Сетевые модели

Николай Скворцов
nskv@mail.ru

План лекции

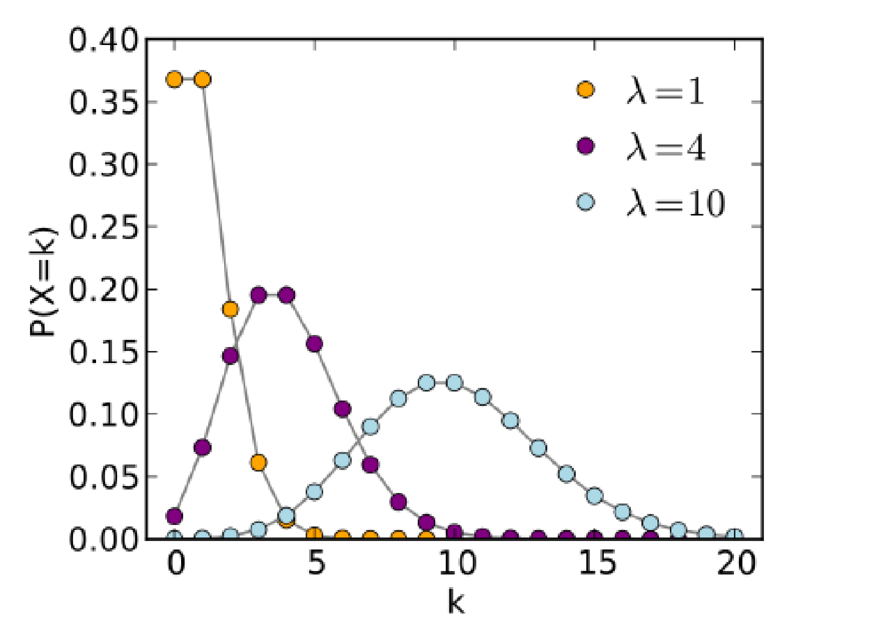
- Моделирование сетей
- Случайный граф
- Модель тесного мира
- Модель предпочтительного присоединения

Прошлая лекция

- Вершины большой степени связываются с вершинами большой степени
  - Сети с предпочтительным соединением
- От распределения степеней зависят характеристики и поведение сети

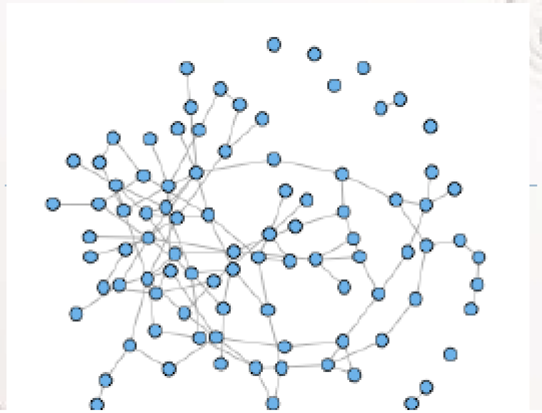
Модель случайной сети Эрдёш-Реньи

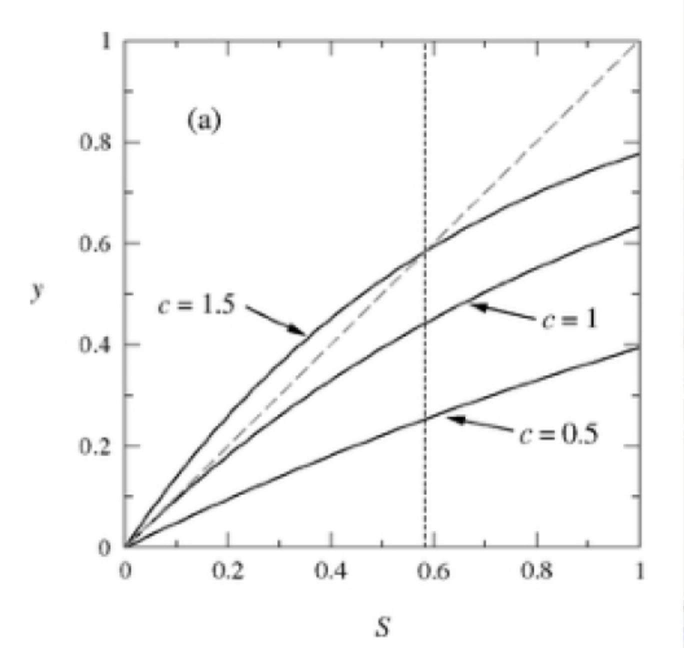
- $G(V,E)$ - случайная сеть из всех возможных сетей с заданным количеством вершин и рёбер
  - $n = |V|$, $\mathrm{m} = |E|$
- $G_{n,p}$, p – вероятность ребра между парой
- Среднее количество рёбер $p\frac{n(n-1)}{2}$
- Средняя степень вершины $\frac{\sum k}{n} = \frac{2m}{n} = p(n-1)$
- Функция распределения степеней вершин
  - $P(k) = C_{n-1}^{k} p^{k} (1-p)^{n-1-k}$ – распределение Бернулли
    - C_{n-1}^{k} – количество способов выбрать k вершин
  - При $n \to \infty$ стремится к распределению Пуассона $\frac{\lambda^{k} e^{-\lambda}}{k!}$
    - где $\lambda = pn$ – средняя степень вершин

