

Сравнительный анализ алгоритмов построения больших коммуникационных сетей со свойствами “малого мира”

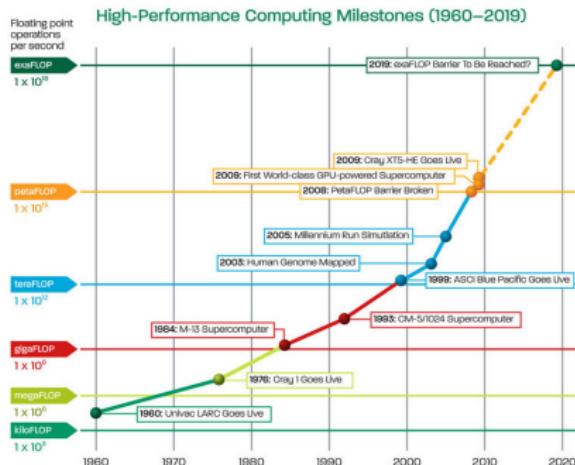
А. Демичев, В. Ильин, А. Крюков, С.Поляков

НИЦ “Курчатовский институт” и НИИЯФ МГУ

Тематический семинар НИЦ "Курчатовский институт"
“Развитие информационно-коммуникационных технологий и
систем, стратегических компьютерных технологий и
программ”

Суперкомпьютеры следующего поколения

- ▶ Появление реальных вычислительных систем следующего поколения - с производительностью порядка экзафлопса (10^{18} FLOPS) - ожидается не ранее 2018–2020 года
- ▶ однако принципы их построения начинают интенсивно разрабатываться уже сейчас
- ▶ на пути к их построению предстоит решить ряд сложных научно-технических задач
- ▶ выработать принципиально новые решения для их архитектуры и аппаратной реализации.



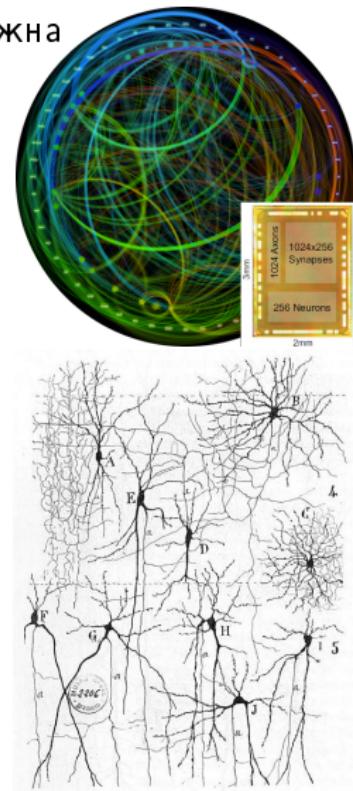
Суперкомпьютеры следующего поколения: коммуникационная сеть

- ▶ Коммуникационная сеть - одна из важнейших составляющих любого суперкомпьютера
 - ▶ в первую очередь определяет возможность увеличения числа вычислительных узлов, что необходимо для достижения желаемой производительности
- ▶ Как ожидается, коммуникационная сеть суперкомпьютеров экзафлопного уровня будет объединять > 100 000 вычислительных узлов ("end points")
- ▶ ⇒ разработка коммуникационных сетей с хорошими свойствами масштабируемости и возможностью эффективно обслуживать огромное число вычислительных узлов



Другой подход к построению СК следующего поколения: нейрокомпьютеры

- ▶ Коммуникационная сеть еще более важна
- ▶ нейроморфные системы: память, распределена в связях между очень простыми процессорами (нейронами)
- ▶ основная нагрузка на выполнение конкретных функций процессорами ложится на архитектуру системы → межнейронные связи
- ▶ Supercomputing 2012: IBM представила результаты работы нейросимулятора Compass - $530 \cdot 10^9$ нейронов и $137 \cdot 10^{12}$ синапсов

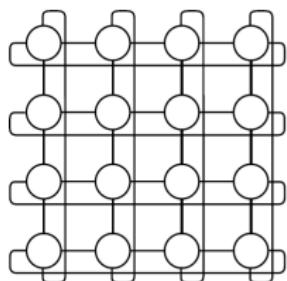


Базовые аспекты функционирования коммуникационных сетей

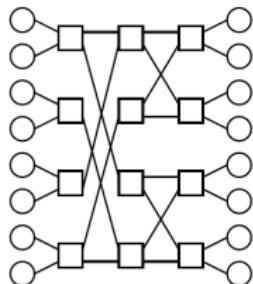
- ▶ Три базовых аспекта = “три кита”, на которых базируется разработка коммуникационных сетей:
 - ▶ топология,
 - ▶ управление потоком,
 - ▶ маршрутизация.
- ▶ в данном сообщении обсуждается, в основном, именно топология коммуникационной сети (в том смысле, в котором термин используется в теории сетей);
- ▶ два других аспекта очень важны, но выходят за рамки текущего обсуждения;
 - ▶ Выбор подходящей топологии жизненно важен для проектирования сети поскольку маршрутизация и механизмы управления потоком в большой степени основаны на ее свойствах

Типы топологий

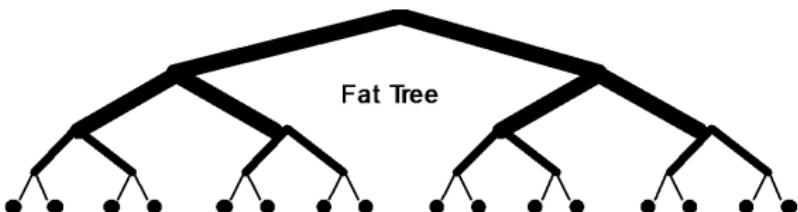
- ▶ прямые сети (direct networks): \forall узел = источник + приемник + рутер (пример (a)); непрямые сети (indirect networks): \exists узлы = рутеры, только для маршрутизации (пример (b)); еще пример непрямой сети - топология типа "толстое дерево"



(a) Torus (4-ary 2-cube)



(b) Butterfly (2-ary 3-fly)



Выбор топологии сети для дальнейшего исследования

- ▶ Непрямые сети имеют свои достоинства для ограниченного числа узлов, но плохо масштабируются
- ▶ Как видно, обычно коммуникационные сети имеют регулярную структуру; преимущества:
 - ▶ равномерная нагрузка на узлы, высокая устойчивость работы сети в целом при выходе из строя некоторых узлов (существование обходных коммуникационных путей),
 - ▶ гиперкуб достаточно высокой размерности для данного числа узлов обеспечивает малое среднее расстояние между узлами (в смысле графа),
 - ▶ при решении специфических задач моделирования 3D-объектов структура вычислительной задачи оптимальным образом отображается на естественную трехмерную решетку вычислительных узлов.

Недостатки регулярной архитектуры для ЭСК ($\gtrsim 10^5$ выч. узлов)

- ▶ решетки невысокой размерности - средняя длина пути очень велика
 - ▶ наилучшие структуры вычислительных систем по различным критериям функционирования (производительность, надежность и т.п.) обладают минимальным средним расстоянием между узлами
- ▶ решетки высокой размерности (гиперкубы), сравнимой с \log числа узлов, - трудно реализовать технически + большая длина физических коммуникационных каналов
- ▶ сети типа толстых деревьев - большое число коммутаторов + необходимость очень высокой пропускной способности каналов у "корня" дерева
- ▶ поэтому обычные сети с простой структурой регулярных решеток окажутся недостаточно эффективными для решения задач более общего типа, не связанных с триангуляцией D -мерных (в частности, трехмерных) объектов

Задача проекта

- ▶ разработать подход к построению и оптимизации коммуникационных сетей более общего вида — с неоднородной структурой
- ▶ в настоящее время проект находится в начальной стадии:
 - ▶ формулировка принципов подхода
 - ▶ выбор методов построения и исследования архитектуры сетей
 - ▶ разработка инструментария для численного моделирования
 - ▶ предварительные первые результаты
- ▶ частично финансируется РФФИ, грант 12-07-00408-а.

