Boeing Optimization Case Research Technology Review

Michael Chan, Brandon To, Mike Truong

November 17, 2016

OSU Senior Computer Science Capstone

Fall 2016

Version 1

Abstract

A focused workforce is able to accomplish multiple tasks within the time allocated. The direction of this project is to define a plan that determines the most efficient way for salaried Boeing employees to get to work on time to minimize the disruption to both their production and personal lives. This project will be experimental using the technique of a case study to examine Boeing employees over a time duration of two work days. Our plan is to use data collection techniques and analytics to develop an algorithm that can be used as demands change. The data collection involves evaluating instances that can be improved within employee work schedules with the use of surveys, questionnaires, observation, and records. Combining faculty documents with the data obtained from the employees of Boeing Everett the findings will determine a method to improve efficiency. At length the results will boost productivity for Boeing Everett.

CONTENTS

I. INTRODUCTION

The Boeing Optimization Research Project is based solely on designing an algorithm rather than the implementation itself. Due to this reason, our group is unable to conduct a proper technology review. Instead, we will be designing a case regarding the research, development, and potential implementation considerations for the algorithm we that are designing. The outline of this document closely follows the Rational Model describing the pre-production design, design during production, and post-production design. Each member was responsible for three sections each within the document. These members authored the following sections:

Brandon To: Design Brief, Analysis of Design Goals, Implementation

Michael Chan: Presentation, Development, Testing

Mike Truong: Research, Development Tools, Method of Analyzing Data

II. PRE-PRODUCTION DESIGN

A. DESIGN BRIEF

At the Boeing Everett Factory, there have issues with inefficient arrivals and departure times cause a direct correlation to the employees ability to balance production efficiency and time spent with their families. Since they have such a large population of approximately 30,000 employees and unorganized shift times, there are many instances where a dense amount of workers would attempt to department the Everett Facility at the same time, bringing about unnecessary traffic congestion. It is imperative that large factories like Boeing remain as efficient to not only ensure that the optimal amount of progress is made at the work environment, but also, ensuring that their employees are content and commute safely to their homes.

The aim of this research project is to effectively raise efficiency of the shift start and stop times of the salaried employees and management team. This is taken into consideration to enhance their quality of work environment as well as enabling them to live a less time sensitive personal life with their families. The qualities that we are working to accomplish in our algorithm are efficiency, sustainability, and versatility. All three of these goals, we believe tie closely together and support each other as we work towards accomplishing them. To maintain efficiency in all areas of Boeing's production, our algorithm must ensure that changing the start and stop times of shifts will not hinder the productivity of employees. We then look to sustainability of this efficiency to ensure that the solution is a long term one instead of creating issues else where within the production. Lastly, we must consider the possibility of new objectives that the company may want to accomplish. Thus, versatility of our algorithm is key so that it as can remain the solution to new problems.

The timing of the project will involve the beginning of Fall term and will conclude at the end of Spring term at Oregon State University. It has not been determined at this time but our team will be required to visit the Boeing Everett Site to gather data of the facility and to interview employees about their perspectives and needs regarding their start and stop times. The design of the algorithm is fairly open as it is primarily in its conceptualization stage. Our client from Boeing Everett has notified our team that the majority of the design choices made are up to our team as long as we keep in consideration that the algorithm must be changeable based on the demand requested by the company. At this current point of time, there is not a designated budget dedicated to the implementation of the

algorithm in development. The algorithm and suggestions disclosed by the final deliverable will be tested and used in the discretion of the management team at Boeing Everett. There will however be a budget that will allow our group members to travel down to Everett, Washington to interview employees and to analyze the facilities. This number has not been disclosed to us at this time as well.

B. ANALYSIS OF DESIGN GOALS

As stated in our design goals for the algorithm, we will primarily be focusing on efficiency, sustainability, and versatility.

