python .\copy\_of\_before\_the\_learning\_begins!!!\_who\_is\_my\_dataset\_\_for\_twibot\_22.py

torch: imported

Random set by seed = 1

695362 0 False

0 695362 693761 0 695362

30.62 % of nodes which has label but no followr/following relationship

100%|███████████████████████████████████████████████████████████████████████████████████████████████████| 693760/693760 [00:00<00:00, 1677888.96it/s]

2196 693760 693761

100%|██████████████████████████████████████████████████████████████████████████████████████████████████| 3743634/3743634 [00:07<00:00, 510499.19it/s]

Number of nodes: 693761

Number of edges: 3711903

Average in-degree: 5.35

Average out-degree: 5.35

Only in-edge nodes ratio: 0.375

Only out-edge nodes ratio: 0.236

Nodes with both in and out edges ratio: 0.389

Isolated nodes ratio: 0.0

Maximum degree: 2221

Number of leaf nodes: 247132

Number of nodes: 693761

Number of edges: 3711903

Average in-degree: 5.35

Average out-degree: 5.35

Only in-edge nodes ratio: 0.375

Only out-edge nodes ratio: 0.236

Nodes with both in and out edges ratio: 0.389

Isolated nodes ratio: 0.0

Maximum degree: 2221

Number of leaf nodes: 247132

Number of nodes: 612384

Number of edges: 2997878

Average in-degree: 4.895

Average out-degree: 4.895

Only in-edge nodes ratio: 0.375

Only out-edge nodes ratio: 0.197

Nodes with both in and out edges ratio: 0.369

Isolated nodes ratio: 0.059

Maximum degree: 1853

Number of leaf nodes: 200383

Number of nodes: 81377

Number of edges: 36466

Average in-degree: 0.448

Average out-degree: 0.448

Only in-edge nodes ratio: 0.102

Only out-edge nodes ratio: 0.095

Nodes with both in and out edges ratio: 0.065

Isolated nodes ratio: 0.738

Maximum degree: 171

Number of leaf nodes: 12853

Standard Deviation of Degree: 53.5758 and avg is 9.52

Standard Deviation of Degree: 53.5758 and avg is 9.52

Standard Deviation of Degree: 4.3380 and avg is 0.79

Standard Deviation of Degree: 49.3630 and avg is 8.7

In-Degree Std: 24.8680 and avg is 24.87

Out-Degree Std: 40.2681 and avg is 40.27

In-Degree Std: 24.8680 and avg is 24.87

Out-Degree Std: 40.2681 and avg is 40.27

In-Degree Std: 2.5930 and avg is 2.59

Out-Degree Std: 2.7751 and avg is 2.78

In-Degree Std: 22.2668 and avg is 22.27

Out-Degree Std: 37.7379 and avg is 37.74

693761 3711903

Graph is directed.

Sparsity ratio: 0.000008

693761 3301025

Graph is undirected.

Sparsity ratio: 0.000014

693761 3301025

Graph is undirected.

Sparsity ratio: 0.000014

Number of nodes in largest component: 693759

Number of edges in largest component: 3301024

#edge/#node = 4.76

Number of nodes in largest weakly connected component: 693759

Number of edges in largest weakly connected component: 3711902

#edge/#node = 5.35

SCC Analysis for original G:

Total number of SCCs: 425036

Number of SCCs with size < 10: 424976

Sizes of SCCs with size ≥ 10 (sorted): [267250, 38, 36, 32, 31, 28, 26, 25, 22, 22, 21, 20, 19, 17, 17, 16, 16, 16, 15, 15, 15, 15, 15, 14, 14, 14, 14, 14, 14, 14, 14, 13, 13, 13, 13, 13, 13, 13, 13, 13, 12, 12, 12, 12, 12, 11, 11, 11, 11, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10]

SCC Analysis for reversed G:

Total number of SCCs: 425036

Number of SCCs with size < 10: 424976

Sizes of SCCs with size ≥ 10 (sorted): [267250, 38, 36, 32, 31, 28, 26, 25, 22, 22, 21, 20, 19, 17, 17, 16, 16, 16, 15, 15, 15, 15, 15, 14, 14, 14, 14, 14, 14, 14, 14, 13, 13, 13, 13, 13, 13, 13, 13, 13, 12, 12, 12, 12, 12, 11, 11, 11, 11, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10]

C:\Users\u7746112\OneDrive - Australian National University\Desktop\data\_shenasi\twi22\copy\_of\_before\_the\_learning\_begins!!!\_who\_is\_my\_dataset\_\_for\_twibot\_22.py:359: UserWarning: Glyph 10141 (\N{TRIANGLE-HEADED RIGHTWARDS ARROW}) missing from current font.

plt.tight\_layout()

