МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»**

Кафедра математической кибернетики и компьютерных наук

# РАЗРАБОТКА И АНАЛИЗ МОДЕЛЕЙ РАСТУЩИХ СЕТЕЙ, ПРЕДСТАВЛЕННЫХ ДВУДОЛЬНЫМИ ГРАФАМИ

АВТОРЕФЕРАТ МАГИСТЕРСКОЙ РАБОТЫ

студента 2 курса 273 группы
направления 02.04.03 — Математическое обеспечение и администрирование информационных систем
факультета КНиИТ
Ковешникова Даниила Вадимовича

Научный руководитель
зав. каф., к. ф.-м. н., доцент ______________ С. В. Миронов

Заведующий кафедрой
к. ф.-м. н., доцент ______________ С. В. Миронов

Саратов 2026

# ВВЕДЕНИЕ

Многие современные системы (социальные интернет-сети, рекомендательные сервисы, системы «клиент-сервер», базы данных совместного использования, биологические сети «ген–болезнь») естественным образом описываются с помощью сложных сетей. В основе таких сетей лежат графы, узлы которых взаимодействуют друг с другом. Особый интерес представляют системы, которые не являются статичными: в них постоянно появляются новые узлы и связи, а старые отношения могут изменяться. Такие системы называют растущими сетями.

Многочисленные исследования показывают, что распределение степеней узлов во многих реальных сетях подчиняется степенному закону $P(k) \sim k^{-\gamma}$ с показателем $\gamma$, обычно лежащим в интервале от 2 до 3. Первые эмпирические подтверждения степенного закона были получены при анализе топологии Интернета ,веб-графа , а также других сложных сетей. Одной из первых и наиболее успешных моделей, объясняющих возникновение таких распределений, стала модель Барабаши–Альберт (BA). В ней рост сети происходит за счёт добавления новых узлов, которые соединяются с существующими с вероятностью, пропорциональной их степени (принцип «предпочтительного присоединения»). Однако классическая модель BA предполагает фиксированное количество рёбер $m$ для каждого нового узла и рассматривает обычные графы. В реальности многие сети имеют двудольную структуру: например, в социальной сети взаимодействие между пользователями происходит не напрямую, а через серверы или сервисы, что соответствует двудольному графу $(U, V, E)$, где $U$ — пользователи, $V$ – сервисы. Кроме того, количество связей, которое устанавливает новый узел, часто является случайной величиной (новый пользователь может добавить от одного до нескольких друзей).

Целью данной работы является анализ растущих сетей, представленных двудольными графами, на основе трёх моделей: базовой эвристической модели, классической модели Барабаши–Альберт и её расширенной версии со случайным числом присоединяемых рёбер.

Для достижения цели были поставлены следующие задачи:

1. Изучить модели растущих сетей (классическую БА, расширенную со случайным числом рёбер, модель триадного замыкания) и основные характеристики двудольных графов.

2. Построить четыре модели растущих двудольных сетей и реализовать их программно на языке Python.
3. Провести вычислительные эксперименты для каждой модели, собрать статистические данные о распределениях степеней на различных итерациях роста.
4. Проанализировать полученные распределения с помощью линейной регрессии в логарифмических координатах, оценить динамику коэффициента наклона.
5. Сравнить модели между собой, выявить их особенности и предложить рекомендации по применению для моделирования реальных систем.

Актуальность работы обусловлена необходимостью разработки и анализа моделей роста, которые адекватно отражали бы свойства реальных двудольных сетей. Такие модели востребованы для прогнозирования эволюции онлайн-платформ, оптимизации работы рекомендательных систем, анализа устойчивости сетей к атакам и перегрузкам. В отличие от однодольных сетей, двудольные сети имеют специфику: связи существуют только между долями, и каждая доля может обладать различными статистическими свойствами.

Научная новизна работы заключается в следующем:

1. Предложена и реализована расширенная модель Барабаши–Альберт для двудольных графов со случайным числом присоединяемых рёбер. Число рёбер $m(t) = 1 + \zeta(t)$, где $\zeta(t)$ – случайная величина с заданным распределением (Пуассона, геометрического или биномиального) и математическим ожиданием $\mathbb{E}\zeta = \lambda$.
2. Проведено систематическое сравнение четырёх моделей растущих двудольных сетей: базовой эвристической модели, классической двудольной модели БА, расширенной модели БА и модели триадного замыкания для двудольных графов. Сравнение выполнено по таким характеристикам, как коэффициент наклона распределения степеней в логарифмических координатах, скорость выхода на стационарный режим и разница между долями.
3. Для анализа динамики структуры сети применён подход линейной регрессии на каждом шаге роста, что позволило отследить эволюцию коэффициента наклона во времени, а не только финальное распределение.
4. Установлено, что расширенная модель даёт меньшие по модулю коэффи-

циенты наклона ($\approx -0.94$ для верхней доли, $\approx -0.61$ для нижней) по сравнению с классической двудольной моделью BA ($\approx -2.06$ и $\approx -2.68$), что свидетельствует о более равномерном распределении связей и может лучше соответствовать некоторым типам реальных двудольных сетей.

