

Модель сложных сетей с придирчивостью

Zarvanskiy Igor, Snarskii

NTUU KPI and

IPRI

(Dated: May 01, 2014)

Бла-бла-бла

I. Введение

Значительное количество реальных сложных сетей являются безмасштабными сетями, такими, степени которых распределяются по степенному закону. К таким сетям относятся WWW-сети, сети метаболизма, сети питания (food webs), социальные сети и многие другие [1].

В настоящее время свойства таких безмасштабных сетей достаточно подробно изучено, установлены их сетевые характеристики (средняя степень узла, минимальный средний путь, коэффициент кластеризации и т.д.) [?]. Необходимо заметить, что сложные сети построенные согласно предположенным сценариям [2] являются идеализацией реальных сетей, характеристики которых могут иногда значительно отличаться от идеальных [?]. Тем не менее степенная зависимость степени узлов реальных сложных сетей встречается достаточно часто и особенно для тех сетей, которые образованы (возможно само организованными) развивающимся по времени процессом [].

Одним из таких процессов, который начал изучаться задолго до появления понятия сложная сеть, был процесс распределения между людьми «богатства» (под которым можно понимать деньги, вложения, недвижимость...). В [3] Парето был установлен т. н. Закон Парето — степенное распределение богатства — когда, число людей ν , владеющих μ - долей богатства является степенной функцией $\nu \sim \mu^{-\gamma}$, при $\gamma = \dots$ получается так, что 20% людей владеют 80% богатства, что часто называется законом 80/20.

В работе [4] был найден сценарий образования сложной сети, т. н. сценарий Барабаши-Альберт, со степенным законом распределения степеней узлов, основанном на двух принципиально важных положениях:

1. Сеть является растущей, начиная с некоторого затравочного числа узлов m_0 , на каждом временном шаге появляется некоторое число новых узлов с n связями.
2. Вероятность подсоединения связей от нового узла к уже существующим прямо пропорциональна по степени.

Коротко говоря модель Барабаши-Альберт — это растущая сеть с предпочтительным подсоединением.

В дальнейшем появилось много модификаций алгоритма Барабаши-Альберт, в [2] их перечислено около 20-ти. Все они приводят к безмасштабным сетям с различным значением показателя степени распределения узлов по их степеням. На первый взгляд представляется, что растущая сеть с различным типом предпочтительного соединения обязательно вырастет в безмасштабную сеть.

В настоящей работе показано, что возможна такая, незначительная на первый взгляд, модификация закона предпочтительного соединения, при которой степенное распределение модели Барабаши-Альберт принципиально нарушается. В функции распределения при этом появляется провал, означающий отсутствие узлов сети для некоторого диапазона значений степени. Как показали подробные исследования, введенный параметр r , определяющий модификацию закона предпочтительного подсоединения, имеет пороговое значение r_c , так что при $r < r_c$ сеть остается безмасштабной сетью, а при $r \geq r_c$ появляется провал. Величина провала степенным образом зависит от близости параметра r к своему пороговому значению, что позволяет говорить об аналогии с фазовым переходом второго рода.

Модификация закона подсоединения также была нами опробована на детерминированных иерархичных безмасштабных сетях, т. н. детерминированных (u,v)-flowers [5]. При этом также наблюдалось нарушение степенной функции распределения аналогично фазовым законам второго рода.

Статья построена следующим образом. Вначале рассматривается модель подсоединения с т. н. придиричностью, связь и отличие от модели Барабаши-Альберт. Далее рассматривается модель (u,v)-flowers, приводятся результаты численного моделирования и исследования различных характеристик растущих (u,v)-flowers при подсоединении с придиричностью. Аналогичные исследования проведены для случайно растущих сетей, где к алгоритму Барабаши-Альберт добавили параметр, связанный с «придиричностью». И в этом случае численное моделирование приводит к явлениям аналогичным фазовым переходам второго рода.

