УДК 621.395.7

МОДЕЛИРОВАНИЕ ИНФОРМАЦИОННЫХ ПОТОКОВ ДАННЫХ В БОЛЬШИХ СЕТЯХ

Шахтурин Д. В.

Казанский государственный технический университет им. А.Н.Туполева
Тел.: (843) 238-94-16. E-mail: shakhturin@riit.kstu-kai.ru

Аннотация. Рассмотрены понятия фрактал и фрактальная размерность применительно к телекоммуникационным технологиям для моделирования топологии больших сетей. Предложен критерий оптимального согласования топологии большой сети с геометрией объекта, на котором она размещается. Рассмотрено приложение методов фрактальной геометрии для определения задержек сообщений в распределенных в топологическом смысле сетях.

Ключевые слова: фрактальная геометрия, сложная распределенная сеть, основная количественная характеристика сети, критерий согласования, моделирование, задержка сообщений.

Введение

Существующие тенденции развития современных связанных и распределенных сетей свидетельствуют о необходимости разработки эффективных методов моделирования и оптимизации сетей огромных размеров. Возможности применение к сложным сетям, которым свойственны непрерывное расширение и динамические характеристики, обычных методов моделирования существенно ограничены.

Свойства распределенных в топологическом смысле сетей (компьютерных, телекоммуникационных, сотовых станций и т.д.) существенно зависят от геометрии, способа размещения узлов, расстояния между ними и т.д. Например, с топологических позиций глобальную информационную сеть Интернет можно рассматривать как совокупность большого числа распределенных точек (узлов), взаимодействующих между собой через каналы связи [1, 2].

Информационные и коммуникационные свойства большой совокупности (ансамбля) распределенных объектов качественно и количественно отличаются от аналогичных свойств отдельного объекта [1, 2]. Например, в сетевых структурах со множеством узловых станций появляются совершенно новые свойства, такие как живучесть, надежность, множественность маршрутов доставки сообщений до пользователя, неустойчивость, конфликтность и т.д. Здесь можно привести некоторую физическую аналогию. Например, свойства газа в сосуде определяются как результат коллективного взаимодействия большого числа (ансамбля) молекул или частиц. При этом изолированная молекула или их малое количество сами по себе вовсе не обнаруживают общих макроскопических свойств, обусловленных только их большой совокупностью и геометрией сосуда (давление, температура, диффузия, объем, ламинарное или неустойчивое турбулентное течение и т.д.) [3-5].

Следовательно, можно утверждать, что свойства этого класса сетей главным образом зависят от геометрии или топологии. При этом одной из фундаментальных характеристик сетей является их топологическая размерность D_c. Поэтому множество свойств сетей зависят от размерности D_c, и эту зависимость можно выразить функциональным соотношением вида “свойство сети = $f(D_c)$”. Подсчитав размерность топологии сети D_c, можно количественно выразить, например, системные свойства сети и найти общие информационные закономерности движения потоков данных как функцию $f(D_c)$.

Размерность топологии сложных систем и сетей нельзя вычислить, используя обычную евклидову размерность, имеющую только целочисленные значения. На практике рассматриваемые большие сети, несмотря на их внешне нерегулярную структуру, характеризуются некоторым основополагающим порядком, обусловленным внешними ограничениями и моделью их роста. Данное обстоятельство позволяет использовать метод определения размерности топологии этих сетей, основанный на приложении свойства самоподобия, присущего фракталам. При этом топология глобальной информационной сети является примером случайного фрактала, поскольку ее малая часть подобна целой. Так, топология сети сотовой связи в масштабе отдельного городского района подобна топологии сети городского масштаба, а топология городской сети – топологии сети регионального масштаба (и т.д., по восходящей иерархии масштабов) [1]. Следовательно, узлы информационной сети можно рассматривать как множество точек, вложенных в пространство. Размерность этой совокупности точек имеет дробную размерность, или фрактальную размерность.

1. Фрактальный характер топологии большой сети

Рассмотрим вопрос о фрактальности геометрии городских сетей коммуникаций, имеющих структуру графов. Антропогенная деятельность человека, как показывают проведенные исследования, может иметь фрактальный характер. Это, в частности, относится к транспортной сети метро в Париже [6] и развитию некоторых мегаполисов [7].

