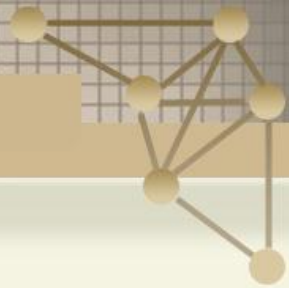




**Выявление пересекающихся
сообществ
в социальных сетях**

Коршунов Антон, Бузун Назар ИСП РАН



Содержание

- Определение и примеры
- Применение
- Требования к методам
- Классы методов
- Fuzzy обобщение
- Тестирование

Примеры сообществ

В контакте

- Лучшие друзья
- Родственники
- Коллеги
- Друзья по вузу
- Друзья по школе

Создать список



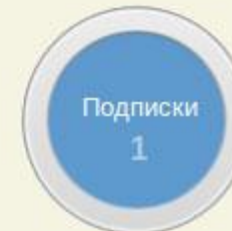
Друзья
8



Семья
0



Знакомые
7



Подписки
1

twitter

- Followers >
- Favorites >
- Lists >
- Recent images >



Name: **Social Computing**
Type: Organizations
Members: 14 members



Name: **Social Computing**
Type: Internet & Technology
Members: 12 members



Name: **Social Computing Magazine**
Type: Internet & Technology
Members: 34 members



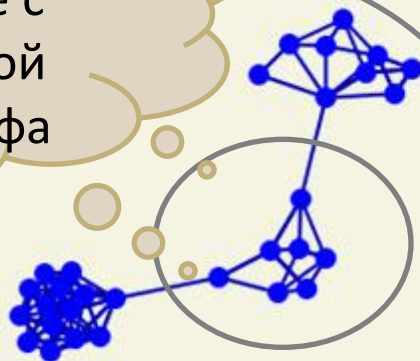
Name: **Trustworthy Social Computing**
Type: Internet & Technology
Members: 28 members

