

## УСТОЙЧИВОСТЬ РАБОТЫ РЕГУЛЯРНЫХ И СТОХАСТИЧЕСКИХ КОММУНИКАЦИОННЫХ СЕТЕЙ СО СВОЙСТВАМИ МАЛОГО МИРА

А.П. Демичев<sup>1</sup>, В.А. Ильин<sup>1,2</sup>, А.П. Крюков<sup>1</sup>, С.П. Поляков<sup>1</sup>

### Аннотация

В работе исследована устойчивость важнейших характеристик стохастических и регулярных (детерминистских) коммуникационных сетей с малым средним расстоянием между узлами (сети “малого мира”) при выходе из строя части узлов. В случае стохастических сетей используется алгоритм с оптимальными значениями числа перемычек и параметра распределения их длин, а в качестве регулярных сетей рассмотрена iVT-сеть (Interlaced Bypass Torus Networks), обладающая наилучшими характеристиками в классе сетей, построенных на основе детерминистских алгоритмов. Показано, что в широком диапазоне значений относительного числа вышедших из строя узлов рассмотренные сети являются весьма устойчивыми к выходу узлов из строя, причем iVT-сети ведут себя несколько лучше, чем стохастические сети.

---

**Ключевые слова:** суперкомпьютеры; коммуникационные сети; сети “малого мира”, устойчивость, каскадные отключения

**1. Введение.** Если нынешняя скорость прогресса суперкомпьютеров будет сохраняться, то следующий рубеж производительности — в 1 эксафлопс, или  $10^{18}$  операций в секунду (FLOPS), — будет достигнут к 2018–2020 годам [1], [2]. Подходы и принципы построения таких суперкомпьютеров интенсивно разрабатываются уже сейчас, поскольку на пути к их построению предстоит решить ряд сложных научно-технических задач и выработать принципиально новые решения для их архитектуры и аппаратной реализации. Суперкомпьютеры экзамасштаба должны удовлетворять большому набору требований к их основным характеристикам, большинство из которых почти в 100–1000 раз превышают аналогичные показатели лучших существующих суперкомпьютеров (см., например, [3]).

Одной из важнейших составляющих любого суперкомпьютера является коммуникационная сеть, которая в первую очередь определяет возможность увеличения числа вычислительных узлов, что необходимо для достижения желаемой производительности. Как ожидается (см., например, [4]), коммуникационная сеть суперкомпьютеров экзафлопсного уровня будут иметь порядка  $10^5$ – $10^6$  вычислительных узлов (без учета числа ядер каждого ЦПУ). Таким образом, одной из ключевых задач,

---

<sup>1</sup>НИИ ядерной физики имени Д.В. Скобельцына МГУ имени М.В. Ломоносова (НИИЯФ МГУ), Ленинские горы, дом 1, строение 2,119991, ГСП-1, Москва; А.П. Демичев, ст. науч. сотр., e-mail: demichev@theory.sinp.msu.ru; В.А. Ильин, вед. научн. сотр., e-mail: ilyin@theory.sinp.msu.ru; А.П. Крюков, вед. научн. сотр., e-mail: kryukov@theory.sinp.msu.ru; С.П. Поляков, мл. науч. сотр., e-mail: s.p.polyakov@gmail.com

<sup>2</sup>Национальный исследовательский центр “Курчатовский институт”, пл. Академика Курчатова, д. 1, 123182, Москва; начальник отделения, e-mail: ilyin@theory.sinp.msu.ru

которую предстоит решить на пути к построению суперкомпьютеров следующего поколения, является разработка коммуникационных сетей с хорошими свойствами масштабируемости, возможностью эффективно и надежно обслуживать огромное число вычислительных узлов.

В классе прямых коммуникационных сетей, в которых каждый узел является терминальным, действующим как источник и приемник для сообщений, а также и как маршрутизатор для управления входящими сообщениями, весьма популярными являются коммуникационные сети со структурой регулярных решеток и топологией  $D$ -мерных торов. Для них часто используется термин “ $k$ -ary  $D$ -cube” [5] ( $D$  — размерность тора;  $k$  — число узлов вдоль каждого из измерений). Такие сети обладают рядом преимуществ, особенно при решении специфических задач моделирования  $D$ -мерных объектов, когда структура вычислительной задачи оптимальным образом отображается на  $D$ -мерную решетку вычислительных узлов, а именно такого типа задачи, как предполагается, будут составлять значительную долю задач, решаемых на суперкомпьютерах экзафлопсного уровня. Однако при огромном числе узлов, характерном для компьютеров следующего поколения, архитектура простых регулярных решеток имеет и существенные недостатки. В частности, решетки невысокой размерности имеют весьма большую среднюю длину пути между узлами, а решетки высокой размерности, сравнимой с логарифмом числа узлов, трудно реализовать технически из-за большой длины физических коммуникационных каналов. С другой стороны, известно, что наилучшими структурами вычислительных систем по различным критериям функционирования, например, производительности и надежности, при одинаковом числе вычислительных узлов и каналов связи являются структуры с минимальным средним расстоянием между узлами (см., например, [6]). Поэтому обычные сети с простой структурой регулярных решеток окажутся недостаточно эффективными для решения задач более общего типа, не связанных с триангуляцией  $D$ -мерных объектов.

В связи с этим представляется перспективным использовать для построения коммуникационных сетей для экзафлопсных компьютеров решеточные сети с дополнительными связями — перемычками — между узлами (в англоязычной литературе для таких перемычек используется два термина: “shortcuts” — обычно в теории сложных сетей, см., например, [7] и “bypass” — обычно в литературе, посвященной собственно коммуникационным сетям, см., например, [3]). При правильном построении такие перемычки обеспечивают малое среднее расстояние  $d$  между узлами и малый диаметр  $\mathcal{D}$  сети. В частности, для ряда таких сетей с перемычками среднее расстояние  $d$  между узлами растет как логарифм числа узлов:  $d \sim \ln N$ , в то время как для обычной  $D$ -мерной решетки оно растет существенно быстрее — как степень числа узлов:  $d \sim N^{1/D}$ . В некоторых сетях с перемычками среднее расстояние зависит от числа узлов степенным образом, как и в случае регулярных решеток, однако показатель степени существенно меньше:  $d \sim N^\gamma$ ,  $\gamma \ll 1/D$ .