Исходные принципы подхода

- ▶ Сеть должна обладать свойствами “малого мира”, а именно, с медленным (логарифмическим) ростом среднего расстояния между узлами с ростом их числа
 - ▶ \Rightarrow масштабируемость
- ▶ Сеть должна иметь базовую структуру D -мерной регулярной решетки
 - ▶ Например, для моделирования D -мерных объектов
- ▶ Эта базовая структура должна иметь дополнительные связи таким образом, чтобы могли достаточно оптимальным образом решаться и общие задачи \rightarrow малая средняя сетевая длина коммуникаций, оптимальная навигация и маршрутизация по сети
 - ▶ Желательно, чтобы характеристики новых связей зависели от варьируемых параметров для возможности оптимизации

Важность малой средней длины: простая модель потоков сообщений

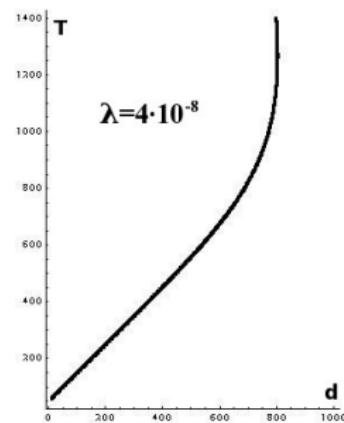
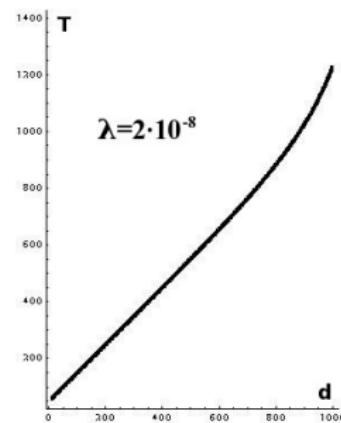
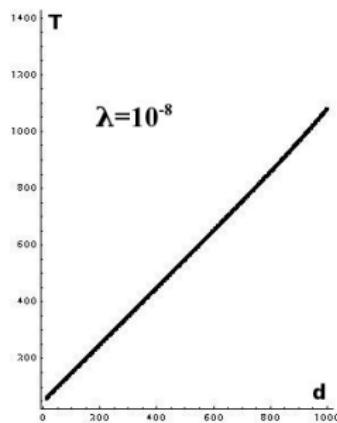
- ▶ Моделируя каждое измерение D -мерного тора как модель системы обслуживания типа $M/G/1$
 - ▶ из-за блокировок каналов при одновременной передаче многих сообщений в сети время обслуживания предполагается произвольным
- ▶ Используя приближение двух моментов в теории массового обслуживания + еще ряд упрощающих приближений
 - ▶ например, random traffic pattern
- ▶ удается получить замкнутое выражение для времени задержки в сети, в частности, для 1D-тора:

$$\bar{T}_1 = M + \bar{k} \left\{ 1 + \frac{\lambda_g \bar{k}^2 \bar{T}_1^3 [1 + (\bar{T}_1 - M)^2 / \bar{T}_1^2]}{8(1 - \lambda_g \bar{k} \bar{T}_1 / 2)} \right\}.$$

- ▶ из-за простоты и существенных приближений модель корректно работает только для достаточно малых значений интенсивности λ_g генерации сообщений каждым узлом

Важность малой средней длины: простая модель потоков сообщений

- модель предсказывает резкий рост временных задержек с ростом средней длины пути для сообщений даже для весьма низких значений λ_g



- Поэтому важно разработать подход к созданию коммуникационных сетей с малой средней длиной (сохраняя решеточную структуру)

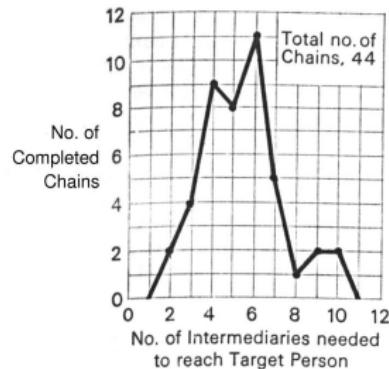
Базовая идея: построить сеть со свойствами “малого мира”

- ▶ Понятие “малого мира” появилась в рамках исследований социальных сетей
- ▶ Классический эксперимент Милграма со знаменитым результатом: “Six degrees of separation” (Milgram, Psych Today 2, 60 (1967)):
- ▶ Он разослал ~ 300 писем с одинаковым содержанием случайно выбранным людям в двух разных городах США (Wichita (Kansas) и Omaha (Nebraska)), в которых сообщалось, что данное письмо должно достигнуть определенного человека в Бостоне (точнее - в Sharon (Massachusetts)), с которым эти люди не были знакомы, и адрес этого человека им не был известен



Базовая идея: построить сеть со свойствами “малого мира” (2)

- ▶ Основное содержание писем:
 - ▶ “If you do not know the target person on a personal basis, do not try to contact him directly. Instead, mail this folder to a personal acquaintance *who is more likely than you* to know the target person.”
- ▶ В эксперименте С. Милграма из 296 писем, 69 достигли цели, что составляет 29% от общего числа.
- ▶ Средняя длина пути оказалась равной 6.2.
- ▶ Этот эксперимент считается первым эмпирическим доказательством «явления малого мира» («small world phenomenon»).



In the Nebraska Study the chains varied from two to 10 intermediate acquaintances with the median at five.

Базовая идея: построить сеть со свойствами “малого мира” (3)

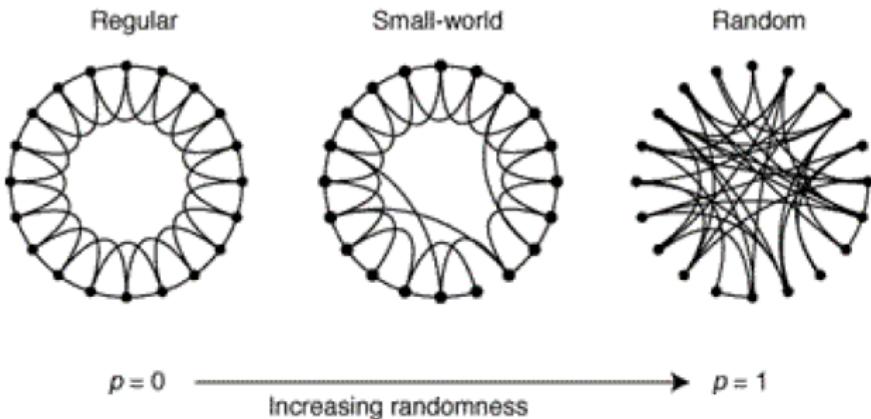
- ▶ В 2003 г. - эксперимент с использованием электронной почты:
 - ▶ 24163 отправителей; получатели - 18 человек из 13 стран;
 - ▶ завершены только 384 (!) цепочки (социальные сети на глобальном уровне весьма сильно разрежены) средняя длина пути ~ 4
- ▶ Простыми словами концепция малого мира означает, что несмотря на большие размеры сети расстояние между двумя узлами может быть малым
- ▶ Более точное выражение этого свойства заключается в следующем: для регулярной D -мерной решетки среднее расстояние d между узлами растет как степень числа узлов: $d \sim N^{1/D}$, а для сети со свойствами малого мира существенно медленнее: $d \sim \ln N$.



Классическая модель “малого мира” (Watts-Strogatz)

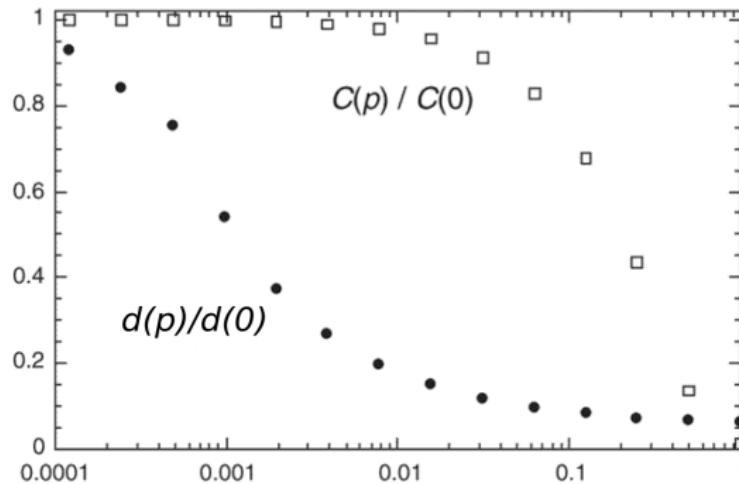
► Алгоритм построения:

- ▶ Задается число узлов N , средняя степень K и вероятность $0 \leq p \leq 1$ ($N \gg K \gg \ln N \gg 1$)
- ▶ Конструируется регулярная решетка с K соседями
- ▶ Для каждого узла берется ребро (n_i, n_j) ($i < j$) и с вероятностью p «перекидывается» на узел n_k (выбранный с равномерной вероятностью), исключая петли (соединения с самим собой) и двойные ребра



WS-модель «малого мира»: основные свойства

- ▶ Основная особенность — сочетание
 - ▶ Высокой степени кластеризации — как у регулярных решеток
 - ▶ Малой средней длины — как у случайных графов
- ▶ В результате алгоритма образуются порядка $pNK/2$ нерегулярных ребер
- ▶ Варьирование p от 0 до 1 осуществляет интерполяцию между решеткой и случайными графиками

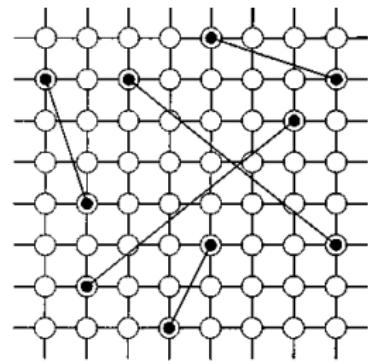
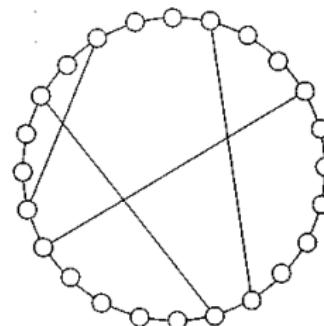