II. Модель Барабаши-Альберта с придиричностью

A. Алгоритм Барабаши-Альберта

Рассмотрим растущую сеть. В стандартном варианте модели Барабаши-Альберт [4] на первом шаге по времени существует m_0 узлов связанных между собой. На каждом следующем шаге возникает m новых узлов с q связями. p_i - вероятность подсоединения(создание связи между узлами) нового узла к уже существующему узлу i пропорциональна по степени (числу связей узла i) - k_i :

$$p_i = \frac{k_i}{\sum_j k_j}, \quad (1)$$

где суммирование происходит по всем “старым” узлам.

Такой алгоритм, при большом числе шагов по времени приводит к степенной функции распределения степеней $P(k)$:

$$P(k) \sim k^{-\gamma}, \quad (2)$$

с показателем $\gamma = 3$ [4].

В [2] приведено много модификаций правила предпочтительного подсоединения, которые приводят к различным значениям показателя γ . Однако сама степенная зависимость (2) остается.

В. Модификация алгоритма Барабаши-Альберта

Здесь мы предлагаем обобщение модели, основанной на правиле предпочтительного соединения, введенного Барабаши-Альберт. Новое правило предпочтительности будем для краткости называть подсоединением с “придирчивостью”(exceptive). Согласно этой модели вводится новый параметр exceptive - r , принимающей значения в диапазоне $(0, 1)$. В том случае, когда выбор подсоединения новой связи выпал на узел i со степенью k_i , подсоединение происходит с вероятностью p_i , но только в том случае, когда выполняется условие:

$$p_i = \frac{k_i}{\sum_j k_j}, \quad k_i \geq r \langle k \rangle, \quad (3)$$

где $\langle k \rangle$ – среднее значение степени узлов в сети на момент присоединения, $\langle k \rangle = \sum_j k_j / N$.

Введения дополнительного условия (3) в процессе роста сети отсекает часть узлов (делает их невалидными), то есть к ним в данный момент не может присоединиться новая связь. Необходимо заметить, что если в данный момент времени некий узел не удовлетворяет условию (3), это еще не значит, что в следующие моменты времени к нему не смогут присоединиться новые узлы. Валидность или невалидность узла меняется со временем, так как с течением времени изменяется значение $\langle k \rangle$.

С. Функция распределения степеней узлов

При значении параметра придирчивости $r = 0$ предлагаемая модель переходит в стандартную модель Барабаши-Альберт. Удивительным является наличие порогового значения параметра придирчивости r_c . При значении параметра придирчивости меньше некоторого порогового r_c , то есть при $r < r_c$ функция распределения степеней узлов $P(k)$ остается степенной, а сама сеть, тем самым, безмасштабной сетью. При значениях параметра придирчивости больше порогового значения $r \geq r_c$ сеть меняет свою структуру, а именно в сети исчезают узлы со «средним» количеством связей, что мы можем увидеть на рис. ПС.

Определим пороговое значение параметра придирчивости. Для этого рассчитаем пороговое значение для сети в 100 узлов, а дальше с шагом в 100 узлов будем увеличивать размер сети до 2000 узлов, при этом будем усреднять значения по 10 экспериментам. Для сети в 100 узлов $r_c = 0.62$, далее происходит насыщение r_c . Для сети в 1100 узлов r_c полностью насыщается и мы получаем $r_c = 0.51$. В последующих расчетах мы будем использовать $r_c = 0.51$.

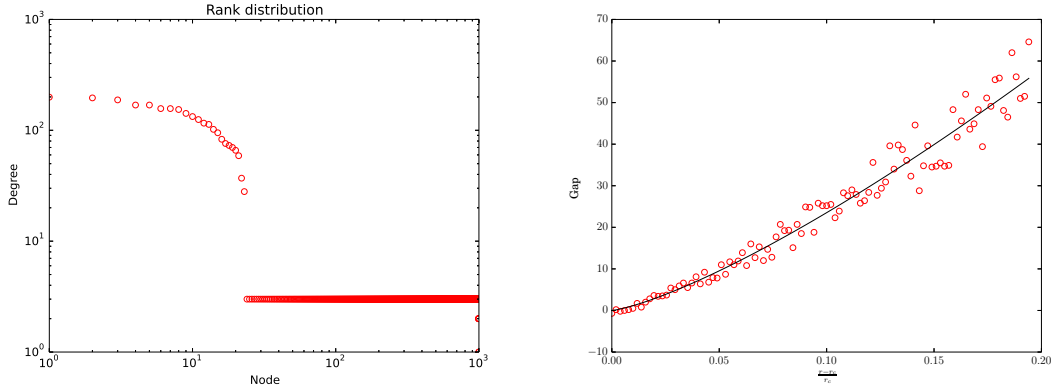


Рис. 1. Ранжированное распределение при $r = 0.6$. По горизонтальной оси отложено порядковый номер узла, по вертикальной оси отложена степень узла. Изменение величины разрыва

