МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»**
Кафедра математической кибернетики и компьютерных наук

ИССЛЕДОВАНИЕ НЕВОЗРАСТАЮЩИХ СЛУЧАЙНЫХ СЕТЕЙ

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студентки 4 курса 411 группы
направления 02.03.02 — Фундаментальная информатика и информационные технологии
факультета КНиИТ
Белавиной Виктории Олеговны

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

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

Саратов 2026

Актуальность темы. Системно теория случайных графов сложилась в конце 1950-х благодаря работам П. Эрдёша и А. Реньи. Среди же современных моделей случайных графов выделяются алгоритмы Барабаши-Альберт и триадного замыкания.

Тем не менее, таковые в основном описывают рост сети – добавление вершин и рёбер. В реальных системах, напротив, сети могут сжиматься: узлы и связи теряются, что требует анализа невозрастающих случайных сетей.

Вместе с тем, невозрастающим структурам в науке уделено сравнительно незначительное внимание, хотя проблема, связанная с тем, что многие сети на определённом этапе перестают расти или сокращаются, существует. Поэтому и не предложено до сих пор целостной модели эволюции сетей в условиях отсутствия их роста. В этой связи настоящая работа претендует на то, чтобы восполнить данный пробел за счёт синтеза механизмов предпочтительного присоединения, триадного замыкания и реструктуризации.

Цель бакалаврской работы – разработка и исследование модели невозрастающей случайной сети, объединяющей механизмы предпочтительного присоединения, триадного замыкания и реструктуризации, которые приводят к сокращению или стабилизации размера сети.

Для достижения поставленной цели необходимо решение **следующих задач**:

1. описать и проанализировать механизмы, которые ведут к остановке роста или сокращению сети: удаление вершин и рёбер, затухание предпочтительного присоединения, уплотнение связей внутри групп;
2. построить математическую и алгоритмическую модель невозрастающей сети, объединяющую начальный рост, триадное замыкание, слияние вершин и новые механизмы сокращения;
3. осуществить программную реализацию модели на Python с библиотеками NetworkX, NumPy и Matplotlib для расчётов и графиков;
4. изучить с помощью серии вычислительных экспериментов, как параметры модели влияют на ключевые характеристики сети: распределение степеней, коэффициент кластеризации, индекс дружбы, диаметр и связность;
5. сравнить результаты с классическими растущими моделями, выявить, в чем принципиальное отличие невозрастающих сетей, и оценить, насколько предложенная модель подходит для описания реальных процессов.

Методологической основой послужили такие классические работы по теории случайных графов, как исследования П. Эрдёша и А. Реньи, заложившие основы данного направления, а также модели А.-Л. Барабаши и Р. Альберт, описывающие механизм предпочтительного присоединения. Дополнительно использована концепция триадного замыкания, предложенная Г. Зиммелем. Программная реализация выполнена на языке Python с использованием библиотек NetworkX, NumPy и Matplotlib.

Практическая значимость работы заключается в создании программного комплекса на языке Python для моделирования невозрастающих случайных сетей. Разработанный инструмент может применяться для исследования структурных свойств сетей, находящихся в фазах стабилизации или сокращения (социальные группы, виртуальные сообщества, видовые популяции), и позволяет гибко управлять ключевыми характеристиками сети за счёт варьирования параметров модели. Предложенная модель также может быть использована для анализа динамики реальных сетей, где баланс между процессами добавления и удаления элементов определяет итоговую структуру.

Структура и объём работы. Бакалаврская работа состоит из введения, двух глав, заключения, списка использованных источников и шести приложений. Общий объём работы – 76 страниц, из них основное содержание – 55 страниц, включающее 5 рисунков и 28 таблиц. Список использованных источников содержит 20 наименований.