Сложность описания больших сетей заключается в том, что интенсивность информационного потока данных изменяется в различных областях зоны обслуживания. Интенсивность информационного потока в отдельных областях может быть оценена при помощи географических и демографических характеристик зоны обслуживания. При этом распределение плотности населения в зоне обслуживания однозначно определяет интенсивность информационного потока. Следовательно, узлы сети сосредоточены в областях с высокой интенсивностью информационного потока (высокой плотностью пользователей) и редки в областях с низкой интенсивностью (низкой плотностью пользователей). Например, для достижения эффективной конфигурации сети мобильной связи основные объекты сети (базовых станций и центров коммуникации) должны быть расположены близко к предполагаемым источникам трафика.

Таким образом, сети коммуникаций (транспортные сети, сети сотовых станций и других телекоммуникаций, сети магазинов, больниц и систем обслуживания населения и т.д.) обуславливают развитие и рост городов, причем с сильной обратной связью. Следовательно, возможно представление коммуникационных сетей города имеющих тесную корреляционную связь с географической геометрией улиц и районов. Геометрия городских застроек непременно обуславливает топологию сетей коммуникаций.

В качестве примера проведено исследование фрактальности транспортной сети города Казани. Использовалась карта города масштабом 1:100000 издания 2006 г. Подсчитывалось число $N(R)$ узлов сети, расположенных внутри окружностей радиусом R с центром в Казанском Кремле, являющимся географическим центром города (рис. 1). При равномерном распределении вдоль прямых линий значение $N(R)$ пропорционально расстоянию R.

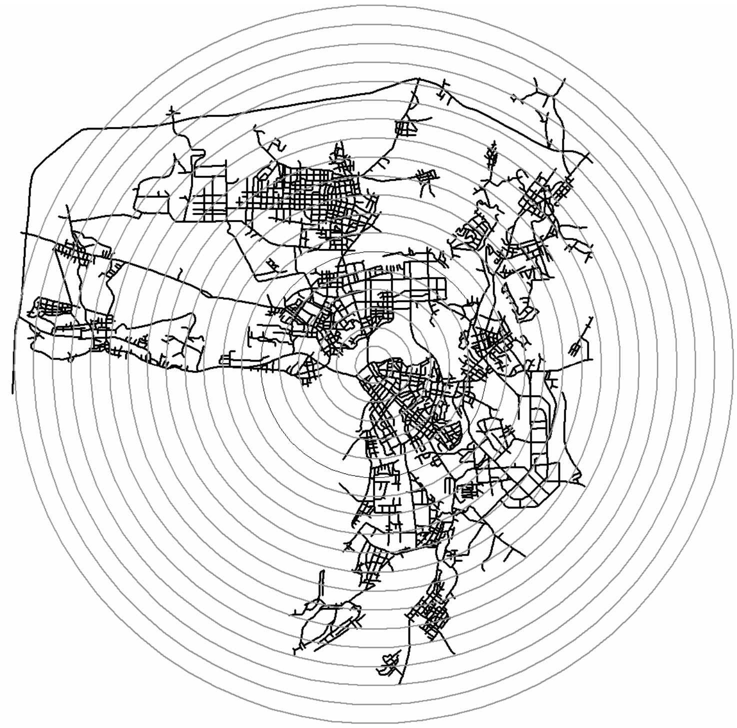


Рис. 1. Топология транспортной сети г. Казани

Если пространственное распределение узлов сети компактно, то есть если их плотность (число узлов на единицу площади) постоянна, значение $N(R)$ пропорционально R^2. При фрактальном распределении значение $N(R)$ пропорционально R^{D_c}, где D_c – фрактальная размерность коммуникационной сети. Это означает, что плотность числа узлов уменьшается с возрастанием R.

Таким образом, для окружности радиуса R и площади, пропорциональной R^2, плотность $\rho(R)$ есть

$$\mathrm{c}(R) \sim N(R)/R^2 \sim R^{D_c-2}. \tag{1}$$

Из (1) следует, что плотность узлов сети при увеличении радиального расстояния убывает по степенному закону. При этом если фрактальная размерность сети меньше евклидовой размерности пространства, вмещающего сеть ($D_0 = 2 < D_c$), то плотность узлов сети асимптотически стремится к нулю с ростом R. Однако плотность узлов сети будет в среднем постоянна, если фрактальная размерность D_c приближается к евклидовой размерности пространства.

Результаты топологических расчетов для г. Казани показаны в полулогарифмическом масштабе на рис. 2 ($R_{\max} = 21$ *км*). Для сравнения на этом же рисунке приведены аналогичные данные для линий железнодорожной транспортной системы Парижа [5].

