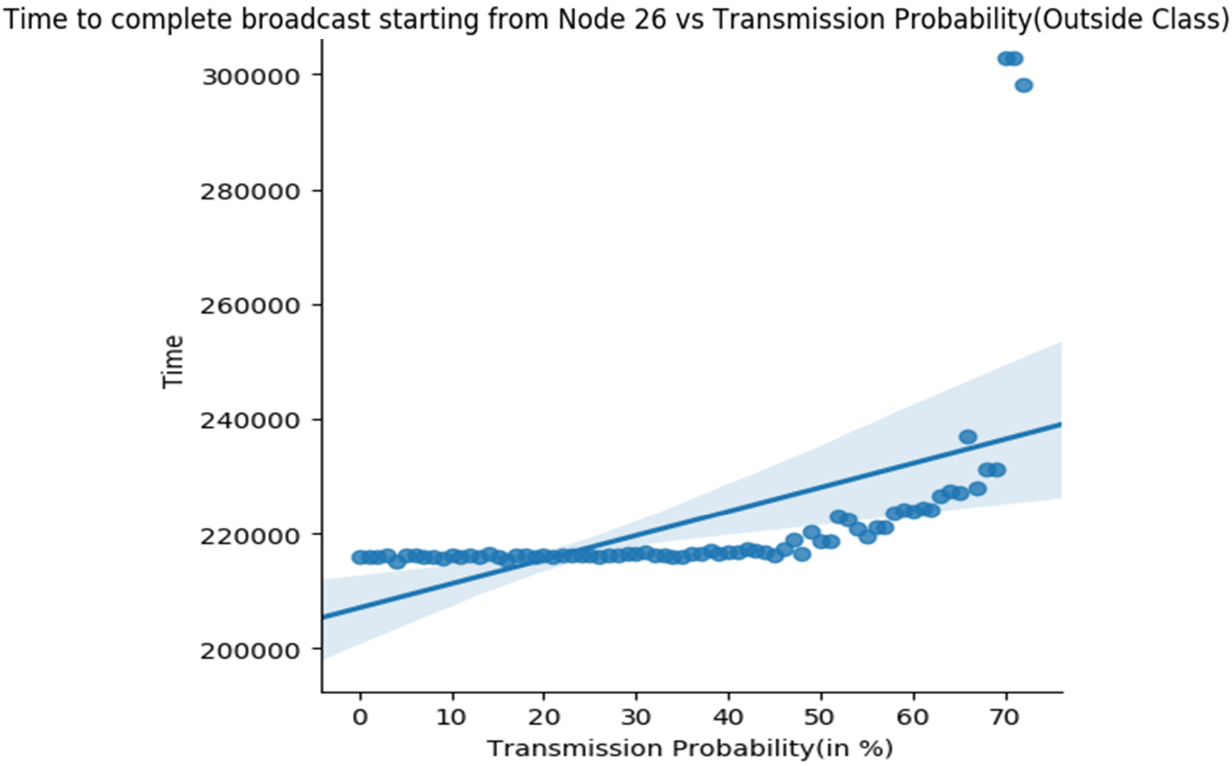
Part 3 Analysis

Modularity measures how well a **network decomposes** into modular **communities/subnetworks**. Classes represent **tightly knit communities** which have a lot more **interconnections** between their nodes than with nodes in other communities.



