГУ
0	0,387 (0,11)	0,385 (0,03)
0,1	0,429 (0,15)	0,429 (0,10)
0,2	0,569 (0,003)	0,567 (0,002)
0,3	0,809 (0,005)	0,828 (0,005)
0,4	–	–
0,5	–	–
0,6	–	–

Для второй модели проведен ряд численных экспериментов со следующими параметрами: N=5; 10; 20; 50; 100; 200; Z=0...100; k = 1000. Оба параметра представляют собой расширенный диапазон размерностей общей сети и подсети: от малых размеров до больших. Значения порога перколяции представлены на рис. 3.

Полученные результаты показывают, что при увеличении количества узлов в подсетях порог перколяции снижается достаточно быстро и постепенно приближается к значению 0.5.

С увеличением значения параметра *N* кривые на рисунке становятся сглаженнее и описывают практически одинаковы значения, что говорит о том, что количество узлов в подсетях не является определяющим при большом количестве испытаний.

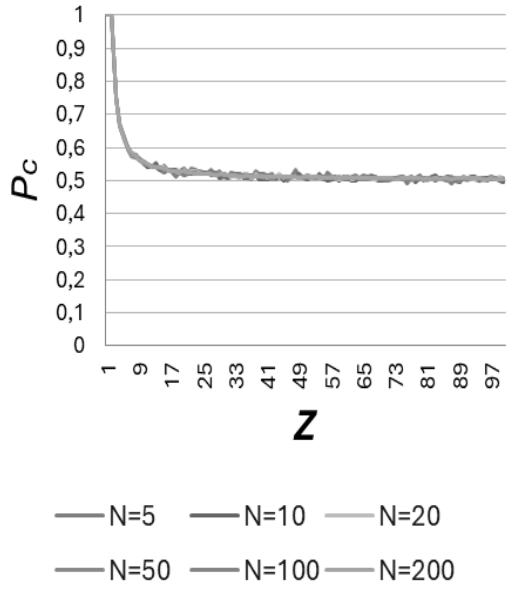

Рис. 3. *Сравнительный график значений порога перколяции между различными N*

Исходя из вышеописанного, значение p_c = 0.5 можно взять за значение порога перколяции для бесконечного случая.

Заключение

В рамках данной работы проведено моделирование распространения сетевого вируса в локальной компьютерной сети с применением теории перколяции. Были разработаны две модели: модель поведения сетевого вируса в беспроводной локальной сети с учетом способности вычислительного узла к восстановлению; модель, использующая смешанную топологию локальной сети.

Результаты исследования первой модели показывают, что при отсутствии возможности восстановления зараженных узлов компьютерная сеть довольно быстро подвергается заражению и теряет свою работоспособность. Однако, если компьютеры имеют механизм восстановления после заражения, то риск полного заражения локальной сети значительно снижается.

По результатам второй модели был определен порог перколяции p_c = 0.5 для бесконечной системы.

Список источников

1. *Penrose S.* Se1f-reproducing machines // Scientific American. 1959. Vol. 200. P. 105–114.
2. *Von Neumann's* self-reproducing automata / Burks A.W. // THE UNIVERSITY OF MICHIGAN, 1969. 113p.
3. *Компьютерные* вирусы и антивирусы: взгляд программиста / Климентьев К.Е. // М.: ДМК Пресс, 2013. 656 с.
4. *Минаев В.А., Сычев М.П., Вайц Е.В., Киракосян А.Э.* Имитационное моделирование эпидемий компьютерных вирусов // Вестник Российского нового университета. Серия "Сложные системы... ". 2019. № 3. С. 3–12.
5. *Семёнов С.Г., Давыдов В.В.* Математическая модель распространения компьютерных вирусов в гетерогенных компьютерных сетях автоматизированных систем управления технологическим процессом // Вестник НТУ "ХПИ". 2013. № 38. С. 163–171.
6. *Гусаров А.Н., Жуков Д.О., Косарева А.В.* Описание динамики распространения компьютерных угроз в информационно-вычислительных сетях с запаздыванием действия антивирусов // Вестник МГТУ им. Н.Э. Баумана. Сер. "Приборостроение". 2010. № 1. С. 112–120.
7. *Лесько С.А., Алёшкин А.С., Филатов В.В.* Стохастические и перколяционные модели динамики блокировки вычислительных сетей при распространении эпидемий эволюционирующих компьютерных вирусов // Российский технологический журнал. 2019. Т. 7, № 3. С. 7–27.
8. *Moore C.* and Newman M. E. J. Epidemics and percolation in small-world networks // Phys. Rev. E. 2000. № 61. P. 5678.
9. *Michele Garetto,* Weibo Gong and Don Towsley, Modeling Malware Spreading Dynamics // Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies. 2003. Vol. 3. P. 1869–1879.
10. *Hoshen J.,* and Kopelman R. Percolation and cluster distribution: I. Cluster multiple labeling technique and critical concentration algorithm // Phys. Rev. B. 1976. I. 14 (October). P. 3438–3445.

