

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«Московский государственный университет имени М. В.
Ломоносова»

Факультет вычислительной математики и кибернетики

Кафедра исследование операций

Оценка эффективности модельной сети воздушного
транспорта России с помощью средневзвешенного
коэффициента кластеризации

Курсовая работа

Обучающийся

М.А.Класон

Руководитель

Н.М.Новикова

Москва 2022

Содержание

| | | |
|---|---|----|
| 1 | Введение | 3 |
| 2 | Средний коэффициент кластеризации для некоторых простейших сетей | 7 |
| 3 | Подсчет средневзвешенного коэффициента кластеризации для модели авиаперевозок внутри России | 9 |
| 4 | Заключение | 13 |
| 5 | Список литературы | 14 |

1 Введение

Целью данной работы является оценка модельной сети воздушных перевозок России, существующей на момент апреля 2022 года, с помощью средневзвешенного коэффициента кластеризации, применение которого описано в [1] для другой страны.

Сеть воздушных перевозок (СВП) состоит из отдельных аэропортов (городов) и регулярных рейсов между парами аэропортов. Для ее описания был введен граф $G(V, E)$, где множество узлов V представляет все N аэропортов, а множество ребер E представляет все m прямых маршрутов между аэропортами. Предполагается, что этот граф является взвешенным, и для его описания используется взвешенная матрица смежности $W = (w_{ij})_{i,j}$, где $w_{i,j}$ представляет собой общее количество регулярных рейсов на маршруте (i, j) в течении недели, и $w_{ij} = 0$, если нет регулярных рейсов между аэропортом i и аэропортом j . В частности, $w_{ii} = 0$ и $w_{ij} = w_{ji}$, так как предполагается, что СВП симметрична. Также обозначим $A = (a_{ij})_{i,j}$ невзвешенную матрицу смежности, где $a_{i,j}$ определяется просто как 1, если $w_{i,j} > 0$, и 0 в противном случае. Наконец, d_i представляет собой степень узла i , то есть число его прямых соседей, а t_i - это количество ребер среди его прямых соседей, т.е

$$t_i = \frac{1}{2} \sum_{j,h} a_{ij} a_{ih} a_{jh}.$$

Хорошая связность и устойчивость сети к отказам узлов и связей являются ключевыми факторами для транспортных сетей в целом и для СВП в частности.

Устойчивость - это способность обеспечивать и поддерживать приемлемый уровень работы сети при отказе ребра или вершины.

Для их измерения используются различные параметры:

Алгебраическая связность (например в [3], [4]), которая равна второму минимальному собственному значению матрицы Кирхгофа:

$$K = k_{ij} = \begin{cases} d_i & , i = j \\ -a_{ij} & , i \neq j \end{cases},$$

поскольку минимальное собственное значение этой матрицы всегда равно нулю, причем кратность нулевого собственного значения равна количеству компонентов связности в графе. При этом учитываются кратные собственные значения, то есть если кратность нулевого собственного значения больше одного, то алгебраическая связность этой сети будет равна нулю. Алгебраическая связность отражает, насколько хорошо связан граф.

Степень центральности (например в [5]), которая равна:

$$C_D(G) = \frac{\sum_{i=1}^N |\max_i d_i - d_i|}{H};$$

$$H = \max_G \sum_{i=1}^N |\max_i d_i - d_i|, \text{ } G\text{—связный граф состоящий из } N \text{ вершин.}$$

Причем H максимально в случае, когда одна вершина связана со всеми остальными, и остальные связаны только с ней, поскольку в связном графе из N вершин $1 \leq d_i \leq N - 1 \forall i \in (1, N)$, и, следовательно, в графе, где $\max_i d_i = N - 1$, а все остальные степени равны 1, сумма $\sum_{i=1}^N |\max_i d_i - d_i|$ будет максимальна. Для данного графа $H = (N - 1)((N - 1) - 1) = N^2 - 3N + 2$, то есть степень центральности любого связного графа равна:

$$C_D(G) = \frac{\sum_{i=1}^N |\max_i d_i - d_i|}{N^2 - 3N + 2}.$$

Очевидно, что $C_D(G) \in [0, 1]$, и, как правило, чем выше степень центральности графа, тем выше его связность.