Часто сети с медленным (медленнее чем  $N^{1/D}$ ) ростом среднего расстояния называют сетями “малого мира”. Такие сети могут быть построены как с помощью детерминистских [8]-[19], так и с помощью стохастических алгоритмов [20]-[22], [14]. Оригинальный алгоритм получения сложной сети со свойствами малого мира [23] является стохастическим: на каждом шаге алгоритма ребра графа меняют свое по-

ложение с некоторой вероятностью. В результате многократного применения такого алгоритма возникает ансамбль графов с некоторым распределением их характеристик, в частности, средней длины пути между узлами экземпляра графа. Для многих реальных сетей стохастический процесс оказывается внутренне присущим (так, это справедливо для сети Интернет; другие примеры см., например, в [24]). Однако проектирование коммуникационной сети суперкомпьютера находится под контролем разработчика, и поэтому стохастичность не является внутренне присущим элементом этого процесса. Поэтому важным вопросом является следующий: существует ли такой регулярный (детерминистский) алгоритм модификации решетки с помощью перемычек, превращающей ее в сеть “малого мира”, чтобы характеристики получившейся сети были лучше, чем при использовании стохастических алгоритмов. В работе [14] проведено сравнение для ряда детерминистских и стохастических алгоритмов с точки зрения оптимизации соотношения “цены” и “качества” причем в качестве “цены” выступает удельная длина дополнительных перемычек (общая длина перемычек в единицах базовой решетки, деленная на число узлов сети), а “качество” — это глобальная или локально-навигационная (см. [25]) средняя длина пути между узлами. Показано, что существуют детерминистские алгоритмы, которые по этому показателю не уступают стохастическим сетям или превосходят их.

Важно отметить, однако, что для сверхбольших коммуникационных сетей, какими ожидаются сети экзафлопсных суперкомпьютеров, практически неизбежны случаи отказов части узлов коммуникационной сети в силу физической неисправности или просто перегрузки. Поэтому весьма важным для таких сетей является вопрос об устойчивости значений важнейших характеристик сети при выходе из строя части узлов (fault tolerance). Выход из строя части узлов коммуникационной сети приводит к некоторой стохастичности даже регулярных коммуникационных сетей, построенных по детерминистским алгоритмам. При этом может оказаться, что выход из строя узлов сильнее сказывается на характеристиках регулярных сетей, чем в случае сетей, которые изначально были стохастическими и нерегулярными.

Цель данной работы — сравнить устойчивость детерминистских (регулярных) и стохастических (нерегулярных) сетей. В данной работе для такого сравнения рассмотрены сети на основе двумерных регулярных решеток  $D = 2$ , к которым стохастическим или детерминистским образом добавляются перемычки. При этом рассматриваются только топологические (статические) характеристики — среднее навигационное расстояние (то есть расстояние при выбранной маршрутизации), нагрузка относительно маршрутизации (см. определение ниже), а также характеристики сетей при учете каскадных отключений узлов за счет перегрузки. Динамические модели передачи сообщений в рассматриваемых коммуникационных сетях, в частности, на основе теории систем массового обслуживания (см., например, [26], [27], [28]), будут изучены в следующих работах.

**2. Выбор сетей для сравнения и способа маршрутизации сообщений.** Стохастический алгоритм построения сетей малого мира с сохранением базовой решетки предложен в ряде работ, в частности, в работах [20], [21], [22] (см. также обзор [7] и ссылки в нем). Мы будем использовать следующий вариант стохастического алгоритма (сравнение различных вариантов см. в [14]):

1. Исходным объектом является двумерная решетка с  $N = L \times L$  узлами и топо-

логией двумерного тора;

2. последовательно перебираются все узлы решетки и к каждому узлу  $(i, j)$  подсоединяют первый конец перемычки ( $1 \leq i, j \leq L$ );
3. второй конец перемычки (то есть, узел решетки  $(k, l)$ , в который она входит) не может совпадать с соседями исходного узла в смысле базовой решетки и приводить к дублированию уже существующей перемычки, а в остальном выбирается случайно с вероятностью

$$P(r) \sim r^{-\alpha}, \quad (1)$$

которая является степенной функцией решеточного расстояния  $r = r_{(ij),(kl)}$  между узлами  $(i, j)$  и  $(k, l)$ .

Напомним, что *сетевое расстояние*  $d_{(ij),(kl)}$  между узлами сети определяется как число ребер кратчайшего пути *с учетом* перемычек. Это сетевое расстояние является функцией *решеточного расстояния*  $r_{(ij),(kl)}$  между узлами, то есть расстояния, измеренного по исходной решетке *без* перемычек. Если имеется в виду решеточное расстояние между произвольной парой узлов, то используется обозначение  $r$ . Зависимость вероятности появления перемычки от расстояния между узлами отражает корреляцию между топологическими и пространственными свойствами сети [7].

Помимо сетей, построенных по указанному базовому алгоритму, мы рассмотрим также сети, построенные по аналогичному алгоритму, но в котором к требованиям п. 3 добавляется требование фиксированной степени узлов (то есть, фиксированного числа ребер, присоединенных к вершинам графа, соответствующего сети) при условии возможно минимального искажения распределения (1).

Среди множества детерминистских сетей с малым средним расстоянием наиболее предпочтительной (см. [18], [19], [14]) по своим общим характеристикам представляется сеть iBT (Interlaced Bypass Torus Networks) [17]. Поэтому именно эту детерминистскую сеть мы будем сравнивать со стохастической сетью — с точки зрения устойчивости значений ее характеристик при выходе из строя части узлов. Детальное описание и анализ этих сетей можно найти в работах [18], [19]. Здесь очень кратко приведем алгоритм построения iBT-сетей для случая двумерного тора с перемычками в обоих измерениях и с использованием перемычек двух разных длин. В работах [18], [19] такие сети обозначаются следующим образом:  $iBT(L \times L; 2; s = \langle s_1, s_2 \rangle)$ :

1. Исходным объектом является двумерная решетка с  $N = L \times L$  узлами и топологией двумерного тора;
2. для каждого узла  $\vec{i} = \{i_1, i_2\}$  вычисляются две характеристики:
  - номер измерения  $d(\vec{i})$  (1 или 2), вдоль которого направлены перемычки, выходящие из данного узла:

$$d(\vec{i}) = [(i_1 + i_2) \bmod 2] + 1; \quad (2)$$

- номер длины перемычек  $h$  (1 или 2; соответствует длинам  $s_1$  и  $s_2$ ), которые выходят из данного узла:

$$h(\vec{i}) = \left\lfloor \frac{(i_1 + i_2) \bmod 4}{2} \right\rfloor + 1 ; \quad (3)$$

3. к каждому узлу  $\vec{i}$  присоединяются две перемычки, которые идут в двух направлениях измерения  $d(\vec{i})$  до узлов, находящихся на решеточном расстоянии  $s_{h(\vec{i})}$ .

