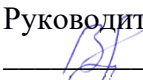





МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
Федеральное государственное автономное образовательное учреждение высшего образования  
**«Дальневосточный федеральный университет»**  
(ДФУ)

**ИНСТИТУТ МАТЕМАТИКИ И КОМПЬЮТЕРНЫХ ТЕХНОЛОГИЙ (ШКОЛА)**

СОГЛАСОВАНО  
Руководитель ОП  
  
\_\_\_\_\_ Величко А.С.  
(подпись) (ФИО)

УТВЕРЖДАЮ  
И.о. директора департамента  
  
\_\_\_\_\_ Заболотский В.С.  
(подпись) (ФИО)



«\_13\_» сентября 2021 г.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**  
Математические модели социальных сетей  
**Направление подготовки 01.04.04 Прикладная математика**  
(Аналитические, социальные и экономические сети)  
Форма подготовки: очная

курс 1 семестр 2  
лекции 18 час.  
практические занятия 0 час.  
лабораторные работы 36 час.  
в том числе с использованием МАО лек. 0 час. / пр. 0 час. / лаб. 36 час.  
всего часов аудиторной нагрузки 54 час.  
в том числе с использованием МАО 36 час.  
самостоятельная работа 54 час.  
в том числе на подготовку к экзамену 0 час.  
контрольные работы (количество) 3  
курсовой проект не предусмотрен  
зачет 2 семестр  
экзамен не предусмотрен

Рабочая программа составлена в соответствии с требованиями Федерального государственного образовательного стандарта по направлению подготовки 01.04.04 Прикладная математика, утвержденного приказом Министерства образования и науки Российской Федерации от 10 января 2018 г. № 15 (с изменениями и дополнениями)

Рабочая программа обсуждена на заседании департамента математики, протокол № 1 от 13 сентября 2021 г.

И.о. директора департамента математики Заболотский В.С.

Составитель: профессор, канд. техн. наук, доцент А.Л. Абрамов

Владивосток  
2021

**Оборотная сторона титульного листа РПД**

**I. Рабочая программа пересмотрена на заседании департамента:**

Протокол от « \_\_\_\_\_ » \_\_\_\_\_ 20\_\_ г. № \_\_\_\_\_

Директор департамента \_\_\_\_\_  
(подпись) (И.О. Фамилия)

**II. Рабочая программа пересмотрена на заседании департамента:**

Протокол от « \_\_\_\_\_ » \_\_\_\_\_ 20\_\_ г. № \_\_\_\_\_

Директор департамента \_\_\_\_\_  
(подпись) (И.О. Фамилия)

**III. Рабочая программа пересмотрена на заседании департамента:**

Протокол от « \_\_\_\_\_ » \_\_\_\_\_ 20\_\_ г. № \_\_\_\_\_

Директор департамента \_\_\_\_\_  
(подпись) (И.О. Фамилия)

**IV. Рабочая программа пересмотрена на заседании департамента:**

Протокол от « \_\_\_\_\_ » \_\_\_\_\_ 20\_\_ г. № \_\_\_\_\_

Директор департамента \_\_\_\_\_  
(подпись) (И.О. Фамилия)

## АННОТАЦИЯ

Дисциплина «Математические модели социальных сетей» предназначена для студентов направления подготовки 01.04.04 «Прикладная математика», магистерская программа «Аналитические, социальные и экономические сети».

Общая трудоемкость освоения дисциплины составляет 3 зачетные единицы (108 часов). Дисциплина реализуется на 1 курсе в 2-м семестре. Дисциплина входит в обязательные дисциплины части, формируемой участниками образовательных отношений, блока «Дисциплины (модули)».

Особенности построения курса: лекции (18 часов), лабораторные работы (36 часов), самостоятельная работа (54 часа).

Содержание дисциплины охватывает следующий круг вопросов: стратегии формирования сетей и модели распространения потоков, волн, объектов в экономических сетях; равновесные и экстремальные задачи на сетях и графах.

**Цель** - развитие способности и готовности использовать стратегии формирования сетей и модели распространения потоков, волн, объектов в экономических сетях, формулировать равновесные и экстремальные задачи на сетях и графах, обнаруживать соответствующие явления в экономических сетях, обосновывать адекватность используемых моделей.

### **Задачи:**

- совершенствовать и развивать свой интеллектуальный и общекультурный уровень;
- демонстрировать и использовать фундаментальные и прикладные знания передового рубежа науки;
- способностью организовывать работу исполнителей, находить и принимать управленческие решения в области организации труда;
- готовностью к принятию ответственности за свои решения в рамках профессиональной компетенции, способен принимать нестандартные решения, разрешать проблемные ситуации;

- способностью и готовностью проводить научные эксперименты, оценивать результаты исследований.

Для успешного изучения дисциплины «Математические модели социальных сетей» у обучающихся желательны следующие предварительные компетенции:

- способность использовать основы экономических знаний в различных сферах деятельности;
- способность определять экономическую целесообразность принимаемых технических и организационных решений;
- способностью выявить естественнонаучную сущность проблем, возникающих в ходе профессиональной деятельности, готовностью использовать для их решения соответствующий естественнонаучный аппарат.

В результате изучения данной дисциплины у обучающихся формируются следующие универсальные, общепрофессиональные, профессиональные компетенции (элементы компетенций).

Универсальные компетенции выпускников и индикаторы их достижения:

Наименование категории (группы) универсальных компетенций	Код и наименование универсальной компетенции выпускника	Код и наименование индикатора достижения компетенции	Результаты обучения по дисциплинам (модулям), практикам
Межкультурное взаимодействие	УК-5 Способен анализировать и учитывать разнообразие культур в процессе межкультурного взаимодействия	УК-5.3 Анализирует и оценивает явления в социальных сетях	<i>Знает</i> современные методы исследований в области анализа социальных сетей <i>Умеет</i> самостоятельно обучаться новым методам исследования <i>Владеет</i> приемами выбора методов, наиболее подходящих к выбранной области исследования

**Профессиональные компетенции выпускников и индикаторы их достижения:**

<b>Код и наименование профессиональной компетенции</b>	<b>Код ПК (при наличии ПК) или ссылка на иные основания</b>	<b>Код трудовой функции</b>	<b>Индикаторы достижения компетенции</b>
<b>Тип задач профессиональной деятельности: научно-исследовательский</b>			
ПК-5 Способен к разработке и исследованию математических методов и моделей для проведения многовариантных аналитических расчетов и подготовки принятия решений	Анализ требований, предъявляемых к выпускникам	-	<p>ПК-5.1 Формулирует модели, применяет методы анализа объектов, систем, процессов и технологий на основе математических моделей и методов прикладной математики</p> <p>ПК 5.2 Проводит сценарные аналитические расчеты для обоснования принимаемых решений по вариантам в том числе на основе программных средств</p>

## **I. СТРУКТУРА И СОДЕРЖАНИЕ ТЕОРЕТИЧЕСКОЙ ЧАСТИ КУРСА**

### **Раздел I. Основы моделирования социальных сетей**

#### **Тема 1. Основные понятия и определения**

История возникновения (работы Jacob Moreno, Anatol Rapoport, William Horvath). Первые графовые модели. Работы Stanley Milgram – эффект «маленького мира». Введение в теорию шести рукопожатий. Работы Barabási Albert-László введение в теорию «Управляемость сложных сетей» («Controllability of Complex Networks»). Результаты эксперимента Mark Granovetter. Предположение о важности слабых связей (weak ties). Применение в торговле товарами и услугами, транспортными, энергетическими, городских региональных, международных сетях. Сети ОЭЗ и свободных портов. Результаты Alfred Lotka, закон Лотки (сети цитирования). Всемирная паутина (World Wide Web) - циклическая сеть. Результаты Steven Strogatz и Duncan Watts - феномен тесного мира. Работы Reka Albert и Laszlo Barabasi - распределение вершин по числу связей. Сети предпочтений (Preference networks) - двусторонние информационные сети. Радиус, эксцентриситет, геодезическая цепь. Диаметр. Диаметр и деревья. Диаметры в случайных графах. Диаметры в мире. Теорема о структуре сети. Распределение степеней. Кластеризация. Модель гомофилии. Динамика и сила слабых связей.

Центральность. Возможности измерения центральности: степень – связность, близость и простота достижения других вершин. Маршруты роль промежуточных вершин и ребер. Влияние. Престиж. Центральность в сети собственные вектора. Применение мер центральности (Centrality). Диффузия центральности. Случайные сети. Случайные сети: пороги и фазовые переходы. Теорема Threshold. Модель «маленького мира». Другие статические модели сетей: модели для генерации кластеров, модели для получения другого распределения степеней, отличного от распределения Пуассона, модель подгонки данных.

## **Тема 2. Модель «маленького мира»**

Эксперимент Stanley Milgram. Теория шести рукопожатий - модель «маленького мира» (small world). Модель Duncan Watts и Steve Strogatz с высокой степенью кластеризации и малой средней длиной пути между вершинами. Свойства социальной сети, как гомофилия (homophily) и слабые связи (weak ties).

## **Раздел II. Модели случайных сетей**

### **Тема 3. Случайные сети**

Рост случайных сетей. Аппроксимация. Гибридные модели. Формирование гибридных моделей. Блочные модели. Случайные сетевые модели: Эрдеша (Erdos) – Реньи (Renyi). Другие модели случайных сетей: Watts and Strogatz, Barabasi and Albert, Jackson and Rogers. Стохастические блочные модели: модели дополнения Эрдеша (Erdos) – Реньи (Renyi). Набор моделей: ERGMs и новые: SERGMs / SUGMs.

### **Тема 4. Стратегия формирования сети**

Стратегия формирования сети. Равновесие и эффективность. Модель соединения сети. Эффективность модели соединения: попарное равновесие и модель соединений. Внешние эффекты: формирование сети и трансферы. Неоднородность в стратегии формирования сети. Модель SUGMs и стратегия формирования сети. Равновесие по Нэшу. Динамические стратегии формирования сети. Эволюция и стохастика. Режиссура формирования сети.

