Модель сложных сетей с придирчивостью

Zarvanskiy Igor, Snarskii
NTUU KPI and
IPRI
(Dated: May 01, 2014)

В статье рассматриваются безмасштабные сложные сети, в частности модель Барабаши-Альберта и (u,v)-цветки. Вначале статьи приводится модификация правила предпочтительного соединения - введение параметра “придирчивости”, которое приминяется к классам безмасштабных сетей. Приводяться результаты численного моделирования предложенных моделей и рассматриваются значения различных характеристик моделируемых сетей. В результате было получено, что предложенные модели ведут себя аналогично фазовым переходам второго рода, расчитано пороговое значение параметра “придирчивости”, при котором происходит фазовый переход.

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 [4].

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 [1] 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 [7] Парето был установлен т. н. Закон Парето — степенное распределение богатства – когда, число людей ν, влдеющих μ - долей богатсва является степенной функцией $\nu \sim \mu^{\gamma}$, при $\gamma = ...$ получается так, что 20% людей владеют 80% богаства, что часто называется законом 80/20.

В работе [2] был найден сценарий образования сложной сети, т. н. сценарий Барабаши-Альберт, со степенным законом распределения степенней 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 подсоединение 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 [1] 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 распределении 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 безмаштабной 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. детерманированых (u,v)-flowers [3]. При этом также наблюдалось нарушение степенной функции распределения аналогично фазовым законам второго рода.

II. Модель Барабаши-Альберта с придирчиостью

A. Алгоритм Барабаши-Альберта

Рассмотрим растущую сеть. В стандартном варианте модели Барабаши-Альберт [2] 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:

$$p_i = \frac{k_i}{\sum\limits_j k_j}, \qquad (1)$$

где суммирование происходит по всем “старым” узлам.

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

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

с показателем $\gamma = 3$ [2].

В [1] приведено много модификаций правила предпочтительного подсоединения, которые приводят к различным значениям показателя γ. Однако сама степенная зависимость (2) остается.

B. Модификация алгоритма Барабаши-Альберта

Здесь мы предлагаем обобщение модели, основанной на правиле предпочтительного соединения, введеного 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:

$$p_i = \frac{k_i}{\sum\limits_j k_j}, \quad k_i \geq r\langle k\rangle, \tag{3}$$

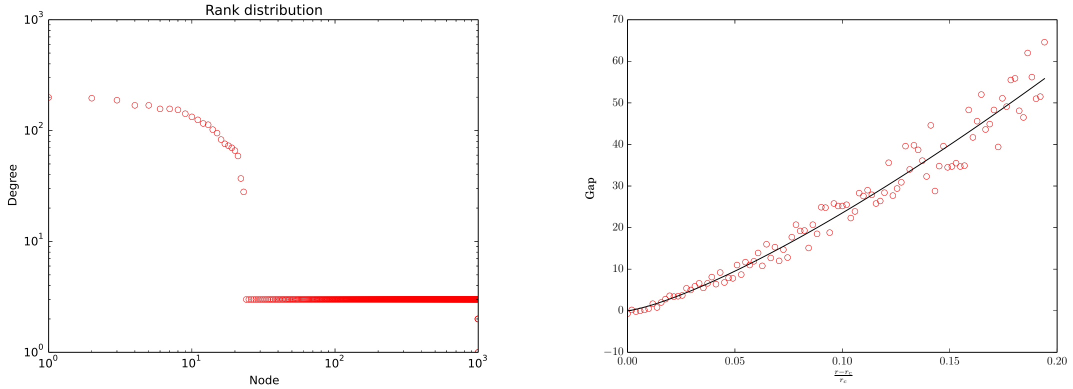
где $\langle k\rangle$ – среднее значение степени узлов в сети на момент присоединения, $\langle k\rangle = \sum\limits_j k_j/N$.

Введения дополнительного условия (3) в процесе 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 $\langle k\rangle$.

C. Функция распределения степеней узлов

При значение 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, тоесть 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. 1(a).

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 $r_c = 0.51$.