**2.1. Способ маршрутизации сообщений.** Как уже упоминалось, длина пути (расстояние) между узлами понимается в сетевом смысле: как минимальное число ребер, по которым надо пройти, чтобы попасть из одного узла в другой. Соответственно среднее расстояние между узлами определяется как среднее по всем парам узлов данной сети. Однако для больших сетей, какой ожидается коммуникационная сеть экзафлопсных компьютеров, и при учете того факта, что неизбежны случаи отказов части узлов коммуникационной сети в силу физической неисправности или просто перегрузки, определенная таким образом длина пути между узлами может оказаться неадекватной характеристикой, поскольку для нахождения кратчайших маршрутов необходимо знать текущую глобальную структуру сети (с учетом вышедших из строя узлов). Соответственно маршрутизация сообщений, использующая кратчайшие пути, может оказаться слишком сложной и неэффективной, так как связана с хранением и обработкой большого объема информации.

Более эффективным подходом является использование того ли иного адаптивного алгоритма (см., например, [5]). Однако, как правило, эффективность конкретного адаптивного алгоритма существенно зависит от типа и архитектуры сети, для которой он используется. Поскольку в данной работе мы сравниваем существенно разные — стохастические и детерминистские — коммуникационные сети, мы используем максимально общий адаптивный алгоритм маршрутизации, основанный на принципах локальной навигации [25]. При локальной навигации узел “знает” только географическое положение (другими словами, положение в базовой решетке) всех узлов и своих ближайших сетевых соседей с учетом перемычек. Используя только такую информацию (информация обо всех перемычках в сети не используется) необходимо доставить сообщение в узел назначения по возможно кратчайшему пути. В простейшем варианте эту задачу решает так называемый жадный алгоритм (англ. greedy algorithm; иногда называется также алгоритмом экономного продвижения): текущий узел пересылает сообщение тому из своих соседей, который географически — то есть в смысле координат на решетке — ближе всего к цели (узлу назначения).

Важнейшей характеристикой процесса навигации является средняя длина навигационного пути между узлами сети, которая определяется аналогично обычной средней длине:

$$\ell = \frac{2}{L(L-1)} \sum_{i>j} \langle \ell_{ij} \rangle = \frac{1}{L(L-1)} \sum_{i,j} \langle \ell_{ij} \rangle , \quad (4)$$

где  $\ell_{ij}$  — расстояние (число ребер) между узлами  $i$  и  $j$  при локальной навигации. Угловые скобки  $\langle \cdot \rangle$  означают усреднение по ансамблю сетей, полученных с помо-

щью стохастических алгоритмов. Для детерминистских алгоритмов такое усреднение, естественно, отсутствует. Однако некоторые сети обладают весьма большой средней длиной пути между узлами при навигации, основанной на простейшем жадном алгоритме. Для исправления такой ситуации можно использовать модифицированный алгоритм локальной навигации — двухуровневую локальную навигацию [14]. В этом варианте просматриваются не только ближайшие соседи, но и соседи соседей. При этом сообщение на следующем шаге пересылается в тот соседний узел, один из соседей которого ближе всего к узлу назначения в смысле решеточной метрики. Хотя при таком алгоритме на каждом шаге объем вычислений несколько увеличивается, алгоритм остается локальным (не вычисляется весь путь до адресата, и объем вычислений не зависит от размеров системы). Поэтому этот алгоритм является хорошо масштабируемым и приемлем для сверхбольших коммуникационных сетей. Среднюю длину пути при двухуровневой локальной навигации обозначим  $\ell^{(2)}$ . Глобальная средняя длина пути между узлами  $d$  может рассматриваться как предельный случай навигационной длины при бесконечной глубине просмотра.

Заметим, что если для регулярной сети используется маршрутизация, привязанная к конкретной регулярной структуре, то при выходе из строя узлов (разрушении этой регулярной структуры) длина путей должна меняться сильнее, чем просто для жадного алгоритма, который является универсальным для любых сетей. Другими словами, представляется, что жадный алгоритм может быть не самым оптимальным для регулярных сетей, но наиболее устойчивым по отношению к выходу из строя части узлов сети.

Конечно, при возможности выхода из строя части узлов жадный алгоритм требует уточнения — на случай, когда узел, куда сообщение должно перейти в соответствии с (простым или двухуровневым) жадным алгоритмом маршрутизации оказывается неисправным. Такое уточнение представлено в следующем разделе.

**2.2. Маршрутизация и зависимость нагрузки от размера сети.** В качестве первой характеристики, которая позволяет сравнить детерминистские и стохастические сети, рассматривается нагрузка на узлы. Глобально вычисленная нагрузка на узлы (betweenness centrality; см., например, [7]) — важная общая характеристика сложных сетей. Однако для коммуникационных сетей, как уже отмечалось, реальный интерес представляют не глобально-кратчайшие пути, а пути, определяемые конкретной маршрутизацией. Поэтому и нагрузку необходимо определять по отношению к используемой маршрутизации (нагрузка относительно маршрутизации, НОМ) [29]:

$$f(v) = \sum_{s \neq v \neq t} \sigma_{st}^r(v), \quad (5)$$

где  $\sigma_{st}^r(v)$  — число кратчайших путей из узла  $s$  в узел  $t$  для данной маршрутизации, проходящих через узел  $v$ . Для максимального значения  $f_{max} = \max_{v \in V} f(v)$  ( $V$  — множество всех узлов сети) в англоязычной литературе используется термин “forwarding index” (см., например, [29]). Мы будем использовать термин “максимальная нагрузка относительно маршрутизации” (МНОМ). В данной работе мы исследуем нагрузку относительно маршрутизации, основанной на жадном алгоритме глубины просмотра два.

На рис. 1 показаны МНОМ в стохастической сети и в iBT-сети в зависимости от

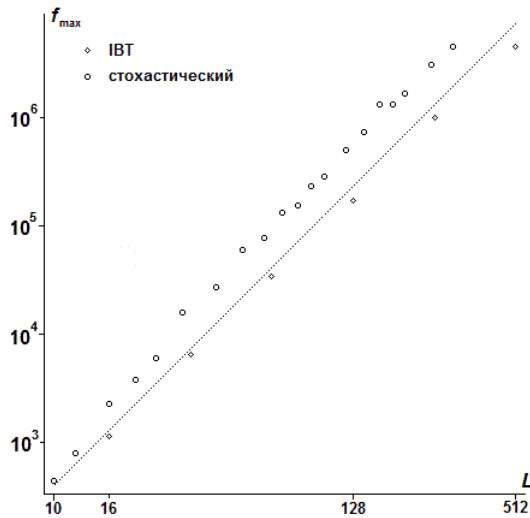


Рис. 1. Максимальная нагрузка относительно маршрутизации на основе жадного алгоритма глубины два (МНОМ) в стохастической сети (окружности) и iBT (ромбы) в зависимости от линейного размера сети; для сравнения: пунктирная линия —  $f_{max} \sim L^{(5/2)}$ ; данные на осях представлены в логарифмическом масштабе

их размеров, полученные с помощью численного моделирования. Результаты представлены для сетей с наилучшими значениями  $\ell^{(2)}$ : в случае стохастической сети отобраны сети из 100 образцов с параметром  $\alpha = 1$ ; а в случае iBT-сети выбраны наилучшие схемы с двумя длинами перемычек. Видно, что МНОМ для для обоих типов сетей МНОМ растет примерно как  $f_{max} \sim L^{(5/2)}$ , но значения нагрузки в стохастической сети несколько выше, чем в iBT.

