HUMAN DISEASE NETWORK

Network Analysis
IT University of Copenhagen

Christina Aftzidis (caft@itu.dk), Germán Buttiero (gebu@itu.dk), Hubert Wójcik (huwo@itu.dk),
Laurids Holme Pedersen (lrpe@itu.dk), Paula Menshikoff (pmen@itu.dk)

NETWORK DESCRIPTION

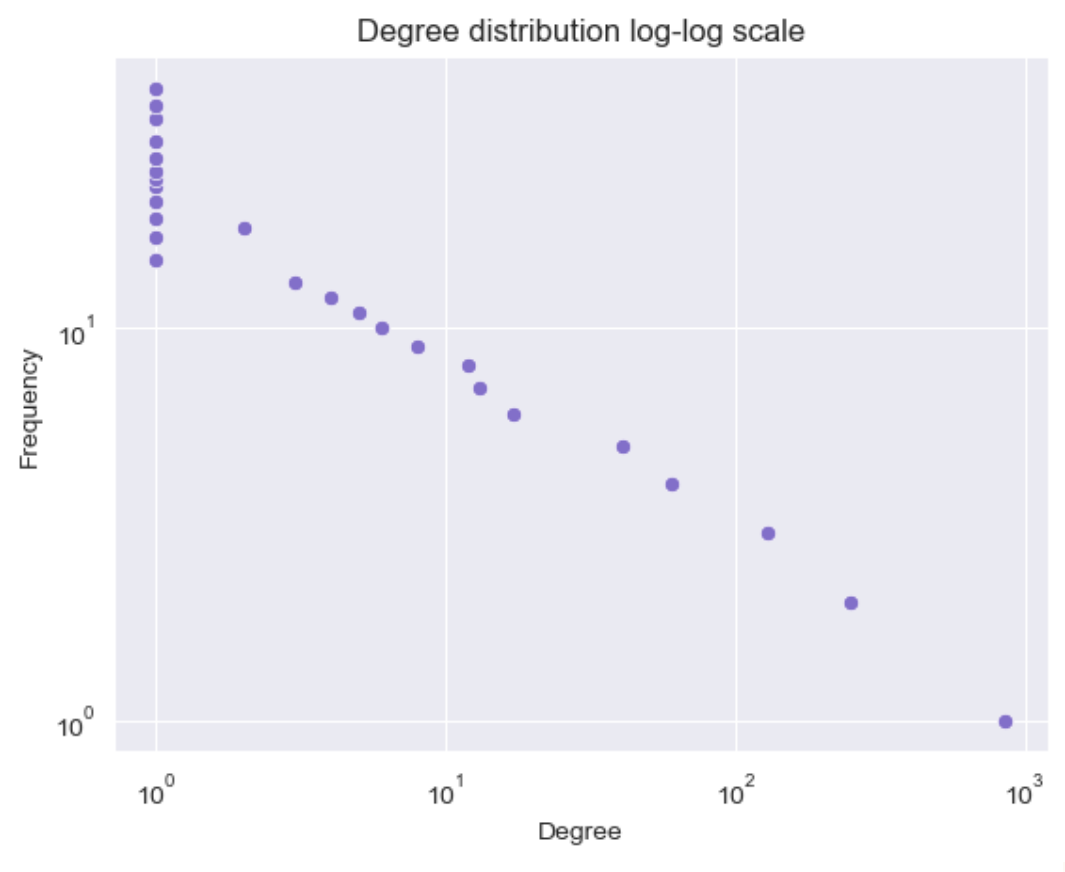
- Describes the association between human genes and diseases
- Extracted from the Morbid Map (MM) of the Online Mendelian Inheritance in Man (OMIM) in 2005

