Authors

Ethan Chapman, Max Fan, Ian Haile, Aiyu Kamate, Jeffrey Zhou

Introduction

In our Simulation Technique Directed Study, we decided to create a housing lottery system that can optimize students' housing choices, using the simulation techniques we learned, such as Monte Carlo Simulations. It is important to understand that the purpose of our project is not to completely replace the current system on hand. We wish to create a tool that the Deans can use to *assist* them in their task to create the housing allocation. Our goal is to create an algorithm that can create a housing allocation within seconds and have it satisfy the majority of the student's preferences. The finishing touches are still expected to be finished with human hands.

Dean's Algorithm

We define the Dean's Algorithm to be the current system that Choate uses for the housing allocations. During the span of our project, our goal was to improve on the current system and provide aid to it. To do that, it was important to understand the current system.

The Dean's method starts by asking all students to rank their top four preferred houses. The students are then randomly distributed through the available houses using a computer system, then the allocation is finished by hand based on how the student's ranked their preferences. Of course, there are many other nuances that change the results.

The Lottery system

There is a lottery system put in place to help your chances of obtaining the house you want. Each student chooses a random number, and those who choose a better number are given priority over those who chose a worse number - the better numbers are decided by a lottery by the deans. When students are on leave (medical or otherwise) for a year, students are unable to participate in the lottery. The school identifies what would be a good fit for the student based on what is known about them.

The purpose of the lottery system is to not only make it fun for the students to bet their luck, but it is also to create an element of randomness. The school wishes for diversity within each house in hopes that different groups of people can gather together and become friends.

Nuances

Fourth formers are normally given priority over 3rd formers and new students, if their first choice dorm is the same dorm they lived in. So, fourth formers are in a tiered lottery that can be separated into three different tiers, as indicated below (refers to roommate status).

- Tier 1 status. Both students choose to stay in the same house and have priority for beds
- Tier 2 status. One student chooses the stay in the same house and their roommate is coming from a different house
- Tier 3 status. Neither student are returning members of the house

The school does not allow students to fill out all their living choices as singles. A student is only allowed to fill out one or two houses as singles if they would like to at all.

Markov chain Monte Carlo (MCMC)

In this project, we used Markov chain Monte Carlo methods to distribute students into their respective housing. MCMC is a generic method for approximate sampling from an arbitrary distribution. The main idea is to generate a Markov chain whose limiting distribution is equal to the described distribution.

Metropolis-Hastings Algorithm is a specific type of MCMC method that we used in this project.

Metropolis-Hastings Algorithm:

The algorithm of Metropolis-Hastings MCMC is as the following: (P(a) is the probability of a)

1. Initialize with some random state X_0
2. From the current state, generate a new state X'
3. To decide whether X' is accepted as the next state: it is accepted with a probability of $P(X')/P(X_1)$ % and rejected otherwise.
4. Repeat the process of 2 and 3 as many times as needed.

Methodology

Swap and Gibbs

The swap and gibbs algorithm used a very simple variation of MCMC - the probability of accepting is 100% if the new state is better, and is 0% if the new state is worse, where our state is the weight of the student ballots.

Ethan Swap Naive

Max Network

Ian Minimax

Friends

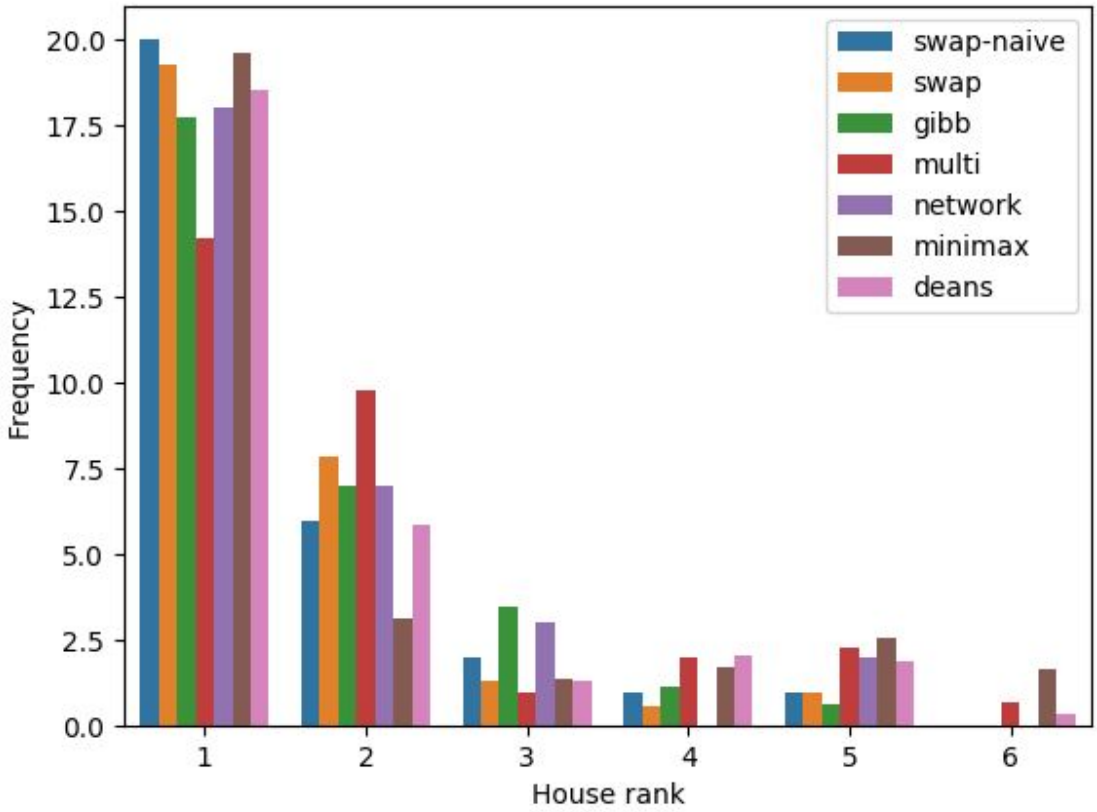
During our project, we considered important variables that went into housing allocations and we believed that friends could play a role into how much you valued your housing