В **первой главе «Основы теории случайных графов»** рассматриваются основные понятия теории случайных графов: степень вершины, распределение степеней, коэффициент кластеризации и индекс дружбы. Степень вершины определяется как количество рёбер, инцидентных данной вершине. Распределение же степеней показывает, какая доля вершин графа имеет заданную степень. Для безмасштабных сетей распределение степеней подчиняется степенному закону $P(k) \sim k^{-\gamma}$, где γ – постоянный параметр. Коэффициент кластеризации, в свою очередь, характеризует тенденцию вершин графа образовывать плотно связанные группы (треугольники). А индекс дружбы (коэффициент ассортативности) определяет корреляцию между степенями соседних вершин.

Известно, что случайные графы находят широкое применение в различных областях. Они используются при моделировании социальных сетей, где вершины соответствуют людям, а рёбра – отношениям между ними. В биологии случайные графы применяются для анализа взаимодействий между белками и генами. В компьютерных науках они используются в задачах кластеризации, ранжирования, в рекомендательных системах и при решении задач класса NP.

Особый интерес представляют модели, описывающие эволюцию сетей во времени. Классическая модель Эрдёша-Реньи предполагает фиксированное число вершин и случайное добавление рёбер. Однако реальные сети, как правило, развиваются: в них появляются новые узлы и связи, тогда как другие узлы могут терять связи или покидать сеть. Именно поэтому модели растущих сетей, такие как Барабаши-Альберт и триадного замыкания, лучше отражают динамику реальных процессов. В то же время, для описания сетей на этапах стабилизации или сокращения требуются модели, учитывающие не только рост, но и потерю элементов структуры.

Далее исследуются ключевые модели растущих сетей – Барабаши-Альберт и триадного замыкания. Модель Барабаши-Альберт представляет собой алгоритм генерации случайных безмасштабных сетей с использованием правила предпочтительного связывания, которое соответствует принципу «богатый становится богаче». Алгоритм формирования сети заключается в следующем: первоначально берётся граф из m вершин, затем на каждой итерации добавляется одна новая вершина, которая соединяется m рёбрами с уже имеющимися в соответствии с принципом предпочтительного присоединения. Это приводит к формированию безмасштабных сетей со степенным распределением $P(k) \sim k^{-3}$.

Реальные сети, будь то социальные взаимодействия, биологические системы или информационные структуры, обладают рядом общих свойств. Они, как правило, являются безмасштабными, то есть распределение степеней в них подчиняется степенному закону. Кроме того, для них характерна высокая кластеризация – вершины образуют плотно связанные группы. Важно также, что многие реальные сети со временем меняют свою структуру: одни узлы появляются, другие исчезают, связи перестраиваются. Именно поэтому важно разрабатывать модели, которые учитывают не только рост, но и возможное сокращение или реорганизацию сети.

Однако классическая модель Барабаши-Альберт обладает рядом ограничений. Во-первых, результирующий граф зависит от начального параметра m: при $m = 1$ модель описывает генерацию дерева. Во-вторых, модель не учитывает высокую кластеризацию и динамическое изменение структуры, характерные для многих реальных сетей.

Модель триадного замыкания, развивающая идеи Барабаши-Альберт, в свою очередь, вводит дополнительный механизм образования связей между соседями уже связанных узлов. Её ключевая особенность заключается в том, что с заданной вероятностью p новые связи образуются не только по принципу предпочтительного присоединения, но и между соседями уже связанных узлов. Это приводит к созданию большого числа треугольников и значительному росту коэффициента кластеризации, при этом распределение степеней сохраняет степенной закон с показателем $\gamma = 3$. Модель может создавать сети с различными уровнями кластеризации в зависимости от выбора параметров p и m.

Важно отметить, что обе рассмотренные модели – Барабаши-Альберт и триадного замыкания – описывают только растущие сети. В них число вершин монотонно увеличивается с каждой итерацией. Однако в реальности многие системы не растут бесконечно. Рано или поздно наступает насыщение: новые элементы перестают появляться, а старые могут исчезать. Кроме того, внутренняя структура сети может перестраиваться – группы сливаются, связи перераспределяются. Это требует расширения существующих моделей и разработки новых подходов.

