

Austin Kelly, Brandon Kindrick, Luis Montes
CS499 Network Analysis
21 Feb 2018

Project Proposal

The problem our group seeks to solve is finding the most relevant board games that a user would be interested in compared to any given game they already like, based on mechanics of the game or based on the category (theme). In this network, the nodes would be board games and the edges would be the containing of a tagged mechanism or category node.

This project is important because it can help individuals find board games relevant to their play style or to their interests. We can analyze the top board games on boardgamegeek and look for common trends in category(theme) and or mechanics (ways the game is played). Advertisers (or boardgamegeek) could use our results to better suggest games to a user based on their interests or games they are currently viewing on the site.

To approach this, we are going to use boardgamegeek's database and their API. This database of board games contains over 15000 different board game entries, which is too large for the scope of our project. To make our project smaller in size and more manageable we decided to analyze the top 1000 board games. These top 1000 games are provided by board game geek and are based on the average geek rating. This data will be split into two sets, one subnet of board games connected by category and one subnet of board games connected by mechanism. Recently, a JavaScript API was developed that collects XML from the site and turns into the more flexible JSON format and from game JSON objects we can draw all kind of conclusions for our question. This API named "bgg" is hosted on npm.org, the Node.js package manager, and is very easy to use within the Node.js runtime environment. We will use this API for sampling purpose.

The algorithm for network analysis we plan to use extensively for determining the most similar games to each other either mechanically or categorically is the Jaccard Similarity Index. For our uses, a board game will be represented by a node connected to many different category nodes and different mechanism nodes, but similarity discovery is completely separate for desiring mechanically similar games or categorically similar games. Take discovering mechanically similar games, for example. If we have a single game with six mechanism tags, than that single game is handled as a set with six mechanism properties, and we then compare the overlap of that set to every of other game's set of tagged mechanism properties. Sets that are 90% similar (our chosen Jaccard Similarity Index) or more in mechanism properties then have their respective board game nodes clustered together as games that a user could enjoy based on similar mechanics. The same methodology applies for category discovery. The end result would be a single game node having a clustering of none (e.g. the game had not high enough overlap in mechanism or category tags with any other game) or more games that are similar to itself in mechanics or categories. For example, we might take the find the union First, we will apply these

algorithms to multiple game qualities such as genre price. For visual representations, we will use python's networkx library to display graphs, as we had some success using this in class.

We will consider the project a success if we complete each task required while also having data representation we are completely satisfied with aesthetically and true to the Jaccard Similarity Index method. If we can successfully find similar games based on category or mechanisms using our method we will be satisfied as a whole.

By the end of the semester, we will have created graphs to visually represent similarity in board games, and used the algorithm mentioned above to generate similarity clusters and present them in a professional way. Our results will be expressed in detail on the project write-up and website. A stretch goal of ours is to share our results and program with boardgamegeek.com's development team in hopes of improving their game suggestion to users. Overall, we hope to learn rigid discipline for gathering professional research, while finding out interesting information about board games along the way.

Milestone Timeline

- Bi-weekly group meetings throughout the rest of the semester will be held to stay productive, Sunday and Wednesday around 4pm.
- Project Proposal completed on ~~2/12~~ 2/21
- Data Collected in mid/late March before Milestone report
- Data Analyzed in mid/late March before Milestone report
- Milestone Report completed in mid/late March
- Website completed by mid April
- Final project report completed by mid/late April