experience. As a result, several of our algorithms considered friends and would place more importance on a dorm with friends compared to a dorm without. While every algorithm placed a different amount of importance on friends, each tried to group friends together.

It is important to note that the data that we use for friends is computer generated, since real data has never been collected, and thus it will not necessarily be a good representation of real data.

Results

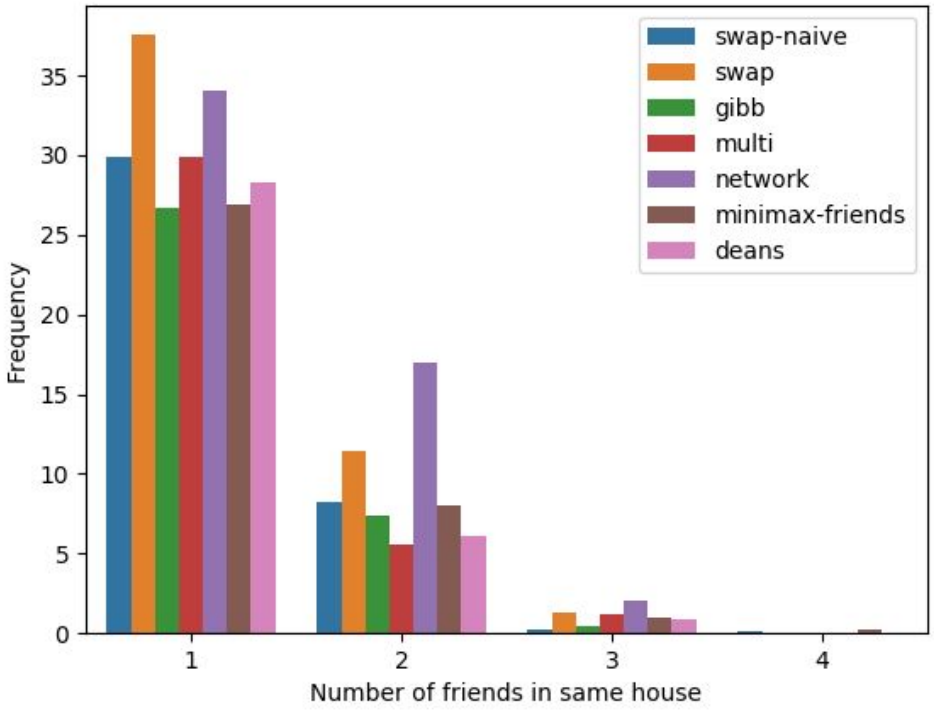
Housing distribution based on real ballots(year? grade?)



Above is a visual representation of the results of several algorithms after taking in real data in the year of (year?). The graph shows the distribution of students who were able to get their first through sixth choice depending on the algorithm used. One particular algorithm to look for is the dean's. As explained above, the dean's algorithm was implemented to be as close as possible to the current system that Choate uses for housing allocations. Consequently, this method sets a good baseline for comparison to see the strengths and weaknesses of our own algorithms. From the graph above, it can be seen that the dean's algorithm shows strong results in putting students into their first and second choices. Unfortunately, there are also an adequate amount of students in low ranked choices.

Swap-naive and swap are both notable methods that outperformed the others. Not only does their first choice eclipse all other algorithms, but they also minimize the amount of students going into lower ranked houses.

Friends distribution based on randomly generated ballots (100 students)


swap-naive
swap
gibb
multi
network
minimax-friends
deans
Frequency
0
5
10
15
20
25
30
35
1
2
3
4
Number of friends in same house

Conclusion