Далее предлагается модифицированная модель, которая включает три новых механизма. Так, первый механизм – объединение вершин: на каждой итерации случайно выбираются две вершины u и v, создаётся новая вершина $w = u \cup v$, и сохраняются все связи $\Gamma(w) = \Gamma(u) \cup \Gamma(v)$. Второй механизм – триадное замыкание: после объединения выбираются два соседа из $\Gamma(u)$ и $\Gamma(v)$, и с вероятностью p добавляется ребро между ними. Третий же механизм – динамический рост: добавление новой вершины с m рёбрами, что компенсирует объединения сохранением среднего размера сети. Сочетание этих механизмов позволяет моделировать более богатую динамику, приводящую к увеличению средней степени вершин, плотности рёбер и коэффициента кластеризации, что лучше соответствует свойствам многих реальных эволюционирующих сетей.

В завершении главы формулируется концепция невозрастающих случайных сетей. Невозрастающая случайная сеть определяется как сеть, в которой число вершин не увеличивается со временем. Такая динамика может проявляться в двух формах. Первая – стабилизация: процессы добавления и удаления вершин уравновешивают друг друга, благодаря чему общее число узлов сохраняется примерно на постоянном уровне. Вторая – сокращение: удаление вершин преобладает над их добавлением, что приводит к уменьшению размера сети. Таким образом, моделирование невозрастающих сетей требует учёта как механизмов роста, так и процессов, приводящих к потере элементов структуры.

Вторая глава «Реализация модификации классических растущих сетей для получения неубывающих сетей» содержит практическую реализацию предложенной модели и анализ результатов вычислительных экспериментов. Все расчёты производятся на компьютере с процессором Intel Core i7-9750H и 16 ГБ оперативной памяти. Модель реализована на Python 3.11.2 с использованием библиотек NetworkX, NumPy, random, collections.defaultdict и typing.

В целях оптимизации вычислений на больших графах разрабатывается структура данных `FenwickTree` (дерево Фенвика), позволяющая выполнять операции предпочтительного выбора вершин за время $O(\log N)$. Дерево Фенвика реализует операции добавления значения к элементу, вычисления суммы на префиксе и поиска элемента по накопленной сумме. Это необходимо для эффективной реализации механизма предпочтительного присоединения. Кроме того, реализуется класс `FastGraphState` для эффективного хранения состояния графа, который обеспечивает быстрое обновление и получение информации о степенях вершин, активных узлах, а также выполняет операции равномерного и предпочтительного выбора вершин.

Для ускорения вычислений на графах большого размера реализуются функции аппроксимации среднего коэффициента кластеризации и коэффициента ассортативности с использованием случайной выборки. Функция `approximate_average_clustering` возвращает приближённое значение среднего коэффициента кластеризации графа. Если число вершин не превышает размера выборки, вычисляется точное значение через `nx.average_clustering(G)`. В противном случае случайно выбираются вершины, для них считаются локальные коэффициенты кластеризации и возвращается среднее арифметическое. Функция `approximate_degree_assortativity` возвращает приближённое значение коэффициента ассортативности по степеням с использованием резервуарной выборки рёбер.

На основе разработанных структур данных выполняются три основные функции, соответствующие трём механизмам модели.

Так, функция merge_nodes объединяет две случайные вершины в одну. Выбор вершин осуществляется в зависимости от параметра select_type (равномерно или предпочтительно). Функция возвращает объединённую вершину и множества её соседей и соседей удалённой вершины, что необходимо для последующего триадного замыкания. В оптимизированной версии с состоянием, если степени вершин различаются, вершина с большей степенью остаётся, что позволяет минимизировать количество операций обновления.

