



О том, как Computer Science нам
жить помогает или современные
приложения теории графов

Калачёв Максим Александрович

Разработчик

kalachevmax@gmail.com

Cloud

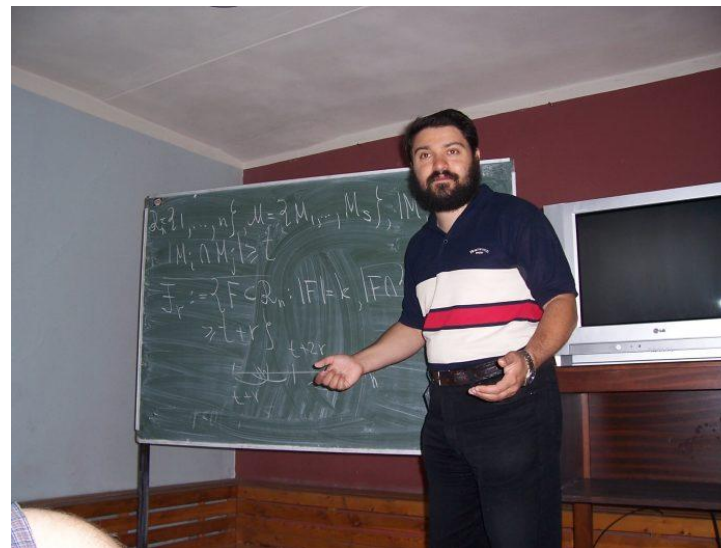
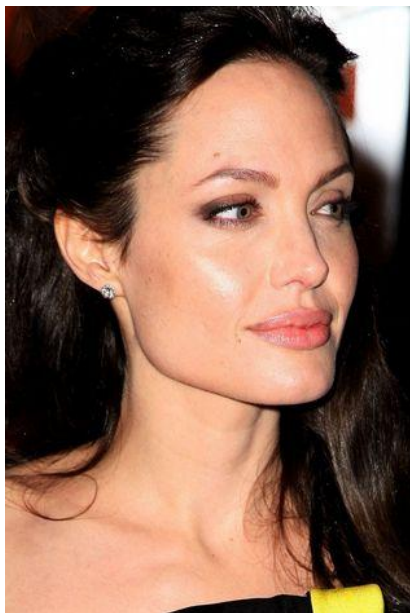
Software

Services

Agenda

- веб-графы
- методы моделирования
- ранжирование
- неестественные структуры
- shortest path problem
- нерешённые проблемы

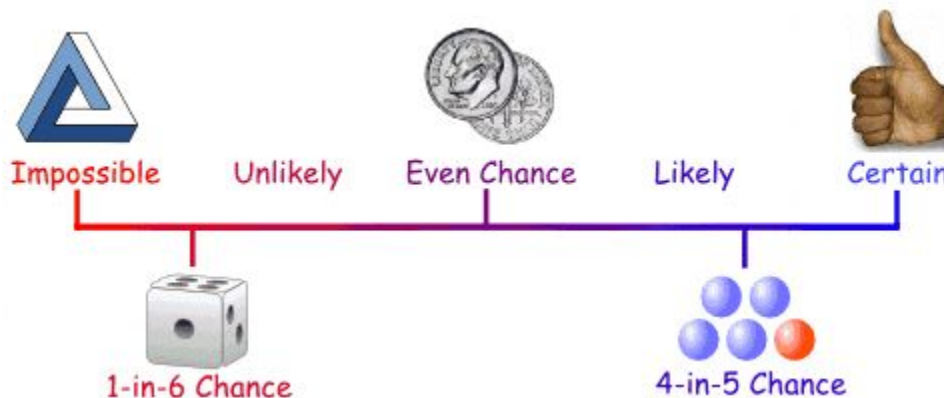
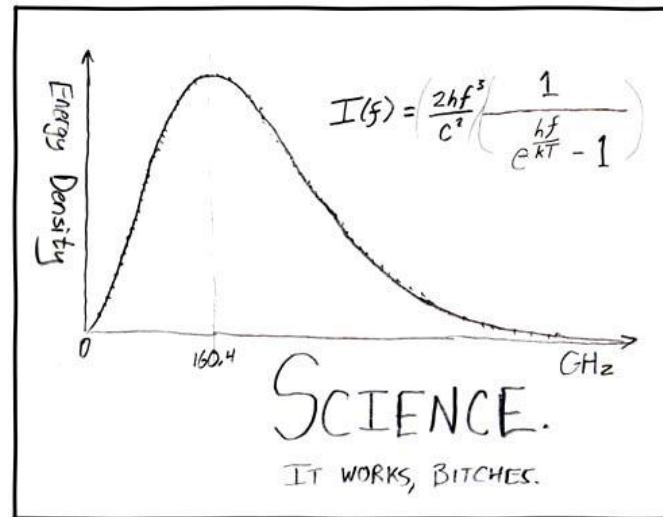
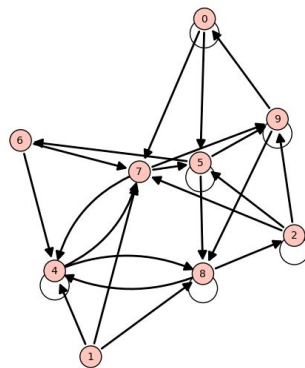
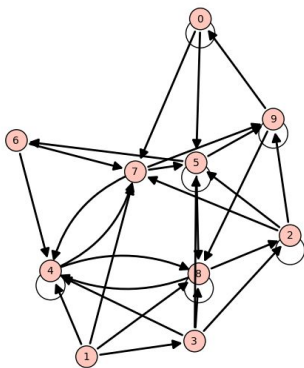
Метафизический вопрос №1



