
Графовые модели, контент и динамические процессы

Анализ социальных сетей
Лекция 1

Михаил Пожидаев

Институт прикладной математики и компьютерных наук ТГУ

27 октября 2026 г.

Социальные сети как объект анализа

Люди важны, но связи между ними — не меньше

Анализ социальных сетей (Social Network Analysis, SNA) — совокупность методов исследования структуры отношений между участниками некоторой социальной системы.

- В качестве участников могут выступать люди, аккаунты, организации, команды, публикации, хештеги или другие сущности.
- В качестве отношений рассматриваются дружба, подписка, переписка, упоминание, совместная работа, репост, финансовая операция и другие взаимодействия.

Основная идея

Поведение и возможности участника определяются не только его свойствами, но и контентом, который он публикует, а также его положением среди других участников.

Структура связей может объяснять распространение информации, влияние, изоляцию, конфликты и формирование сообществ.

От анкет к графам

Чем SNA отличается от традиционного анализа данных

Традиционный анализ:

- объектом исследования часто является отдельный респондент, клиент или пользователь;
- признаки объекта описывают возраст, доход, интересы, покупки или ответы анкеты;
- наблюдения обычно рассматриваются как независимые.

Сетевой анализ:

- объектом исследования становится система отношений;
- связь между двумя участниками является самостоятельным наблюдением;
- зависимость между участниками рассматривается как содержательное свойство данных;
- выводы строятся на основе структуры графа.

Пример

Два сотрудника могут иметь одинаковые должности и опыт работы. Но один из них связан с несколькими подразделениями, а другой общается только внутри собственной команды. Их возможности влиять на организацию и получать новую информацию будут различаться.

Графовая модель

Узлы, рёбра и математическое представление

Граф — это пара $G = (V, E)$, где V — множество вершин, а E — множество рёбер между вершинами. В SNA вершины обычно соответствуют акторам, а рёбра — отношениям или взаимодействиям между ними.

Вершины:

- пользователи социальной платформы;
- сотрудники компании;
- организации и города;
- сообщения, документы и хештеги.

Рёбра:

- дружба и подписка;
- звонки и сообщения;
- упоминания и репосты;
- участие в одном проекте;
- совместное посещение события.

Типы связей

Направление, вес и временная привязка

Ненаправленный граф:

- связь симметрична;
- если $\{u, v\} \in E$, то u связан с v , а v связан с u ; афы
- пример: взаимная дружба или совместная работа над проектом.

Направленный граф:

- связь имеет источник и получателя;
- если $(u, v) \in E$, то u направляет действие к v ;
- пример: подписка, сообщение, упоминание или ретвит.

Взвешенный граф:

- Каждой связи может соответствовать вес w_{uv} .
- Весом может быть число сообщений, количество репостов, длительность общения, денежный объём или оценка силы связи.

Сообщества в графе

Группы с плотными внутренними связями

Сообщество — группа вершин, связи внутри которой плотнее или значимее, чем связи с остальной сетью. Сообщество не обязано совпадать с формальной группой, географией, профессией или самоидентификацией участников.

Примеры:

- тематические группы вокруг интереса;
- команды внутри организации;
- политические аудитории;
- группы клиентов с общими покупками;
- кластеры совместной работы.

Задачи:

- сегментация аудитории;
- поиск информационных пузырей;
- выявление скрытой структуры организации;
- анализ конкурирующих нарративов;
- выбор каналов коммуникации.

Эго-сети

Анализ окружения отдельного участника

Эго-сеть состоит из центрального участника, называемого *ego*, его контактов *alter* и связей между этими контактами. Эго-сетевой подход полезен, если невозможно получить данные обо всей сети.

Можно измерять:

- размер окружения;
- разнообразие контактов;
- плотность связей между контактами;
- число мостов в другие сообщества;
- изменение окружения во времени.

Практические сценарии:

- сеть клиента вокруг продукта;
- профессиональные контакты кандидата;
- окружение лидера мнений;
- коммуникации сотрудника;
- связи автора в научной сети.

Динамика сетей

Связи появляются, исчезают и меняют смысл

Недостаток!

Статический граф отражает сеть только в выбранный момент или за выбранный период.

Динамический анализ сети изучает изменение вершин, рёбер, сообществ и центральностей во времени.

Подход временных срезов:

- сеть строится отдельно для каждой недели, месяца или квартала;
- метрики сравниваются между периодами;
- подход прост для интерпретации.

Временные рёбра:

- каждой связи назначается временная метка;
- анализируется порядок событий;
- учитывается длительность и давность контакта.

Распространение информации

Каскады, заражение и влияние структуры

Диффузия — процесс распространения информации, поведения, инновации, вирусного контента или слуха по сети.

Independent Cascade:

- активированная вершина получает одну попытку активировать каждого соседа;
- вероятность передачи по ребру задаётся заранее;
- процесс развивается дискретными шагами.

Linear Threshold:

- вершина активируется, если суммарное влияние активных соседей превышает её порог;
- модель подходит для социального доказательства;
- пороги могут различаться у разных участников.

Нельзя делать сильный вывод

Наблюдение последовательности репостов не доказывает причинное влияние одного пользователя на другого. На поведение могут одновременно влиять внешние новости, алгоритмы рекомендаций и общие интересы.

Контент и структура

SNA особенно полезен вместе с анализом текстов

- Связи отвечают на вопрос «кто с кем взаимодействует».
- Контент отвечает на вопрос «о чём и в каком тоне взаимодействуют».

Методы анализа контента:

- тематическое моделирование;
- извлечение ключевых слов;
- анализ тональности;
- распознавание именованных сущностей;
- кластеризация текстовых эмбедингов.

Графовые представления:

- автор — автор через упоминание;
- автор — хештег;
- автор — тема;
- публикация — ссылка;
- сообщество — нарратив.

Спасибо за внимание!

Домашняя страница:

<https://marigostra.ru>

Электронная почта:

mvp@luwrain.org

Telegram:

@Marigostra