(a) (b)

Рис. 1. (a) Ранжированое распределение сети при $r = 0.6$. По горизонтальной оси отложено порядковый номер узла, по вертикальной оси отложена степень узла. (b) Величина разрыва при увеличении r от r_c до $r_c + 0.01$ с шагом 0.001. По горизонтальной оси отложено $\frac{r-r_c}{r_c}$, по вертикальной оси отложено значение величины разрыва.

Введем новую характеристику сети - величину разрыва η рис. 1(a), 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.

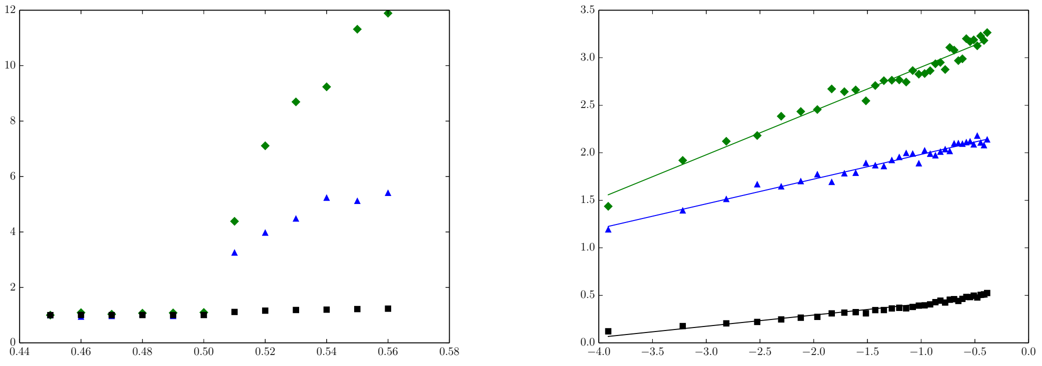
Как следует из численного моделирования, поведение параметра η аналогично поведению параметра порядка в теории фазовых переходов второго рода [5]. Как известно, параметр порядка η, например намагниченность, при приближении температуры к критическому значению T_c уменьшается степенным образом $\eta \sim (r - r_c)^\beta$, где β – критический индекс. При проведении численного эксперимента были выбраны следующие начальные параметры: количество узлов $N = 5000$, начальное количеств узлов $m_0 = 20$, количество связей у каждого нового узла $m = 3$.

На рис. 1(b) показана полученная зависимость $\eta = A \cdot (r - r_c)^\beta$, где $\beta \sim 1.3$

D. Коэфициент 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 показал численный эксперимент для сети с $N = 5000$ узлов, коэфициент 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 [1, 6].

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. 2(a)). Для нормализации зависимости возьмем отношение параметра к его значению при $r < r_c$, а также приведем зависимости к возрастающим функциям: $A = \frac{A}{A_0}$, $C = \frac{C}{C_0}$, $l = -\frac{l}{l_0}$.



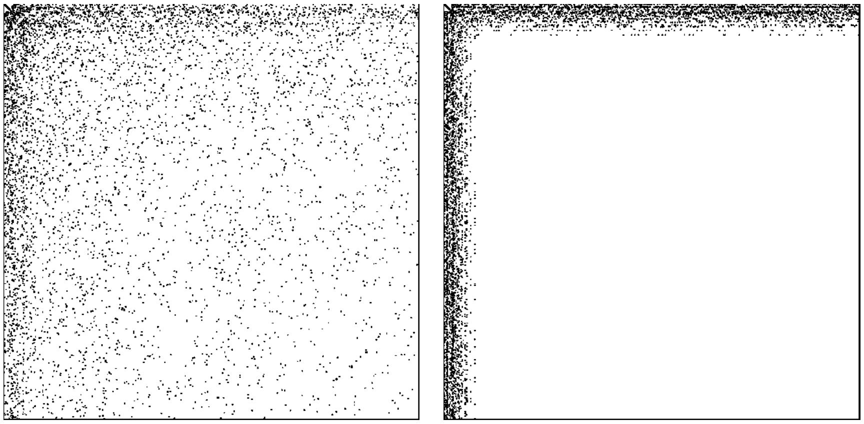
(a) (b)

Рис. 2. По горизонтальной оси отложено значение параметра придирчивости, по вертикальной оси отложено значение соответсвующей характеристики. (a) Изменение кластеризации,ассортативности, минимального среднего пути при $r = \lceil 0,45; 0,56 \rceil$ с шагом 0.01. (b) Изменение кластеризации,ассортативности, минимального среднего пути при $r = \lceil 0,5; 0,56 \rceil$ с шагом 0.001. В двойном логарифмическом масштабе.

