

Измеряем океан мензуркой (оценка характеристик глобальной сети)*

Александра Коконова, Алексей Городилов, Russian Federation

7 августа 2019 г.

Аннотация

The experiment, aimed at finding the distribution of path lengths between nodes in the global network and estimation of parameters of that distribution is described. At particular, described the method of measurement of path length from node α to the node β with traceroute utility of the GNU/Linux system, and limitations on the selection of nodes α and β , imposed by traceroute. Measurements results are given and analyzed. Simulation model of this experiment was developed to test the experiment validity in the determination of distribution parameters in the global network. This model is also described. It is shown that the global network could not be described by Barabási-Albert model, but it could be described with some approximation by developed and proposed a pre-fractal model.

Введение

Разработка эффективных сетевых приложений невозможна без сведений о структуре сети, в которой они будут работать. Так, в 2011 году, для получения параметров методики МУЧОС [1] передачи мультимедийных данных с учётом особенностей сети, был начат эксперимент по оценке распределения длин путей сетевого уровня модели TCP/IP глобальной сети.

Определим понятие длины пути и распределения длин путей. Из-за свойств стека протоколов TCP/IP, если от узла α к узлу β

*illinc@bk.ru, <https://lvee.org/???>

в принципе возможна передача данных, маршрут передачи данных на сетевом уровне будет единственным. Пусть это $\alpha \rightarrow \alpha_1 \rightarrow \dots \rightarrow \alpha_{k-1} \rightarrow \beta$.

Тогда длина $s(\alpha, \beta)$ пути от узла α к узлу β — количество k пересылок узел-узел в этом маршруте. При этом $s(\alpha_{k-1}, \beta) = 1, \dots, s(\alpha_1, \beta) = k - 1$.

Со временем глобальная сеть изменяет свою структуру, и как маршрут передачи данных, так и его длина $s(\alpha, \beta)$ могут изменяться. Кроме того, из-за особенностей маршрутизации может быть $s(\alpha, \beta) \neq s(\beta, \alpha)$.

Таким образом, под распределением длин путей графа G понимается ряд чисел $\zeta(x)$, показывающий, как часто длина x встречается среди всех возможных путей в G :

$$\zeta(x) = p(s(\alpha, \beta) = x \mid \alpha, \beta \in G).$$

Непосредственно измерить распределение длин путей $\zeta(x)$ для глобальной сети (то есть измерить и проанализировать длины $s(\alpha, \beta)$ для всех возможных пар (α, β)) невозможно как из-за её гигантских размеров, так и из-за отсутствия в стеке TCP/IP средств для измерения расстояния между двумя произвольно взятыми узлами.

Выбираем мензурку

Для исследования маршрутов сетевого уровня в GNU/Linux предназначена утилита `traceroute`. При запуске в командной строке узла α команды `traceroute` β результатом будет маршрут передачи данных $\alpha \rightarrow \alpha_1 \rightarrow \dots \rightarrow \alpha_{k-1} \rightarrow \beta$, причём одна строка вывода `traceroute` соответствует одному узлу маршрута. Далее путём синтаксического анализа можно определить длину пути $s(\alpha, \beta)$.

Таким образом, для измерения длины пути $s(\alpha, \beta)$ необходимо иметь право на запуск программ на узле α (в дальнейшем будем называть его корневым, рис. 1) и знать IP-адрес или имя узла β (будем называть его конечным). Также необходима возможность соединения α с β , то есть конечный узел β должен иметь белый IP-адрес либо находиться в одной подсети с α .

Ограничения на выбор конечного узла гораздо мягче, чем для корневого. Соответственно, количество возможных конечных узлов может быть намного больше (множество конечных узлов $Q = \{\beta_1, \beta_2, \dots, \beta_n\}$ на рис. 1).