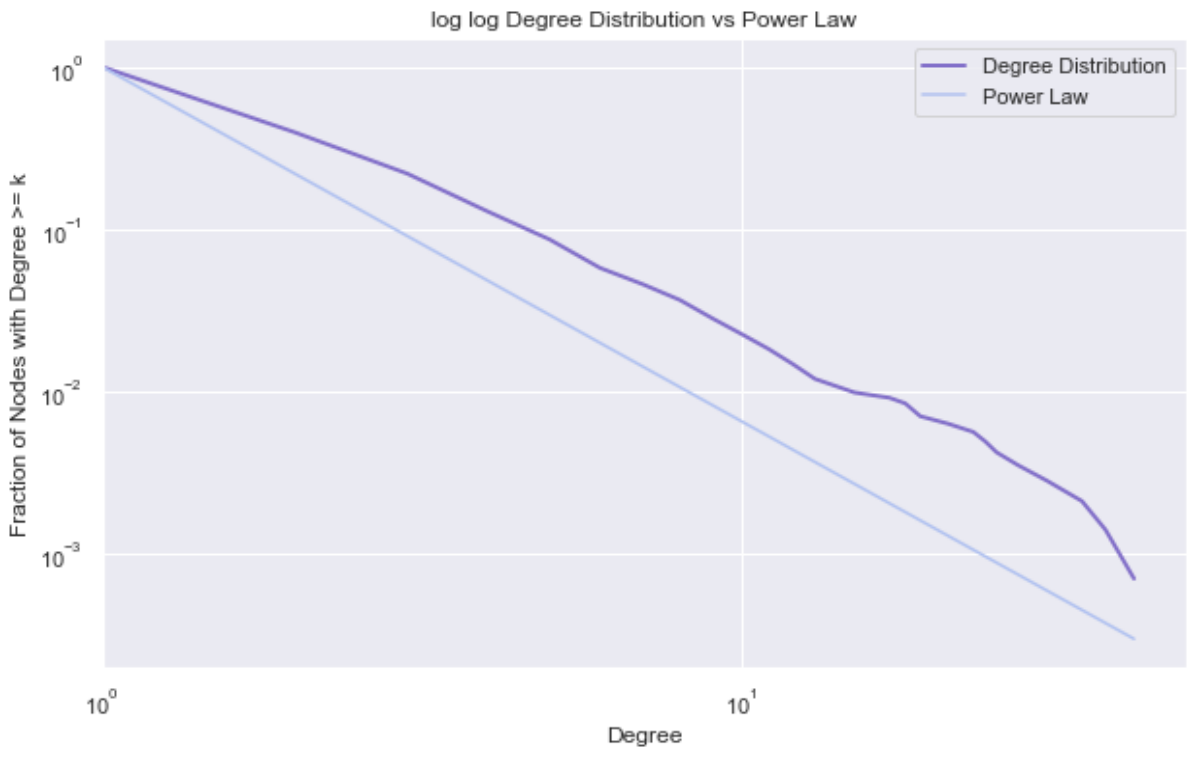
NETWORK DESCRIPTION

Type of network	Bipartite, undirected, unweighted with attributes
Number of nodes that are genes	903
Number of nodes that are diseases	516
Total number of nodes	1419
Total number of edges	1550
Number of classes	22
Edge type	Edges represent association between a gene and a disease
Density	0.00154

NETWORK DESCRIPTION

DEGREE DISTRIBUTION AND POWER LAW





RESEARCH QUESTION

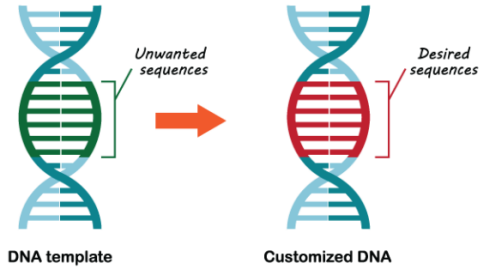
If we could treat gene mutations with the CRISPR method, which genes should be targeted first?