Данная работа базируется на фундаментальных результатах:

- Модель Барабаши–Альберти её аналитические свойства (степенной закон, динамика степени).
- Расширения модели со случайным числом рёбер для однодольных сетей.

В отличие от указанных работ, где рассматривались однодольные сети или двудольные графы без учёта роста со случайным числом рёбер, в настоящей работе выполнена адаптация и сравнительный анализ как классической, так и расширенной моделей BA именно для двудольных растущих сетей. Кроме того, новым является детальный анализ эволюции коэффициента наклона распределения степеней на каждой итерации роста.

Результаты работы могут быть использованы для моделирования и прогнозирования эволюции реальных двудольных сетей (социальные сети, рекомендательные системы). Предложенные реализации моделей и методы анализа могут служить основой для дальнейших исследований, например, для изучения динамики других локальных характеристик (индекс дружбы, коэффициент кластеризации) в двудольных растущих сетях.

# 1 Растущие сети, представленные двудольными графами

## 1.1 Двудольные графы

Двудольный граф представляет собой особый тип графа, который можно разделить на два таких непересекающихся подмножества вершин, что каждое ребро соединяет вершины только из разных подмножеств. Это означает, что в двудольном графе нет рёбер, соединяющих вершины внутри одного и того же подмножества. Двудольные граф обозначаются как $G = (U, V, E)$, где $U$ и $V$ являются непересекающимися подмножествами вершин, $E$ представляет собой множество ребер, соединяющих вершины из $U$ с вершинами из $V$.

## 1.2 Модели растущих сетей

Состояние сложной сети в момент $t$ можно описать в виде графа $G_t$, представляющего собой пару $(V_t, E_t)$б где набор вершин, являющихся узлами сети, $V_t = \{v_1, \ldots v_t\}$, $E_t$ – набор ребер, описывающих связи улов сети.

### 1.2.1 Модель Барабаши-Альберт

Пусть натуральное число $m$ определяет количество ребер, присоединяемых к сети на каждой итерации. В начальный момент времени $t = m, G_m = \{V_m, E_m\}$ представляет собой полный граф.

В модели Барабаши-Альберт граф $G_{t+1}$ получается следующим образом:

1. к графу присоединяется узел $v_{t+1}$;
2. добавляется $m_{t+1}$ ребер, $m_{t+1} <= m$ соединяющих новый узел с уже существующими в сети, которые появляются в соответствии с величиной $\zeta^{t+1}$, принимающей значение $i$ с вероятностью $\frac{d_i(t)}{2mt}$. Если $\zeta^{t+1} = i$, то к графу добавляется ребро $(v_{t+1}, v_i)$. Данная операция проводится $m$ раз. Если будет выбрано одно и то же значение несколько раз, то добавляется только одно ребро.

В результате присоединения новых ребер значение суммарной степени соседей узла $v_i$ может измениться в следующих случаях:

1. новый узел присоединяется к $v_i$ и величина $s_i$ увеличится на $m$, когда $d_i$ изменится только на 1;
2. новый узел присоединяется к соседу $v_i$, тогда $s_i$ увеличивается на 1, $d_i$ остается неизменной;
3. новый узел соединяется одним из своих ребер с узлом $v_i$ и одним из оставшихся соединяется с соседом узла $v_i$, тогда $s_i$ вырастет на $m + 1$, $d_i$

станет больше на 1.

### 1.2.2 Модель триадного замыкания

Модель триадного замыканияимитирует рост сетей на шаге $t+1$ следующим образом:

— к графу добавляется узел $v_{t+1}$;

— к сети добавляется $m$ ребер, соединяющих новый узел с уже существующими по следующим правилам:

  1. первое ребро соединяется с помощью предпочтительного присоединения (вероятность создания ребра с $v_i$ пропорциональна его степени);
  2. оставшиеся ребра присоединятся следующим образом:

     *а*) с вероятностью $p$ ребро создается с соседом $v_i$;

     *б*) с вероятностью $1-p$ новый узел соединяется с помощью предпочтительного присоединения.