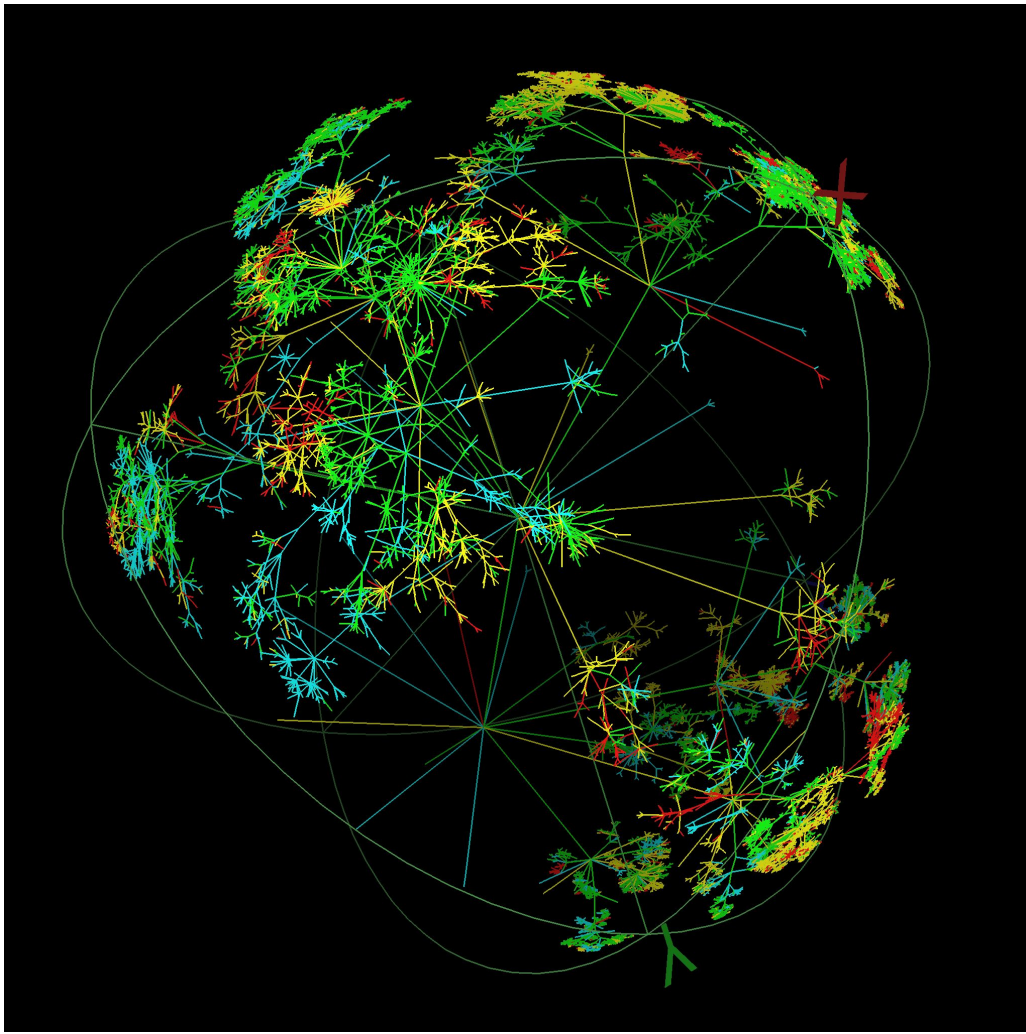
Метафизический вопрос №2



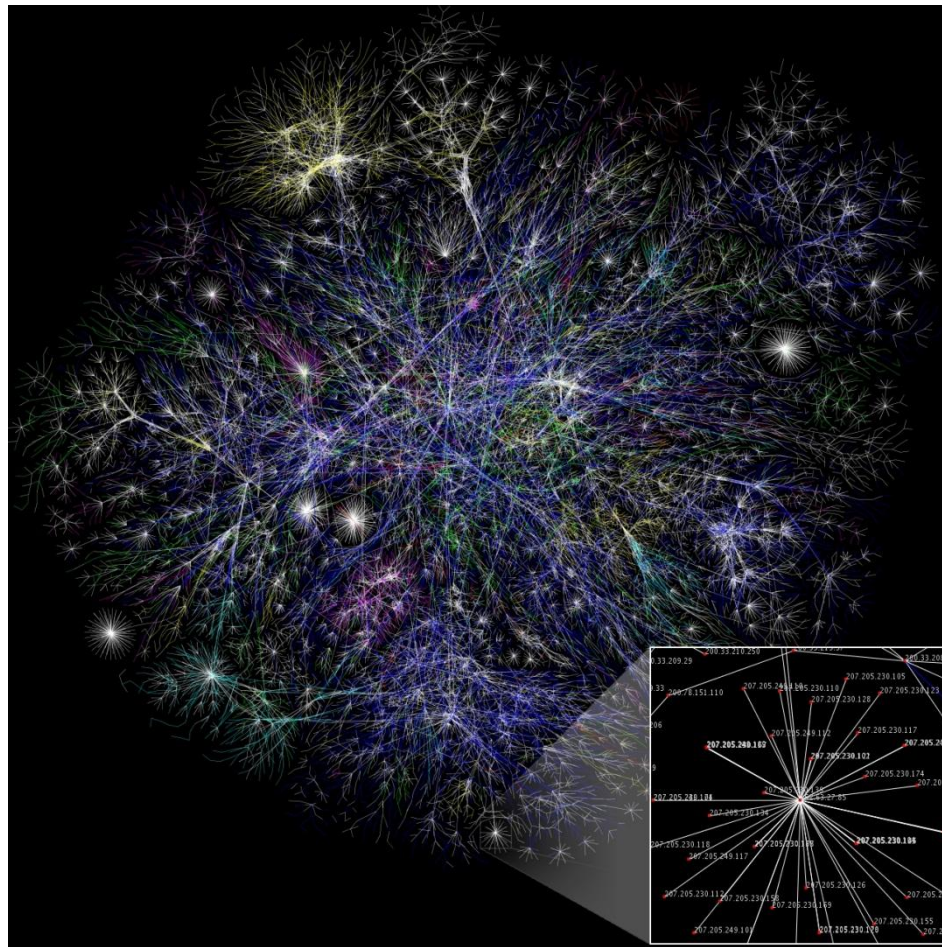
Графы, вероятность, приложения



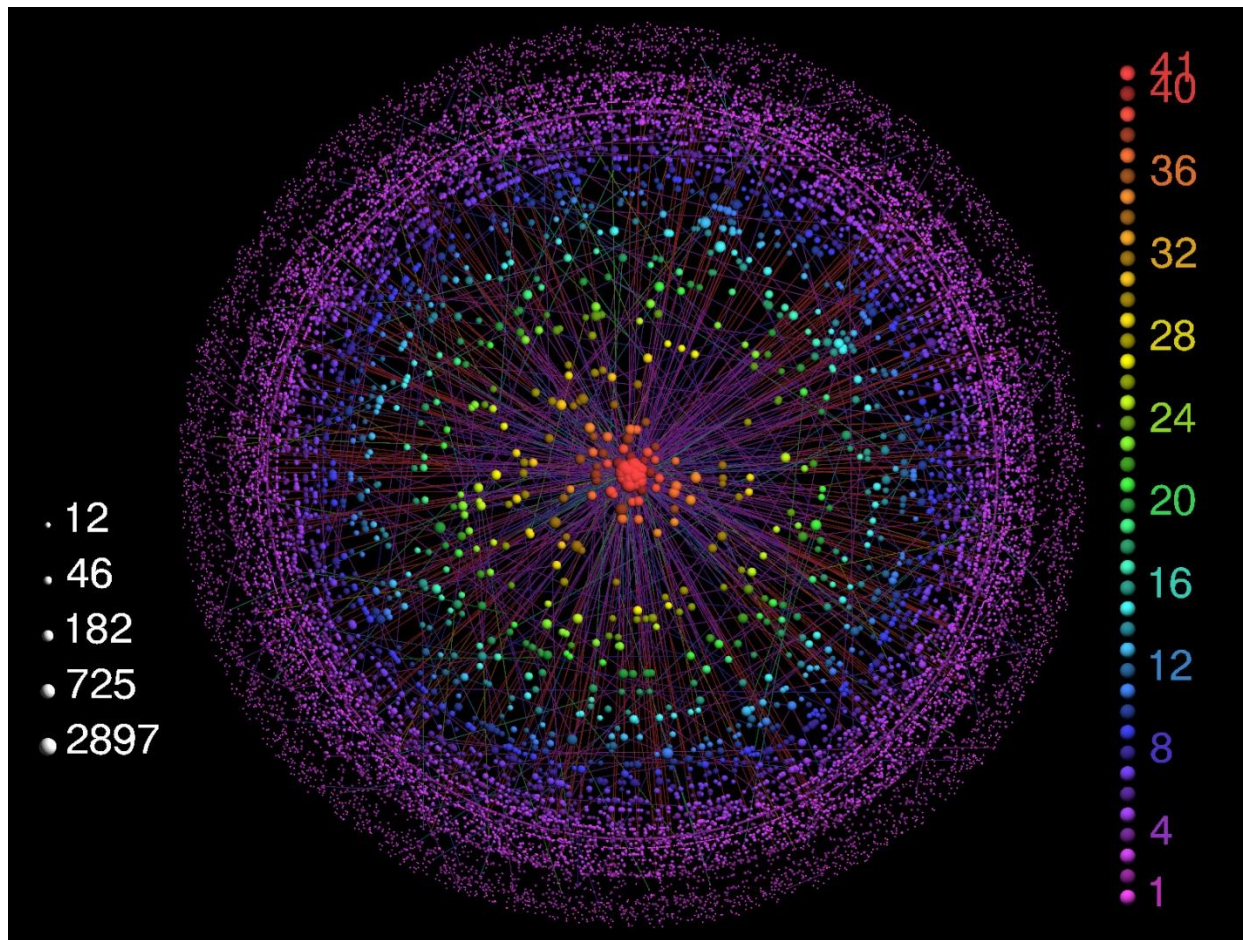