Функция add_triad_closure с вероятностью p добавляет ребро между соседями объединённых вершин, что приводит к образованию треугольника и повышению коэффициента кластеризации сети. Формируются списки кандидатов из вершин, которые являются соседями объединённых вершин. При равномерном выборе вершины выбираются случайно и независимо. В случае предпочтительного выбора вычисляются веса – текущие степени вершин, и вершина выбирается с вероятностью, пропорциональной её степени. После выбора двух вершин проверяется, что они различны, обе присутствуют в графе, и между ними ещё нет ребра. Если все условия выполнены, ребро добавляется.

Функция add_new_node добавляет новую вершину в граф и соединяет её с m существующими вершинами согласно выбранной стратегии (равномерной или предпочтительной). При предпочтительном выборе, чтобы избежать повторного выбора одной вершины, выбранная вершина временно удаляется из дерева Фенвика, а после завершения цикла все временно удалённые степени возвращаются обратно.

Основная же функция run_modified_ba_model запускает процесс генерации графа в соответствии с модифицированной моделью. На каждой итерации выполняются три шага в строгой последовательности: объединение двух случайных вершин согласно выбранной стратегии, вероятностное добавление ребра между их соседями (триадное замыкание) и добавление новой вершины с m рёбрами. В том числе на каждом из этих шагов может использоваться либо равномерная стратегия выбора вершин, либо предпочтительная.

Для анализа свойств полученных сетей реализуется функции вычисления основных метрик: распределения степеней, глобального и локального коэффициентов кластеризации, индекса дружбы, коэффициента ассортативности, а также показателя степенного закона γ. Функция calculate_friendship_index_by_degree вычисляет индекс дружбы для вершин с одинаковой степенью. Функция `calculate_power_law_exponent` вычисляет показатель степени для распределения степеней с использованием линейной регрессии в логарифмических координатах.

С целью исследования влияния различных комбинаций стратегий выбора на свойства генерируемых сетей воплощается функция, запускающая все 8 возможных комбинаций стратегий для трёх операций: объединения, триадного замыкания и роста. Каждая стратегия может быть двух типов: равномерная (выбор вершины случайным образом) или предпочтительная (выбор вершины с вероятностью, пропорциональной её степени).

Причем эксперименты проводятся при следующих параметрах: начальное число вершин $n_{\text{start}} = 100000$, количество итераций 200000, вероятность триадного замыкания $p_{\text{triad}} = 0.5$. Параметр m (число рёбер, добавляемых с каждой новой вершиной) варьируется от 2 до 5. Такой диапазон значений m позволяет проследить, как изменение этого параметра влияет на структуру сети в сочетании с различными стратегиями.

Для каждого набора параметров вычисляются основные характеристики сети: число вершин и рёбер, средняя, максимальная и минимальная степени, глобальный и локальный коэффициенты кластеризации, коэффициент ассортативности, показатель степенного закона γ, а также средний индекс дружбы.

Результаты вычислительных экспериментов

Анализ результатов экспериментов показывает, что стратегия объединения вершин является наиболее значимым фактором, определяющим структурный тип формируемой сети. Все 8 комбинаций стратегий чётко разделяются на две группы: случаи 1-4 (равномерное объединение) и случаи 5-8 (предпочтительное объединение).

Для группы с равномерным объединением характерна высокая плотность связей. При $m = 2$ число рёбер составляет примерно 696000-699000, что почти в 4 раза больше, чем во второй группе. Средняя степень держится в узком диапазоне от 13,93 до 13,99. Максимальная степень варьируется от 1006 до 2892. Это говорит об отсутствии доминирующих хабов.

У группы с предпочтительным объединением складывается иная картина. Плотность сети значительно ниже. При $m = 2$ число рёбер составляет 175000-226000. Средняя степень падает до 3,51-4,54. Сеть становится разреженной. Максимальная степень достигает 83728 при средней степени 3,66. Это означает, что отдельные вершины-хабы аккумулируют большую часть всех связей. Отношение максимальной степени к средней составляет от 12500 до 23800.