Топология для ЭСК: решеточные сети с перемычками (РСП) - базовая модификация WS-алгоритма

- исходный объект - D -мерная регулярная решетка с топологией тора, содержащая $N = L^D$ узлов ;
- последовательно перебираются все узлы решетки и с вероятностью $0 < p \leq 1$ к каждому узлу i подсоединяют первый конец перемычки;
- второй конец перемычки выбирается случайно с вероятностью

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

- функция решеточного расстояния r между узлами



Замечание: РСП и кластеризация

- ▶ Как отмечалось, помимо малой средней длины пути между узлами еще одним общим отличительным свойством сетей со свойствами малого мира является высокая степень кластеризации.
- ▶ Высокая кластеризация обеспечивает локальную устойчивость сети: существование локальных обходных путей при выходе из строя какого-либо узла сети.
- ▶ в случае РСП с $D > 1$ такую локальную устойчивость (существование локальных обходных путей) обеспечивает решеточная основа
 - ▶ Поэтому в дальнейшем это свойство не обсуждается

Выбор вида распределения длин перемычек: ренормгрупповой анализ

- ▶ Исходная D -мерная решетка разбивается на блоки размером $b : 1 \ll b \ll L$, в результате получается новая решетка (узлы = блоки),
- ▶ узлы которой I и J связаны ребром любая пара $\{i \in I, j \in J\}$ соединена перемычкой
- ▶ естественно предположить, что "блочная" сеть имеет те же свойства, что и исходная; это предположение - аналог ренормгруппового подхода к анализу фазовых переходов.
- ▶ можно показать, что инвариантной относительно такой процедуры является распределение:

$$p(r) = \left(1 - e^{-\rho/r^\alpha}\right) , \quad (1)$$

где параметр ρ преобразуется так

$$\tilde{\rho} = b^{2D-\alpha} \rho . \quad (2)$$

Выбор вида распределения длин перемычек: ренормгрупповой анализ (2)

- ▶ Вероятность иметь перемычку любой длины равна (напомним, что p - вероятность возникновения перемычки из узла)

$$p = S_D \int_1^L p(r) r^{D-1} dr , \quad (3)$$

Поэтому сеть в целом содержит в среднем pL^D перемычек.

- ▶ соотношение (3) позволяет выразить ρ через физически понятный параметр p
- ▶ при малом p инвариантное распределение можно аппроксимировать более простым степенным; в результате:

$$p(r) = \begin{cases} \frac{\alpha-D}{S_D} \frac{p}{r^\alpha} & \text{для } \alpha > D \\ \frac{D-\alpha}{S_d L^{D-\alpha}} \frac{p}{r^\alpha} & \text{для } \alpha < D \end{cases} . \quad (4)$$

Ренормгрупповой анализ: общее поведение сети

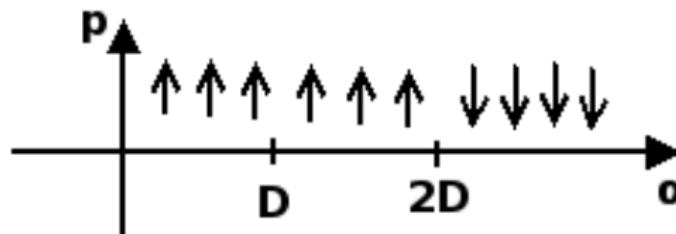
- из преобразования $\rho \Rightarrow$, что p преобразуется при РГ-преобразованиях:

$$\tilde{p} = b^{y_p} p , \quad (5)$$

где

$$y_p = \begin{cases} D & \text{for } \alpha \leq D \\ 2D - \alpha & \text{for } \alpha > D \end{cases} , \quad (6)$$

- Значит $p = 0$ является линией РГ-стационарных точек в пространстве параметров α и p ,
 - при $\alpha < 2D$ эти точки являются неустойчивыми ($b > 1 \Rightarrow b^{2D-\alpha} > 1$ при $\alpha < 2D$),
 - при $\alpha > 2D$ - устойчивыми ($b > 1 \Rightarrow b^{2D-\alpha} < 1$ при $\alpha > 2D$).



Ренормгрупповой анализ: общее поведение сети (2)

- ▶ Это означает, что при $\alpha > 2D$ в результате РГ-преобразований (то есть, огрубления решетки) перемычки исчезают.
- ▶ Другими словами: $\alpha_c = 2D$ является верхней границей, выше которой перемычки в соответствии с распределением (4) (или (1)) являются слишком короткими и не оказывают существенного влияния на свойства сети.

Усредненные характеристики расстояния между узлами в РСП

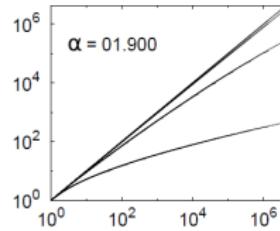
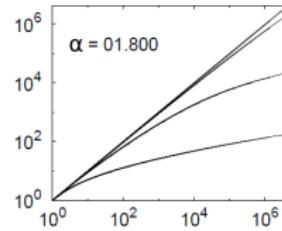
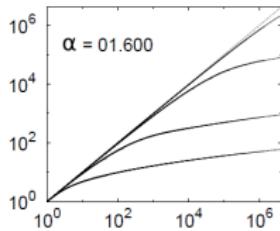
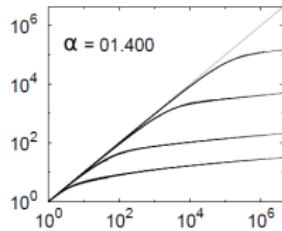
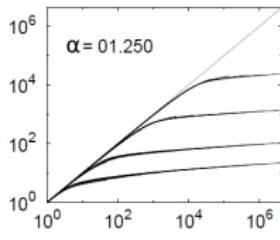
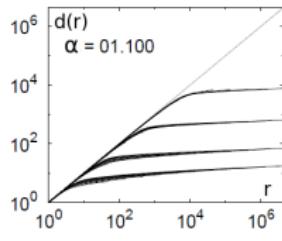
- ▶ При использовании стохастических алгоритмов построения РСП средняя сетевая длина пути между узлами d получается после двойного усреднения:
 - ▶ по стохастическому процессу создания перемычек = по статистическому ансамблю случайных графов \Rightarrow угловые скобки: $\langle \cdot \rangle$;
 - ▶ по всем парам узлов графа:

$$d = \frac{2}{L(L-1)} \sum_{i>j} \langle d_{ij} \rangle .$$

- ▶ Средняя сетевая длина пути между узлами $d(r)$ при данном решеточном расстоянии между узлами $r \Rightarrow$ двойное усреднение, но по узлам усреднение осуществляется при фиксированном r :

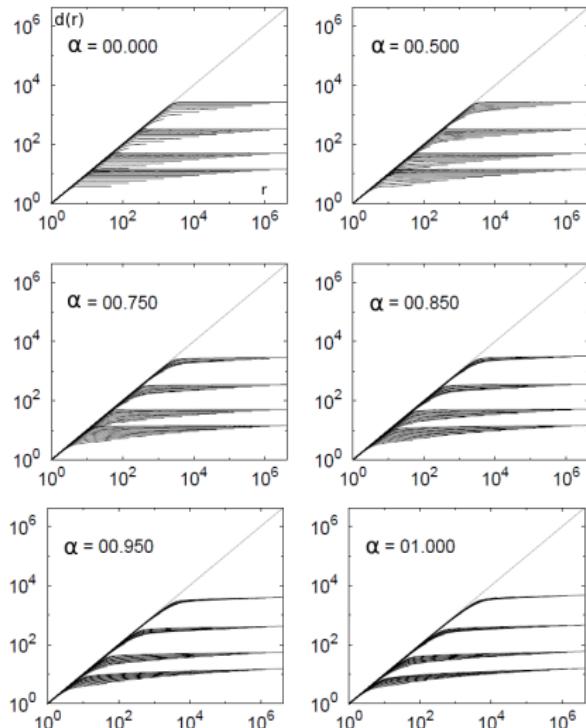
$$d(r) = \frac{\sum_{i,j} \langle d_{ij} \rangle \delta(r_{ij} - r)}{\sum_{ij} \delta(r_{ij} - r)} .$$

Численные результаты для $D = 1$ при $1 < \alpha < 2$



- ▶ $d(r)$ для $1 < \alpha < 2$
- ▶ Плотности перемычек $p = 10^{-3}, 10^{-2}, 10^{-1}, 1$ (чем выше значение p , тем меньше величина $d(r)$ на каждом графике)
- ▶ $\forall \alpha < 2$, d сначала растет линейно с r , потом при некотором характеристическом $r = \xi$ (зависит от p и α для $1 < \alpha < 2$, а также и от L при $\alpha < 1$) начинает расти медленнее - как r^{θ_s} , $\theta_s < 1$.