Введем новую характеристику сети - величину разрыва η рис. II C, расстояние между узлами, ближайшими к разрыву (разница значений степени узла до разрыва и после разрыва). Величина разрыва, указывает по вертикальной оси те значения степеней узлов, которые отсутствуют.

Как следует из численного моделирования, поведение параметра η аналогично поведению параметра порядка в теории фазовых переходов второго рода [6]. Как известно, параметр порядка η , например намагниченность, при приближении температуры к критическому значению T_c уменьшается степенным образом $\eta \sim (r - r_c)^\beta$, где β - критический индекс. При проведении численного эксперимента были выбраны следующие начальные параметры: количество узлов $N = 5000$, начальное количество узлов $m_0 = 20$, количество связей у каждого нового узла $m = 3$.

На рис. II C показана полученная зависимость $\eta = A \cdot (r - r_c)^\beta$, где $\beta \sim 1.3$

D. Коэффициент кластеризации, ассортативность, минимальное среднее расстояние, величина разрыва

Появление разрыва η в распределении степеней узлов $P(k)$ свидетельствует о значительном изменении структуры сети, что не может не сказаться на её характеристиках. Ниже рассмотрено поведение C - коэффициента кластеризации, A - ассортативности и l - минимального среднего расстояния, как функции коэффициента придиричивости r , при $r \geq r_c$. Как показал численный эксперимент для сети с $N = 5000$ узлов, коэффициент кластеризации C , ассортативность A , минимальное среднее расстояние l при $r < r_c$ от r не зависит и равна $C_0 \approx 0.01$, $l_0 \approx 3.98$, $A_0 \approx -0.096$, что, как и должно быть, совпадает с расчетами приведенными в [2, 7].

При увеличении r от r_c до $r_c + 0.01$ с шагом 0.001 коэффициент кластеризации увеличивается от 0.04 до 0.14, ассортативность уменьшается от -0.3 до -0.6 , среднее минимальное расстояние уменьшается от 3.5 до 2.9 (согласно рис. II D). Для нормализации зависимости возьмем отношение параметра к его значению при $r < r_c$, а также приведем зависимости к возрастающим функциям: $A = \frac{A}{A_0}$, $C = \frac{C}{C_0}$, $l = -\frac{l}{l_0}$.

При $r \geq r_c$ такая зависимость появляется, и она оказывается степенной, а именно $C \sim (r - r_c)^\alpha$, $A \sim (r - r_c)^\gamma$, $l \sim (r - r_c)^\sigma$, где $\alpha \approx 0.46155027$, $\gamma \approx 0.26025569$, $\sigma \approx 0.11761072$

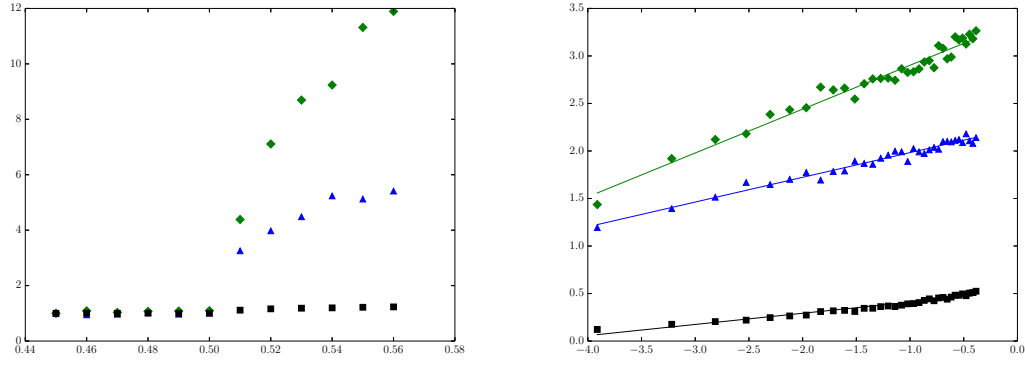


Рис. 2. Изменение кластеризации,ассортативности, минимального среднего пути при $r = [0, 51; 0, 52]$ с шагом 0.001