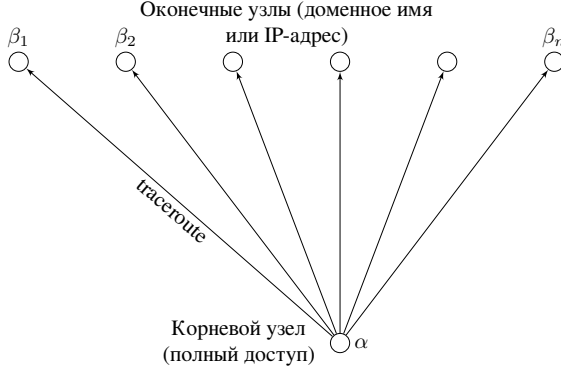


Рис. 1: Схема измерений выборочных длин путей

Соответственно, задавшись парой из корневого узла α и множества конечных узлов Q и измеряя при помощи traceroute расстояния $s(\alpha, \beta_j)$ от α до каждого из конечных $\beta_j \in Q$, можно получить некоторую оценку распределения длин путей $\zeta(x)$ глобальной сети. Эта оценка зависит от выбора корневого узла, от множества конечных узлов и от времени (от текущей конфигурации глобальной сети).

В ходе этого эксперимента были сформированы три различных множества конечных узлов (P, R, T размерами 564, 5222 и 540 уникальных белых IP-адресов соответственно) и три корневых узла (A, H, M).

Так как все конечные узлы имели белые IP-адреса, а корневые — серые, каждый маршрут передачи данных $\alpha \rightarrow \alpha_1 \rightarrow \dots \rightarrow \alpha_{k-1} \rightarrow \beta$ от корневого узла α до конечного узла β делился на две части: маршрут от корневого узла α до шлюза α_{gw} его провайдера $\alpha \rightarrow \dots \rightarrow \alpha_{gw}$ и маршрут от шлюза провайдера корневого узла до конечного узла $\alpha_{gw} \rightarrow \dots \rightarrow \alpha_{k-1} \rightarrow \beta$.

Для каждой пары корневого узла α и конечного узла β получались два значения длины пути: полная длина $\tilde{s}(\alpha, \beta) = k$, включающая подсеть провайдера корневого узла и скорректированная длина $s(\alpha, \beta) = k - gw$, исключая подсеть провайдера корневого узла (то есть включающая только пересылки между узлами с белыми IP-адресами).

После окончания измерений множества полученных значений

длин путей \tilde{s} и s обрабатывались GNU Octave. Соответственно, для каждой пары корневого узла α и множества оконечных узлов Q получались две оценки распределения длин путей: $\tilde{\zeta}(x)$, включающая подсеть провайдера корневого узла и $\zeta(x)$, исключающая подсеть провайдера корневого узла.

Оцениваем

От корневых узлов A и M было выполнено по четыре разнесённых во времени измерения расстояний до каждого из наборов P, R, T ; от узла H , из-за ограниченного времени доступа — только по одному измерению до каждого из наборов P, R, T . Таким образом, всего было выполнено 27 измерений и получено 27 оценок $\zeta(x)$ распределения длин путей, включающих подсеть провайдера корневого узла (рис. 2, а) и 27 оценок $\zeta(x)$ распределения длин путей, исключающих подсеть провайдера корневого узла (рис. 2, б).

В соответствии с множеством оконечных узлов, корневым узлом и порядковым номером измерения обозначены как $PA_0, PA_1, PA_2, PA_3, PH_0, PM_0, \dots, PM_3, RA_0, \dots, RM_3, TA_0, \dots, TM_3$ и перечислены на рис. 3 именно в таком порядке.

Рис. 3, а) показывает разброс характеристик (моды Mo , среднего μ , среднеквадратического отклонения σ , асимметрии A и эксцесса E) оценок $\zeta(x)$ распределения длин путей, включающих подсеть провайдера корневого узла. Ось абсцисс соответствует номеру измерения, как сказано выше; границы между измерениями, соответствующими множествам оконечных узлов P, R, T показаны линиями сетки. Слева от оси ординат разброс показан в виде ящика с усами (усы соответствуют минимальному и максимальному значениям, края ящика — первой и третьей квартилям, линия внутри ящика — медиане, кружок — среднему значению характеристики).

Рис. 3, б) аналогично показывает разброс характеристик оценок $\zeta(x)$ распределения длин путей, исключающих подсеть провайдера корневого узла.