Определение сообщества

Высокая степень

Расхождение с
вероятностной
моделью графа

Каждая компонента
является сообществом

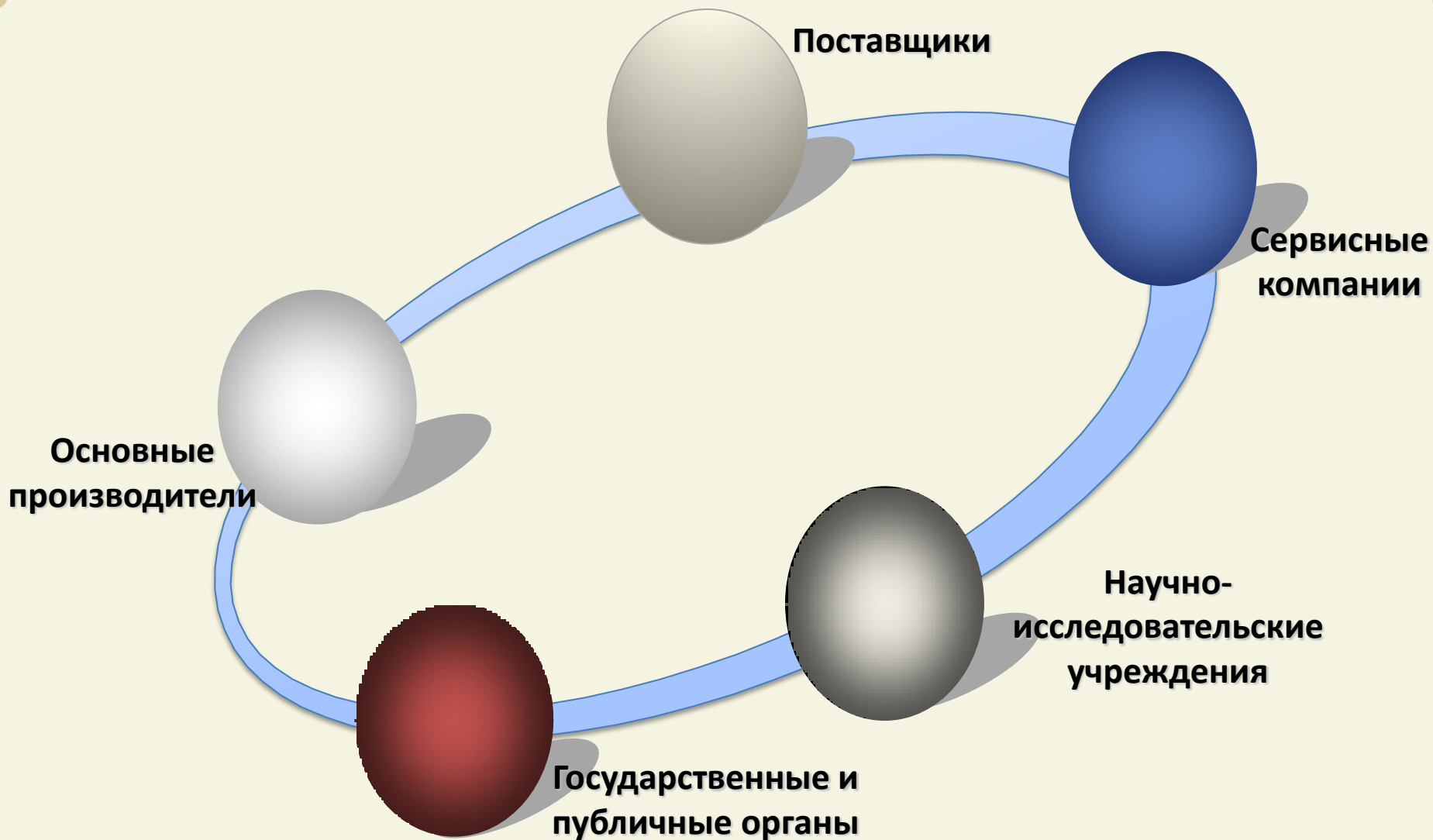




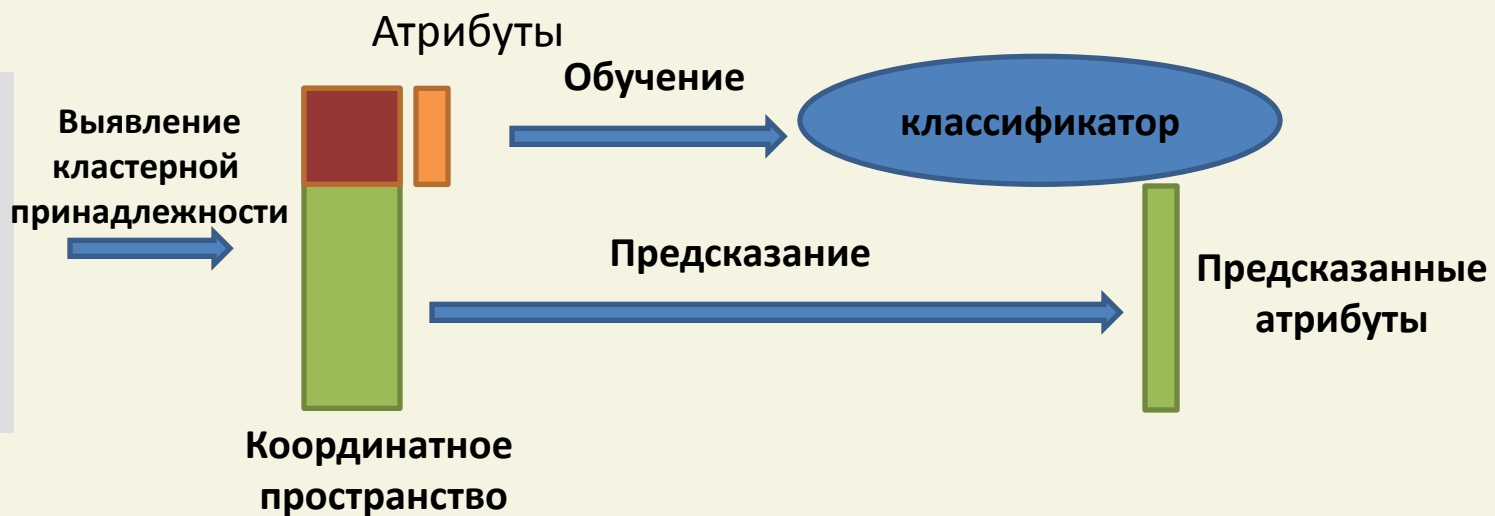
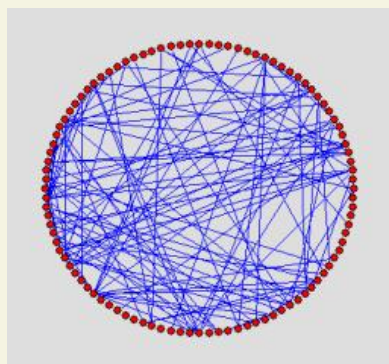
Области применения