Е. Матрица смежности для сети с придиричивостью

Рассмотрим матрицу смежности A_{ij} для сети с придиричивостью. Для удобства нумерации узлов в матрице смежности будем вести в порядке спадания количества связей. То есть $k_i = \sum_j A_{ij}$ не возрастает с увеличением i .

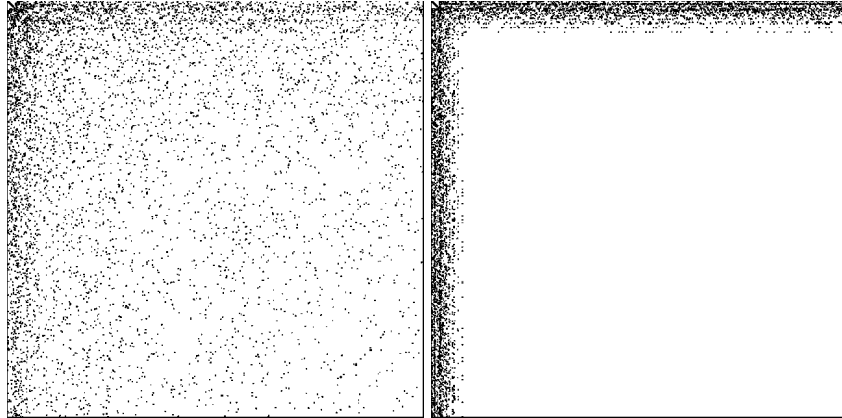


Рис. 3. Матрица смежности при $r = 0$ и $r = 0.6$

Изменение структуры сети при $r \geq r_c$ отражается и на виде матрицы смежности. Для сети с $N = 5000$ были построены две матрицы смежности: для $r < r_c$ - рис. ПЕ и для $r > r_c$ - рис. ПЕ. Для удобства элементы матрицы смежности $A_{ij} = 1$ отображены в виде черной точки. Обе матрицы были ранжированы, то есть узлы сети пронумерованы в порядке спадания количества связей k_i . Из рис. ПЕ можно заметить, что в матрице смежности при $r > r_c$ в правом нижнем углу появляется значительная квадратная область, заполненная 0, то есть теми парами узлов, которые не связаны друг с другом. Эта область, как показывает численный эксперимент, прямопропорционально зависит от величины r .

Таким образом такие характеристики сети, как коэффициент кластеризации, ассортативность, среднее минимальное расстояние ведут себя аналогично “параметру порядка” η .

III. Иерархические сети (U,V)-FLOWERS с придиричивостью

Кроме случайных безмасштабных сетей, построенных по алгоритму Барабши-Альберта(и их обобщений), известен класс простых детерминированных сетей, которые также являются безмасштабными сетями [8] - Fractal and Transfractal Recursive Scale-Free Networks. В частности это так называемые (u,v)-flowers, рис. III А.

Ниже мы обобщим модель Fractal and Transfractal Recursive Scale-Free сетей, введя фактор “придирчивости” r и случайность в закон роста сети. Как оказывается в этом случае поведение характеристик сети аналогично фазовому переходу.

А. Алгоритм (u,v)-flowers

Детерминированные растущие SF-сети, называемые (u,v)-flowers и (u,v)-trees были предложены и исследованы в [5, 8, 9].