Значение эксцесса E здесь и далее вычисляется по формуле с коррекцией на три (так, чтобы эксцесс нормального распределения был бы равен нулю).

Видно, что значения асимметрии и эксцесса оценок распределения длин путей (как $\tilde{\zeta}(x)$, так и $\zeta(x)$) существенно больше нуля.

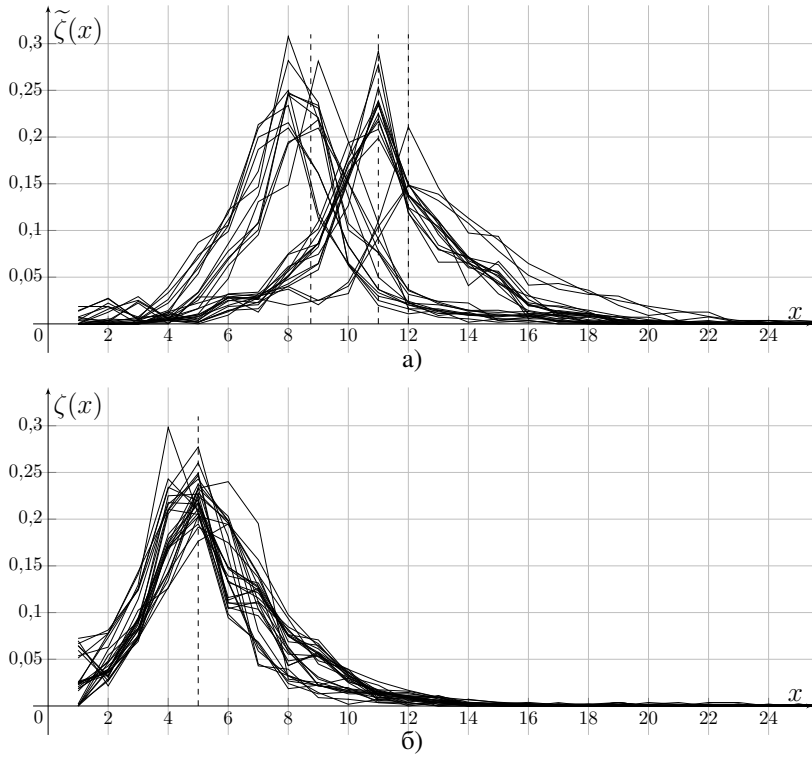


Рис. 2: Экспериментально полученное распределение длин путей: а) включая подсеть провайдера корневого узла, б) исключая подсеть провайдера корневого узла

Поверяем

Насколько корректны полученные оценки $\zeta(x)$? Для их проверки было сгенерировано некоторое количество графов G , размеры которых позволяют:

- рассчитать расстояние $s(\alpha, \beta)$ для всех возможных пар узлов $\alpha, \beta \in G$ и рассчитать истинное (полное) распределение длин путей $\zeta(x)$;
- выполнить имитацию эксперимента и оценить $\zeta(x)$ по заданным корневному узлу r и множеству конечных узлов Q .