- Наглядное представление
- Классификация
- Рекомендации
- Предсказание
- Хранение и расположение данных
- Сжатие данных
- Выявление атипичных элементов
- Распространение информации

Инновационная инфраструктура



Предсказание атрибутов профиля





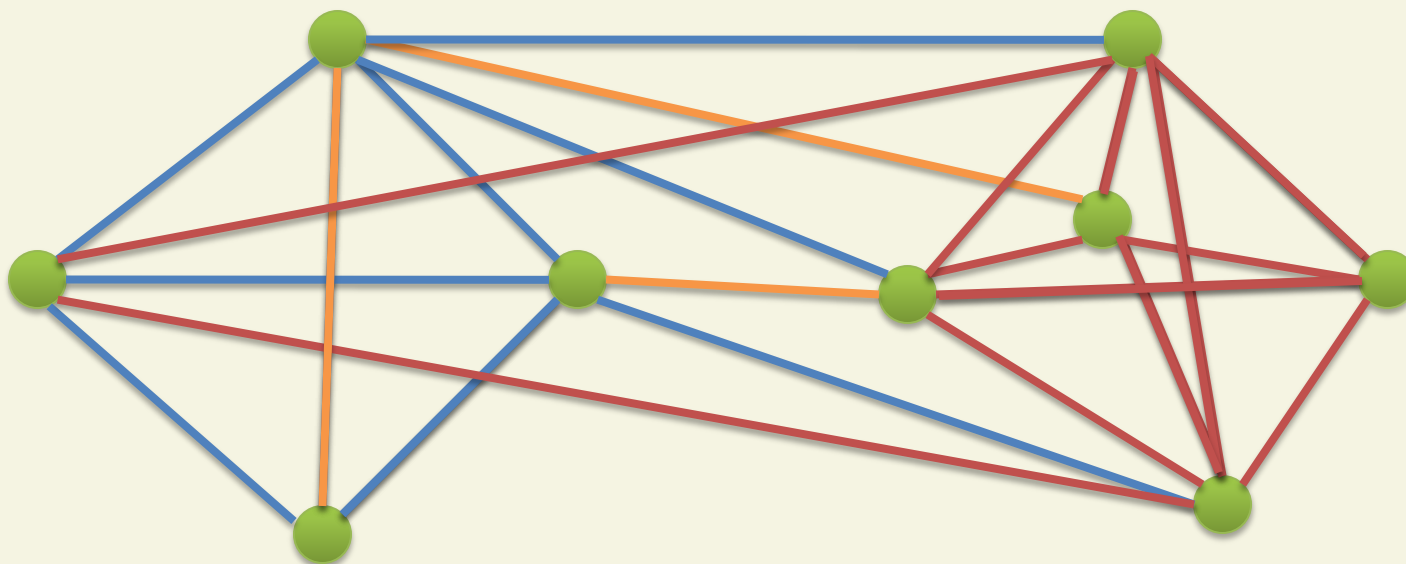
Требования

- Пересекающиеся сообщества
- Иерархическая кластеризация
- Возможность параллельного запуска
- Статистическая значимость

