THE NFL TRUE FAN PROBLEM

by

SCOTT WHITTLE

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Industrial and Manufacturing Systems Engineering
College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2014

Approved by:

Major Professor
Dr. Todd Easton

Copyright

Abstract

Throughout an NFL season, 512 games are played in 17 weeks. For a given fan that follows one team, only 16 of those games usually matter, and the rest of the games carry little significance. The goal of this research is to provide substantial reasons for fans to watch other games.

This research finds the easiest path to a division championship for each team. This easiest path requires winning the least number of games. Due to NFL's complicated tiebreaker rules, games not involving the fan's team can have major implications for that team. The research calls these games critical because if the wrong team wins, then the fan's team must win additional games to become the division champion.

To identify both the easiest path and the critical games, integer programming is used. Given the amount of two-team, three-team, and four-team division tie scenarios that can occur, 31 separate integer programs are solved for each team to identify the easiest path to the division championship. A new algorithm, Shortest Path of Remaining Teams (SPORT) is used to iteratively search through every game of the upcoming week to determine critical games.

These integer programs and the SPORT algorithm were used with the data from the previous 2 NFL seasons. Throughout these 2 seasons, it was found that the earliest a team was eliminated from the possibility of winning a division championship was week 12, and occurred in 2012 and 2013. Also, throughout these 2 seasons, there was an average of 65 critical games per season, with more critical games occurring in the 2013-2014 season. Additionally, the 2012 season was used to compare flexed scheduled games with the critical games for those weeks and it was found that the NFL missed three weeks of potentially scheduling a critical game.

Table of Contents

List of Tables

Acknowledgements

I would like to personally thank Dr. Todd Easton for working with me on this thesis. His patience and insights were invaluable throughout the research and writing process.

I would also like to thank the entire Department of Industrial Engineering at Kansas State University for providing a learning environment that inspired me to begin this journey.

Finally, I would like to thank my parents for their continued support throughout my college career and for their understanding my desire to pursue a Master's Degree

Chapter 1 - Introduction

Sports fans everywhere root for their teams. In modern day sports, the relationship between a fan and their team is quite passionate. However, throughout a season, fans usually only watch their team, which is a fraction of the total games played in a particular sport. In the majority of games where their team does not play, absence of attention is due to lack of interest. However, if a rooting interest were provided, then those fans would have a reason to watch another team play for one or more games. Answering the question what teams to cheer for in certain games for fans of a specific team is the purpose of this research and is called the True Fan Problem (TFP).

1.1 Money and the Importance of Fans

Money is everywhere in sports. Owners buy teams and are in control of every aspect of a franchise, from ticket and concession prices to coach and player salaries. Teams in the 4 major professional sports in the United States (football, basketball, baseball, and hockey) cost in the hundreds of millions, and even some in the billions of dollars (Smith). These owners spend hundreds of millions of dollars every year, and expect the team to perform up to a certain level so the owners can hopefully gain a net profit from the success of the team. There are countless revenue streams where an owner can recoup their investment into the team, but the most lucrative revenue stream comes from the fans.

Fans are the most important client to a sports team. Without fans, nobody would watch the games. If no one is watching, no one is buying tickets or merchandise, and TV ad sales decline until the TV network decides not to pay the team to show games. Soon, the owner cannot afford to pay the players and staff, and must sell the team or fold it into non-existence.

Thankfully, the world is full of fans that fervently cheer on their teams. Fans invest great time, money, and emotional energy into experiencing the thrill of every victory and the agony of every defeat. From a sports league perspective, the best way to make more money is to capitalize this connection between fans and their teams. By creating reasons to watch other teams and games, the league creates added viewership and interest, and fans are getting more reasons to watch the sport. Creating these extra moments where fans have a vested interest in the outcome of a game, reacting with extreme joy or despair on every play, would create more popularity for the league, and thus generate more revenue. By giving the fans more reasons to watch, the fans in turn give more to the league in terms of added potential revenue. This research provides knowledge reasons for fans to care about games not played by their team.