Численные результаты для $D = 1$: $0 \leq \alpha < 1$



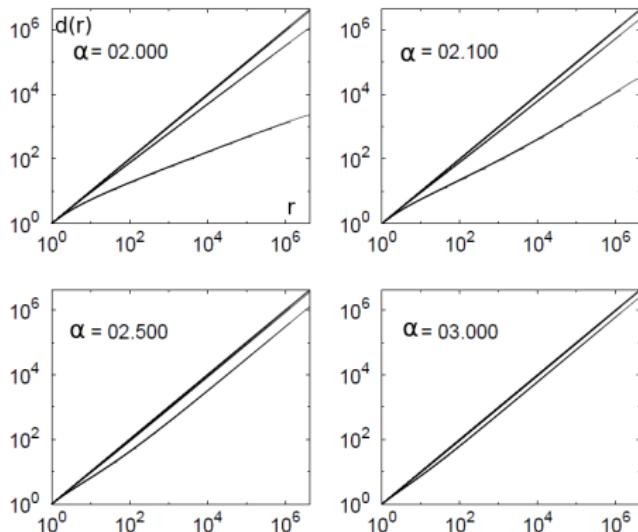
- сетевое расстояние зависит от L
- $\alpha = 0 \Rightarrow$ любые перемычки равновероятны; аналитические вычисления и численные эксперименты:

$$d(r) = \begin{cases} r & r < r_c \\ r_c & r > r_c \end{cases},$$

- предложена эмпирическая формула:

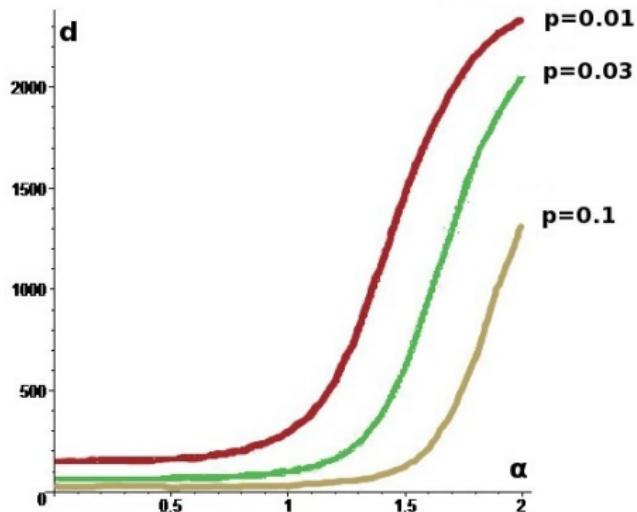
$$r_c = \frac{\log [4(pL)^{(1-\alpha)}]}{p \cdot \log(4)}.$$

Численные результаты для $D = 1$ при $\alpha > 2$



- ▶ Средний кратчайший путь $d(r)$ между узлами, разделенными решеточным расстоянием r для $\alpha > 2$.
- ▶ Большие расстояния; отсутствие свойств малого мира

Зависимость средней длины от параметра распределения длин перемычек $P(r) \sim r^{-\alpha}$ и вероятности перемычек p



- ▶ Увеличение среднего пути между узлами с ростом параметра α
- ▶ Наивный вывод: надо использовать сети с **минимальным** α

Уменьшение степени стохастичности РСП

- ▶ Коммуникационным сетям СК внутренне не присуща стохастичность \Rightarrow может быть удобно использовать модификацию базового алгоритма:
 - ▶ фиксируется не вероятность p появления перемычки из каждого узла, а полное число перемычек t , которые должны быть добавлены к решетке;
- ▶ в базовом алгоритме случайным является как расположение, так и число перемычек;
- ▶ в модифицированном алгоритме случайным является только положение перемычек.
- ▶ Однако:

$$\bigcup_t [\text{пространства событий}] = \bigcup_p [\text{пространства событий}]$$

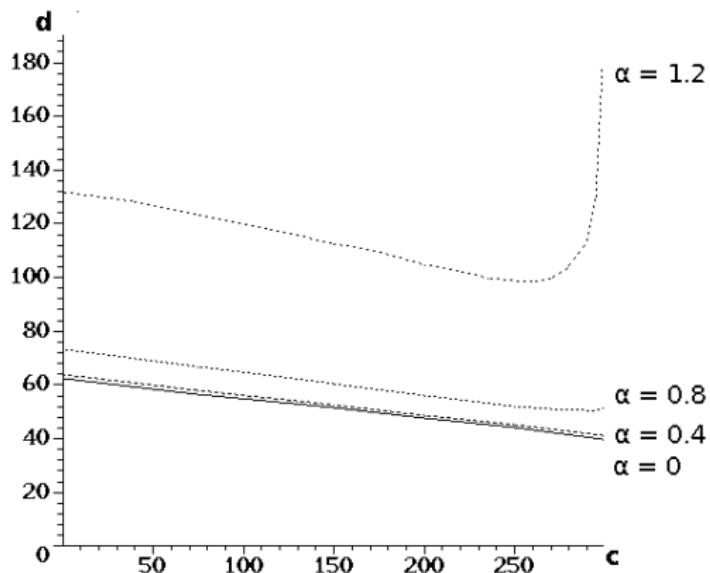
- ▶ \Rightarrow поиск оптимальной РСП, построенной по базовому и модифицированному алгоритмам, при достаточно репрезентативной выборке должен приводить к одинаковым результатам.

Дальнейшее уменьшение степени стохастичности РСП

- ▶ В дополнение к предыдущей модификации:
- ▶ $c \leq t$ перемычек добавляются специальным образом; перемычки добавляются последовательно одна за другой, при добавлении каждой перемычки:
 - ▶ составляется список узлов, к которым уже присоединены перемычки;
 - ▶ из полученного списка узлов случайным равновероятным образом выбирается один и нему подсоединяется новая перемычка;
 - ▶ второй конец дополнительных перемычек выбирается так же как в базовом алгоритме.
- ▶ теперь не только число перемычек детерминировано, но положение их не вполне случайно, степень этой детерминированности определяется параметром c (при $c = 0$ совпадает с предыдущим алгоритмом).
- ▶ В определенной степени этот алгоритм является гибридным: частично стохастическим, а частично детерминистским

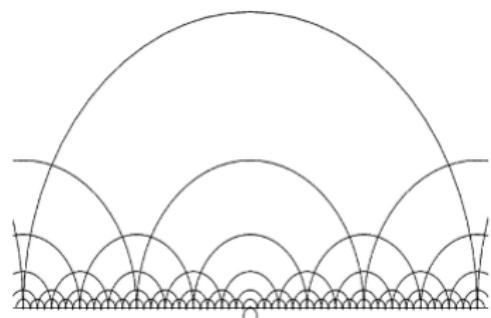
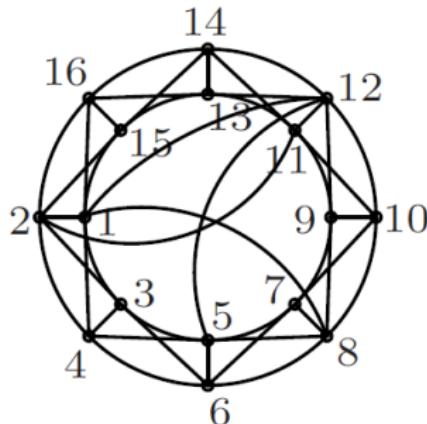
Дальнейшее уменьшение степени стохастичности РСП (2)

- ▶ Среднее расстояние уменьшается с ростом числа заведомо соединенных между собой перемычек
- ▶ → с точки зрения глобальной средней длины этот алгоритм является предпочтительным по сравнению с исходным стохастическим алгоритмом



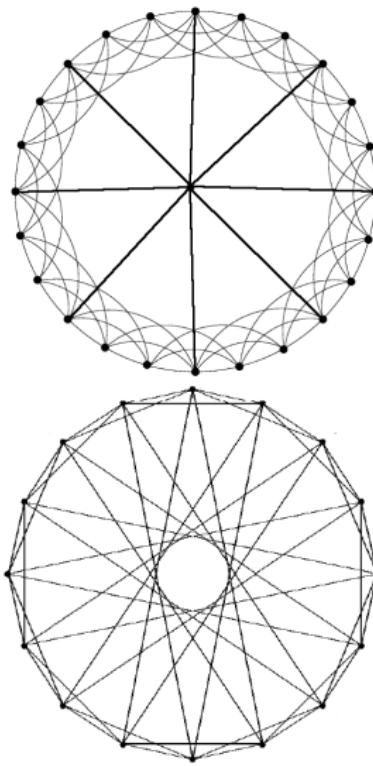
РСП без стохастичности: детерминистские алгоритмы

