

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 ПОВЕДЕНИЯ УБЫВАЮЩИХ СЕТЕЙ В МОДЕЛИ
ТРИАДНОГО ЗАМЫКАНИЯ**

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студента 4 курса 411 группы
направления 02.03.02 — Фундаментальная информатика и информационные
технологии
факультета КНиИТ
Безлюдного Владимира Ильича

Научный руководитель

зав. каф., к. ф.-м. н., доцент

С. В. Миронов

Заведующий кафедрой

к. ф.-м. н., доцент

С. В. Миронов

Саратов 2023

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	3
1 Модели случайных графов.....	5
1.1 Модели растущих сетей.....	5
1.2 Убывающие сети	7
2 Экспериментальные результаты.....	10
ЗАКЛЮЧЕНИЕ	16

ВВЕДЕНИЕ

Многие технологические, биологические и социальные системы могут быть представлены лежащими в их основе сложными сетями. Такие сети состоят из множества узлов, и если взаимодействие между парой элементов в системе имеет место, то предполагается, что соответствующая пара узлов имеет связь. Недавние исследования в области теории графов и сетевых технологий показывают, что различные сложные сетевые структуры, такие как социальные сети, биологические сети и сети связи, могут быть представлены в виде графов. Графы являются мощным инструментом исследований, который помогает исследователям понять и анализировать структуру и связи между различными элементами сетей.

Экономические системы являются примером сложных систем, где элементами являются фирмы или компании, а их взаимодействия отражаются в экономических, информационных и финансовых связях. Обычно компании, появившиеся на ранних стадиях развития сети, получают серьезное преимущество перед компаниями, появившимися позже, из-за эффекта преимущества первопроходца. Однако существуют примеры успешных технологических компаний, где элементы, появившиеся позже, занимают более доминирующее положение в системе, чем элементы, появившиеся на ранних стадиях развития. Этот эффект также может быть наблюдаемым при диффузии инноваций, производстве, распространении патентов и технологий, информационном взаимодействии и социальных сетях.

Основой для анализа исходных данных в теории графов служат случайные графы, которые имеют широкий спектр применений в различных областях, включая инженерию систем, социологию, экономику, биологию и другие. В данной работе мы будем рассматривать методы и алгоритмы генерации случайных графов, а также, в частности, рассмотрим убывающие сети и их характеристики.

Целью данной работы является исследование поведения убывающих сетей, построенных в соответствии с моделью триадного замыкания.

Для достижения этой цели необходимо решить следующие задачи:

- изучить понятия теории графов и её приложение к теории случайных графов;
- для программной реализации изучить возможности библиотеки `networkx` языка Python;

- рассмотреть и реализовать модель триадного замыкания построения случайного графа;
- рассмотреть и реализовать модели убывающей сети, построенной по модели триадного замыкания;
- провести серию экспериментов по построению убывающих сетей;
- исследовать поведение локальных характеристик узлов убывающей сети;
- проанализировать результаты эксперимента.

1 Модели случайных графов

1.1 Модели растущих сетей

Объектом исследования в данной работе является случайный граф. Случайный граф — общий термин для обозначения вероятностного распределения графов. Случайные графы можно описать просто распределением вероятности или случайным процессом, создающим эти графы.

Граф представляет из себя пару множеств

$$G = (V, E),$$

где $V = \{v_0, v_1, \dots, v_N\}$ — множество вершин (узлов) графа, $E \subseteq V \times V$ — множество ребер графа.

В работе будем рассматривать сети, изменяющиеся в дискретном времени $t = 0, 1, \dots$.

Состояние растущей сети в момент времени t будем описывать графом

$$G(t) = (V(t), E(t)).$$

В модели растущей сети будем в каждый момент времени t добавлять в граф одну вершину v_t . Кроме того, будем считать, что на момент времени $t + 1$ в графе уже присутствуют вершины v_0, v_1, \dots, v_t .

У вершины есть несколько атрибутов (локальных характеристик), которые понадобятся в ходе работы:

- $d_i(t)$ — степень вершины с номером i в момент времени t ;
- $s_i(t) = \sum_{j:(v_i, v_j) \in E(t)} d_j(t)$ — сумма степеней соседей вершины v_i в момент времени t ;
- $\alpha_i(t) = \frac{s_i(t)}{d_i(t)}$ — средняя степень соседей вершины v_i в момент времени t .