1.2 The Importance of Quantitative Analysis is Sports

Beginning in the 1960's, baseball took to statistics and mathematical simulations to gain advantages on the field. Slowly but surely, sports analytics have gained traction across the sporting community and now most sports are in a full-fledged analytic revolution. The rise in popularity of fantasy sports has further spurred this revolution. Ardent fans of fantasy sports use the basic statistics available to make roster decisions day in and day out. Soon, more advanced statistics began to appear in fantasy sports, and many fans wanted those to be included in the box scores of games as they portray a more accurate performance by teams and players. Currently, these advanced statistics can be seen in most major sports in one way or another.

Major league baseball (MLB) has adopted an entirely new set of statistics with which to judge players. These "new" statistics are not revolutionary in their mathematical makeup, but rather in their application to validating how much a player is contributing. Gone are the days of using batting average, runs batted in, and stolen bases determining who are the best offensive players. Now, on-base percentage, slugging percentage, and wins above replacement (WAR) shape a better picture of how good a baseball player is on offense. There are similar new measurements for running, fielding, throwing, and pitching. By using the data available, new insights and powerful simulations can be created to more accurately analyze the value of a player.

The National Basketball Association (NBA) has installed advanced tracking cameras in every arena. These cameras track touches, shots, passes, distance traversed by players; they capture almost every movement on the court. These movements are then processed into statistical measurements of efficiency that reveal, among many things, the shooting tendencies, defensive shifts, and rebounding rates. By knowing what is likely to occur given the placement and tendencies of players, it will help to create better diagrammed offensive and defensive schemes to maximize the talent of the players on the court.

The NFL has developed new ways to gauge quarterback performance and measure win probability based on field position and play-call selection. These new statistics have created better ways to assess a quarterback's performance throughout a game, not just at the end. Additionally, probability charts are being used to help determine play calls in specific situations. These insights have led to better and more informed decisions throughout games, during drafts, and when weighing trades. Consequently, a better team is assembled (hopefully), a better game plan is created, and the probability of winning games increases.

1.3 Playoffs and Tie Breakers

For the fans, the use of advanced statistics by teams is a great positive. The ability to make more informed decisions about players and plays should create a better chance of winning, which is the primary concern of the fans, players, and personnel. Everyone cares about winning because winning gets a team to the playoffs, and only teams that make the playoffs have a chance to win the league championship, which is the ultimate goal for each team. However, playoff format and determining which teams make the playoffs is quite different in each sport.

The National Hockey League (NHL), NFL, MLB, and NBA all have different playoff formats, and the goal of the playoff games is to determine the two best teams that play for the championship. The NHL, MLB, and NBA all have the playoff games between two teams played in a series, usually best of 3, 5, or 7 depending on the sport and playoff round; the NFL consists of single game elimination rounds. Seeding in the playoffs is determined by different factors. All of the previous sports leagues use a division format to determine some of the playoff qualifiers, with the best team in each division automatically qualifying. Other teams qualify for the playoffs through league specified rules, usually called wildcard teams, and involve division and conference dominance stipulations. There may come a point when determining which teams have qualified for the playoffs when two teams are tied in one or more playoff-qualifying categories. When this occurs, each league has a unique way to handle these tiebreaker scenarios. The NFL is a particularly interesting league for tiebreakers. The league has so few games that tiebreakers frequently determine division and champions and wildcard teams.

The NFL currently has 32 teams in 8 divisions which are mostly established based on geography, with each division containing 4 teams. The NFL also has two conferences, with each conference comprising of 4 divisions, for 16 teams per conference. The NFL does not have every team play every other team, but instead uses a rotating 4-year schedule. The only games guaranteed every year for every NFL team are the 6 games against the other 3 teams in its division. For the playoffs, the NFL takes the best team in each division, based on Win/Lose Percentage (WLP), and those teams automatically make the playoffs. An example of WLP is for a team that has 10 wins and 6 losses has a WLP of .625 (10/16). However, if two or more teams are tied in their WLP, then a series of tiebreakers are applied until only one dominant team remains able to qualify for the playoffs. Given the NFL's small sample size of 16 games, there is a higher probability, compared to other sports, for teams in a division to finish with similar records. Additionally, since 6 out of 16 games occur within a team's division, there is a great emphasis placed on those games, and this is reflected in the NFL's tiebreaking procedures. A few of these tiebreakers are head to head, division, conference, and games played in common WLP.

