Allocation System Experiment Report: Key Findings and Plans

T039

OLIVER PINEL – ALEXANDER KEARNEY – RAPHAEL FAHEY – CALUM INGRAM

Abstract

Based on the experimental system we developed over the past few weeks and the current allocation process used in IFB398, our team conducted an experiment aimed at identifying the key facets of an effective allocation process. The experiment compared the student project pairing data from this semester using both the experimental system and the existing allocation process. This comparison allowed us to study both methods, highlighting successful elements, identifying flaws, and recognizing areas that require further research. From this analysis, we identified the critical components, requirements, and considerations that should be clearly defined as goals for a successful allocation system. Our team will now undertake a focused investigation into these key facets, producing detailed reports on the most significant aspects of the allocation system.

Contents

Introduction

Through our team's research over the past semester, we attempted to identify the key requirements of a system to assist in allocating projects to teams. We found three main considerations that we believed formed the basis for the success of such a system:

1. Accuracy of allocations:
   Each allocation of a pairing, of a project allocated to a team, should be as high as possible according to the benefit function. The system should aim to assist each team in being allocated a pairing that was as suitable as possible, particularly without impacting other pairings
2. Speed of allocations:
   While allocating a project to a team should not be a quick action, and indeed is the central focus of such a system requiring proper consideration, debating the potential pairing should not be complex, with all the relevant data easily accessible.
3. Retaining the human element:
   The system must avoid making decisions for the user, and while some algorithmic suggestions or user flow directions are essential in meeting the previous two points, all the allocations must be fully justifiable by the user.

Our team attempted to develop a system that could serve as an improvement on the current method, based on these considerations and the research undertaken through the last semester. To that end, we developed a system that aimed to provide options for sorting through the different available pairings quickly, simplify the comparisons between options, ensure that all the available data could quickly be accessed, and provide an assistive algorithm to present the theoretical best allocation set available as the allocations progressed.

In order to gain greater understanding of both the overall goals of our project, as well as the most significant aspects of such a system, our team conducted an experiment, comparing our system to the current method used to allocated projects to teams. Through this, we aimed to redefine the most important factors of the allocation process, to hopefully lead to a new focus for research and development in the future.

Experiment

Overview

To test, compare and gain insight from the two systems, an experiment was conducted wherein they were both utilised in the allocation process, which was recorded and then analysed to develop a deeper understanding.

This experiment was done using the allocation data for the IFB398 2024 semester 2 class. This data was collated into an excel spreadsheet consisting of three sheets representing the impact, capability and preference, which made up the benefit function. It was with this data that we expected the allocations to be made

Through the two tests, our team was specifically attempting to identify qualitative and quantitative information, which could be used to characterise the success of each system, as well as identify what are the most important elements of the allocation process. The overall goal of this experiment was to gain a deeper understanding of the allocation process, how well we understood it from simple descriptions, and determine how successful of an implementation our system was overall in meeting both the original requirements, and where it fell short.

By identifying key factors in each experiment and comparing notable elements of each system, we can redefine how an allocation system should function and clarify the essential core components for further investigation.

Current System

The current allocation system involved two main views, and extra data that was not initially incorporated when developing the experimental system. The data for the benefit function was loaded into an RStudio instance, from which it could be manipulated to view in different ways. These included lists sorted and filtered for several different categories, best pairings for a specific team or project, or overall best pairings. The RStudio system was complemented by a google sheets document which expanded on each pairing by displaying the full data behind the derived benefit values. By switching between the two, potential pairings can be identified and then verified in depth, before confirming them by pushing them to a shared google sheet.