Применение структурной модели формирования стратегии.

### **Тема 5. Диффузия и обучение в сетях. Игры на сетях**

Диффузия. Bass-модель диффузии. Диффузия на случайных сетях. Главная компонента (Пуассона). SIS – модель. Решения SIS -модели – примеры. Подготовка данных для модели диффузии. Пример распространения эпидемии. Обучение. Модель ДеГрута (DeGroot). Конвергенция в модели ДеГрута (DeGroot). Дополнения и заменители. Свойства равновесий. Несколько равновесий. Применения. Дискретный (бинарный) выбор. Линейные и квадратичные модели. Многошаговые игры на сетях.

## **II. СТРУКТУРА И СОДЕРЖАНИЕ ПРАКТИЧЕСКОЙ ЧАСТИ КУРСА**

### **Лабораторные работы**

**Лабораторная работа №1. Математические модели социальных сетей**

**Лабораторная работа № 2. Модель «маленького мира» для экономических сетей**

**Лабораторная работа № 3. Случайные сети для экономических сетей**

**Лабораторная работа № 4. Стратегии формирования сетей для экономических сетей**

**Лабораторная работа № 5. Диффузия и обучение в сетях. Игры на сетях. Применение для экономических сетей**

## **III. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ**

Учебно-методическое обеспечение самостоятельной работы обучающихся по дисциплине «Математические модели социальных сетей» представлено в Приложении 1 и включает в себя:

план-график выполнения самостоятельной работы по дисциплине, в том

числе примерные нормы времени на выполнение по каждому заданию;

характеристика заданий для самостоятельной работы обучающихся и методические рекомендации по их выполнению;

требования к представлению и оформлению результатов самостоятельной работы;

критерии оценки выполнения самостоятельной работы.

#### **IV. КОНТРОЛЬ ДОСТИЖЕНИЯ ЦЕЛЕЙ КУРСА**

Контролируемые разделы дисциплины, этапы формирования компетенций, виды оценочных средств, зачетно-экзаменационные материалы, комплекты оценочных средств для текущей аттестации, описание показателей и критериев оценивания компетенций на различных этапах их формирования, методические материалы, определяющие процедуры оценивания знаний, умений и навыков и (или) опыта деятельности, а также критерии и показатели, необходимые для оценки знаний, умений, навыков и характеризующие этапы формирования компетенций в процессе освоения образовательной программы, представлены в Приложении 2.

#### **V. СПИСОК УЧЕБНОЙ ЛИТЕРАТУРЫ И ИНФОРМАЦИОННО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ**

##### **Основная литература**

1. Проблемы сетевого управления : монография / Н. О. Амелина, М. С. Ананьевский, Б. Р. Андриевский [и др.] ; под редакцией А. Л. Фрадков. — Ижевск : Регулярная и хаотическая динамика, Институт компьютерных исследований, 2015. — 392 с. — ISBN 978-5-4344-0344-3. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/69364.html>.

2. Pizzuti, C.; Socievole, A. Computation in Complex Networks. *Entropy* 2021, 23, 192. <https://doi.org/10.3390/e23020192>.



3. Kuikka, V. Modelling community structure and temporal spreading on complex networks. *Comput Soc Netw* 8, 13 (2021). <https://doi.org/10.1186/s40649-021-00094-z>.

### **Дополнительная литература**

1. Yang, H.; An, S. Critical Nodes Identification in Complex Networks. *Symmetry* 2020, 12, 123. <https://doi.org/10.3390/sym12010123>.

2. C. Li and S. Pang, "The Role of Individual Edges in Edge Controllability of Complex Networks," in *IEEE Access*, vol. 8, pp. 63559-63566, 2020, doi: <https://doi.org/10.1109/ACCESS.2020.2984358>.

### **Перечень ресурсов информационно-телекоммуникационной сети «Интернет»**

1. Электронно-библиотечная система - <http://e.lanbook.com/>;
2. Студенческая электронная библиотека - <http://www.studentlibrary.ru/>;
3. Электронно-библиотечная система - <http://znanium.com/>;
4. Электронная библиотека - <http://www.nelbook.ru/>;
5. База данных Scopus <http://www.scopus.com/home.url>;
6. База данных Web of Science <http://apps.webofknowledge.com/>;
7. База данных полнотекстовых академических журналов Китая <http://oversea.cnki.net/>;
8. Электронная библиотека диссертаций Российской государственной библиотеки <http://diss.rsl.ru/>;
9. Электронные базы данных EBSCO <http://search.ebscohost.com/>;

### **Перечень дополнительных информационно-методических материалов**

1. Newman M. E. J. *Networks: An Introduction* [Книга]. - University of Michigan and Santa Fe Institute: Oxford Press, 2010. – 790 с.

2. Jackson M. O. Social and economic networks [Книга]. - Princeton : Princeton University Press, 2008. – 648 с.
3. D. J. Watts and S. H. Strogatz. Collective dynamics of ‘small-world’ network. Nature, 393(6684):
4. 440-442(1998)Albert-László Barabási & Réka Albert (October 1999). «Emergence of scaling in random networks». Science 286 (5439): 509–512. DOI:10.1126/science.286.5439.509.
5. Ozik, J., Hunt, B. R. & Ott, E. *Growing networks with geographical attachment preference: Emergence of small worlds*. Phys. Rev. E **69**, 026108 (2004).
6. Aoyuan Peng, Lianming Zhang. Deterministic multidimensional growth model for small-world networks. CoRR abs/1108.5450 (2011)
7. Yilun Shang. Geometric Assortative Growth Model for Small-World Networks // <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3919055/>
8. Barrat A. Barthelemy M., Pastor-Satorras R., Vespignany A. The architecture of complex weighted networks [Журнал] // Proc. Natl. Acad. Sci. USA. - 2004 г.. - Т. 101. - стр. 3747-3752.
9. Barrat A. Weigt M. On the properties of small-world network model [Журнал] // The European Physical Journal B. - [б.м.] : EDP Sciences, 2000 г.. - 13. - стр. 547-560.
10. Bolanos M. Bernat E. M., He B., Aiyente S. A weighted small world network measure for assessing functional connectivity [Журнал] // Journal of Neuroscience Methods. - 2013 г.. - 212. - стр. 133-142.
11. Gleiser P. M. Danon L. Community structure in Jazz [Журнал] // Advanced in Complex Systems. - 2003 г.. - 4 : Т. 6. - стр. 565-573.
12. Granovetter M. The Strength of Weak Ties. [Журнал] // The American Journal of Sociology. - 1973 г.. - стр. 1360-1380.
13. Neman M. E. J. Models of the Small World: A Review [Журнал]. - [б.м.] : ARXIV, 2000 г..
14. Neman M. E. J. The structure and function of complex networks [Журнал]. - 2003 г..

15. Watts D. J. Small worlds: the dynamics of networks between order and randomness [Книга]. - Princeton : Princeton University Press, 1999.
16. Watts D. J. Strogatz S. H. Collective dynamics of 'small-world' networks [Журнал] // Nature. - 1998 г.. - 393. - стр. 440-442.
17. Wenyuan L. Yongjing L., Ying L. The structure of weighted small-world networks [Журнал] // Physica A. - 2006 г.. - 376. - стр. 708-718.
18. Райгородский А. М. Модели случайных графов и их применения [Журнал] // Труды МФТИ. - 2010 г.. - 4 : Т. 2. - стр. 130-140.
16. Onnela J. P. et al. Intensity and coherence of motifs in weighted complex networks // Physical Review E. – 2005. – Т. 71. – №. 6. – С. 065103.
17. BATAGELJ V., MRVAR A. Pajek Manual: Program for Analysis and Visualisation of Large Networks. – 2006.
18. De Nooy W., Mrvar A., Batagelj V. Exploratory social network analysis with Pajek. – Cambridge University Press, 2011. – Т. 27.
19. Kolaczyk E. D., Csárdi G. Statistical analysis of network data with R. – New York, NY : Springer, 2014. – С. 1-5.
20. Jeong H., Néda Z., Balabási A. L. Measuring pre-ferential attachment for evolving network, 2001 //arXiv preprint cond-mat/0104131.

## **VI. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ**

### **Рекомендации по планированию и организации времени, отведенного на изучение дисциплины, описание последовательности действий обучающихся**

Освоение дисциплины следует начинать с изучения рабочей учебной программы, которая содержит основные требования к знаниям, умениям и навыкам. Обязательно следует учитывать рекомендации преподавателя, данные в ходе установочных занятий. Затем – приступать к изучению отдельных разделов и тем в порядке, предусмотренном программой.

Получив представление об основном содержании раздела, темы, необходимо изучить материал с помощью рекомендуемой основной литературы. Целесообразно составить краткий конспект или схему, отображающую смысл и связи основных понятий данного раздела и включенных в него тем. Обязательно следует записывать возникшие вопросы, на которые не удалось ответить самостоятельно.

Подготовку к началу обучения включает несколько необходимых пунктов:

1) Необходимо создать для себя рациональный и эмоционально достаточный уровень мотивации к последовательному и планомерному изучению дисциплины.

2) Необходимо изучить список рекомендованной основной и дополнительной литературы и убедиться в её наличии у себя дома или в библиотеке в бумажном или электронном виде.

3) Необходимо иметь «под рукой» специальные и универсальные словари, справочники и энциклопедии, для того, чтобы постоянно уточнять значения используемых терминов и понятий. Пользование словарями и справочниками необходимо сделать привычкой. Опыт показывает, что неудовлетворительное усвоение предмета зачастую коренится в неточном, смутном или неправильном понимании и употреблении понятийного аппарата учебной дисциплины.

4) Желательно в самом начале периода обучения возможно тщательнее спланировать время, отводимое на работу с источниками и литературой по дисциплине, представить этот план в наглядной форме (график работы с датами) и в дальнейшем его придерживаться, не допуская срывов графика индивидуальной работы и «аврала» в предсессионный период. Пренебрежение этим пунктом приводит к переутомлению и резкому снижению качества усвоения учебного материала.