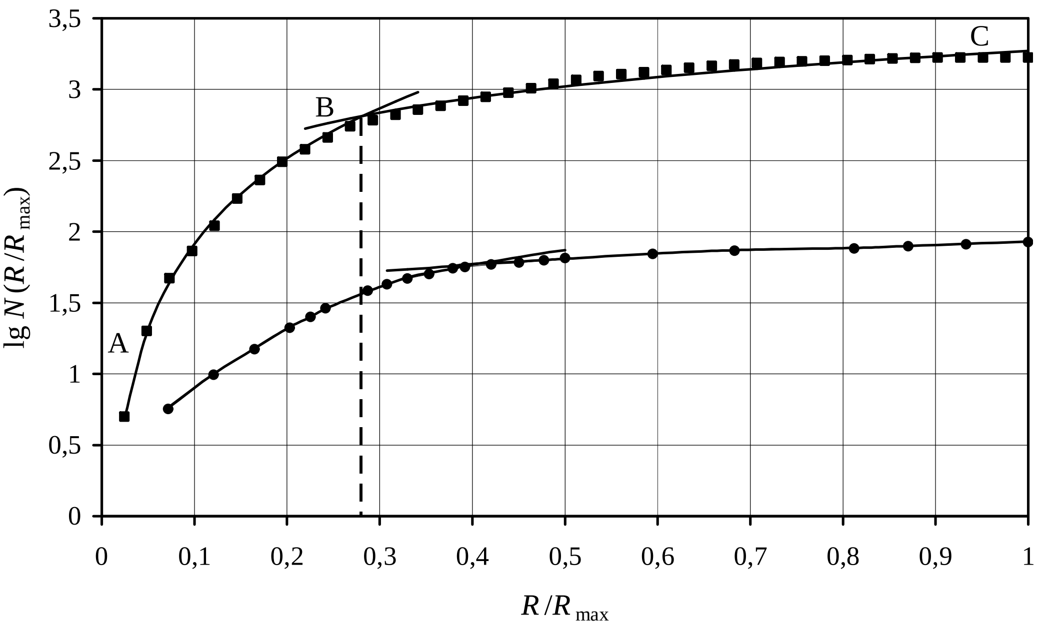


Рис. 2. Графики зависимости числа узлов $N(R/R_{\max})$ транспортной сети г. Казани (-■-■-■-■-), числа станций железнодорожной сети г. Парижа (-●-●-●-●-) от относительного радиального расстояния $R/R_{\max}$ в полулогарифмическом масштабе

Вблизи центра ($R \le 6$ *км*) плотность узлов сети пропорциональна значению R^2 (рис. 2, участок АВ). При удалении большем, чем $R \approx 6$ *км*, наблюдается резкий переход к закономерности вида $N(R) \sim R^{0,8}$ (рис. 2, участок ВС).

Значение $R = 6$ *км* (рис. 2, точка В) примерно соответствует радиусу исторического центра города (рис. 1). Следовательно, при $R \le 6$ *км* число узлов городской транспортной сети изменяется как квадрат расстояния от центра. Для $R > 6$ *км* сеть образует фрактальную размерность около 0,8, плотность узлов при этом определяется уравнением

$$\mathrm{c}(R) \sim R^{-1,2}. \tag{2}$$

Сравнение графиков транспортных сетей г. Казани и г. Парижа (рис. 2) свидетельствует о существовании общих глобальных закономерностей развития топологии сетей коммуникаций крупных мегаполисов. При этом одной из фундаментальных характеристик подобных сетей является их топологическая размерность D_c.