C:\Users\u7746112\OneDrive - Australian National University\Desktop\data\_shenasi\twi22\copy\_of\_before\_the\_learning\_begins!!!\_who\_is\_my\_dataset\_\_for\_twibot\_22.py:362: UserWarning: Glyph 10141 (\N{TRIANGLE-HEADED RIGHTWARDS ARROW}) missing from current font.

plt.savefig(filename, bbox\_inches='tight', dpi=300)

Figure figure\_0.png saved

Figure figure\_1.png saved

Isolated benign nodes in undirected graph of whole nodes: 0, whole benign nodes are 612384

Isolated sybil nodes in undirected graph of whole nodes: 0, whole benign nodes are 81377

Figure figure\_2.png saved

Figure figure\_3.png saved

Figure figure\_4.png saved

Figure figure\_5.png saved

Figure figure\_6.png saved

BFS Level Summary (excluding virtual node):

Level 1: 693761 nodes, average degree: 10.52

Non-reachable nodes: 0 (0.0%)

BFS Level Summary (excluding virtual node):

Level 1: 612384 nodes, average degree: 10.69

Level 2: 75263 nodes, average degree: 8.81

Level 3: 6114 nodes, average degree: 1.04

Non-reachable nodes: 0 (0.0%)

BFS Level Summary (excluding virtual node):

Level 1: 81377 nodes, average degree: 9.22

Level 2: 197849 nodes, average degree: 25.3

Level 3: 414083 nodes, average degree: 2.24

Level 4: 452 nodes, average degree: 1.0

Non-reachable nodes: 0 (0.0%)

# of sybils and benigns in biggest component is 81376, 612383 respectively.

# of sybils and benigns in other components is 1, 1 respectively.

1 component(s) with size 693759, avg edges: 3301024, avg benigns: 612383, avg sybils: 81376

1 component(s) with size 2, avg edges: 1, avg benigns: 1, avg sybils: 1

Figure figure\_7.png saved

# of sybils and benigns in biggest component is 0, 576202 respectively.

# of sybils and benigns in other components is 0, 36182 respectively.

1 component(s) with size 576202, avg edges: 2664142, avg benigns: 576202, avg sybils: 0

36182 component(s) with size 1, avg edges: 0, avg benigns: 1, avg sybils: 0

# of sybils and benigns in biggest component is 20673, 0 respectively.

# of sybils and benigns in other components is 60704, 0 respectively.

1 component(s) with size 20673, avg edges: 31744, avg benigns: 0, avg sybils: 20673

1 component(s) with size 30, avg edges: 30, avg benigns: 0, avg sybils: 30

1 component(s) with size 21, avg edges: 20, avg benigns: 0, avg sybils: 21

1 component(s) with size 19, avg edges: 18, avg benigns: 0, avg sybils: 19

1 component(s) with size 15, avg edges: 14, avg benigns: 0, avg sybils: 15

3 component(s) with size 14, avg edges: 13, avg benigns: 0, avg sybils: 14

2 component(s) with size 13, avg edges: 12, avg benigns: 0, avg sybils: 13

2 component(s) with size 12, avg edges: 11, avg benigns: 0, avg sybils: 12

2 component(s) with size 11, avg edges: 10, avg benigns: 0, avg sybils: 11

1 component(s) with size 10, avg edges: 9, avg benigns: 0, avg sybils: 10

6 component(s) with size 9, avg edges: 8, avg benigns: 0, avg sybils: 9

10 component(s) with size 8, avg edges: 7, avg benigns: 0, avg sybils: 8

2 component(s) with size 7, avg edges: 6, avg benigns: 0, avg sybils: 7

9 component(s) with size 6, avg edges: 5, avg benigns: 0, avg sybils: 6

12 component(s) with size 5, avg edges: 4, avg benigns: 0, avg sybils: 5

9 component(s) with size 4, avg edges: 3, avg benigns: 0, avg sybils: 4

13 component(s) with size 3, avg edges: 2, avg benigns: 0, avg sybils: 3

32 component(s) with size 2, avg edges: 1, avg benigns: 0, avg sybils: 2

60094 component(s) with size 1, avg edges: 0, avg benigns: 0, avg sybils: 1

Tree component with 2 nodes and 1 edges is a Path. Sybils: 1, Benigns: 1.

Connected component 1 has 693759 nodes: Benigns = 612383, Sybils = 81376

(Skipped 0 nodes in small clusters ≤ 5; Benigns = 0, Sybils = 0)

Number of sybil-dominated clusters (sybils ≥3× benigns and sybils ≥30 ): 0

Number of benign-dominated clusters (benigns ≥3× sybils and benigns ≥30): 44

While there are 70 clusters,

Top 7 clusters contain 80% of all sybils.