- ▶ Самый простой и естественный алгоритм: на каждом шаге алгоритма осуществляется соединение перемычкой наиболее удаленных друг от друга в сетевом смысле узлов
- ▶ Иерархический алгоритм HN4 основан на представлении номеров узлов $0 < n < L$:
 $n = 2^i(2j + 1)$



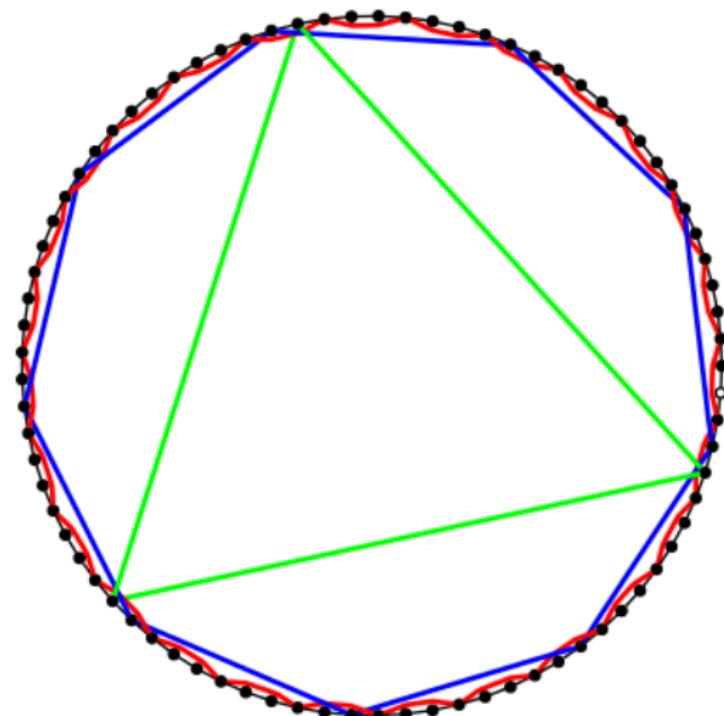
РСП без стохастичности: детерминистские алгоритмы (2)

- ▶ Выделение узлов-хабов, которые соединяются ребрами (перемычками) так, что получается некоторый граф H со специальными свойствами
- ▶ Мультипликативные циркулянтные графы $C(s^k; 1, s, s^2, \dots, s^{k-1})$



РСП без стохастичности: субциркулянтные (хордовые) сети

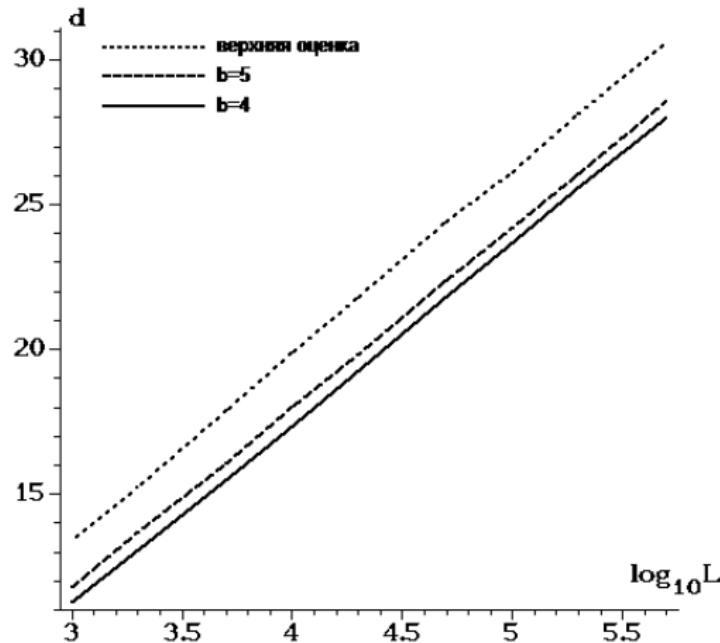
- ▶ узлы
 $0, b, 2b, \dots, L - b$
соединяются
перемычками в цикл
(узел 0 соединяется
с b , b с $2b$ и т.д.,
 $L - b$ с 0);
- ▶ для каждого
 $i = 2, \dots, k$ узлы
 $0, b^i, 2b^i, \dots, L - b^i$
соединяются
перемычками в цикл



Субциркулянтные сети \leftrightarrow малый мир

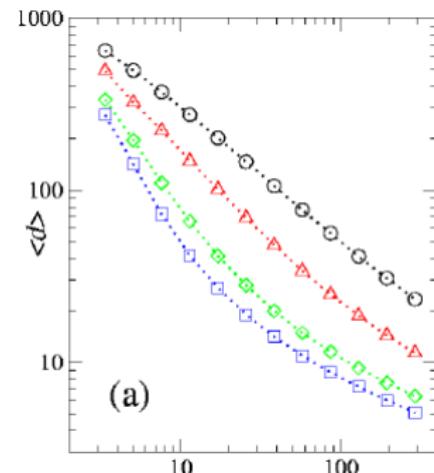
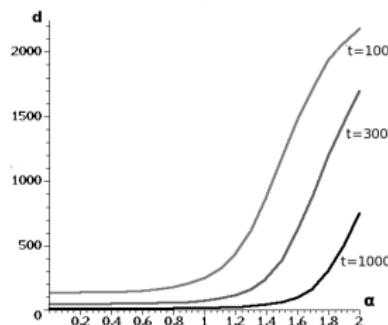
- приближенная верхняя оценка:

$$d \lesssim \lambda \left[\frac{k(b+4)}{2} + \frac{L}{4b^k} - 2 \right] \quad \text{где } \lambda \simeq 1$$



Учет “цены” построения РСП

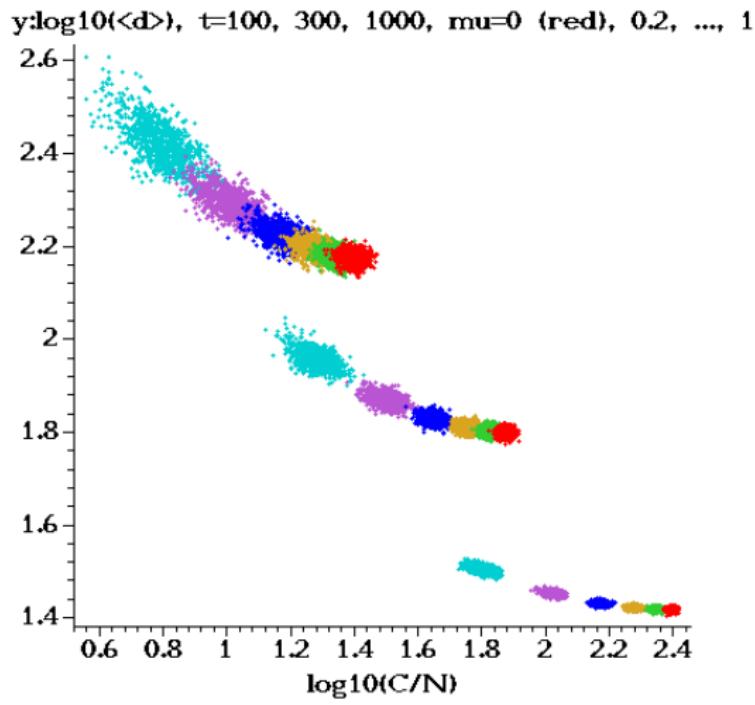
- Необходимо учитывать “цену” построения сети; иначе - тривиальное решение = полный граф
- Естественная функция “цены” - удельная длина перемычек = (суммарная длина)/(число узлов)
- Petermann & De Los Rios (2006): при фиксированном значении C_W наименьшая средняя длина пути уменьшается с ростом α
 - $\alpha = 0$: $\circ \rightarrow 0$, $\blacktriangleright \rightarrow 1$ $\diamond \rightarrow 1.5$, $\square \rightarrow 1.75$
- грубая картина: усреднение по ансамблю + “теоретическая цена” (усреднение по распределению)



(a)

Учет “цены” построения РСП (2)

- ▶ Реальная ситуация еще сложнее
- ▶ $\alpha = 0$ (красные точки),
0.2, 0.4, 0.6, 0.8, 1.0
- ▶ при одинаковой цене для уменьшения d
⇒ max t (или p) и α
 - ▶ пока это возможно для данной удельной длины перемычек
- ▶ большое количество не слишком длинных перемычек является более дешевым решением.