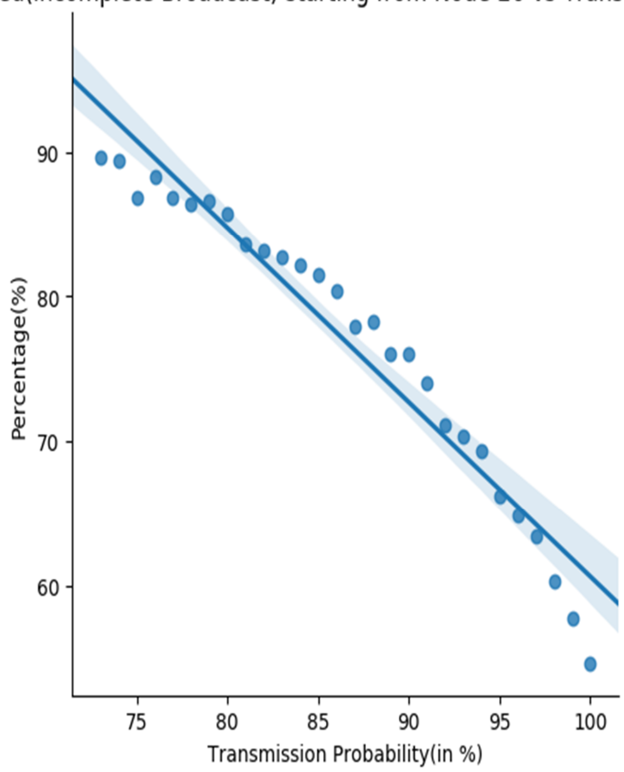
Starting broadcast from **node 26**, we observe that time for broadcast completion remains **nearly constant till X=40%**, and thereby starts rising, so much so that broadcast **fails to complete** for transmission probability **above ~70%**.

<u>Insights-</u> To understand this effect we analyse what a community class is. As mentioned, it represents nodes **tightly knitted**, hence a node has **many of its neighbour** in this class. As X increases, transmission probability inside the class starts to fall. As such, **expected transmissions** inside a community class experiences a **plummet**.

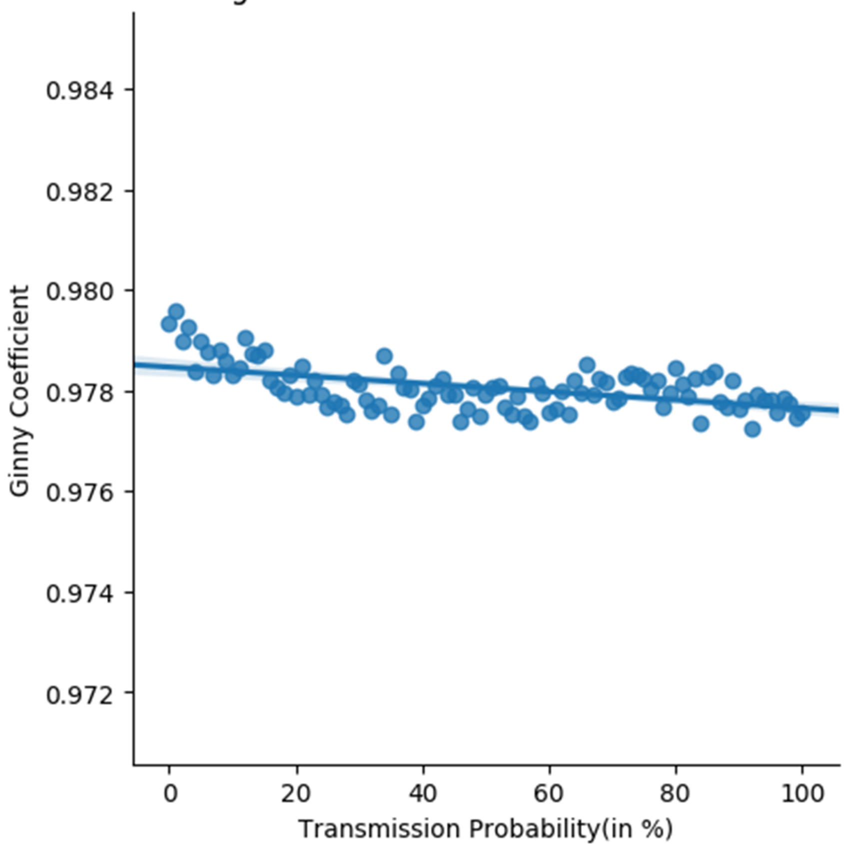
One may argue that the class might see an **increased inter-class communication**, but modularity classes are such designed that **intra-class communications** are much larger, and decreasing intra-class transmission rate is decreasing in an **expected sense**, the **potential** of the network to transmit. **Average number of transmissions** go down and graph reflects it.

A **linear regression** fit shows similar trend.