## Рекомендации по работе с литературой

1) Всю учебную литературу желательно изучать «под конспект». Чтение литературы, не сопровождаемое конспектированием, даже пусть самым кратким – бесполезная работа. Цель написания конспекта по дисциплине – сформировать навыки по поиску, отбору, анализу и формулированию учебного материала. Эти навыки обязательны для любого специалиста с высшим образованием независимо от выбранной специальности.

2) Написание конспекта должно быть творческим – нужно не переписывать текст из источников, но пытаться кратко излагать своими словами содержание ответа, при этом максимально структурируя конспект, используя символы и условные обозначения. Копирование и «заучивание» неосмысленного текста трудоемко и по большому счету не имеет большой познавательной и практической ценности.

3) При написании конспекта используется тетрадь, поля в которой обязательны. Страницы нумеруются, каждый новый вопрос начинается с нового листа, для каждого экзаменационного вопроса отводится 1-2 страницы конспекта. На полях размещается вся вспомогательная информация – ссылки, вопросы, условные обозначения и т.д.

4) В итоге данной работы «идеальным» является полный конспект по программе дисциплины, с выделенными определениями, узловыми пунктами, примерами, неясными моментами, проставленными на полях вопросами.

5) При работе над конспектом обязательно выявляются и отмечаются трудные для самостоятельного изучения вопросы, с которыми уместно обратиться к преподавателю при посещении установочных лекций и консультаций, либо в индивидуальном порядке.

6) При чтении учебной и научной литературы всегда следить за точным и полным пониманием значения терминов и содержания понятий, используемых в тексте. Всегда следует уточнять значения по словарям или энциклопедиям, при необходимости записывать.

7) При написании учебного конспекта обязательно указывать все прорабатываемые источники, автор, название, дата и место издания, с указанием использованных страниц.

### **Подготовка к промежуточной аттестации по дисциплине: экзамену (зачету)**

К аттестации допускаются студенты, которые систематически в течение всего семестра посещали и работали на занятиях и показали уверенные знания в ходе выполнении практических заданий и лабораторных работ.

Непосредственная подготовка к аттестации осуществляется по вопросам, представленным в рабочей учебной программе. Тщательно изучите формулировку каждого вопроса, вникните в его суть, составьте план ответа. Обычно план включает в себя:

— определение сущности рассматриваемого вопроса, основных положений, утверждений, определение необходимости их доказательства;

— запись обозначений, формул, необходимых для полного раскрытия вопроса;

— графический материал (таблицы, рисунки, графики), необходимые для раскрытия сущности вопроса;

— роль и значение рассматриваемого материала для практической деятельности, примеры использования в практической деятельности.

## **VII. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ**

Для осуществления образовательного процесса по дисциплине желательна учебная аудитория для проведения занятий лекционного типа.



МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ  
Федеральное государственное автономное образовательное учреждение  
высшего образования  
«Дальневосточный федеральный университет»  
(ДВФУ)

---

---

ИНСТИТУТ МАТЕМАТИКИ И КОМПЬЮТЕРНЫХ ТЕХНОЛОГИЙ (ШКОЛА)

**УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ  
РАБОТЫ ОБУЧАЮЩИХСЯ**

**по дисциплине «Математические модели социальных сетей»**

**Направление подготовки 01.04.04 Прикладная математика**

**магистерская программа «Аналитические, социальные и экономические сети»**

**Форма подготовки очная**

**Владивосток  
2021**

## План-график выполнения самостоятельной работы по дисциплине

№ п/п	Дата/сроки выполнения	Вид самостоятельной работы	Примерные нормы времени на выполнение	Форма контроля
1	4 неделя	Повторение теоретического и практического материала дисциплины, заслушиваемого и конспектируемого в ходе аудиторных занятий; изучение основной и дополнительной литературы, указанной в рабочей учебной программе дисциплины.	10 часов	Собеседование
2	6 неделя	Самостоятельный разбор заданий и задач, решаемых на практических занятиях.	8 часов	Проект
3	10 неделя	Повторение теоретического и практического материала дисциплины, заслушиваемого и конспектируемого в ходе аудиторных занятий; изучение основной и дополнительной литературы, указанной в рабочей учебной программе дисциплины, самоконтроль ответов на основные проблемные вопросы по темам лекций	10 часов	Собеседование
4	12 неделя	Самостоятельный разбор заданий и задач, решаемых на практических занятиях; самостоятельный повтор действий, осуществляемых в ходе выполнения лабораторных работ, в том числе при работе со специальным программным обеспечением	8 часов	Проект
5	16 неделя	Повторение теоретического и практического материала дисциплины, заслушиваемого и конспектируемого в ходе аудиторных занятий; изучение	10 часов	Собеседование



		основной и дополнительной литературы, указанной в рабочей учебной программе дисциплины, самоконтроль ответов на основные проблемные вопросы по темам лекций		
6	18 неделя	Самостоятельный разбор заданий и задач, решаемых на практических занятиях; самостоятельный повтор действий, осуществляемых в ходе выполнения лабораторных работ, в том числе при работе со специальным программным обеспечением	8 часов	Проект

## Характеристика заданий для самостоятельной работы обучающихся и методические рекомендации по их выполнению

**Термин «сложные сети»**, а именно так обобщенно называются аналитические, социальные и экономические сети, возник в начале 21-го века и относится к сетям с более сложной архитектурой, чем, классические сети с заданным числом узлов и связей (транспортные, энергетические, коммунальные, сети связи и т.д.).

**Сложные сети** являются предметом нашей повседневной жизни, находящей отражение в современных экономических и социальных сетях, и, одновременно, объектом изучения в биологии, экономике, математике, психологии, социологии, инженерии, компьютерных науках. Таким образом, сложные сети – это бурно развивающееся явление сетевой экономики и социальных сетей, требующее от науки нового междисциплинарного подхода, а от образования новых образовательных дисциплин, таких как «Математические модели социальных сетей».

**Государство** применяет сложные сети для организации управления, обеспечения безопасности, выявления потенциальных угроз и выстраивания политики. Последним уровнем государственного управления является правительства краев и областей России, где сетевые технологии управления применяются для стратегического планирования, оценки регулирующих воздействий региональных и федеральных законов, создания соответствующей законодательной базы; анализа и внедрения государственно-частного партнерства.

**Бизнес** использует сети для управления производством, продвижения товаров и услуг, в целеполагании, прогнозировании и стратегическом планировании.

**Домашние хозяйства** применяют сети для организации собственной жизни, развития человеческого капитала, организации взаимодействия с властью и бизнесом.

**Социальные сети**, примерами которых являются Instagram, Google, Twitter, Yelp, WhatsApp, Facebook и др., образуют «моментальную нацию», наиболее типичные фабрики данных в новой цифровой экономике. В отличие от бывших промышленных производств фабрики 21-го века действуют повсюду, где есть подключённый к сети девайс. Труд на этих маленьких девайсах — непрерывный поток твитов, постов, поисковых запросов, обновлений, просмотров, комментариев и снимков — создает всю стоимость в сетевой экономике. Для управления этим хозяйством нужны аналитики, содержанием подготовки которых является дисциплина «Социально-экономические и информационные сети: модели и методы анализа».

**История возникновения (работы Jacob Moreno, Anatol Rapoport, William Horvath).** Социальные сети представляют собой множество людей или групп людей, которые обладают некоторой системой контактов или взаимодействий между собой, такие как: дружба между индивидуумами, деловые отношения между компаниями, брачные отношения между семьями и т.д. [9]

Одним из первых исследователей социальных сетей был Якоб Морено. В 20-30 годах прошлого века он заинтересовался характером дружбы в рамках небольших групп. На рисунке 2 представлена сеть содружества между мальчиками и девочками, учащихся четвертого класса, построенная Якобом Морено. Визуально можно заметить, как сильно эти две группы разделены по половому признаку; за исключением лишь одного мальчика, выбравшего в качестве друга девочку.

В 1961 году математик Anatol Rapoport в совместной работе с William Horvath также исследовали товарищеские взаимоотношения в школе среди младших классов. Они одни из первых попытались описать математически полученную графическую модель, и подчеркнули значимость распределения степеней в сетях.

Экономические сети, в свою очередь, представляют собой набор экономических агентов, имеющих между собой экономические отношения. Примерами таких отношений

являются: внешняя торговля между странами, долговые обязательства между банками или людьми.

Часто социальные сети часто служат средством передачи информации в экономической сфере. Например, при распространении товара и услуг, не обращающихся на рынке. В роли такого «товара» может выступать информация о вакансиях, технологиях, продуктах [9].

В дальнейшем сети будут представляться в виде графа, где набор элементов сети (индивидуумы, экономические агенты) обозначаются вершинами графа, а отношения между ними будут обозначаться как ребра графа. Если ребру графа сопоставлено определенное число, то такой граф будем называть взвешенным графом.

**Свойства сетей.** Сети в реальном мире обладают определенными свойствами, которыми не всегда обладают модели случайных графов. Основной задачей исследователей является получение модели сети, соответствующую тем или иным свойствам реальных сетей.

**Эффект маленького мира.** Эффект маленького мира представляет собой идею, что большие сети склонны иметь маленький диаметр и маленькую среднюю длину геодезической цепи. Геодезической цепью графа называется кратчайшая простая цепь, принадлежащая сети соединяющая две вершины  $u$  и  $v$ . Цепью называется последовательность из чередующихся вершин и ребер, начинающаяся, и заканчивающаяся вершиной. Диаметр графа называется длина (количество ребер) самой длинной геодезической цепи между любыми двумя вершинами.

