The Impact of Graph Structure on Small-World Shortest Paths

Pegah Hozhabrierdi and Reza Zafarani
Data Lab, College of Engineering & Computer Science,
Syracuse University
{pegah, reza}@data.syr.edu

Motivation

Motivation: Adding "social" to social networks analysis!

Idea:

- **Small-world shortest paths:** Merging the concept of shortest paths in a social network with small-world phenomenon.
- **Small-world representation:** The flow of information is NOT always unidirectional!

Analysis:

- *Global* and *local* network measures vs. small-world shortest paths

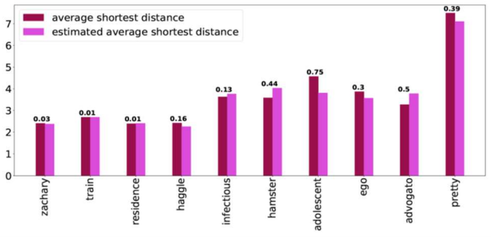
Network Measures and Average Shortest Path

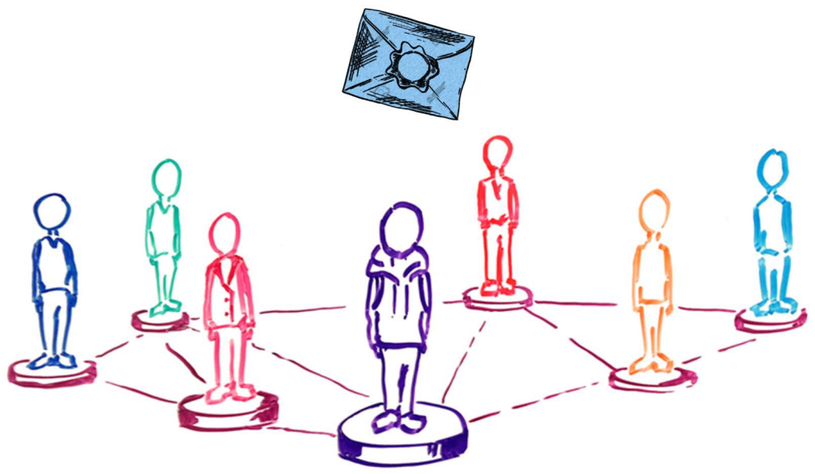
Goal:

- Identification of network measures with highest average shortest path prediction capability

Data: 10 real-world online and physical social networks

Influential Network Measures: Maximum Degree Centrality and Diameter



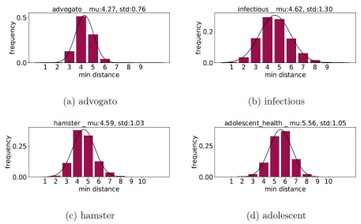


References:

- http://www.hunter.cuny.edu/advising/my-success-network/my-success-network
- https://www.123rf.com/photo_93679307_stock-vector-envelope-in-sketch-style-with-sealing-wax-outline-or-contour-drawing-hand-drawn-vector-isolated-vint.html
- **P. Hozhabrierdi, R. Zafarani, "The Impact of Graph Structure on Small-World Shortest Paths," In Proceedings of International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation (SBP-BRiMs), Springer, 2019.**

Local Measures and Shortest Path Distribution

- **Goal:** Can we approximate shortest path distribution using local network measures?
- **Shortest path node ranking:** $\mathrm{SPN}(i) = \frac{\sum_{j \neq i} d_{\min}(i,j)}{|V|-1}$
- **Distributions:**
  - Naïve Unif (1, $\mathrm{SPN}_{\max}$) $\quad \mathrm{SPN}_{\max} = \frac{1+2+\ldots+|V|-1}{|V|-1} = \frac{|V|}{2}$
  - Small-World Unif (1, $\mathrm{SPN}_{\max}$) $\quad \mathrm{SPN}_{\max} \approx \mathrm{SPN}(n_{mcc})$
  - Small-World Normal (SPN (n_{mcc}), 1)



(a) advogato (b) infectious (c) hamster (d) adolescent

- **Results:**

Dataset	deg. dist.	deg. cent.	eig. cent.	local clust. coeff.	naïve uniform	uniform	norma
zachary	0.32	0.32	0.24	0.29	0.14	0.04	**0.03**
train_bombing	0.44	0.44	0.88	0.57	0.17	0.09	**0.06**
residence_hall	0.13	0.13	0.26	0.09	0.21	0.08	**0.04**
haggle	1.29	1.29	1.13	0.65	0.23	0.07	**0.04**
infectious	0.28	0.28	1.66	0.19	0.24	0.1	**0.02**
hamster	0.67	0.67	1.53	0.19	0.25	0.1	**0.02**
adolescent_health	0.19	0.19	1.43	0.35	0.24	0.11	**0.01**
ego_facebook	1.57	1.57	1.78	1.82	0.27	0.07	**0.04**
advogato	1.06	1.06	1.24	0.32	0.25	0.09	**0.03**
pretty_good_privacy	0.78	0.78	0.38	0.19	0.26	0.15	**0.02**

Maximum Clustering Coefficient + **Standard Normal**

- **Big Picture:**
  - Approximating global properties of social networks using only network local information