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 $\alpha \approx 0.46155027$, $\gamma \approx 0.26025569$, $\sigma \approx 0.11761072$ (2(b)).

E. Матрица смежности для сети с придирчивостью

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



(a) (b)

Рис. 3. Матрица смежности для сети с $N = 5000$ узлов: (a) при $r = 0.0$ (b) при $r = 0.6$

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. 3(a) и для $r > r_c$ - рис. 3(b). Для удобвства элементы матрицы смежности $A_{ij} = 1$ отображены в виде черной точки. Обе матрицы были ранжированы, тоесть узлы сети пронумерованы в порядке спадания количества связей k_i. Из рис. 3(b) можно заметить, что в матрице смежности при $r > r_c$ в правом нижнем углу появляется значительная квадратная область, заполненная 0, тоесть теми парами узлов,которые не связаны друг с другом. Эта область, как показывает исленный 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 придирчиостью

Кроме случайных безмасштабных сетей, построенных по альгоритму Барабши-Альберта(и их обобщений), известен класс простых детерменированных сетей, которые также являются безмасштабными сетями [9] - Fractal and Transfractal Recursive Scale-Free Networks. В частности это так называемые (u,v)-flowers, рис. 4.

Ниже мы обощим 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.

A. Алгоритм (u,v)-flowers

Детерминированные растущие SF-сети, называемые (u,v)-flowers и (u,v)-trees были предложены и исследованы в [3, 8, 9].

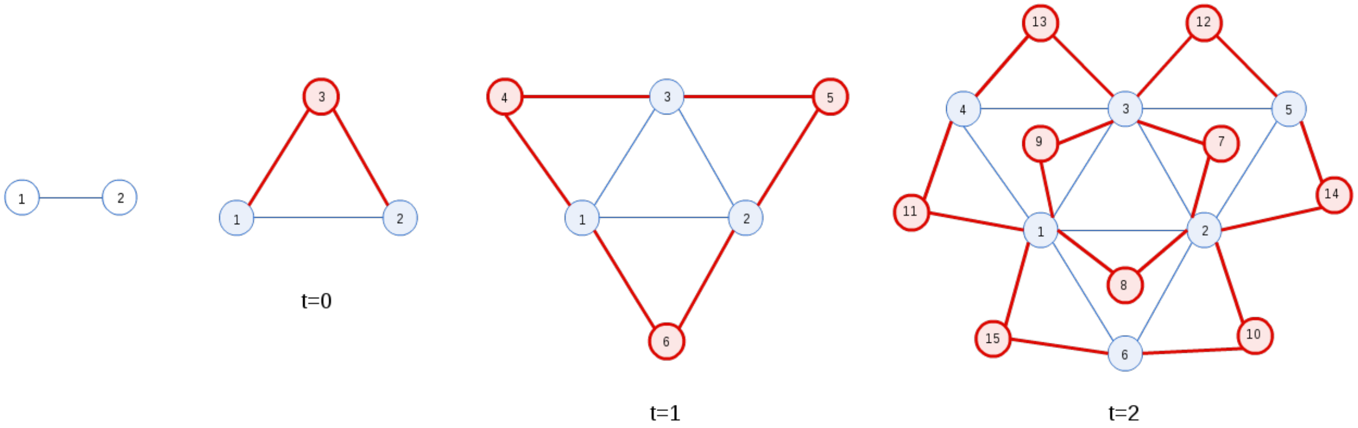
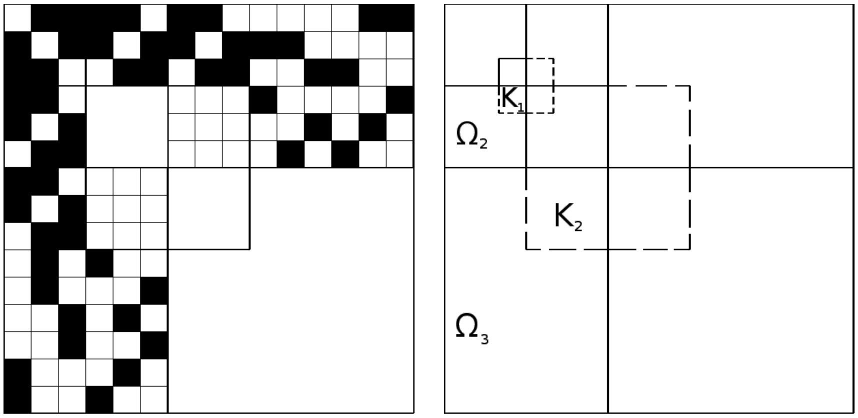