As an example, team *a* is currently in first place in their division and team *b* is currently in third. Team *b* is playing second place team *c* this week, which is only one game behind team *a*, and team *c* has beaten team *a* once already. In this scenario, team *a* would want team *b* to beat team *c* for two reasons: first, team *c* losing would add *a* loss and hurt its chances of tying team *a* for the division lead, and second, in the event of *a* division leading tie, this loss also hurts team *c*'s in division record, giving team *a* yet another opportunity to win a tiebreaker. Extending this example, knowing all teams to root for every week or series in a sport so that one's favorite team (team *a*) would have an easier path to the playoffs would significantly add interest and intrigue for fans of those teams. Determining these games and associated division champions is the NFL True Fan Problem and is the focus of this research.

The research in this paper aims to help a true fan root for their team in the NFL each week by identifying the easiest path to winning the division, and what games are critical to that path. These critical games are defined as games for the upcoming week that would negatively affect the easiest path to winning the division. These are games outside of the team's control, so the fan must rely on cheering for certain teams to beat other teams to aid in the easiest path possible to win the division. It is important to note that no game would ever positively affect the easiest path as that would contradict the given fact that the easiest path was already found. Again, the NFL is being used primarily due to the unique nature of its division tiebreakers and season schedule. Additionally, the NFL is far and away the most popular sport in the United States both in terms of revenue generated and TV viewership. This is easily evidenced by the fact that the National Football League brought in $9.5 billion in 2012, $2 million more than the second-place MLB (Statistic Brain), and the top 8 most-watched shows in 2012 according to the Nielsen ratings were NFL shows or games (Nielsen).

1.4 Research Motivation

In every sport at every level, rivalries exist among teams. A rival could be that cross-town, same state, same division, or historic matchup team that every fan gets especially excited for. The most famously touted rivalry is the Boston Red Sox and New York Yankees from the MLB. These two teams have played each other 2,121 times since 1901, with the Yankees having a .543 winning percentage, which makes it a very competitive rivalry. This rivalry is also one of the oldest, having been around for 113 years, which lends itself to generations of fans being raised to love one team and hate the other. The NBA and NHL also have storied rivalries of their own, most notably the Boston Celtics vs. Los Angeles Lakers for the NBA and the Philadelphia Flyers vs. New York Rangers for the NHL. However, unlike the MLB, NBA, and NHL, the NFL rivalry scene is very unique.

As noted earlier, the NFL's divisions are largely divided by the geographic proximity of the teams in the divisions. Intra-division games are very important to each team hoping to qualify for the playoffs as a division winner. Unlike the MLB, NBA, and NHL where each team plays every other team multiple times, each NFL team does not play every other team, playing only 42% of the possible teams each season. Additionally, since only divisional teams play every year, intense rivalries have developed for almost every team in the NFL within their divisions. The last thing most fans do is cheer for a rival team, no matter what; however, given the NFL tiebreaking procedures for multiple teams being tied for the division championship, cheering for a rival might be the best thing to do in certain scenarios.

For even the casual fan, the NFL TFP aims to find teams to cheer for other than one's favorite team in the hope that the other team ultimately aids the favorite team's chances of qualifying for the playoffs. The surest way to make the playoffs is to win every single game, but that task proves very difficult over a 16 game season. In fact, only one team has managed an undefeated regular season since 1978, and only four teams have ever accomplished this feat. The objective becomes how to win the least amount of games, but still win the division championship and qualify for the playoffs. By winning the least number of games possible, the path to the playoffs usually goes through many tiebreaker scenarios. Given the NFL's emphasis on intra-division games, the best result in most games is for a rival to lose; however, given that head-to-head WLP, division WLP, strength of opponent WLP and strength of schedule WLP are all playoff tiebreaker factors, having the rival win as many games as possible and not win the division is one of the easier paths to the playoffs as a division champion.