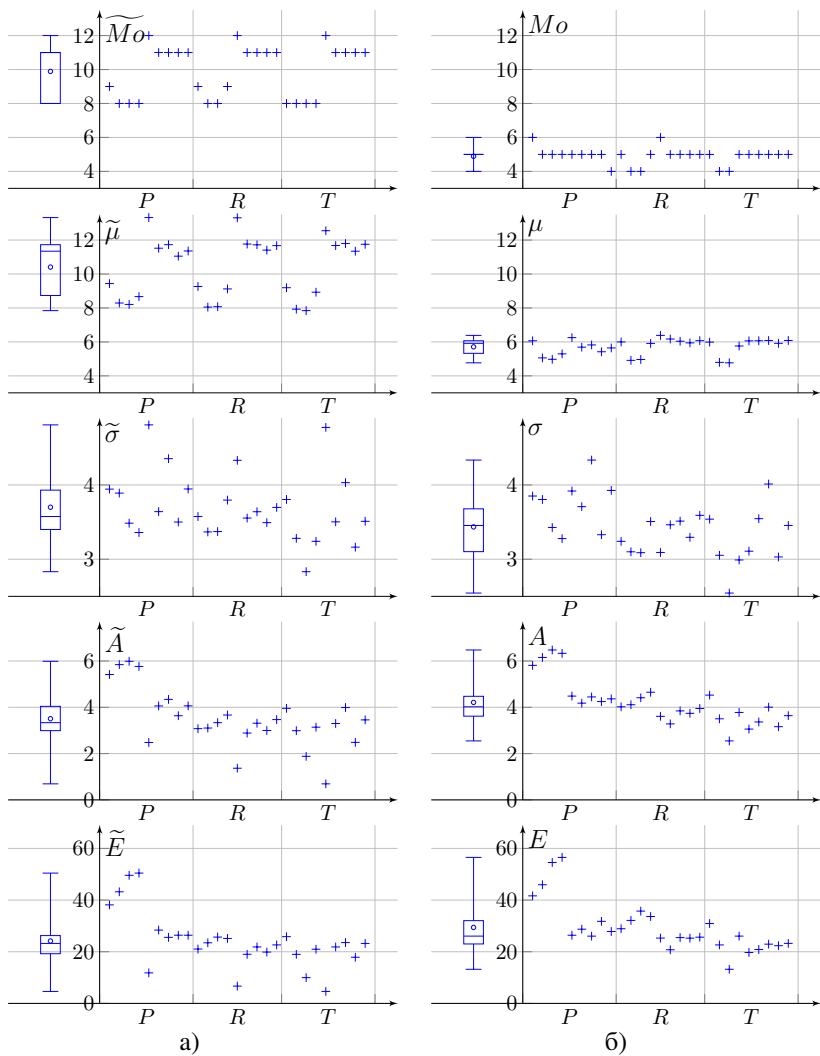


Рис. 3: Характеристики экспериментально полученных оценок распределения длин путей: а) включая подсеть провайдера корневого узла, б) исключая подсеть провайдера корневого узла

Размеры множеств Q , использованных в экспериментальном исследовании, составляют 500 – 5000 узлов. Соответственно, минимальным количеством вершин в G в расчётах было $|G| = 10^3$ узлов. Максимальным размером G , при котором можно выполнить расчёт менее чем за месяц – $|G| = 10^6$ узлов.

Граф G из $|G|$ вершин представлялся в памяти компьютера в виде сжатой матрицы смежности, то есть вектором ρ_G из $|G|$ списков, каждый из которых соответствует вершине графа и содержит номера смежных с ней вершин.

Для расчёта длин путей использован поиск в ширину, позволяющего для заданной вершины $\alpha \in G$ рассчитать расстояния до всех прочих вершин графа G . Таким образом, результат соответствует экспериментальным измерениям для корневой вершины α и множества конечных вершин $Q = G$, включающего все вершины графа. Сложность поиска в ширину для представления графа в виде ρ_G равна $O(|G| + |E_G|)$, где $|E_G|$ – количество рёбер [2]. Таким образом, при расчёте полного распределения длин путей все вершины $\alpha \in G$ графа поочерёдно рассматриваются в качестве корневых, и сложность возрастает в $|G|$ раз. Для ускорения расчёта использован компилируемый язык C++ (стандарт C++11) и распараллеливание OpenMP.

Кроме полного распределения длин путей $\zeta(x)$, для каждого графа G рассчитывались ещё несколько оценок $\zeta(x)$ по разным парам (r, Q) корневой узел-множество конечных узлов. Эти оценки обозначались $\zeta_i(x)$.

Корневая вершина r , как и конечные вершины $\beta_j \in Q$ для каждого i выбиралась случайно. Количество конечных вершин $|Q|$ во время предварительных вычислений варьировалось в диапазоне 500 – 5000 узлов. Так как не было выявлено существенного различия результатов, для итогового моделирования было принято $|Q| = 500$.