Данные параметры не подходят для сравнения эффективности модельной сети воздушного транспорта России, поскольку, например, не подходят для несвязных графов, а при сравнении сетей, действующих на одинаковом наборе городов, могут возникнуть ситуации, в которых в одной сети будут существовать ребра из города N , а в другой - не будут. И тогда ни алгебраическая связность, ни степень центральности не подойдут для второй сети.

В [1] предлагается рассмотреть средний коэффициент кластеризации (СКК) для измерения связности и устойчивости сети к отказам вершин и ребер. В СВП устойчивость выше всех других факторов: когда маршрут полета перестает работать, компании хотят направить пассажиров с наименьшим количеством возможных пересадок. Именно это и определяет взвешенный коэффициент кластеризации: он показывает, какую часть пассажиров можно пересадить на другие стороны каждого треугольника, состоящего из трех, соединенных друг с другом, городов, и поэтому СКК можно рассматривать, как способ измерения устойчивости к отказам для СВП.

В отличие от других показателей, которые используются для измерения устойчивости сети к отказам, СКК, как правило, имеет обратную зависимость со средним количеством пересадок в нерегулярном графе (нерегулярный граф - граф, не все степени вершин которого равны между собой), и, таким образом, гарантирует, что сеть является графом "Мир тесен", что, как было доказано в [2], важно для СВП.

Граф "Мир тесен" - граф, две произвольные вершины которого с большой вероятностью не будут являться смежными, однако одна будет достижима из другой посредством небольшого количества переходов через другие вершины.

Таким образом, в СВП, где среднее количество пересадок очень важно, и где устойчивость к отказам является необходимостью, не потребуется ничего делать для ограничения максимального количества пересадок. Это еще одна причина, по которой СКК был выбран для сравнения эффективности.

Коэффициент кластеризации (невзвешенный) узла i , c_i , определим, как:

$$c_i = \begin{cases} 0 & , d_i = 0 \\ 1 & , d_i = 1 . \\ \frac{2t_i}{d_i(d_i-1)} & , d_i \geq 2 \end{cases}$$

Данная формула демонстрирует, что СКК в первую очередь характеризует устойчивость к отказам узлов, а не ребер. В случае "тупикового" узла, при обрыве его единственного ребра надежность сети должна падать, так как из него не существует альтернативных маршрутов, но при этом коэффициент кластеризации этого узла равен одному, т.е. максимален. А значит рассматривается скорее ситуация, когда обрывается сам тупиковый узел, и тогда логично, что его отсутствие не влияет на другие маршруты.

Используя невзвешенную матрицу смежности A , представим c_i в виде:

$$c_i = \frac{1}{d_i(d_i - 1)} \sum_{j,h} a_{ij}a_{ih}a_{jh}, \quad d_i \geq 2. \quad (1)$$

Определим взвешенный коэффициент кластеризации узла i , c_i^w в виде:

$$c_i^w = \frac{1}{(d_i - 1) \sum_j w_{ij}} \sum_{j,h} \frac{w_{ij} + w_{ih}}{2} a_{ij}a_{ih}a_{jh}. \quad (2)$$

Для упрощения вычислений и используя определения матриц, определим c_i^w как:

$$c_i^w = \frac{(A^2W)_{i,i}}{(d_i - 1) \sum_j w_{ij}}.$$

Средневзвешенный коэффициент кластеризации $C^w(G)$ имеет вид:

$$C^w(G) = \frac{1}{N} \sum_i c_i^w.$$

Поскольку мы рассматриваем только графы заданного порядка N , мы можем использовать сокращенный средневзвешенный коэффициент кластеризации, $C_R^w(G)$, определяемый следующим образом:

$$C_R^w(G) = N \times C^w(G) = \sum_i c_i^w.$$

2 Средний коэффициент кластеризации для некоторых простейших сетей

Невзвешенный коэффициент кластеризации для ряда простых сетей по (1).