Веб-графы



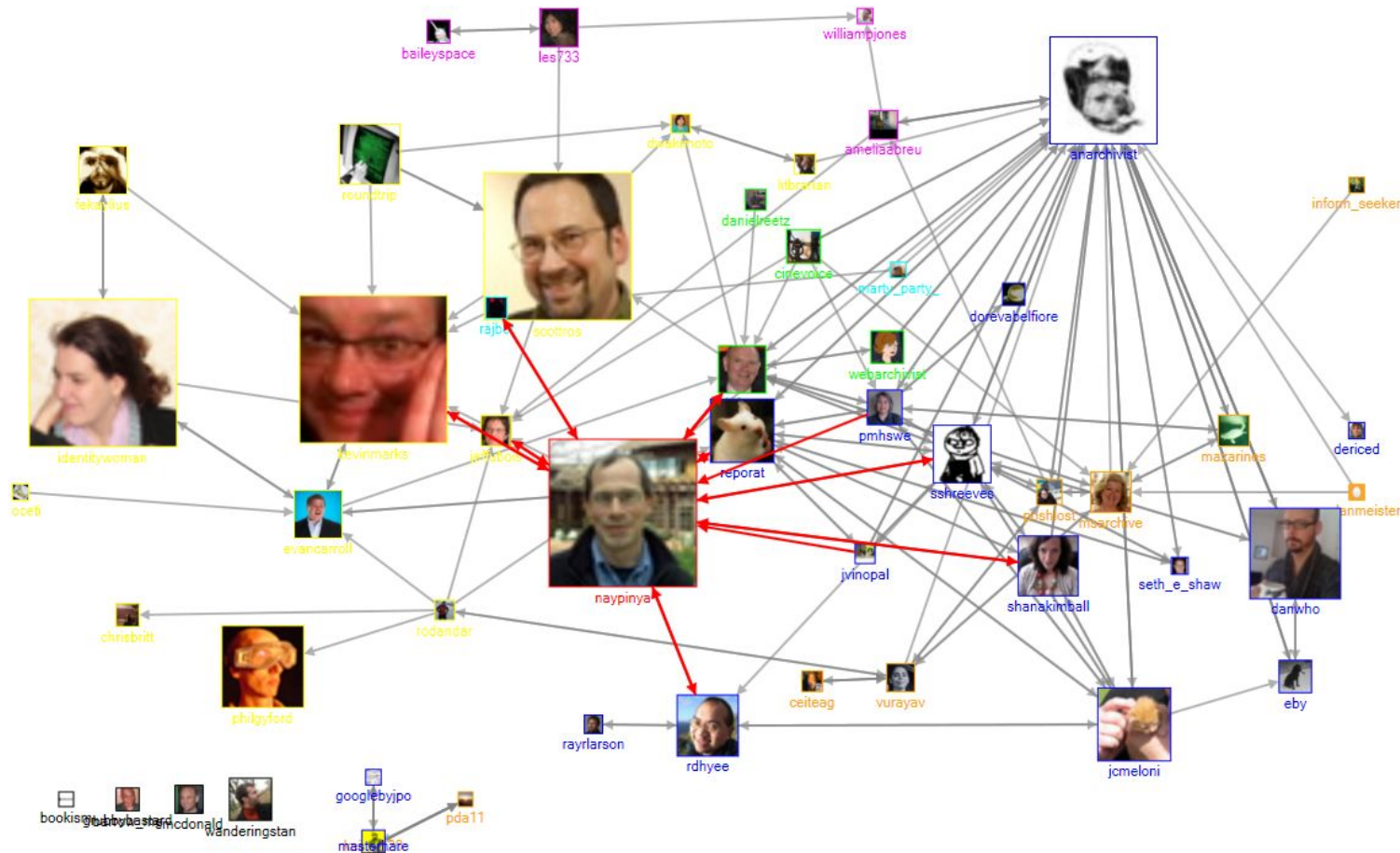
Веб-графы



Веб-графы



Социальные сети

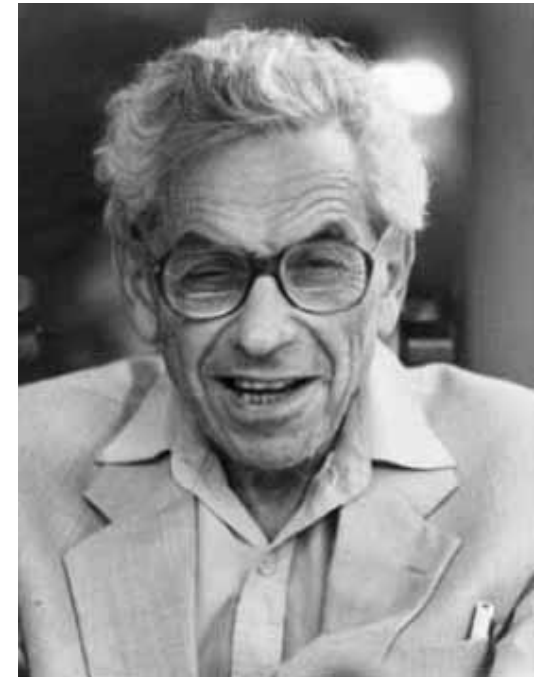
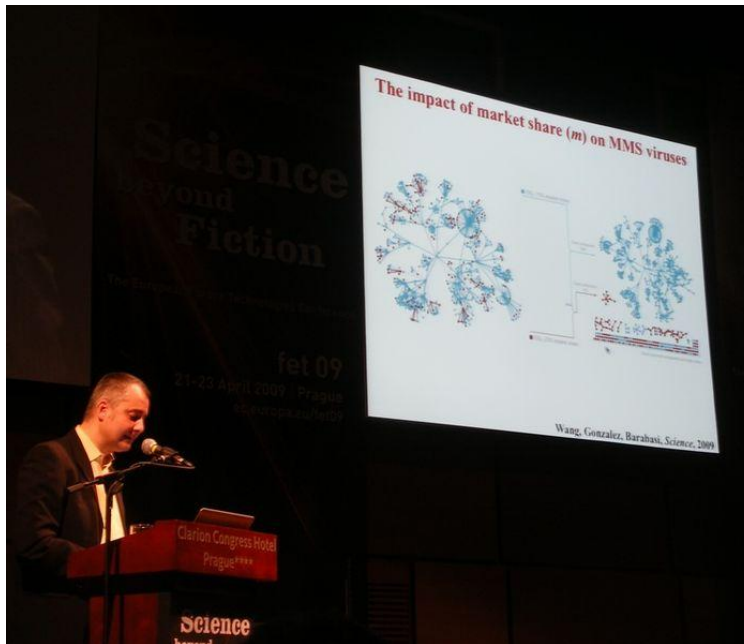


