



Сетевые метрики

Николай Скворцов
nsv@mail.ru

План лекции

- ▶ Представление сетей
- ▶ Основные сетевые метрики

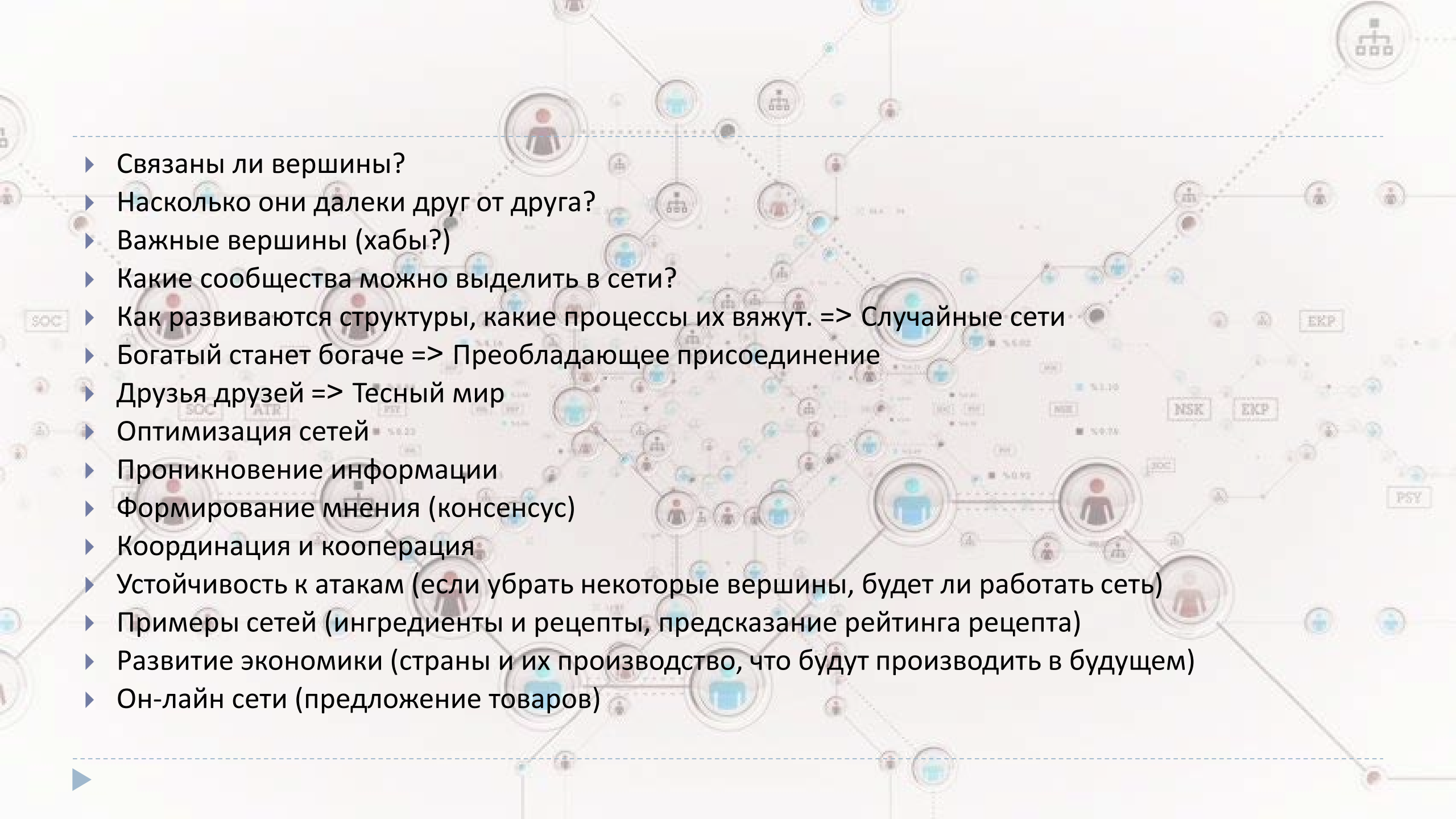


Составляющие сетей

- ▶ Участники (персоны, организации, сообщества)
- ▶ Ресурсы (книги, видео, фотографии)
- ▶ Атрибуты (профили, анкеты, текст комментариев)
- ▶ Отношения (друзья, оценки, метки, посещения, публикации, цитирования, комментарии, транзакции, взаимодействия, роли)