Таким образом, аналогом выполненных в эксперименте измерений (оценивания распределения длин путей глобальной сети по нескольким корневым узлам и нескольким выборкам конечных узлов) является один прогон модели, включающий генерацию графа G (модели сети), расчёт полного распределения длин путей $\zeta(x)$ в нём и его оценок $\zeta_i(x)$ (моделей экспериментальных оценок). Вид полного $\zeta(x)$ и его оценок $\zeta_i(x)$ для одного из прогонов модели представлен на рис. 4).

Целью вычислительного эксперимента является сопоставление полного $\zeta(x)$ и его оценок $\zeta_i(x)$, а также их характеристик, в част-

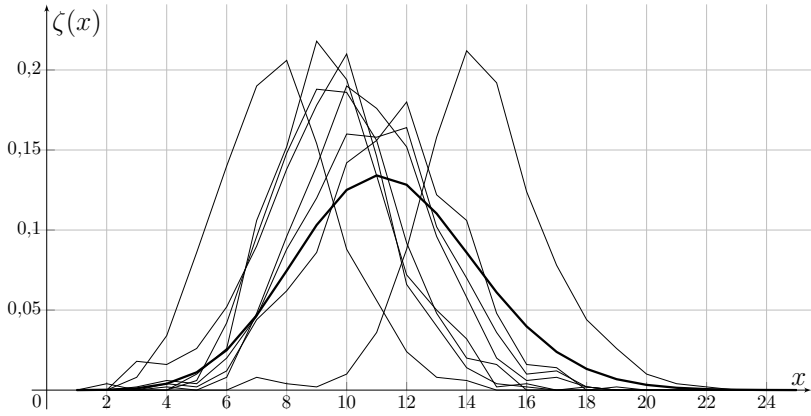


Рис. 4: Распределение длин путей между узлами (жирная линия — полное $\zeta(x)$, тонкие — его оценки $\zeta_i(x)$) в дереве Барабаши—Альберт

ности, асимметрии и эксцесса.

Так как граф G является моделью сетевого уровня сети, использующей стек протоколов TCP/IP, G должен быть связным деревом. Сложность поиска в ширину для дерева составляет $O(|G|)$, таким образом, сложность расчёта всех возможных путей составляет $O(|G|^2)$.

Были использованы несколько алгоритмов для генерации связного дерева G . Прежде всего исследована знаменитая модель Барабаши—Альберт [3] (типичные результаты прогона модели, то есть полное $\zeta(x)$ и его оценки, показаны на рис. 4).

Полное $\zeta(x)$ для деревьев Барабаши—Альберт очень похоже на дискретизацию нормального. Все они унимодальны. Большинство оценок $\zeta_i(x)$ рис. 4 выглядит более островершинными, чем полное $\zeta(x)$, но менее островершинными, чем распределения, полученные экспериментально (рис. 2, а, б). Рассмотрим характеристики полных распределений и их оценок подробнее (рис. 5).

Так как один прогон модели соответствует множеству оценок, а в процессе исследования было выполнено множество прогонов, визуализировать характеристику каждой конкретной оценки аналогично рис. 3 нерационально. Соответственно, один прогон на рис. 5 соответствует одной вертикали (по оси абсцисс откладывается но-

