PreCog Task 3: ATLAS

Sudhanva Joshi - 2023102022
ECE, IIIT Hyderabad

***Abstract*—This report details the analysis of a directed ALTAS Graph on various parameters in an attempt to correlate them to real-life strategies. The pipeline to test this model has been expounded upon below.**
Note: All corresponding files are included in their respective folders.

I. Graph Properties

Establishing a crucial *Win-Condition* can be defined as follows:
A player wins if they are able to reach the node which only has neighbouring nodes with 0 outdegree.

A. Degree

As mentioned prior, Node *Degrees* is key in developing and reaching a winning strategy for the game.

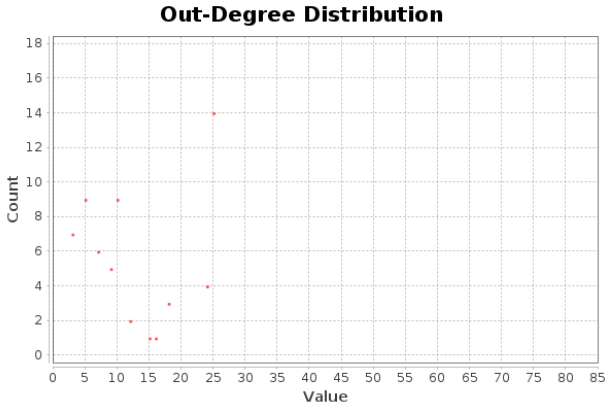

Fig. 1: Country Only

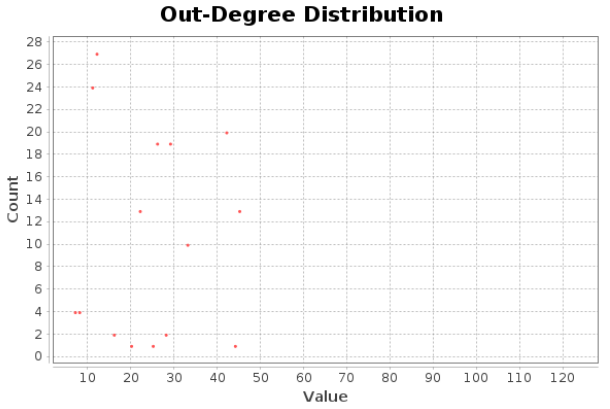

Fig. 2: City Only

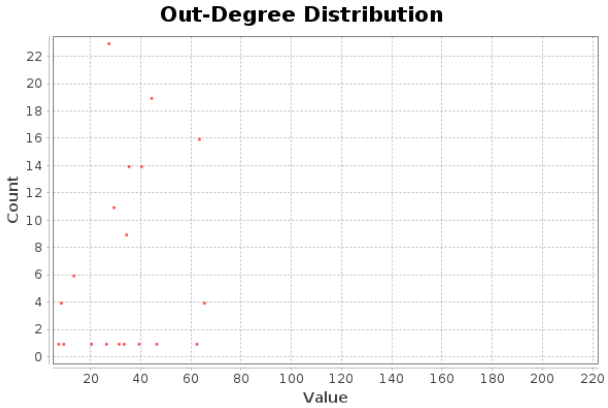

Fig. 3: Country+City

It is evident that as the sample sample size increases, the distribution of the outdegree of the nodes tends to follow a power-law distribution, which indicates that the game path may 'fan out' more as the game progresses, making it difficult to corner the opponent.

B. Centrality

We can lock the opponent out of a cluster of nodes by observing the *Centrality* of nodes. H

- **Betweenness** - The number of shortest paths that pass through a node *(Bottleneck)*.
- **Closeness** - The average distance of a node to all other nodes *(Options to Move yourself)*.
- **Eccentricity** - The maximum distance of a node to all other nodes *(Distance from Objective-Low Outdegree neighbours)*.

C. Diameter

The obtained values for the *Diameter* of all 3 graphs are as follows:

$$d_{country} = 7,\ d_{city} = 6,\ d_{country+city} = 5$$

The diameter of the combined graph is the lowest, which indicates that the game may be more 'compact' and the player may have to make more strategic moves to win.