Рис. 4. Схема построения (1,2)-flowers на шагах $t = 0, 1, 2$. Утолщенные(красные online) - узлы появивщиеся на данном шаге,не утолщенные (синие online) - узлы, которые появились на предыдущих шагах.

Нумерация узлов вообще говоря может быть любой, однако, в рассматриваемом примере можно занумеровать узлы таким образом (рис. 4), 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.



(a) (b)

Рис. 5. Матрицы смежости для 3-го шага (1,2)-flower: (a) матрица смежности сети с выбранной нами нумерацией (b) схематическое представление

На рис. 5 - $N \times N$ 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. 4). На втором шаге к ней добавляются новые элементы, и матрица состоит из 6×6 элементов(шаг $t = 1$ на рис. 4). На третьем шаге шаге добавляются новые элементы, и матрица состоит из 15×15 элементов(шаг $t = 2$ на рис. 4). Как видно из рис. 5 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. 5 белым обозначены места матрицы смежности, где $A_{ij} = 0$.

При выбранной нами нумерации появляются дополнительные к построению матрицы смежности в работе [3] области K_1, K_2, в которых также $A_{i,j} = 0$.

На каждом шаге t имеется N_t узлов и L_t связей [8]

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

Т.е. на каждом шаге t появляется $N_t - N_{t-1}$ узлов и $L_t - L_{t-1}$ связей. Например, рис. 4, на шаге $t = 1$ появляется 3 узла и 6 связей.

Сделаем замену $(u + v) = w$ и раскроем рекурентные формулы 4 [8]:

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

На каждом шаге t есть Ω_t клеток, которые могут быть заполнены – рис. 5. Их число равно:

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

Вероятность заполнения ячейки матрицы равна:

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

Так с каждым шагом вероятность падает и матрица смежности становится все более разряженной. В [3] было показано, что (u,v)-flowers являются безмасштабными сетями. Для изображенного на рис. 4 (1,2)-flower распределение узлов по степеням имеет вид $P(k) \sim k^{-(1+\frac{\ln 3}{\ln 2})}$. В общем случаи для (u,v)-flowers [8]

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

B. 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.

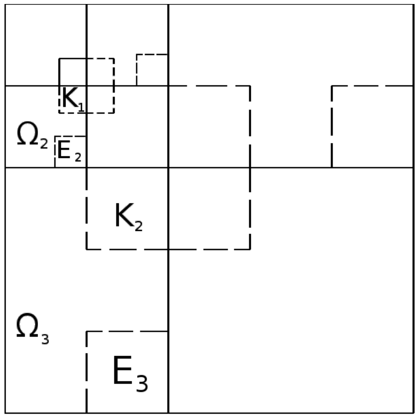


Рис. 6. Схема матрицы смежности для 3-го шага (1,2)-flower.