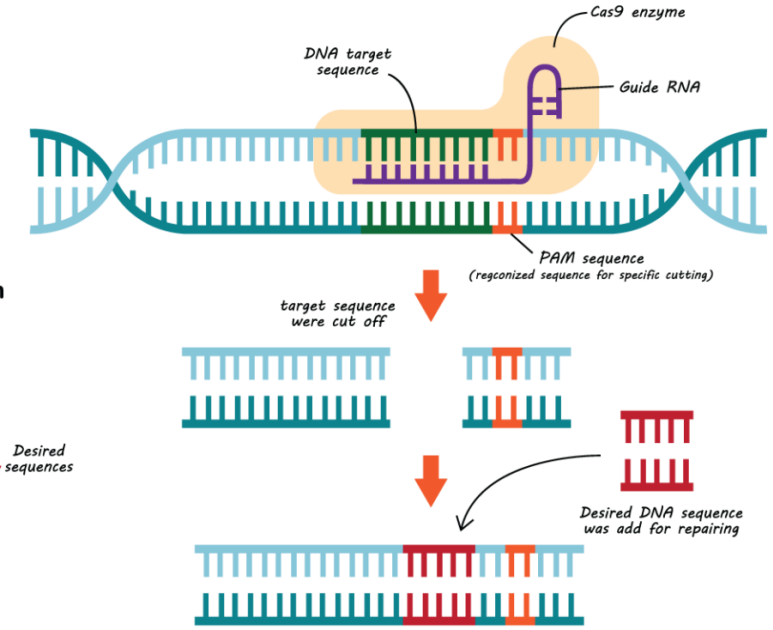
CRISPR-CAS9

Science

How does CRISPR-Cas9 work?

- Adapted from defense mechanism against virus of bacteria
- Cas9 is an enzyme using guide RNA leading to cut target DNA sequence
- Desired genetic sequence could add in repairing system for customize DNA