To be efficient on successfully allocating salaried workers and to reduce traffic congestion, it is important to find a means to do this effectively without compromising their work quality and productivity. Employees must be able to arrive without difficulties so that they will not be late for important meetings, can attend the programs they are participating in, etc. They must also be able to leave at desirable times so that they may minimize disruption in their personal lives. With this, efficiency is one of the top priorities for the design of this algorithm because it also shares a direct correlation to net happiness of the employees. With net happiness higher, this allows for a stronger and more proactive work environment enabling workers to be more productive and deliver higher quality work.

Our aim of the algorithm is to design it to offer a long term fix rather than a temporary one. Sustainability is crucial in this respect. An issue that we might come across while focusing on reducing the traffic congestion is causing an inefficiency somewhere else within the site. Thus, the efficiency is nullified within the company since this would just allocate the problem to another portion of the company. To effectively avoid this, we must carefully analyze the various causes of the efficiency lapses at the Boeing Everett site. An example of this could involve the start and stop times of people within the same project groups. Regarding the employees that work on similar projects that need to consult with one another, we must ensure that at least one project manager is present at the time so that decisions can be made to push their processes forward. On the other hand, if we focus solely on raising the efficiency of start and stop times based on the preferences of the employees, important personnel may not be present when necessary. For situations such as this one, it is imperative to understand the cause and effect of the decisions that the algorithm makes so that there will be no compromise to efficiency.

The algorithm designed must be versatile to change in order to accommodate new goals in the future. Since demand is an ever-changing problem within the company, we must ensure that the design that we use can take into account the new factors that can impact the allocation of salaried employees and project managers. This goes hand an hand with sustainability as it allows for the long term use of the algorithm. Some ways that may accomplish this may be to take in account a specific amount of factors that can be modified or interchangeable within the algorithm. This will be the most difficult of all of the goals to accomplish as we will be required to predict what might be beneficial to Boeing in this aspect. We plan to gather data in this portion as we interview employees at the Boeing Everett Site.

C. RESEARCH

Research is a fundamental factor for our Project at Boeing. The goals that the Boeing group want to cover when it comes to research is population coverage, detail of the specifics, and easily understood form of communication with those at Boeing who will participate in this research. As a group we have chosen a few techniques, including: survey, case studies, and archival analysis. The importance of research is centered by the project requiring us to commute down to Boeing in Seattle, Washington to perform a various amount of test to gather data for our propose plan which should help Boeing in efficiency in the work place.

Surveys are used as a sampling units from a population, this data gathering technique can be categorized into different methods including written questionnaire, web survey, face-to-face interview and telephone surveys. Surveys can have factors in itself including cost, facilities, time and personnel. [2]

Our primary technique will be case studies because the technique will cover a majority of our goals with Boeing which includes coverage. We will be covering 20,000-30,000 employees at Boeing Everett, case studies allows us to scale the project down to a group of 20-30 individuals. The negative of the case studies is scheduling time with the individuals at Boeing that would participate in the study, I would assume that Boeing management would help us in this scenario however. Focusing on case studies as a median of communication among Boeing employees; University of Manchester has written that the technique itself allows engineers to be more open and collaborative. [3]

The last technique that can be examine is archival analysis, which involves seeking out and extracting evidence from original archive records. This is important because we are attempting to solve an issue of efficiency in the workforce. Being able to look within the past allows us to accomplish many of our goals including population coverage, details of species and the technique does not constraint the current Boeing employees.

The best techniques to use here are case studies and archival analysis, using a mix of these techniques will allow the team to pull information from multiple sources to build the case studies. Analyzing and choosing past records that contribute to the team's work will help build a direction of the project and also eliminate factors that could be negative. With the enormous factor involved in this research project, using a case study is beneficial, it allows the team to cover more individuals in the limited time slot we have visiting Boeing Everett.

D. DEVELOPMENT TOOLS

Currently the team knowledge of the presentation will be on white paper. However if there is a situational need for simulation there needs to be a program in-place for that use. Our goal as a team when selecting a developing tool is to evaluate the best tool for this research project. The two choices that have been selected and will be reviewed are: MATLAB and R. A brief description of MATLAB and R is too follow. MATLAB is a multi-paradigm numerical computing environment and a fourth-generation programming language. The program itself is written in C, C++, and Java. It's main use to my knowledge is for numerical computing. R is a statistical computing and graphic program. R also consist of a language with a run-time environment with graphics, a debugger and access to certain system functions.

