Examinations of network dynamics through random walks

Peter Larsen, Nathaniel Ogilvie

November 2020

The below is in the format of the final paper, but with only some of the sections filled out. As we get more results, and get more input, we will continue to update the final paper.

1 Introduction

“A drunk man may find his way home, but a drunk bird will be lost forever.”(Shizuo Kakutani) Most often, when studies of random walks are conducted, one maintains a constant degree/dimension throughout the entire process. The ”drunk man”, a random walk in two dimensions, stays in two dimensions, save for perhaps falling down the stairs. The ”drunk bird”, a random walk in three dimensions, similarly stays in three dimensions. But what happens when dimension is non-uniform? What happens when the dimension changes for every step? How does a random walk perform when the space it exists in is finite? Answers to these questions and more are the goal of this paper. Below, we will attempt to see whether shifts in dimension/degree every time step changes the overall distance one can travel. We will also attempt to influence network coverage by choosing specific starting points based on criteria explained later. Finally, we will attempt to guide a random walk through a network from a chosen point to a destination using a paired network.

For this project, we will be using two networks: Peer to Peer (referred to as P2P) and Group to Group (referred to as G2G). These networks were created using an bipartite adjacency matrix of characters in the Marvel Comic Book Universe and the groups and organizations they belong to, and applying matrix multiplication to them to create unipartite adjacency lists. The matrix was constructed by hand by Peter Larsen, and the information was gathered from the Marvel Comics Database at: https://marvel.fandom.com/wiki/Marvel_Database. The concept for creating this matrix was taken from Kieran Healy’s work here: https://kieranhealy.org/blog/archives/2013/06/09/using-metadata-to-find-paul-revere/.

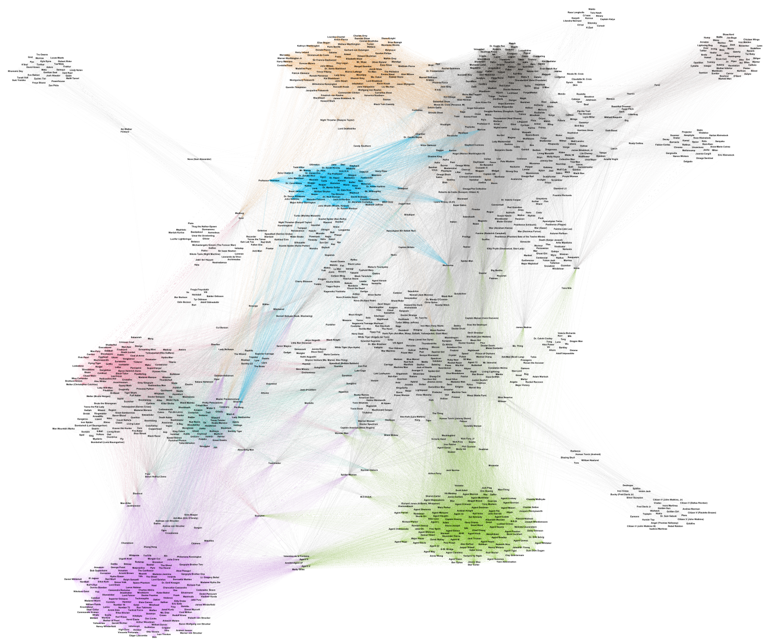

Figure 1: An image of the P2P network

2 Non-standard degree dynamics

2.1 Explanation of Set-up

Using a simple three-dimensional random walk as a comparison, we will begin each run at a random node. Then for 100 time steps, we will walk through the network, recording each node visited. This will then allow us to compare: by examining the number of unique nodes visited divided by the number of nodes it was possible to visit (ideally, this would the identical to number of time steps plus our original node), we would arrive at a coverage fraction, or the percentage of nodes it was possible to visit. We would then perform this action 1000 times to smooth over any outliers, and average the fractions, and compare them. This would allow us to see how uniformity informs the movement of the random walk. We could then examine the lists of nodes visited to determine how many of the uniform dimension trials returned to the origin, and how many of the non-uniform dimension trials did the same. These two comparative measures, being Rate of Return (RoR) and Coverage Fraction (CF) will allow us to say whether a non-uniform distribution of degree/dimension alters a random walk's mathematical behavior or not.

2.2 Results

3 Random walks using fixed starting positions

3.1 Explanation of Set-up

Centralities allow us to determine the most important vertices/nodes in a network, but how do they affect random walks? Using the two measures described above in RoR and CF, we will take the top 5 nodes with the highest betweenness centrality, the middle 5, and the bottom 5, and perform the dynamics tests described above: 1000 trials, 100 time steps each, but this time, beginning at the specific nodes. Then we will average the results together to see if betweenness centrality has any affect on a random walk's mathematical behavior or not. With luck, we will also do the same with eigenvector centrality.

3.2 Results

4 Co-dependent networks operating in tandem to perform a task

4.1 Explanation of Set-up

The goal of this final part is to choose 2 groups, an origin group and a destination group, and then choose a node in the origin group such that it does not belong to the destination group and denote it as $v*$. We will then take the adjacency list of $v*$, and weight each of its neighbors based upon their interaction with the destination group. The G2G network provides us with the strength of interaction between each group. Using these weights, we will alter the adjacency lists, and then randomly walk until we hit the target group. We will then compare this to the shortest path calculation between the two groups, and see how alike they are.

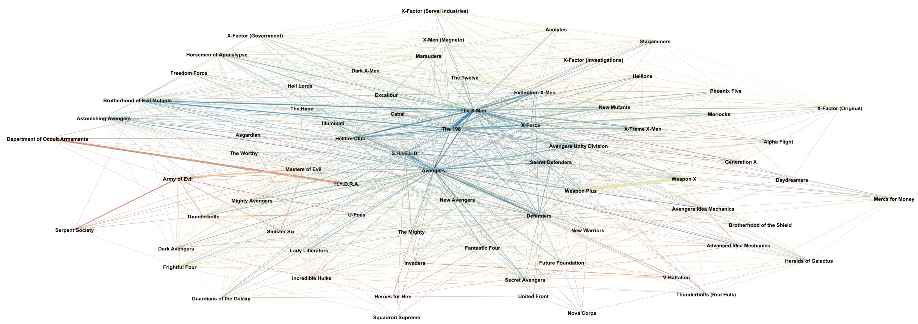

Figure 2: An image of the G2G network

4.2 Results

5 Code

6 Works Cited