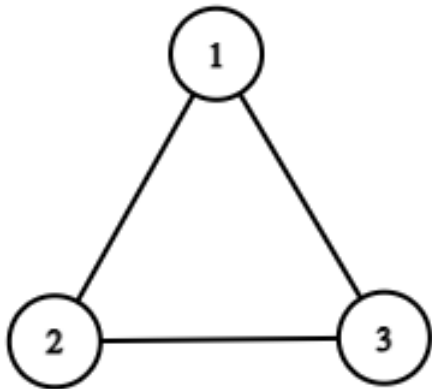


Рис. 1
 $N=3$, $m=3$;
 $c_1 = c_2 = c_3 = 1$;
 $C(G) = 1$;

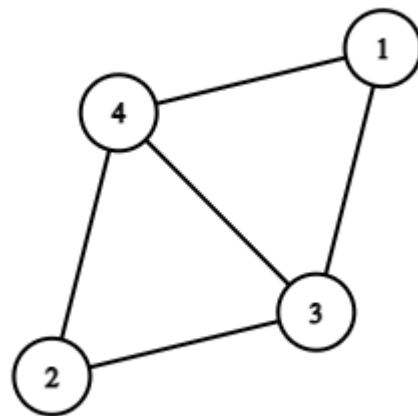


Рис. 3
 $N=4$, $m=5$;
 $c_1 = c_2 = 1$, $c_3 = c_4 = 0.67$;
 $C(G) = 0.835$;

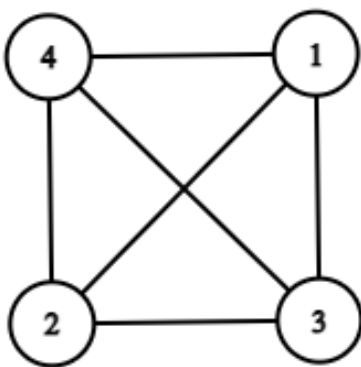


Рис. 2
 $N=4$, $m=6$;
 $c_1 = c_2 = c_3 = c_4 = 1$;
 $C(G) = 1$;

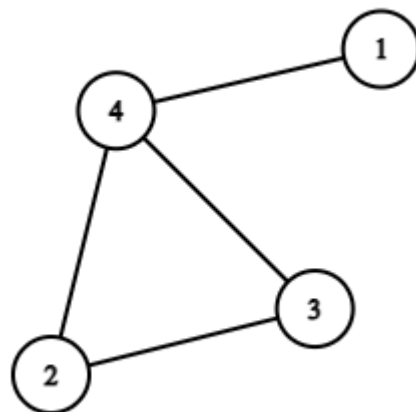


Рис. 4
 $N=4$, $m=4$;
 $c_1 = c_2 = c_3 = 1$, $c_4 = 0.33$;
 $C(G) = 0.835$;

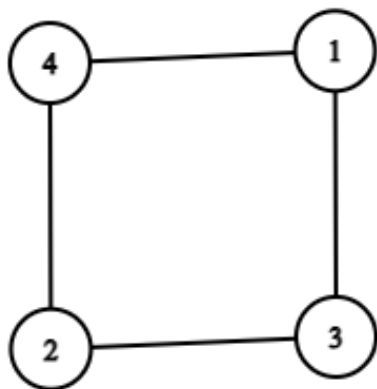


Рис. 5

$N=4$, $m=4$;

$c_1 = c_2 = c_3 = c_4 = 0$;

$C(G) = 0$;

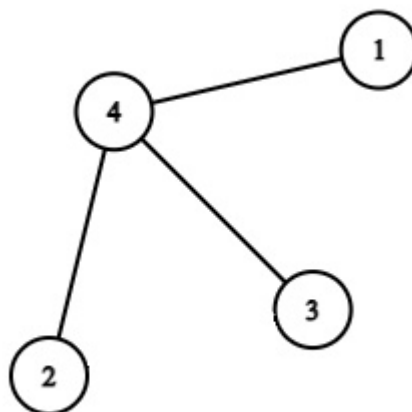


Рис. 6

$N=4$, $m=3$;

$c_1 = c_2 = c_3 = 1$, $c_4 = 0$;

$C(G) = 0.75$;