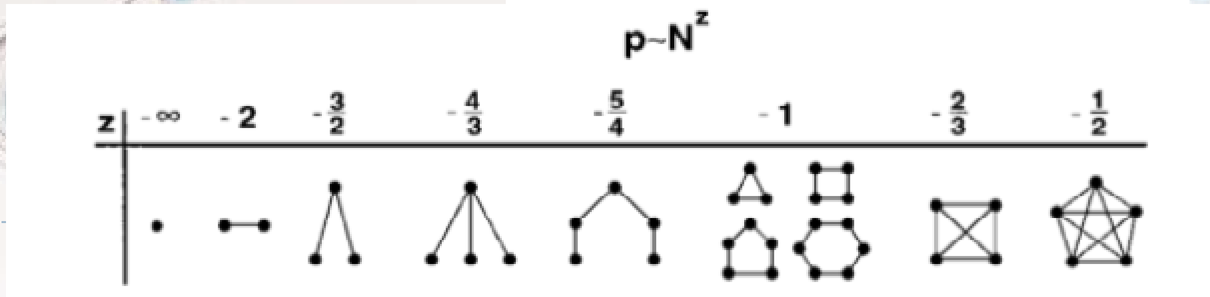


Параметры сети

- Гигантский компонент
  - При предельной вероятности p_c возникает один связный компонент из большинства вершин
  - Доля узлов в гигантском компоненте $s = 1 - e^{-\lambda s}$
  - $\lambda_c = p_c n = 1$
  - $p_c = \frac{1}{n}$
  - Среднее количество соседей у вершины больше 1
- Диаметр графа при $p > p_c, d = \frac{\ln n}{\ln \lambda}$
- Кластерный коэффициент C(k)= p
- Структуры также предсказуемы

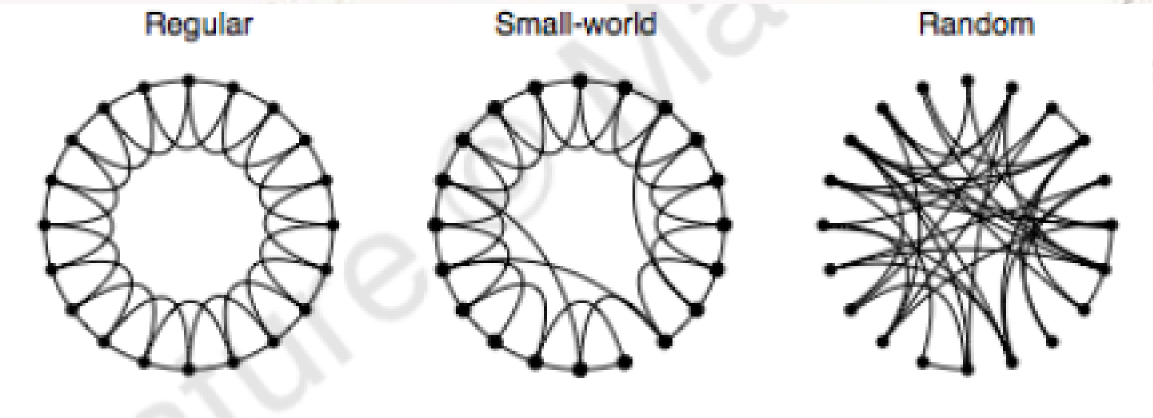




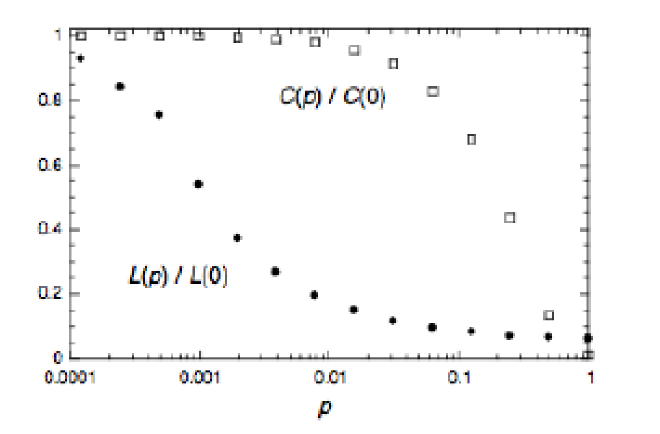


Модель тесного мира

- Эксперимент Милгрема
- Модель Ватца-Строгаца 1998
  - Изучали сети на окружности с соединёнными близкими вершинами
  - Соединения через один делает высоким кластерный коэффициент
  - Хордовые связи, резко понижающие диаметр
  - p – доля рёбер, которые пересоединяются из кольца в хорду, не меняя среднюю степень



