

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»**

Кафедра дискретной математики и информационных технологий

**ЗАДАЧИ МАКСИМИЗАЦИИ ВЛИЯНИЯ В СЛОЖНЫХ
СЕТЯХ**

АВТОРЕФЕРАТ МАГИСТЕРСКОЙ РАБОТЫ

студента 2 курса 271 группы
направления 09.04.01 — Информатика и вычислительная техника
факультета КНиИТ
Алексаненкова Дениса Павловича

Научный руководитель
доцент, к. ф.-м. н.

И. Д. Сагаева

Заведующий кафедрой
доцент, к. ф.-м. н.

Л. Б. Тяпаев

Саратов 2026

ВВЕДЕНИЕ

В современных условиях социальные сети, коммуникационные платформы и информационные системы играют важную роль в распространении информации. Одной из ключевых задач анализа сложных сетей является задача максимизации влияния, которая заключается в выборе такого набора начальных узлов сети, при активации которых распространение информации будет максимальным. Данная задача имеет широкий спектр практических применений, включая вирусный маркетинг, распространение новостей, информационную безопасность, анализ социальных сетей и моделирование эпидемий.

Актуальность исследования обусловлена постоянным ростом объемов сетевых данных и необходимостью разработки эффективных алгоритмов поиска наиболее влиятельных узлов в больших графах. Классические методы решения задачи максимизации влияния обладают высокой вычислительной сложностью, что делает необходимым исследование более эффективных алгоритмов, включая эвристические подходы и методы машинного обучения.

Целью магистерской работы является исследование и программная реализация алгоритмов решения задачи максимизации влияния в сложных сетях.

Для достижения поставленной цели были сформулированы следующие задачи:

1. проанализировать материалы по моделям распространения информации и алгоритмам максимизации влияния в сложных сетях;
2. программно реализовать модели распространения информации и алгоритмы максимизации влияния в сложных сетях;
3. выяснить особенности моделей распространения информации и алгоритмов максимизации влияния в сложных сетях;

В первой части рассматриваются модели сложных сетей, алгоритмы максимизации влияния и модели распространения информации. Описаны сети Уоттса–Строгаца, сети Барабаши–Альберта и сети Холма–Кима. Исследуются жадные алгоритмы CELF, CELF++ и FASTCELF++, эвристические методы на основе центральностей, а также алгоритмы, основанные на машинном обучении с подкреплением. Описаны эпидемиологические модели SI, SIS и SIR, используемые для моделирования процессов распространения инфор-

мации.

Во второй части программно реализованы все рассмотренные алгоритмы максимизации и модели распространения информации в сложных сетях и выполнен сравнительный анализ результатов алгоритмов симуляций для различных сценариев распространения информации.

Работа состоит из введения, четырех глав, заключения, списка литературы и приложений.

Структура глав работы:

1. Сложные сети
 - а) Сети Уоттса–Строгаца
 - б) Сети Барабаши–Альберта
 - в) Сети Холма–Кима
2. Алгоритмы решения задачи максимизации влияния
 - а) Жадные алгоритмы
 - i. CELF
 - ii. CELF++
 - iii. FASTCELF++
 - б) Эвристические алгоритмы
 - в) Алгоритмы, основанные на машинном обучении
 - i. Обучение с подкреплением
3. Модели распространения информации
 - а) Модель SI
 - б) Модель SIS
 - в) Модель SIR
4. Программная реализация
 - а) Сбор реальных данных
 - б) Описание программной реализации
 - i. Описание создания сложных сетей
 - ii. Описание поиска начальных узлов
 - iii. Описание симуляции заражения сети
 - iv. Описание построения графиков
 - в) Примеры работы программы

В первой главе рассматриваются основные модели сложных сетей и их свойства. Описаны принципы построения сетей Уоттса–Строгаца, Барабаши–Альберта и Холма–Кима. Приведены основные характеристики сложных сетей, включая коэффициент кластеризации, среднюю длину пути и распределение степеней узлов.

В подразделе 1.1 рассматриваются сети Уоттса–Строгаца, обладающие свойствами малых миров. Описаны механизмы их построения и особенности применения для моделирования социальных взаимодействий.

В подразделе 1.2 представлены безмасштабные сети Барабаши–Альберта. Рассмотрен механизм предпочтительного присоединения и особенности степенного распределения степеней узлов.

В подразделе 1.3 описываются сети Холма–Кима, сочетающие свойства безмасштабных сетей и высокий коэффициент кластеризации.

Во второй главе рассматриваются алгоритмы решения задачи максимизации влияния. Выполнен анализ жадных, эвристических и обучаемых методов выбора наиболее влиятельных узлов в сети.

В подразделе 2.1 рассматриваются жадные алгоритмы CELF, CELF++ и FASTCELF++. Описаны их принципы работы, механизмы оптимизации вычислений и особенности применения для больших графов.

В подразделе 2.2 представлены эвристические алгоритмы, основанные на структурных характеристиках сети, включая Degree Centrality, Betweenness Centrality и PageRank.

В подразделе 2.3 рассматриваются алгоритмы, основанные на машинном обучении, применяемые для решения задачи максимизации влияния в сложных сетях. Описаны основные принципы обучения моделей, этапы предсказания влияния узлов и особенности использования методов машинного обучения для анализа структуры графов.