Устойчивость при росте степени
пересечения сообществ

Нулевая модель графа

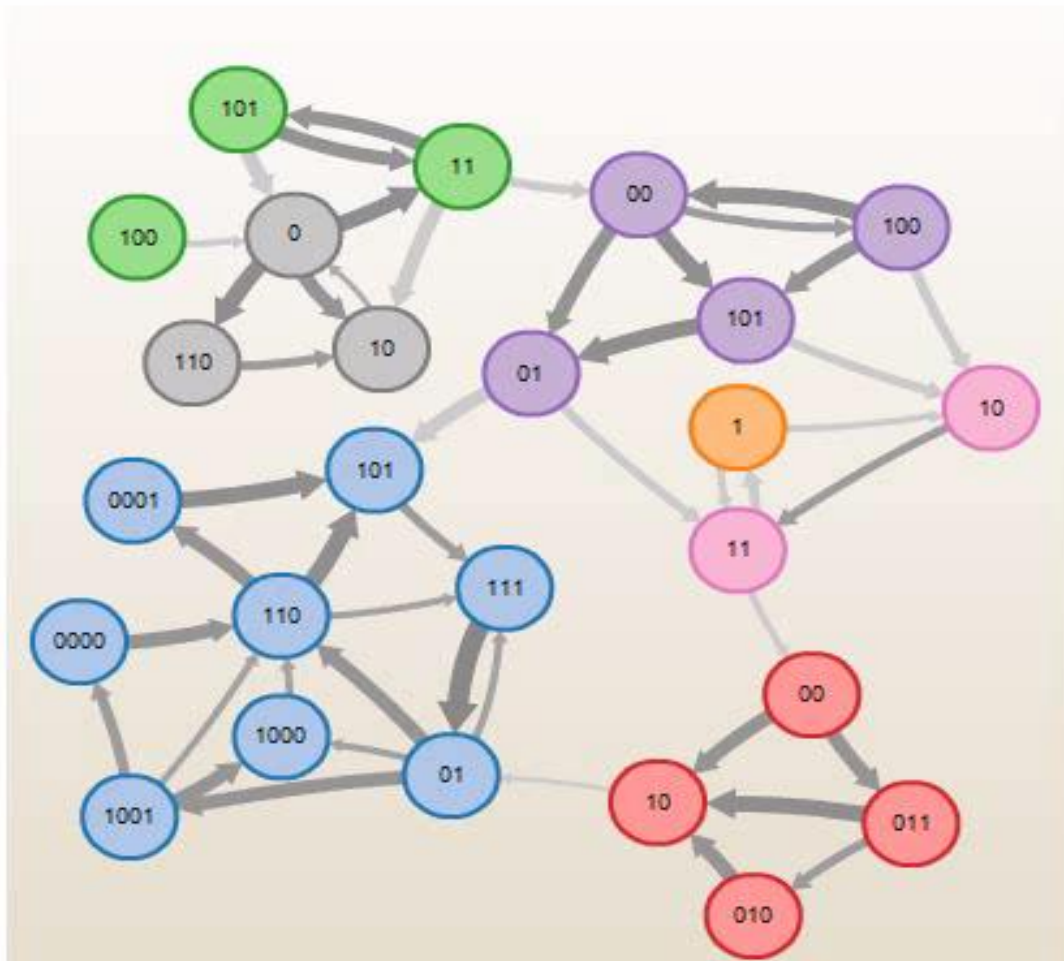
Степени вершин фиксированы



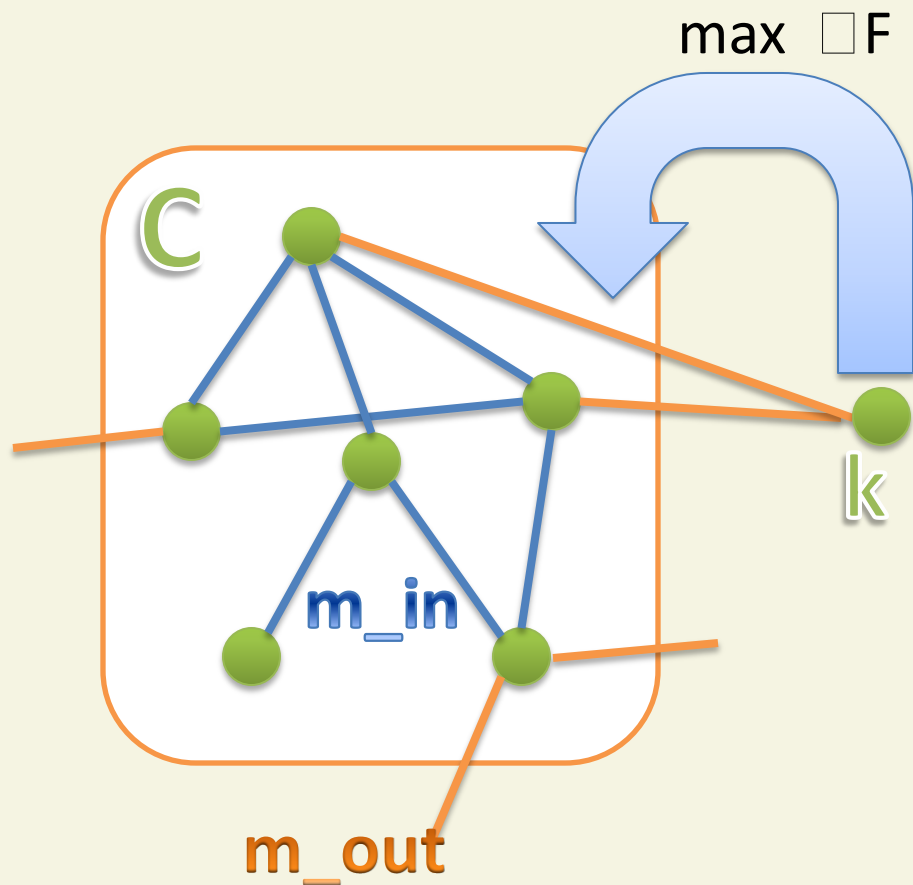
$$P_{ij} = \frac{k_i k_j}{2m}$$

вероятность \square ребра (i,j)