Note: Simulated using *GEPHI*.

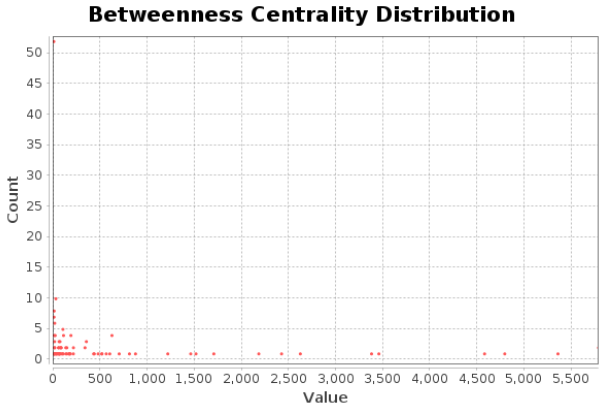

Fig. 4: Country Only

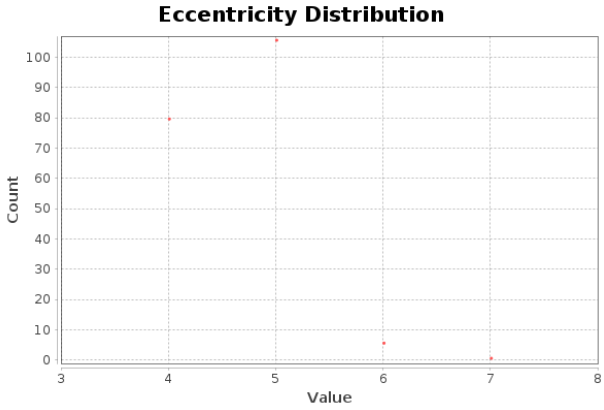

Fig. 7: Country Only

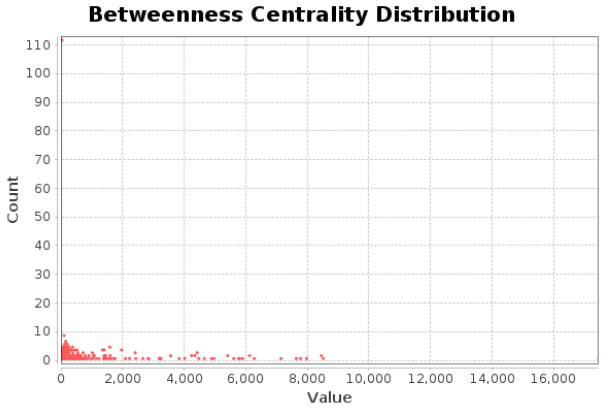

Fig. 5: City Only

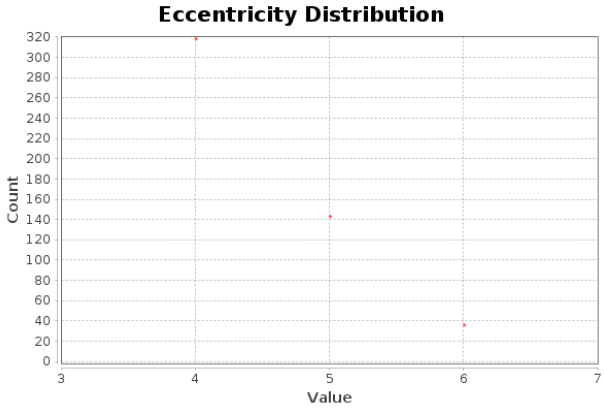

Fig. 8: City Only

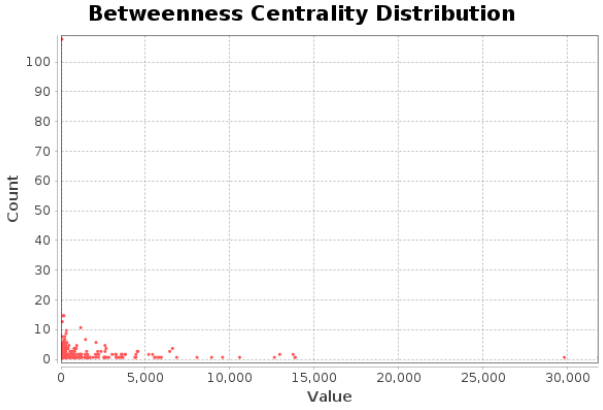

Fig. 6: Country+City

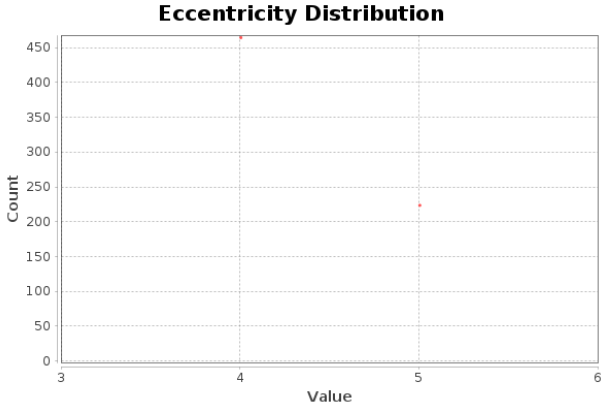

Fig. 9: Country+City

The addition of more chokepoints to defend or 'capture' *(traverse first)* makes it more difficult to establish control over the opponent.