В подразделе 2.3.1 рассматриваются методы обучения с подкреплением, позволяющие адаптивно формировать стратегию выбора узлов на основе взаимодействия агента с сетевой средой. Описаны принципы работы алгоритмов на основе обучения с подкреплением, этапы обучения агента и особенности применения методов обучения с подкреплением для максимизации распространения информации.

В третьей главе рассматриваются модели распространения инфор-

мации в сложных сетях. Описаны модели SI, SIS и SIR, используемые для моделирования процессов распространения информации и эпидемий.

В подразделе 3.1 рассматривается модель SI, в которой узлы переходят из состояния восприимчивости в состояние заражения.

В подразделе 3.2 описывается модель SIS, учитывающая возможность повторного заражения узлов.

В подразделе 3.3 представлена модель SIR, включающая состояния восприимчивости, заражения и выздоровления.

В четвертой главе представлена программная реализация системы моделирования процессов распространения информации и решения задачи максимизации влияния в сложных сетях. Программный комплекс реализован на языке программирования C# с использованием платформы .NET и технологии Windows Forms. Для реализации и экспериментального исследования алгоритмов, основанных на машинном обучении и обучении с подкреплением, использовался язык Python.

В главе рассматриваются архитектура программной системы, основные модули и особенности реализации алгоритмов генерации сложных сетей, поиска наиболее влиятельных узлов и моделирования процессов распространения информации. Описаны средства визуализации результатов моделирования и анализа динамики распространения информации.

В подразделе 4.1 описывается процесс сбора и подготовки реальных данных, использованных для экспериментального исследования алгоритмов. Рассматриваются наборы данных из репозитория Stanford Network Analysis Project (SNAP), а также методы предварительной обработки графов и переиндексации узлов.

В подразделе 4.2 рассматривается применение методов машинного обучения для решения задачи максимизации влияния. Описаны используемые структурные характеристики узлов, процесс формирования обучающей выборки и особенности реализации подхода на основе обучения с подкреплением.

В подразделе 4.3 приводится описание программной реализации системы. Рассматриваются основные классы программы, реализующие генерацию сложных сетей, алгоритмы максимизации влияния, модели распространения информации и построение графиков.

В подразделе 4.4 представлены примеры работы программы и результаты экспериментального исследования. Выполнено сравнение жадных, эвристических и RL-алгоритмов на реальных и синтетических сетях различной структуры и размерности. Проведен анализ времени работы алгоритмов и качества распространения информации.

В результате проведённых экспериментов установлено, что жадные алгоритмы обеспечивают высокое качество решений, однако обладают значительной вычислительной сложностью. Эвристические методы и алгоритмы, основанные на обучении с подкреплением, демонстрируют более высокую скорость работы при сохранении приемлемого качества результатов, что делает их перспективными для применения в крупных сложных сетях.

ЗАКЛЮЧЕНИЕ

В ходе выполнения магистерской работы были изучены модели сложных сетей, алгоритмы максимизации влияния и модели распространения информации. Проведен анализ жадных, эвристических и обучаемых методов решения задачи максимизации влияния.

Результатом работы является программная система для моделирования процессов распространения информации и исследования алгоритмов максимизации влияния в сложных сетях. Программа реализована на языке программирования C# с использованием технологии Windows Forms, а для реализации методов машинного обучения использовался язык Python.

Разработанная система обеспечивает возможность генерации сложных сетей, запуска алгоритмов выбора влиятельных узлов, моделирования процессов распространения информации и визуального анализа результатов. Проведенные эксперименты показали эффективность реализованных алгоритмов и возможность их применения для исследования процессов распространения информации в социальных и коммуникационных сетях.

Основные источники информации:

- 1 Евин, И.А. Введение в теорию сложных сетей [Текст] // Компьютерные исследования и моделирование. — 2010. — Т. 2, № 2. — С. 121–141
- 2 Newman, MARK EJ. Networks: an introduction [Текст]. — 2010
- 3 Kempe, D. Maximizing the spread of influence through a social network [Текст] / Kempe, D., Kleinberg, J. и Tardos, E. // Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and datamining. — Washington : ACM, 2003. — С. 137–146.
- 4 Schoenebeck, G. Limitations of Greed: Influence Maximization in Undirected Networks Re-visited [Текст] / Schoenebeck, G., Tao, B. и Yu, F.-Y. // CoRR. — 2020. — Т. abs/2002.11679.
- 5 FastCELF++: A Novel and Fast Heuristic for Influence Maximization in Complex Networks [Текст] / do Carmo, V.E., Vieira, V.F., Oliveira, R.S. и Xavier, C.R. // International Conference on Computational Science and Its Applications. — Athens : Springer Cham, 2023. — С. 453–470.
- 6 A survey on meta-heuristic algorithms for the influence maximization problem in the social networks [Текст] / Aghaee, Z., Ghasemi, M.M., Beni, H.A., Bouyer, A. и Fatemi, A. // Computing. — 2021. — Т. 103. — С. 2437–2477.
- 7 Banerjee, S. A survey on influence maximization in a social network [Текст] / Banerjee, S., Jenamani, M. и Pratihar, D. K. // Knowledge and Information Systems. — 2020. — Т. 62. — С. 3417–3455.