

**MGSC 662-075**

Decision Analytics

Presented to Professor Javad Nasiry

**Final Project – Seattle Seahawks Scheduling Problem**

Prepared by Group 10:

Gautami Edara 260973353

Cristina Esposito 260744222

Nathan Murstein

Hisham Salem 261003138

Sunday, December 5th, 2021

Contents

[**1.** **Introduction** 1](#_Toc89410705)

[**2.** **Problem Description** 2](#_Toc89410706)

[**3.** **Problem Formulation** 3](#_Toc89410707)

[**4.** **Numerical Implementation and Results** 7](#_Toc89410708)

[**5.** **Problem Extensions** 8](#_Toc89410709)

[**6.** **Recommendations and Conclusion** 10](#_Toc89410710)

[**Works Cited** 12](#_Toc89410711)

[**Appendix** 13](#_Toc89410712)

List of Figures

[Figure 1: Seattle Seahawks schedule, TV Viewership is not considered. 13](#_Toc89410734)

[Figure 2: Seattle Seahawks schedule, TV Viewership is considered. 14](#_Toc89410735)

[Figure 3: TV viewership comparison for Model 1 and Model 2 15](#_Toc89410736)

# **Introduction**

Football is the most popular sport in the United States, enjoying higher viewership and revenue than any other major sport. The National Football League (NFL) is the major football league in the United States, and the largest sports league in the world by revenue, grossing $13 billion in 2020 and attracting 7.3 million viewers on average per game so far in 2021 (Gough, 2021) .

The NFL’s league structure is like that of the NHL – the league is split into two conferences: American Football League Conference (AFC) and National Football League Conference (NFC), each of which is split into four divisions (North, South, East, West). Each division has four teams, with a total of 32 franchises in the NFL. The Seattle Seahawks belongs to the NFC West. Each year, there is a 4-week preseason, an 18-week season, followed by the playoffs.

Creating the schedule for each team in the 18-week season is a large, complex, and interesting problem. According to NFL.com, it is “one of the hardest and most scrutinized scheduling problems in existence”. One reason for this is that there are around a quadrillion game combination (Amazon, 2021), and the number of possibilities makes the problem computationally intensive. In fact, it takes 12 weeks for the NFL to run this optimization every year using several computers. Additionally, there are thousands of constraints. For example, regulating stadium availability, travel considerations, competitive equity, and internal rules about divisional structure can require formulating hundreds of thousands of constraints and bringing in many sources of data.

Additionally, there are many options of objective functions to optimize. Different formulations could be made to maximize value based on viewership, time slot advertisement cost, competitive equity (based on previous results/rankings), minimize travel costs, and more. In practice, the NFL uses optimization to create several viable schedules and manually selects from the generated options. Our team was interested in applications of optimization to scheduling, and the unique demands of this use case intrigued us.

Due to the computational, data gathering, and problem size requirements of scheduling for all 32 teams, to make our project viable in the given time, we have limited the scope to schedule for one team: the Seattle Seahawks. Our goal is to research the complex rules used by the NFL and translate them into constraints, and experiment with different formulations using various objective functions. Rather than generating several viable solutions, we aim to generate one schedule that lists which team the Seahawks face each week, which day of the week they play on each week, which of the 18 weeks they will take off as a bye-week, and more.

# **Problem Description**

Some of the scheduling problems we have seen in class apply integer programming to determine the schedule. The way we approached the design of this schedule applied mixed-integer programming, using a combination of both binary and integer variables. Our goal was to try and incorporate as many actual constraints the NFL experiences as possible to demonstrate the feasibility of applying a mixed integer programming problem. Granted, the design we have used does not allow for certain constraints regarding specific home and away game rules that must be respected during the NFL regular season, and therefore were left out of this project. However, it is import to keep in mind these rules for future iterations and design improvements of this project.

The problem consists of a total of 716 decision variables and 205 constraints. For us to generate a schedule that has 17 games, we have set our objective function to maximize the sum of all days of the week for each week and applied a strict constraint that forces this sum to be equal to 17. Since we have this constraint, we could have also chosen to minimize the objective function; there is no reason why one should be chosen over the other. However, thanks to the feedback provided by Professor Nasiry, there is a way in Gurobi to set this objective function to any random number, since the end goal is to generate a schedule that meets all the constraints.

