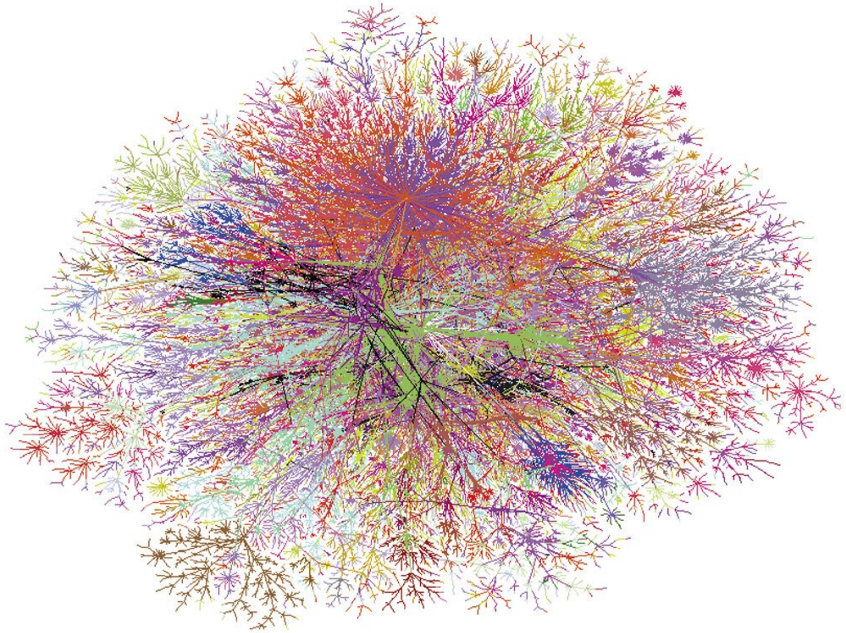
Network Destroy-Repair Game

December 5, 2022

Vikas Kashyap, Tony Huang (Group 8)



PROJECT GOAL

Recap of the Project

“Networks play a key role in the robustness of biological, social and technological systems. Whenever nature seeks robustness, it resorts to networks.”

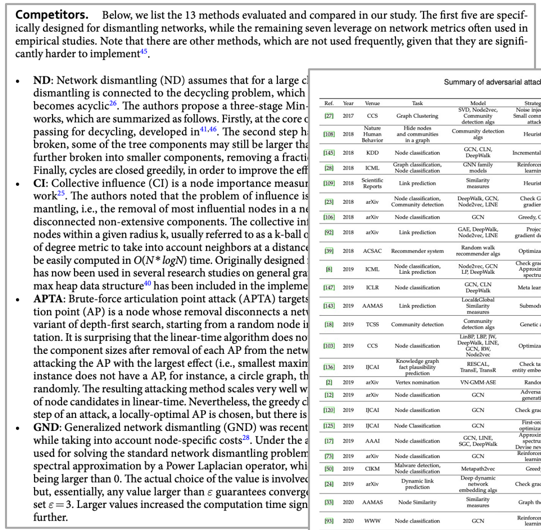
Network Science - Albert-László Barabási

RELATED WORK

Prior Work

ALGORITHMS

1. Graph data are vulnerable to attack, and **reinforcement learning**-based approaches have been proposed, given a properly designed reward mechanism
2. Researchers have also studied adversarial **attacks on nodes**, using heuristics on special networks



Competitors. Below, we list the 13 methods evaluated and compared in our study. The first five are specifically designed for dismantling networks, while the remaining seven leverage on network metrics often used in empirical studies. Note that there are other methods, which are not used frequently, given that they are significantly harder to implement[45].