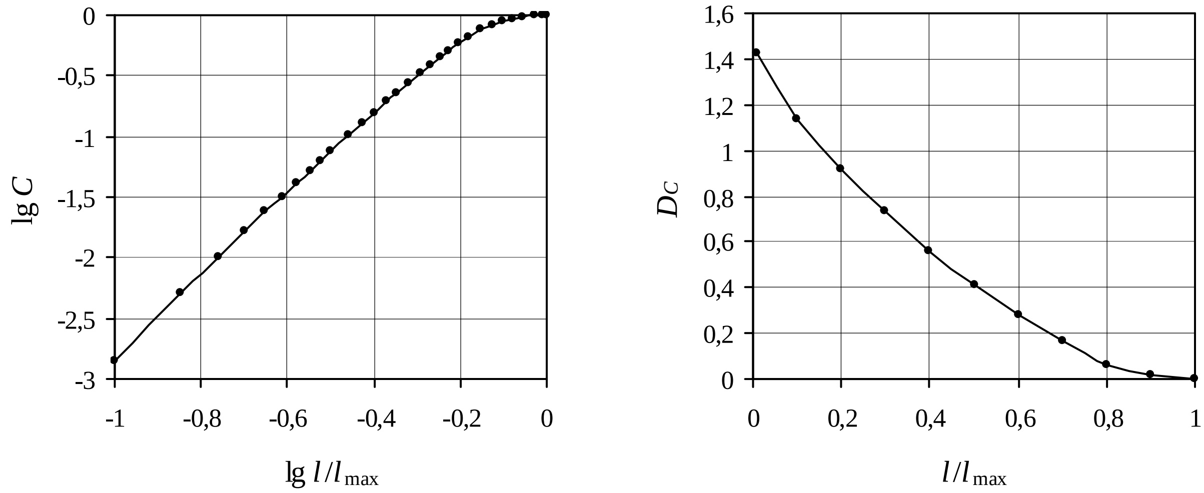
Применение функции плотности узлов дает нужную методику моделирования и синтеза больших сетей. Дополнительное преимущество этого подхода состоит в том, что число узлов $N(R)$ монотонно спадает в радиальном направлении от некоторой центральной точки. Кроме того, хорошо согласуется с эмпирическим законом Зипфа о распределении населения вокруг городских центров [8]. Можно предположить, что сети, узлы которых распределены случайным образом, согласно соответствующей функции плотности с одинаковой фрактальной размерностью, хотя и различны, но качественно подобны и, следовательно, обладают некоторыми общими глобальными свойствами.

Таким образом, проведенное исследование позволило выявить несколько важных характеристик фрактального моделирования больших сетей узлы, которых распределены случайным образом в соответствии функции плотности. Во-первых, по мере удаления от начальной точки плотность узлов монотонно убывает, поскольку топологическая размерность D_E больше фрактальной размерности D_c. Во-вторых, по мере приближения фрактальной размерности к топологической ($D_c \to D_E$) узлы сети будут равномерно покрывать круг соответствующим образом заданного радиуса, тогда как за пределами этого радиуса не окажется ни одного узла.

В работе определена фрактальная размерность топологии узлов транспортной сети г. Казани. На рис. 3, *б* приведена зависимость корреляционной размерности $D_C(l/l_{max})$ от относительного расстояния l/l_{max}. Корреляционная размерность определялась следующим образом [9]:

$$D_C = \lim_{e \to 0} \frac{\ln C(e)}{\ln e}, \tag{3}$$

где $C(\varepsilon)$ – корреляционная функция, рассчитываемая как отношение числа точек n, попарные расстояния между которыми меньше ε, к квадрату общего числа точек N. Как видно из графика, с увеличением расстояния значение D_C убывает.



а) *б*)

Рис. 3. Распределение корреляционная функция $C(l/l_{\max})$ (*а*) и корреляционной размерности $D_C(l/l_{\max})$ (*б*) в зависимости от относительного расстояния $l/l_{\max}$

На рис. 3, *а* в двойном логарифмическом масштабе показана зависимость корреляционной функции $C(l/l_{max})$ от относительного расстояния l/l_{max}. Средний наклон кривой оказался равным 1,32, что и дает оценку фрактальной размерности $D_C = 1{,}32$ для системы узлов транспортной сети г. Казани.

Однако транспортная сеть г. Казани представляет собой неоднородный фрактальный объект – мультифрактал, для полного описания которого, в отличие от регулярного фрактала, недостаточно введения всего лишь одной величины, его фрактальной размерности. Причина этого заключается в том, что наряду с чисто геометрическими характеристиками, определяемыми величиной D, такие фракталы обладают и некоторыми статистическими свойства.

Под словом “неоднородный” здесь понимается географически неравномерное распределение узлов сети. Причина неоднородности в данном случае разные вероятности заполнения геометрически одинаковых элементов фрактала, или в общем случае несоответствие вероятностей заполнения геометрическим размерам соответствующих областей.