Eccentricity lowering as the sample size increases implies that the game map is smaller; the player can reach *(or force their opponent)* towards fringe nodes faster.

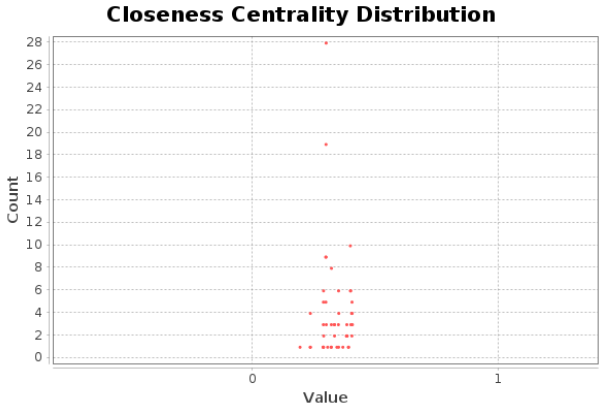

Fig. 10: Country Only

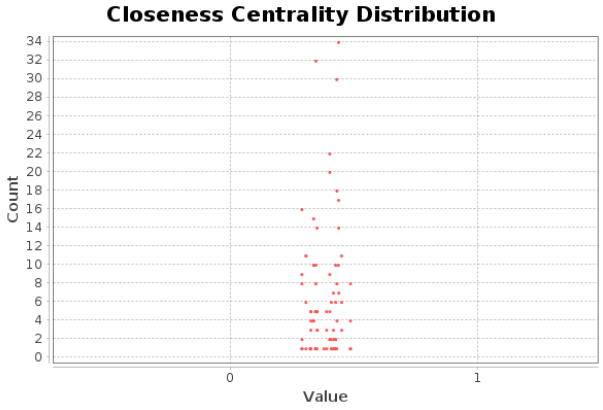

Fig. 11: City Only

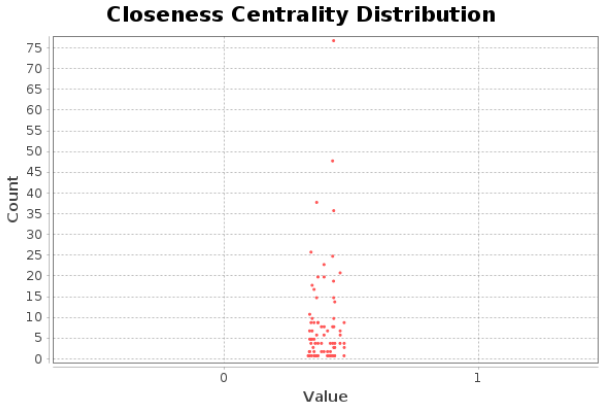

Fig. 12: Country+City

Tighter spread of scatterplot implies that the network is more densely connected, which makes it easier to manuveur yourself.