- **ND**: Network dismantling (ND) assumes that for a large c
dismantling is connected to the decycling problem, which
becomes acyclic[26]. The authors propose a three-stage Min-
works, which are summarized as follows. Firstly, at the core c
passing for decycling, developed in[41,46]. The second step h
broken, some of the tree components may still be larger tha
further broken into smaller components, removing a fracti
Finally, cycles are closed greedily, in order to improve the eff
- **CI**: Collective influence (CI) is a node importance measur
work[25]. The authors noted that the problem of influence is
mantling, i.e., the removal of most influential nodes in a ne
disconnected non-extensive components. The collective inf
nodes within a given radius k, usually referred to as a k-ball c
of degree metric to take into account neighbors at a distance
be easily computed in $O(N * logN)$ time. Originally designed
has now been used in several research studies on general gra
max heap data structure[40] has been included in the impleme
- **APTA**: Brute-force articulation point attack (APTA) targets
tion point (AP) is a node whose removal disconnects a netw
variant of depth-first search, starting from a random node i
tation. It is surprising that the linear-time algorithm does no
the component sizes after removal of each AP from the netw
attacking the AP with the largest effect (i.e., smallest maxim
instance does not have a AP, for instance, a circle graph, th
randomly. The resulting attacking method scales very well w
of node candidates in linear-time. Nevertheless, the greedy c
step of an attack, a locally-optimal AP is chosen, but there is
- **GND**: Generalized network dismantling (GND) was recent
while taking into account node-specific costs[28]. Under the a
used for solving the standard network dismantling problem
spectral approximation by a Power Laplacian operator, whi
being larger than 0. The actual choice of the value is involved
but, essentially, any value larger than ε guarantees converge
set $\varepsilon = 3$. Larger values increased the computation time sign
further.

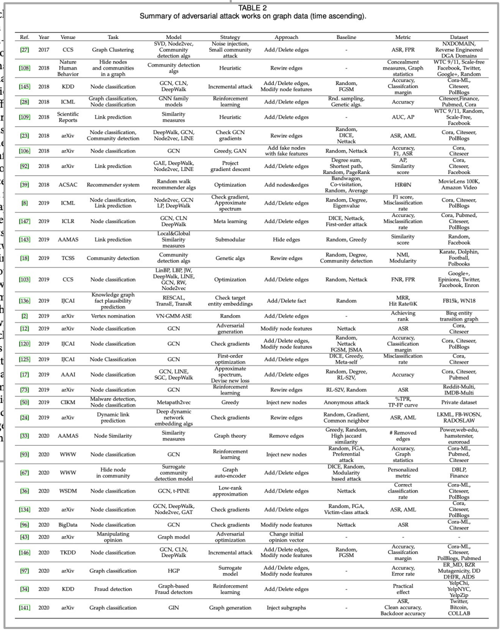


TABLE 2
Summary of adversarial attack works on graph data (time ascending).