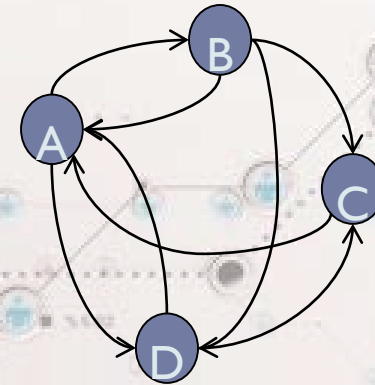
Рассматриваемые классы задач над сетями

- ▶ Участники и ресурсы сети не равнозначны из-за различий в рассматриваемых связях между ними
 - ▶ Ранжирование участников по различным критериям
- ▶ Участники и ресурсы сети образуют связанные структуры
 - ▶ Обнаружение и анализ сообществ
- ▶ Социальные сети – живые системы
 - ▶ Моделирование развития сетей с разными свойствами
 - ▶ Устойчивость сети к воздействиям
 - ▶ Моделирование распространения событий различной природы
- ▶ Необходимость анализа больших сетей
 - ▶ Подходы к редуцированию и распараллеливанию вычислений над сетями

- 
- ▶ Связаны ли вершины?
 - ▶ Насколько они далеки друг от друга?
 - ▶ Важные вершины (хабы?)
 - ▶ Какие сообщества можно выделить в сети?
 - ▶ Как развиваются структуры, какие процессы их вяжут. => Случайные сети
 - ▶ Богатый станет богаче => Преобладающее присоединение
 - ▶ Друзья друзей => Тесный мир
 - ▶ Оптимизация сетей
 - ▶ Проникновение информации
 - ▶ Формирование мнения (консенсус)
 - ▶ Координация и кооперация
 - ▶ Устойчивость к атакам (если убрать некоторые вершины, будет ли работать сеть)
 - ▶ Примеры сетей (ингредиенты и рецепты, предсказание рейтинга рецепта)
 - ▶ Развитие экономики (страны и их производство, что будут производить в будущем)
 - ▶ Он-лайн сети (предложение товаров)

Представление сетей

- ▶ Граф $G = (V, E)$
 - ▶ V – Вершины (Участники, Ресурсы)
 - ▶ E – Рёбра (Отношения)
 - ▶ Ориентированные рёбра (дуги)
 - ▶ Атрибуты вершин и рёбер
 - ▶ Например, вес, тип, метрики структуры графа
- ▶ Матрица смежности M
 - ▶ Элементы множества V по горизонтали и вертикали
 - ▶ $M_{ij} = 1$, если между вершинами i и j есть ребро, иначе 0

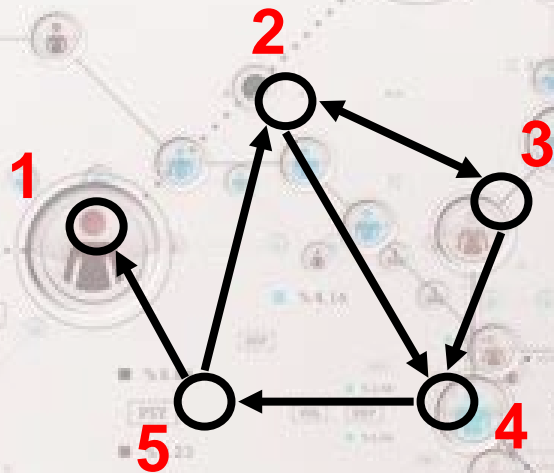


	A	B	C	D
A	0	1	0	1
B	1	0	1	1
C	1	0	0	0
D	1	0	1	0

Другие представления

▶ Список дуг

- ▶ 2, 3
- ▶ 2, 4
- ▶ 3, 2
- ▶ 3, 4
- ▶ 4, 5
- ▶ 5, 2
- ▶ 5, 1



▶ Список смежности

- ▶ 1:
- ▶ 2: 3 4
- ▶ 3: 2 4
- ▶ 4: 5
- ▶ 5: 1 2

Используемые термины

Область	Точка	Линия
Математика	Вершина vertex	Дуга, ребро edge, arc
Физика	Место site	Связь bond
Социология	Участник, актер, агент actor, agent	Связи, отношения tie, relation
Информатика	Узел node	Связь link
Мы	Вершины (элементы графов) Агенты (социальные)	Дуги, рёбра (элементы графов) Связи, деятельности (социальные)