Случайные блуждания



Локальная оптимизация



$$F(m_{out}, m_{in}, k) \rightarrow \max$$

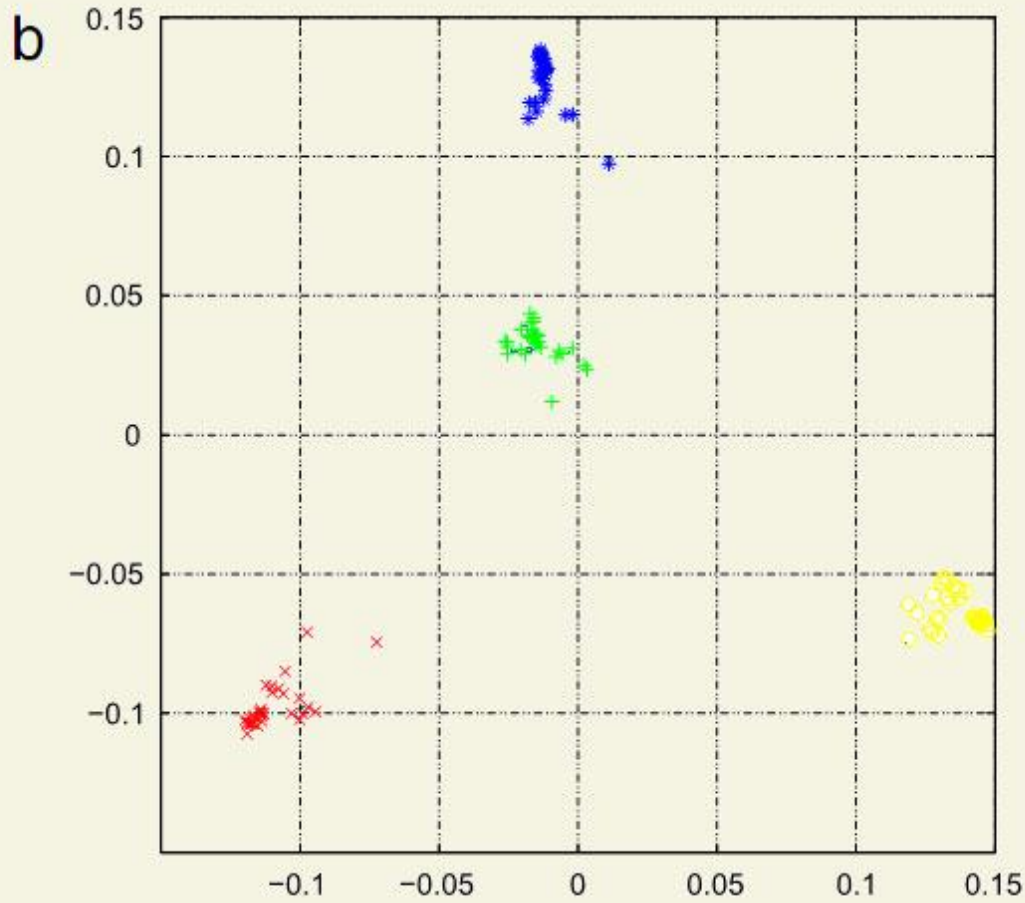
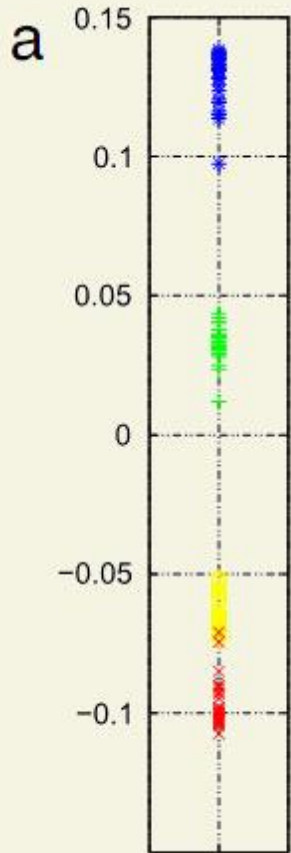
m_{out} – внешние ребра C

m_{in} – внутренние ребра C

k – степень вершины

Увеличение статистической
значимости C

Введение координат



Компоненты
собственных
векторов
Лапласиана

$$\mathcal{L}_{ij} = \begin{cases} 1 \\ -\frac{1}{\sqrt{k_i k_j}} \\ 0 \end{cases}$$