This research into the easiest possible path to win a division in the NFL and which games are considered critical week to week had many motivating factors. One factor is that for any fan of a team, it would make the NFL much more meaningful if the list of teams who needed to win and lose each week was available. Usually, people watch their favorite team, and then do not pay as much attention to the other games or have minimal rooting interest in the other games. Giving fans another reason to deliberately watch football can only be good thing for the NFL.

Instead of having other games on in the background or paying little attention to them, fans would know which games to watch and who to cheer for. A direct result of having more fans intently tuned in to more football games is that the NFL should have more fans watching its product. Having more fans watching means more people are seeing ads, which the NFL as a business can leverage the increased popularity for higher prices on TV contracts.

More money into the league means more opportunity to improve the league and the product on the field, which leads back to better fan interest. Additionally, knowing how a team can make the playoffs the easiest way possible is a new take on how to make the playoffs. There may be only a small chance for a team to make the playoffs, but knowing that one exists surely give fans hope for that dream scenario. This research aims to minimize that effort and difficulty my focusing on the easiest path, not the most difficult.

1.5 Research Contribution and Significance

This research introduces the NFL TFP. As stated earlier, the NFL TFP aims to determine which games are critical for each week for each team. A SPORT (Shortest Path of Remaining Teams) algorithm was developed to solve the TFP. This algorithm solves 31integer programs to determine whether or not a game is critical. These 31 integer programs comprise every possible scenario of two-team, three-team, and four-team divisional ties for playoff qualification. Each integer program (IP) determines the easiest path to winning the division and thus qualify for the playoffs for a given team. The integer program also determines the number of wins that correspond to the easiest path. The integer program takes into account all possible NFL division tiebreaker scenarios up through strength of schedule to determine the easiest path. Among the 31 IPs, the minimum solution and corresponding divisional tie scenario are isolated. The SPORT algorithm uses them to determine critical games, if any exist.

The critical games are determined by changing the results of each game for the upcoming week, and then re-solving the easiest path integer program. If the solution changes, that game is a critical game since a change in the outcome of that game requires additional wins or ties by the desired team.

The developed algorithm SPORT uses CPLEX 12.5 (CPLEX) to solve the IPs. On average, SPORT requires only about 70 minutes to determine all of the critical games in a given week for every team. The analysis showed that in the 2012 season the NFL had 67 critical games. Furthermore, every week had at least one critical game and week 15 had the most critical games with 10. Comparing the second half of the 2012 season to the second half of the 2013 season, similar results were obtained. The 2013 season featured many more critical games in that span than the 2012 season, but did not have as many teams eliminated throughout the second half of the season.

As has been previously touched on, there are significant benefits that solving the True Fan Problem can deliver for the NFL. One obvious benefit is the ability to better market and sell games. For a given week, if one or two games pop up as critical for many teams, it would be in the NFL's best interest to hype those games and sell the fact that the game has many implications for many teams' playoff hopes. This idea would help all teams across the league. Some of the less popular teams by television market such as the Jacksonville Jaguars and Buffalo Bills, and even more popular teams, such as the Dallas Cowboys would receive a boost in viewer interest in unrelated television markets when they are involved in a critical game (Klis). For some games that are not "big" games to the general viewing public, having the ability to sell the game with certain playoff implications would create additional interest to otherwise ambivalent fans.

Currently, the NFL sells its weekly games to the NFL Network for one Thursday night game per week, ESPN for one (two on opening night) Monday night game per week, NBC for one Sunday night game per week, and CBS and FOX for the remaining games, with televised games being based on region. Beginning in Week 11, the NFL uses a flex schedule, where it can place any Sunday game into the Sunday night primetime slot. The advantage of this slot is that it is the only night game on Sunday, and it is on a basic cable network, allowing the most people to watch without any other games to worry about. If the NFL used the SPORT Algorithm, it could determine which matchup had the most critical game implications, or which teams were in a must win scenario. With this knowledge, a more informed decision could be made as to what the best game for the Sunday night primetime slot would be. A better game with more people watching creates a more attractive product to viewers, which will only increase the value of the NFL. Even thinking beyond just the NFL, knowing critical games for playoff hopeful teams would be significant knowledge for fans, teams, and the league.