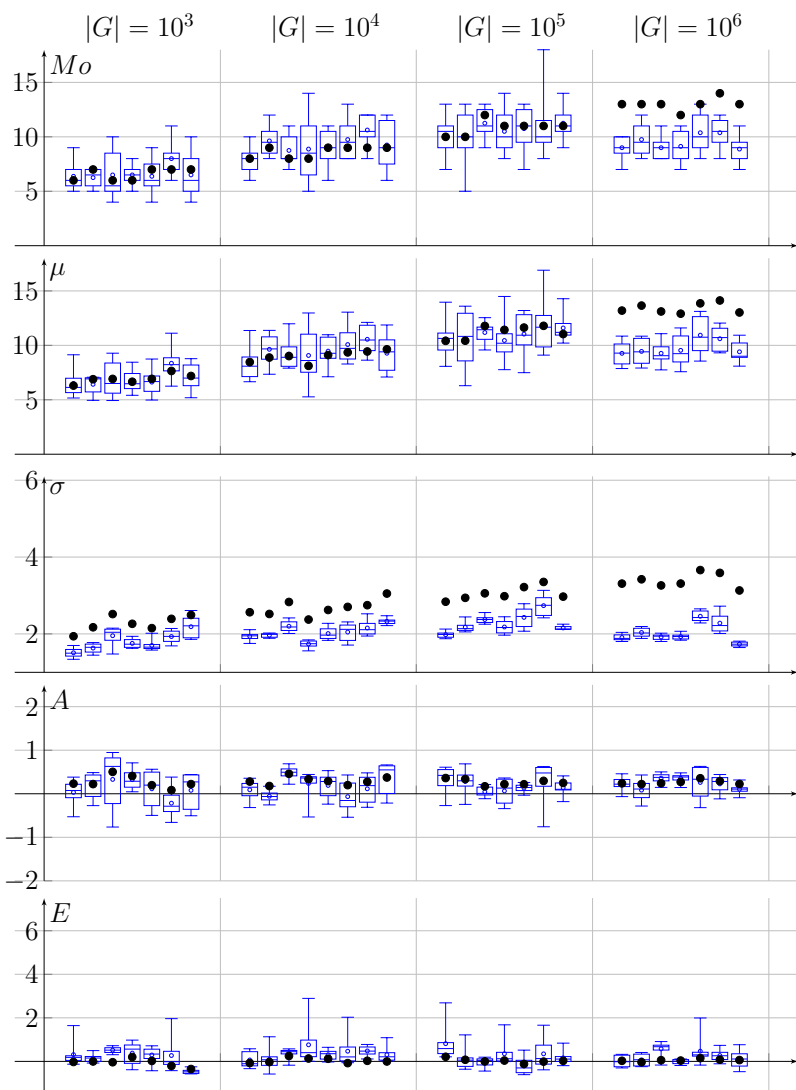


Рис. 5: Характеристики полного распределения длин путей в дереве Барабаш–Альберт (чёрные точки) и разброс характеристик его оценок (ящики с усами)

мер смоделированного графа; измерения сгруппированы по количеству $|G|$ вершин в графе). Для каждого прогона и каждой характеристики показан разброс значений характеристики $\zeta_i(x)$ в виде ящика с усами (аналогично показанному слева на рис. 3); кроме того, чёрной точкой на рис. 5 показано значение характеристики $\zeta(x)$ (для экспериментальных данных его аналог — характеристика $\zeta(x)$ глобальной сети — недоступна и поэтому отсутствует на рис. 3).

Видно, что асимметрия полных распределений $\zeta(x)$ для деревьев Барабаши—Альберт несколько больше нуля, но не достигает полученных в результате эксперимента значений 3 — 4.5. Асимметрия оценок $\zeta_i(x)$ может быть как больше, так и (что чаще) ещё меньше асимметрии $\zeta(x)$. С ростом количества $|G|$ вершин в графе асимметрия как $\zeta(x)$, так и $\zeta_i(x)$ не растёт.

Экссесс полных распределений $\zeta(x)$ для деревьев Барабаши—Альберт близок к нулю. Экссесс оценок $\zeta_i(x)$ может быть как больше (что чаще), так и меньше эксцесса $\zeta(x)$, при этом ни одна из этих величин не достигает полученных в результате эксперимента значений 20 — 30. С ростом количества $|G|$ вершин в графе эксцесс как $\zeta(x)$, так и $\zeta_i(x)$ не растёт.

Погрешность оценивания асимметрии достигает 100%, а эксцесса — 1000% (последнее обусловлено близостью к нулю эксцесса $\zeta(x)$).

Таким образом, различие между асимметрией и эксцессом экспериментально полученных для сетевого уровня глобальной сети данных и распределениями длин путей в деревьях Барабаши—Альберт обусловлено не случайными отклонениями, не методикой измерения и не отличием в размерах (асимметрия и эксцесс мало изменяются при росте $|G|$).

