

## MiniProject 2: Pandemaniac++

Assigned: 03/01/2023

Due: 03/12/2023

*We encourage you to talk to other people about general strategies, but it's important to keep in mind that using the same exact strategy will end up hurting yourself. Keep a record of what things you tried; it'll make the report write up that much easier.*

### 1 Pandemaniac++ [100 points]

Welcome to our second mini-project, Pandemaniac plus ~~plus~~ a twist! This assignment combines the competing cascades of Pandemaniac with an important topic that shapes a wide variety of applications: data markets. Specifically, in Pandemaniac we learned effective ways to take over as much of the network as possible when we had access to the whole graph. Now, you will do the same thing, but you will have to buy access to the graph by participating in data markets covered by second-price auctions.

**The Setting:** Like in Pandemaniac your goal is to take over as much of the graphs as possible through placing your seed nodes strategically. However, now you don't initially know the graph and instead have to participate in auctions to buy information about the graphs. Specifically, the graphs have been divided into sections, and there will be auctions held for each of the sections of the graphs. The information won in the auctions will be used to determine where seeds will be placed in the graph during that round. The key components of the problem are the data market, where you bid in a set of sequential auctions for access to the graph information, and the seed node selection, where you use the graph information to select your seed nodes. We discuss these two components in detail below.

**The Data Market:** The data market will consist of a set of sequential auctions for each section of the graph. All of the auctions will be second-price auctions, so the winner will pay the price of the second-highest bid. If there is a tie for section bids across different teams, then both teams will get the section and pay the higher price. If no one else bids, the winner pays exactly \$1. Every team will have a budget of \$1,000 to bid per graph, shared across all bidding rounds. You are never allowed to place bids that would overspend your total budget for a graph.

When determining your bids for each section, you will be given basic specifications of each section such as the number of nodes, max degree of a node, and number of unique edges. Once the auctions for each of the sections have been run, the winner of the auction for each section will then have won the full graph information for that section and will get to use that information for all future rounds of that graph.

Each graph will have multiple rounds of bids in the data market, so you will have four opportunities to bid on each section of the graph sequentially. You will be able to see the bids and outcomes of each section of the graph in previous rounds when determining your bid in each round for each section of the graph.

**The Seed Node Selection:** The spread of the epidemics and the termination condition will work the same as in Pandemaniac. For reference, please refer to the previous spec. However, unlike Pandemaniac, you will not be selecting your own seed nodes. All teams will use the same seed selection algorithm in order to keep the focus on how you participate in the data market.

The selection algorithm used by all teams will sample from the nodes in the graph, weighted by *known* degree. By *known* degree, we mean information that each team has won by participating in the data markets. There will be different seed nodes picked for 50 rounds per graph to reduce variance.

Additionally, students will be able to pick how many nodes to select from the "known" graph via the algorithm, and how many to select completely randomly (sampling from the known and unknown parts of

the graph). So, teams can choose how many seeds to distribute in the part of the graph they have data for using the weighted degree algorithm, and how many seeds to distribute uniformly at random.

**Your Task:** In groups of 2 or 3 (**not** 4), decide how much to bid on each section of the graph in each round. You will also set how many seeds are selected using the algorithm over the known graph and how many are selected randomly from the remaining part of the graph. Your team's goal is to obtain enough information that your seed nodes allow you to take over more of the network (i.e., a larger number of nodes) than your competitors do with theirs.

### The Details:

- (a) *Website:* You will submitting your bids for each round of this project through a website. **Please provide your team username and members of your team through a private post on Piazza so we can set your account up by Friday 11:59PM March 3rd.** You are welcome to use the same groups and team username from Pandemaniac. Once you post your team username and list of team members, you will have access to the Pandemaniac++ website: <http://52.9.240.217/>.
- (b) *Rounds:* There will be 2 seasons, with 2 days per season, from **March 6-9 (Monday-Thursday)**. Each season will have 7 graphs, and consist of 2 rounds per day, for a total of 4 rounds per season. Since there are 2 rounds of submissions for each day, **the first submission will be due at 6:00PM and the second submission will be due at 11:59PM.** The results will be available on the website by 7:00pm and 8:00am, respectively. **Note that you will not receive money committed to a round back until that round's results are available.** After each round, the clearing price and any information about the previous round's bids will be released.
- (c) *Graphs:* The graphs are split into 10 sections by a graph partitioning algorithm. All of the graphs will be from the Stanford Large Network Dataset Collection: <https://snap.stanford.edu/data/>. The graphs are named in the following (semantic) way:

*competition\_format . num\_nodes . unique\_id*

- *competition\_format* is a string "G", "O", or "J" representing the type of competition style,
- *num\_nodes* is an integer representing the number of seed nodes for the graph,
- *unique\_id* is an integer that distinguishes between different types of graphs.