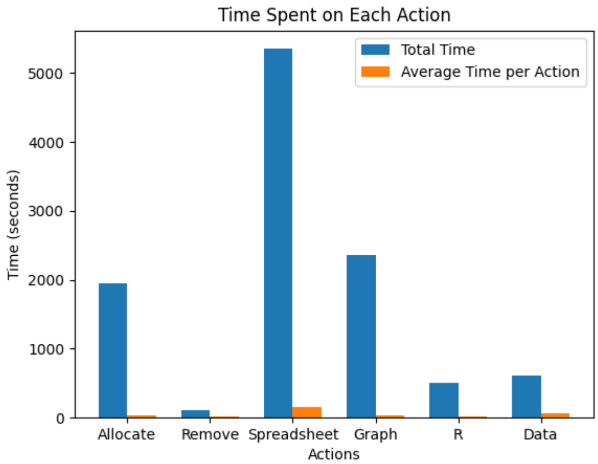
The most significant element of this system, being the central method for identifying potential pairings to be allocated, is the scatter plot. This graph displays all the non-zero impact pairings by graphing them across a scale of b value and sorting the teams by the highest b values. This created a visual distinction between the available pairings, and a distinct hierarchy of the teams. In addition, the preferences of the teams could easily be seen as extensions on projects. Most notably, the graph would update as allocations were made, only showing the remaining pairings.

Overall, the allocation process for this system ran without major issue. By studying the recording of the process, there were three major aspects noted as flaws in the system. In addition to the visual observations, the actions taken during the process were recorded, providing qualitative data that was used to graph some of these observations.

Most notably, was the unavoidable issue of unfairness and greedy allocation. Beginning at the top, the initial teams were afforded as much possibility as they could, being paired with their best fitting projects. As the process went through the teams and the available project pool dropped, the requirements for a pairing being accepted dropped to match, making the satisfaction of each pairing lower the later in the process it was. Notably, it should be considered though that 20 of the 43 teams had their preference allocated, meaning that this biased human method could favour teams, even in the greedy process.

Additionally, most of the pairings were identified initially from the scatter plot graph. This was, for most pairings, the only source of data considered before a pairing was allocated, potentially missing any improvements that could be made by viewing the data in another format.

Finally, there was a significant split in the time taken when viewing the lists and graph in RStudio and comparing and confirming the pairings in the google spreadsheet. While the accuracy of the recorded actions is not high, the graphs of the data is enough to show that the user flow of the process can have a significant impact on the time taken for each allocation.



It is worth noting that due to the simplicity of the system, while each allocation took time due to the switching between views, the process ran smoothly without issue. In addition, the final allocation set produced was, while not as good as an algorithmically produce one, still a good matching of teams to projects, while respecting the preferences.

Overall, the system performed as expected, making clear the benefit of a clear user flow, whilst highlighting the significant impact of the greedy nature on the success of the allocations. The graph provided a large amount of value, and an expansion of that to consider more pairings, in conjunction with a minimisation of greedy allocation could produce an improved system.

Experimental System

Based on research undertaken last semester, the experimental system was designed to meet the considerations identified as most important for an allocation process. In order to do that, our team focused on removing direct contact with the data, and instead providing a number of different methods of viewing and sorting through data. To reduce greedy flow, increase methods of comparison, and provided the user with several ways to sort through options, our team provided a system with several lists that could be sorted in a variety of ways, a colour coded spreadsheet and an algorithm that produced the best possible allocation set based on the current state of the process. The core foundation of the system was the pairing view, which displayed all the relevant information for a pairing of a team and project. This pairing view was incorporated into the list views, with an aim of making it visually clear to separate and identify good pairings initially with colour-relative backgrounds and allow for deeper understanding with the data displayed.

Initially, it was clearly obvious that the experimental system suffered from a poor UI. The simplicity of the current system was much more functional than the wide range of sorting and comparing methods that were not always helpful. In addition, the main feature of the system, the pairing display, was limiting in the poor layout choices for the data being displayed. While they were visually distinct, the significant data values were less than obvious, and when the rabbit hole process that was intended to ensure each allocation was thoroughly considered was followed, it became easy to lose track of progress.

In addition, the system suffered from an overreliance on the algorithm. It was simple to produce the best set, and then for each suggested pairing only consider them, which as it was an algorithm output, will be 'good enough' to approve the pairing. The full algorithm was more of a crutch that removed the human element than an assistive tool to the allocation process.