MATLAB is a high-productivity developing tool; allowing for extensive toolboxes, apps, fast performance and easy development. MathWorks the developers of MATLAB state "As a result, scientists, engineers and their IT

colleagues have often found they are more productive using MATLAB than R for data analytics"[1]. MATLAB also has staff that work full-time verifying these algorithms. MATLAB is a free program that can be download from their website or accessible for our team because of the license we have with Oregon State using citrix. MATLAB also have built-in multithreading for computation on a multicore-enabled machine. MathWork have also stated that their program is 3 to 120 times faster than R. Ron Hochreiter, assistant professor at WU Vienna University Economics and Business wrote about his experience from switching to MATLAB to R.[2] In the case where he would use MATLAB over R, he states that the main reason is the optimization tasks.

R an open-source language, which includes over 6,000 available packages which have been reviewed by the community are all free to access. [3] R has a more simple interface that allows tools like dplyr, tidyr and Reshape to transfer data easily. The friendly interface allows transformation of graphs and data to PDF which will be a great resources when presenting data physically for the Expo in spring and also for the client. R however is stressed more on statistics.

Choosing this situational need for a simulation program MATLAB will benefit the team better than R. MATLAB is a more designing algorithms, simulations and prototyping program. A big part of this research project is able to gather data from multiple perspective and transfer that into an algorithm/plan that improves the work efficiency in Boeing Everett. Including the information I have now about my team, I believe that MATLAB is a more comfortable program to use compared to R. Even though R is a free open source software, being an Oregon State student in the engineering field allows us to access MATLAB and it's other program that are similar for free.

E. METHOD OF ANALYZING DATA

Methods of analyzing is an important factor when we begin logging data into our system for the Boeing research project. There are many ways that we can analyze data, here the two methods that we propose are white paper and modeling. As a research group our primary goal is gathering data from Boeing Everett and from other numerous variables to develop a plan that can increase efficiency in the workforce. At this point the team is unaware if this plan will be white paper or modeled based, the team is leaning towards a white paper methodology. The term algorithm can be defined numerically or theoretical. For this review we will consider both definition of algorithm in the use of each of their median.

Aim of a white paper algorithm is too develop the most basic understanding concept for the target audience. This simple white paper algorithm is more theoretical based off the data that we will receive in the following months of the research project. Data provided by Boeing and also our own case studies will need to be sorted to determine data that is accurate and sufficient for our use. Looking at the cost of developing a white paper algorithm we will need to be more organized in time and also documentation. With a white paper algorithm you can mistakenly lose a document which can become an issue after signing an NDA (non-disclosure agreement). We will be dealing with document and data which is privately owned by Boeing.

Modeling an algorithm is more of a high level view compared to a white paper algorithm. This method is more visual and wouldn't necessary need multiple physical documentation like the white paper algorithm. Saving space,

time and also the security aspect. The data would be stored virtually, allowing the team to eliminate variables of violating our NDA.

Suggesting the best method to analyze data currently is up for debate. The team currently waits on confirmation with our clients, however preparing for both of these factor will be important. In the manner of physical versus virtual, documentation will need to be organized strictly to hold privacy.

F. PRESENTATION

What our client currently wants is to have white paper and the other formats are up for debate as of right now. The paper's purpose is to display our findings and how we arrived at the conclusions that we drew from the various methods of data that we used. It will most likely be a LaTeX document that displays our algorithm in detail. The algorithm will be presented in a programming-like display with functions and other aspects that make it easy to glance over and understand. There will also be a separate section that will describe the algorithm in words so non-technical persons may read and understand the purpose and function of the algorithm. Example simulations may also be described to show how to apply the algorithm and what the outcome is expected to look like. The finalized paper may look similar to a thesis or similar dissertation used for presenting research findings.