**3. Устойчивость сетей при выходе из строя части узлов.** Для сравнения устойчивости из набора сетей каждого вида были выбраны лучшие образцы с точки зрения средней навигационной длины  $\ell^{(2)}$  и МНОМ (forwarding index)  $f_{max}$  для сетей без вышедших из строя узлов. Узлы, которые считаются вышедшими из строя, выбирались случайно, несколько раз (10 для представленных графиков) для каждого экземпляра сети и каждого значения числа  $b$  неисправных узлов. Результаты усреднены по этим выборкам.

Возможность выхода узлов из строя требует уточнения алгоритма маршрутизации, основанного на локальной навигации: необходимо определить поведение сообщения при встрече с вышедшим из строя узлом. Такой уточненный алгоритм маршрутизации выглядит следующим образом:

- просматриваются соседи текущего узла на глубину два;
- сообщение перемещается в тот соседний узел, один из соседей которого ближе всего к узлу-адресату (пока это просто жадный алгоритм с глубиной просмотра два; дальнейшие шаги алгоритма уточняют его на случай выхода из строя части узлов);

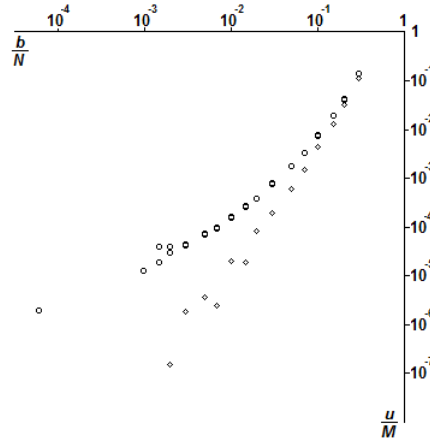


Рис. 2. Зависимость доли  $u/M$  недоставленных при выбранном алгоритме сообщений от доли вышедших из строя узлов в стохастических сетях (окружности) и в сетях iVT (ромбы); линейный размер всех сетей  $L = 128$

- запрещается возвращаться в узел, из которого сообщение пришло на предыдущем шаге (полный список пройденных узлов не хранится, поэтому возможны закливания (live-lock));
- перемещение сообщения заканчивается
  1. если оно попадает в узел-адресат;
  2. если некуда идти;
  3. если пройденный путь больше  $2\mathcal{D}^{(2)}$ , где  $\mathcal{D}^{(2)}$  — навигационный диаметр, т.е. максимальная длина пути в сети при использовании жадного алгоритма с глубиной просмотра два.

Во всех случаях, кроме первого, сообщение считается потерянным;

- источником и адресатом сообщений могут быть только исправные узлы.

Зависимость доли  $u/M$  недоставленных при выбранном алгоритме сообщений ( $M$  — полное число сообщений, сгенерированных в моделируемой системе) от доли  $b/N$  вышедших из строя узлов в стохастических сетях с  $\alpha = 1$  (окружности) и в сетях iVT (ромбы) показана на рис. 2. Линейный размер всех сетей  $L = 128$ , а полное число узлов  $N = 16384 \approx 1.6 \times 10^4$ . Видно, что при относительно небольшой доле неисправных узлов в iVT-сетях оказывается меньше потерянных сообщений, а при достижении доли неисправных узлов  $\sim 10\%$  относительное число потерянных сообщений сравнивается для обоих типов сетей.

**3.1. Устойчивость значений среднего навигационного расстояния при выходе из строя части узлов сети.** На рис. 3 показана зависимость среднего навигационного расстояния  $\ell^{(2)}$  от доли числа вышедших из строя узлов  $b/N$  в стохастических сетях при  $\alpha = 1$  (сплошные линии соответствуют лучшим образцам из



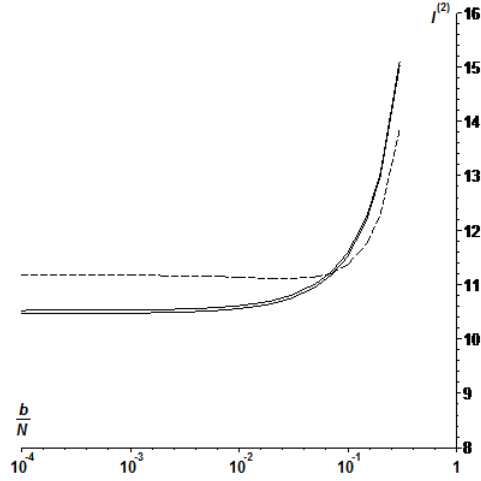


Рис. 3. Зависимость среднего навигационного расстояния  $\ell^{(2)}$  от числа вышедших из строя узлов  $b$  в стохастических сетях (сплошные линии) и в iVT-сетях при  $k = 8, 32$  (прерывистая линия); сообщения, не достигшие узла-адресата, не учитываются; линейный размер сетей  $L = 128$

ансамбля в 100 экземпляров) и в iVT-сетях при  $s_1 = 8, s_2 = 32$  (прерывистая линия) в случае, когда сообщения, не достигшие узла-адресата, не учитываются. Очевидно, это дает нижнюю оценку для реальной величины  $\ell^{(2)}$ . Было осуществлено также численное моделирование зависимости навигационной длины пути от числа вышедших из строя узлов, при котором для не достигших узла-адресата сообщений длина пути считается равной нескольким навигационным диаметрам, например  $3\mathcal{D}^{(2)}$  или  $10\mathcal{D}^{(2)}$ . Результаты моделирования показывают, что

- качественное поведение результатов не зависит от величины пути, которое приписывается недоставленным сообщениям (если этот путь  $\gtrsim 2\mathcal{D}^{(2)}$ ) и практически совпадает с нижней оценкой, представленной на рис. 3;
- оба лучших экземпляра стохастических сетей и iVT-сеть ведут себя примерно одинаково;
- пока доля вышедших из строя узлов  $\lesssim 10\%$ , средняя навигационная длина практически не меняется,
- когда доля вышедших из строя узлов  $\gtrsim 10\%$ , средняя навигационная длина резко возрастает.