SOURCE: HTTPS://CAS9.COM/WHAT-IS-CRISPR-CAS9-AND-HOW-DOES-IT-WORK/

CASE

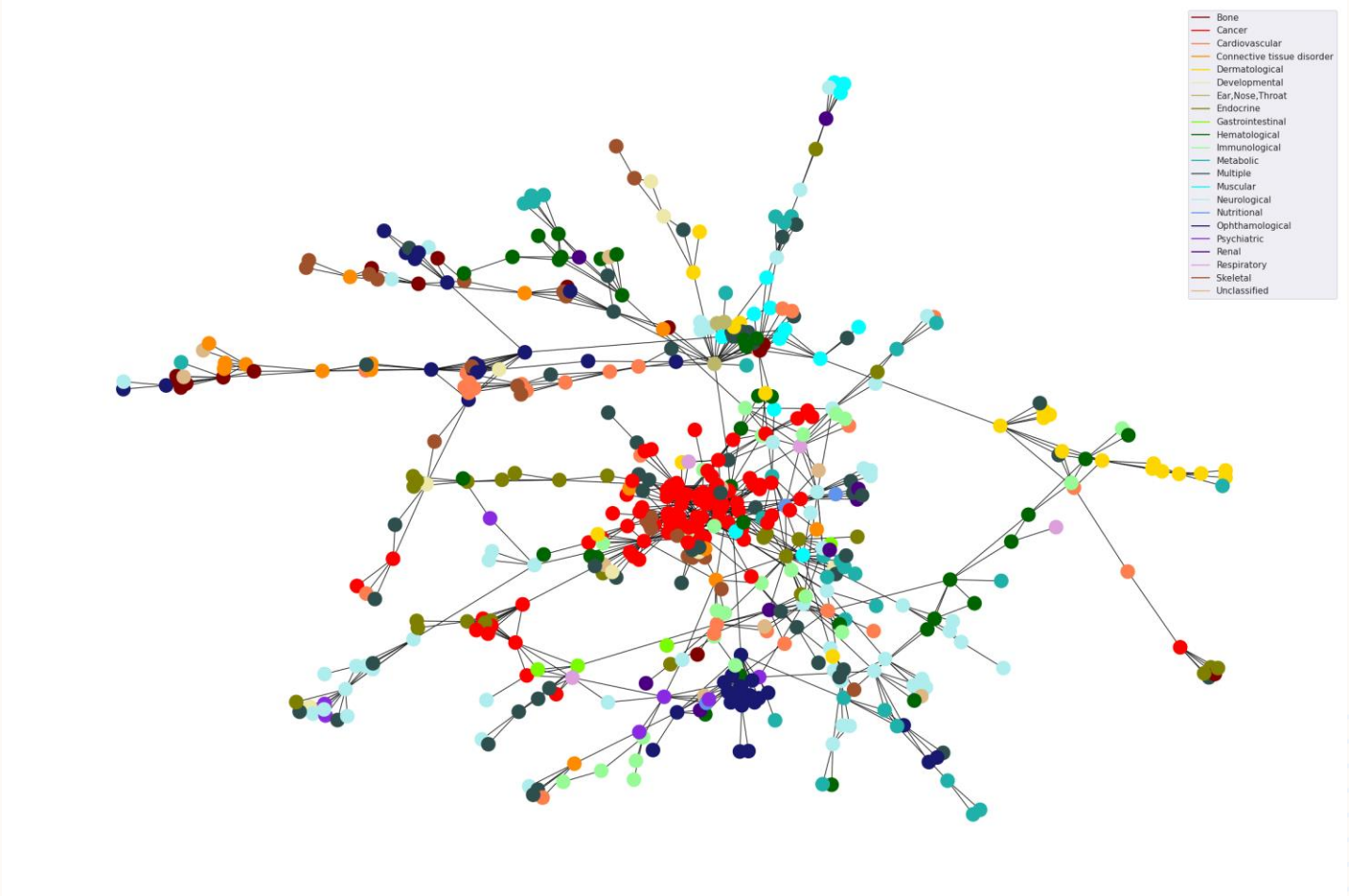
We are researchers hired by biologists and investigating the relationship between genes and diseases. We want to know which genes we should focus on with the CRISPR-Cas9 method in order to have the most impact on the diseases.