- Между регулярным графом и случайным находится сеть малого мира
  - Кластерный коэффициент ещё высок
  - Средняя длина пути уже низкая

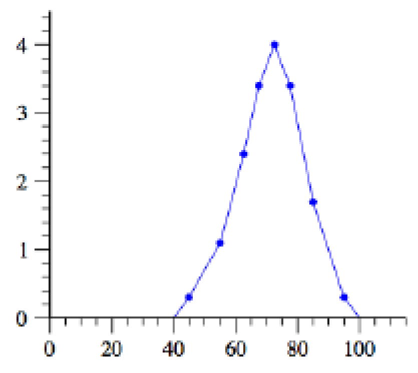


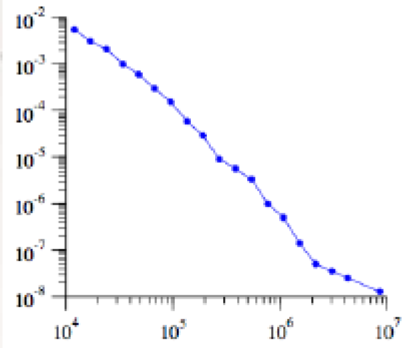
Сети с предпочтительным присоединением

- Агенты предпочитают быть соседями тех, у кого много соседей
- Реальные комплексные сети обычно именно таковы
- Экспоненциальный закон распределения степеней вершин
  - Есть много вершин с редкими соседями
  - И немного – с большим количеством соседей
- Вначале есть n0 вершин
  - Вероятность новой вершины присоединиться к пропорциональна степени старой вершины
  - Модель Барабаши-Альберта

Распределение величин

- Функция плотности распределения непрерывной случайной величины
  - $P(x \in [a,b]) = \int_a^b f(x)dx\,, f(x) \geq 0, \int f(x)dx = 1$
- Функция плотности распределения дискретной случайной величины
  - $p(x), p(x) \geq 0, \sum p(x) = 1$
- Кумулятивная функция распределения
  - Вероятность, что величина меньше заданной
- Распределение Гаусса
  - Среднее значение и стандартное отклонение
  - Рост человека
  - Скорости машин
- Степенное распределение
  - Население городов
  - Количество цитирований статей
  - Имение людей
  - $1/x^\alpha$
  - Используются логарифмические координаты