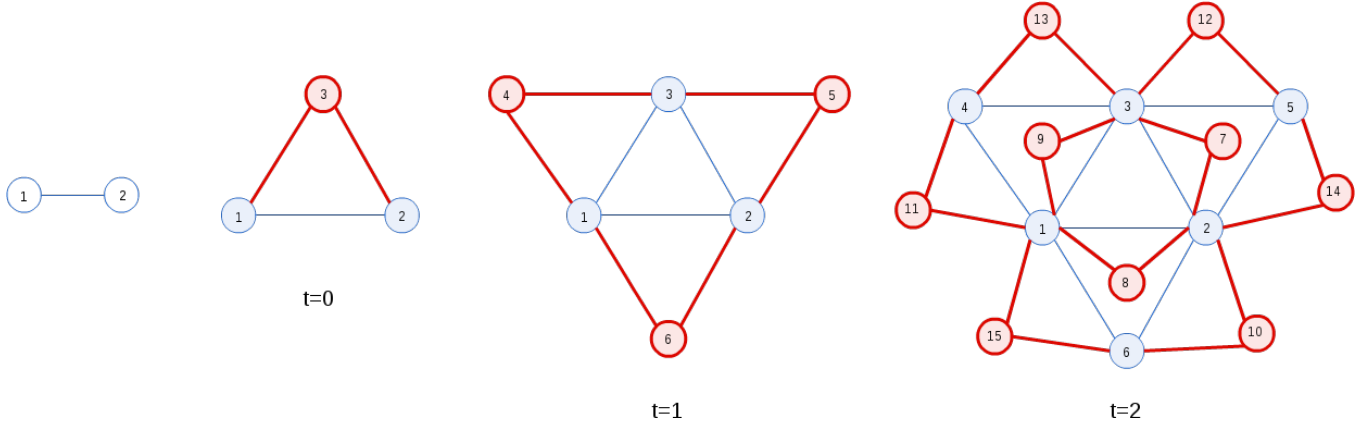


Рис. 4. Схема построения (1,2)-flowers на шагах $t = 0, 1, 2$. Утолщенные(красные online) - узлы появившиеся на данном шаге, не утолщенные (синие online) - узлы, которые появились на предыдущих шагах.

Нумерация узлов вообще говоря может быть любой, однако, в рассматриваемом примере можно занумеровать узлы таким образом (рис. III A), что матрица смежности A_{ij} станет наиболее простой. Под простой A_{ij} мы, в данном случае, понимаем такую ее структуру, что наибольшее число наибольших квадратных областей $N \times N$ в нем остаются пустыми.

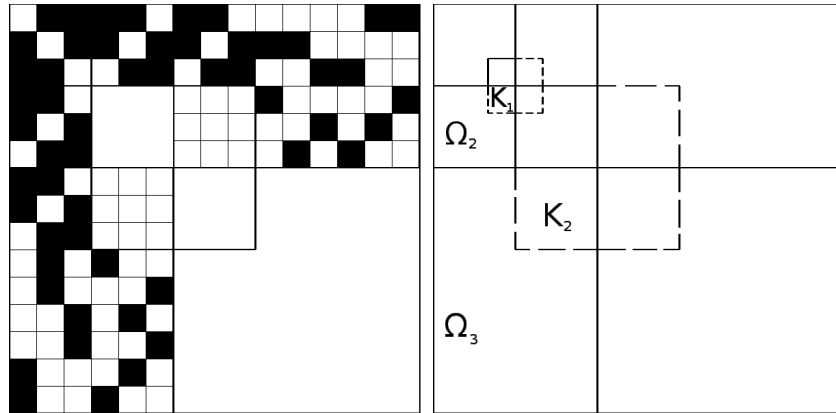


Рис. 5. Схема матрицы смежности для 3-го шага (1,2)-flower.

На рис. III A - $N \times N$ матрица смежности, где черным обозначены элементы матрицы с $A_{ij} = 1$. На первом шаге матрица смежности \hat{A} состоит из 3×3 элементов(верхний левый угол, шаг $t = 0$ на рис. III A). На втором шаге к ней добавляются новые элементы, и матрица состоит из 6×6 элементов(шаг $t = 1$ на рис. III A). На третьем шаге добавляются новые элементы, и матрица состоит из 15×15 элементов(шаг $t = 2$ на рис. III A). Как видно из рис. III A правый нижний квадрат первого шага свободный от связей и состоит из одного элемента. На втором шаге добавляется правый нижний квадрат, состоящий из 3×3 элементов, а на третьем из 9×9 элементов. На рис. III A белым обозначены места матрицы смежности, где $A_{ij} = 0$.