# **Problem Formulation**

**Decision Variables:**

* Xij (binary) game played/not played on week i (from 1-18) on day j (from Monday to Sunday)
* Yi (binary) game played at home or not at home (need to account for the bye-week) on week i (from 1-18)
* Zij (binary) game played/not played on week i (from 1-18) against team j (from 1-31)
* Ni (integer) total number of games played in the NFC division i (North, East, South), not including the NFC West division
* Ai (integer) total number of games played in the AFC division i (North, East, South, West)
* Ti (binary) dummy variable for Ni games (1-3)
* Si (binary) dummy variable for Ai games (1-4)

**Objective Equation:** Maximize the number of games played in a season =

**Constraints** (Breech, 2021), (“Creating the NFL Schedule”):

1. *Can only play at most 1 game a week, both xij and zij must reflect this:*

and

1. *Must have 1 week off for the bye week between the 6th week and 14th week inclusively. This means that the total number of games that can be played between weeks 6 and 14 must be 8:*

1. *Must have 17 Games in a season, both xij and zij must reflect this:*

and

1. *If there is a game during the week of xij, then there must be a game during zij. This is to account for the bye-week:*

1. *Games are not played on Tuesdays, Wednesdays, and Fridays:*

1. *Games are not played on Saturday during weeks 1-13, and 17:*
2. *If the team plays on Sunday on week i, they cannot play on Monday week i+1:*
3. *1 Monday and 1 Thursday in the schedule:*

and

1. *At most 1 Saturday in the schedule:*

1. *If the team plays on Saturday on week i, they cannot play on Monday week i+1 (avoid games being too close):*
2. *There is a total of 8 home games:*

1. *There cannot be more than 2 back-to-back away/road games in the season. This means that within 3 weeks, the sum of y should be at minimum equal to 1 home game:*

1. *If there is a game on xij, then the game is either home or away. This is to account for the bye-week:*

1. *Must have 2 games per team in the same division:*

1. *Must have 6 games against teams in their conference. 4 must be with teams in another division that is in the same conference. 2 must be against a team from each of the 2 other divisions in their conference. Should not play the same team twice:*
2. *The total number of games from all the three divisions within the same conference must be 6*
3. *The sum of the North division games is equal to the total North division games*

1. *The sum of East division games is equal to the total East division games*

1. *The sum of South division games is equal to the total South division games*

1. *Should not play the same team more than once with the 3 divisions:*

1. *Need to choose the combination of games to be played with the 3 divisions to have the 6 games. There are 3 combinations:*

* 1 game with North, 1 game with East, and 4 games with the South
* 1 game with North, 4 games with East, and 1 game with the South
* 4 games with the North, 1 game with East, and 1 game with the South

For this type of problem, we applied the “N possible values” method:

For example, if t1 is equal to 1, that means the North division will have 1 game, the East will have 1 game, and the South will have 4 games.

1. *Must have 5 games against divisions in the other conference. 4 games must be with 1 division and 1 game from any other division. Cannot play the same team twice:*
2. *The total number of games from all divisions in the other conference must be 5:*
3. *The sum of AFC North division games is equal to the total AFC North division games*

1. *The sum of AFC East division games is equal to the total AFC East division games*

1. *The sum of AFC South division games is equal to the total AFC South division games*

1. *The sum of AFC West division games is equal to the total AFC West division games*

1. *Can’t play the same team twice in the AFC:*

1. *Need to choose which division will have 4 games. Since the sum of games within the AFC is 5, this would mean that the combinations of games between the divisions would have 1 division with 4 games, 1 division with 1 game, and 2 divisions with 0 games. We can apply the big M method to identify which division in the AFC will get the 4 games:*

For example, if S1, S2, and S3 are all equal to 1, the resulting division that would get the 4 games is the AFC West division.

1. *Non-negative constraints for N and A variables:*

and

Please refer to the Jupyter notebook for the full problem formulation.

# **Numerical Implementation and Results**

The formulation of the problem is based on the scheduling of the NFL. For this problem, we are focusing on the scheduling of the Seattle Seahawks team. The data we are using in the model contains a list of names of the 31 teams in the NFL that the Seattle Seahawks may play against. With this list of names, we created smaller lists with the indices that match these names for team in the AFC and NFC. These conferences are further split into 4 divisions with their respective list of indices: AFC East, AFC North, AFC South, AFC West, NFC East, NFC North, NFC South, NFC West. Since the Seattle Seahawks are in the NFC West division, the list of indices for the NFC West division only lists their 3 opposing teams (since we do not want to schedule Seattle with itself). To track the days of the week, we have created a list with Monday, Tuesday, etc. Apart from this, we have variables defining the total number of weeks (18) and the total number of games (17).

