ST5225 Statistical Analysis of Networks

Academic Year 2024/25, Semester 1

Adrian Röllin

Course overview

Week 1: Introduction to Networks

Objectives: Introduce basic network terminology, network types, and examples. Discuss various basic concepts and network properties. Discuss how to turn raw data into a network.

Keywords: Networks/graphs, nodes/vertices, links/edges, adjacency matrix, directed and undirected networks, weighted and unweighted networks, components, paths, cycles, loops.

References: [N2018, Sections 6.1–6.9, 6.11–6.12], [K2009, Section 2.1]

Week 2: Basic Quantities and Properties of Networks

Objectives: Introduce various quantities to measure the importance of vertices and edges.

Keywords: Degree, edge density, degree distribution, centrality, degree centrality, closeness centrality, betweenness centrality, eigenvector centrality, PageRank, triangles, cluster coefficient, similarity, assortativity, and mixing.

References: [N2018, Sections 6.10, 7.1–7.3, 7.6–7.7], [K2009, Sections 4.1–4.3]

Week 3: Basic Phenomena Around Real Networks

Objectives: Explore fundamental network phenomena like small-world effect, six degrees of separation, and Braess's Paradox.

Keywords: Small-world, six degrees of separation, snowball sampling, powerlaw degree distribution, Braess's Paradox

References: [N2018, Sections 4.6–4.7, Chapter 10], [EK2010, Sections 8.1–8.2]

Week 4: Basic Network Models

Objectives: Introduce the concept of random graph models, in particular the Erdős-Rényi random graph and the configuration model. Simulate these networks and analyze their properties.

Keywords: Random graphs, Erdős-Rényi random graph, configuration model.

References: [N2018, Chapters 11–12], [K2009, Section 6.2]

Week 5: Advanced Network Models

Objectives: Explore advanced network models like the preferential attachment model and spatial models. Simulate these networks and analyze their properties.

Keywords: Barabási-Albert Model and preferential attachment, spatial networks

References: [N2018, Chapters 13], [K2009, Section 6.4]; M. Barthelemy (2011). Spatial networks. *Physics Reports* **499**.

Week 6: Community Detection

Objectives: Learn about community detection algorithms and their application. Use software tools to detect and analyze communities in networks.

Keywords: Community detection, modularity, spectral clustering, hierarchical clustering.

References: [N2018, Sections 14.1–14.3, 14.5]; U. Von Luxburg U. (2007). A tutorial on spectral clustering. *Statistics and Computing* **17**. P. Doreian et al. (2020). *Advances in Network Clustering and Blockmodeling*, John Wiley and Sons.

Week 7: Stochastic Block Models (SBM) and Degree-Corrected SBMs (DCSBMs)

Objectives: Understand SBMs and DCSBMs for modeling community structure. Fit these models to data using statistical software.

Keywords: Stochastic Block Models, Degree-Corrected SBMs, Model Fitting

References: [N2018, Section 14.4]; E. Abbe (2018). Community detection and stochastic block models: recent developments. *Journal of Machine Learning Research* **18**. P. Doreian et al. (2020). *Advances in Network Clustering and Blockmodeling*, John Wiley and Sons.

Week 8: Exponential Random Graph Models (ERGMs)

Objectives: Learn about ERGMs for modeling complex network structures. Fit ERGMs to networks and interpret the results.

Keywords: Exponential Random Graph Models, Markov Chain Monte Carlo, Maximum Pseudolikelihood Estimation.

References: [K2009, Section 6.5]; D. R. Hunter et al. (2008). ERGM: A package to fit, simulate and diagnose exponential-family models for networks. *Journal of Statistical Software* **24**; G. Robins et al. (2007). An introduction to exponential random graph models for social networks. *Social Networks* **29**.

Week 9: Latent Space Models

Objectives: Introduce Latent Space Models for inferring hidden structures in networks. Estimate models and interpret latent spaces.

Keywords: Latent Space Models

References: https://sites.stat.washington.edu/raftery/Research/latent.html

Week 10: Network Visualization

Objectives: Focus on network visualization techniques and tools. Apply principles to effectively represent network data visually.

Keywords: Network Visualization, Gephi, Cytoscape, D3.js, Python, R, Data Visualization, Graph Layouts.

References: https://medium.com/@vespinozag/16-network-visualization-tools-that-you-should-know-2c26957b707e

Week 11: Machine Learning for Networks

Objectives: Learn about different machine learning tasks for networks. Understand the basic concepts like message passing and embedding.

Keywords: Node classification, edge prediction, clustering and community detection, Random walk embeddings, node embeddings, message passing

References: https://www.datacamp.com/tutorial/comprehensive-introduction-graph-neural-networks-gnns-tutorial; William L. Hamilton (2020). *Graph Representation Learning*

Week 12: Graph Neural Networks

Objectives: Explore Graph Neural Networks architectures and their applications. Implement GNNs for real-world data.

Keywords:

Week 13: Revision and Wrap-up

Objectives: Review key concepts and prepare for final assessments.

Resources

- https://networkrepository.com
- https://snap.stanford.edu/data/

References

- [N2018] *Networks*, 2nd edition. M. Newman, Oxford University Press, 2018.
- [K2009] *Statistical Analysis of Network Data: Methods and Models*. E. D. Kolaczyk, Springer, 2009.
- [KC2018] *Statistical Analysis of Network Data with R*, 2nd edition. E. D. Kolacyk and G. Csardi, Springer, 2018.
- [EK2010] *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. D. Easley and J. Kleinberg, Cambridge University Press, 2010.