При выбранной нами нумерации появляются дополнительные к построению матрицы смежности в работе [5] области K_1 , K_2 , в которых также $A_{i,j} = 0$.

На каждом шаге t имеется N_t узлов и L_t связей [9]

$$N_t = (u + v) \cdot N_{t-1} - (u + v), \quad L_t = (u + v)^t \quad (4)$$

Т.е. на каждом шаге t появляется $N_t - N_{t-1}$ узлов и $L_t - L_{t-1}$ связей. Например, рис. III А, на шаге $t = 1$ появляется 3 узла и 6 связей.

Сделаем замену $(u + v) = w$ и раскроем рекуррентные формулы 4 [9]:

$$N_t = \frac{w-2}{w-1} \cdot w^t + \frac{w}{w-1}, \quad L_t = w^t \quad (5)$$

На каждом шаге t есть Ω_t клеток, которые могут быть заполнены – рис. III А. Их число равно:

$$\begin{aligned} \Omega_t &= (N_t - N_{t-1}) \cdot N_t - N_{t-2}^2 = \frac{w^3 - w^2 - 1}{w^4} N_t^2 + \frac{w^3 - 2w^2 - 2w - 2}{w^3} N_t - \frac{2w^2 + 2w + 1}{w^2} = \\ &= \frac{w^{2t+4} - 5w^{2t+3} + 8w^{2t+2} - 5w^{2t+1} + 4w^{2t} + w^{t+5} - 3w^{t+4} + 4w^{t+2} - w^5}{w^3 \cdot (w-1)^2} \end{aligned} \quad (6)$$

Вероятность заполнения ячейки матрицы равна:

$$W_t = \frac{L_t - L_{t-1}}{\Omega_t} = \frac{w^{t+5} - 3w^{t+4} + 3w^{t+3} - w^{t+2}}{w^{2t+4} - 5w^{2t+3} + 8w^{2t+2} - 5w^{2t+1} + 4w^{2t} + w^{t+5} - 3w^{t+4} + 4w^{t+2} - w^5} \quad (7)$$

Так с каждым шагом вероятность падает и матрица смежности становится все более разреженной. В [5] было показано, что (u,v)-flowers являются безмасштабными сетями. Для изображенного на рис. III А (1,2)-flower распределение узлов по степеням имеет вид $P(k) \sim k^{-(1+\frac{\ln 3}{\ln 2})}$. В общем случае для (u,v)-flowers [9]

$$P(k) \sim k^{-\alpha}, \quad \alpha = 1 + \frac{\ln(u+v)}{\ln 2} \quad (8)$$

В. Модификация алгоритма (u,v)-flowers

Рассмотрим теперь ситуацию, когда при построении (u,v)-flower не все связи реализуются. При чем вероятность нереализованности связи тем больше, чем меньше степени узлов она соединяет.

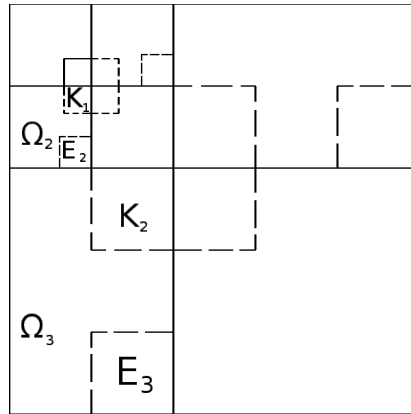


Рис. 6. Схема матрицы смежности для 3-го шага (1,2)-flower.