В 60-х годах прошлого века Стенли Милграм провел эксперимент по исследованию длины цепи в обществе, в котором люди должны были передать письмо через своих знакомых человеку, которого они не знали. Письма были распределены между людьми проживающих в штатах Канзас и Небраска, им было сообщено имя, профессия, и некоторые приближенные данные о нахождении адресата, проживающего в штате Массачусетс. Участники эксперимента должны были передать письмо своим знакомым, вероятно знающим адресата, либо которые могли бы передать его другому знакомому. Приблизительно, только четверть писем достигло цели, и среднее число людей, участвующих в передаче письма было равно 5, а максимальное число – 12. Данные письма не обязательно следовали кратчайшему пути, но, тем не менее, число участников в цепи достаточно малое, что демонстрирует эффект маленького мира [6]. В настоящее время эффект маленького мира изучен и установлен в большом количестве различных сетей. Цепью называется чередующаяся последовательность вершин и различных ребер, начинающаяся и заканчивающаяся в вершине.

**Кластеризация.** Одним из важных наблюдений в социальных сетях является то, что они склонны иметь высокий коэффициент кластеризации. Идея кластеризации важна в социологии, где уделяют внимание треугольникам (трем взаимно связанным вершинам) в социальной сети.

Если в графе вершина  $A$  является смежной вершинам  $B$  и  $C$ , то коэффициент кластеризации сети показывает вероятность того, что вершины  $B$  и  $C$  смежны между собой. Коэффициент кластеризации для вершины  $i$  графа  $G$  определяется по формуле:

$$C_i = \frac{|\{jk \in G: k \neq j, j \in N_i(G), k \in N_i(G)\}|}{|\{jk: k \neq j, j \in N_i(G), k \in N_i(G)\}|} = \frac{|\{jk \in G: k \neq j, j \in N_i(G), k \in N_i(G)\}|}{d_i(G)(d_i(G) - 1)/2}, \quad (1)$$

где  $N_i(G)$  обозначает множество вершин смежных вершине  $i$  в графе  $G$ ,  $d_i(G)$  – степень вершины  $i$  в графе  $G$ .

По-другому формулу 1 можно представить как отношение количества треугольников с вершиной  $i$  к количеству троек с центром в вершине  $i$ .

$$C_i = \frac{\text{количество треугольников с вершиной } i}{\text{количество троек с центром в вершине } i}, \quad (2)$$

где тройкой с центром в вершине  $i$  будем называть цепь длины 2 с центром в вершине  $i$ .

Средний коэффициент кластеризации равен:

$$C = \frac{1}{n} \sum_i C_i \quad (3)$$

Существует глобальный коэффициент кластеризации, или его еще называют коэффициентом транзитивности, который вычисляется по следующей формуле:

$$C = \frac{3 * \text{количество треугольников в графе}}{\text{количество троек в графе}}.$$

**Гомофилия.** Многие социальные сети обладают свойством гомофилии, названной Lazarsfeld и Merton. Данное свойство отражает тот факт, что люди более склонны поддерживать отношения с людьми из схожей группы. Группы могут быть различными, отличающиеся признаками такие как возраст, раса, пол, религия или профессия. На пример, основываясь на национальных опросах, Marsden определил, что только 8% людей обсуждают «важные вопросы» с людьми другой расы. Гомофилия представляет собой важный аспект социальных сетей, так как он означает, что социальные сети могут быть сильно разделены на разные группы [6].

Существует 2 различные формы гомофилии: гомофилия, связанная с предоставленными возможностями, и гомофилия, связанная с выбором. Например, большинство детей в школе имеют лучших друзей близких по возрасту к собственному. В основном это обусловлено дружественными отношениями с детьми, с которыми они постоянно взаимодействуют в школе. Этот пример показывает гомофилию, связанную с предоставленными возможностями. Но даже в таком случае, присутствует тенденция формировать непропорциональную долю отношений с детьми своего возраста, что основано ряде факторов, включающих зрелость детей и их интересы. Такие факторы представляют собой гомофилию, связанную с выбором индивида.

**Слабые связи в сети.** В 1973 году Марк Грановеттер опубликовал одну из самых влиятельных работ в социальных сетях по исследованию слабых связей. В своей работе он определил силу связи как комбинацию продолжительности, эмоциональной интенсивности, близости или взаимного доверия [5]. Также он изучил роль силы связи в социальной сети в нахождении работы участниками сети. Он опросил людей разных профессий одного из городов штата Массачусетс о том, как они нашли свою работу, при этом он отмечал не только социальные контакты при поиске работы, но и силу социальных отношений, измеренную в частоте взаимодействий между индивидуумами за последний год. Из 64 опрошенных человек 16,7% нашли работу через сильную связь (взаимодействие с человеком не реже, чем 2 раза в неделю), 55,7% через среднюю связь (менее 2 раз в неделю, но чаще, чем раз в год) и 27,6% через слабую связь (реже, чем раз в год).

Идея Грановеттера заключалась в том, что индивидуумы, имеющие между собой слабую связь, наименее вероятно имеют между собой общих знакомых, чем индивиды, имеющие сильную связь. Такие связи более вероятно формируют «мосты» между группами, имеющими малое количество связей между собой, и таким образом играют ключевую роль в распространении информации. Грановеттер приходит к заключению, что слабые связи играют связывающую роль во взаимодействиях между группами, в то время, как сильные связи отвечают за сплоченность внутри группы и увеличивают общее разделение в обществе.

**Модели случайных графов.** Первые модели, которые пытались выделить тип графов, отвечающих свойствам и структурам сетей в реальном мире появились из теории графов, и предполагали, что сети формируются случайным образом.

**Модель случайного графа Эрдеша-Реньи.** На рубеже 50-х и 60-х годов XX века модель случайного графа предложили математики П. Эрдеш и А. Реньи независимо с А. Рапопортом. Данная модель также имеет название «Случайный граф Пуассона» или «Граф Бернулли». Данная модель не позволяет построить графы с кратными ребрами, графы с петлями и ориентированные графы.

Пусть дано множество вершин  $V = \{1, \dots, n\}$ . На этом множестве вершин построим случайным образом множество ребер. Максимальное количество ребер, которое может быть построено на  $n$  вершинах равно  $C_n^2$ . Предположим, что ребро между любыми двумя

вершинами  $i$  и  $j$  образуется с вероятностью  $p \in [0,1]$  независимо от всех остальных  $C_n^2 - 1 = \frac{n(n-1)}{2} - 1$  пар вершин. Таким образом, ребра появляются в соответствии со стандартной схемой Бернулли, в которой  $C_n^2$  испытаний, и вероятность появления ребра  $p$ . Такой случайный граф Эрдеша-Реньи обозначается как  $G_{n,p}$ , где  $p$  обозначает вероятность возникновения ребра. Обозначив через  $E$  случайное множество ребер, которое возникает в результате реализации данной схемы, а за  $M$  – множество всех возможных ребер графа, получим вероятность формирования  $|E|$  ребер в графе равно  $p^{|E|} p^{|M|-|E|}$ , где  $|M| = \frac{n(n-1)}{2}$ . Такая модель случайного графа Эрдеша-Реньи обозначается  $G = (V, E)$  [12]. На самом деле,  $G = (V, E)$  представляет собой один из множества случайных графов с  $n$  вершинами, и количеством ребер равным  $|E|$ .

Для случайного графа легко находится распределение степеней вершин графа. Степень вершины – это число инцидентных ей ребер. Распределение степеней вершин в случайном графе описывает вероятность того, что любая заданная вершина имеет степень не меньше  $d_i$ . Таким образом, вероятность, что любая вершина  $i$  имеет строго степень  $d$  равна:

$$p_{d_i} = \binom{n-1}{d_i} p^{d_i} (1-p)^{n-1-d_i}. \quad (4)$$

Несмотря на то, что ребра формируются независимо друг от друга, существует некоторая корреляция между степенями различных вершин. При увеличении числа вершин  $n$  корреляция между степенями любых двух вершин исчезает, так как существует только одно ребро между ними из  $n-1$  возможных инцидентных ребер для каждой вершины. При большом  $n$  и маленьком  $p$  биномиальная формула вероятности (4) приближенно равна распределению Пуассона, таким образом доля вершин, имеющих  $d_i$  инцидентных ребер приближенно равна:

$$p_{d_i} \approx \frac{e^{-(n-1)p} ((n-1)p)^{d_i}}{d_i!} = \frac{e^{-k} k^{d_i}}{d_i!}, \quad (5)$$

Где  $k = (n-1)p$  – это средняя степень вершины при  $n \gg 1, p \ll 1$ .

Ожидаемая структура случайного графа зависит от значения вероятности  $p$ . Анатолий Рапопорт, Рэй Соломонофф и Пол Эрдеш, Эдгар Реньи продемонстрировали важное свойство случайных графов: фазовый переход (phase transition). Фазовый переход подразумевает собой переход случайного графа от одного типа структуры к другой при изменении вероятности формирования ребра. Предположим, что в модели случайного графа Пуассона вероятность формирования ребра зависит от количества вершин, и обозначим вероятность  $p(n)$ . Пусть структура графа является его свойством, которое описывается перечислением всех возможных графов, обладающих этим свойством, на множестве графов с заданным множеством вершин  $N$ , и обозначается  $A(N) \subset G(N)$ . Например, такое свойство, что граф не имеет изолированных вершин (вершин со степенью равной нулю), описывается как  $A(N) = \{g | d_i(g) \neq \emptyset \forall i \in N\}$ , где  $d_i$  – степень вершины  $i$ . Тогда для определенного свойства графа задается пороговая функция  $t(n)$  следующим образом:

$$P[A(N)|p(n)] \rightarrow 1, \text{ если } \frac{p(n)}{t(n)} \rightarrow \infty, \text{ и}$$

$$P[A(N)|p(n)] \rightarrow 0, \text{ если } \frac{p(n)}{t(n)} \rightarrow 0.$$

Если существует такая пороговая функция, то говорят, что возникает фазовый переход в значении пороговой функции.

Обозначим основные свойства случайного графа Пуассона и их пороговые функции:

1) При пороговой функции  $t(n) = 1/n^2$  возникают первые ребра, т.е. при  $p(n)$  большего порядка, чем  $1/n^2$ , существует хотя бы одно ребро в графе с вероятностью, стремящейся к 1;



2) При  $p(n)$  большей, чем  $t(n) = n^{-\frac{3}{2}}$  вероятность того, что в графе присутствует хотя бы одна компонента, имеющая как минимум три ребра, стремиться к 1;

3) При  $p(n)$  большей, чем  $t(n) = 1/n$  в графе образуются циклы, и возникает гигантская компонента, содержащая нетривиальную долю вершин;

4) При  $p(n)$  большей, чем  $t(n) = \log(n)/n$ , граф становится связным.