**3.2. Устойчивость значений нагрузки при выходе из строя части узлов сети.** На рис. 4 показаны результаты численного моделирования для зависимости МНОМ  $f_{max}$  от доли вышедших из строя узлов  $b/N$  в стохастических и iVT-сетях. МНОМ, созданная недоставленными сообщениями, учитывается так же, как и дошедшими до адресата. Для вышедших из строя узлов НОМ считается равной нулю. Результаты показывают, что до значений  $\approx 10\%$  в случае стохастических и iVT-сетей

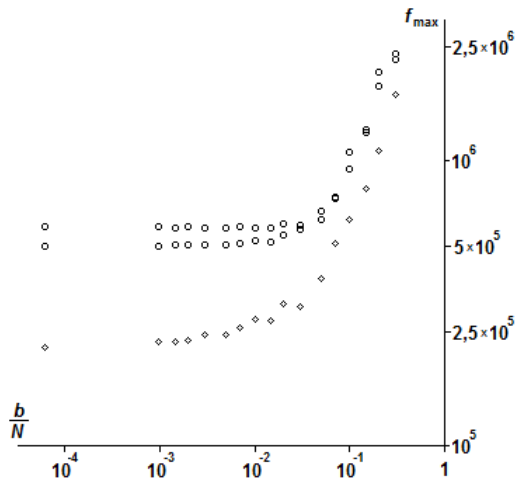


Рис. 4. Зависимость МНОМ от доли выведенных из строя узлов в стохастических сетях (окружности) и в сетях iVT с  $k = 8, 32$  (ромбы);  $L = 128$

МНОМ практически не меняется, и оба типа сетей (стохастические и детерминистские) ведут себя примерно одинаково (но нагрузка для iVT-сетей несколько ниже). При достижении доли неисправных узлов  $\gtrsim 10\%$  максимальная нагрузка  $f_{max}$  резко возрастает — как и в случае средней навигационной длины.

МНОМ является важной, но неполной характеристикой нагрузки на узлы сети. Более детальную картину дает распределение нагрузки относительно маршрутизации (НОМ) по узлам сети. На рис. 5 показаны распределения НОМ в стохастических и детерминистских сетях при разном числе вышедших из строя узлов. Результаты показывают:

- в случае, когда все узлы исправны, распределение нагрузок в стохастических сетях, как и можно было ожидать, существенно шире, чем в iVT-сетях, причем МНОМ в первых примерно в два раза выше, чем во вторых; это означает, что для одинаковых по размеру сетей пропускная способность маршрутизаторов в стохастических сетях должна быть существенно выше, чем в iVT-сетях;
- распределение нагрузок в iVT-сетях продолжает оставаться более узким вплоть до значения доли вышедших из строя узлов  $\sim 30\%$  (на последнем графике на рис. 5 показаны для большей наглядности только результаты для стохастических и iVT-сетей);
- весьма наглядным является сравнение распределения нагрузок для сетей со свойствами “малого мира” с нагрузками в случае обыкновенной решетки с топологией тора: для полностью исправной решетки нагрузка для всех узлов одинакова, но существенно выше, чем для сетей малого мира (это непосредственно связано с существенно большей величиной среднего расстояния между узлами для обычной решетки); при выходе из строя даже небольшой части узлов (на рисунке представлены результаты для 1% и 5% вышедших из строя узлов)

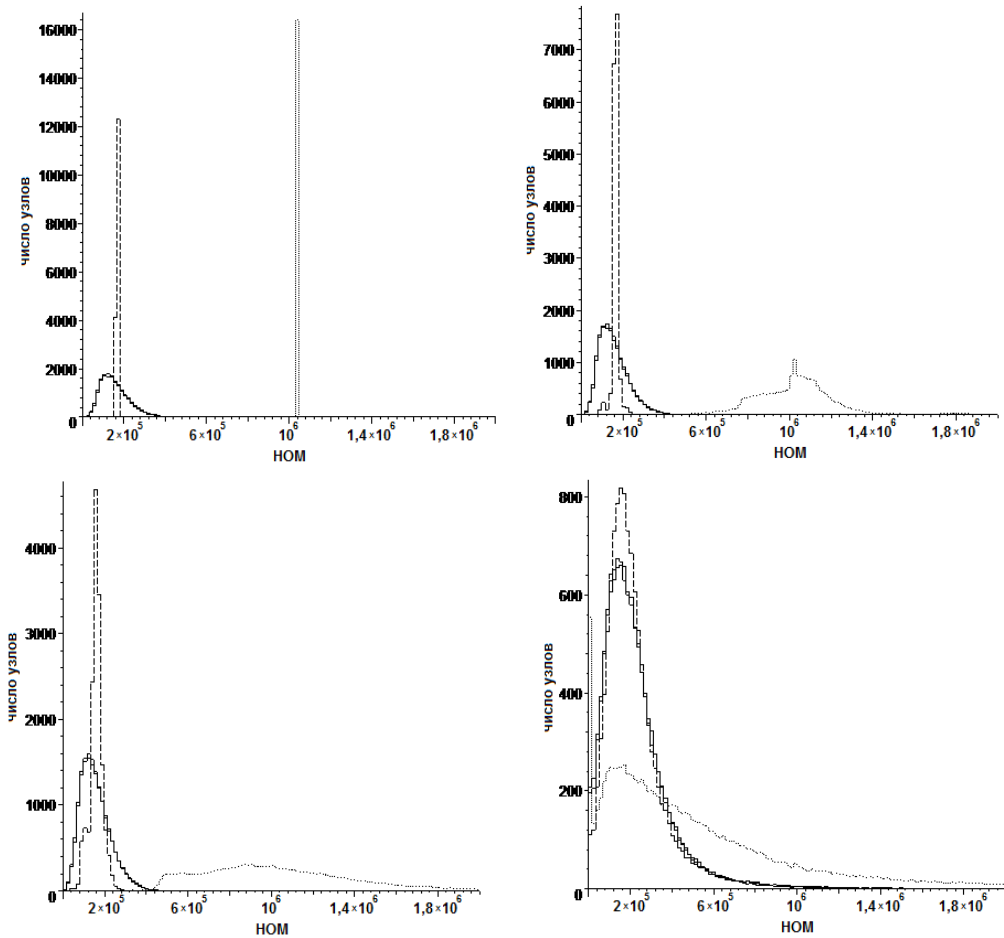


Рис. 5. Распределение нагрузок (НОМ) для стохастических сетей (сплошные линии), iVT (прерывистая линия) и тора без перемычек (пунктир) с долями вышедших из строя узлов (слева направо и сверху вниз):  $b = 0, 0,01N, 0,05N, 0,3N$ ;  $L = 128$

ширина распределения нагрузки резко увеличивается, так что значительная часть узлов может испытывать существенные перегрузки; это показывает, что сети малого мира не только обеспечивают меньшую временную задержку при передаче сообщений за счет малого среднего расстояния между узлами, но и являются существенно более устойчивыми при выходе из строя части узлов сети (например, при 5% вышедших из строя узлов средняя нагрузка на узлы оказывается примерно в пять раз выше чем для стохастических сетей с перемычками).