Все эти определения необходимы для сбора информации о локальных характеристиках графе:

- количество вершин имеющих степень d ;
- среднее значение суммы степеней соседей для вершин имеющих степень d ;
- стандартное отклонение суммы степеней соседей для вершины имеющих степень d — показывает, как распределены значения относительно сред-

него в нашей выборке. Другими словами, можно понять на сколько велик разброс величины;

- коэффициент вариации суммы степеней соседей для вершины имеющих степень d — показывает степень изменчивости по отношению к среднему показателю выборки.

Многие сложные сети, такие как социальные, экономические, сети цитирования или Интернет, с течением времени развиваются, путем добавления новых узлов, которые соединяются с уже существующими. Результаты исследований реальных сетей показывают, что распределение степеней в таких сложных сетях подчиняется степенному закону. То есть, для конкретной фиксированной сети количество вершин заданной степени d можно предсказать как

$$M(d) = Kd^{-\gamma}, \quad (1)$$

где γ и K — некоторые фиксированные вещественные значения.

Исследования реальных сетей показывают, что в большинстве из них распределение степеней подчиняется степенному закону с параметром γ из диапазона от 2 до 3.

В каждый момент времени в сеть добавляется новый узел. Каждый новый узел соединяется с m существующими узлами. Выбор узла для установления связи между ним и новой вершиной происходит с вероятностью, пропорциональной его степени. Формально, вероятность p , что новый узел v_t соединится с узлом v_i равна:

$$p = \frac{d_i(t)}{\sum_{j=0}^{t-1} d_j(t)}.$$

В настоящей работе мы рассмотрим модель триадного замыкания.

Триадное замыкание — понятие в теории социальных сетей, которое первым предложил немецкий социолог Г. Зиммель. Это свойство трех узлов А, В и С, например, представляющих людей, и заключающееся в том, что если существуют соединения АВ и ВС, то существует предрасположенность к формированию нового соединения АС. Триадное замыкание можно использовать для понимания и прогнозирования роста сетей, хотя это только один из многих механизмов, с помощью которых формируются новые связи в сложных сетях.

Модель триадного замыкания генерирует растущие сети следующим об-

разом.

На начальном шаге берется граф из m вершин, где m — первый параметр модели.

На каждом шаге $t + 1$:

1. добавляется новая вершина v_{t+1} к уже существующим;
2. проводится m рёбер, соединяющих v_{t+1} и уже существующие m вершин следующим образом:
 - первое ребро на шаге соединяет новую вершину с вершиной v_i с вероятностью пропорциональной её степени $d_i(t + 1)$ (по принципу предпочтительного присоединения);
 - оставшиеся $m - 1$ рёбер проводятся от вершины v_{t+1} по следующему принципу:
 - с вероятностью p , формируется триада, что означает, что ребро проводится к произвольному соседу вершины v_i (к которой было проведено первое ребро). Выбор соседа вершины v_i производится равновероятно. Здесь и далее p — второй параметр модели;
 - с вероятностью $1 - p$ ребро проводится к одной из вершин графа по принципу предпочтительного присоединения.

Главная особенность триадного замыкания заключается в том, что оно создаёт большое количество треугольников в процессе роста сети, а эволюция степени каждой вершины следует степенному закону:

$$E(d_i(t)) = m \left(\frac{t}{i} \right)^\beta$$

с одной и той же экспонентой $\beta = \frac{1}{2}$ для всех вершин.

Модель может создавать сети с различными уровнями кластеризации, в зависимости от выбора параметров модели p и m . С другой стороны, их распределения степеней подчиняются степенному закону с показателем $\gamma = 3$ для любого p , т.е. с тем же показателем, что и в модели Барабаши—Альберт.

1.2 Убывающие сети

На сегодняшний день убывающие сети практически не изучались. Но графы, соответствующие убывающим сетям, представляют собой реальные взаи-

моотношения в реальной жизни. Например, каждый человек может прекратить дружбу с другим, компании могут перестать сотрудничать друг с другом, могут обанкротиться и закрыться. Все эти свойства характерны для больших социальных сетей и в данной работе ставится задача рассмотреть характеристики таких сетей.