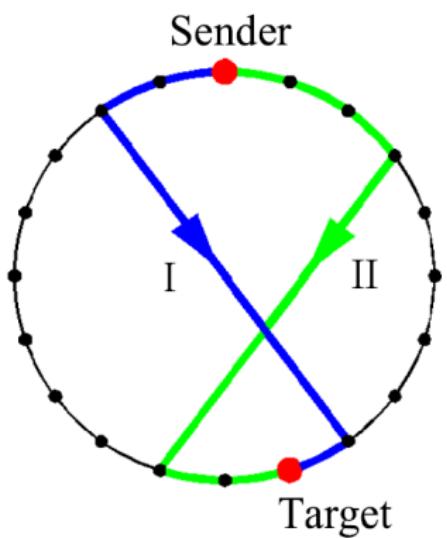
Качественные характеристики сетей

- ▶ Какую качественную характеристику выбрать для оптимизации?
- ▶ Простейший вариант - среднее расстояние между узлами d
- ▶ Для больших сетей d может оказаться неадекватной характеристикой
 - ▶ для нахождения кратчайших маршрутов необходимо знать глобальную структуру сети
 - ▶ маршрутизация сообщений, использующая кратчайшие пути, может оказаться слишком сложной и неэффективной, так как связана с хранением и обработкой большого объема информации
- ▶ Возможное решение - локальная навигация
 - ▶ соответствующая качественная характеристика - навигационная длина ℓ

Динамические свойства малого мира на решетке: навигация при наличии только локальной информации

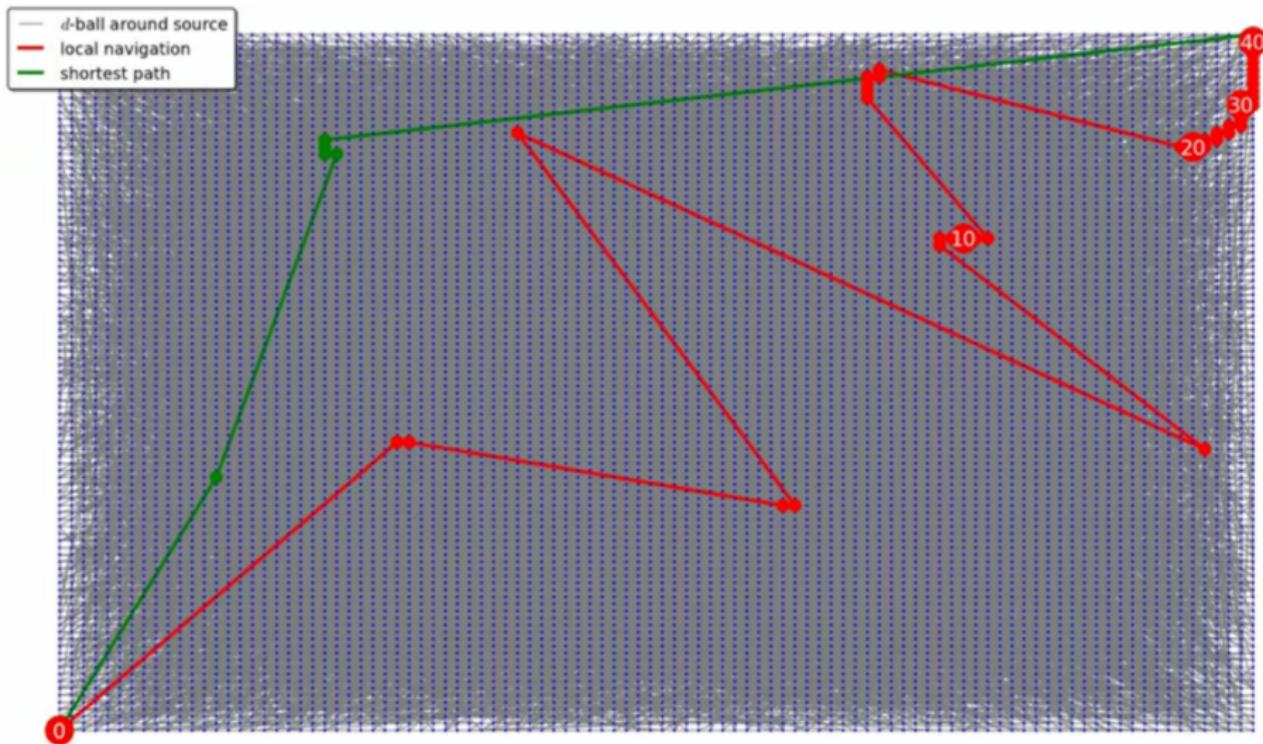
- ▶ Пионерскими статьями в этой области являются работы Клейнберга (2000 г.), а сама постановка задачи и экспериментальные результаты восходят к классическому эксперименту Милграма (1967 г.)
- ▶ Задача навигации в пространственно-вложенных сетях
 - ▶ сообщение передается от узла к узлу по ребрам (связям);
 - ▶ узел "знает" географическое положение (другими словами, положение в оставной решетке) всех узлов и своих ближайших сетевых соседей;
 - ▶ необходимо доставить сообщение в узел назначения по возможно кратчайшему пути.
- ▶ Эту задачу решает **жадный алгоритм** (*greedy algorithm*; также: *алгоритм экономного продвижения*): данный узел пересыпает сообщение тому своему соседу, который географически ближе всего к цели (узлу назначения).

Отступление: простой пример, когда глобально кратчайший путь не совпадает с путем при навигации с локальной информацией

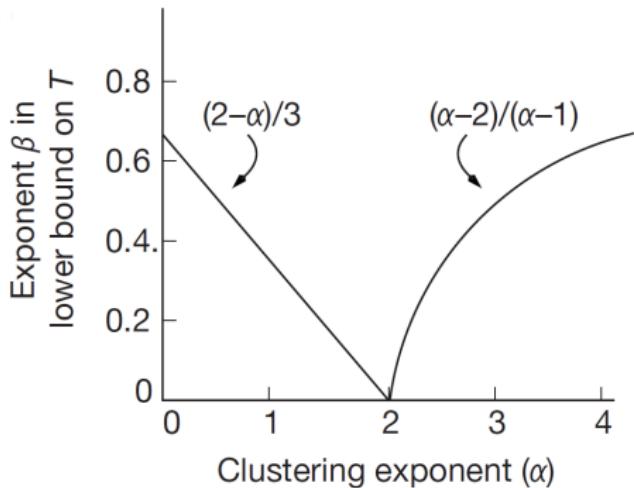


- ▶ Простой пример, когда знание глобальной (всей) информации о сети позволило бы передать сообщение адресату по более короткому пути, чем в случае знания только локальной информации. При знании глобальной информации сообщение можно передать за 5 шагов по пути I. Знание только локальной информации с $v = 1$ (возможны обобщения на $v = n$) и жадный алгоритм позволяет передать сообщение за 6 шагов (путь II).

Отступление: более реалистичный пример

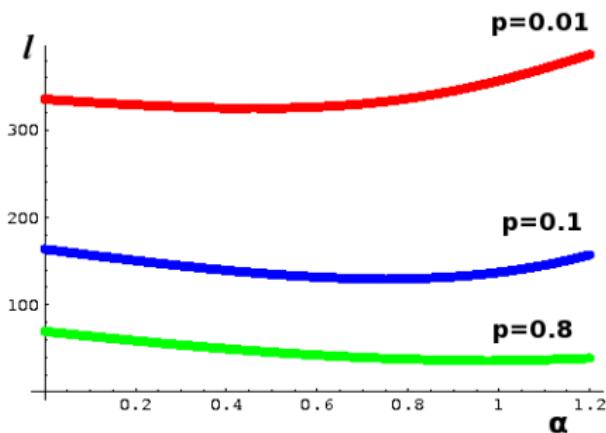


Отступление: основной результат Клейнберга (2000)



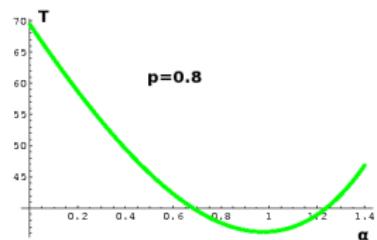
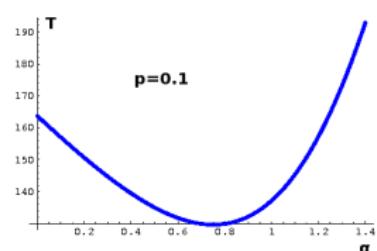
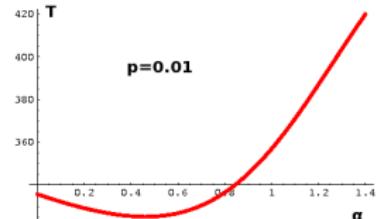
- ▶ Нижняя граница для β в $\bar{T} \sim L^\beta$ ($N = L^2$) в зависимости от α ($p = 1$)
- ▶ Наиболее важным выводом работы Клейнберга является то, что оптимальным значением показателя распределения длин перемычек является $\alpha = D$, которое приводит к логарифмической зависимости времени доставки от размера сети:
$$\bar{T} \sim \log^2 N.$$

Отступление: обобщение результата Клейнберга



- ▶ Используя слегка упрощенный алгоритм локальной навигации (Moukarzel & Menezes, Phys. Rev. E, 65:056709, 2002), в нашей работе показано, что для $p \neq 1$ положение минимума средней навигационной длины зависит от p и достигается при $\alpha_{min} < D$
- ▶ Для $D = 1$ показаны графики зависимости навигационной длины от α для 3-х значений p