The PowerPoint Presentation will be a part of the required demonstration where we present our work and findings for both our clients and the professors. As of right now the plan is just a simple voice-over with PowerPoint slides recorded onto a video that the interested parties may watch. Another possibility would be a physical meeting with our client to demonstrate our research, but due to the distance between us and the Boeing Everett site, this will most likely not happen.

A simulation is an imitation of a system. In this case the simulation will be a program that will implement and demonstrate our designed algorithm so our clients can see that it is actually functional without actually needing to enact physical change in their current work flow. Testing a new process takes time and risk, if the algorithm somehow proves in suboptimal then no actual man-hours will have been wasted. However there are many more variables that take place in real life than in theoretical situations so a simulation may prove ineffective at actually demonstrating the correctness of an algorithm. Confounding factors will affect physical tests that simulations won't be able to catch, so the absolute and effective method of proving our algorithm will be to implement it in a real world situation and record the results, which is known as an experiment. After reviewing the possible methods of presenting our research, we will most likely stick with a white paper recommendation for the final presentation method for the clients and the presentation as a requirement in the Senior Project Design process. A simulation would most likely prove to be ineffective for multiple reasons so it would be within our and our client's interest to not spend time on such methods. White paper is the most effective way to present an algorithm and as such, should be the main source of information. The presentation will be a step in the project where the client and our professors can see what work we have accomplished.

III. DESIGN DURING PRODUCTION

A. DEVELOPMENT

Development is the continued improvement of our designed solution. Our proposed solution will be an algorithm that when applied should improve arrival and departure for employees. In a time and motion study, the complex task/activity of interest will be broken down into steps which can then be analyzed individually. Wasteful and redundant movements will be reduced as well as enforcing any other improvements. From these changes the response variable, time, can be measured and from that it can be determined if a real statistically significant improvement has occurred. At the current time we are unsure about specifics in what our finalized solution may look like so much of this section is speculative. Due to time constraints my group may not be able to test the algorithm on multiple iterations but Boeing will be able to take our paper recommendation and develop it as they see fit.

First step would be to see if an improvement is possible or necessary. Is it worth the time or resources to continue changing the currently proposed solution? If the answer to these questions is yes then the next step would be to see where improvements can be made. The simplest way to determine where to start would be to look at the beginning of the broken down steps and go from there. A possible first step in the solution would be the arrival time step. The salaried employees want a good window to arrive at the Boeing site without being held up by traffic from other employees and any other factors. A solution proposed with the algorithm might be for each salaried worker in a specified window to arrive in a specific order but never at the same time. This is one way to improve traffic flow, but it's definitely not the best or most optimal. Say some employees want to arrive before other specific persons so they may start work sooner. The proposed solution does not cover that factor, so this could be implemented into the algorithm. The new block in the algorithm would organize employees into time windows as before, but after that it takes into account new factors such as preferred order. These small improvements can be applied in any of the broken down sections of the algorithm.

The other possible outcome is that the algorithm is someone proven to be fundamentally incorrect or there are not noticeable differences between applying the solution or not. This can be done through testing by implementing the algorithm or through simulation. Assuming either testing methods were done correctly and the solution is indeed flawed, then development is definitely necessary. Another option that is possible from this conclusion is to redesign or create a new algorithm if development is determined to be ineffective.

In conclusion, the best way to continue development of our proposed solution would be to analyze individual steps/blocks of the algorithm that will be affected most by the factors that are needed to be changed.

B. TESTING

Testing is the validation that is done at the end of a product's development cycle to ensure that it does what it is supposed to do and there are no errors that need to be addressed. In the Agile software development cycle, testing is done after the main parts of the project are finished[1]. In the testing block the software is vigorously tested for functionality and correctness, then released for approval for use. In our project we plan on developing our algorithm in a similar development cycle, but without programming, since the Agile development style is a familiar testing development style that we are used to.

Before any vigorous testing the solution must be analyzed for any obvious defects in logic or reasoning. These are the issues that are easiest to see and should be fixed first, such as faulty or incorrect outcomes that are produced from running through the algorithm. After that testing can begin.