Координаты
вершин

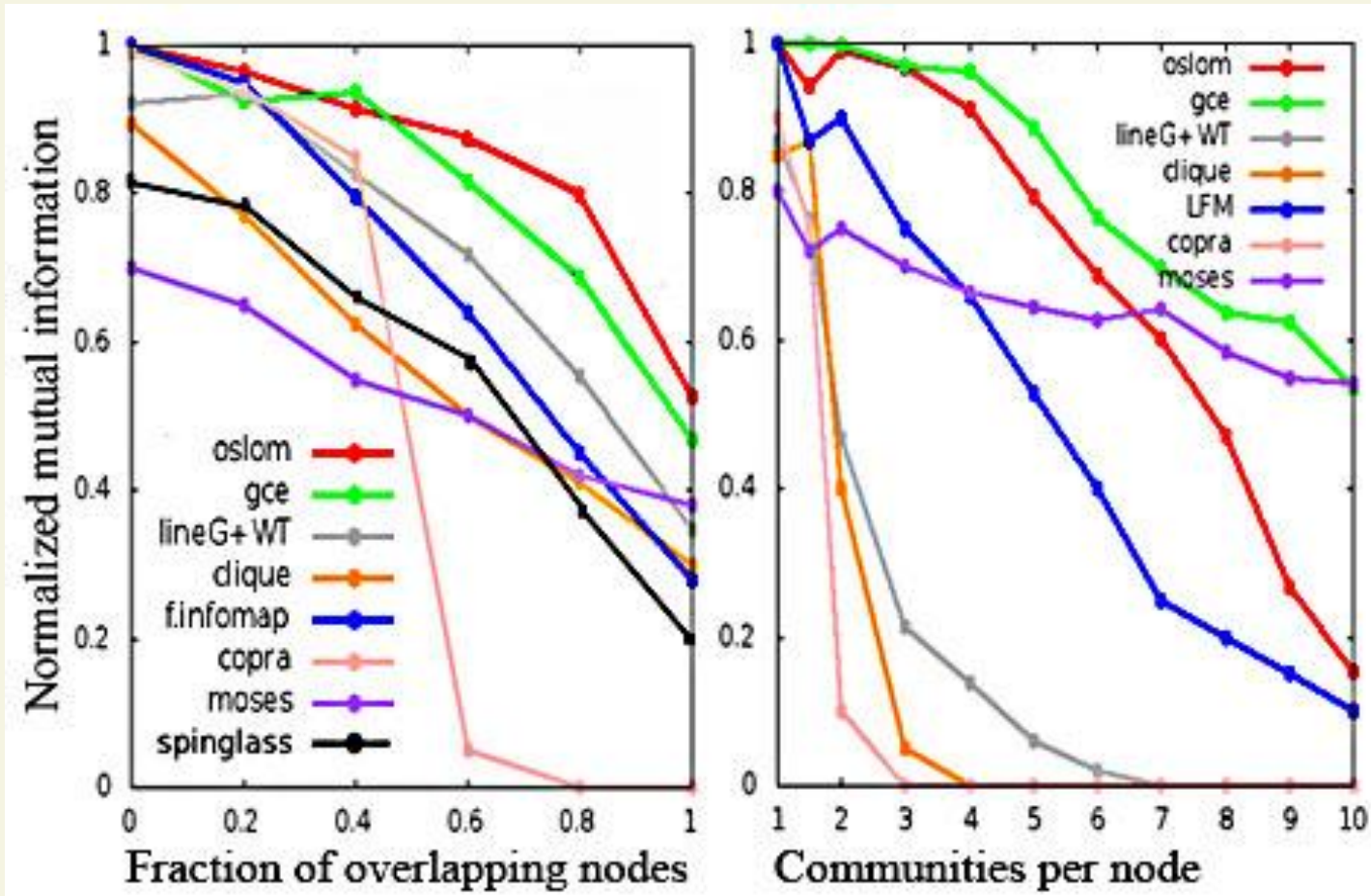


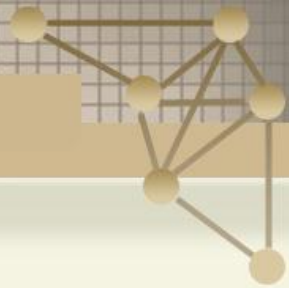
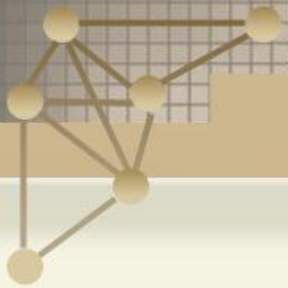
Fuzzy обобщение

- Линейные графы
- Клонирование вершин
- Коэффициенты принадлежности

Тестирование

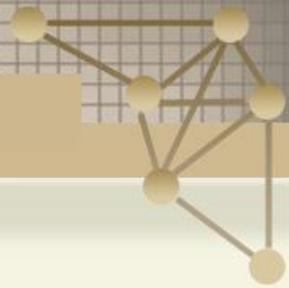
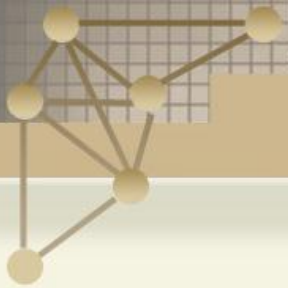
Искусственные сети с заданными сообществами





Заключение

- Использование в алгоритмах машинного обучения
- Сильно пересекающиеся сообщества
- Распознавание псевдосообществ
- Четыре класса подходов
- Результативность локальной оптимизации



Благодарю за внимание!