Ref.	Year	Venue	Task	Model	Strategy	Approach	Baseline	Metric	Dataset
[27]	2017	CCS	Graph Clustering	SVD, Node2vec, Community detection algs	Noise injection, Small community attack	Add/Delete edges	-	ASR, FPR	NXDOMAIN, Reverse Engineered DGA Domains
[108]	2018	Nature Human Behavior	Hide nodes and communities in a graph	Community detection algs	Heuristic	Rewire edges	-	Concealment measures, Graph statistics	WTC 9/11, Scale-free Facebook, Twitter, Google+, Random
[145]	2018	KDD	Node classification	GCN, CLN, DeepWalk	Incremental attack	Add/Delete edges, Modify node features	Random, FGSM	Accuracy, Classifcation margin	Cora-ML, Citeseer, PolBlogs
[28]	2018	ICML	Graph classification, Node classification	GNN family models	Reinforcement learning	Add/Delete edges	Rnd. sampling, Genetic algs.	Accuracy	Citeseer,Finance, Pubmed, Cora
[109]	2018	Scientific Reports	Link prediction	Similarity measures	Heuristic	Add/Delete edges	-	AUC, AP	WTC 9/11, Random, Scale-Free, Facebook
[23]	2018	arXiv	Node classification, Community detection	DeepWalk, GCN, Node2vec, LINE	Check GCN gradients	Rewire edges	Random, DICE, Nettack	ASR, AML	Cora, Citeseer, PolBlogs
[106]	2018	arXiv	Node classification	GCN	Greedy, GAN	Add fake nodes with fake features	Random, Nettack	Accuracy, F1, ASR	Cora, Citeseer
[92]	2018	arXiv	Link prediction	GAE, DeepWalk, Node2vec, LINE	Project gradient descent	Add/Delete edges	Degree sum, Shortest path, Random, PageRank	AP, Similarity score	Cora, Citeseer, Facebook
[39]	2018	ACSAC	Recommender system	Random walk recommender algs	Optimization	Add nodes&edges	Bandwagon, Co-visitation, Random, Average	HR@N	MovieLens 100K, Amazon Video
[8]	2019	ICML	Node classification, Link prediction	Node2vec, GCN LP, DeepWalk	Check gradient, Approximate spectrum	Add/Delete edges	Random, Degree, Eigenvalue	F1 score, Misclassification rate	Cora, Citeseer, PolBlogs
[147]	2019	ICLR	Node classification	GCN, CLN DeepWalk	Meta learning	Add/Delete edges	DICE, Nettack, First-order attack	Accuracy, Misclassification rate	Cora, Pubmed, Citeseer, PolBlogs
[143]	2019	AAMAS	Link prediction	Local&Global Similarity measures	Submodular	Hide edges	Random, Greedy	Similarity score	Random, Facebook
[18]	2019	TCSS	Community detection	Community detection algs	Genetic algs	Rewire edges	Random, Degree, Community detection	NMI, Modularity	Karate, Dolphin, Football, Polbooks
[103]	2019	CCS	Node classification	LinBP, LBP, JW, DeepWalk, LINE, GCN, RW, Node2vec	Optimization	Add/Delete edges	Random, Nettack	FNR, FPR	Google+, Epinions, Twitter, Facebook, Enron
[136]	2019	IJCAI	Knowledge graph fact plausibility prediction	RESCAL, TransE, TransR	Check target entity embeddings	Add/Delete fact	Random	MRR, Hit Rate@K	FB15k, WN18
[2]	2019	arXiv	Vertex nomination	VN·GMM·ASE	Random	Add/Delete edges	-	Achieving rank	Bing entity transition graph