Основным отличием данной модели от модели Барабаши-Альберта состоит в создании большего числа треугольных соединений в процессе роста. Если новая вершина $v_{t+1}$ на шаге 1 выбирает в качестве соединяющего узла вершину $v_j$, являющуюся смежной с вершиной $v_i$, и после $v_{t+1}$ соединяется оставшимися ребрами с соседями $v_j$ на шаге $2a$), то с очень высокой вероятностью эти вершины будут являться соседями $v_i$. Это связано с генерацией множества треугольников в процессе роста сети для данной модели, что приводит к большей локальной и глобальной кластеризации.

## 1.3 Модели растущих сетей со случайным количеством присоединяемых ребер

При росте реальных сетей каждый новый узел присоединяется к нескольким существующим узлам, число которых изначально неизвестно. Но классические модели предполагают, что количество связей, которые имеет новый объект при добавлении в сеть, постоянно. Поэтому «случайные» модели имеют особенность по сравнению с классическими, ведь они позволяют добавление в сеть новых узлов с большим значением степени на любой итерации. Данная особенность отличает их от классических моделей, в которых узлы, появившиеся на ранних этапах развития сети имеют преимущество.

### 1.3.1 Расширенная модель Барабаши-Альберт

Данная модель со случайным числом присоединяемых ребер расширяет классическую модель Барабаши-Альберт. Основные теоретические результаты для однодольного случая получены в работе. В расширенной модели количество добавленных ребер представляет собой случайную величину, распределенную по заданному закону распределения . Рост сети определяется следующими правилами:

- на каждой итерации $t$ добавляется один узел $v_t$;
- в граф добавляется одно ребро и новая вершина соединяется этим ребром с одним из узлов с вероятностью, пропорциональной степени данного узла $d_i(t)$. Данный процесс повторяется $m(t) = 1 + \zeta$ раз за итерацию, $\zeta = \zeta(t)$ есть случайная величина, принимающая неотрицательные целочисленные значения с ограниченным математическим ожиданием $\mathbb{E}(\zeta) = \lambda$.

## 1.4 Методы анализа распределений степеней

Для количественного описания распределений степеней вершин в полученных моделях используется линейная регрессия в логарифмических координатах.

Степенные распределения широко распространены в сложных сетях (модели Барабаши-Альберт, реальные интернет-графы, социальные сети). Линейная регрессия в логарифмических координатах является простейшим и наглядным способом оценки показателя $\gamma$, несмотря на наличие смещения для малых $k$ и чувствительность к хвосту распределения. В данной работе регрессия строится для значений степени $k$, не превышающих порогов $k_{\max}^{\text{upper}} = 10$ для верхней доли и $k_{\max}^{\text{lower}} = 6$ для нижней доли. Эти пороги выбраны эмпирически, чтобы исключить области с малым количеством наблюдений (шумы на хвосте).

Для каждой итерации $t$ и для каждой доли графа:

1. Из распределения степеней выбираются пары $(k, n_k)$, где $n_k$ – среднее по экспериментам количество вершин степени $k$.
2. Выполняется логарифмическое преобразование: $x = \log_{10} k$, $y = \log_{10} n_k$.
3. Строится линейная регрессия $y = \beta x + \alpha$ методом наименьших квадратов.
4. Коэффициент $\beta$ принимается за оценку $\gamma$ (параметр степенного распределения).

## 2 Постановка экспериментов со случайными сетями, представленными двудольными графами

### 2.1 Параметры вычислительных экспериментов

Все эксперименты проводились на языке Python с использованием библиотек NumPy, Matplotlib, scikit-learn. Вычисления выполнялись на процессоре Intel Ultra 7 155H, 32 ГБ ОЗУ. Для обеспечения воспроизводимости в каждом эксперименте фиксировались начальные значения генераторов случайных чисел.

### 2.2 Базовая модель двудольной сети

В качестве базовой модели определена модель двудольного графа со следующими правилами:

— в начальный момент времени в каждом множестве графа находится по $n$ вершин; каждая из вершин $V_c$ соединена с 1 вершиной из множества $V_s$;

— на каждой итерации $t$:

1. множество $V_c$ разбивается на множество пар $\{v_{ck}, v_{cl}\}$ случайным образом каждая вершина из $V_c$ участвует в одной паре;
2. с вероятностью $p$ обе вершины в каждой паре объединятются в одну вершину $v_{new}$, соседи $v_{ck}$ и $v_{cl}$ присоединяются к вершине $v_{new}$;
3. с вероятностью $1 - p$ обе вершины в каждой паре объединятются в одну вершину $v_{new}$, вершина $v_{new}$ присоединяется к одной из «свободных» вершин множества $V_s$;
4. к множеству $V_c$ добавляется $n/2$ новых вершин, каждая из которых присоединяется к одной из вершин множества $V_s$ случайным образом;

Модель реализована на языке Python. Для отображения распределнений программа строит графики с помощью библиотеки sklearn.