Percentage of Nodes Reached(Incomplete Broadcast) starting from Node 26 vs Transmission Probability(Outside Class)



Ginny Coefficient starting from Node 26 vs Transmission Probability(Outside Class)

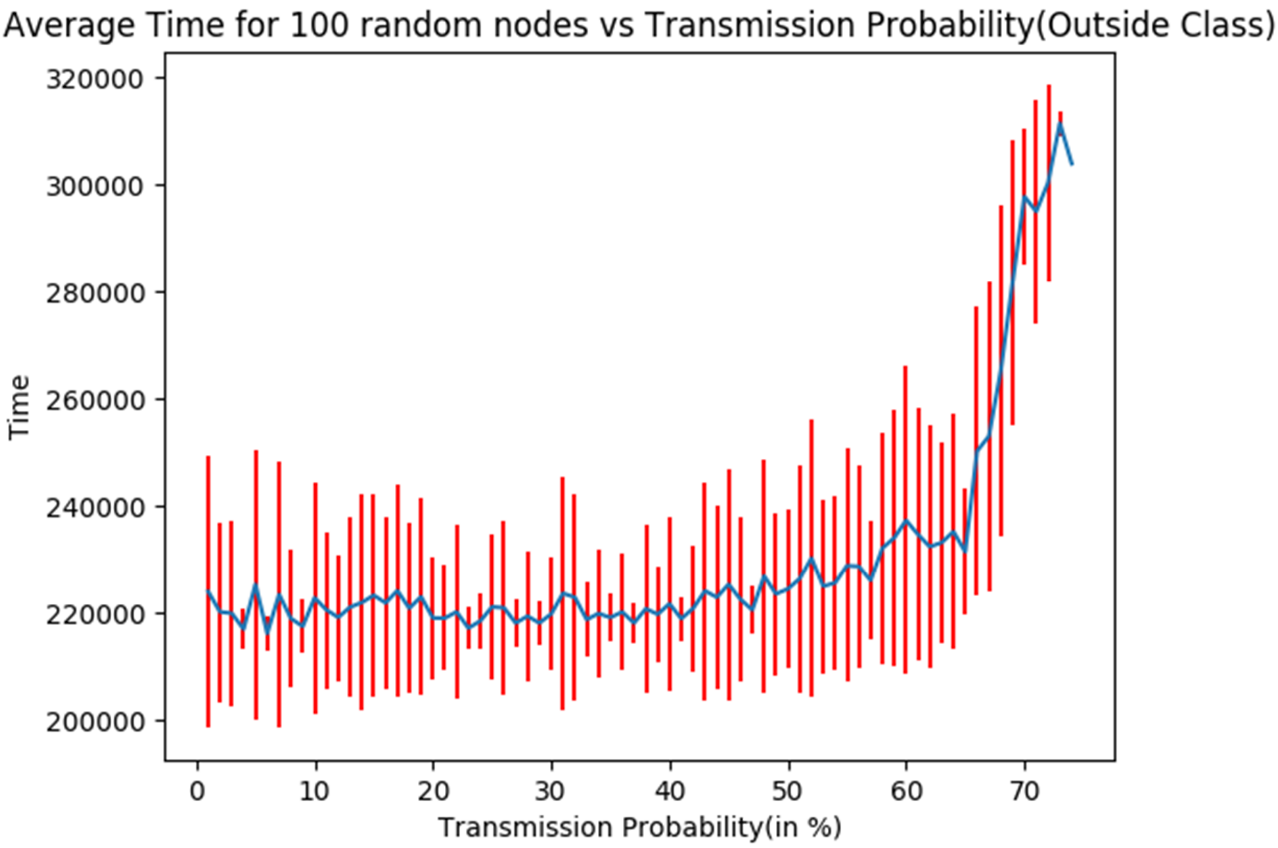


Former claims are **reinforced** when we analyse the graph of percent of nodes reached in case of **incomplete broadcast**. Increasing inter-class communication probability here too, **decreases network potential** to communicate and transfer data. The fall can be quite fairly modelled by a **linear fit as apparent**.