Ориентированность связей

- ▶ Ориентированные

- ▶ $A \rightarrow B$

- ▶ Дуга, arc

- ▶ A нравится B, A звонит B, A - ребёнок B

- ▶ Неориентированные

- ▶ $A \leftrightarrow B$ или $A - B$

- ▶ Ребро

- ▶ A брат или сестра B, A соавтор статьи B

Атрибуты

- ▶ Атрибуты вершин

- ▶ Вес (приоритет)
- ▶ Возраст
- ▶ Свойства в графе (значение центральности)

- ▶ Атрибуты

- ▶ Вес (частота взаимодействия)
 - ▶ Ранг (лучший друг, позитивное/негативное отношение)
 - ▶ Тип (друг, коллега)
 - ▶ Свойства в графе (значение метрики посредничества)
-

Метрики элементов сетей

- ▶ Сеть G : n вершин, m рёбер
- ▶ Степень вершины
 - ▶ количество рёбер, соединяющих вершину с другими вершинами
 - ▶ Для ориентированных сетей

Входящая и исходящая полустепени вершины

$$\sum_{j=1}^n A_{1j}$$

- ▶ Плотность сети
 - ▶ $\rho = 2m / (n(n-1))$
 - ▶ Плотная/разреженная сеть
- ▶ Посредничество вершин или рёбер
 - ▶ количество кратчайших путей графа, проходящих через ребро
- ▶ Диаметр графа
 - ▶ наибольшее расстояние между двумя любыми его вершинами
- ▶ Средняя степень вершины
- ▶ Средняя длина пути

$$l_{av} \sim \frac{\log N}{\log z}$$

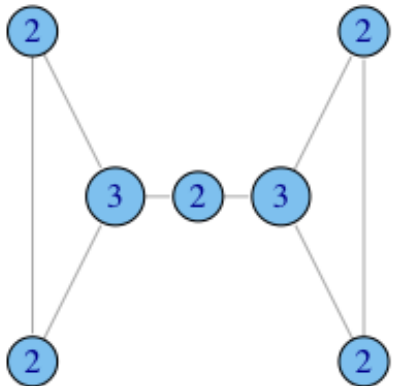
Центральность вершин сети

- ▶ Центральность вершин – ранжирование их важности по некоторому критерию, учитывающему структуру всего графа
 - ▶ по степеням вершин – по количеству соседей вершины
 - ▶ по посредничеству – по количеству минимальных путей, проходящих через вершину
 - ▶ по близости к центру – по средней длине минимальных путей от вершины к другим вершинам сети
 - ▶ по собственному вектору – с учётом центральности соседних вершин

Центральность ориентированных вершин

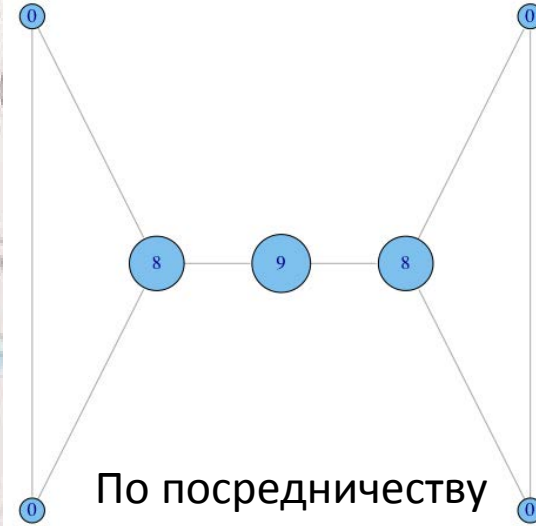
- ▶ Престиж вершины – метрика центральности вершины по входным дугам
- ▶ Влияние вершины – метрика центральности вершины по выходным дугам

Центральность вершин сети (2)