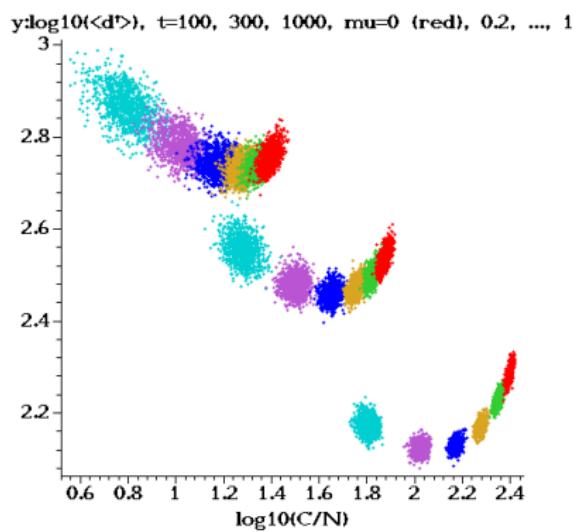
Отступление: обобщение результата Клейнберга (2)



- ▶ Для наглядности графики с предыдущего слайда приведены по отдельности
- ▶ Наглядно видна зависимость минимума навигационного расстояния от p
- ▶ Численное моделирование подтверждает эти результаты

Зависимость навигационной длины от “цены” для стохастических РСП

- ▶ При фиксированном числе перемычек t существует минимум ℓ , достигаемый при различных значениях α , зависящих от t
- ▶ “цена” \iff параметры алгоритма с наименьшей навигационной длиной
- ▶ \Rightarrow надо решать оптимизационную задачу “цена – качество”.

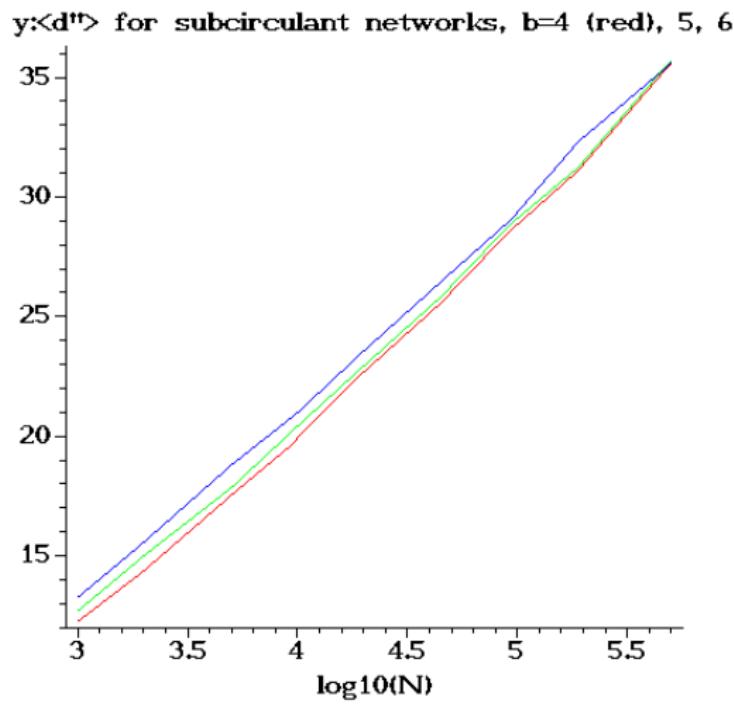


Навигационная длина пути в иерархических РСП

- ▶ Иерархические РСП (численное моделирование): между узлами
 - ▶ весьма малое среднее расстояние d
 - ▶ большой навигационной средней длиной пути ℓ
- ▶ \Rightarrow модифицированный алгоритм локальной навигации:
 - ▶ двухуровневая локальная навигация: просматриваются не только ближайшие соседи, но и соседи соседей
 - ▶ объем вычислений несколько увеличивается, но алгоритм остается локальным и хорошо масштабируемым на большие сети
- ▶ 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: $\ell^{(\infty)}$

Навигационная длина пути в иерархической субциркулянтной РСП

- ▶ Результаты численного моделирования для $\ell^{(2)}$ при двухуровневой локальной навигации
 - ▶ зависимость $\ell^{(2)}$ от \log числа вершин с $b = 4, 5, 6$ и k , минимизирующими $\ell^{(2)}$ для выбранных L



Многокритериальная оптимизация “цена – качество”

- ▶ Оптимизация с векторной целевой функцией
- ▶ В нашем случае — два показателя:
 - ▶ “качество” РСП = среднее расстояние d или ℓ или $ell^{(2)}$
 - ▶ “цена” построения РСП = удельная длина перемычек C/L
- ▶ метод взвешенных сумм \Rightarrow скалярные целевые функции

$$G_w = wd + (1 - w)C/L , \quad (7)$$

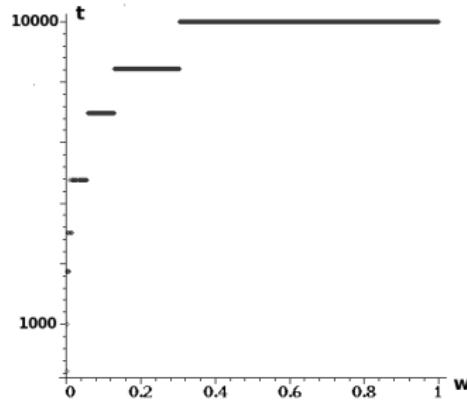
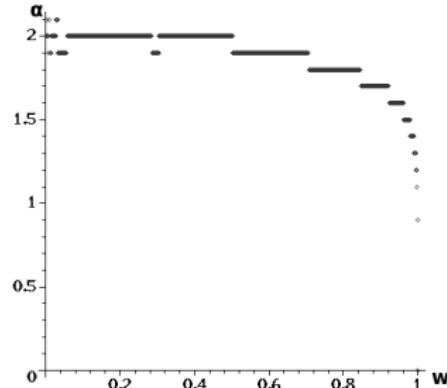
$$G'_w = w\ell + (1 - w)C/L , \quad (8)$$

$$G''_w = w\ell^{(2)} + (1 - w)C/L . \quad (9)$$

- ▶ параметр $0 \leq w \leq 1$ характеризует относительную значимость каждого из критериев (качество и цена)
- ▶ оптимизация = $\forall w$ подобраны значения параметров алгоритмов ($p/t, \alpha, c, b, k, \dots$): $\min G_w, G'_w$ или G''_w
 - ▶ размер сети L считаем заданным

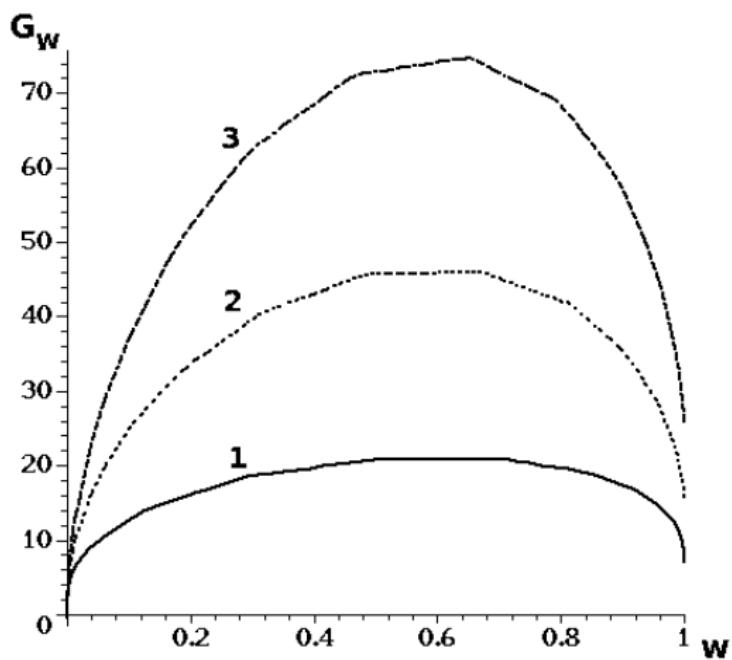
Оптимальные параметры для стохастических РСП

- ▶ параметры минимизации G_w
- ▶ при $w \lesssim 0.5$ (доминирует цена) $\Rightarrow \alpha \sim 2$ (критическое значение)
- ▶ при $w \gtrsim 0.35 \Rightarrow$ число перемычек должно быть максимально возможным ($= L$);
 - ▶ при $w \lesssim 0.35 \Rightarrow$ важность минимизации длины перемычек \rightarrow уменьшение числа перемычек (значения t).