The decision variables and constraints were defined to capture the details of the game for the optimization model. Once the model was built and optimized it was able to produce the schedule for the Seattle Seahawks. For each week the model specifies the day on which the game will be played against which team. It will also indicate whether the team must travel for an away or home game. As shown in figure 1 of the appendix, the first game will be scheduled on Monday which will be a home game and will be played against ‘San Francisco 49ers’, second will be again a home game on Monday played against ‘Los Angeles Rams’ and so on.

Further, when we go through the list of constraints, we observe that all the constraints have been satisfied. Some of the examples are - as per NFL rule there can be only one game in a week which will be played by Seattle Seahawks and one week should be a bye week between the 6th and 14th week which is the 8th week in our league schedule. As per the results, games have not been scheduled on Tuesdays, Wednesdays and Fridays and games are not played on Saturdays during weeks 1-13 and week 17th all these clauses are imposed by the NFL.

By analyzing the model’s output, we can say that it follows all the game scheduling rules which are made mandatory by the constraints. This problem can be further extended to various team schedules. As an extension to this problem, we have used television viewership data to create a schedule that maximizes viewership. The viewership data that is used for the league is based on the NFL 2020 season. One modification was performed in this dataset where one of the week’s data was missing. We filled this week’s data with the NFL 2018 season’s dataset (because the website link to the 2019 data was broken) and took the average of the dataset as a list.

# **Problem Extensions**

There are several possible extensions to the base formulation that can yield interesting results. While the constraints are fairly set in stone in the NFL ruleset and cannot be changed, there are many creative options with the objective function, allowing the creation of different schedules that prioritize competitive fairness, revenue, or cost reduction. For example, a recent article in the Journal of Sports Analytics proposes ‘reducing the variability of teams’ strength of schedule or minimizing the number of pairwise comparisons needed to differentiate team quality to make each teams’ regular-season schedule as fair as possible’ (Bouzarth et al., 2021). Over 50% of the NFL’s extremely large revenue amount comes from broadcasting contracts (Dixon , 2020). Therefore, TV viewership is the most important driver of the NFL’s financial value proposition, and from the company’s perspective, should be maximized. Therefore, our problem extension will aim to maximize viewership, while respecting the same constraints from the original problem formulation.

Acquiring the necessary data for this extension was difficult because TV viewership is estimated by Neilsen, and we do not have access to this data. One idea was to assume that when viewership is higher, the cost of TV advertising slots is also higher. This data is publicly available, and by looking at historical ad placement costs, one could estimate viewership. The approach we took was to aggregate information from sportsmediawatch.net, a website that reports viewership estimates each week. We looked at every game in every week for 2020 (the last full season to be played) and averaged the viewership for each day of the week in each week (for example, average viewership among all games played on Thursday in week 7). There is a trade-off between using Seahawks-specific data and recent data. Since the Seahawks only played on one day each week in the 2020 season, to find data on Seahawks viewership on days they didn’t play, we would have to go back and use older data, which is a drawback. Therefore, we opted to use less specific (all NFL teams), but more recent data. Future Improvements on this problem may consider other variables correlated with viewership, including specific matchups, competing television programming, and specific time slot, which we did not consider in our scope.

After obtaining the data, we implemented the extension by changing our objective function to: . When Xij is 1 (meaning we have a game on day i in week j), we look at the 2020 viewership data for the same day of the week in the same week of the season and sum those numbers. When we run the problem, we obtain a different schedule from the original model, as shown in Figure 2. The new schedule has a completely rearranged set of opponents, a different bye-week, and different days of the week scheduled. Interestingly, the new formulation schedules the Seahawks to play on Thursday of week 12, which is Thanksgiving. NFL Thanksgiving games are always the highest viewed of any Thursday games, this makes sense intuitively.

Comparing the results of the original formulation and the extensions, we see that under the original model, we would estimate 273.23 million viewers, while under the new model, we would estimate 315.46 viewers, for an increase of around 15%. As shown in Appendix 3, the new allocation of games increases viewership on Monday, Thursday, and Sunday games.

# **Recommendations and Conclusion**

To conclude, one can see that the design of the scheduling problem is feasible given the constraints that were used, and the scope focusing only on one team. However, the schedule is not made at the team level, but rather from the league itself. Therefore, this type of problem would need to be replicated for all 32 teams and the schedules between the teams must make sense to ensure that they are not contradicting each other, adding another layer of complexity to the problem (i.e. if the schedule for team A says they will play team B during week 2 on Monday, then this must also be the same in team B’s schedule).

For this problem to be replicated for all 32 teams, continuing to use our design would not be feasible, as it focuses specifically on one team. Instead, the problem would have to incorporate a binary decision variable xijklmno, which would be interpreted as a game happening between team i and team j, during week k, on day l, during the time slot m (i.e. incorporate different start times such as 2 pm, 6 pm, 8 pm, etc.), with broadcaster n (i.e. FOX, CBS, NBC, etc.), at home or away. Attempting to conceptualize such a variable in gurobi goes beyond our capabilities but would be very interesting to see in terms of feasibility.