1.6 Thesis Outline

Chapter 2 contains background information to help grasp the contributions of this research. It describes a generalized sports league in a formal manner, as well as prior research contributions to sports in various mathematical disciplines. Some background research is presented on general applications to sports, sports scheduling, and playoff qualification.

Chapter 3 starts with a description of the NFL True Fan Problem. The descriptions of the 31 different integer programs that are used to solve this problem are presented. The chapter concludes with the SPORT algorithm which iteratively identifies the critical games.

Chapter 4 describes the computational results. First, the results for the previous 2 years of the NFL are summarized by week. Significant results for each year are also given. Along with year by year results, specific scenarios are selected and discussed in detail as being particularly interesting.

Chapter 5 gives the final conclusions from the research, the insights gained from the research and findings, and results of the NFL True Fan Program. Also, future work is discussed to provide other researchers with additional ideas and concepts on how to further this research or apply it to other areas.

Chapter 2 - Background Information and Prior Research

This chapter describes the background information necessary to understand the contributions of this thesis. The first section provides a formal description of a sport's league. The second section describes prior Operations Research as applied to sports. The final section has a detailed explanation of the NFL tiebreaker procedure.

2.1 Sports Leagues

All sports leagues have a number of teams. This set of p teams is denoted by $T = \{t_1,\ldots, t_p\}$. Frequently these teams are partitioned into divisions, denoted by $\mathcal{D} = \{D^1,\ldots,D^q\}$. It is assumed that $\bigcup_{u=1}^{q} D^u = T$ and that $D^u \cap D^v = \emptyset$ for all $u, v \in \{1,\ldots,q\}$ and $u \neq v$. Several divisions can be combined together to make a conference $\mathcal{C} = \{C^1,\ldots,C^r\}$ where $C^v \subseteq \mathcal{D}$ for each $v=1,\ldots,r$. Furthermore, $C^u \cap C^v = \emptyset$ and $\bigcup_{v=1}^{r} C^v = \mathcal{D}$ for all $u, v \in \{1,\ldots,r\}$ and $u \neq v$. For notational convenience D_i and C_i will denote the division and conference that team i plays in, respectively.

The NFL in 2013 is used to help clarify this notation. The NFL contains 32 teams as shown in figure 1 and $T = \{t_1,\ldots,t_{32}\}$. For convenience, team 1 is the Baltimore Ravens, further known as $t_1 = t_{Bal} = t_{Baltimore} = t_{Ravens}$ and similar notation is used for the other teams. These teams are grouped into 8 divisions $\{D^1,\ldots,D^8\}$ of 4 teams each as shown in Figure 2. These divisions are further grouped into 2 conferences $\{C^{AFC}, C^{NFC}\}$ with $C^{AFC}=\{D^1, D^2, D^3, D^4\}$ and $C^{NFC}=\{D^5, D^6, D^7, D^8\}$. As noted $D_{Bal}= D^{AFC\ North} = D^1$ and $C_{Bal}= C^{AFC} = C^1$. For the NFL, the conferences and divisions are shown below in Table 2.1.