Убывающий граф — это граф, в котором наряду с ростом, присутствует механизм исчезновения рёбер и/или узлов.

Мы рассмотрим модели убывающих сетей, которые получаются из модели растущей сети, где в каждый момент времени кроме добавления новой вершины и проведения новых ребер внедрена, в соответствии с некоторыми принципами, последовательность действий, направленная на «убывание» графа.

Будем использовать следующие механизмы убывания сети.

1. Модель удаления Пакт. В этой модели два узла сливаются в один. При этом вершины, смежные с исходными, остаются смежными с образовавшейся вершиной. Рассмотрим два способа выбора пар вершин для слияния:
 - обе вершины выбираются из всего множества вершин равновероятно;
 - одна вершина выбирается из всех вершин с равной вероятностью, а другая из множества соседей первой с вероятностью пропорциональной степени каждого соседа.
2. Удаление случайной вершины. Выбирается случайная вершина и удаляется из графа вместе с инцидентными ей ребрами.
3. Удаление в зависимости от количества соседей. Случайно равновероятно выбирается вершина графа и удаляется если у неё менее r соседей.
4. Удаление в зависимости от количества удаленных соседей.
 - для каждой вершины хранится количество удалённых соседей;
 - вершина для удаления выбирается случайно с вероятностью пропорциональной количеству удаленных у неё соседей.

В данной работе модель убывания будет внедряться в модель триадного замыкания растущей сети. Эволюцию сети разобьем на два этапа.

На первом этапе будем реализовывать чистый рост сети по модели триадного замыкания.

На втором этапе, перед началом итерации в момент времени t по алгоритму триадного замыкания реализуем «убывание» по одному из методов, перечисленных выше. То есть эволюция сети будет происходить следующим

образом.

На начальном шаге берется граф из m вершин, где m — первый параметр модели.

На каждом шаге t :

1. если $t > N_1$, где N_1 — параметр модели, то выполняется шаг убывания (возможное удаление одной из имеющихся в графе вершин);
2. добавляется новая вершина v_t к уже существующим;
3. проводится m рёбер, соединяющих v_t и уже существующие m вершин следующим образом:
 - первое ребро на шаге соединяет новую вершину с вершиной v_i с вероятностью пропорциональной её степени $d_i(t)$ (по принципу предпочтительного присоединения);
 - оставшиеся $m - 1$ рёбер проводятся от вершины v_t по следующему принципу:
 - с вероятностью p , формируется триада, что означает, что ребро проводится к произвольному соседу вершины v_i (к которой было проведено первое ребро). Выбор соседа вершины v_i производится равновероятно. Здесь и далее p — второй параметр модели;
 - с вероятностью $1 - p$ ребро проводится к одной из вершин графа по принципу предпочтительного присоединения.

2 Экспериментальные результаты

В ходе экспериментов строился один граф по каждой модели с заданными параметрами. В качестве модели роста сети использовалась модель триадного замыкания с параметрами:

- $m = 5$, то есть на каждом шаге добавления вершины, она соединялась с пятью другими вершинами;
- $p = 0.5$, вероятность, что ребро проводится к произвольному соседу вершины (к которой было проведено первое ребро).

В качестве начального графа брался полный граф из пяти вершин. В результате роста получался граф, представляющий из себя сеть с 5005 узлов.

В ходе первого эксперимента к полученной сети из 5005 узлов применялись 5000 итераций метода триадного замыкания, в котором в начале каждой итерации выполнялся шаг модели убывания Пакт, в которой вершины для объединения выбирались случайно, равновероятно.

График распределения степеней получившегося графа представлен на рисунке 1.

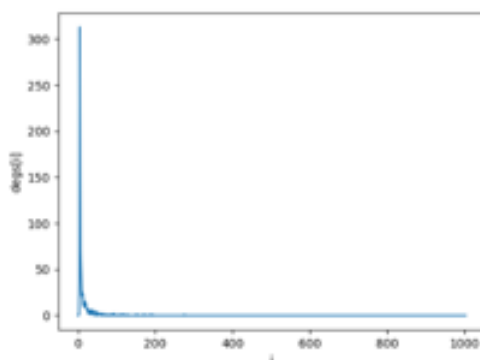


Рисунок 1 – График распределения степеней в модели Пакт, первый вариант

Этот же график, но в логарифмической шкале представлен на рисунке 2.