This system would benefit from changing the focus from numerically sorted comparisons to a heavier balance of graphical comparisons and visual differentiation. The algorithm could be minimised or redirected to provide a more assistive service, and a full investigation and redesign of the main UI components could have a large impact both on minimising greedy allocation and improving the user ability to properly compare potential pairings for allocation.

Redefinition of Project Objectives

Based on all the identified factors, good and bad, in each system through this experiment, our team has identified the most important factors of an allocation system to assist in the allocation process for IFB398. While there are some similarities to the original considerations we defined in our initial research report, our team has deepened our understanding of how the allocation should work and moved forward from abstract concepts of accuracy and speed and focuses now on actual functionality facets of our system.

Based on the results of the experiment, we have determined the following to be the four most important topics involved with the allocation system that should both be understood in greater detail as well as identify multiple solutions to avoid the inherent problems:

1. Greedy Allocation
   The inherent problem with the allocation process is that each pairing must be done by a singular person, leading to a greedy one-at-a-time allocation. This was very inherent in both systems, as one team was done, until the remaining teams at the end were left with the remaining projects, despite the unfairness of this method. Initially, our team had identified the greedy nature of the process as unavoidable and aimed to minimise the impact through other methods, ignoring it in any greater depth. Through this experiment, it has been noted as the biggest impact on the fairness of allocations, and as such proper research into methods of mitigation, be it through phycological techniques in displaying data, or changing the allocation process itself.

2. Algorithmic Identification
   Due to the large number of potential combinations, using an algorithm to provide a proposed allocation set was something we identified early on in our research as a beneficial factor. Despite this, in the experimental system it was used too much and took away from the human element requirement of the project. While the benefit of assistive algorithms for sorting the data is still worthwhile, alternative options are worth exploring that do not directly provide allocation sets, but rather identify specific pairings that would have a low impact or suggest groups of pairings that would work well.

3. Representation of a Pairing
   The most significant aspect of the allocation process is being able to visually identify the potential of a pairing. This should be quick and obvious. The attempt in the

experimental system was lacking due to a flawed UI layout design, and the current system requires extra data that was used to derive the b values to confirm each allocated pairing. The representation of a pairing should somehow communicate all the relevant information required to the user quickly so that they can easily discern whether its acceptable, as well as ensure that they can be quickly compared. Proper research and plans into how pairings can be compared without significant flipping between datasets is essential to speeding up the allocation and reducing the onus on the user's memory.

4. Graphically visualising and sorting through large datasets
   While inherently limiting in that it enhanced the greedy allocation process, the scatter plot used in the current system provided an incredible benefit in making comparisons between data very clear, much more than lists of numbers. As such, visual comparisons are clearly more effective than numerical comparisons, and further research into this should be conducted. Identifying ways to compare data, pairings and impacts through simple visuals could provide a large benefit to the allocation process and enhance the ability of the user to sort through data and identify good pairings.

Plans

To make the most of the remaining semester of work, our team aims to do an in-depth analysis of the four major facets identified in the previous section. By performing a thorough examination into how they impact this project, the potential mitigations or improvements of the problems they pose, we can identify a range of solutions to improve the overall nature of the allocation process by ensuring that the key components are improved.

To that end, our team is planning to spend the next three weeks individually investigating one of the topics each, reporting the results in the form of a comprehensive report. These reports will aim to make clear the depth of understanding on each facet of the system, and by breaking them down into components identify how they can be managed. Following this, our team will attempt to implement these improvements either in a vacuum for individual testing, or by updating the experimental system. How this research is implemented and tested will depend on the results of the research.

At the end of the semester, our team will deliver a comprehensive report detailing our progress throughout the year, including our new understanding of the project and any lessons learned. Ideally, this report will encompass all our work so that any new team can pick up where we left off and be able to understand all of our work and justification. Our team will produce a GitHub repository, which contains our experimental system, and our reports on our original research, the functionality of the system before testing, this report and the final overview. This final report and repository will serve as a comprehensive resource, ensuring that future teams can seamlessly continue to build on our progress.