Компьютерное моделирование проводилось путем заполнения в матрице смежности областей Ω_1, Ω_2... с соотетсвующим 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 5. Тогда вероятность заполнения ячейки матрицы:

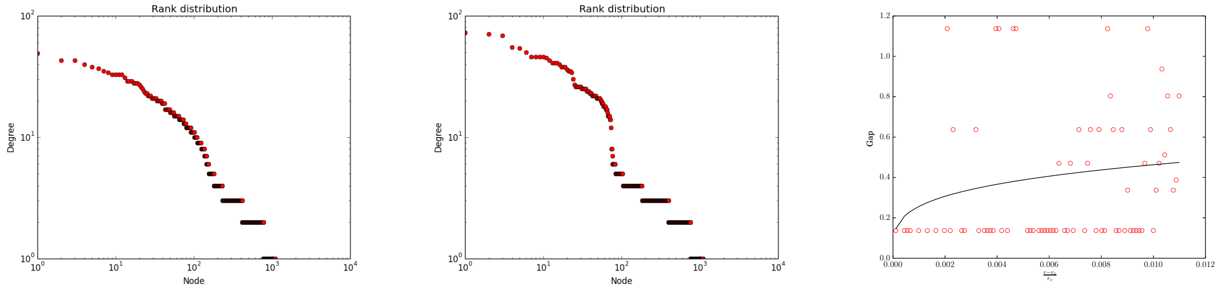
$$W_t = \frac{L_t - L_{t-1}}{\Omega_t - E_t} \tag{9}$$

$$E_t = (r \cdot N_{t-1}) \times (r \cdot (N_t - N_{t-1})) \tag{10}$$

C. Функция распределения степеней узлов

При значение параметра придирчивости $r = 0$ предлагаемая модель принимает вид стандартной (u,v)-flowers. Как и в случае с моделью Барабаши-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.

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 50 экспериментам. Как и в модели Барабаши-Альберт с придирчивостью происходит насыщение порогового значения. Для восьмого поколения r_c полностью насыщается и мы получаем $r_c = 0.6$. Необходимо отметить, что при $r >= 0.6$ происходит изменение структуры и характеристик сети, однако видимый разрыв появляеться при $r >= 0.8$, что мы можем увидеть на рис. 7(b)



(a) (b) (c)

Рис. 7. (a) Ранжированое распределение сети при $r = 0.6$. По горизонтальной оси отложено порядковый номер узла, по вертикальной оси отложена степень узла. (b) Ранжированое распределение сети при $r = 0.8$. По горизонтальной оси отложено порядковый номер узла, по вертикальной оси отложена степень узла. (c) Величина разрыва при увеличении r от r_c до $r_c + 0.2$ с шагом 0.01. По горизонтальной оси отложено $\frac{r-r_c}{r_c}$, по вертикальной оси отложено значение величины разрыва.

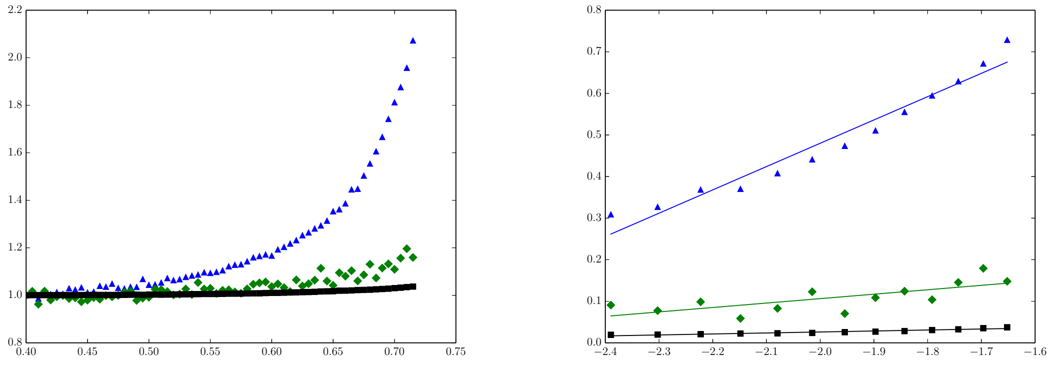
Как следует из численного моделирования, поведение параметра η аналогично поведению параметра порядка в теории фазовых переходов второго рода [5]. Параметр порядка η при приближении к критическому значению уменьшается степенным образом $\eta \sim (r - r_c)^\beta$, где β – критический индекс. При проведении численного эксперимента были выбраны следующие начальные параметры: (1,2)-flowers восьмого поколения, тоесть количество узлов $N = 3282$.