### 2.3 Модель Барабаши-Альберт для двудольных сетей

Для данной модели применим правила роста модели Барабаши-Альберт на сети, представленные двудольными графими. Модель реализована на языке Python.

## 2.4 Расширенная модель Барабаши-Альберт для двудольных сетей

Для данной модели применим правила роста расширенной модели Барабаши-Альберт на сети, представленные двудольными графими. Модель реализована на языке Python.

## 2.5 Модель триадного замыкания для двудольных сетей

Для данной модели были определены дополнительные подсети хранящие связи, полученные в результате алгоритма триадного замыкания и применены правила роста модели триадного замыкания. Модель реализована на языке Python.

## 2.6 Сравнительный анализ моделей

Для наглядного сопоставления характеристик всех рассмотренных моделей (базовая эвристическая, классическая двудольная модель Барабаши-Альберт, расширенная двудольная модель со случайным числом рёбер и модель триадного замыкания для двудольных сетей) в таблице сведены основные параметры и полученные значения распределения степеней в логарифмических координатах.

| **Модель** | $\gamma_{\text{верх}}$ | $\gamma_{\text{низ}}$ | **Стабилизация (итер.)** |
|---|---|---|---|
| Базовая эвристическая | $\approx -3{,}48$ | $\approx -4{,}85$ | $< 10^3$ |
| Классическая БА (двудольная) | $\approx -1{,}07$ | $\approx -1{,}70$ | $\approx 25000$ |
| Расширенная БА (двудольная) | $\approx -1{,}47$ | $\approx -2{,}86$ | $\approx 50000$ |
| Триадного замыкания (двудольная) | $\approx -1{,}34$ | $\approx -2{,}81$ | $\approx 25000$ |

Базовая модель демонстрирует наибольшее значение распределения степеней ($\gamma \approx -3{,}48$ и $-4{,}85$) при очень быстрой стабилизации ($< 10^3$ итераций). Однако она не воспроизводит классический степенной закон.

Классическая модель Барабаши-Альберт, переложенная на двудольные графы даёт умеренные по модулю коэффициенты ($\gamma_{\text{верх}} \approx -1{,}07$, $\gamma_{\text{низ}} \approx -1{,}70$) и разницу между долями $\Delta\gamma = 0{,}63$. Стабилизация после $25000$ итераций. Модель хорошо подходит для систем с фиксированным числом связей и умеренной асимметрией долей.

Расширенная модель, переложенная на двудольные графы демонстрирует более крутое распределение, чем классическая БА ($\gamma_{\text{верх}} \approx -1{,}47$, $\gamma_{\text{низ}} \approx -2{,}86$), и большую разницу между долями ($\Delta\gamma = 1{,}39$). Стабилизация требует около

$50000$ итераций. Модель подходит для систем с варьирующейся активностью узлов, но асимметрия между долями возрастает.

Модель триадного замыкания, переложенная на двудольные графы даёт значения $\gamma_{\text{верх}} \approx -1{,}34$, $\gamma_{\text{низ}} \approx -2{,}81$, разница $\Delta\gamma = 1{,}47$ — наибольшая среди всех моделей. Стабилизация после $25000$ итераций. Вспомогательные графы внутри долей создают дополнительные связи в нижней доле ($p_{\text{ino}} = 2/3$), что усиливает асимметрию.

Все четыре модели выходят на постоянные значения $\gamma$, но с разной скоростью. Расширенная модель требует больше итераций для стабилизации (около $50000$) из-за случайного числа рёбер. Классическая БА и модель триадного замыкания стабилизируются быстрее ($\approx 25000$ итераций). Базовая модель стабилизируется практически мгновенно ($< 10^3$ итераций), но её распределения наиболее крутые и не соответствуют степенному закону.

— Если важна быстрая стабилизация и не требуется точное воспроизведение степенного закона – можно использовать базовую эвристическую модель.
— Для классических задач с фиксированным числом связей и известной степенной экспонентой $\gamma \approx 2\breve{}3$ следует применять классическую модель BA, переложенная на двудольные графы.
— Для систем с варьирующейся активностью узлов (новые пользователи добавляют разное количество связей) предпочтительна расширенная модель BA, переложенная на двудольные графы, особенно если ожидается более равномерное распределение степеней.
— Если требуется моделировать сети с высокой внутренней кластеризацией внутри долей (например, социальные группы с сильными связями между участниками), можно использовать модель триадного замыкания, переложенная на двудольные графы, но следует учитывать её большую параметрическую сложность.