Поэтому для полной характеристики транспортной сети г. Казани были определены обобщенные фрактальные размерности D_q

$$D_q = \lim_{e \to 0} \frac{1}{q-1} \cdot \frac{\ln \sum_{k}^{N} p_k^q}{\ln e}, \quad -\infty \le q \le \infty, \tag{4}$$

и мультифрактальный спектр $f(a)$

$$f(a(q)) = \tau(q) + q \cdot a(q), \tag{5}$$

где $\tau(q) = -\lim\limits_{\varepsilon\to 0} \dfrac{\ln \sum\limits_{k}^{N} p_k^q}{\ln \varepsilon}$ – последовательность показателей массы;

$a(q) = -\dfrac{d}{dq}\tau(q) = \lim\limits_{\varepsilon\to 0} \dfrac{\sum\limits_{k}^{N} p_k^q \ln p_k}{\ln \varepsilon \sum\limits_{k}^{N} p_k^q}$ – показатель массы Липшица-Гельдера.

На рис. 4 приведены обобщенные размерности и мультифрактальный спектр транспортной сети г. Казани.

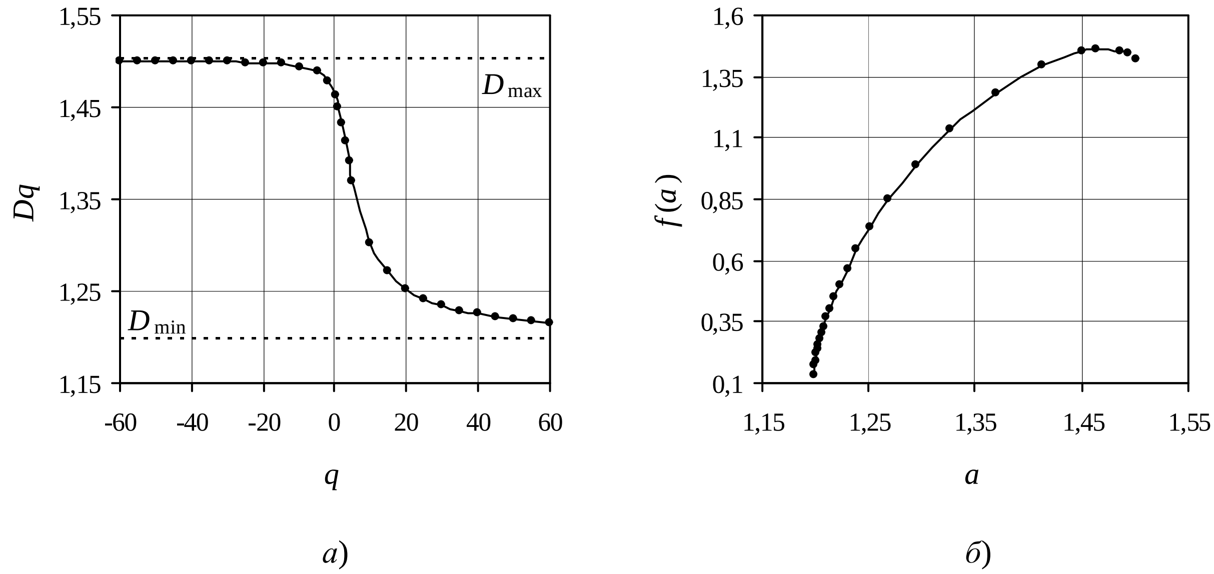


Рис. 4. Обобщенные размерности (*а*) и мультифрактальный спектр (*б*) транспортной сети г. Казани

Следует отметить, что мультифрактальный спектр транспортной сети г. Казани коррелирует со спектром модели диффузионно-ограниченной агрегации, описывающей образование кластеров из коллоидов [9, 10]. В этой модели одномерные цепочки молекул (мономеры) поступают из удаленного источника и диффундируют, совершая случайные блуждания. Достигнув кластера, блуждающие мономеры прилипают к нему. Процесс агрегации такого типа порождает характерные кластеры с фрактальной размерностью D = 1,71 в случае диффузии на плоскости.

2. Критерий согласования топологии большой сети с геометрией объекта, на котором она размещается

Введенное на основе концепции фракталов понятие размерности топологии сети D_c является мощным математическим инструментом для количественного сравнения, анализа и синтеза различных топологий сетей [1, 2]. Например, задача покрытия заданной территории сетью узлов сотовых станций, обеспечивающих устойчивую связь, может быть решена как задача синтеза соответствующей топологии сети нужной фрактальной размерности.