There are limitations to our current design if it were to be used just on an individual team level. For instance, when formulating the problem, if one wanted to add an additional constraint to consider having 1 game at home and 1 game away against each of the teams that would be linked to constraint 14 of section 3 of this report, the current design does not allow for this to be formulated without applying some type of conditional statement between the yi and zij variables. This is where redesigning the problem to potentially have a variable xijk (where I is the week, j is the day of the week, and k is at home or away) could solve this problem.

Overall, we had a fun time coming up with the design of this scheduling problem. Making a schedule is not easy, especially when there are hundreds of constraints that need to be respected all while making sure the teams will be happy with the results.

# **Works Cited**

Breech, J., 2021. *NFL 17-game schedule: Here's how the complicated scheduling formula will work with the extra game*. [online] CBSSports.com. Available at: <https://www.cbssports.com/nfl/news/nfl-17-game-schedule-heres-how-the-complicated-scheduling-formula-will-work-with-the-extra-game/> [Accessed 29 November 2021].

Bouzarth, E., Cromer, A., Fravel, W., Grannan, B. and Hutson, K., 2021. *Dynamically scheduling NFL games to reduce strength of schedule variability*. [online] IOS Press. Available at: <https://content.iospress.com/articles/journal-of-sports-analytics/jsa200428> [Accessed 29 November 2021].

“Creating the NFL Schedule.” *NFL Football Operations*, operations.nfl.com/gameday/nfl-schedule/creating-the-nfl-schedule. Accessed 12 Nov. 2021.

Dixon, S. (2020, June 15). The NFL - with or without fans this season - is on decent financial footing because of its lucrative TV contracts. chicagotribune.com. Retrieved December 2, 2021, from <https://www.chicagotribune.com/sports/breaking/ct-nfl-tv-money-20200614-ghh5urtcfvawbdlh5xv4pp3w6m-story.html>.

Gough, Christina. “National Football League (NFL) - Statistics &amp; Facts.” Statista, 29 Mar. 2021, <https://www.statista.com/topics/963/national-football-league/#dossierKeyfigures>. Accessed 2 Dec. 2021.

Staff, Amazon. “With Nearly Limitless Schedule Options, NFL Turns to AWS for Help.” US About Amazon, Amazon, 13 May 2021, <https://www.aboutamazon.com/news/aws/the-nfl-teams-up-with-aws-to-create-2021-schedule>.

# **Appendix**

Graphical user interface, text, application, website

Description automatically generated

Figure 1: Seattle Seahawks schedule, TV Viewership is not considered.

Graphical user interface, table, website

Description automatically generated

Figure 2: Seattle Seahawks schedule, TV Viewership is considered.

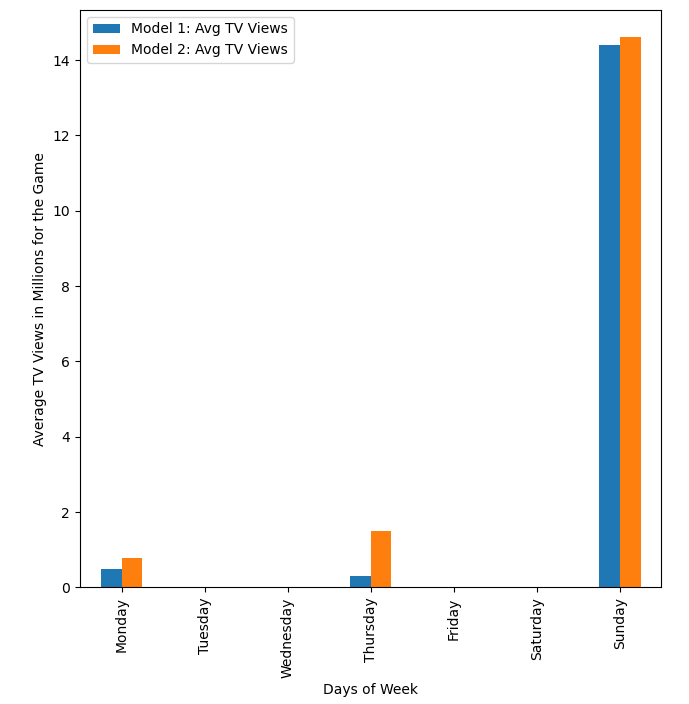


Figure 3: TV viewership comparison for Model 1 and Model 2