Как видно из этих результатов, в случае стохастических сетей распределение нагрузки на узлы является достаточно широким. Это является нежелательным свойством, так как означает, что узлы должны быть либо различными с точки зрения их пропускной способности, либо все должны быть рассчитаны на максимальную нагрузку, соответствующую “хвосту” распределения. В последнем случае значительное число узлов будет работать с существенной “недогрузкой”. Можно предполо-

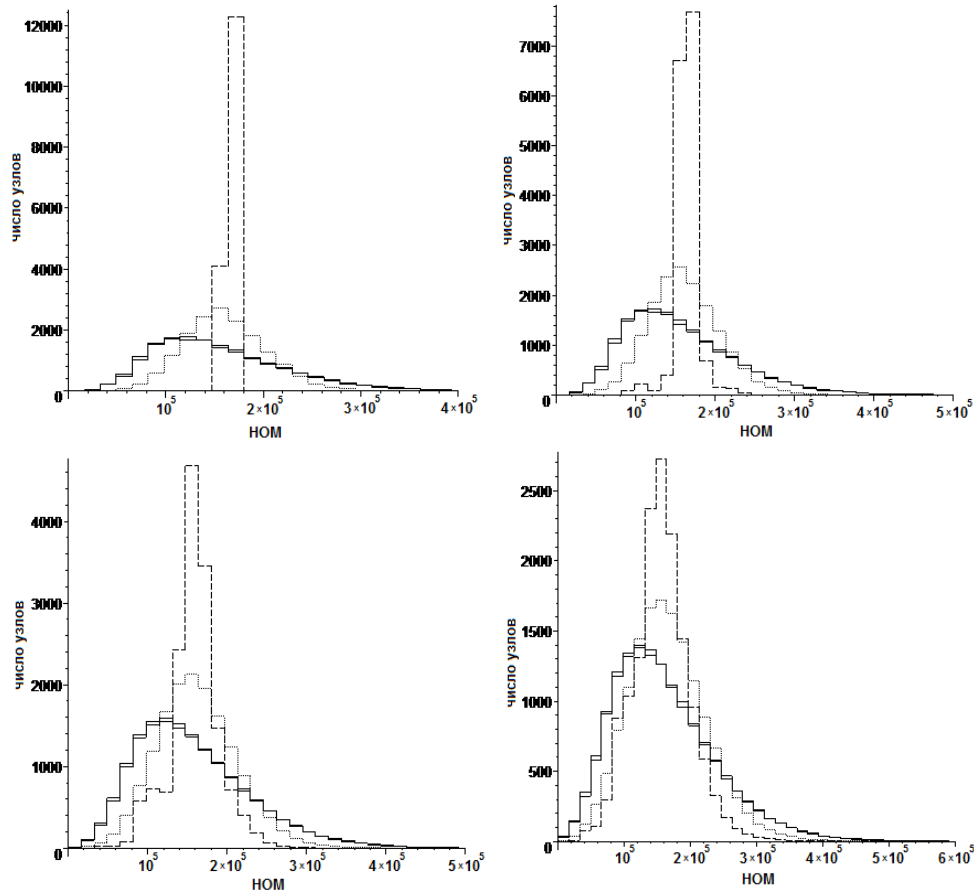


Рис. 6. Распределение нагрузок (НОМ) для стохастических сетей (сплошные линии), iVT (прерывистая линия) и стохастических сетей со степенью узлов 6 (пунктир) с долями вышедших из строя узлов (слева направо и сверху вниз):  $b = 0, 0,01N, 0,05N, 0,1N$ ;  $L = 128$

жить, что нагрузка на данный узел зависит от числа его связей (величины степени узла в терминологии теории графов). Поэтому представляет интерес выяснить — изменится ли ширина распределения нагрузки, если строить стохастические сети с фиксированной степенью узлов при условии минимального искажения распределения длин перемычек (1). Результаты соответствующего численного моделирования представлены на рис. 6, где обычные стохастические сети и iVT-сети сравниваются со стохастическими сетями, в которых все узлы имеют степень шесть: по четыре связи, образующие регулярную решетку, и ровно две перемычки — так же как и iVT-сети.

К сожалению, как видно из представленных результатов, фиксация величины степеней узлов в стохастических сетях практически не влияет на ширину распределения нагрузок: стохастичность расположения перемычек оказывается достаточной для того, чтобы узлы испытывали существенно разную нагрузку.

**3.3. Устойчивость сетей при каскадных отключениях.** Когда часть узлов выходит из строя, нагрузка на узлы, оставшиеся работоспособными, возрастает. Если

для некоторого узла эта возросшая нагрузка достигнет некоторого порога, а именно максимальной допустимой нагрузки для данного узла, то этот узел фактически тоже перестает быть работоспособным. Таким образом может развиваться каскадный процесс отключения узлов (blackout). Другими словами, каскадные отключения узлов сети есть результат лавинообразного распространения перегрузки, возникшей из-за первоначального выхода из строя некоторой части узлов (см., например, [30], [31]).

В данном разделе представлены результаты численного моделирования устойчивости стохастических и детерминистских сетей с переключками к таким каскадным отключениям узлов. Как и в предыдущем разделе, под нагрузкой узла понимается нагрузка относительно маршрутизации (НОМ). Как всюду в данной работе маршрутизация основана на жадном алгоритме глубины просмотра два. Для каждого узла сетей вводится предельная (пороговая) нагрузка  $f_{th}$ , после достижения которой узел перестает передавать сообщения — фактически выходит из строя. Каскадное отключение в этой модели выглядит следующим образом:

1. из сети случайным образом удаляется некоторая начальная доля узлов (предполагается, что эти узлы вышли из строя в силу аппаратных отказов или случайной перегрузки (не связанной с перераспределением нагрузки));
2. это приводит к перераспределению нагрузок на другие узлы; определяется значение предельной (пороговая) нагрузки  $f_{th}$  и находятся все перегруженные узлы в сети;
3. все перегруженные узлы удаляются из сети, что вновь приводит к перераспределению нагрузки и процедура предыдущего шага повторяется до тех пор, пока остаются перегруженные узлы.