При этом, если фрактальная размерность геометрии объекта (район города, город, регион и т.д.) равна D_0, тогда справедливо следующее утверждение [2]: для оптимального покрытия объекта сложной коммуникационной сетью ее топология должна быть согласована с топологией объекта, т.е. должно выполняться соотношение

$$D_c \geq D_0 . \tag{6}$$

Для доказательства воспользуемся оценкой фрактальной размерности через размерность Реньи (информационную размерность) D_I, связанную с энтропией $H(\varepsilon)$ исследуемого объекта (количеством информации, необходимым для описания состояния объекта с погрешностью ε) выражением [9]:

$$D_I = \lim_{e \to 0} \frac{H(e)}{\ln(1/e)}. \tag{7}$$

где $H(e) = -\sum_{i=1}^{N} p_i \ln p_i$ – информационная энтропия по Шеннону; p_i – вероятность попадания в i-ю сферу диаметром ε из $N(\varepsilon)$ сфер покрывающих объект. П. Грасбергер и И. Прокачча показали, что информационная размерность D_I является нижней оценкой для D_F [11]:

$$D_I \le D_F. \tag{8}$$

Действительно, если топология объекта характеризуется информационной энтропией H_0 – потенциальной способностью описываться количеством информации $I_{0max} = H_0$, то в результате неполного описания топологии энтропия уменьшается на величину H_z и характеризующая объект информация равна

$$I = H_0 - H_z = \text{Д}H. \tag{9}$$

Количество информации о геометрии объекта I без учета ее потерь должно соответствовать информации, необходимой для покрытия топологии объекта узлами коммуникационной сети:

$$I_c \ge I. \tag{10}$$

Пусть межузловое расстояние равно ε. Характеризующая топологию объекта максимальная информация при этом будет зависеть от ε:

$$I_{0\max}(e) = \text{Д}H_{\max}(e) = H_0(e). \tag{11}$$

Для согласованного покрытия геометрии сложного объекта сетью узлов, при заданном межузловом расстоянии ε, согласно (10), должно удовлетворяться условие

$$\text{Д}I(e) = I_c(e) - I_{0\max}(e) \ge 0. \tag{12}$$

Разделим (12) на ln (1/ε) и если при ε → 0 существует предел выражения

$$\lim_{e \to 0} \frac{\text{Д}I(e)}{\ln(1/e)} = \lim_{e \to 0} \frac{H_c(e)}{\ln(1/e)} - \lim_{e \to 0} \frac{H_0(e)}{\ln(1/e)}, \tag{13}$$

то, исходя из определения информационной размерности (7) и соотношения (8), получим подтверждение соотношения (6)

$$\text{Д}D = D_c - D_0 \ge 0 \text{ или } D_c \ge D_0. \tag{14}$$

Таким образом, предложенный фрактальный критерий покрытия объекта большой сетью позволяет утверждать, что большое или избыточное количество узлов сети не гарантирует оптимальную распределенность узлов на топологии сложного объекта без учета фрактальности его геометрии.

3. Моделирование задержек сообщений во фрактальных структурах

Время задержки сообщений является одним из основных факторов, определяющих производительность сети. Здесь под задержкой сообщений понимается время, за которое сообщение проходит путь от его источника через сеть до места назначения, т.е. время, проводимое сообщением в сети. Представляет теоретический и практический интерес зависимость задержки сообщений от топологии сети.

В работе было проведено имитационное моделирование задержки сообщений на участке топологической сети г. Казани (рис. 5). При этом была построена идеализированная модель сети и введен ряд допущений:

1. отсутствие помех в каналах и отсутствие ненадежных узлов и каналов;

2. отсутствие задержки передачи внутри узлов;
3. единственность адресата каждого сообщения и отсутствие потерь;
4. бесконечная емкость хранения на каждом узле;
5. необходимость полного приема сообщения для ретрансляции;
6. пуассоновский входящий поток сообщений.

Длины сообщений ν_n и промежутки между их поступлением a_n распределены экспоненциально со значениями λ и 1/? соответственно:

$$p(a_n)=\text{л}\cdot e^{-\text{л}\,a_n}\,, \tag{15}$$

$$p(\text{н}_n)=\text{м}\cdot e^{-\text{м}\,\text{н}_n}\,. \tag{16}$$

Применена случайная путевая процедура, согласно которой выбор следующего узла, который будет посещен, производится с равной вероятностью по множеству свободных каналов, выходящих из данного узла.