Социальные сети



Моделирование веб-графов

- Случайные графы
- Исследования Varabasi-Albert
- Модель Vollobas-Riordan
- Модификации модели



Как устроен веб-граф?

Albert-Laszlo Barabasi and Reka Albert. Emergence of scaling in random networks. Science, 286:509, 1999.

5 млрд вершин, псевдомультиторграф

Ключевые свойства веб-графа:

- Разрежённость

на k вершин k^2 рёбер, $k \geq 1$

- Диаметр графа $\in \{5, 6\}$

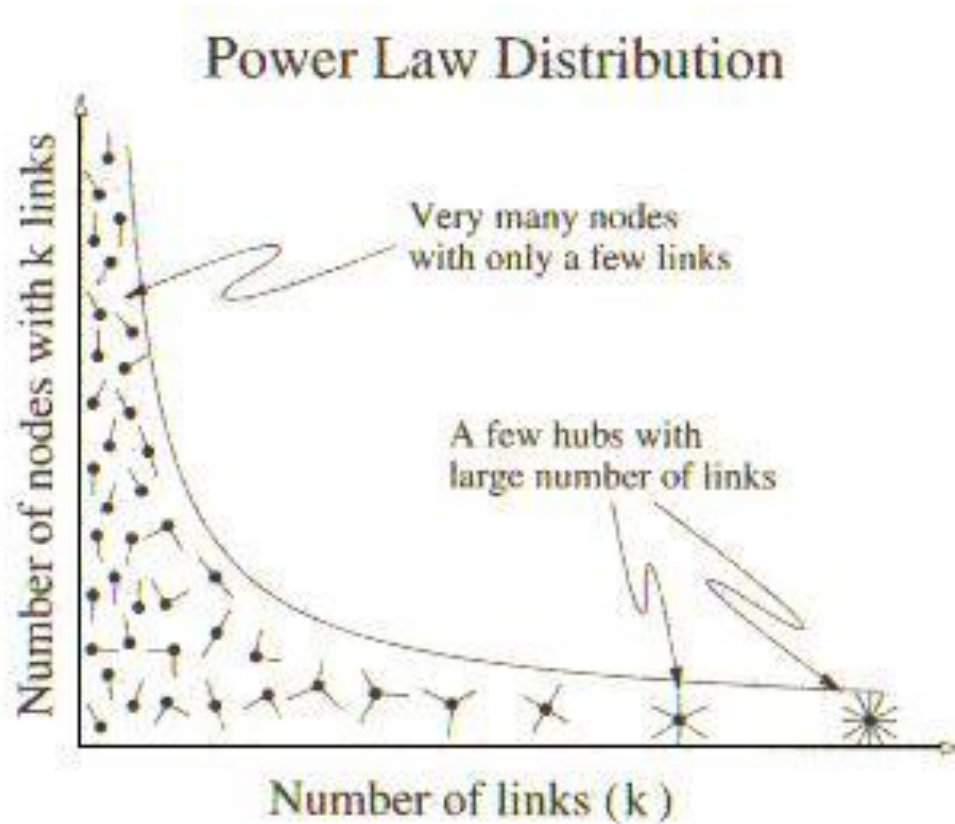
Теория о шести рукопожатиях

- Степенное распределение степеней вершин

$$P(d) \sim c / d^\lambda$$

$\lambda \approx 2.1$, c – нормирующий множитель

Степенной закон распределения

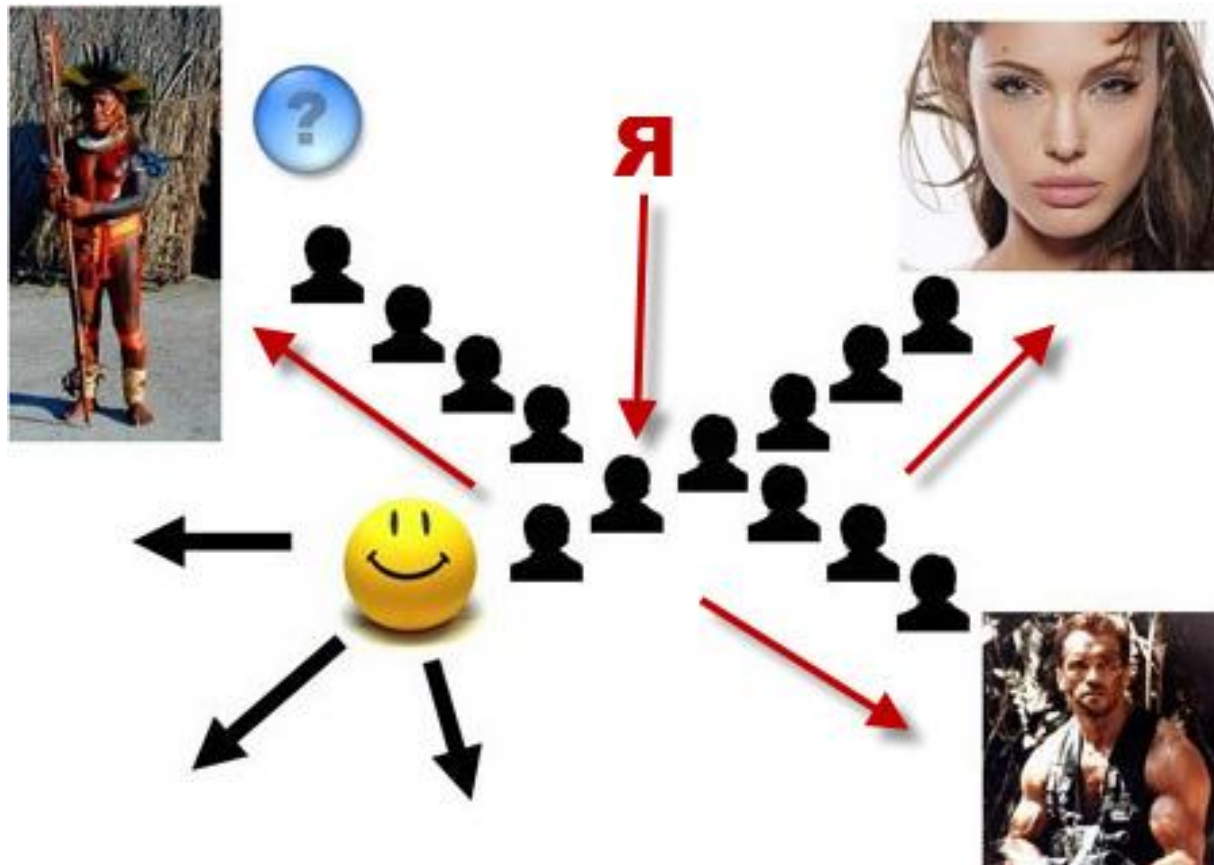


Эволюция веб-графа

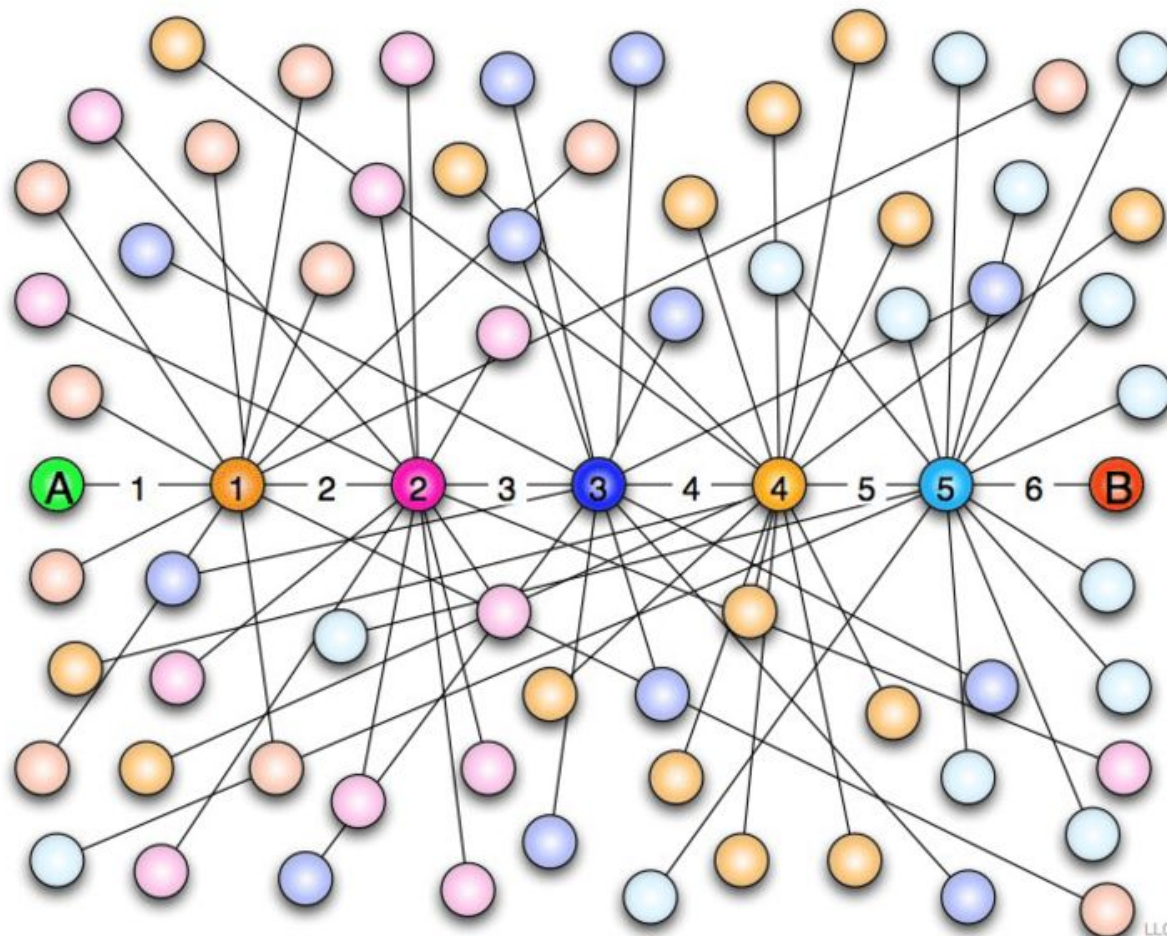
Модель предпочтительного соединения (preferential attachment)