ASSUMPTIONS

- CRISPR works perfectly (no side effects)
- Our data set is complete (we are missing no genes that could be impactful)
- Every gene has the same weight
- Every “fixed” gene is permanently altered in the genome

PROJECTION

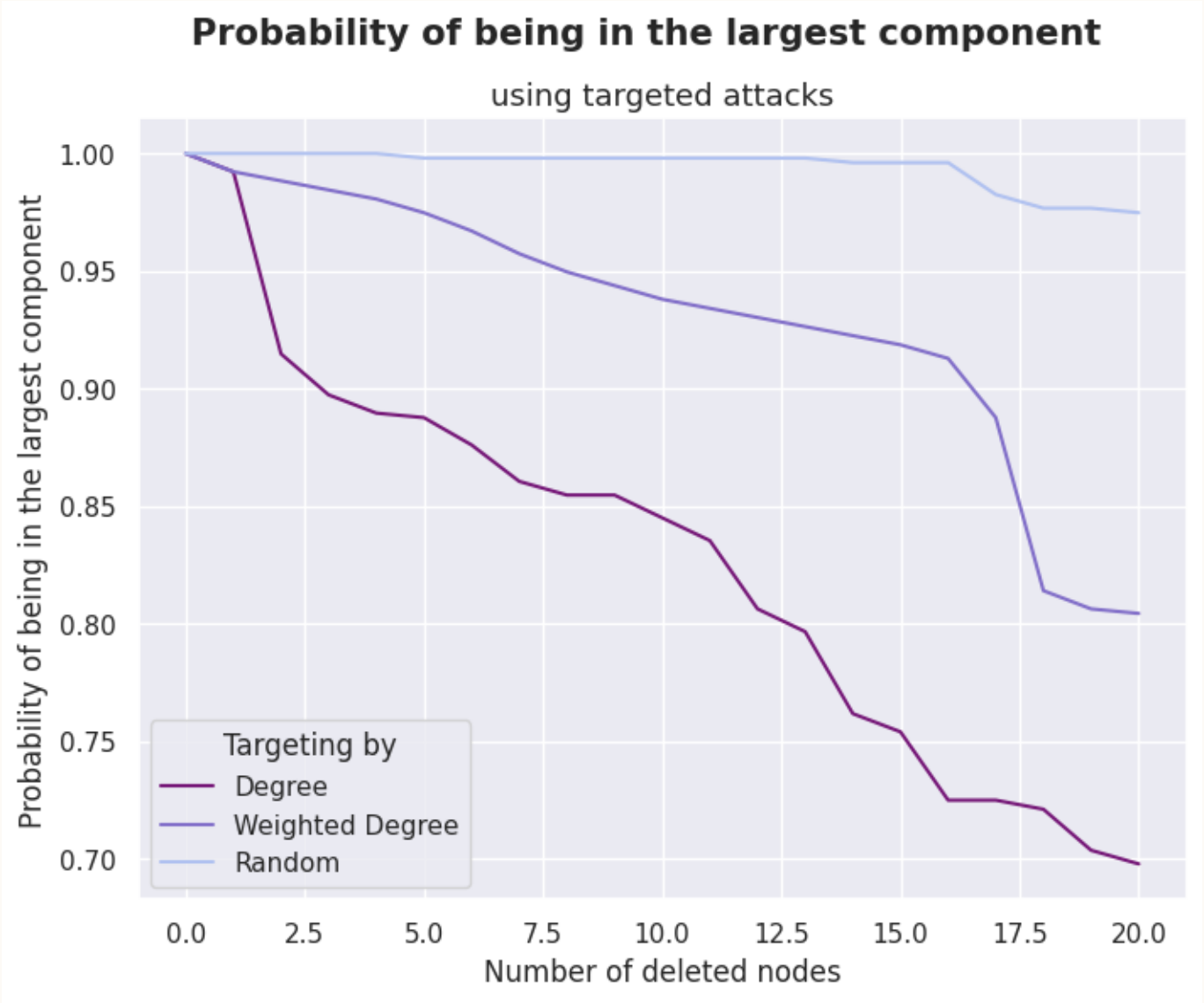
- Projection on the diseases after every deleted gene node
- Used for community discovery, mutual information, modularity and assortativity per group