[12]	2019	arXiv	Node classification	GCN	Adversarial generation	Modify node features	Nettack	ASR	Cora, Citeseer
[120]	2019	IJCAI	Node classification	GCN	Check gradients	Add/Delete edges, Modify node features	Random, Nettack FGSM, JSMA	Accuracy, Classification margin	Cora, Citeseer, PolBlogs
[125]	2019	IJCAI	Node Classification	GCN	First-order optimization	Add/Delete edges	DICE, Greedy, Meta-self	Misclassification rate	Cora, Citeseer
[17]	2019	AAAI	Node classification	GCN, LINE, SGC, DeepWalk	Approximate spectrum, Devise new loss	Add/Delete edges	Random, Degree, RL-S2V,	Accuracy	Cora, Citeseer, Pubmed
[73]	2019	arXiv	Node classification	GCN	Reinforcement learning	Rewire edges	RL-S2V, Random	ASR	Reddit-Multi, IMDB-Multi
[50]	2019	CIKM	Malware detection, Node classification	Metapath2vec	Greedy	Inject new nodes	Anonymous attack	%TPR, TP-FP curve	Private dataset
[24]	2019	arXiv	Dynamic link prediction	Deep dynamic network embedding algs	Check gradients	Rewire edges	Random, Gradient, Common neighbor	ASR, AML	LKML, FB-WOSN, RADOSLAW
[33]	2020	AAMAS	Node Similarity	Similarity measures	Graph theory	Remove edges	Greedy, Random, High jaccard similarity	# Removed edges	Power,web-edu, hamsterster, euroroad
[93]	2020	WWW	Node classification	GCN	Reinforcement learning	Inject new nodes	Random, FGA, Preferential attack	Accuracy, Graph statistics	Cora-ML, Pubmed, Citeseer
[67]	2020	WWW	Hide node in community	Surrogate community detection model	Graph auto-encoder	Add/Delete edges	DICE, Random, Modularity based attack	Personalized metric	DBLP, Finance
[36]	2020	WSDM	Node classification	GCN, t-PINE	Low-rank approximation	Add/Delete edges	Nettack	Correct classification rate	Cora-ML, Citeseer, PolBlogs
[134]	2020	arXiv	Node classification	GCN, DeepWalk, Node2vec, GAT	Check gradients	Add/Delete edges	Random, FGA, Victim-class attack	ASR, AML	Cora, Citeseer, PolBlogs
[96]	2020	BigData	Node classification	GCN	Check gradients	Modify node features	Nettack	ASR	Cora-ML, Citeseer
[43]	2020	arXiv	Manipulating opinion	Graph model	Adversarial optimization	Change initial opinion vector	-	-	-
[146]	2020	TKDD	Node classification	GCN, CLN, DeepWalk	Incremental attack	Add/Delete edges, Modify node features	Random, FGSM	Accuracy, Classifcation margin	Cora-ML, Citeseer, PolBlogs, Pubmed
[97]	2020	arXiv	Graph classification	HGP	Surrogate model	Add/Delete edges, Modify node features	-	Accuracy, Error rate	ER_MD, BZR Mutagenicity, DD DHFR, AIDS
[34]	2020	KDD	Fraud detection	Graph-based Fraud detectors	Reinforcement learning	Add/Delete edges	-	Practical effect	YelpChi, YelpNYC, YelpZip
[141]	2020	arXiv	Graph classification	GIN	Graph generation	Inject subgraphs	-	ASR, Clean accuracy, Backdoor accuracy	Twitter, Bitcoin, COLLAB