Компьютерное моделирование проводилось путем заполнения в матрице смежности областей $\Omega_1, \Omega_2 \dots$ с соответствующим им количеством связей L_t . Таким образом «придирчивость» в матрице смежности отображается правыми нижними углами областей Ω . При $r > 0$ в правых нижних углах областей Ω появляются пустые области. В нашем моделировании мы использовали пустые области E_t III A. Тогда вероятность заполнения ячейки матрицы:

$$W_t = \frac{L_t - L_{t-1}}{\Omega_t - E_t} \quad (9)$$

$$E_t = (r \cdot N_{t-1})^2 \quad (10)$$

С. Функция распределения степеней узлов

При значении параметра придирчивости $r = 0$ предлагаемая модель принимает вид стандартной (u,v)-flowers. Как и в случае с моделью Барабаш-Альберта, в модели (u,v)-flowers присутствует пороговое значения параметра придирчивости r_c . При значении параметра придирчивости меньше некоторого порогового $r < r_c$ функция распределения степеней узлов $P(k)$ остается степенной, а сама сеть, тем самым, безмасштабной сетью. При значениях параметра придирчивости больше порогового значения $r \geq r_c$ сеть меняет свою структуру, а именно в сети исчезают узлы со «средним» количеством связей, что мы можем увидеть на рис. III С.

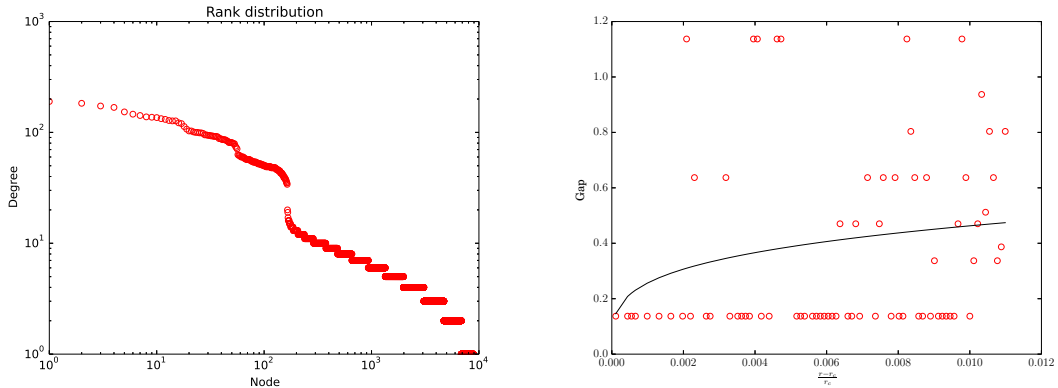


Рис. 7. Ранжированное распределение узлов

Определим пороговое значение параметра придирчивости. Для этого рассчитаем пороговое значение для (u,v)-flowers с 1 по 14 поколение, при этом будем усреднять значения по 10 экспериментам. Как и в модели Барабаш-Альберт с придирчивостью происходит насыщение порогового значения. Для 9 поколения r_c полностью насыщается и мы получаем $r_c = 0.91$.

D. Коэффициент кластеризации, ассортативность, минимальное среднее расстояние, величина разрыва

Появление разрыва η в распределении степеней узлов $P(k)$ свидетельствует о значительном изменении структуры сети, что не может не сказаться на её характеристиках. Ниже рассмотрено поведение C - коэффициента кластеризации, A - ассортативности и l - минимального среднего расстояния, как функции коэффициента придирчивости r , при $r \geq r_c$, как оказалось при увеличении r от r_c до $r_c + 0.01$ с шагом 0.001 происходит существенные изменения этих характеристик. Коэффициент кластеризации увеличивается от 0.005 до 0.007, ассортативность уменьшается от -0.18 до -0.195 , среднее минимальное расстояние колеблется около значения 1.34, величина разрыва увеличивается от 4 до xxx (согласно рис. III D).