Соответственно, различие должно быть обусловлено моделью роста дерева. В исследовании были рассмотрены несколько моделей роста. Наиболее перспективной выглядит предфрактальная модель, описывающая рост сети как рост множества связанных друг с другом подсетей (рис. 6 и 7).

Вид полного $\zeta(x)$ и его оценок $\zeta_i(x)$ для одного из прогонов модели представлен на рис. 6. На рис. 7 этот прогон — первый в группе $|G| = 10^5$.

Так как подсети формируются при росте графа случайным образом, полные $\zeta(x)$ для предфрактальных деревьев более разнообразны, чем для деревьев Барабаши—Альберт. Некоторые из них не унимодальны. Соответственно, на рис. 7 в числе прочих ха-

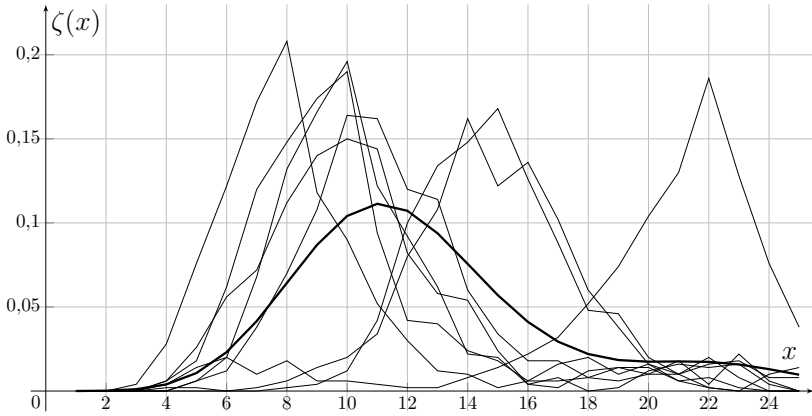


Рис. 6: Распределение длин путей между узлами (жирная линия — полное $\zeta(x)$, тонкие — его оценки $\zeta_i(x)$) в предфрактальном дереве

рактических характеристик показано и количество выраженных максимумов $\zeta(x)$. В остальном организация рис. 7 аналогична рис. 5.

Распределение $\zeta(x)$ на рис. 6 отличается от рис. 5 наличием толстого правого хвоста. В целом, в зависимости от соотношения размеров подсетей $\zeta(x)$ предфрактального дерева может иметь как более или менее толстый правый хвост, аналогичный рис. 6, так и один или несколько ярко выраженных побочных максимумов (второй прогон в группе $|G| = 10^5$), либо, наоборот, может иметь тонкий хвост, как для $\zeta(x)$ деревьев Барабаши—Альберт (первый прогон в группе $|G| = 10^6$ — в процессе роста сформировалась только одна крупная подсеть).

Соответственно, разброс как характеристик $\zeta(x)$ для разных прогонов, так и их оценок $\zeta_i(x)$ для одного прогона модели на рис. 7 больше, чем на рис. 5.

Асимметрия и эксцесс $\zeta(x)$ первого прогона в группе $|G| = 10^6$ (с единственной крупной подсетью) ожидаемо близки к асимметрии и эксцессу $\zeta(x)$ деревьев Барабаши—Альберт, а их оценки имеют столь же маленький разброс.

Асимметрия полного $\zeta(x)$ второго прогона в группе $|G| = 10^5$ (с двумя конкурирующими подсетями, то есть практически одинаковыми максимумами) положительна, при этом асимметрии её оценок могут принимать и отрицательные значения. Эксцесс полного