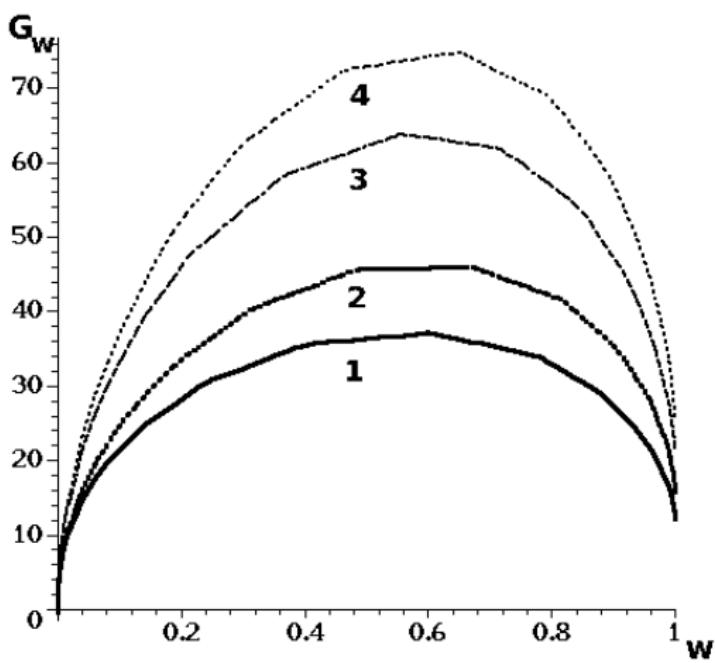
Минимизация целевых функций для стохастического алгоритма

- ▶ Каждая точка → ансамбль сетей с t, α ($L = 10^4$): минимум G_w (1), G'_w (3) и G''_w (2)
- ▶ $\forall w \min G_w \leq \min G''_w \leq \min G'_w$
 - ▶ больше информации о структуре сети (большая глубина просмотра) \Rightarrow меньшее взвешенная сумма



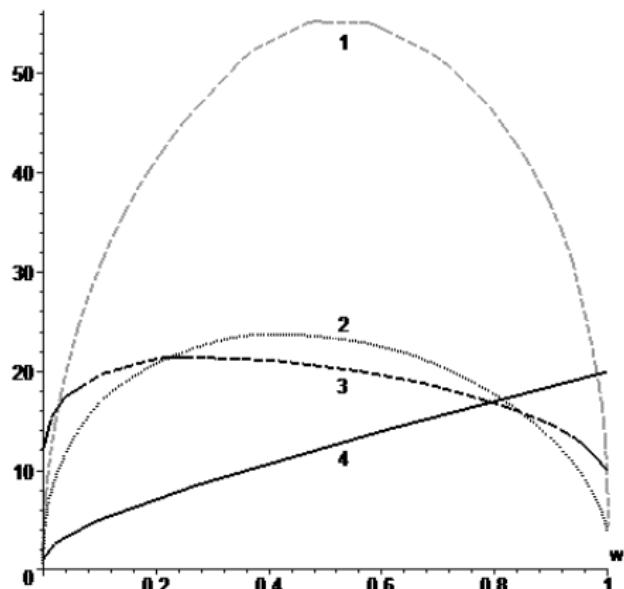
Сравнение стохастических алгоритмов

- ▶ 1 — $\min G''_w$, 3 — $\min G'_w$ для алгоритма с соединенными перемычками
- ▶ 2 — $\min G''_w$, 4 — $\min G'_w$ для базового алгоритма
- ▶ ➡ в классе стохастических предпочтительным является алгоритм с соединенными перемычками



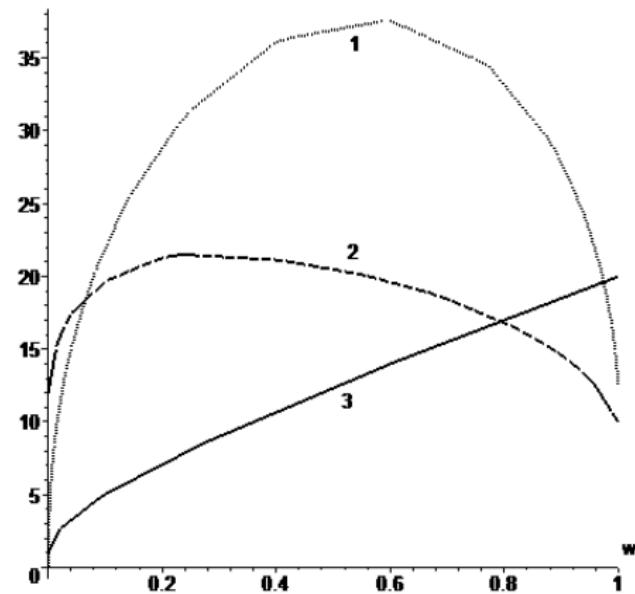
Сравнение детерминистских алгоритмов

- ▶ 1 — $\min G_w$ для мультипликативных циркулянтных графов
- ▶ 2 — $\min G_w$ для сетей со звездообразным, а 3 — с двухпетлевым соединением хабов
- ▶ 4 — $\min G''_w$ для субциркулянтных сетей
- ▶ \Rightarrow предпочтительными являются субциркулянтные сети



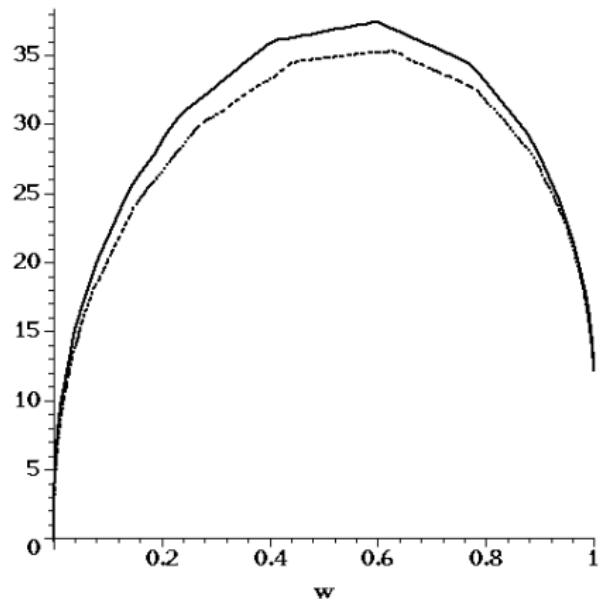
Сравнение детерминистских и стохастических алгоритмов

- ▶ 1 — $\min G_w''$ для стохастических сетей с соединенными перемычками
- ▶ 2 — $\min G_w$ для детерминистских сетей с двухпетлевым (double-loop) соединением хабов
- ▶ 3 — $\min G_w''$ для субциркулянтных сетей
- ▶ → предпочтительными являются субциркулянтные сети



Выбор экземпляра в стохастическом ансамбле

- ▶ использовалось сравнение характеристик, усредненных по ансамблям
- ▶ экземпляр с наилучшими характеристиками?
- ▶ Численное моделирование:
 - ▶ сплошная — для средних по ансамблю
 - ▶ прерывистая — экземпляры в выборке из 100: 5-е в порядке возрастания среднее расстояние
- ▶ ⇒ большого выигрыша нет



Заключение

- ▶ При использовании стандартных (регулярных) топологий для построения коммуникационных сетей компьютеров следующего поколения (Exascale Supercomputers, ESC) могут возникнуть недопустимо большие временные задержки при передаче сообщений
 - ▶ ⇒ необходимость разработки новых подходов к построению коммуникационных сетей
- ▶ В данной работе предложено использовать сеть с нерегулярной структурой
 - ▶ малая средняя длина пути между вычислительными узлами достигается минимальным добавлением новых связей к регулярной решетке
 - ▶ сохраняется оптимальная структура для моделирования D -мерных объектов

Заключение (2)

- ▶ Рассмотрен ряд как известных в литературе, так и оригинальных алгоритмов построения сложных сетей со свойствами “малого мира”
- ▶ Предложена методика сравнения эффективности алгоритмов различных типов на основе оптимизации соотношения “цены” и “качества”
 - ▶ В качестве “цены” выступает удельная длина дополнительных перемычек (общая длина перемычек в единицах базовой решетки, деленная на число узлов сети)
 - ▶ “качество” — это глобальная или локально-навигационная средняя длина пути между узлами.
- ▶ Двумя основными классами алгоритмов являются стохастические и детерминистские алгоритмы ⇒ сравнение на основе соотношения “цена – качество”
- ▶ Предпочтительным типом сетей в широком диапазоне параметра относительной значимости указанных критериев являются субциркулянтные сети, которые строятся с помощью детерминистского алгоритма

Дальнейшие планы

- ▶ Рассмотрение свойств предложенных в данной работе сетей, как с точки зрения других характеристик (в частности, нагрузки на узлы, устойчивости и т.п.)
- ▶ Обобщения на более высокие размерности базовой решетки.
- ▶ Построение модели маршрутизации потоков сообщений; разработка
 - ▶ ⇒ аналитических моделей
 - ▶ ⇒ программного инструментария для численного моделирования потоков в решеточных сетях с перемычками и свойствами малого мира
- ▶ Разработка конкретного алгоритма построения оптимальной коммуникационной сети ЭСК в классе решеточных сетей со случайными перемычками