На рис. 7 показана зависимость доли перегруженных узлов от доли первоначально испорченных узлов в стохастических и iBT-сетях. При этом максимальная допустимая нагрузка  $f_{th}$  выбирается исходя из распределения нагрузок в исправных сетях малого мира (первый график на рис. 5). А именно,  $f_{th}$  выбирается кратной максимальному значению нагрузки в iBT-сетях:  $f_{th} = k f_{max}^{(iBT)}$  (то есть, при проектировании коммуникационной сети маршрутизаторы выбираются с  $k$ -кратным запасом прочности по отношению к iBT-сетям). Численное моделирование показывает, что если взять  $k = 2$ , то часть узлов в стохастической сети сразу оказывается перегруженной и эта перегрузка сразу вызывает каскадное отключение всей сети. На рис. 7 представлены результаты для  $k = 3$ .

Результаты показывают, что с точки зрения возможности каскадных отключений оба типа сетей ведут себя примерно одинаково: до значений доли вышедших из строя узлов  $\sim 10\% \div 15\%$  каскадных отключений нет, а выше этих значений происходит резкий рост перегруженных, и как следствие, вышедших из строя узлов. При этом iBT-сети являются даже несколько более устойчивыми к каскадным отключениям, чем стохастические сети. Как видно, фиксация величины степеней узлов в стохастических сетях не оказывает существенного влияния на каскадные процессы.

**4. Заключение.** В данной работе исследована устойчивость важнейших характеристик стохастических и регулярных (детерминистских) коммуникационных сетей

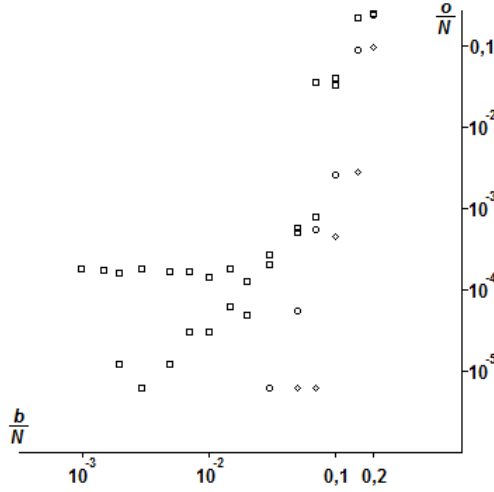


Рис. 7. Доля перегруженных узлов от доли испорченных узлов в стохастических сетях (прямоугольники), в стохастических сетях со степенью 6 (круги) и в сетях iVT с  $k = 8, 32$  (ромбы);  $L = 128$ , навигация с глубиной просмотра 2; максимальная допустимая нагрузка равна  $3f_{max}^{(iVT)}$ ; логарифмическая шкала по обеим осям

с малым средним расстоянием между узлами (сети малого мира) при выходе из строя части узлов. Рассмотрены сети на основе двумерных регулярных решеток, к которым стохастическим или детерминистским образом добавляются перемычки. В случае стохастических сетей используется алгоритм с оптимальными значениями числа перемычек и параметра распределения их длин [14], а в качестве регулярных сетей рассмотрена iVT-сеть (Interlaced Bypass Torus Networks) [17], обладающая наилучшими характеристиками в классе сетей, построенных на основе детерминистских алгоритмов. Показано, что максимальная нагрузка относительно маршрутизации (МНОМ) для для обоих типов сетей МНОМ растет примерно как  $f_{max} \sim L^{(5/2)}$ , но значения нагрузки в стохастической сети несколько выше, чем в iVT.

При относительно небольшой доле неисправных узлов в iVT-сетях оказывается меньше потерянных сообщений, а при достижении доли неисправных узлов  $\sim 10\%$  относительное число потерянных сообщений сравнивается для обоих типов сетей. С точки зрения устойчивости значений среднего навигационного расстояния между узлами  $\ell^{(2)}$  стохастические и iVT-сети ведут себя примерно одинаково, причем пока доля вышедших из строя узлов  $\lesssim 10\%$ , средняя навигационная длина практически не меняется, а когда доля вышедших из строя узлов  $\gtrsim 10\%$ , средняя навигационная длина резко возрастает. Аналогично среднему навигационному расстоянию, МНОМ практически не меняется пока доля неисправных узлов не достигает  $\approx 10\%$ , и оба типа сетей (стохастические и детерминистские) ведут себя примерно одинаково (но нагрузка для iVT-сетей несколько ниже); при достижении доли неисправных узлов  $\gtrsim 10\%$  максимальная нагрузка  $f_{max}$  резко возрастает.

Распределение нагрузок в стохастических сетях, как и можно было ожидать, существенно шире, чем в iVT-сетях, причем МНОМ в первых примерно в два раза

выше, чем во вторых; это означает, что для одинаковых по размеру сетей пропускная способность маршрутизаторов в стохастических сетях должна быть существенно выше, чем в iBT-сетях. Распределение нагрузок в iBT-сетях продолжает оставаться более узким вплоть до значения доли вышедших из строя узлов  $\sim 30\%$ . Фиксация величины степеней узлов в стохастических сетях несколько уменьшает ширину распределения нагрузок, однако стохастичность расположения перемычек остается достаточной для того, чтобы узлы испытывали существенно разную нагрузку.

Весьма наглядным является сравнение распределения нагрузок для сетей со свойствами “малого мира” с нагрузками в случае обыкновенной решетки с топологией тора: для полностью исправной решетки без перемычек нагрузка для всех узлов одинакова, но существенно выше, чем для сетей малого мира (это непосредственно связано с существенно большей величиной среднего расстояния между узлами для обычной решетки); при выходе из строя даже небольшой части узлов ( $\sim 1\% \div 5\%$ ) ширина распределения нагрузки резко увеличивается, так что значительная часть узлов может испытывать существенные перегрузки; это показывает, что сети малого мира не только обеспечивают меньшую временную задержку при передаче сообщений за счет малого среднего расстояния между узлами, но и являются существенно более устойчивыми при выходе из строя части узлов сети.

Результаты исследования устойчивости сетей при каскадных отключениях показывают, что с точки зрения возможности каскадных отключений оба типа сетей ведут себя примерно одинаково: до значений доли вышедших из строя узлов  $\sim 10\% \div 15\%$  каскадных отключений нет, а выше этих значений происходит резкий рост перегруженных, и как следствие, вышедших из строя узлов. При этом iBT-сети являются даже несколько более устойчивыми к каскадным отключениям, чем стохастические сети. Фиксация величины степеней узлов в стохастических сетях не оказывает существенного влияния на каскадные процессы.