On the graphs labeled "G", you will compete 1-on-1 with a greedy TA strategy, on the graphs labeled "O", you will compete 1-on-1 with an opportunistic TA strategy, and on graphs labeled "J", all teams will be competing with each other without a TA.

There will be two seasons of 7 graphs each, there will be 2 greedy graphs, 2 opportunistic graphs, and 3 jungle graphs:

| March 6-7 | March 8-9 |
|-----------|-----------|
| G.10.11   | G.10.21   |
| G.10.12   | G.10.22   |
| O.10.13   | O.10.23   |
| O.10.14   | O.10.24   |
| J.20.15   | J.20.25   |
| J.20.16   | J.20.26   |
| J.30.17   | J.30.27   |

- (d) *Submissions*: You will not have to download any graphs. Instead, you will be given general information for each graph. This includes the graph name, the round number, the number of players, and the number of sections in the graph. You will be given a budget of \$1,000 for how much you have available to bid on each graph, and the balance will decrease accordingly as you bid on each section. If your team wins a section, you will keep that information on the graph for the rest of the game. Therefore, you're gradually collecting data as you don't have to bid for it anymore (though they could bid to try to protect it from others getting the info if they want).

You will also have the ability to select a number of random nodes to be chosen. This way, teams can choose how many seeds to distribute in the part of the graph they have data for using the degree/TA algorithm and how many seeds they want to distribute randomly.

- (e) *Scoring*: The overall winner will be determined by adding up the points from each day of matches. **For each of the rounds** in a game, teams will earn points in the following Olympic format:

| Place  | 1st | 2nd | 3rd | 4th | 5th | 6th | 7th | 8th | 9th and beyond |
|--------|-----|-----|-----|-----|-----|-----|-----|-----|----------------|
| Points | 20  | 15  | 12  | 9   | 6   | 4   | 2   | 1   | 0              |

For each graph, the points across all 50 rounds will be summed. Teams will then receive points for the graph based on rank again according to the table above. Additionally, there will be incremental scoring, so each round within a season will be assigned a particular weight, from the first round having 1/6, to 1/5, to 3/10, to 1/3. This means later rounds matter more within a season! The team with the most points across the 2nd season's jungle graphs will be crowned the Plus Plus Pandemaniac!

### Grading:

- **Report [65 points]**: The most important part of this assignment is the thought that goes into your game theory strategies. So, the bulk of your grade comes from a report detailing your design process. Specifically, your team must submit one detailed report describing your strategy and the reasoning behind it. It must also describe the contributions of each team member as well as citations to any papers or other resources that helped in the formulation of your strategy.

A good report will show that you put significant thought into whether and how you could use different game theoretic principles. We also ask that you include suggestions on how we could change/improve it for next year! **This report, along with your final code, is due 11:59 pm, Sunday, March 12. You should submit the report through Gradescope and add all members of the group to the submission.**

- **Beat the TAs [10 points]**: The TAs will participate every day. You only have to beat each TA team once to get credit.
  - Beat the TA\_opportunistic. (5 Points)
  - Beat the TA\_greedy. (5 Points)
- **Participate [25 points]**: Submit to all graphs for all of the bidding rounds.

### Hints and Notes:

- Even after you have won a section, it can be prudent to keep betting on the section to deny information to other teams.

- Once you've unlocked a section, you can't ignore the information from it, so beware of 'poisonous' sections which can actually reduce your performance. Think about whether or not this phenomenon can happen, and why.
- The TAs have a fixed strategy based on the state of the game. Can you figure it out?
- Consider using different strategies on different graphs for the same TA bot within a season, why put all your eggs in one (strategy) basket?
- Consider writing a simple HTML parser to formalize your strategy with code.