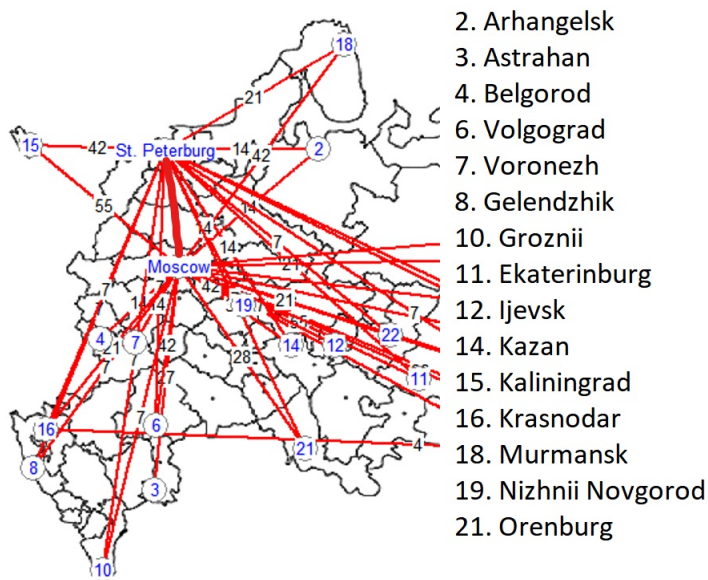
Треугольный граф не является показательным для подсчета СКК, поскольку во всех случаях, отличных от полного графа, коэффициент будет равен 0.

На примере графов с четырьмя вершинами можно сделать следующие выводы:

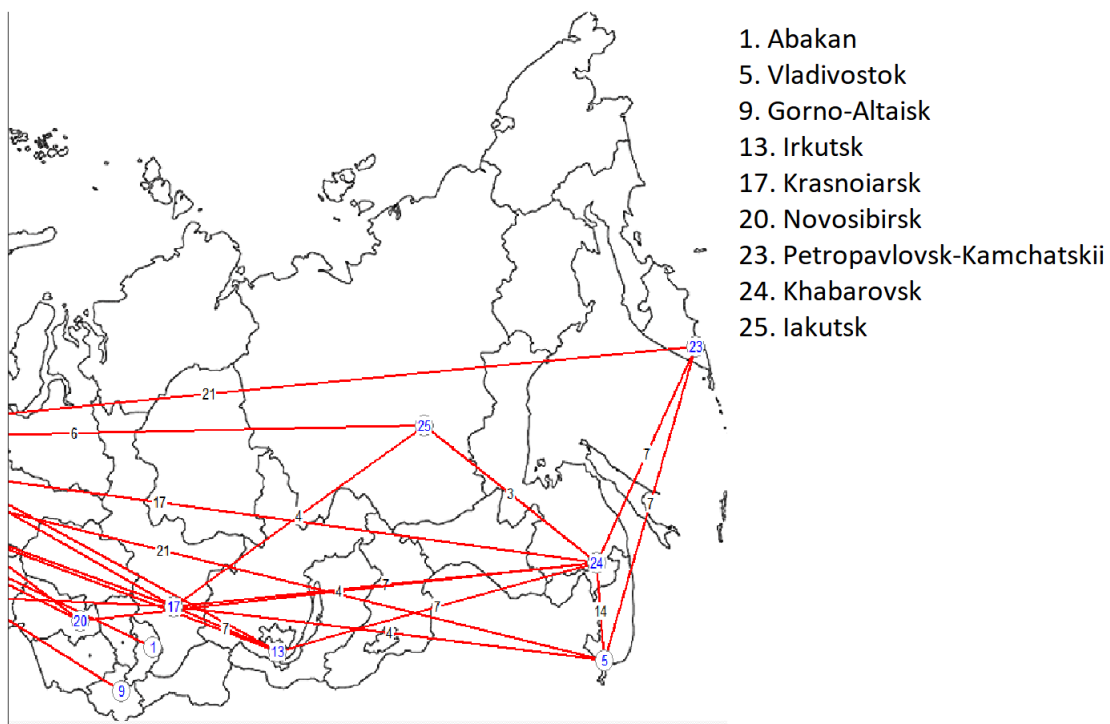
1. У полного графа наибольший СКК среди всех, что логично поскольку он является самым надежным.
2. Как правило, чем больше в графе ребер при одинаковом количестве вершин, тем больше его СКК.
3. СКК при обрыве ребра учитывает обходной путь только через одну вершину, т.к. в графе на рис. 5, например, коэффициент равен нулю.
4. СКК не всегда хорошо характеризует надежность графа. К примеру, очевидно, что граф на рис. 3 является более надежным чем граф на рис. 4, поскольку при прочих равных обрыв вершины 4 в первом случае оставляет граф связным, а во втором - нет, но при этом СКК у этих графов одинаковый.
5. На рис. 6 видно, что граф "звезда" является оптимальным с точки зрения СКК при минимальном количестве ребер.