Top 7 clusters contain 80% of all benigns.

Global Clustering Coefficient: 0.0116

Average Clustering Coefficient: 0.0685

Degree Assortativity: -0.4675

Approx. Diameter: 8.0000

Reciprocity: 0.2214

K-core Max: 80.0000

PageRank Entropy: 12.5016

Global Clustering Coefficient: 0.0116

Average Clustering Coefficient: 0.0685

Degree Assortativity: -0.4675

Approx. Diameter: 8.0000

Reciprocity: 0.2214

K-core Max: 80.0000

PageRank Entropy: 12.5016

C:\Users\u7746112\OneDrive - Australian National University\Desktop\data\_shenasi\twi22\copy\_of\_before\_the\_learning\_begins!!!\_who\_is\_my\_dataset\_\_for\_twibot\_22.py:913: RuntimeWarning: overflow encountered in scalar multiply

gini = (2 \* np.sum((np.arange(1, n + 1) \* sorted\_degrees))) / (n \* np.sum(sorted\_degrees)) - (n + 1) / n

gini is -0.5695126695700021

0.22138401784744913

0.22138401784744913

0.22264815312697847

0.2313387813305545

Avg: 5.1662, Std: 1.0656, Max: 15.00

Degree Entropy: 3.1914

In-Degree Entropy: 3.2434

Out-Degree Entropy: 2.4816

K-shell index distribution:

Shell 1: 315977 nodes

Shell 2: 110537 nodes

Shell 3: 61449 nodes

Shell 4: 39939 nodes

Shell 5: 27557 nodes

Shell 6: 20544 nodes

Shell 7: 16089 nodes

Shell 8: 12684 nodes

Shell 9: 10222 nodes

Shell 10: 8433 nodes

Shell 11: 7014 nodes

Shell 12: 5847 nodes

Shell 13: 5002 nodes

Shell 14: 4291 nodes

Shell 15: 3743 nodes

Shell 16: 3348 nodes

Shell 17: 2941 nodes

Shell 18: 2557 nodes

Shell 19: 2282 nodes

Shell 20: 2051 nodes

Shell 21: 2098 nodes

Shell 22: 1648 nodes

Shell 23: 1516 nodes

Shell 24: 1332 nodes

Shell 25: 1296 nodes

Shell 26: 1169 nodes

Shell 27: 1081 nodes

Shell 28: 936 nodes

Shell 29: 871 nodes

Shell 30: 847 nodes

Shell 31: 846 nodes

Shell 32: 720 nodes

Shell 33: 680 nodes

Shell 34: 704 nodes

Shell 35: 609 nodes

Shell 36: 593 nodes

Shell 37: 544 nodes

Shell 38: 487 nodes

Shell 39: 503 nodes

Shell 40: 544 nodes

Shell 41: 504 nodes

Shell 42: 573 nodes

Shell 43: 403 nodes

Shell 44: 378 nodes

Shell 45: 456 nodes

Shell 46: 462 nodes

Shell 47: 403 nodes

Shell 48: 377 nodes

Shell 49: 464 nodes

Shell 50: 415 nodes

Shell 51: 466 nodes

Shell 52: 372 nodes

Shell 53: 297 nodes

Shell 54: 336 nodes

Shell 55: 380 nodes

Shell 56: 323 nodes

Shell 57: 432 nodes

Shell 58: 194 nodes

Shell 59: 414 nodes

Shell 60: 292 nodes

Shell 61: 255 nodes

Shell 62: 416 nodes

Shell 63: 437 nodes

Shell 64: 836 nodes

Shell 65: 176 nodes

Shell 66: 143 nodes

Shell 67: 59 nodes

Shell 68: 649 nodes

Shell 69: 69 nodes

Shell 70: 24 nodes

Shell 71: 74 nodes

Shell 72: 48 nodes

Shell 73: 68 nodes

Shell 74: 55 nodes

Shell 75: 63 nodes

Shell 76: 88 nodes

Shell 77: 71 nodes

Shell 78: 57 nodes

Shell 79: 91 nodes

Shell 80: 610 nodes

Influence Concentration (Gini over PageRank): 0.5488

Periphery Edge Ratio: 0.0957

100%|██████████████████████████████████████████████████████████████████████████████████████████████████████████████| 200/200 [06:29<00:00, 1.95s/it]

100%|██████████████████████████████████████████████████████████████████████████████████████████████████████████████| 200/200 [04:48<00:00, 1.44s/it]

Approx Avg Path Length (Real): 4.2842

Approx Avg Path Length (Rand): 6.2290

Approx Rewiring Probability: 0.6878

kglkjglsd

(base) PS C:\Users\u7746112\OneDrive - Australian National University\Desktop\data\_shenasi\twi22>