11. URL: https://ru.wikipedia.org/wiki/Поиск_в_глубину (дата обращения: 20.04.2024).
12. *M. Matsumoto* and T. Nishimura Mersenne Twister: A 623-Dimensionally Equidistributed Uniform Pseudo-Random Number Generator // ACM Transactions on Modeling and Computer Simulation. 1998. Vol. 8, № 1. P. 3–30.
13. *Stauffer D.* Introduction to percolation theory. London: Taylor & Francis, 1985. 192 p.
14. *Тарасевич Ю.Ю.* Перколяция: теория, приложения, алгоритмы. М.: Едиториал УРСС, 2002. 112 с.

References

1. Penrose, S. (1959), "Se1f-reproducing machines", *Scientific American*, 1959, vol. 200, pp. 105-114.
2. Burks, A.W. (1969), *Von Neumann's self-reproducing automata*, THE UNIVERSITY OF MICHI-GAN, 113p.
3. Kliment'ev, K.E. (2013), *Komp'yuternye virusy i antivirusy: vzglyad programmista*, M.: DMK Press, 656 p.
4. Minaev, V.A., Sychev, M.P., Vajc, E.V., Kirakosyan, A.E'. (2019), "Imitacionnoe modelirovanie e'pidemij komp'yuternyx virusov", *Vestnik Rossijskogo novogo universiteta. Seriya "Slozhnye sistemy...* ", no. 3, pp. 3-12.
5. Semenov, S.G., Davydov, V.V. (2013), "Matematicheskaya model' rasprostraneniya komp'yuternyx virusov v geterogennyx komp'yuternyx setyax avtomatizirovannyx sistem upravleniya texnologicheskim processom", *Vestnik NTU "XPI*", no. 38, pp. 163-171.
6. Gusarov, A.N., Zhukov, D.O., Kosareva, A.V. (2010), "Opisanie dinamiki rasprostraneniya komp'yuternyx ugroz v informacionno-vychislitel'nyx setyax s zapazdyvaniem dejstviya antivirusov", *Vestnik MGTU im. N.E'. Baumana. Ser.* "*Priborostroenie*", no. 1, pp. 112-120.
7. Les'ko, S.A., Alyoshkin, A.S., Filatov, V.V. (2019), "Stoxasticheskie i perkolyacionnye modeli dinamiki blokirovki vychislitel'nyx setej pri rasprostranenii e'pidemij e'volyucioniruyushhix komp'yuternyx virusov", *Rossijskij texno-logicheskij zhurnal*, vol. 7, no. 3, pp. 7-27.
8. Moore, C. and Newman, M. E. J. (2000), "Epidemics and percolation in small-world networks", *Phys. Rev. E*, no. 61, pp. 5678.
9. Garetto, M., Weibo G. and Donald F. T. (2003), "Modeling malware spreading dynamics", *IEEE INFOCOM 2003. Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE Cat. No.03CH37428)*, no. 3, pp. 1869-1879.
10. Hoshen, J., and Kopelman, R. (1976), "Percolation and cluster distribution: I. Cluster multiple labeling technique and critical concentration algorithm", *Phys. Rev. B*, i. 14 (October), pp. 3438-3445.
11. https://ru.wikipedia.org/wiki/Poisk_v_glubinu (accessed data: 20.04.2024).
12. Matsumoto, M. and Nishimura, T. (1998), "Mersenne Twister: A 623-Dimensionally Equidistributed Uniform Pseudo-Random Number Generator", *ACM Transactions on Modeling and Computer Simulation*, vol. 8, no. 1, pp. 3-30.
13. Stauffer, D. (1985), *Introduction to percolation theory*, London: Taylor & Francis, 192 p.
14. Tarasevich, Yu. Yu. (2002), *Perkolyaciya: teoriya, prilozheniya, algoritmy*, M.: Editorial URSS, 112 s.

Информация об авторах:

М. М. Бузмакова – кандидат физико-математических наук, доцент, доцент кафедры прикладной математики и информатики физико-математического института Пермского государственного национального исследовательского университета (614068, Россия, г. Пермь, ул. Букирева, 15), AuthorID: 642701;

Е. А. Воробьёв – магистр первого года обучения по направлению "Прикладная математика и информатика" физико-математического института Пермского государственного национального исследовательского университета (614068, Россия, г. Пермь, ул. Букирева, 15).

Information about the authors:

Maria M. Buzmakova – Candidate of Physical and Mathematical Sciences, Associate Professor, Associate Professor of the Applied Mathematics and Informatics Department, Institute of Physics and Mathematics, Perm State University (15, Bukireva St., Perm, Russia, 614068), AuthorID: 642701;

Egor A. Vorobyev – first-year Master of Applied Mathematics and Informatics at the Physics and Mathematics Institute of Perm State University (15, Bukireva St., Perm, Russia, 614068).