3 Подсчет средневзвешенного коэффициента кластеризации для модели авиаперевозок внутри России

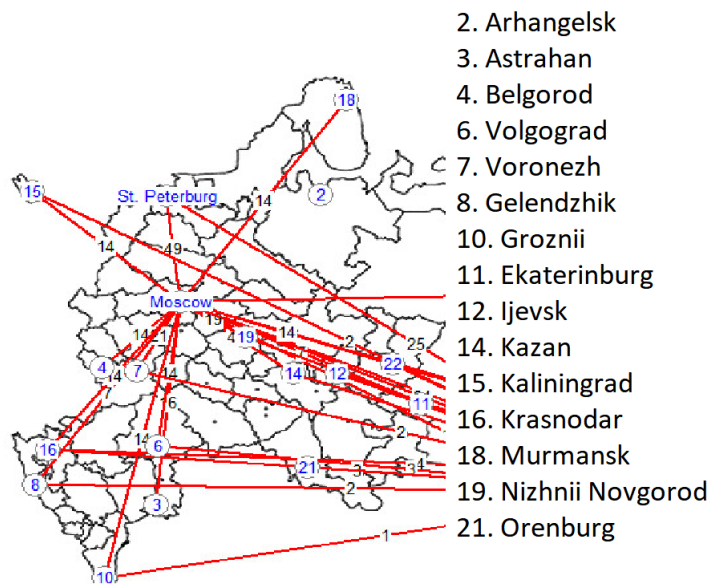
Для сравнения были взяты 2 авиакомпании, которые занимаются авиаперевозками внутри России, Aeroflot и S7. Было составлено 2 графа - состоящих из 27 вершин(крупнейшие города России) и ребер, где наличие ребра означает существование регулярного рейса между двумя городами, а его вес - количество таких рейсов в неделю на момент апреля 2022.



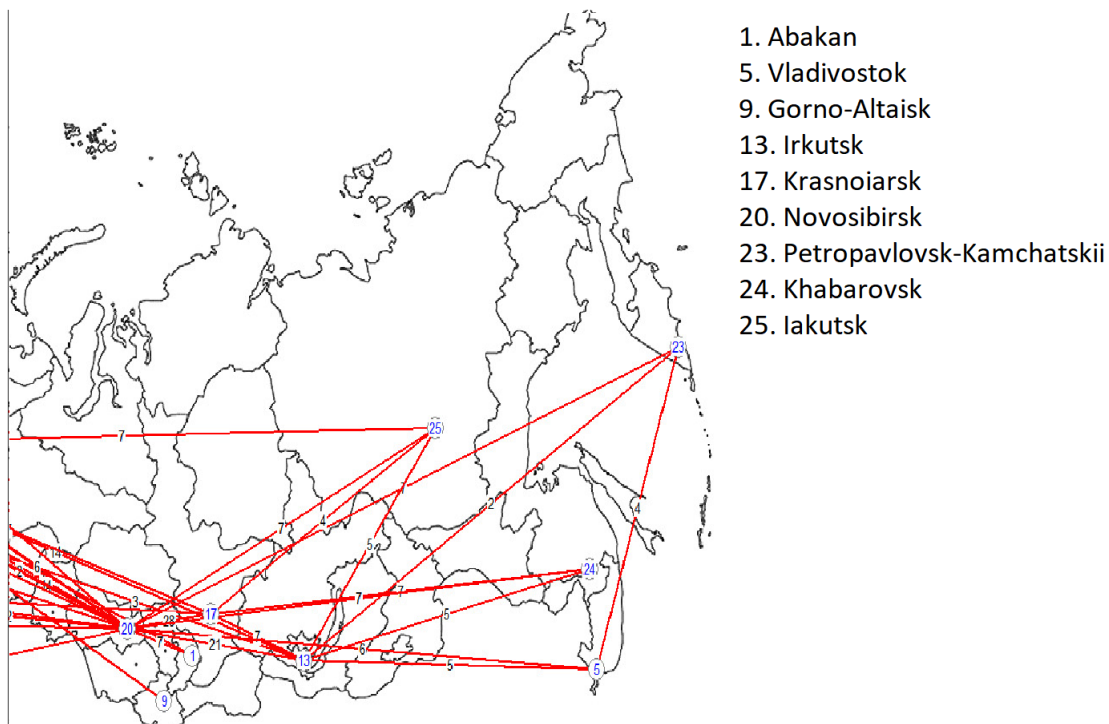
СВП Аэрофлота в европейской части России