Six degrees of separations



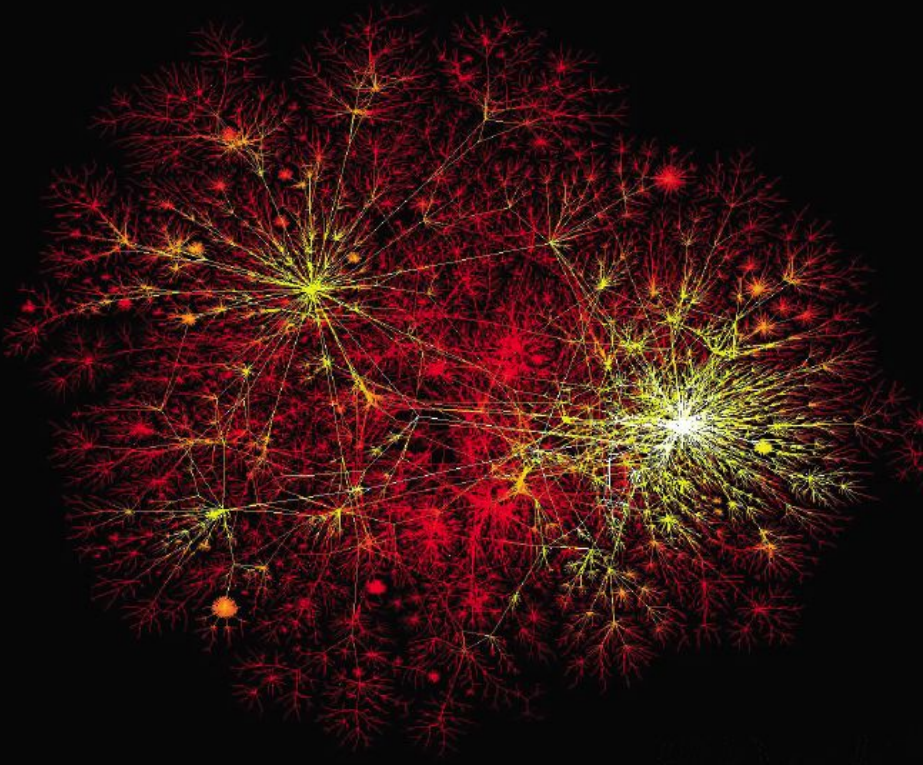
Six degrees of separations



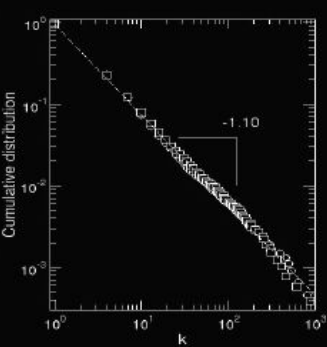
LL07

Масштабная инвариантность

Complex networks are **scale-free**

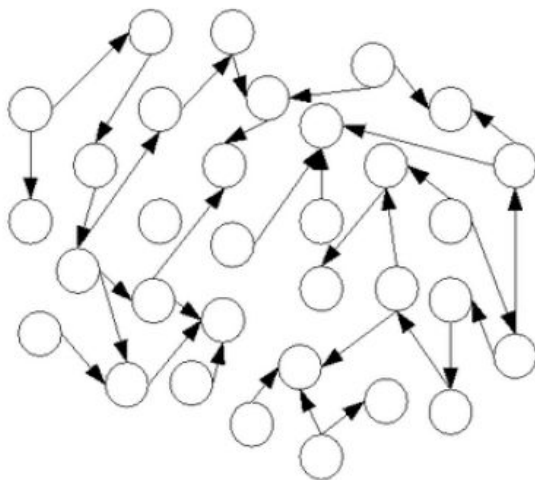


$P(k) \sim k^{-\gamma} \phi(k/\xi)$

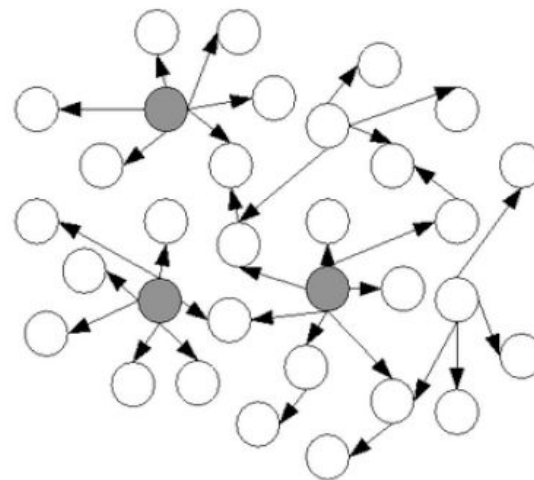


Scale-free networks

- Техника: Сети электропередачи, VLSI, Интернет, Веб
- Социум: контакты, связи, организации, язык, дороги, авиамаршруты
- Биология: нейроны; пищевые, экологические, метаболические сети
- Физика: молекулы, галактики