Таким образом, стратегия объединения является главным бинарным переключателем режимов сети: равномерное объединение порождает однородные плотные графы, а предпочтительное – экстремально неравномерные разреженные структуры с супер-хабами.

Эффективность триадного замыкания существенно зависит от того, на фоне какой стратегии объединения оно применяется. В группе с равномерным объединением различие между равномерной и предпочтительной стратегиями триадного замыкания невелико и во многих случаях находится в пределах погрешности эксперимента. В тоже время в группе с предпочтительным объединением влияние триадного замыкания проявляется значительно отчётливее. Предпочтительная стратегия триадного замыкания даёт прирост глобального коэффициента кластеризации примерно в 1,5-2 раза. Например, при $m = 2$ для случая ПРР (предпочтительное объединение, равномерная триада, равномерный рост) кластеризация равна 0,2165, а для ПРП (предпочтительное объединение, равномерная триада, предпочтительный рост) – 0,4518. Объяснение этого эффекта заключается в том, что предпочтительное объединение создаёт высокостепенные узлы-хабы, соседи которых сами обладают относительно высокой степенью, и добавление рёбер между такими соседями с высокой вероятностью формирует треугольники.

Локальный коэффициент кластеризации, усреднённый по всем вершинам, качественно повторяет динамику глобального. Для равномерного объединения его значения крайне малы и снижаются с ростом m. В случае предпочтительного объединения локальная кластеризация на порядок выше. Наибольшие значения достигаются при $m = 2$. Предпочтительная стратегия триадного замыкания даёт прирост локальной кластеризации примерно в 1,5-2 раза по сравнению с равномерной.

Стратегия добавления новой вершины оказывает наименьшее влияние на итоговую структуру сети по сравнению с двумя другими механизмами. В группе с равномерным объединением предпочтительный рост приводит к некоторому увеличению максимальной степени и, как следствие, к росту неравенства распределения. В группе с предпочтительным объединением влияние стратегии роста практически нивелируется, поскольку в условиях существования сверхсильных хабов новая вершина с высокой вероятностью присоединяется к одному из них независимо от способа выбора.

Параметр m закономерно увеличивает среднюю степень сети. Однако его влияние на другие метрики противоположно для двух групп стратегий.

Для равномерного объединения с ростом m коэффициент кластеризации снижается. Например, в случае, когда все три стратегии (объединение, триадное замыкание и рост) выбираются равномерно, кластеризация падает с 0,0296 при $m = 2$ до 0,0069 при $m = 5$. Ассортативность остаётся около нуля. Средний индекс дружбы растёт.

Для предпочтительного объединения с ростом m кластеризация тоже снижается, но остаётся значительно выше, чем в первой группе. Например, в случае, когда объединение и рост выбираются предпочтительно, а триадное замыкание – равномерно, кластеризация падает с 0,4518 при $m = 2$ до 0,3355 при $m = 5$. Отрицательная ассортативность ослабевает и стремится к нулю. Средний индекс дружбы, напротив, уменьшается.

Сравнение полученных результатов с классической моделью Барабаши-Альберт показывает принципиальные различия. В классической модели средняя степень растёт линейно с числом вершин, а коэффициент кластеризации стремится к нулю при увеличении размера сети. В предложенной модификации поведение сети зависит от выбранных стратегий.

При равномерном объединении сеть сохраняет высокую плотность, сравнимую с классической моделью, но при этом кластеризация остаётся крайне низкой. При предпочтительном объединении сеть становится разреженной, однако кластеризация оказывается на порядок выше, чем в модели Барабаши-Альберт. Кроме того, в классической модели распределение степеней всегда подчиняется закону с показателем $\gamma = 3$, тогда как в предложенной модели этот показатель может варьироваться от 2,5 до 3,4 в зависимости от m.

