1 "Entropic Centrality for non-atomic Flow Networks"

1.1 Introduction:

Given a graph to analyze, the vertex centrality is a notion meant to identify the most important vertices within a graph. The closeness centrality gives the highest weight to the vertices that are closer to all than the others. The betweenness centrality characterizes the number of times a vertex acts as a bridge along the shortest path between two other vertices. However, the authors take an information theoretic approach to quantify the importance of a node in this work. They also generalize their work by extending it to non-atomic flow graph instead of traditional split and flow graph.

1.2 Author's Contribution:

FIrstly, T=the authors define flow model which allows the flow in a graph to split in multiple direction and based on that they generalize their idea of entropic centrality. Secondly, a case of entropic centrality gives a new and more powerful interpretation of the centrality measure. Finally, The authors provide real world examples and explore them to measure the applicability and limitations of the entropic centrality measures.

1.3 The split and Transfer Flow model:

Let d : V $\rightarrow$ R be a function such that $\sum_{v\epsilon V} d(v) = 0$. It is called a demand function. Vertices for which d(v) = 0 are called transit vertices. By a flow, we mean a function f : E $\rightarrow$ R with the property, called flow conservation, that $d(w) = \sum_{v,(w,v)\epsilon E} f(w,v) - \sum_{v,(w,v)\epsilon E} f(v,w)$ for all $w\epsilon V$

1.4 Notion of Entropic Centrality:

Now to every vertex v, they attach a probability that describes how the flow that arrives at v will behave next. Now, we want the flow to not only to be transferred from one vertex to another but also to be possibly split across different vertices. The probability of a vertex having a flow starting at u and ending at w is: $P_{uw} = \sum_{p\epsilon P_s} \prod_{(w,v)\epsilon P_s}^{n} \tau P_v(V) \frac{f(V',V)}{|S(P_v)|}$. The entropic centrality $C_H(u)$ of a vertex is: $C_h(u) = \sum_{w\epsilon V} p_{uw} \log p_{uw}$

1.5 Applications:

There are several examples of networks for which the flow splits. The bitcoin transaction network is one of them. In this network, the vertices are users identified by their address, and transactions occur among users. A transaction with no input is a mining operation. The authors took the Bitcoin transaction history for the month of December 2016 and derived the transaction graph. They chose two connected components which includes 5,206 and 5,251 vertices respectively, to study.

1.6 Conclusion:

The authors generalized the notion of entropic centrality to non atomic flows and showed that this metric could give exciting and meaningful results that other centrality measures fail to derive.

1.7 Discussion on presented questions:

Q1 :In the first network(figure 2, slide 12), we see a lot of vertices with a high entropy centrality. What do you interpret about the structure of the graph from this data?

If there is a high entropic centrality, it means there are major flow activity in those parts of the graph which means that the component as a whole must be a major activity point. This usually happens in case of exchanges and marketplaces in the web.

Q2: A large out-degree would help the vertex spread the flow widely, which would lead to a higher entropy centrality. Is this argument always true?

No, it is not true. The flow may not be evenly distributed among the outgoing edges. If majority of the outgoing edges have low probability assigned, then the entropic centrality would be low.

2 "The Graph Structure of Bitcoin"

2.1 Introduction:

In this paper, the authors analyze the bow tie graph structure, attributed to the Web in the seminal paper. They link the connectivity structure of the Bitcoin users graph to the economical activity of its nodes. The bow tie structure groups the nodes of the graph depending on their reachable sets of nodes. The nodes in the biggest strongly connected component are called SCC. The remaining nodes reaching (resp. reached by) the ones in the SCC are called IN. The remaining nodes in the biggest weakly connected component are called TUBE, TENDRIL, or FRINGE. Other nodes of the graph are called DISCONNECTED. In the graph proposed by the authors, nodes represent users and edges model the flow of value between such users.Nodes of Graph are augmented with their balance, and the edges are weighted according to the Bitcoin value exchanged. We have access to the creation dates of the edges, which allow us to perform temporal analysis.

2.2 Contribution by authors:

- This paper gives thorough description of data acquisition and formal definitions of all entities involved which makes it easier to understand.
- The deanonymized node analysis is very insightful as it shows that the known entities in the IN group was from a miners pool and the bitcoin exchange and hubs mostly are grouped in the SCC component.
- A clear advantage of building a graph from historical data is that it would make it possible to perform a temporal analysis of the graph structure and to study the evolution of its components. The temporal analysis shows the size of components and number of nodes in components grow steadily over time. The balance analysis is also interesting as it supports authors conjecture that value mostly tends to get exchanged inside SCC.
- Temporal analysis of in and out dimension analysis was interesting as out shows rapid growth compared to ins.

2.3 Future Directions:

The economical meaning of the nodes in the different components of the bow-tie can be further investigated by exploiting more sophisticated deanonymization techniques. This is a research that will be more focused on the privacy of the cryptocurrency transaction issue. The authors are currently performing the same analysis for the graph obtained from the Ethereum blockchain, with the goal of comparing the economies of the two cryptocurrencies. However, it should be noted that usage of cryptocurrency is often backed by the use-case it provides, such as ZCash is exclusively used by mostly privacy aware users and Ripple is mostly used by financial institutions. Therefore, comparing two cryptocurrencies based on economy might not provide comprehensive analysis based on context.

2.4 Conclusion:

The authors analyzed the bow-tie structure of the bitcoin graph and investigated the graph from different perspectives including temporal analysis.

2.5 Questions presented by the presenter:

Q1: We saw that the nodes in the SCC have the highest average value for the measures introduces by the authors(Figure 3 slide 26). what does this data represent in terms of economic activity?
It tells us that major economic activity is happening within the SCC which means it can be an exchange, pool or online marketplace where transactions are always taking place.
Q2: What does the evident difference between the first two columns in figure3(slide 26) tell us?
They tell us that number of coinbase outputs to clusters have been greater than number of unique clusters receiving coinbase outputs. Which is evident as the addresses who receive the coinbase transactions may belong to the same cluster, which will allow the unique cluster number much lower.