Модель случайного графа Эрдеша-Реньи позволяет вычислить размер гигантской компоненты с помощью следующих эвристических вычислений. Сгенерируем случайный граф Пуассона с  $n - 1$  вершиной и вероятностью формирования ребра  $p > 1/n$ . Добавим в граф вершину, и соединим её с остальными вершинами с вероятностью  $p$ . Пусть  $q$  – это доля вершин, принадлежащих гигантской компоненте графа из  $n - 1$  вершины. При большом значении  $n, q$  будет также долей вершин, принадлежащих гигантской компоненте графа из  $n$  вершин. Вероятность того, что добавленная вершина не принадлежит гигантской компоненте равна вероятности того, что ни одна из смежных ей вершин не принадлежит гигантской компоненте. Если добавленная вершина имеет степень  $d_i$ , то данная вероятность сходиться к  $(1 - q)^{d_i}$ , при больших  $n$ . Общую долю вершин, не принадлежащих гигантской компоненте может быть найдена усреднением величины  $(1 - q)^{d_i}$  по всем вершинам:

$$1 - q = \sum_d (1 - q)^d P(d). \quad (6)$$

Подставляя в (6) формулу (5), то получаем приближенную долю вершин, не принадлежащих гигантской компоненте:

$$1 - q = \sum_d \frac{e^{-(n-1)p} ((n-1)p)^d}{d!} (1 - q)^d = e^{-(n-1)p} \sum_d \frac{((n-1)p(1 - q))^d}{d!} = e^{-q(n-1)p}.$$

Таким образом, получаем:

$$q = 1 - e^{-q(n-1)p}. \quad (7)$$

Одно из решений уравнения (7) при  $q = 0$ . Тогда и только тогда, когда средняя степень вершины больше единицы, т.е. при  $p > 1/(n - 1)$ , существует решение для  $q$  лежащее между 0 и 1. Данные вычисления показывают наличие фазового перехода, означающим возможность появления гигантской компоненты при пороге в  $(n - 1)p = 1$ . Это показывает различие в структуре случайных графов при средней степени больше или меньше 1.

Случайные графы отражают одно из свойств реальных сетей – маленькая средняя длина кратчайшей цепи между вершинами, называющийся эффектом малого мира. Поскольку  $k$  – это среднее значение степени вершины, а это значит, что, выбирая произвольную вершину, она будет иметь в среднем  $k$  смежных вершин. В свою очередь, эти вершины будут иметь  $k^2$  смежных вершины (их количество  $k$ , и еще  $k$  смежных ребер). После  $l$  шагов, получаем  $k^l$  по отношению к первоначально выбранной вершине. Очевидно, что добавление новых вершин продолжается не бесконечно, процесс останавливается по достижению  $n$  вершин:

$$k^l = n,$$

откуда диаметр случайного графа равен:

$$l = \frac{\log n}{\log k}$$

Однако, для современного мира  $n \approx 6 * 10^9$ , в соответствии с полученным значением  $l \leq 6$  в ходе эксперимента Stanley Milgram, получаем, что  $k \approx 42,6$ , что не соответствует реальности, так как каждый взятый в отдельности человек по большей части имеет более чем 42 знакомых.

Это говорит о том, что, нельзя утверждать, что реальные жизненные сети можно смоделировать, в частности, пуассоновским случайным графом. Тем не менее, все последующие исследования начались с изучения базовых понятий пуассоновского случайного графа.

**Конфигурационная модель (configuration model) случайного графа.** Случайные

графы могут быть расширены множеством способов для того, чтобы они лучше соответствовали реальным сетям. Одна из таких моделей – конфигурационная модель, позволяющая создавать сети с точно заданным распределением степеней вершин (degree distribution), что в свою очередь позволяет генерировать различные сети с тем же распределением, как заданная модель.

Зададим распределение степени в сети  $p_d$  так, что  $p_d$  – вероятность того, что случайно выбранная вершина имеет степень  $d$ . Назовем последовательностью степеней (degree sequence) набор значений степеней  $d_i$  вершин  $i = 1 \dots n$  из этого распределения. Таким образом, можно считать, что у каждой вершины  $i$  в рассматриваемом графе есть «корешки ребер», или их еще называют «хвосты ребер» в соответствии со значением  $d_i$ . Далее, случайным образом выбираем пары «хвостов» из сети и соединяем их в ребра и продельваем эту операцию до тех пор, пока не исчезнут все «хвосты».

Таким образом, ребра формируются случайным образом, исходя из того, что вероятность, что вершины  $i$  и  $j$  окажутся смежными:

$$p_{ij} = \frac{d_i d_j}{\sum_{i=1}^n d_i} = \frac{d_i d_j}{2m}$$

Заметим, что чем выше степень у вершин  $i$  и  $j$ , тем больше вероятность того, что вершины являются смежными.

**Модели сети маленького мира.** Модели маленького мира ставят перед собой цель описать существующие сети, обладающие свойствами эффекта маленького мира и высокой степени кластеризации. Модели случайных графов отражают эффект маленького мира, но коэффициент кластеризации в таких сетях остается на низком уровне. Далее описываются различные модели маленького мира.

**Модель сети маленького мира Уоттса и Строгаца.** Первая модель маленького мира в графовой постановке была предложена Уоттсом и Строгацом (Watts and Strogatz 1998) и генерация графа представляет собой следующую процедуру: пусть имеем одномерную решетку, имеющую  $N$  вершин, с периодическими граничными условиями (т.е. кольцо), где каждая вершина соединена с её ближайшими  $k$  вершинами, далее каждое ребро первоначального графа переопределяем с вероятностью  $p$ . Процесс переопределения включает проход по часовой стрелке через каждое ребро, и с вероятностью  $p$  перемещения одного конца ребра от одной вершине к другой случайной вершине решетки, но при этом запрещено появление петель и кратных ребер.

В модели требуется условие  $n \gg k \gg \ln(n) \gg 1$ , где условие  $k \gg \ln(n)$  гарантирует связность графа. В такой модели при  $p = 0$  мы получаем регулярную решетку. В регулярной решетке коэффициент кластеризации равен  $C = \frac{3(k-2)}{4(k-1)}$ , и стремится к  $\frac{3}{4}$  при  $k$  стремящимся к бесконечности. Тем не менее, в регулярной решетке длина средней геодезической цепи между двумя вершинами вычисляется как  $L = \frac{n(n+k-2)}{2k(n-1)}$ , и при больших  $k$  стремится к  $\frac{n}{2k}$ .

Баррат и Вейгт показали, что при переопределении ребра с вероятностью  $p$  коэффициент кластеризации равен  $C(p) = \frac{3(k-2)}{4(k-1)}(1-p)^3$ . При  $p = 0$  коэффициент кластеризации равен коэффициенту кластеризации в регулярной решетке. При  $p > 0$  две соседние вершины  $i$ , которые были смежными при  $p = 0$  все еще являются соседними вершинами по отношению к  $i$ , и являются смежными по отношению друг к другу с вероятностью  $(1-p)^3$ . Тогда коэффициент кластеризации определяется, как отношение среднего числа ребер между соседними вершинами и среднего числа возможных ребер между соседними вершинами.

Для определения средней геодезической цепи на настоящий момент точной формулы не существует, но известно, что  $L \sim \frac{\ln(n)}{\ln(k)}$ , при  $P \rightarrow 1$ . Т.е. в модели маленького мира средняя геодезическая цепь возрастает логарифмически с увеличением  $n$ . На рисунке 3 показан

график зависимости отношения  $\frac{C(p)}{C(0)}$  и  $\frac{L}{\max(L)}$  от вероятности  $p$ . Как показано на рисунке 3, между  $p = 0$  и  $p = 1$  существует область, где коэффициент кластеризации высок, а средняя длина геодезической цепи между вершинами одновременно мала. Эта область соответствует определению модели маленького мира.

Как было показано Уоттсом граф является графом маленького мира с  $n$  вершинами и средней степенью вершин  $k$ , если одновременно выполняются следующие условия:

1.  $L \approx L_{random}(n, k)$ , где  $L_{random}(n, k)$  – это средняя длина геодезической цепи случайного графа «Эрдеша – Реньи»;
2.  $C \gg C_{random} \approx \frac{k}{n}$ .

Таким образом, для выявления, является сеть – сетью маленького мира, достаточно проверить выполнимость вышперечисленных условий.

Высокий коэффициент кластеризации и маленькая средняя длина кратчайшей цепи отражают такие свойства социальной сети, как гомофилия (homophily) и слабые связи (weak ties).

**Модель маленького мира Ньюмана и Уоттса.** Модель маленького мира Ньюмана и Уоттса (Newman and Watts) является модернизацией модели Уоттса-Строгаца. Генерация обеих моделей начинается с одномерной регулярной решетки на  $N$  вершинах, где каждая вершина смежна с  $k$  её ближайшими вершинами. Но вместо переопределения ребер, в данной модели между случайно выбранными парами вершин добавляются дополнительные ребра, называемые «сокращениями» (shortcuts), и в тоже время ни одно ребро не удаляется из основополагающей решетки.

Рассмотрим данную модель в случае двумерной решетки, и покажем, как модель отражает свойства гомофилии и слабых связей в сети. Предположим, что человек живет в двумерной сети – это может быть примером географического расстояния и социальной близости, благодаря чему в сети могут возникнуть ребра между вершинами. Положим, что две вершины находятся в шаговой доступности (grid step) в сети, если они смежные по отношению друг к другу либо в горизонтальном, либо вертикальном направлениях. Теперь создадим сеть, где каждой вершине инциденты два вида ребер: те, которые будут представлять свойство гомофилии, и тех, которые представляют слабые связи. Пусть для некоторой постоянной величины  $r$  каждая вершина создает ребро со всеми другими вершинами, которые лежат в радиусе шаговой доступности меньшей или равной  $r$  - это соответствует свойству гомофилии.