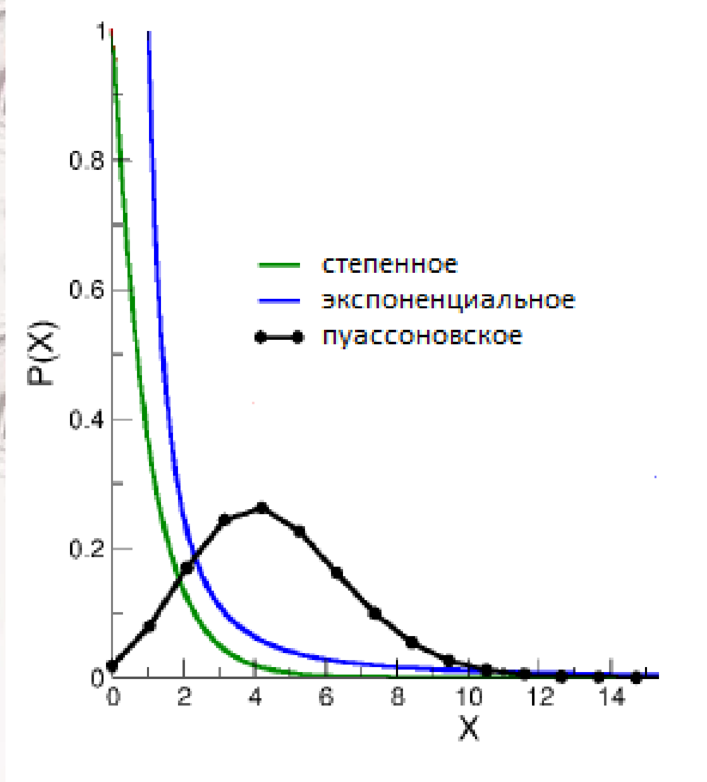


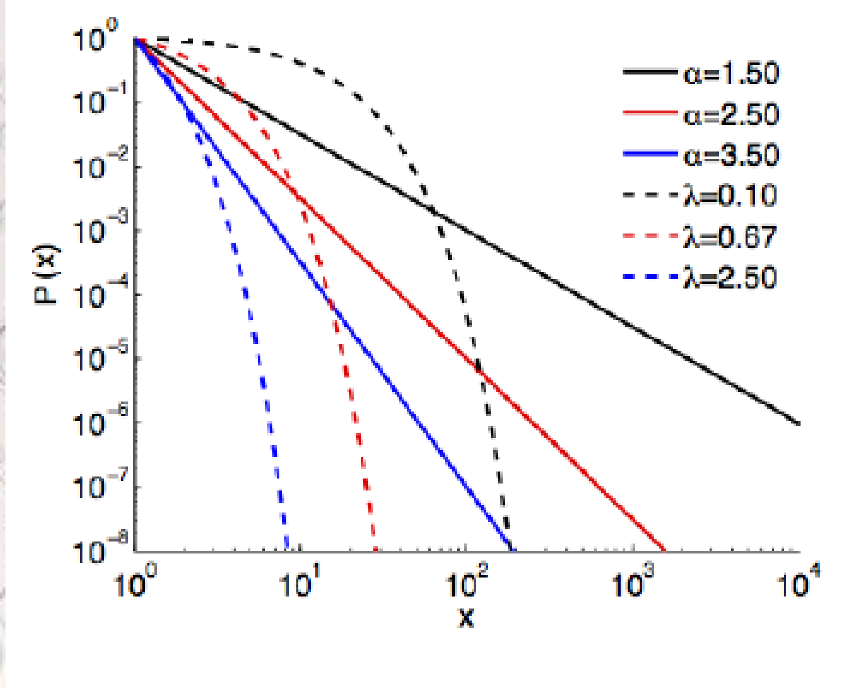
Степенной закон распределения

- Распределение
  - $p(x) = C \cdot \frac{1}{x^{\alpha}}, \alpha > 1$
- Нормализация
  - $C \cdot \int_{x_{min}}^{\infty} \frac{1}{x^{\alpha}} dx = 1$
  - $C = (\alpha - 1) \cdot x_{min}^{\alpha - 1}$
- Среднее
  - $\langle x \rangle = \int x p(x) dx = C \int \frac{dx}{x^{\alpha-1}} = \frac{\alpha-1}{\alpha-2} x_{min}, \alpha > 2$
- Стандартное отклонение
  - $\langle x^2 \rangle = \int x^2 p(x) dx = C \int \frac{dx}{x^{\alpha-2}} = \frac{\alpha-1}{\alpha-3} x_{min}^2, \alpha > 3$
- Лучше использовать медиану
  - $\langle x^2 \rangle = \int x^2 p(x) dx = C \int \frac{dx}{x^{\alpha-2}} = \frac{\alpha-1}{\alpha-3} x_{min}^2, \alpha > 3$

Сравнение распределений

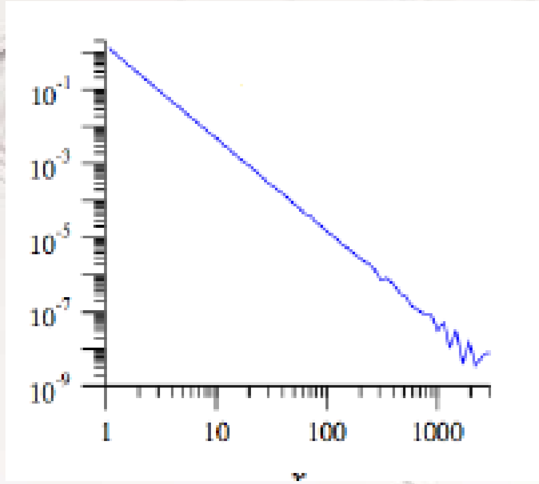
- $p(x) = C \cdot \frac{1}{x^{\alpha}}$
- $\log p(x) = \log C - \alpha \cdot \log x$





Безмасштабные сети

- В степенном законе при изменении масштаба пропорции сохраняются
  - $p(Kx) \sim p(x)$
  - Отношение количества людей, имеющих достаток 150 тыс., к количеству людей, имеющих достаток 50 тыс., такое же, как отношение количества людей, имеющих достаток 30 тыс., к количеству людей, имеющих достаток 10 тыс.
- Вывод
  - Исследовать отношения в можно на меньших масштабах



Структура сетей с предпочтительным присоединением

- Большое количество вершин малой степени
- Присутствие вершин очень большой степени
  - Хабы (актёры, цитируемость)
  - Ассортативность по степеням вершин
- Малый диаметр сети
  - Тесный мир
- Устойчивость к случайным воздействиям, чувствительность к атакам на хабы
- Свойства не зависят от размера сети
- Поиск: направленный в сторону вершин с максимальной степенью