Из рисунка 2 можно сделать предположение, что значения на графике сосредоточены вокруг прямой

$$y = -3x + C,$$

а следовательно, можно выдвинуть гипотезу, что имеет место распределение значений по степенному закону (1) со значением $\gamma = -3$ и получившийся граф

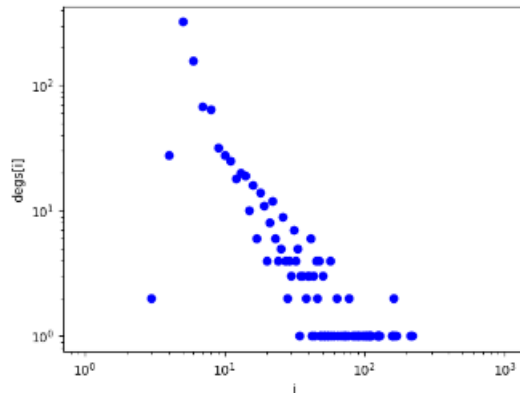


Рисунок 2 – График распределения степеней в логарифмической шкале модели Пакт, первый вариант

по этому свойству схож с реальными сетями.

Во втором эксперименте также первоначально был получен случайный граф из 5005 вершин по модели триадного замыкания с параметрами $m = 5$ и $p = 0.5$. К этому графу затем 5000 раз применялись шаги метода триадного замыкания, где перед добавлением каждой новой вершины применялся шаг объединения вершин согласно модели Пакт, в которой первая вершина для объединения выбирается случайно, а вторая выбирается из соседей первой с предпочтением выбора вершины с наибольшей степенью.

График распределения степеней получившегося графа представлен на рисунке 3.

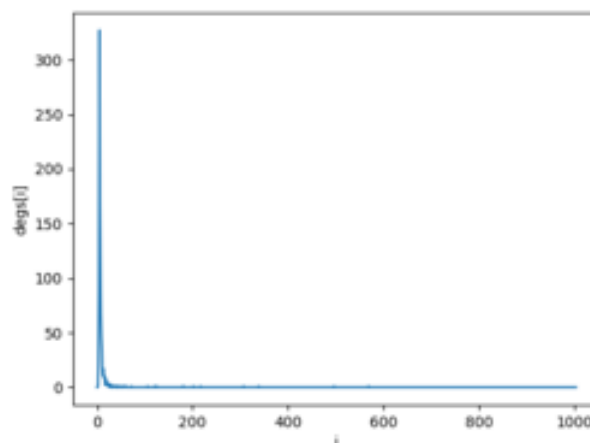


Рисунок 3 – График распределения степеней в модели Пакт, второй вариант

График распределения степеней этого же графа, но в логарифмической

шкале, представлен на рисунке 4.

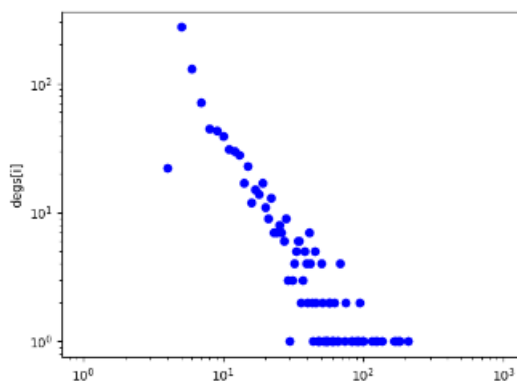


Рисунок 4 – График распределения степеней в логарифмической шкале модели Пакт, второй вариант

Так же, как и в первом эксперименте, в последнем случае можно сделать предположение, что распределение степеней в итоговом графе подчинено степенному закону (1) со значением $\gamma = 3$.

Третий эксперимент начался с построения случайного графа по модели триадного замыкания, которое стартовало с полного графа из пяти вершин и продолжилось 5000 шагами метода триадного замыкания. В методе использовали параметры $m = 5$, $p = 0.5$. Результирующий граф из 5005 вершин был передан на вход модели убывающей сети, полученной из модели триадного замыкания и модели удаления случайной вершины графа.

В конечном итоге был получен граф, график распределения степеней которого представлен на рисунке 5.