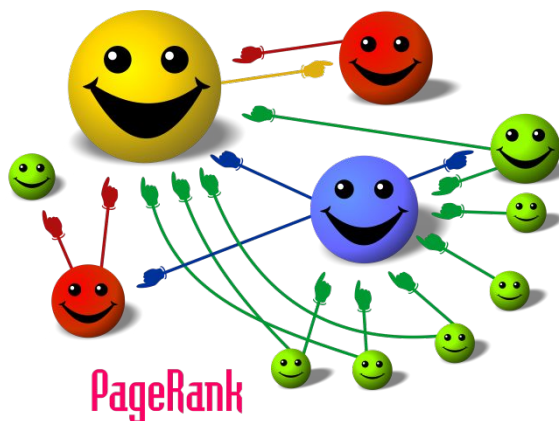


(a) Random network

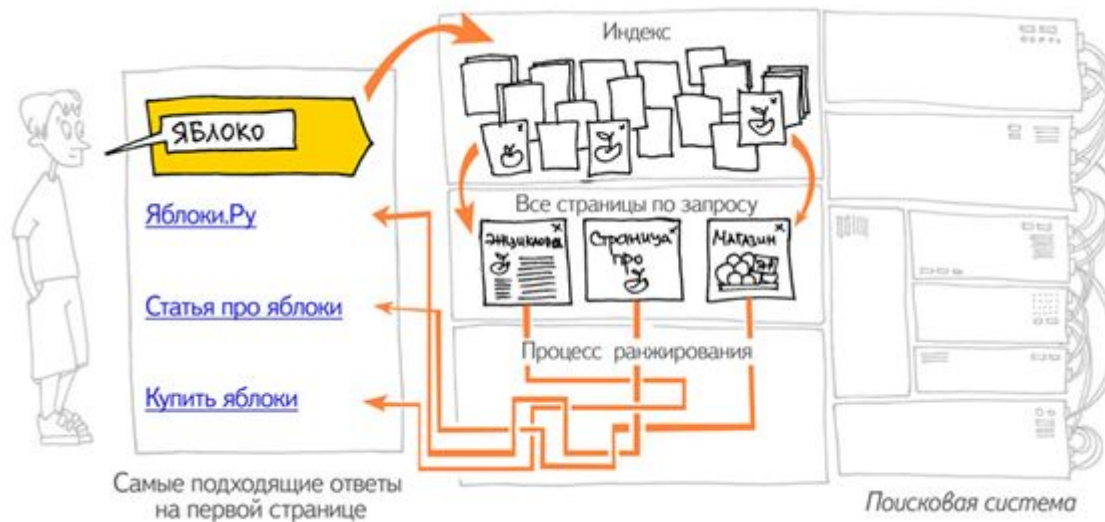


(b) Scale-free network

Ранжирование в поисковых системах

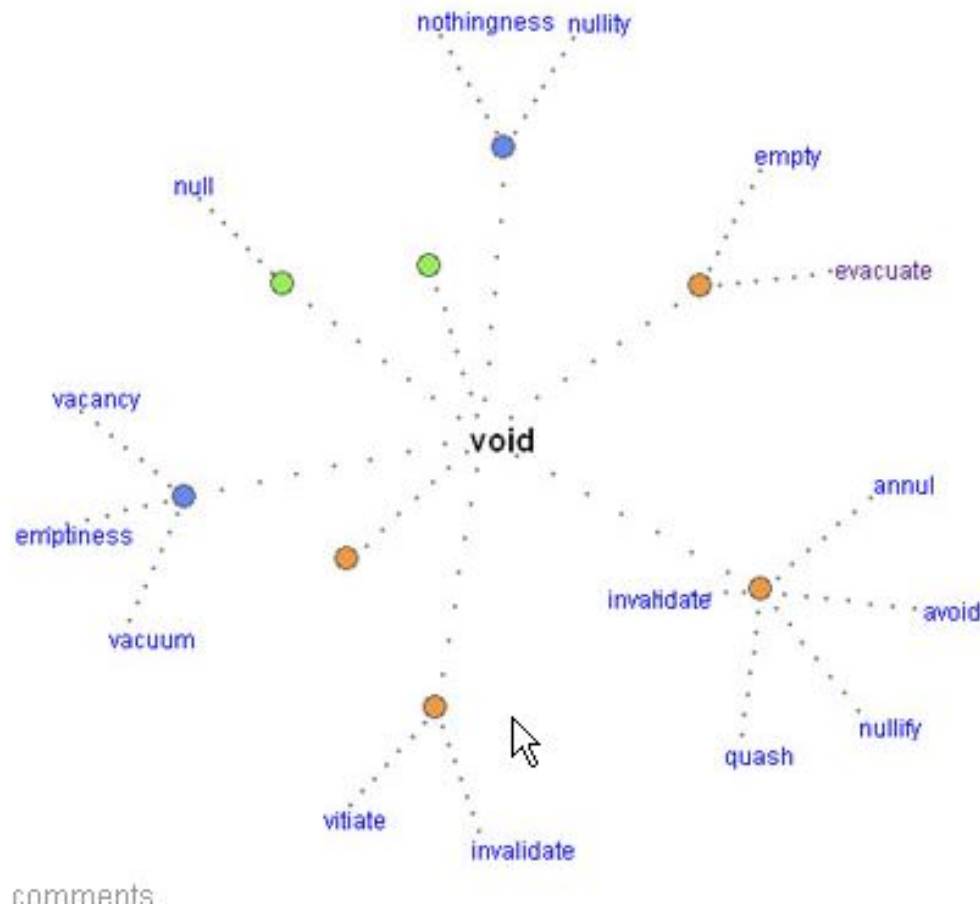