Показатель степенного закона γ во всех экспериментах находится в пределах 2,5–3,4. При $m = 3$, $m = 4$ и $m = 5$ значения γ располагаются в интервале от 2 до 3. Такой интервал характерен для большинства реальных безмасштабных сетей. Следовательно, модель с этими параметрами можно использовать для описания реальных систем. При $m = 2$ значения γ оказались больше 3, что несколько отличается от типичных показателей реальных сетей. При этом ни от m, ни от комбинации стратегий систематической зависимости не наблюдается.

Это подтверждает, что предложенная модель сохраняет свойство безмасштабности независимо от выбранных параметров.

Для оценки неравномерности распределения степеней используется отношение максимальной степени к средней $k_{\max}/\langle k\rangle$. В группе с равномерным объединением это отношение находится в диапазоне от 41 до 208 в зависимости от m. Такие значения указывают на умеренное или выраженное неравенство, но не свидетельствуют о доминировании отдельных узлов. В группе с предпочтительным объединением отношение достигает величин от 6177 до 23791, что на два-три порядка выше. Это говорит о том, что в сети формируется иерархическая структура, в которой небольшая группа узлов-хабов сосредотачивает основную долю связей.

Полученные результаты могут быть полезны для моделирования реальных систем. Например, социальные сети часто проходят этапы стабилизации, когда число пользователей перестаёт расти, но внутренняя структура продолжает меняться. Предложенная модель позволяет описывать такие процессы за счёт механизмов объединения вершин и триадного замыкания. Аналогично, в биологических популяциях численность особей может сокращаться, но взаимодействия между оставшимися особями перестраиваются. Предложенный подход даёт инструмент для анализа подобных сценариев.

По итогам дипломной работы была разработана и исследована модель случайной сети для описания структур на этапах стабилизации или сокращения, а также в полном объеме выполнены соответствующие задачи. Это, в свою очередь, подтвердило то, что классические модели (Барабаши-Альберт, триадного замыкания) описывают только рост, тогда как реальные системы могут сжиматься и реорганизовываться.

Так, в теоретической части работы была проведена систематизация основных понятий теории случайных графов, рассмотрены механизмы предпочтительного присоединения и триадного замыкания и сформулирована концепция невозрастающей сети.

На этой основе, в практической плоскости был разработан программный комплекс на Python для модифицированной модели классических растущих сетей с операциями объединения вершин, триадного замыкания и добавления узлов. Для оптимизации вычислений на больших графах реализована структура данных на основе дерева Фенвика, позволяющая выполнять операции предпочтительного выбора вершин за логарифмическое время.

В ходе вычислительных экспериментов приведены следующие выводы:

1. стратегия объединения вершин определяет тип сети. При равномерном объединении формируются плотные сети с низкой кластеризацией и умеренным неравенством степеней. При предпочтительном – разреженные сети с высокой кластеризацией, сильным неравенством и отрицательной ассортативностью;
2. триадное замыкание повышает кластеризацию только при предпочтительном объединении. При равномерном объединении его влияние на кластеризацию незначительно;
3. стратегия роста слабо влияет на итоговую структуру. Её роль заметна только при равномерном объединении. В случае предпочтительного объединения наличие доминирующих хабов делает способ присоединения новой вершины практически несущественным;
4. параметр m действует по-разному в зависимости от стратегии объединения. При равномерном объединении увеличение m приводит к росту среднего индекса дружбы и снижению кластеризации. При предпочтительном объединении с ростом m уменьшаются как кластеризация, так и индекс дружбы, а отрицательная ассортативность ослабевает;
5. безмасштабность сохраняется во всех режимах (показатель степенного закона $\gamma \approx 3$). При $m = 3$, $m = 4$ и $m = 5$ значения γ располагаются в интервале от 2 до 3, что характерно для большинства реальных безмасштабных сетей.

Таким образом, предложенная модель позволяет управлять свойствами сети через выбор стратегий и параметра m. Приведенные результаты могут быть применены для моделирования реальных сетей в фазах стабилизации или сокращения, где баланс между ростом и процессами внутренней реорганизации определяет итоговую структуру.