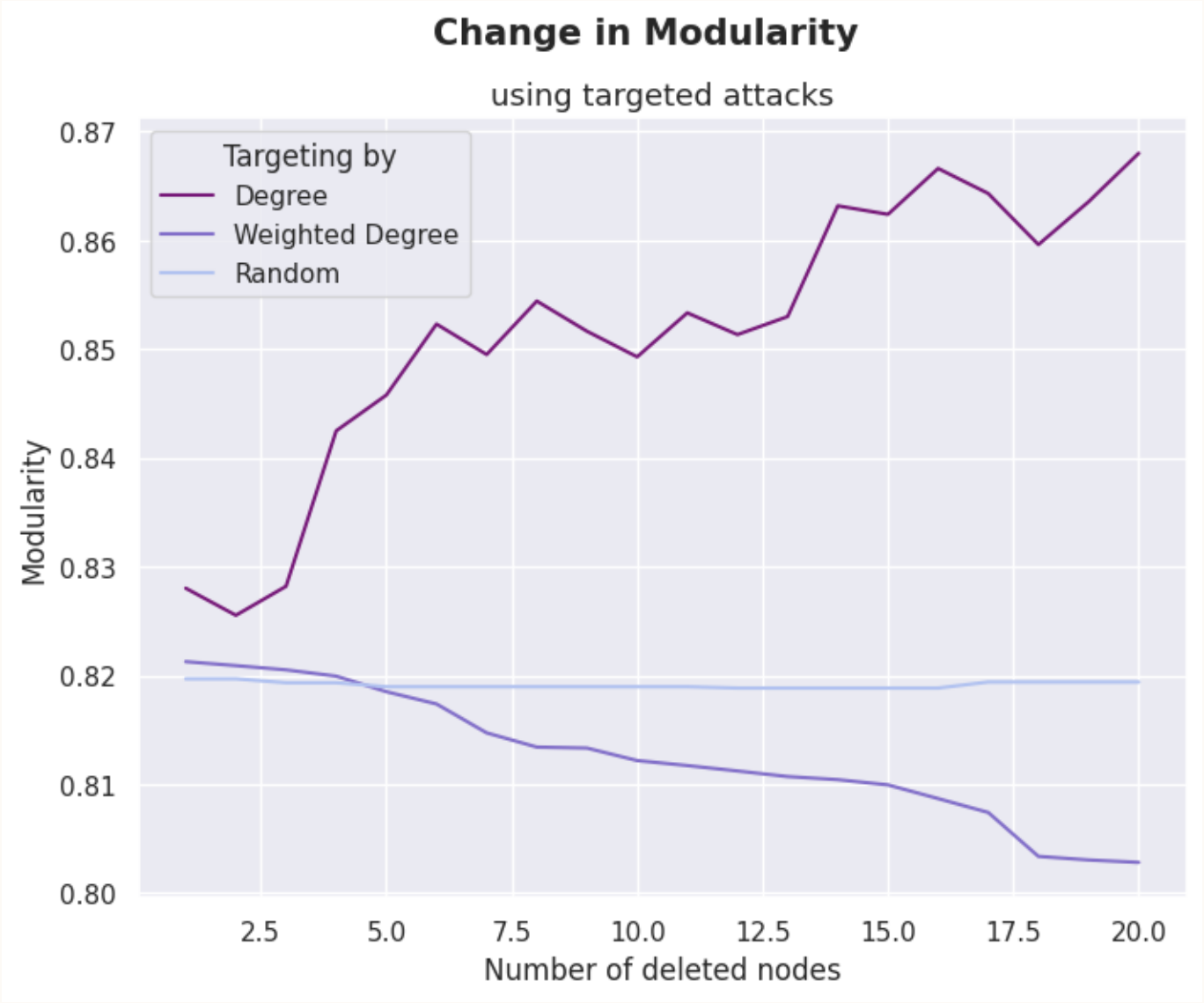
DIFFERENT SCENARIOS

1. We delete 20 genes randomly from the network
2. We delete the 20 genes with the highest degree
3. We delete the 20 genes with the highest degree weighted by how much they contribute to the diseases they are connected to