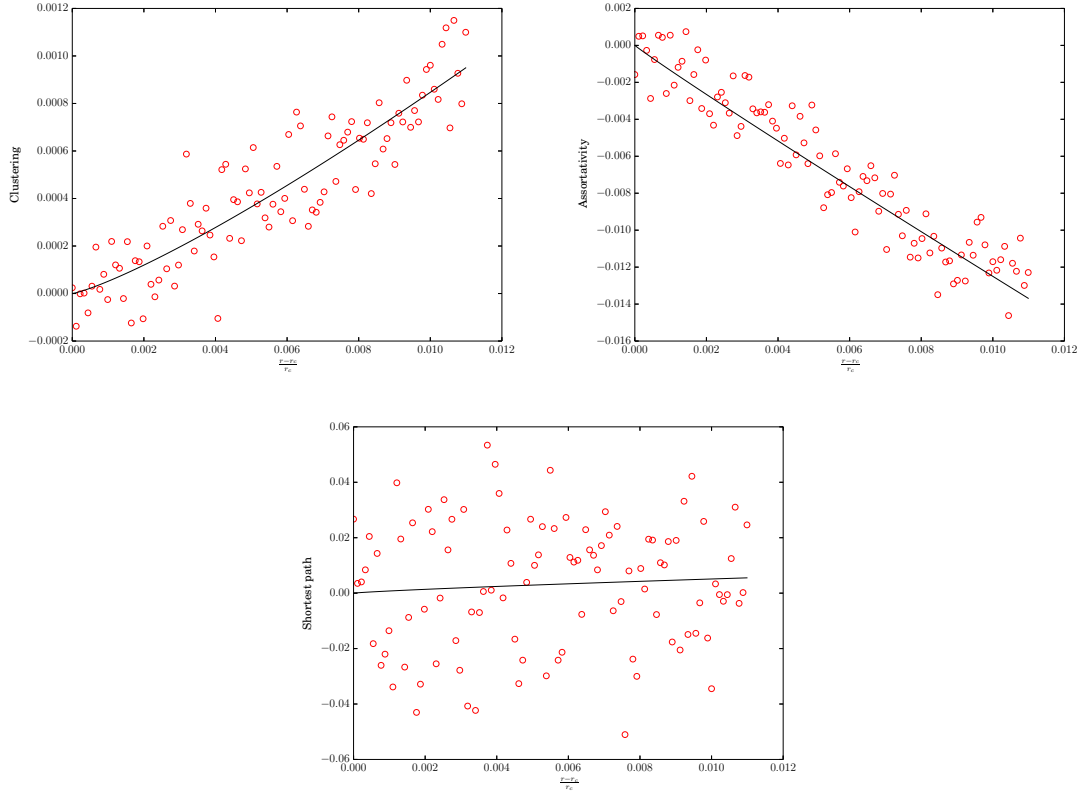


Рис. 8. Изменение кластеризации,ассортативности, минимального среднего пути при $r = [0, 51; 0, 52]$ с шагом 0.001

Зависимость этих характеристик (C, A, l) от $r - r_c$ близка к степенной $C = r - r_c^\alpha$, $A = r - r_c^\gamma$, $l = r - r_c^\sigma$, где $\alpha \sim$, $\gamma \sim$, $\sigma \sim$:

IV. Заключение

V. Литература

-
- [1] S. N. Dorogovtsev and J. F. F. Mendes, *Advances in Physics* **51**, 1079 (2002).
 - [2] R. Albert and A.-L. Barabási, *Reviews of Modern Physics* **74**, 47 (2002).
 - [3] V. Pareto, *Manual of political economy* (Augustus M Kelley Pubs, 1969) p. 504.
 - [4] A.-L. Barabási, R. Albert, and H. Jeong, *Physica A* **272**, 173–187 (1999).
 - [5] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, *Physical Review E* **65** (2002).
 - [6] L. D. Landau, *Zh. Eksp. Teor. Fiz.* **7**, 19–32 (1937).
 - [7] M. E. J. Newman, *Physical Review Letter* **89** (2002).
 - [8] H. D. Rozenfeld, S. Havlin, and D. ben Avraham, *New Journal of Physics* **9**, 175 (2007).
 - [9] H. D. Rozenfeld and D. ben Avraham, *Physical Review E* **75** (2007).

Пометки

1. Битвинес в БА
2. Придирчивостью структурирует сеть(делает из неё иерархическу)

3. Новые графики БА - 5000 узлов, 5 экспериментов, $r = [0.51; 0, 52]$, шаг 0.001
4. сделать аналогичные расчеты для (2,3)-цветков
5. Объяснить про смещения и т.д. для получения формул.
6. Цветки, написать про η