Для выбранного участка сети был выполнен имитационный цикл, состоящий из 450 серий. В каждой серии осуществлялось последовательное во времени формирование n = 50 сообщений при фиксированном значении скорости сообщений λ = const. Значения λ были равны $\overline{1{,}10}$ *сообщ./с*.

Рис. 5. Топологическая сеть г. Казани

Методика измерения величины задержки сообщений в сети T состояла в фиксации момента $t_н$ поступления сообщения в сеть и момента $t_к$ получения сообщения адресатом:

$$T=t_к-t_н\,. \tag{17}$$

Было выдвинуто предположение о том, что время, проводимое сообщением в сети, зависит от фрактальной размерности сети D_c с точностью до числового множителя α следующим образом

$$T_{ср}=\text{б}\,N^{D_c}\,, \tag{18}$$

где α – числовой множитель, имеющий размерность времени; N – количество узлов сети.

По множеству полученных значений T были вычислены оценки плотности вероятности $p(T/\sigma)$ (где σ – среднеквадратическое отклонение), графики которых для различных значений λ приведены на рис. 6.

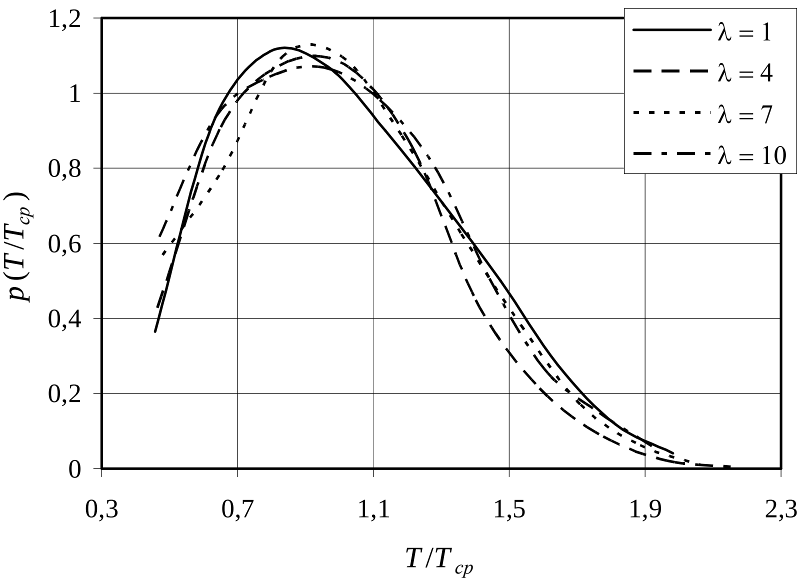


Рис. 6. Графики плотности вероятности $p(T/\sigma)$ полученные при различных значении λ

Используя полученные экспериментальные результаты также было определено среднее время задержки сообщений $T_{cp} = \frac{1}{N}\sum_{i=1}^{N} T_i$ (где T_i – время задержки сообщений i-й серии измерений) для различных значений интенсивности входного потока λ. На рис. 7 представлены графики зависимости средней задержки сообщений T_{cp} как функция интенсивности поступления потока сообщений λ.

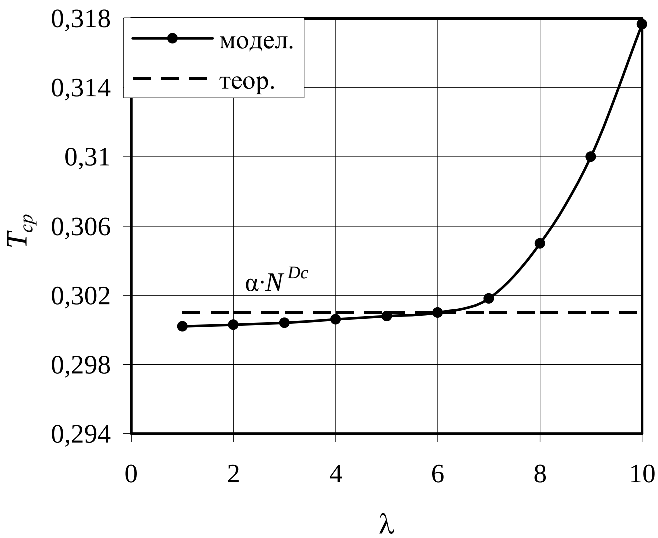


Рис. 7. Зависимости между средней задержкой сообщений T_{cp} и интенсивностью потока λ

Исследуемый участок имеет фрактальную размерность $D_c = 1{,}08$ сети и содержит $N = 116$ узлов. При этом оценка среднего времени задержки сообщений в данном участке сети с точностью до числового множителя α = 1,8 *мс* в соответствии с (18) $T_{cp} = 0{,}301$ *с*.