## ЗАКЛЮЧЕНИЕ

В ходе выполнения работы были решены поставленные задачи: изучены модели растущих сетей (классическая модель Барабаши-Альберт, её расширение со случайным числом рёбер, модель триадного замыкания), адаптированы для двудольных графов, реализованы программно на языке Python, проведены вычислительные эксперименты и выполнен анализ полученных распределений степеней с помощью линейной регрессии в логарифмических координатах.

1. Базовая эвристическая модель показала быструю стабилизацию (уже после нескольких сотен итераций). Параметр распределения для верхней доли стабилизировался в районе $-1{,}2 \ldots -1{,}4$, для нижней – около $-1{,}0$. Разница между долями минимальна ($\Delta\gamma = 0{,}3$). Модель не опирается на предпочтительное присоединение и не воспроизводит степенной закон, но полезна для быстрого эвристического моделирования.
2. Классическая двудольная модель Барабаши-Альберт продемонстрировала степенное распределение с параметром степепнного распределения $\gamma_{\text{верх}} \approx -2{,}06$, $\gamma_{\text{низ}} \approx -2{,}68$ (при $p_{\text{ino}} = 2/3$). Коэффициент $\gamma$ выходит на постоянный уровень после 25000 итераций. Разница между долями $\Delta\gamma = 0{,}62$ объясняется асимметрией вероятности попадания новой вершины.
3. Расширенная двудольная модель Барабаши-Альберт дала значительно меньшие по модулю параметры распределения: $\gamma_{\text{верх}} \approx -0{,}94$, $\gamma_{\text{низ}} \approx -0{,}61$ (при $\lambda = 1, p_{\text{ino}} = 2/3$, распределение Пуассона). Меньшая разница между долями ($\Delta\gamma = 0{,}33$) свидетельствует о более равномерном распределении связей. Модель требует большего времени на стабилизацию (около 50000 итераций) из-за случайного числа рёбер.
4. Модель триадного замыкания для двудольных сетей адаптирована с помощью вспомогательных графов для внутридольных связей. Получены параметры распределения $\gamma_{\text{верх}} \approx -2{,}22$, $\gamma_{\text{низ}} \approx -2{,}99$ – наибольшая разница между долями ($\Delta\gamma \approx 0{,}77$). Стабилизация происходит после 25000 итераций, однако модель имеет больше параметров ($p_f, p_{fr}, p_{fd}, p_{\text{triad}}$), что усложняет её настройку, но позволяет управлять внутренней кластеризацией.

— Для быстрой эвристической оценки без строгого воспроизведения степенного закона подходит базовая модель.

— Для систем с фиксированным числом связей и выраженной асимметрией долей (например, «много пользователей – мало сервисов») рекомендуется классическая двудольная модель БА.
— Для систем с варьирующейся активностью узлов (новые пользователи добавляют разное количество связей) предпочтительна расширенная модель БА.
— Для моделирования сетей с высокой внутренней кластеризацией внутри долей (социальные группы, сообщества) можно использовать модель триадного замыкания с учётом её параметрической сложности.

Полученные результаты для классической двудольной модели согласуются с теоретическими предсказаниями для однодольной модели Барабаши-Альберт ($P(k) \sim k^{-3}$), хотя абсолютные значения коэффициентов отличаются из-за двудольной структуры и вероятности $p_{\text{ino}}$. Для расширенной модели в однодольном случае параметр $\gamma$ стремится к $-3$ при больших $\lambda$; в нашей двудольной адаптации при $\lambda = 1$ получены значения около $-0{,}94$ и $-0{,}61$, что является новым результатом. Модель триадного замыкания для двудольных графов не исследовалась ранее; её адаптация и численный анализ выполнены впервые. По сравнению с работами Гийома и Латапи, где изучались статические двудольные сети, данная работа вносит вклад в динамические модели роста.

В качестве продолжения работы предлагается:
— реализовать анализ динамики индекса дружбы для всех четырёх моделей;
— исследовать влияние различных распределений $\zeta$ (геометрическое, биномиальное) на параметры степенного распределения в расширенной модели;
— добавить механизмы удаления узлов и рёбер для моделирования деградации сетей;
— применить модели к реальным наборам данных;
— оптимизировать вычислительную сложность модели триадного замыкания для работы с большими сетями.

Результаты описанной выше работы были представлены на выступлении в рамках Студенческой Научной Конференции СГУ 2026 года.