



Практический анализ данных и машинное обучение series

Небольшая, но очень милая подборка материалов



ВЫСШАЯ ШКОЛА ЭКОНОМИКИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

Тематический план занятий

- Введение в сетевой анализ, немного истории сетевого анализа, практическое использование сетей, ключевые описательные статистики сетей, формат работы с сетевыми данными
- Сообщества в социальных сетях
- Модели роста и формирования социальных сетей, Эгосети
- Сбор и загрузка данных из сетей



Рекомендуемые источники к Части I

- Базовый учебник: Ф.Харари. ТЕОРИЯ ГРАФОВ. М.: Мир, 1973, 300 стр.
- Из этого учебника важно ознакомиться лишь с серией страниц:
 - Страница 13-14 → подробно о задаче Кёнингбергских мостов
 - Страница 21-23 → базовое понятие графа, в том числе описание тех типов графов, которые были представлены в презентации
 - Страница 41-45 → факультивно
 - Страница 178-181 → про матрицы смежности/инцидентности
- По поводу метрик: существует бесчисленное множество ресурсов, но зачастую они все «непрактичны». Тем не менее, они все полезны одинаково. Что можно порекомендовать: ПРОСТОЕ УНТРО
- Elsevier® ScienceDirect® имеет неплохую страницу с представлением метрик: <https://www.sciencedirect.com/topics/computer-science/centrality-measure>
- Очень классная статья Francisco Aparecido Rodrigues о метриках, но с примерами расчетов: <https://arxiv.org/pdf/1901.07901.pdf>
- Начальные действия при работе с сетями в Python с пакетом NetworkX: ссылка на [aksakalli.github.io](https://github.com/aksakalli)



Рекомендуемые источники к Части II

- В Экономической социологии (к примеру) есть очень хорошая статья про природу сообществ и значения сетевых отношений: Грановеттер Марк Сила слабых связей // Экономическая социология. 2009. №4. URL: <https://cyberleninka.ru/article/n/sila-slabyh-svyazey>.
- Есть неплохая диссертация по анализу сообществ (а также пример законченного научного исследования социальных сетей с использованием всем базовых метрик):
<https://www.ccis.northeastern.edu/~amislove/publications/SocialNetworks-Thesis.pdf>
- Еще один эмпирический текст описывает принципы анализа сообществ (без привязки к Python): <https://arxiv.org/pdf/cond-mat/031-2040.pdf>
- Определение сообществ → «заслуга» методов Ньюмана и соавторов. Поэтому очень полезным оказывается работа Ньюмана и Гирвана: <https://arxiv.org/pdf/cond-mat/0308217.pdf>
- Как это гелать в Python? На примере одной из метрик (Louvain): <https://python-louvain.readthedocs.io/en/latest/>



Рекомендуемые источники к Части III

- Большой обзор Matthew O. Jackson по принципам формирования семей:
<http://www.cis.upenn.edu/~mkearns/teaching/NetworkedLife/netsurv.pdf>
- Еще один текст от Lars Backstrom и соавторов про формирование семей: <https://www.cs.cornell.edu/home/kleinber/kdd06-comm.pdf>
- Одна из важнейших статей из Экономической социологии по данной теме, в контексте фирм: <https://www.jstor.org/stable/3094462>
- Одна из важнейших техник анализа формирования и развитие (growth) семей – это использование соответствующих метрик и данных. Поэтому можно в качестве реферативного примера использовать следующие варианты:
 - Прекрасная иллюстрация принципов работы в обзорном препринте-статье Hiroki Sayama и соавт [располагается на arXiv за номером 1301.2561](#)
 - Дополнительно – про эгосети и суб-графы: 1) в учебнике Харари на странице 24-25 немного теории. 2) на сайте O`Reilly есть интересная статья про эгосети <https://www.oreilly.com/library/view/social-network-analysis/9781449311377/ch04.html>



Общие учебники и статьи

- 1) Ф.Харари. ТЕОРИЯ ГРАФОВ. М.: Мир, 1973, 300 стр.
- 2) Wasserman S., Faust K. Social network analysis: Methods and applications. – Cambridge university press, 1994. – Т. 8.
- 3) Scott J. Social network analysis //Sociology. – 1988. – Т. 22. – №. 1. – С. 109-127.
- 4) Granovetter M. S. The strength of weak ties //Social networks. – Academic Press, 1977. – С. 347-367.

Документация по NetworkX

- 1) Hagberg A., Swart P., S Chult D. Exploring network structure, dynamics, and function using NetworkX. – Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2008. – №. LA-UR-08-05495; LA-UR-08-5495.
- 2) https://networkx.github.io/documentation/latest/_downloads/networkx_reference.pdf