По графику (рис. 7) можно отметить сходство между средней задержкой сообщений, полученных в результате моделирования и теоретических расчетов, для выбранного участка сети (рис. 5) при малых значения интенсивности входного потока. При малых значениях λ, когда интенсивность информационного потока в сети и занятость узлов незначительны, время доставки сообщений в основном определяется топологией сети. С увеличением интенсивности входного потока, несмотря на множественность маршрутов доставки сообщений до получателя, нагрузка сети возрастает и на первое место выходят такие сетевые параметры как время поступления в сеть и длина сообщений, пути, проходимые этими сообщениями по сети (учитывая очереди, образуемые ими в соответствии с принятой дисциплиной очереди, когда это необходимо) и т.д.

Заключение

Системные свойства (живучесть, надежность, время доставки) больших сетей существенным образом зависят от топологии. При этом основной характеристикой сетей является их фрактальная размерность, удобная для моделирования и описания больших распределенных в топологическом смысле сетей. Эти положения показаны на примере транспортной сети г. Казани, имеющей структуру фрактальных графов.

Используя, предложенный фрактальный критерий покрытия объекта сложной сетью определена характерная особенность такого вида сетей: большое или избыточное количество узлов сети не гарантирует оптимальную распределенность узлов на топологии сложного объекта без учета фрактальности его геометрии. Последнее позволило выявить основные свойства фрактального описания топологии сложных сетей.

Полученные результаты позволяют использовать фрактальную размерность как числовую характеристику для анализа информационных свойств сетей. Применение методов фрактальной геометрии для моделирования задержек в больших сетях дает возможность изучать закономерности движения информационных потоков данных с целью предсказания развития и повышения эффективности использования этих сетей.

По результатам проведенных исследований можно сделать вывод о том, что предложенный подход, основанный на методах фрактальной геометрии, позволяет осуществлять количественное сравнение, анализ и синтез сложных сетей.

Литература

1. Евдокимов Ю.К., Шахтурин Д.В. Фрактальный характер топологии сложных сетей // Материалы IV Междунар. конф. “Методы и средства управления технологическими процессами”. – Саранск: Изд-во Мордов. ун-та, 2007, с. 244−251.
2. Евдокимов Ю.К., Шахтурин Д.В. Фрактальное моделирование топологии сложных сетей // Труды Казанского научного семинара “Методы моделирования”. – Казань: Изд-во КГТУ, 2007, вып. 3, с. 218−233.
3. Евдокимов Ю.К., Базлов Е.Ф. Количественная мера топологии множества датчиков методами фрактальной геометрии // Материалы Всероссийской науч.-технич. Конф. “Датчики и преобразователи информации систем измерения, контроля и управления (Датчик – 95)”. – Москва−Гурзуф, 1995, т. 3, с. 538-539.
4. Евдокимов Ю.К. Распределенные измерительные среды. Принципы, топология, моделирование // Вестник КГТУ им. А.Н. Туполева, 2004, № 1, с. 165−191.
5. Евдокимов Ю.К. Распределенные измерительные среды и континуум-измерения: принципы, топология, алгоритмы // Нелинейный мир, 2007, № 10−11, с. 639−656.
6. Benguigui L., Daoud M. Is the Suburban Railway System a Fractal? // Geographical Analysis. 1991, v. 23, № 4, pp. 362−368.
7. Batty M. Cities as Fractals: Simulating Growth and Form // Fractal and Chaos / Ed. By A.Grilly, R Earnashaw, H. Jones. – N. Y.: Springer, 1991, pp. 43−69.
8. Zipf G.K. Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology. Reading, MA: Addison-Wesley Press, 1949, 250 p.
9. Федер Е. Фракталы: Пер. с англ. – М.: Мир, 1991. – 254 с.
10. Смирнов Б.М. Физика фрактальных кластеров. – М.: Наука. Гл. ред. физ.-мат. лит., 1991, 136 с.

11. Grassberger P., Procaccia I. Measuring the strangeness of strange attractors// Physica D, 1983, v. 9. p. 188−208.