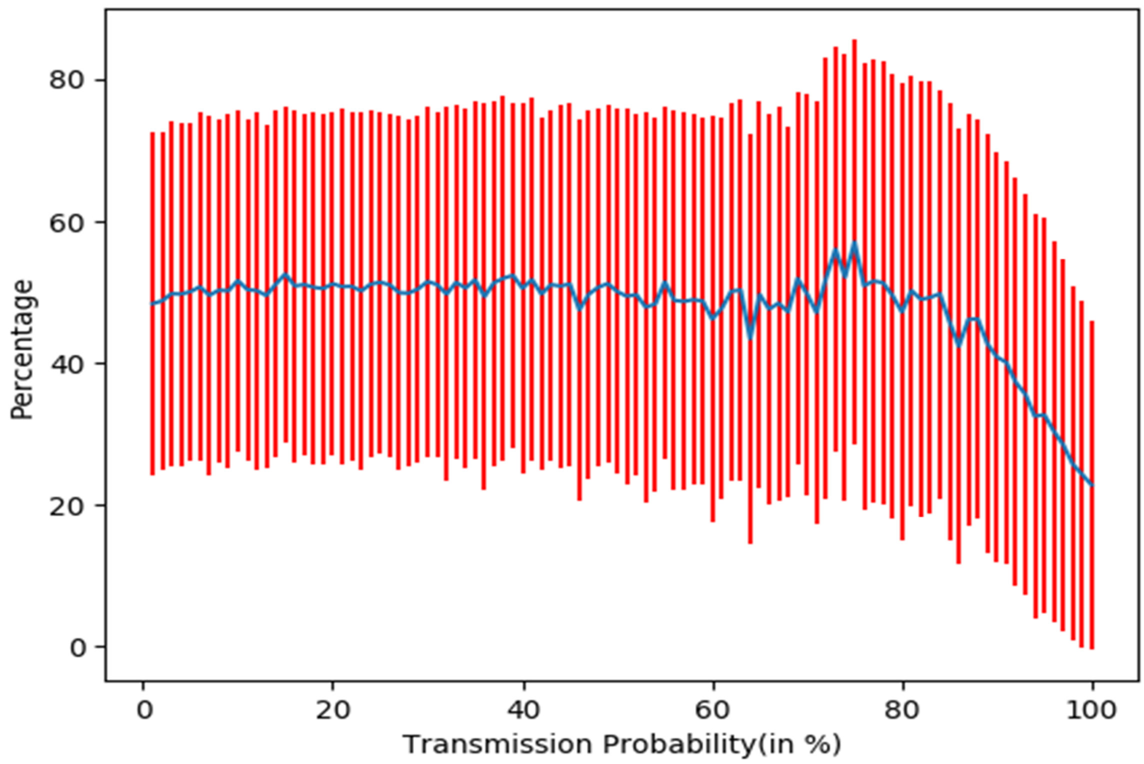
Coming to **gini co-efficient**, we realize that though its value has a very low standard deviation, it has a **slight decreasing trend**. This falls in line because increasing X is taking away networks potential to transmit, so in a way, it's actually **decreasing the stress** as stress is only generated if we work!

So, as X is increased, the intense intra-class communication **starts easing** and **loose** inter class communication become prominent. The **perpetual high** gini coefficient value is because some nodes having **large network coverage** (strong/super) as compared to most with **only 1 neighbour**.