А для некоторого другого постоянного значения  $k$ , узел также формирует ребро со случайно  $k$  выбранными другими узлами из сети, ранее не связанными друг с другом - это соответствует свойству слабых связей.

Такая сеть имеет множество триад: две соседние вершины имеют общего друга и более того, с высокой вероятностью найдутся пары вершин с короткой длиной пути (short paths connecting) между ними.

Коэффициент кластеризации в данной модели был найден Марком Ньюманом и представляет собой:

$$C = \frac{3(k - 1)}{2(2k - 1) + 4kp(p + 2)}$$

**Модель взвешенной сети маленького мира.** Не смотря на распространённое применение моделей невзвешенных графов в моделировании социальных сетей, многие сети ограничиваются описанием не только топологического представления, но и движением информации в сети или транспортными потоками, которые имеют место быть в определенной сети. Например, однородность в частоте взаимодействий может быть важным в понимании социальных систем, аналогично, необходимо учитывать такие фундаментальные показатели как транспортный поток, характеризующий связи в системе коммуникаций, а также огромную транспортную инфраструктуру, являющуюся фундаментальным показателем для полного описания таких сетей.



Под взвешенной сетью граф, в котором каждое ребро имеет количественную характеристику, данная характеристика называется весом ребра. Для описания взвешенной сети используется взвешенный полный граф, в котором ребра, которые соответствуют отсутствующим связям в сети, имеют нулевой вес. В модели предполагается неориентированный граф, в котором вес ребра  $W_{ij}$  между вершинами  $i$  и  $j$  неотрицательны. На рисунке 8 приведен пример взвешенного графа.

Для проверки свойств маленького мира в взвешенной сети, необходимо ввести новую меру взвешенного коэффициента кластеризации ( $WCC_i$ ) и длины взвешенной геодезической цепи. Для коэффициента кластеризации взвешенной сети необходимо, чтобы выполнялось требование  $WCC_i \rightarrow C_i$  в случае бинарного вида весов, т.е. вес ребра равен либо 0, либо 1. В работе Wenyan, Yongjing, Ying, предложили следующую обобщенную формулу вычисления коэффициента кластеризации вершины  $i$  со степенью  $k_i$  для взвешенных графов:

$$WCC = \frac{2}{k_i(k_i - 1)} \sum_{j,h} \min(\tilde{w}_{ij}, \tilde{w}_{jh}, \tilde{w}_{hi}), \quad (8)$$

где  $\tilde{w}_{ij}$ - нормированный вес ребра между вершинами  $i$  и  $j$ , вычисляемая как отношения веса данного ребра к наибольшему весу ребра в сети, т.е.  $\tilde{w}_{ij} = w_{ij}/\max(w_{kl})$ , что позволяет нормировать коэффициент кластеризации, как долю потока между двумя вершинами от максимального потока в треугольнике. Коэффициент кластеризации графа вычисляется как среднее значение коэффициента кластеризации каждого ребра:

$$WCC = \frac{1}{n} \sum_i WCC_i. \quad (9)$$

Другая формула расчета коэффициента кластеризации была предложена Барратом и соавторами. Предложенная формула учитывала важность структуры кластера, беря за основу интенсивность взаимодействий в кластере, и определялась как

$$\tilde{c}_i = \frac{1}{s_i(k_i - 1)} \sum_{j,h} \frac{w_{ij} + w_{jh}}{2} a_{ij} a_{ih} a_{jh}, \quad (10)$$

где  $s_i$  – это сила (strength) вершины  $i$ , определяемая, как  $s_i = \sum_j^N a_{ij} w_{ij}$ , и  $a_{ij}$  – это элемент бинарной матрицы смежности. Количественное измерение силы вершины показывает общий вес инцидентных вершине ребер, другими словами, сила вершины показывает количество потока, проходящего через вершину. Например, Баррат и соавторы в работе рассмотрели структуру сети авиасообщений из базы данных International Air Transportation Association ([www.iata.org](http://www.iata.org)), где вершинами соответствовали аэропорты, а вес ребра отображал количество доступных мест по перелету между аэропортами за 2002 год. Сеть авиасообщений содержала 3 880 вершин (аэропортов) и 18 810 ребер, обозначающих возможность прямого перелета между аэропортами. Баррат и соавторы показали, что данная сеть является сетью маленького мира, если не учитывать вес ребер. Для формулы коэффициента кластеризации в сети авиасообщений сила вершины показывает количество человек, которое потенциально мог обслужить аэропорт в 2002 году.

Также Баррат и соавторы рассмотрели структуру сети соавторства ученых, где вершины обозначили ученых, а ребра соответствовали тому, являлись ли ученые соавторами хотя бы в одной статье. Вес ребра в данном случае показывал интенсивность взаимодействия ученых  $i$  и  $j$ , и определялся как  $w_{ij} = \sum_p \frac{\delta_i^p \delta_j^p}{n_p - 1}$  если  $i \neq j$ ,  $w_{ii} = 0$ , где индекс  $p$  пробегает по всем научным статьям,  $n_p$  – это количество авторов в статье  $p$ , и  $\delta_i^p$  равно 1, если автор  $i$  принимал участие в написании статьи  $p$ , и равно 0 в противном случае. Определение интенсивности взаимодействий в социальной сети всегда зависит от объектов, которые описывает сеть, но в сети соавторства данное определение отражает взаимодействие ученых между собой: вклад каждого ученого в одну статью тем меньше, чем больше авторов статьи.

Коэффициент кластеризации, рассчитанный по формуле (10), является мерой локальной сплоченности, который учитывает важность кластерной структуры базиса,

основанной на интенсивности движения транспорта либо взаимодействий в тройке вершин с центром в вершине  $i$ . Формула (10) учитывает вес только двух ребер, инцидентных вершине  $i$  ( $w_{ij}$  и  $w_{ih}$ ), и игнорирует вес третьего ребра ( $w_{jh}$ ). Более того данное определение коэффициента кластеризации не удовлетворяет требованию  $WCC_i \rightarrow C_i$ , при бинарном виде весов [16]. Множитель  $1/s_i(k_i - 1)$  нормализует коэффициент, учитывая вес каждого ребра, инцидентного вершине, умноженный на максимальное возможное количество треугольников с вершиной  $i$ .

Оннела и соавторы предложили формулу коэффициента кластеризации, основанную на средней геометрической мере весов ребер [11]:

$$\hat{C}_i = \frac{2}{k_i(k_i - 1)} \sum_{j,h} (\tilde{w}_{ij}, \tilde{w}_{jh}, \tilde{w}_{hi})^{\frac{1}{3}}. \quad (11)$$

Формула (11) так же удовлетворяет требованию  $\hat{C}_i \rightarrow C_i$ , при бинарном виде весов.

Коэффициент кластеризации графа вычисляется как среднее значение коэффициента кластеризации каждого ребра:

$$\hat{C} = \frac{1}{n} \sum_i \hat{C}_i. \quad (12)$$

Занг и др. предложили модель расчета коэффициента кластеризации вершины в взвешенной сети, учитывая только веса ребер:

$$C_3^w(i) = \frac{\sum_{i \neq h} \sum_{j \neq i, j \neq h} w_{ij} w_{jh} w_{hi}}{\sum_{j \neq h} w_{ij} w_{ih}}. \quad (13)$$

Формула (13) считает общую интенсивность взаимодействий в тройках вершин для частной вершины  $i$ , нормированный коэффициент между значениями 0 и 1 по взаимодействиям во всех цепях длины 2 с серединой в вершине  $v_i$ . Данная модель склонна давать сбой в идентификации треугольников в сети. Ошибка происходит в случае, когда вес одного из ребер в треугольнике значительно меньше весов остальных ребер и коэффициент кластеризации остается высоким. Ошибка идентификации заключается в том, что треугольник с вершиной  $i$  идентифицируется машиной как тройка с центром в вершине  $i$ .

В работе предлагается модель расчета коэффициента кластеризации в взвешенной цепи на основе формулы 9:

$$C^w(i) = \frac{3 \sum_{j,h \neq i} w_{ij} w_{jh} w_{hi}}{\sum_{j,h \neq i} w_{ih} w_{hj} + \sum_{j,h \neq i} w_{ij} w_{jh} w_{hi}}. \quad (14)$$

Формула (14) трактует вес ребра  $w_{ij}$  как вероятность прохождения потока информации между вершинами  $i$  и  $j$  напрямую, без вершин в цепи между ними. Коэффициент кластеризации рассчитывается как отношение общей вероятности формирования треугольника по отношению к вероятности всех возможных связанных графов с тремя вершинами, содержащих вершину  $i$ . Такое приближение так же гарантирует, что значения коэффициента лежат между значениями 0 и 1.

Формулы (13) и (14) позволяют найти значимые вершины в сети, вершины, через которые наиболее вероятно прохождение информации. Такие вершины могут быть ключевыми в формировании сообществ в социальной сети. В работе была рассмотрена взвешенная сеть Каратэ Клуба Захари (Zachary), представленная на рисунке 10, где вес ребра означал частоту общения между членами клуба. Авторами было замечено, что вершины 1, 33 и 34 имеют более высокий коэффициент кластеризации, и наименьший коэффициент кластеризации имеют вершины 15 и 16, которые являются обычными членами клуба. Вершины 1, 33 и 34 соответствуют наиболее старшим членам клуба, и отношения между этими членами являются ключевыми в формировании двух подгрупп в Каратэ Клубе. Таким образом, формулы взвешенного коэффициента (13, 14) позволяют идентифицировать центральные, с точки зрения важности формирования групп, вершины.

Для расчета длины цепи в взвешенном графе предлагается определить расстояние между двумя вершинами, если между ними существует ребро, как величину обратную весу

ребра между заданными вершинами. Такой метод используется, если вес ребра обозначает силу или интенсивность взаимодействий между вершинами. Таким образом, чем сильнее взаимодействие между вершинами, тем меньше расстояние между ними.