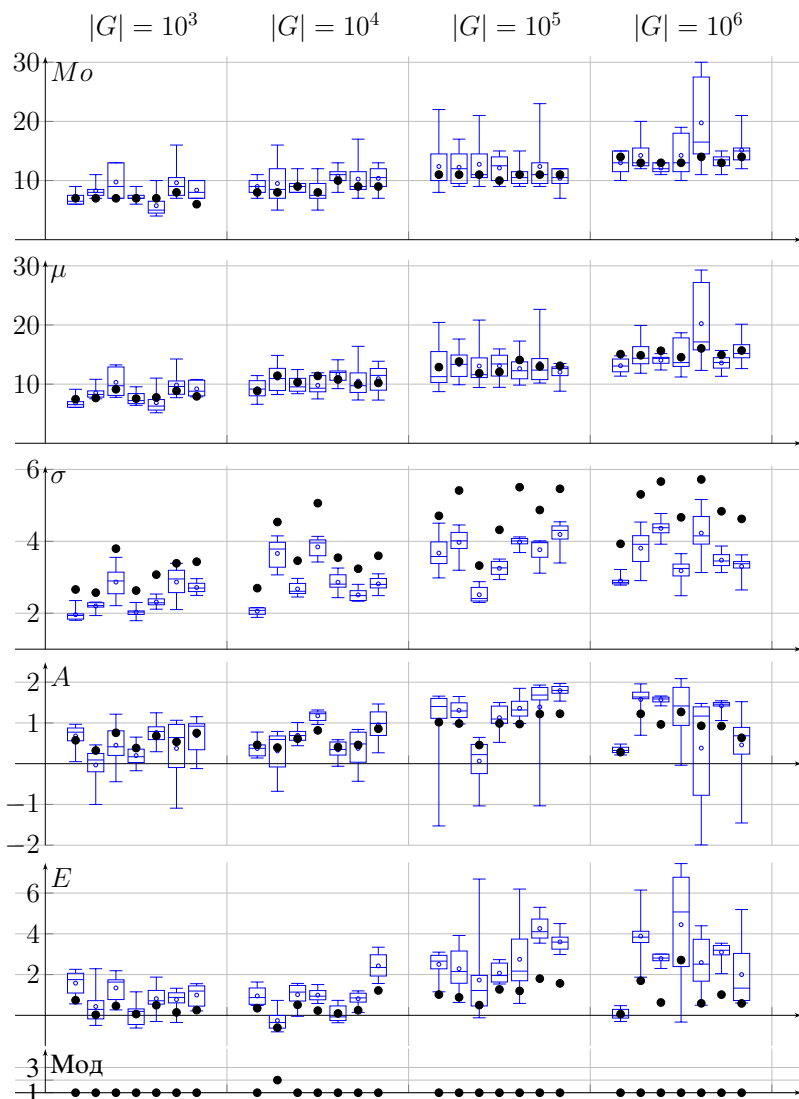


Рис. 7: Характеристики полного распределения длин путей в пред-фрактальном дереве (чёрные точки) и разброс характеристик его оценок (ящики с усами)

$\zeta(x)$ и большинство его оценок отрицательны, так как два близких максимума суммарно массивнее хвостов распределения.

Для всех прочих прогонов асимметрия и эксцесс положительны, а их оценки в подавляющем большинстве завышены. При этом с ростом $|G|$ увеличивается и разброс распределений длин путей от прогонов, и абсолютная величина асимметрии и эксцесса.

Погрешность оценивания асимметрии достигает 100%, а эксцесса — 500%.

Заключение

Инструментальные средства, стандартные для большинства дистрибутивов GNU/Linux, а также GNU Octave, позволяют выборочно оценить распределение длин путей между узлами в глобальной сети.

Имитационное моделирование позволяет утверждать, что, несмотря на гигантскую погрешность оценивания асимметрии и эксцесса описываемым методом, полученные экспериментальные данные позволяют уверенно утверждать, что сетевой уровень глобальной сети не описывается моделью Барабаши—Альберт, но может быть в некотором приближении описана разработанной предфрактальной моделью.

Список литературы

- [1] Кононова А. И., Городилов А. В., Шаньгин В. Ф. Особенности передачи данных в децентрализованных пиринговых сетях // Изв. вузов. Электроника (ВАК). 2012. № 6(98). С. 95.
- [2] Седжвик Р. Фундаментальные алгоритмы на C++: Алгоритмы на графах. СПб: Диа Софт, 2003. 1136 с.
- [3] Jeong H., Tombor B., Albert R. et al. The Large-Scale Organization of Metabolic Networks. 2000. <https://arxiv.org/pdf/cond-mat/0010278.pdf>