СВП Аэрофлота в азиатской части России



СВП S7 в европейской части России



СВП S7 в азиатской части России

С помощью формулы (2) была написана программа на python, которая по графу, заданному с помощью матрицы смежности находит средневзвешенный коэффициент кластеризации. Для компании S7 этот коэффициент оказался равен 0.88680, а для компании Aeroflot - 0.91431.

Для сравнения, в [1], где поставлена задача максимизации средневзвешенного коэффициента кластеризации, максимальным получился коэффициент 0.992, то есть можно сделать вывод, что рассматриваемые в этой работе сети имеют достаточно высокий показатель коэффициента средневзвешенной кластеризации, а значит неплохо приспособлены к ситуациям, когда будет невозможно совершить полет через какой-то город.

У Aeroflot показатель несколько выше, поскольку почти все маршруты проходят сразу через 2 крупных центра (Москва и Санкт-Петербург), и, при невозможности пролета через один из этих городов, будет возможность составить маршрут через другой.

У S7 коэффициент немного ниже, поскольку их СВП имеет больше "тупиковых" маршрутов, и если пропадет возможность перемещаться через один из крупных городов (Москва, Новосибирск), то доступ

к "тупиковым" городам будет закрыт.

4 Заключение

В данной работе был проведен анализ модельной СВП России, существующей на момент апреля 2022 года с помощью средневзвешенного коэффициента кластеризации, а также сделаны наблюдения о применении данного параметра к сетям воздушного транспорта.

В [1] показано, что СКК позволяет не только проводить анализ существующих СВП, но и смоделировать более устойчивую сеть с помощью максимизации средневзвешенного коэффициента кластеризации. Это будет являться предметом моего дальнейшего исследования.

К данной задаче максимизации я планирую применить два следующих ограничения:

- По соображениям безопасности и из-за того, что маршруты имеют пропускную способность, веса ребер имеют верхнее значение β .
- Более того, чтобы установить новый маршрут полета, необходим минимальный объем спроса на трафик α .

Задача максимизации средневзвешенного коэффициента кластеризации для заданного порядка N и заданного размера m может быть записана следующим образом:

$$\max_G C_R^w(G(V, E)), \text{ где } \begin{cases} |V| = N \\ |E| = m \\ \forall(i, j), w_{ij} \in \{0, [\alpha, \beta]\} \end{cases} .$$

Поскольку ставится задача максимизации этого коэффициента, то его можно принять за критерий эффективности, коэффициенты α, β - за неконтролируемые факторы, а граф G - за контролируемый фактор [7].

5 Список литературы

- [1] J. Ponton, P. Wei and D. Sun, "Weighted Clustering Coefficient Maximization For Air Transportation Networks" , 2013 European Control Conference (ECC), Швейцария, Июль 2013.
- [2] L. A. N. Amaral, A. Scala, M. Barthelemy, and H. E. Stanley, "Classes of small-world networks" , Proceedings of the National Academy of Sciences, США, выпуск 97, номер 21, октябрь 2000.
- [3] H. Nagarajan, P. Wei, S. Rathinam, and D. Sun, "Air transportation network robustness optimization under limited legs itinerary constraint" , на 5-ой Международной конференции по исследованиям в области воздушного транспорта, Беркли, США, Май 2012.
- [4] G. Spiers, P. Wei, and D. Sun, "Algebraic Connectivity Maximization for Air Transportation Networks" , in American Control Conference, Монреаль, Канада, июнь 2012.
- [5] R. Guimera, S. Mossa, A. Turtleschi, and L. A. N. Amaral, "The worldwide air transportation network: Anomalous centrality, community structure, and cities global roles" , Proceedings of the National Academy of Sciences, США, выпуск 102, 2005.
- [6] J. Watts and S. H. Strogatz, "Collective dynamics of 'small world' networks" , Nature, выпуск 393, Июнь 1998.
- [7] Морозов В. В., Сухарев А. Г., Федоров В.В., "Исследование операций в задачах и упражнениях Москва, 1986.