В целом результаты работы показывают, что в отличие от регулярных решеток с топологией тора рассмотренные сети со свойствами “малого мира” в широком диапазоне значений относительного числа  $b/N$  вышедших из строя узлов сети являются весьма устойчивыми к выходу узлов из строя. При этом iBT-сети, построенные на основе детерминистского алгоритма, ведут себя несколько лучше, чем стохастические сети.

Как отмечалось во Введении, в данной работе рассмотрены только статические (топологические) характеристики сетей — среднее навигационное расстояние (то есть расстояние при выбранной маршрутизации) и нагрузка относительно маршрутизации. Динамические модели передачи сообщений в рассматриваемых коммуникационных сетях, в частности, на основе теории систем массового обслуживания, будут изучены в следующих работах.

Работа частично финансируется РФФИ, грант 12-07-00408-а.

## СПИСОК ЛИТЕРАТУРЫ

1. *Shainer G., Sparks B., Graham R.* Toward Exascale computing [Electronic resource] // HPC Advisory Council [Official website]. URL: [http://www.hpcadvisorycouncil.com/pdf/Toward\\_Exascale\\_computing.pdf](http://www.hpcadvisorycouncil.com/pdf/Toward_Exascale_computing.pdf) (accessed: 01.08.2013).
2. Концепция создания эксафлопсного суперкомпьютера России: “Эксафлопные технологии. Концепция по развитию технологии высокопроизводительных вычислений на базе суперэвм эксафлопного класса (2012–2020 гг.)” [Электронный ресурс] // Интернет-издание [Официальный сайт]. URL: [http://filearchive.cnews.ru/doc/2012/03/esk\\_tex.pdf](http://filearchive.cnews.ru/doc/2012/03/esk_tex.pdf) (дата обращения: 01.08.2013).
3. *Deng Y., Zhang P.* Perspectives on exascale computing // Jour. of New Computing Architecture and Applications. 2010. 1. 8–22.
4. Report on Institute for Advanced Architectures and Algorithms Interconnection Networks Workshop 2008 [Electronic resource] // Future Technologies Group Technical Report Series, Oak Ridge, Tennessee USA [Official website]. URL: <http://www.csm.ornl.gov/workshops/IAA-IC-Workshop-08> (accessed: 01.08.2013).
5. *Dally W. J., Towles B. P.* Principles and Practices of Interconnection Networks. Amsterdam: Elsevier Science, 2003.
6. *Kleinrock L.* Communication Nets: Stochastic Message Flow and Design. New York: McGraw-Hill, 1964.
7. *Barthelemy M.* Spatial Networks // Phys. Reports. 2011. 499. 1-101.
8. *Zhu Z.-Y., Mao B.-H., Hao H.-M., Gao J.-Z., Yang J.-J.* Regular Small-World Network // Chin. Phys. Lett. 2009. 26. 110502-1–110502-3.
9. *Boettcher S., Goncalves B., Azaret J.* Geometry and Dynamics for Hierarchical Regular Networks // Journal of Physics. 2008. A41. 335003-1–335003-25.
10. *Boettcher S., Goncalves B., Guclu H.* Hierarchical Regular Small-World Networks // Jour. Phys. 2008. A41. 252001-1–252001-7.
11. *Монахова Э. А.* Структурные и коммуникативные свойства циркулянтных сетей // Прикладная дискретная математика. 2011. 3, № 13. 92-115.
12. *Comellas F., Ozona J., Peters J. G.* Deterministic small-world communication networks // Information Processing Letters. 2000. 76. 83-90.
13. *Comellas F., Mitjana M., Peters J.G.* Broadcasting in Small-World Communication Networks // Proc. 9th Int. Coll. on Structural Information and Communication Complexity. 2002. 73-85.



14. Демичев А.П., Ильин В.А., Крюков А.П., Поляков С.П. Сравнительный анализ алгоритмов построения больших коммуникационных сетей со свойствами “малого мира” // Вестник УГАТУ. 2013. 17, № 5. 167-175
15. Inoguchi, Y. and Horiguchi S. Shifted Recursive Torus Interconnection for High Performance Computing // Proceedings of the High-Performance Computing on the Information Superhighway, HPC-Asia '97, IEEE Computer Society. 1997. 61.
16. Yang, Y., Funahashi A., Akiya J., Nishi H., Amano H., Sueyoshi T. Recursive Diagonal Torus: An Interconnection Network for Massively Parallel Computers // IEEE Transactions on Parallel and Distributed Systems. 2001. 12, N 7. 701-715.
17. Zhang, P., Powell R., Deng Y. Interlacing Bypass Rings to Torus Networks for More Efficient Networks // IEEE Transactions on Parallel and Distributed Systems. 2011. 287-295.
18. Zhang p., Deng Y. An Analysis of the Topological Properties of the Interlaced Bypass Torus (iBT) Networks // Appl. Math. Lett. 25. 2147-2155.
19. Powell R. Analysis of Supercomputers and Development of a Novel Network. PhD Thesis. Stony Brook: Stony Brook University. 2010.
20. Moukarzel C.F. , de Menezes M. A. Shortest paths on systems with power-law distributed long-range connections // Phys. Rev. 2002. E65. 056709-1-056709-9.
21. Sen P., Chakrabarti B. Small-world phenomena and the statistics of linear polymer // J. Phys. 2001. A34. 7749-7755.
22. Petermann T., De Los Rios P. Physical realizability of small-world networks // Phys. Rev. 2006. E73. 026114-1-026114-4.
23. Watts D. J., Strogatz D. H. Collective dynamics of small-world networks // Nature. 1998. 393. 440-442.
24. Albert R., Barabasi A.-L. Statistical mechanics of complex networks // Rev. Mod. Phys. 2002. 74. 47-97.
25. Kleinberg J. M. Navigation in the small world // Nature. 2000. 406. 845.
26. Ciciani B., Colajanni M., Paolucci C. Performance evaluation of deterministic wormhole routing in k-ary n-cubes // Parallel Comput. 1998. 24. 2053-2075.
27. Sarbazi-Azad H., Khonsari A., Ould-Khaoua M. Analysis of k-ary n-cubes with dimension-ordered routing // Future Generation Computer Systems. 2003. 19. 493-502.
28. Alzeidi N., Ould-Khaoua M., Khonsari A. A new modelling approach of wormhole-switched networks with finite buffers // The International Journal of Parallel, Emergent and Distributed Systems. 2008. 23, N 1. 45-57.

29. *Xu J.* Topological Structure and Analysis of Interconnection Networks. Dordrecht-Boston-London: Kluwer Academic Publishers, 2001.
30. *Motter A. E., Lai Y.-C.* Cascade-based attacks on complex networks. // Phys. Rev. 2002. E66. 065102-1–065102-4.
31. *Dobson I., Carreras B., Newman D.* Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization // Chaos. 2007. 17. 026103-1–026103-13.