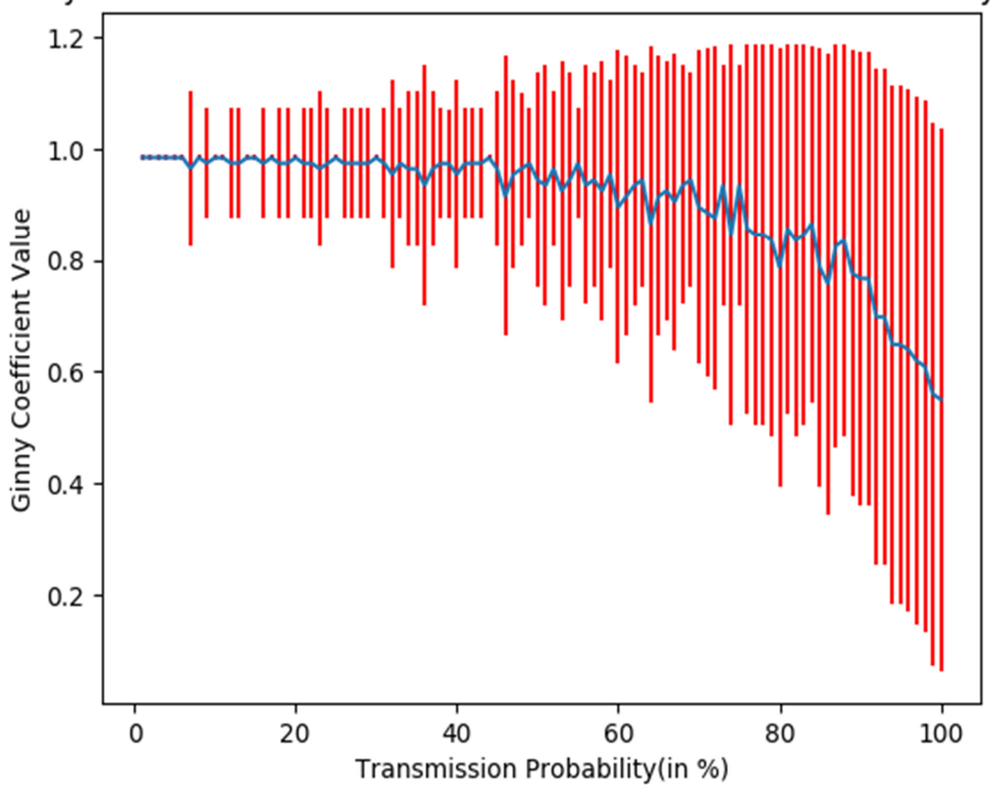
Average Graphs for 100 Random Nodes



Average Percentage of Nodes Reached(Incomplete Broadcast) for 100 random nodes vs Transmission Probability(Outside Class)



Average Ginny Coefficient for 100 random nodes vs Transmission Probability(Outside Class)



For finding the ideal combinations, different graphs were analysed.

The ideal combination appears when X is in **range - (25, 33) %**. Around this value of X, node 26 sees a **global minima** in broadcast time and Gini coefficient is **low** around the range. Also, when experimented with 100 different nodes, average broadcast time graph showed a **global minima** in similar range of its **weakly parabolic** nature. Average percentage of nodes reached in cases of incomplete broadcasted showed **better numbers** in this range. Average Gini for 100 nodes also displays a **fall** in this range.

The average graphs for 100 different nodes test were **in tandem** with observations for node 26. Ginny was a **slight offset**, when in average sense, it **falls sharply** for larger X values, showing how important is **intra class transmissions** in such a **disjoint-union class network**.

Insights behind the combination obtained- Although experimental, we can provide some **intuition** for the ideal combination obtained. Clearly we realised how important is intra-class communication for a network. This is also apparent when X is very low, where the network is still potentially able to complete broadcast. But, that does not give best performance as a community might receive a **message too late**, causing delayed transmissions, and at times, **broadcast failure** as well. As such, the ideal comes around a range, where although intra-class transfers are given **high priority**, inter-class transmissions are not **completely neglected**, and hence the best performance.

Comparison between Algorithms

On comparing across methods, we observe that **part 2 combinations** with low S (0.5%) and high L (85-90%) were ideal, with **high transmission** probability, as because gini co-efficient and time, both were at minimum. On contrast, we **couldn't find** such clear balance between the 2 in 3rd part.

In 3rd case, as gini co-efficient seems to plummet, time to broadcast see a sharp rise, and ideal combination occurs somewhere midway the 2. The equilibrium ideal point combination is not as clean as in 2nd case where one find system minimum gini and time for nearly same values.

Comparing with first method makes little sense as the first one is **deterministic** whereas the other two are **probabilistic**.

Also, keeping a same cut value K for each node **is not fair** as some node never reach such a threshold, whereas others might easily exceed it. Rather introducing a notion of probability in a sense, introduces individual k, which seems more realistic in long run.