II. Clustering

Using **Modularity Classes** to identify and visualize clusters of *Densely Connected Components* in the ATLAS Graph:

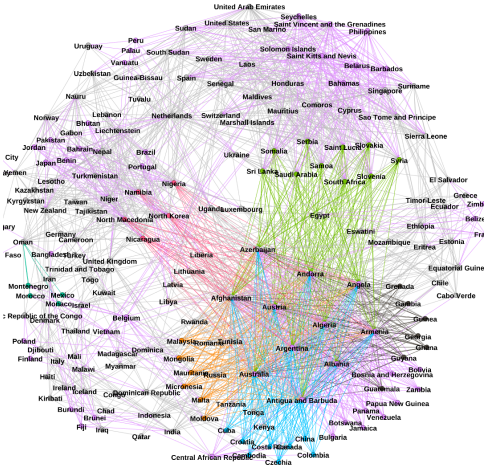

Fig. 13: Country Only

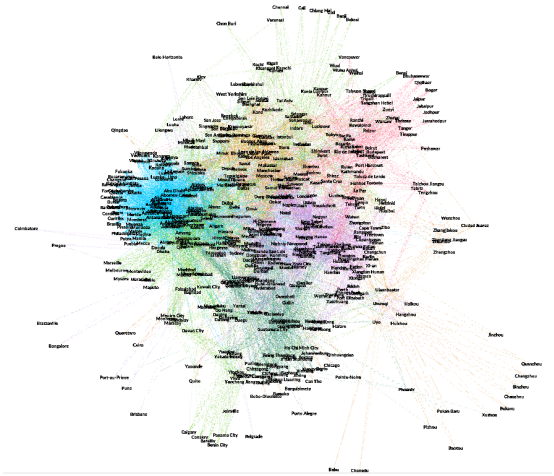

Fig. 14: City Only

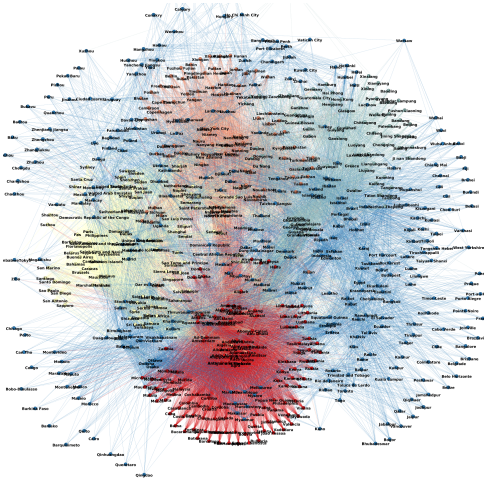

Fig. 15: Country+City

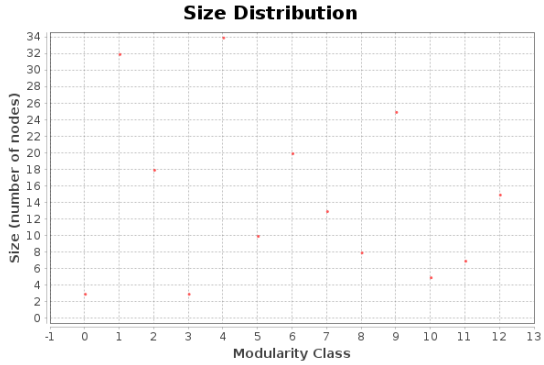

Fig. 16: Country Only

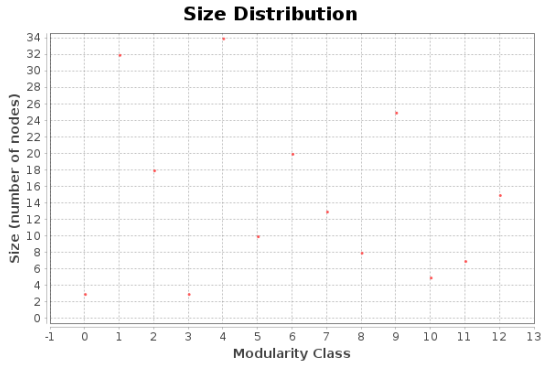

Fig. 17: City Only

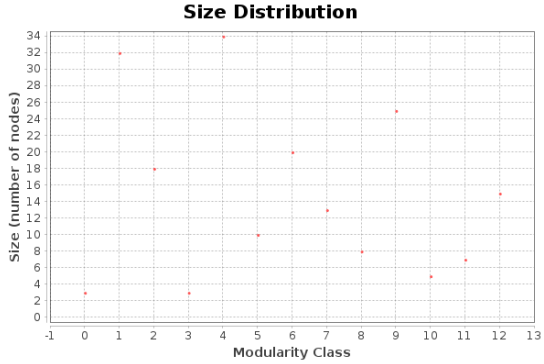

Fig. 18: Country+City

- **Country & City Graph**:
  - High modularity; clusters = regions/cultures.
  - Focus on intra-cluster chains; bridging between clusters is harder.
  - Moderate modularity; clusters = geographic/economic ties.
  - Exploit well-connected clusters; transitions require high-betweenness words.
- **Country + City Graph**:
  - Lower modularity; more interconnected.
  - Longer chains possible; increased competition for central nodes.

Link Prediction

Given below is the comparative result for both the *Node2Vec* and *PyG* approaches to solving the *Edge Masking* problem on the ATLAS Graph. **Note:** 20% of Edges were Masked during testing.

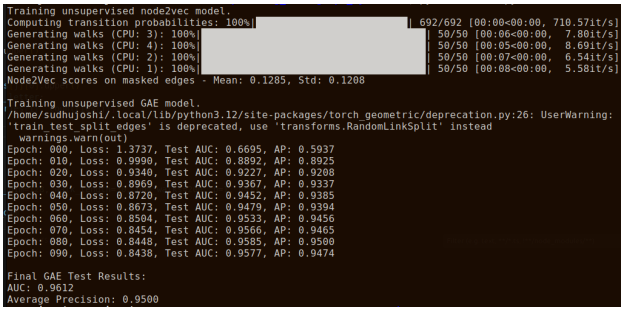

Fig. 19: Result for Link Prediction, 693 Nodes

Parameter Results

- **Mean and Std**:
  - **Mean**: Average performance score across runs.
  - **Std**: Variability of results; low std = consistent performance.
  - High mean + low std = good and consistent performance.
- **AUC (Area Under the Curve)**:
  - Measures ability to distinguish between existing and non-existent edges.
  - Range: 0 to 1 (1 = perfect, 0.5 = random).
  - High AUC = effective edge prediction.
- **Average Precision**:
  - Summarizes precision-recall curve for edge prediction.
  - Range: 0 to 1 (1 = perfect).
  - High average precision = accurate predictions with few false positives.

It is observed that *Node2Vec* trades accuracy for speed due to the usage of **Random Walks**, while *PyG* is sluggish but more accurate due to the usage of **Tensors**.

References

- Ulrik Brandes, A Faster Algorithm for Betweenness Centrality, in Journal of Mathematical Sociology 25(2):163-177, (2001)
- Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, Etienne Lefebvre, Fast unfolding of communities in large networks, in Journal of Statistical Mechanics: Theory and Experiment 2008 (10), P1000
- R. Lambiotte, J.-C. Delvenne, M. Barahona Laplacian Dynamics and Multiscale Modular Structure in Networks 2009
- Robert Tarjan, Depth-First Search and Linear Graph Algorithms, in SIAM Journal on Computing 1 (2): 146–160 (1972)