EVALUATION METRICS

1. One robustness measure is the **Molloy-Reed criterion**, which measures the existence of giant component in the network
2. Other measures (such as shortest path global efficiency metric) are also found, but require higher computation resource

$$\kappa = \frac{\langle k^2 \rangle}{\langle k \rangle} > 2$$

Molloy-Reed Criterion: <k> is the average degree of the network

Our Approach and Assumptions

Main Tasks and Steps

Network Selection

We selected the Gnutella P2P network from Stanford SNAP dataset, consisting of 148K edges and 63K nodes

Algorithm Design

We designed 3 attack and defense algorithms based on randomness, degree, betweenness

Experiment & Analysis

We implemented 9 sets of experiments, each with varying parameters

Main Assumptions

Equal Compute

We assume both attacker and defender have the same computational resources

Equal Access

We assume both attacker and defender have the same access to the network

Equal Magnitude

We assume the same number of edges will be attacked and defended each time

Randomness, Degree and Betweenness Algorithm

Random Algorithm

Randomly destroying and repairing a set of edges from the network

Degree Algorithm

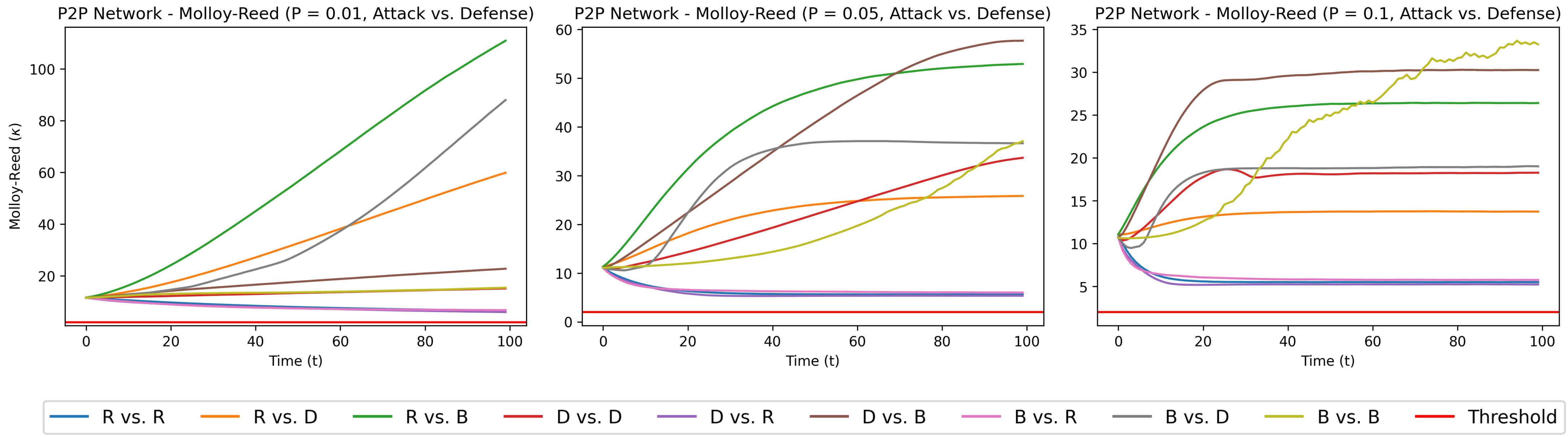
Destroying and repairing the edges based on the degree of the nodes

Betweenness Algorithm

Destroying and repairing the edges based on the betweenness centrality of the nodes and edges

Defense \ Attack	Random	Degree	Betweenness
Betweenness	R vs. B	D vs. B	B vs. B
Degree	R vs. D	D vs. D	B vs. D
Random	R vs. R	D vs. R	B vs. R

Key Results and Findings (Part 1)



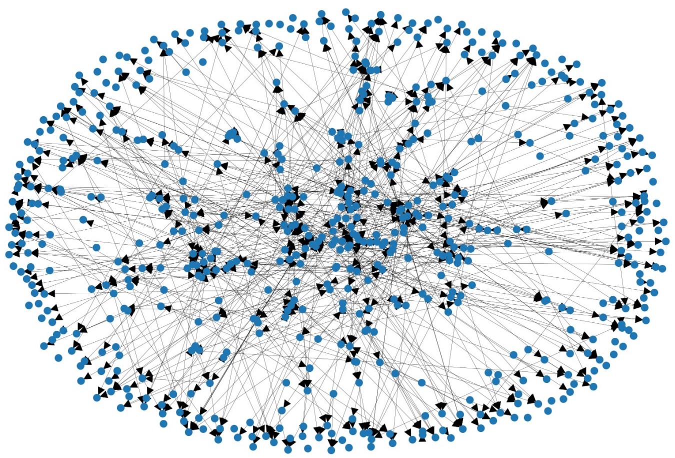
Key Findings

- All defense mechanisms preserve the giant component of the network
- Random defense leads to a decrease in the network robustness as time progresses (blue, purple and pink lines)
- Betweenness defense is the highest performer with the highest network robustness (green, light yellow and brown lines)

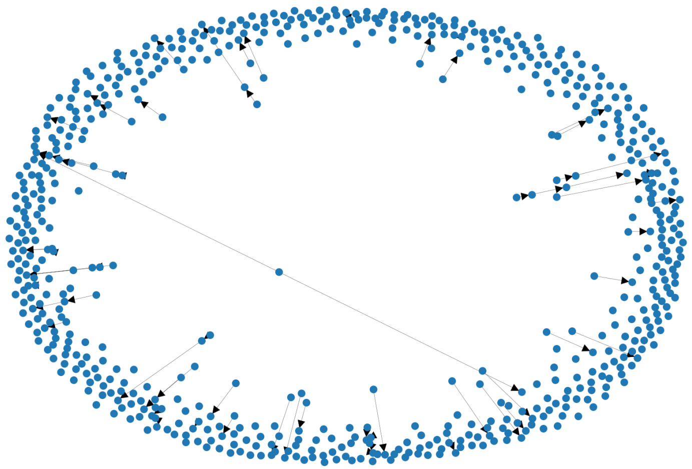
Carnegie Mellon University

Note: Molloy-Reed criterion (red horizontal line) indicates the existence of a giant component is the value is above 2