В работе Bolanos, Bernat, Aviyente предложили следующую модель расчета длины цепи. Предполагается, что от веса ребер в цепи  $P_{ij}$  зависит объем потока информации передаваемый в цепи от вершины  $i$  к вершине  $j$ , где  $P_{ij} = \{w_{i,q_1}, w_{q_1,q_2}, \dots, w_{q_m,j}\}$  – последовательность весов ребер в цепи и  $Q = \{q_1, q_2, \dots, q_m\}$  – это последовательность вершин в цепи  $P_{ij}$ . Далее применяется алгоритм поиска максимального потока от вершины  $i$  к вершине  $j$ , где функция потока задается как  $F_{ij} = w_{i,q_1} w_{q_m,j} \prod_{l=2}^m w_{q_{l-1},q_l}$ . Тогда длина взвешенной цепи определяется как

$$L^w(i) = \frac{1}{N-1} \sum_{j \neq i} \frac{1}{F_{ij}}. \quad (15)$$

Длина взвешенной цепи для вершины  $i$  является средней вероятностью передачи потока информации от вершины  $i$  ко всем остальным вершинам.

Для того, чтобы сеть являлась сетью маленького мира при измерении коэффициента кластеризации по формулам 4-8 и измерении средней длины взвешенной цепи, необходимо чтобы соблюдались те же 2 условия что и для невзвешенной цепи в сравнении с соответствующим ему случайным графом. Соответствующий случайный граф генерируется следующей процедурой, состоящей из двух шагов:

- 1) Для взвешенного графа генерируется невзвешенный случайный граф;
- 2) Веса взвешенного графа случайным образом распределяются по ребрам случайного невзвешенного графа.

Для определения принадлежности сети к с сети маленького мира, при рассчитанной средней взвешенной цепи для каждой вершины по формуле (11) и коэффициента кластеризации по формуле 10 предлагается нормировать коэффициент кластеризации и среднюю длину взвешенной цепи для вершины по средним значениям соответствующих показателей для суррогатного графа, т.е.  $C^{w'}(i) = C^w(i)/\langle C_{rand} \rangle$  и  $L^{w'}(i) = L^w(i)/\langle L_{rand} \rangle$ , где суррогатный граф генерируется как случайная перестановка взвешенных ребер оригинального графа. Мера сети взвешенного маленького мира для каждой вершины определяется как  $\sigma^w(i) = C^{w'}(i)/L^{w'}(i)$ . Тогда существует локальный маленький мир в сети с центром в вершине  $i$  когда  $\sigma^w(i) \gg 1$ .

**Модели роста сети.** В отличие от стационарных сетей, существует особый класс случайных графов, в которых новая вершина появляется с течением времени и при появлении формирует ребра с существующими вершинами. Основной целью создания таких моделей была попытка описания моделей интернета, где новая страница создает ссылки на уже существующие. Подобными моделями можно описать и социальные сети. Например, установление новых знакомств, когда человек приходит в новую школу, нанимается на работу, или знакомится с соседями после переезда.

**Модель Барабаши-Альберт.** Одной из первых моделей, описывающих рост сети, была предложена А.-Л. Барабаши и Р. Альберт как модель, описывающая рост сети интернет. Суть модели состоит в следующем, пусть в каждый момент времени появляется одна вершина, и она формирует фиксированное количество ребер к уже существующим вершинам таким образом, что вероятность формирования ребра к существующей вершине пропорционально степени этой вершины. Модели случайных графов, основанные на описанной идее, называются моделями предпочтительного присоединения.

Рассмотрим модель Барабаши - Альберт более детально. Пусть вершины появляются с течением времени и проиндексированы по времени появления  $i \in \{0, 1, 2, \dots, t, \dots\}$ . Каждая ново появившаяся вершина формирует  $m$  ребер с уже существующими вершинам с вероятностью пропорциональной степени этих вершин. Таким образом, вероятность, что существующая вершина  $i$  сформирует ребро с новой вершиной в момент времени  $t$  равна

$m \frac{d_i(t)}{\sum_{j=1}^t d_j(t)}$ , где  $d_i(t)$  – степень вершины  $i$  в момент времени  $t$ . Так как в сети в момент времени  $t$ , то  $\sum_{j=1}^t d_j(t) = 2tm$ . Следовательно, вероятность, что вершина  $i$  сформирует новое ребро в период  $t$  равно  $\frac{d_i(t)}{2t}$ .

Предполагается, что на начало процесса уже существует группа из  $m$  вершин, каждая смежная с каждой. Далее происходит определенный выше стохастический процесс. Процесс изменения степени вершины в непрерывном времени описывается следующей формулой:

$$\frac{dd_i(t)}{dt} = \frac{d_i(t)}{2t},$$

с начальным условием  $d_i(i) = m$ .

В работе Н. Jeong, Z. Neda и A.L. Barabasi проверили гипотезу о линейности механизма предпочтительного присоединения на 4х реальных сетях: сети цитирования, интернет сети, сети соавторства и сети актеров. Авторы обозначают вероятность того, что вершина со степенью  $d$  сформирует ребро с новопоявившейся вершиной, как  $\Pi(d)$ , тогда изменение степени  $d_i$  вершины  $i$  записывается как:

$$\frac{dd_i(t)}{dt} = m\Pi(d_i(t)),$$

где  $\Pi(d_i(t))$  имеет вид

$$\Pi(d_i(t)) = \frac{d_i^\alpha(t)}{\sum_j d_j^\alpha(t)} = C(t)d_i^\alpha(t),$$

с  $\alpha > 0$  – неизвестный показатель степени. Где  $\alpha = 1$  говорит о линейности функции вероятности от степени вершины, соответствующей стандартной модели предпочтительного присоединения. Для проверки гипотезы Jeong и соавторы [20] проверили значение  $\alpha$ , и для каждой из четырех существующих сетей было подтверждено существование механизма предпочтительного соединения. Было показано, что вероятность присоединения ребра к существующей вершине рассчитывается по более сложной функции, в которой предполагается, что вероятность  $\Pi(d(t))$  рассчитывается по степенному закону. Степень  $\alpha$  уникальна для каждой сети: в то время как для сети интернет и сети цитирования множитель  $\alpha$  близок к единице, и вероятность  $\Pi(d(t))$  является линейной, то для сети актеров и сети соавторства степень  $\alpha < 1$ .

**Модель Боллобаша-Риордана.** Модель Боллобаша-Риордана также была создана с целью описания роста сети интернет, и поэтому по своим свойствам она схожа на модель Барабаша-Альберт.

Построим последовательность случайных графов  $\{G_1^n\}$ , в которой у графа  $n$  вершин и  $n$  ребер. Затем из полученной последовательности сделаем последовательность  $\{G_k^n\}$ , в которой у графа с номером  $n$  число вершин равно  $n$ , а число ребер равно  $kn$ ,  $k \in \mathbb{N}$ . Тогда, пусть  $G_1^1 = (\{1\}, \{(1,1)\})$ , т.е. в начальный момент времени имеем одну вершину и одну петлю. Пусть построен граф  $G_1^{n-1}$ , у него вершины образуют множество  $\{1, \dots, n-1\}$ , а количество ребер равно  $n-1$ . Добавим вершину  $n$  и ребро  $(n, i)$ , у которого  $i \in \{1, \dots, n\}$ . Ребро  $(n, n)$  образуется с вероятностью  $\frac{\deg i}{2n-1}$ , где  $\deg i$  – это степень вершины  $i$  в графе  $G_1^{n-1}$ . Случайный граф  $G_1^n$  построен и удовлетворяет принципу предпочтительного присоединения.

Для перехода к  $G_k^n$  выбираем граф  $G_1^{kn}$ . Этот граф имеет  $kn$  вершин и  $kn$  ребер. Делим множество его вершин на последовательные подмножества размера  $k$ :

$$\{1, \dots, k\}, \{k+1, \dots, 2k\}, \dots, \{k(n-1)+1, \dots, kn\}.$$

Объявляем каждое подмножество вершиной, а ребра сохраняем, т.е. если ребра принадлежали паре вершин, принадлежащих одному подмножеству, то образуются кратные петли, а если они принадлежали паре вершин из разных подмножеств, то получили кратные ребра. В итоге получаем случайный мультиграф.

**Модели роста графов малого мира и их алгоритмы.** Граф «Малого мира» – такой граф, в котором две произвольные вершины  $a$  и  $b$  с большой вероятностью не являются

смежными, однако одна достижима из другой посредством небольшого количества переходов через другие вершины. Типичное расстояние  $L$  между двумя произвольно выбранными вершинами (количество шагов, необходимых, чтобы достичь одну из другой) растёт пропорционально логарифму от числа вершин  $N$  в сети.

В исходной модели Ватца и Строгаца предлагается следующий подход. Пусть  $N$  – число вершин,  $K$  – средняя степень вершины,  $\beta$  – специальный параметр, причем  $0 \leq \beta \leq 1$ . Модель строит ненаправленный граф  $N$  вершин и  $\frac{NK}{2}$  рёбер следующим образом:

1) Строится кольцевая решетка – граф с  $N$  вершин, каждая из которых соединена с  $K$  соседними, по  $K/2$  с каждой стороны. Ребро  $(n_i, n_j)$  существует тогда и только тогда, когда  $0 < |i - j| \bmod (N - 1 - \frac{K}{2}) \leq \frac{K}{2}$

2) Для каждой вершины  $n_i$  берется ребро  $(n_i, n_j)$ , где  $i < j$  и меняется на ребро  $(n_i, n_k)$ , где  $k$  выбирается равновероятно среди всех значений помимо самого  $i$  и номеров уже соединенных с ним вершин (ребра не дублируются)

Традиционный алгоритм генерации сетей, описывающий рост безмасштабных графов, в том числе графов Эрдёша-Реньи и графов «Малого мира» (согласно модели Ватца-Строгаца) – модель Барабаши-Альберт (БА). Она включает в себя две важные общие концепции: рост графа и принцип предпочтительного присоединения. Обе концепции представлены в реальных примерах сетей.