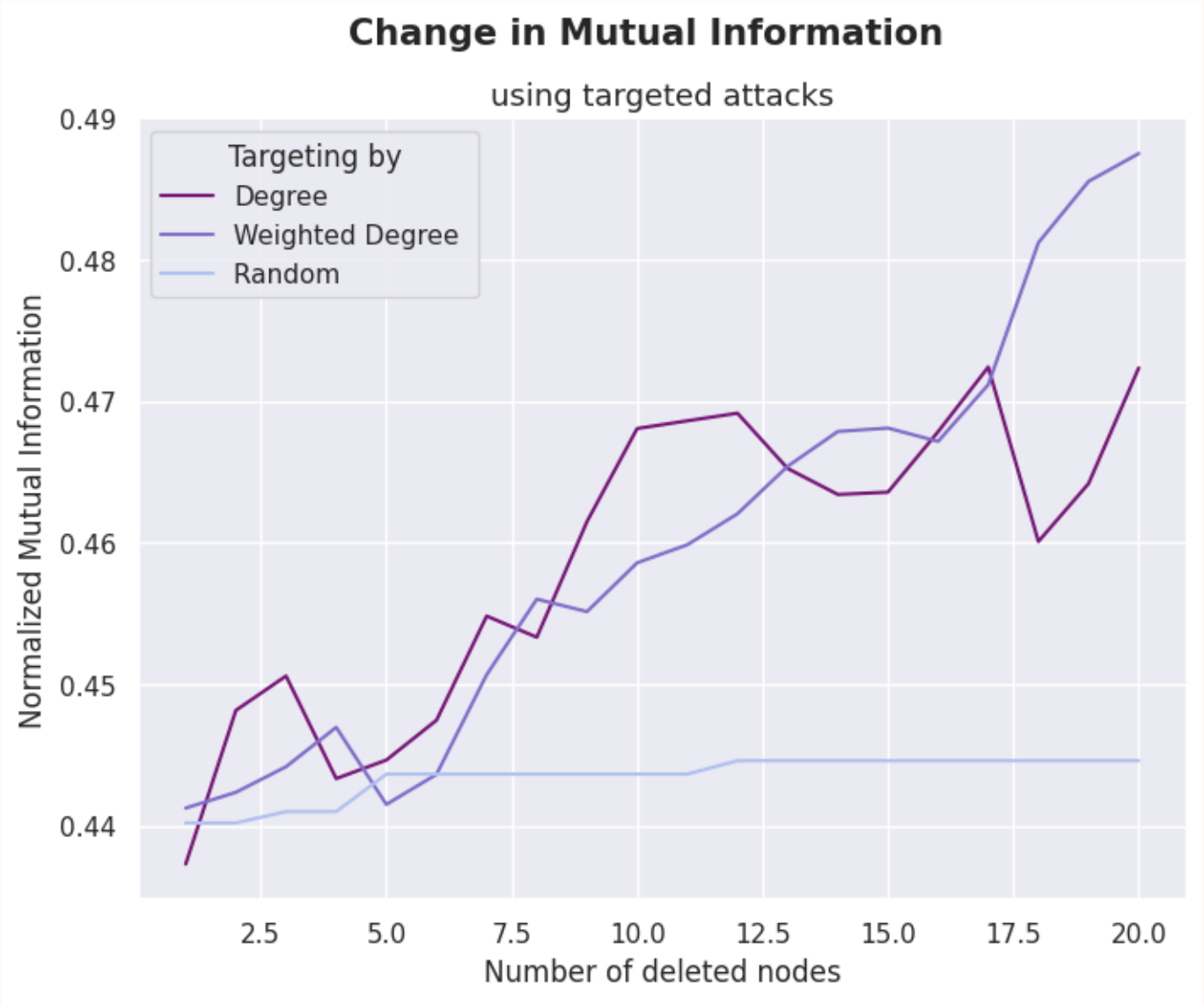
LARGEST COMPONENT



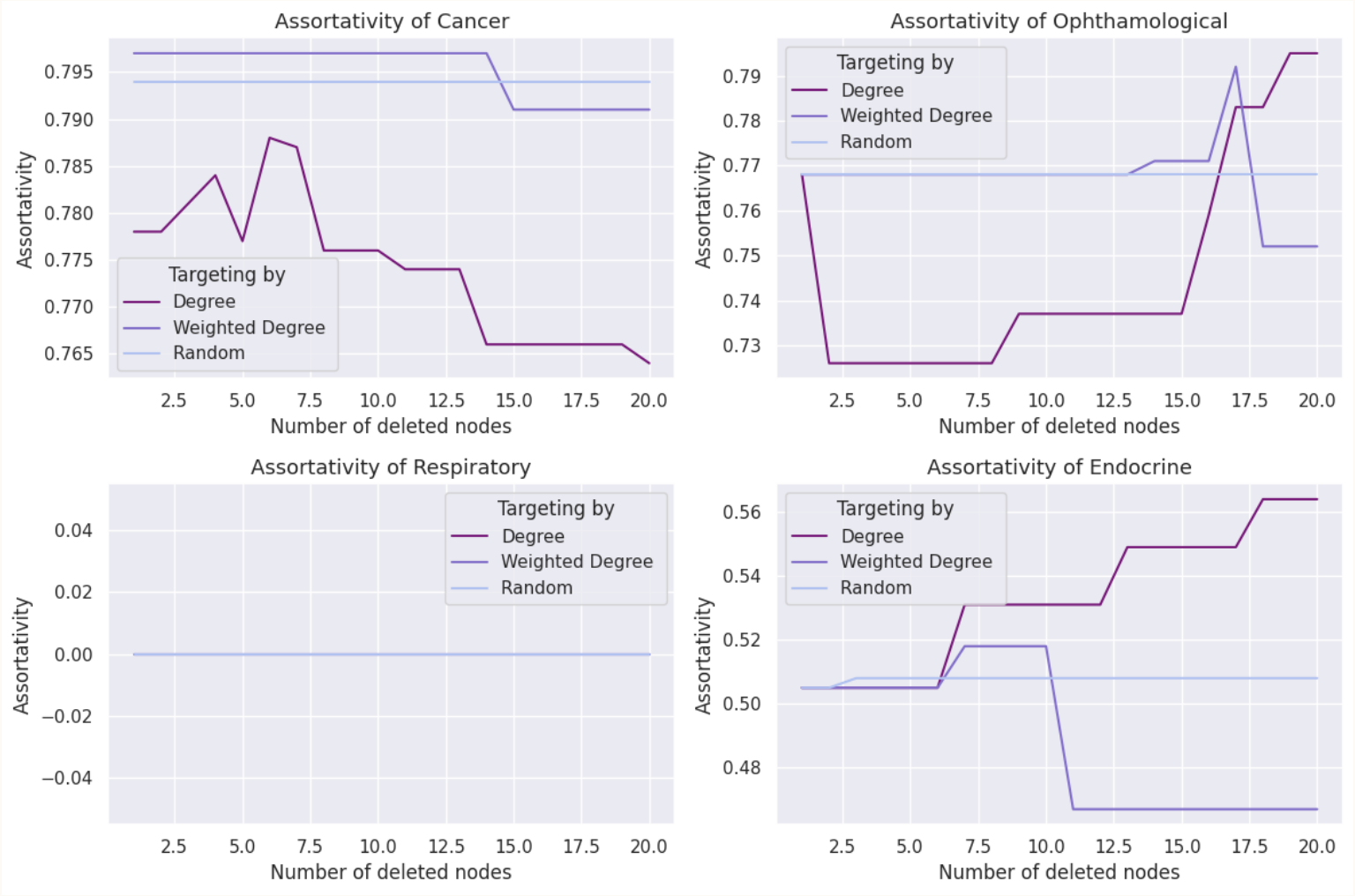
MODULARITY


Change in Modularity
using targeted attacks
Targeting by
Degree
Weighted Degree
Random
Modularity
0.87
0.86
0.85
0.84
0.83
0.82
0.81
0.80
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
Number of deleted nodes

MUTUAL INFORMATION



ASSORTATIVITY PER GROUP



LIMITATIONS

- We are not biologists, so we do not have sufficient knowledge to put our findings into proper context
  - We can only make assumptions

RESULTS

RANDOM DELETION

- Little to no influence

DELETION BY DEGREE

- Big influence on the probability of being in the largest component
- Isolates the communities
- Higher mutual information

DELETION BY WEIGHTED DEGREE

- Influences the probability of being in the largest component
- Connects communities
- Higher mutual information

CONCLUSION

- Some changes might depend on the number of nodes in the group
- Targeting genes with high degrees might lower the chances of having multiple diseases since it disconnects the network
- Weighted degree might have more of an influence on preventing diseases in general since the gene is more important for the expression of the disease

CITATIONS

- (2021, October 3). *What is CRISPR-Cas9 and How Does it Work?* CRISPR Cas9 - Gene Editing. Retrieved December 8, 2022, from https://cas9.com/what-is-crispr-cas9-and-how-does-it-work

- (n.d.). Datasets · gephi/gephi Wiki. GitHub. https://github.com/gephi/gephi/wiki/Datasets

- Goh, K. I., Cusick, M. E., Valle, D., Childs, B., Vidal, M., & Barabási, A. L. (2007). The human disease network. Proceedings of the National Academy of Sciences, 104(21), 8685–8690. https://doi.org/10.1073/pnas.0701361104

THANK YOU

Group 2