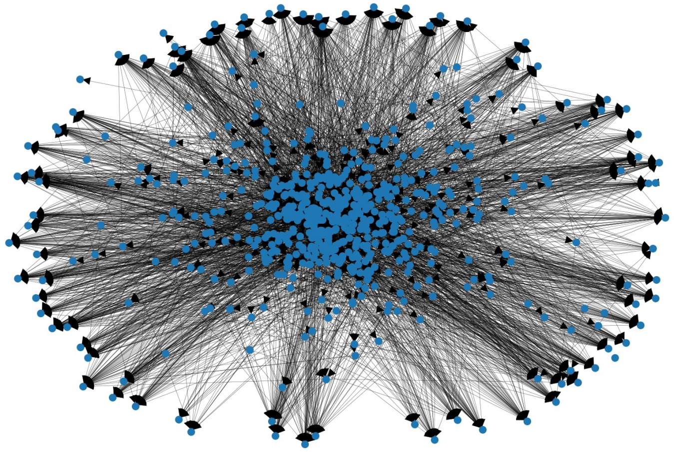
Key Results and Findings (Part 2)



Initial network



Degree attack vs. random defense (p = 0.01)



Random attack vs. betweenness defense (p = 0.01)

Limitation of Our Approach and Outstanding Questions

LIMITATION

1. SNAP Data Set dates to 2002, and we have not applied our analysis to more ***up-to-date networks*** such as TikTok

2. Our algorithm uses brute-force approach to select the edges for attack and defense, more modern-day approach with ***machine learning and deep learning*** was not incorporated

IMPACTS & EXTENSIONS

1. We offered ***3 sets of edge attack and defense algorithms***, which can serve as baseline for future research

2. Deep graph learning and ***reinforcement learning*** approach can be introduced by properly including a reward mechanism

3. More intelligent heuristics about the network can be explored to improve ***computational efficiency***

Division of Labors and Lessons Learned

HOW WE DIVIDED

Tasks	Tony Huang	Vikas Kashyap
Network selection	✓	✓
Attack algorithm design	✓	
Defense algorithm design		✓
Experiment execution	✓	✓
Results analysis	✓	✓
Presentation/report preparation	✓	✓

WHAT WE LEARNED

1. Adverse ***consequences on real-world systems***, such as critical infrastructure, without proper understanding of robustness
2. Rapid development and the ***depth of the research*** on network robustness and adversarial attack/defense

Summary of the Project

1 The world runs on network and an understanding network robustness is key to ensure network's safety

2 This project analyze P2P network's robustness attribute by designing three sets of attack and defense mechanisms

3 Results indicated that betweenness defense provides the highest robustness and the project also identified key limitations of approach and future directions

References

1. **Data source: Stanford SNAP**
   https://snap.stanford.edu/data/#citnets

2. **Network robustness**
   https://barabasi.com/f/619.pdf

3. **Adversarial Attack and Defense on Graph Data: A Survey**
   https://arxiv.org/pdf/1806.02371.pdf

4. **Adversarial Attack on Graph Structured Data**
   https://arxiv.org/pdf/1806.02371.pdf

5. **Network Dismantling**
   https://doi.org/10.1073/pnas.1605083113

6. **A Comparative Analysis Of Approaches To Network-dismantling**
   https://doi.org/10.1038/s41598-018-31902-8

7. **The Robustness of Urban Rail Transit Network based on Complex Network Theory**
   https://doi.org/10.2991/icmemtc-16.2016.211