На рис. 7(c) показана полученная зависимость $\eta = A \cdot (r - r_c)^\beta$, где $\beta \sim$

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

Ниже рассмотрено поведение C - коэффициента кластеризации, A - ассортативности и l - минимального среднего расстояния, как функции коэффициента придирчивости r, при $r \geq r_c$. Как показал численный эксперимент для сети (1,2)-flowers восьмого поколения с $N = 3282$ узлов, коэфициент кластеризации C, ассортативность A, минимальное среднее расстояние l при $r < r_c$ от r не зависит и равна $C_0 \approx 0.02$, $l_0 \approx 4.05$, $A_0 \approx -0.18$, что, как и должно быть, совпадает с расчетами приведенными в [8, 9].

При увеличении r от r_c до $r_c + 0.1$ с шагом 0.01 коэффициент кластеризации увеличивается от 0.02 до 0.04, ассортативность уменьшается от -0.18 до -0.39, среднее минимальное расстояние уменьшается от 4.0 до 3.65(согласно рис. 8(a)). Для нормализации зависимости возьмем отношение параметра к его значению при $r < r_c$, а также приведем зависимости к возрастающим функциям: $A = \frac{A}{A_0}$, $C = \frac{C}{C_0}$, $l = -\frac{l}{l_0}$.



(a) (b)

Рис. 8. По горизонтальной оси отложено значение параметра придирчивости, по вертикальной оси отложено значение соответсвующей характеристики. (a) Изменение кластеризации,ассортативности, минимального среднего пути при $r = \lceil 0, 4; 0, 72 \rceil$ с шагом 0.01. (b) Изменение кластеризации,ассортативности, минимального среднего пути при $r = \lceil 0, 6; 0, 65 \rceil$ с шагом 0.005. В двойном логарифмическом масштабе.

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 $\alpha \approx 0.1066595$, $\gamma \approx 0.56066486$, $\sigma \approx 0.0237942$ (8(b)).

IV. Заключение

В статье предложено модифицированное правило предпочтительного присоединения, а именно присоединение с “придирчивостью”, применительно к классам безмастшабных сетей - модель Барабаши-Альберта и (u,v)-цветки. Модификация правила предпочтительного присоединения заключается в введении параметра “придирчивости”, который в процесе роста сети отсекает часть узлов (делает их невалидными), то есть к ним в данный момент не может присоединиться новая связь

Компьютерное моделирование показало, что в моделируемых классах сетей происходят существенные структурные изменения. Введение новой характеристики сети - величины разрыва, и расчет уже известных характеристик, таких как коэфициент кластеризации, ассортативность и среднее наименьшее расстояние между узлами, позволили сделать вывод о том что в моделируемых классах сетей происходит фазовый переход второго рода. Было вычислено пороговое значение при котором происходит фазовый переход, а также определено, что введенная характеристика “величина разрыва” пропорциональна параметру порядка.

Полученные результаты возможно применить для построения экономических моделей.

V. Литература

[1] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74:47–97, 2002.
[2] A.-L. Barabási, R. Albert, and H. Jeong. Mean-field theory for scale-free random networks. *Physica A*, 272:173–187, 1999.
[3] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes. Pseudofractal scale-free web. *Physical Review E*, 65, 2002.
[4] S. N. Dorogovtsev and J. F. F. Mendes. Evolution of networks. *Advances in Physics*, 51:1079–1187, 2002.
[5] L. D. Landau. On the theory of phase transitions. *Zh. Eksp. Teor. Fiz. 7*, page 19–32, 1937.
[6] M. E. J. Newman. Assortative mixing in networks. *Physical Review Letter*, 89, 2002.
[7] Vilfredo Pareto. *Manual of political economy.* Augustus M Kelley Pubs, 1969.
[8] Hernan D. Rozenfeld and Daniel ben Avraham. Percolation in hierarchical scale-free nets. *Physical Review E*, 75, 2007.
[9] Hernan D. Rozenfeld, Shlomo Havlin, and Daniel ben Avraham. Fractal and transfractal recursive scale-free nets. *New Journal of Physics*, 9:175, 2007.