$$S(V_i) = (1 - d) + d * \sum_{j \in In(V_i)} \frac{S(V_j)}{|Out(V_j)|}$$

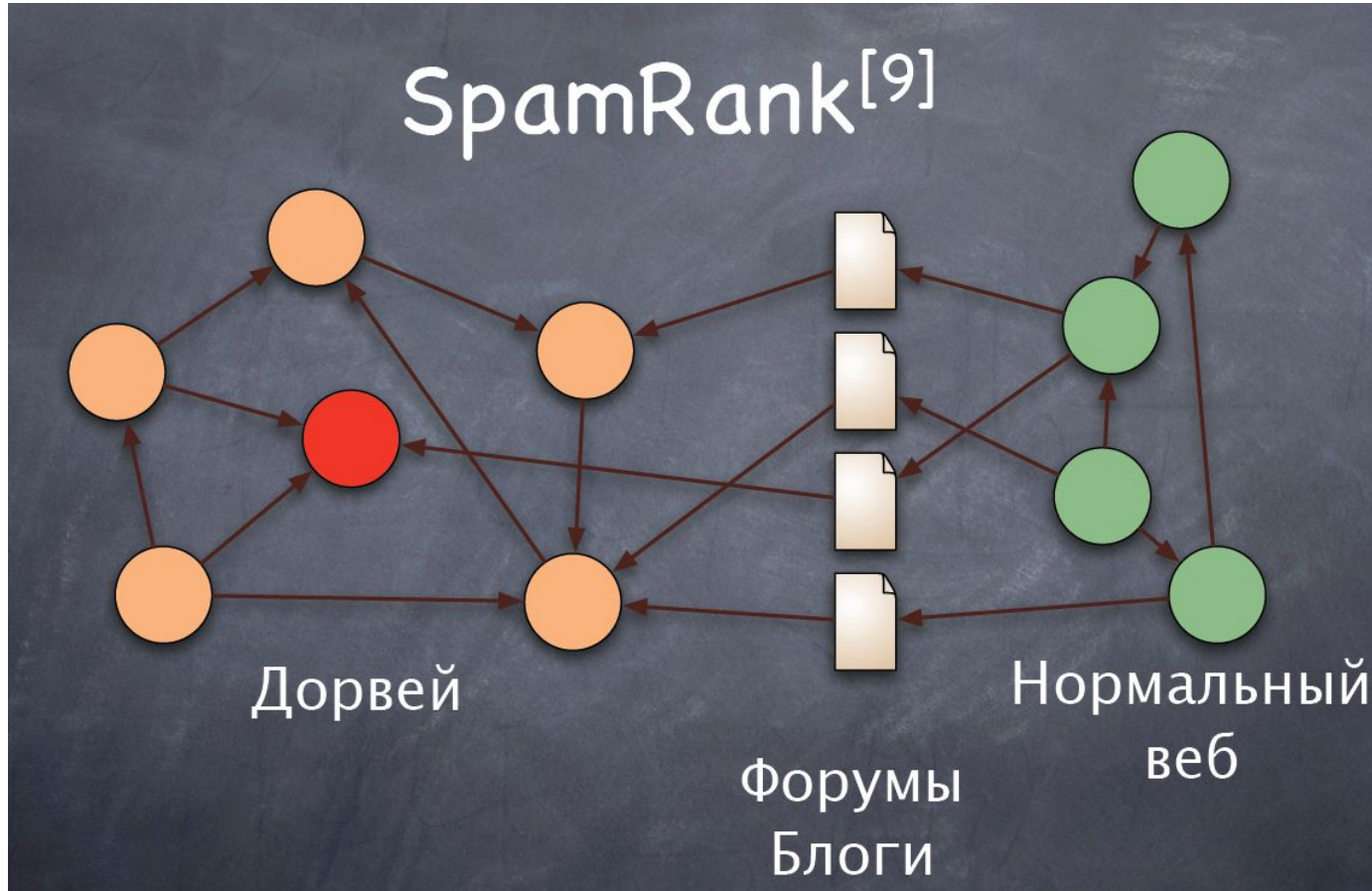


Ранжирование в семантических сетях

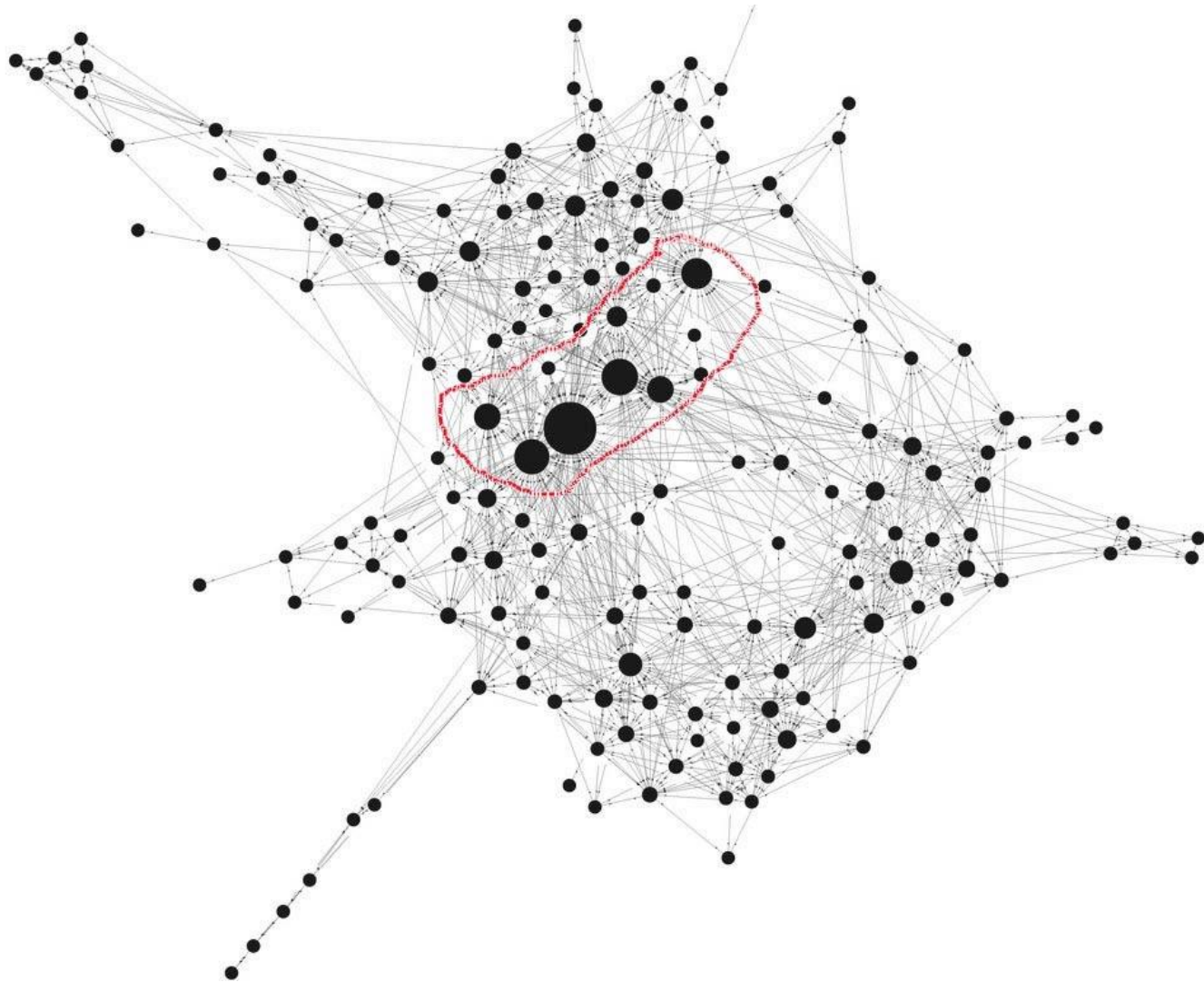
проект WordNet (wordnet.princeton.edu)