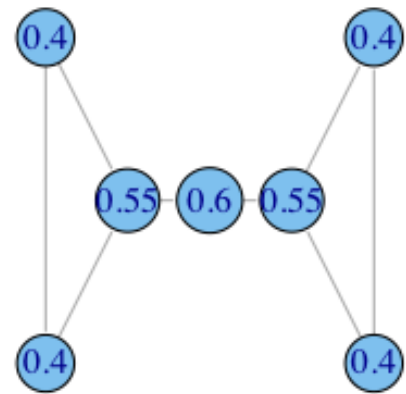
По степеням

$$C_D(v_i) = d_i = \sum_j A_{ij}$$



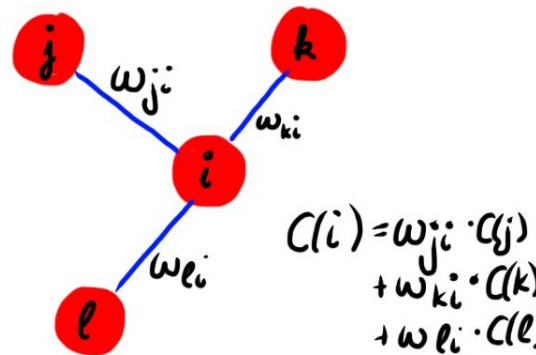
По посредничеству

$$C_B(v_i) = \sum_{v_s \neq v_i \neq v_t \in V, s < t} \frac{\sigma_{st}(v_i)}{\sigma_{st}}$$



По близости

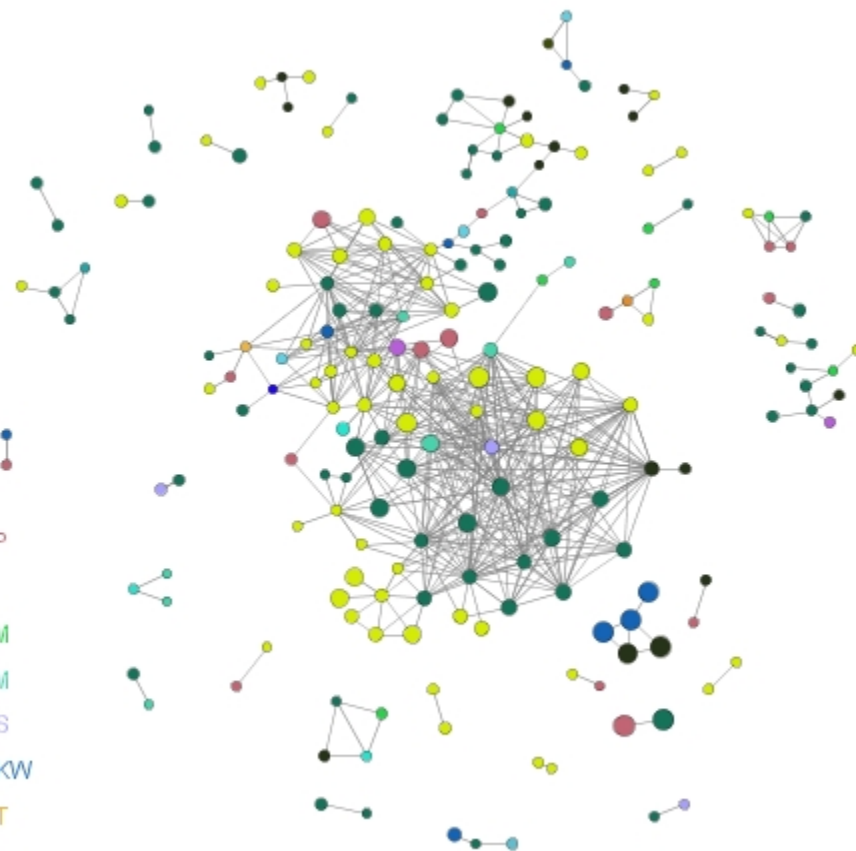
$$C_C(v_i) = \frac{n - 1}{\sum_{j \neq i} g(v_i, v_j)}$$



По собственным векторам

Компоненты сети

- SAM
- SEM
- SHQ
- SAP
- EEM
- EAM
- ENS
- UNKW
- CST



Распределение степеней сети

- ▶ Распределение степеней сети

- ▶ Случайный граф – биномиальный закон

- ▶ p – вероятность связи

- ▶ Реальные графы – экспоненциальный закон

- ▶ $P(k) \approx k^{-\alpha}$

- ▶ Кластерность

- ▶ Вероятность того, что две соседние вершины соединены

- ▶ Для случайного графа составляет:

$$C_{\text{срн}} = p = \frac{\langle k \rangle}{N}$$