График распределения степеней этого же графа в логарифмической шкале представлен на рисунке 6.

Рассматривая график в логарифмической шкале можно сделать предположение, что как и в предыдущих двух экспериментах точки графика распределены, главным образом, вдоль прямой

$$y = -\gamma x + C.$$

Но в отличие от предыдущих случаев здесь значение γ больше, чем 3, что говорит нам о том, что такая модель убывающей сети скорее плохо подходит для моделирования реальной сети.

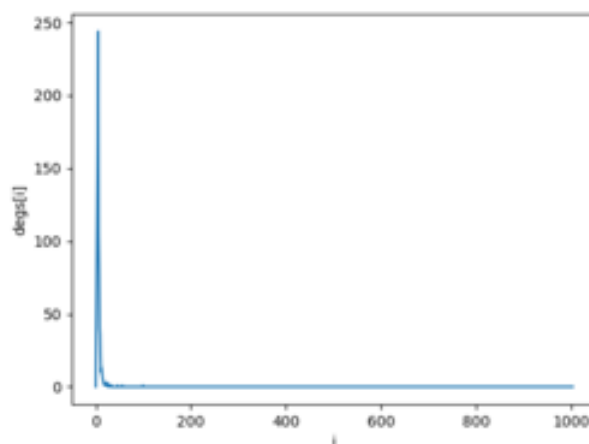


Рисунок 5 – График распределения степеней, удаление случайной вершины

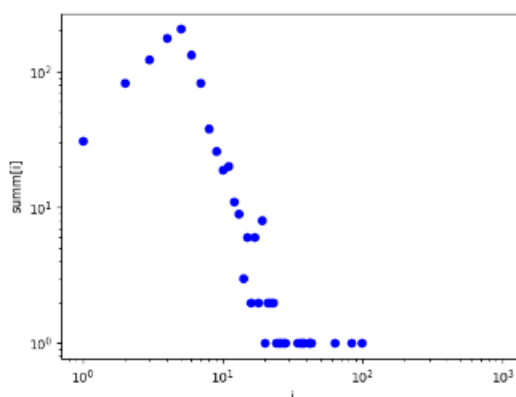


Рисунок 6 – График распределения степеней в логарифмической шкале, удаление случайной вершины

В четвертом эксперименте сначала был получен случайный граф из 5005 вершин по модели триадного замыкания с параметрами $m = 5$ и $p = 0.5$. К этому графу затем 5000 раз применялись шаги метода триадного замыкания, где перед добавлением каждой новой вершины применялся шаг удаления случайной вершины графа, при условии, что количество её соседей меньше чем параметр r .

График распределения степеней получившегося графа представлен на рисунке 7.

График распределения степеней получившегося графа в логарифмической шкале представлен на рисунке 8.

Согласно получившимся графикам можно выдвинуть гипотезу, что как и в предыдущих экспериментах имеет место степенной закон распределения,

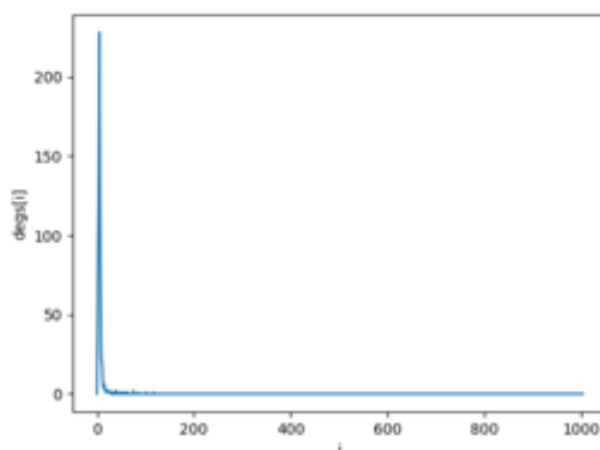


Рисунок 7 – График распределения степеней, удаление в зависимости от соседей вершины