Выявление веб-структур

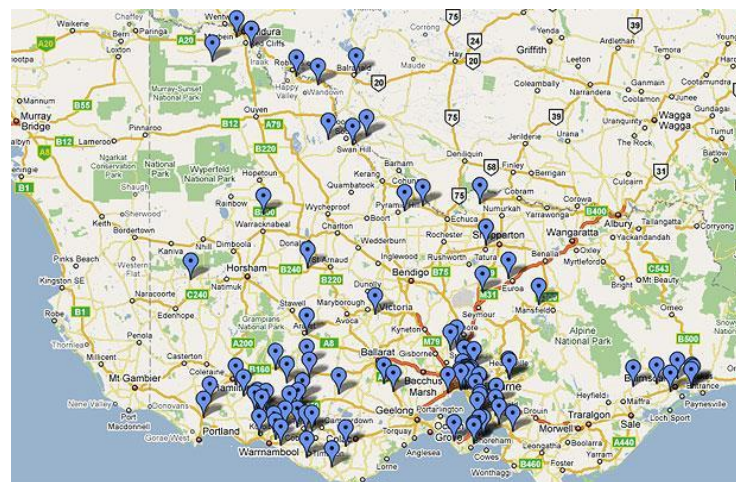
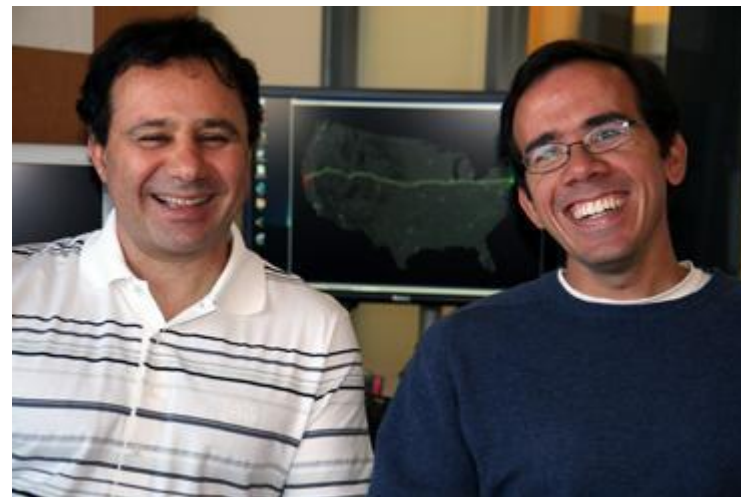


Выявление веб-структур



Shortest path problem

Andrew Goldberg
Microsoft Research



Shortest path problem

Почему современные алгоритмы на картах работают очень быстро

100000 млн вершин
Время работы 10^{-2} с

Интуитивные идеи:

- Указатели на дугах
- Поиск A*
- Достижимость
- Шоссейная и желаемые иерархии
- Перевалочные пункты



Нерешённые вопросы

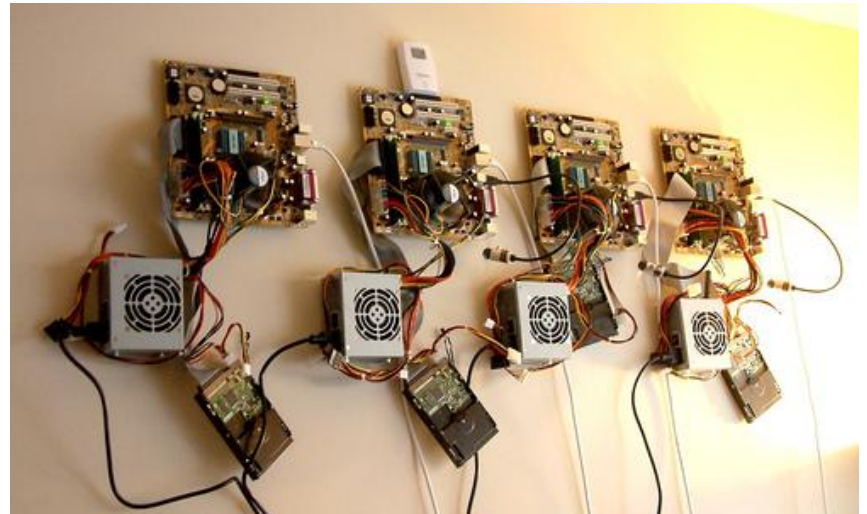
Самое главное, что ученик должен узнать от учителя - это что некоторый вопрос ещё не решён.

Петровский И.Г.

brainware



hardware

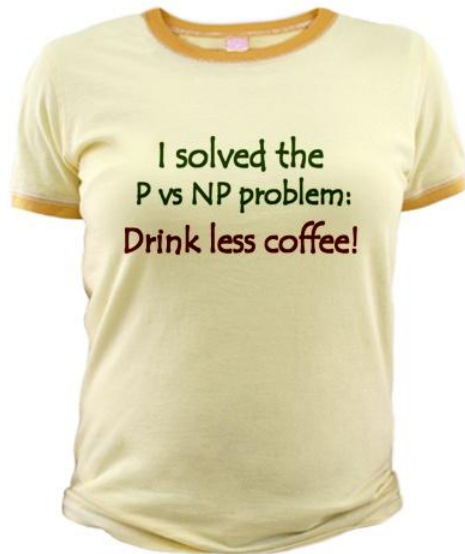


P vs NP

NP – класс всех задач поиска, решение для которых может быть быстро проверено.

P – класс задач поиска, решение для которых может быть быстро найдено.

$P \neq NP$ – верно ли, что каждый раз, когда решение можно быстро проверить, его можно быстро найти.



Послесловие

Я.Б. Зельдович считал, что постановка задачи – искусство куда более тонкое, чем решение. “Стоит точно сформулировать вопрос, - говорил он, - как тотчас найдётся подходящий математик для решения. Ведь математики, они как мухи, - умеют ходить по потолку!”

В.И. Арнольд, Задачи Арнольда.