Алгоритм работает следующим образом: граф начинается в виде начальной сетки с  $m_0$  вершин.  $m_0 \geq 2$  и степень каждой вершины в начальном графе должна быть не меньше 1, т.е. изолированные вершины отсутствуют.

В каждый момент времени в граф добавляется новая вершина. Каждая новая вершина соединяется с существующими с вероятностью, пропорциональной числу связей этих узлов. Формально, вероятность  $p_i$  того, что новая вершина соединится с вершиной  $i$  равна  $p_i = \frac{k_i}{\sum_j k_j}$ , где  $k_j$  – степень вершины  $i$ , а в знаменателе суммируются степени всех

существующих узлов. Наиболее связанные вершины («хабы»), как правило, накапливают ещё больше связей, тогда как узлы с небольшим числом связей вряд ли будут выбраны для присоединения новых узлов. Новые узлы имеют «предпочтение» соединяться с наиболее связанными узлами. Графы, построенные данным алгоритмом, имеют более короткий средний путь, нежели случайный граф, средняя длина пути увеличивается как логарифм размера графа.

**Современные модели роста.** В работах Ozik, J., Hunt, B. и Aoyuan Peng, Lianming Zhang получил развитие следующий алгоритм детерминированного многомерного роста (DMG) графа «малого мира»:

1. На нулевом шаге ( $t=0$ ) даны 4 первых вершины, которые соединяются дугами в подобие треугольной пирамиды

2. Начиная с первого шага ( $t \geq 1$ ) граф DMG( $t$ ) формируется из графа DMG( $t-1$ ), полученного на предыдущем шаге, путем добавления новой вершины напротив каждой дуги исходного графа и соединения её дугами с концами этой дуги.

Очевидно, что на каждом шаге алгоритма, начиная с  $t=1$ , число генерируемых вершин составляет:  $\Delta n_t = 3 \times 2^t$ ,  $t \geq 1$ , а общее число вершин  $n_t$  составляет  $n_t = n_0 + \Delta n_1 + \Delta n_2 + \dots + \Delta n_t = 3 \times 2^{t+1} - 2$ , а число дуг равно  $e_t = 3 \times 2^{t+2} - 6$ .

Средняя степень в таком графе стремится к 4 и составляет:

$$\langle k \rangle (t) = \frac{2e_t}{n_t} = 4 - \frac{2}{3 \times 2^t - 1}$$

Функция распределения степени для такого графа представима в следующем виде:



$$p(k) = P(k' > k - 1) - P(k' > k) = \frac{(\sqrt{2} - 1) \times 2^{-\frac{k}{2}}}{1 - 1/(3 \times 2^t)}$$

Общий коэффициент кластеризации стремится к 0,693 и описывается следующей формулой:

$$\begin{aligned} C &= \frac{1}{n_t} \left[ \sum_{i=1}^t \frac{1}{i} \times \Delta n(t-i+1) + \frac{2}{3t+2} \times 4 \right] \\ &= \frac{1}{3 \times 2^{t+1} - 2} \left[ 1 \times 3 \times 2^t + \frac{1}{2} \times 3 \times 2^{t-1} + \dots + \frac{1}{t} \times 3 \times 2^1 + 4 \times \frac{2}{3t+2} \right] \\ &= \frac{1}{1 - \frac{1}{3 \times 2^t}} \times \ln 2 + \frac{4}{3 \times 2^{t+1} - 2} \times \frac{2}{3t+2} \end{aligned}$$

Средняя длина пути в таком графе приблизительно равна  $L(t) \approx 2t/3$ , диаметр, соответственно, составляет приблизительно  $D \approx 1.46 \ln(n)$ .

В работе Yilun Shang предлагается ассортативная геометрическая модель роста  $G(m,t)$ , во многом аналогичная модели DMG. На каждом шаге к каждой дуге графа из числа появившихся на предыдущем шаге, «прирастает» уже  $m$  вершин, которые соединяются ребрами с вершинами-концами генерирующей дуги.

Таким образом, в данной модели рост графа ускоряется, что имитирует рост сети во многих реальных системах. Порядок, размер графа, средняя степень вершины представимы

$$|V(t)| = \frac{m(2m)^t + 3m - 2}{2m - 1}, \quad |E(t)| = \frac{(2m)^{t+1} - 1}{2m - 1},$$

следующим образом:

$$\bar{\delta}(t) = 4 - \frac{12m - 6}{m(2m)^t + 3m - 2}.$$

В генерируемом таким образом графе, как и в графе DMG, степень стремится к 4.

Сопоставление параметров графов, генерируемых моделью детерминированного многомерного роста (DMG), и социальных графов сообществ сети Facebook

Среди рассмотренных моделей интерес для исследования представляет модель DMG, за малое число итераций генерирующая граф большой размерности со свойствами графа малого мира. Однако, для оценки её применимости и, возможно, модификации, необходимо сопоставить её с существующими социальными графами.

В качестве примера могут быть использованы исследованные в ходе производственной практики с использованием инструментальной среды Gephi графы активности сообществ социальной сети Facebook. Рассмотрим сообщества трёх университетов – ДВФУ, MIT и Кембриджского университета.

Табл. 1. Сравнение параметров графов активности на страницах в Facebook за 2015 год и графов с аналогичным числом вершин, генерируемых алгоритмом DMG

Граф		ДВФУ	MIT	Cambridge
Фактические параметры графа сообщества	Число вершин	1892	89293	219182
	Число дуг	5596	368062	692195
	Диаметр	8	6	7
	Средняя степень	5,915	8,245	6,316
	Средняя длина пути	3,457	14,437	20,183
	Коэфф. кластеризации	0	0	0
Параметры графа DMG с тем же	Число дуг DMG	3776	178543	438416
	Диаметр DMG	7,54	11,39	12,29
	Средняя степень DMG	3,997	3,999	3,999

числом вершин	Средняя длина пути DMG	12,45	20,8	15,157
	Коэфф. кластеризации DMG	0,698	0,694	0,693

Графы, созданные алгоритмом DMG для того же числа вершин, хотя и оказались менее связными, но обладают существенно более высоким коэффициентом кластеризации, чем графы активности сообществ в социальной сети. Таким образом, подтверждается то, что данные графы активности сообществ не обладают свойствами графов малого мира в той мере, в какой ими обладает «эталонная» модель роста и генерации такого графа. Для генерации подобного им графа (с потерей свойств «малого мира») необходимо создавать приблизительно вдвое больше дуг.

**Заключение.** Как мы видим, существующие алгоритмы роста графов малого мира несмотря на детерминированный характер позволяют моделировать рост сетей, в некотором роде сходных с реальными, однако учитывают так называемые «слабые» связи в меньшей мере, чем они фактически проявляются в современных социальных сетях.

Отличия в параметрах полученных графов объясняются также тем, что в графе активности не учитываются пользователи, не проявившие активности в определенный период (что приводит к занижению числа вершин и, вероятно, завышению средней степени связности оставшихся).

Уместно также заметить, что не все сообщества современных социальных сетей в полной мере соответствуют определению «малого мира».

В целом, детерминированные модели, подобные модели DMG и модели геометрического ассортативного роста, пригодны для генерации графов «малого мира» и могут быть использованы для имитации динамики роста сетей реального мира, но лишь после внесения модификаций.

## **Требования к представлению и оформлению результатов самостоятельной работы**

Самостоятельная работа включает в себя повторение теоретического и практического материала дисциплины, заслушиваемого и конспектируемого в ходе аудиторных занятий; изучение основной и дополнительной литературы, указанной в рабочей учебной программе дисциплины, самоконтроль ответов на основные проблемные вопросы по темам занятий; самостоятельный разбор заданий и задач, решаемых на практических занятиях; самостоятельный повтор действий, осуществляемых в ходе выполнения лабораторных работ, в том числе при работе со специальным программным обеспечением.

Результаты самостоятельной работы представляются и оформляются в виде ответов на основные положения теоретического и практического материала дисциплины по темам; письменного разбора процесса решения практических заданий и задач; собственных действий, осуществляемых в ходе выполнения лабораторных работ.

В случае подготовки слайдов для защиты проекта, они должны быть контрастными (рекомендуется черный цвет шрифта на светлом фоне), кегль текста слайдов – не менее 22pt, заголовков – 32pt. Основная цель использования слайдов - служить вспомогательным инструментом к подготовленному выступлению, цитирование больших фрагментов текста на слайдах не допускается. Приветствуется использование рисунков, графиков, таблиц, интерактивного материала, однако, следует предусмотреть выбор цвета и толщину линий.

Слайды должны содержать титульный лист, цели и задачи (не более 2-х слайдов с обзором актуальности, новизны, теоретической и практической значимости работы), основные публикации с их кратким обзором (1-2 слайда), формальную постановку задачи и формулировку моделей (1-2 слайда), краткое тезисное (!) изложение ключевых положений работы (разумное количество слайдов с учетом общего времени выступления), заключение (с изложением результатов работы, подведением выводов, обсуждением практического использования работы, возможностей проведения дальнейших исследований и разработок в данной области).

Как правило, 12-15 слайдов оказывается достаточным для полного представления работы.

### **Критерии оценки выполнения самостоятельной работы**

Общие критерии оценки выполнения самостоятельной работы – правильность ответов на вопросы по темам теоретической части дисциплины, верность получаемых ответов в ходе решения практических заданий и задач,



достижение правильного результата при осуществлении собственных действий по лабораторным работам.

Оценивание знаний в форме собеседования проводится по критериям:

- логичность изложения, знание и понимание основных аспектов и дискуссионных проблем по теме;
- владение методами и приемами анализа теоретических и/или практических аспектов по теме.

Оценивание знаний в форме проекта проводится по критериям:

- завершенность и полнота выполненных заданий в рамках проекта;
- владение методами и приемами решения конкретных задач и самостоятельность использования специализированного программного обеспечения;
- качество оформления письменного отчета в соответствии с правилами и стандартами оформления.