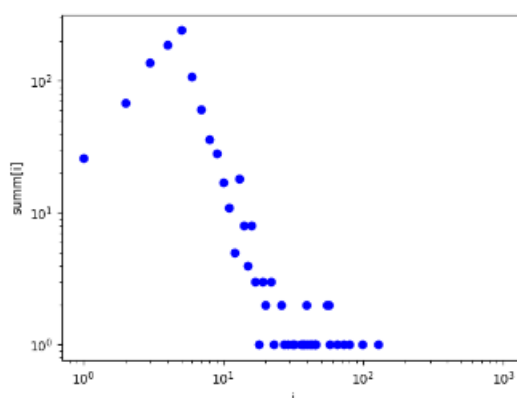


Рисунок 8 – График распределения степеней в логарифмической шкале, удаление в зависимости от соседей вершины

но как и в предыдущем случае значение показателя степени γ больше, чем 3, а следовательно, что такая модель убывающей сети тоже плохо подходит для моделирования реальной сети.

В пятом эксперименте сначала проводилось построение случайного графа по модели триадного замыкания, которое стартовало с полного графа из пяти вершин и продолжилось 5000 шагами метода триадного замыкания. В работе метода использовали параметры $m = 5$, $p = 0.5$. Результирующий граф из 5005 вершин был передан на вход модели убывающей сети, полученной из модели триадного замыкания и модели удаления вершины с вероятностью, пропорциональной количеству её удаленных соседей.

График распределения степеней получившегося графа представлен на рисунке 9.

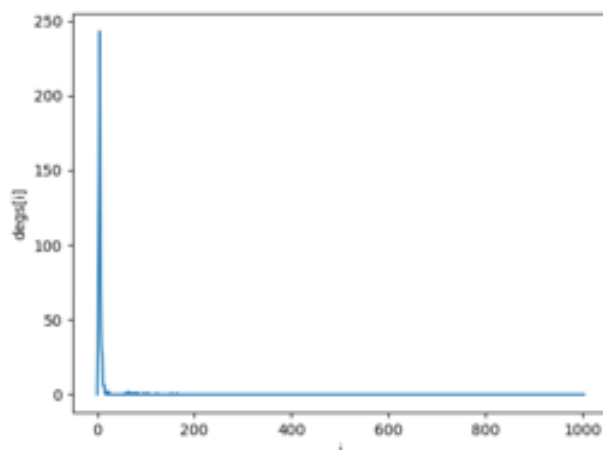


Рисунок 9 – График распределения степеней, удаление в зависимости от количества удаленных соседей

График распределения степеней получившегося графа в логарифмической шкале представлен на рисунке 10.

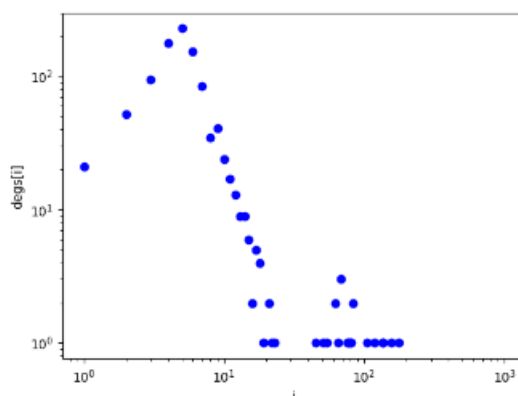


Рисунок 10 – График распределения степеней в логарифмической шкале, удаление в зависимости от количества удаленных соседей

Так же, как и в предыдущих случаях, графики наводят нас на гипотезу, что распределение степеней подчиняется степенному закону, но значение показателя степени γ больше, чем 3, что плохо соответствует реальным сетям.

ЗАКЛЮЧЕНИЕ

В ходе дипломной работы были изучены понятия теории графов и её приложение к теории случайных графов. Рассмотрены модели построения случайных графов и была реализована модель триадного замыкания растущей сети.

Для основного исследования были рассмотрены пять моделей «убывания» сети и реализованы наборы процедур для их работы. Была проведена серия экспериментов, в ходе которых строились растущие сети по модели триадного замыкания, которые подвергались обработке с помощью модели роста сети с внедренным возможным убыванием.

Исследовалось соответствие сетей, полученных в ходе экспериментов, реальным сетям в части распределения степеней вершин итогового графа. Кроме того, получены численные результаты для характеристик (среднее значение, стандартное отклонение, коэффициент вариации) для средней суммы степеней соседей полученных в ходе экспериментов графов.

Таким образом, все задачи, поставленные в задании на выпускную квалификационную работу, выполнены. Цель работы достигнута.