In software, testing means running the program and checking the inputs. For us, testing inputs will be simulating or running through the algorithm by hand and determining correctness by looking at the outputted results, in this case the results will be some sort of schedule for employees. This is the cheapest and easiest way of testing a system/algorithm for correctness so it should be applied before any other more resource intensive tests, such as implementing the algorithm in a real world situation. Ideally a real world test is done eventually, but before that ever happens cheap and fast methods such as this will be tried.

Another viable method of testing is passing our algorithm or schedule to Boeing management and having the employees follow it. Our schedule will tell workers when to arrive and any other procedures that need to be taken. After logging their times for arrival and departure of the facility, we can compare the resulting times with their usual or previous times. It should be clear after a few iterations whether or not the algorithm is sufficient in solving their timing problems, and if it is not, it can be determined mathematically if the result is significant, but it should be very obvious whether or not the algorithm has succeeded or failed in its job. The issue with this method of testing is that there are confounding variables that will affect the results, as well as the time and money that will be used just to change the worker's schedule. Most algorithms won't be correct the first time they are implemented, and that's what the development cycle is for, improving processes and methods within a system. After many iterations of testing, these variables should be found and accounted for in new revisions of the algorithm.

The first method of testing should always be checking and simulating by hand, since it's fast and efficient and can catch obvious issues. After that method has been done enough and management approves, schedules may be implemented into the real world site. That form of testing will be the best for determining correctness, but will cost time and resources so it will most likely not be done as much as the first form of testing.

IV. POST-PRODUCTION DESIGN

A. IMPLEMENTATION

As a team, our overall objective is to compile and analyze the data we have collected in order to develop an algorithm that can be used to optimize the salaried employees arrival and departure times. We were also tasked to consider a multitude of factors that involve general locations of these individuals, methods of transportation, and project groups that these employees participate in. Since this is the case, we have not been tasked with the implementation itself as we are limited to a white paper recommendation as a final deliverable. However, since we are designing the algorithm for Boeing Everett's use, we will be looking at the steps considered in implementing an algorithm and some of the possible situations of implementing algorithms as other companies have done before to offer a recommendation. At this point of time, we have not been disclosed what hardware or software platforms will be used for the algorithm so this is kept in close consideration.

As referenced from Machine Learning Mastery, there are some steps that must be considered before directly implementing the algorithm. They must select a programming language (1), select an algorithm (2), designate a problem that can be tested on (3), research the algorithm's previous uses to obtain multiple perspectives (4), and lastly, write unit tests for each function to ensure that purpose of every piece of the algorithm is understood (5) [8]. The primary significant steps that require attention from the implementer will be steps one and five as two through four will be handled by our team. Since the algorithm is designed to allocate multiple factors and times, a language with recursive features is recommended to avoid redundant code especially since the amount sorted is immensely large. Regarding unit testing, a test driven development (TDD) approach is also recommended since it allows for greater accuracy of implementing the algorithm.

Viewing a case of where this was done in the past, in Implementing Algorithms for Signal and Image Reconstruction on Graphical Processing Units, written by Sangkyun Lee and Stephen Wright, they are given algorithms that they believe to be effective for compressed sensing and image processing applications. With this, they were tasked with implementing the algorithms. They operated on NVIDIA GPUs as a hardware platform and the Compute Unified Device Architecture (CUDA) for a software platform. They noted that in various places in the algorithms that they produced, they had to reduce the data set of size proportional to the problem dimension to a single value. This was crucial in their project as they tried to increase efficiency on a GPU, which involves careful implementation. In result, they divided the computation processes into thread blocks so that each block performs a partial reduction and produces a single number in result [9].

In our case, following this method of dividing tasks so that they handle one specific action is crucial in the implementation of the specific algorithm because it enables close attention to if the factors are taken into account correctly or if it is completely omitted. Also elaborating further on this, parts of the algorithm should be separated accordingly so that each part can be tested individually rather than all at once.