Table 2.1-1. The National Football League by Conference and Division

AFC Conference Teams					NFC Conference Teams			
Team	Logo	Abbrev	Division		Team	Logo	Abbrev	Division
Baltimore	Ravens	BAL	AFC North		Chicago	Bears	CHI	NFC North
Cincinnati	Bengals	CIN	AFC North		Detroit	Lions	DET	NFC North
Cleveland	Browns	CLE	AFC North		Green Bay	Packers	GB	NFC North
Pitssburgh	Steelers	PIT	AFC North		Minnesota	Vikings	MIN	NFC North
Houston	Texans	HOU	AFC South		Atlanta	Falcons	ATL	NFC South
Indianapolis	Colts	IND	AFC South		Carolina	Panthers	CAR	NFC South
Jacksonville	Jaguars	JAX	AFC South		New Orleans	Saints	NO	NFC South
Tennessee	Titans	TEN	AFC South		Tampa Bay	Buccaneers	TB	NFC South
Buffalo	Bills	BUF	AFC East		Dallas	Cowboys	DAL	NFC East
Miami	Dolphins	MIA	AFC East		New York	Giants	NYG	NFC East
New England	Patriots	NE	AFC East		Philadelphia	Eagles	PHI	NFC East
New York	Jets	NYJ	AFC East		Washington	Redskins	WAS	NFC East
Denver	Broncos	DEN	AFC West		Arizona	Cardinals	ARI	NFC West
Kansas City	Chiefs	KC	AFC West		St. Louis	Rams	STL	NFC West
Oakland	Raiders	OAK	AFC West		San Francisco	49ers	SF	NFC West
San Diego	Chargers	SD	AFC West		Seattle	Seahawks	SEA	NFC West

Using this notation, the AFC divisions are $D^{AFC\ North}$ = {t_{Bal}, t_{Cin}, t_{Cle}, t_{Pit}}, $D^{AFC\ South}$ = {t_{Hou}, t_{Ind}, t_{Jax}, t_{Ten}}, $D^{AFC\ East}$ = { t_{Buf} , t_{Mia}, t_{Ne}, t_{Nyj}}, and $D^{AFC\ West}$ = {t_{Den}, t_{Kc}, t_{Oak}, t_{Sd}}. The NFC divisions are $D^{NFC\ North}$ = {t_{Chi}, t_{Det}, t_{Gb}, t_{Min}}, $D^{NFC\ South}$ = { t_{Atl}, t_{Car}, t_{No}, t_{Tb}}, $D^{NFC\ East}$ = {t_{Dal}, t_{Nyg}, t_{Phi}, t_{Was}}, and $D^{NFC\ West}$ = {t_{Ari}, t_{Stl}, t_{Sf}, t_{Sea}}.

Every sports league has certain dates and times when games are played. The days or potentially hours in a day that a game can be scheduled is called a slot. In general, slots that the games can be played are denoted by $S= \{1,\ldots,s\}$. Given the available slots S in a league, the league creates a schedule. This schedule is usually repetitive in terms of days of the week games are played. Overall, the schedule is a collection of games $G= \{1,\ldots,g\}$ with each game occupying exactly one slot $k \in S$. Each game takes a 4-tuple form of (t_i, t_j, k, o) where this indicates that team i plays at team j in slot k with outcome o. The value of o is dependent upon the league and places the value of a win, loss or draw based upon the outcome of the game between t_i and t_j, in slot k. If the game is yet to be played, then o=-1. For the NFL, a win is 2, a tie is 1 and a loss is 0. The schedule for team i, $G_i = \{(t_a,t_b,s_k,o)$: either t_a or $t_b = t_i\}$. Clearly $\bigcup_{i=1}^{p} G_i = G$.

2.2 Operations Research and Sports

Mathematics and sports have gone hand in hand since modern sports have existed. Statistics, in some form or another, are available for even the earliest of baseball, football, soccer, and basketball games. Most of these sports have kept the same core of basic statistics to report to fans, though some have changed and been updated. One branch of mathematics that has been applied in many ways to sports is Operations Research (OR). As a studied topic, OR has been around since the mid 1940's (OR Journal). OR as a field has been applied to countless problems covering massive subjects and disciplines, and eventually the applications of OR found their way into the world of sports.

One of the earliest applications of OR applied to sports was in by Benjamin L. Schwartz. Schwartz (1966) studied baseball and the divisional pennant race. His research answered the question of, "Is a given baseball team eliminated from the pennant race (playoffs)?" Schwartz used network flow for a single team to show if a team was eliminated or still in contention. Soon after this work was published, OR began to appear in many fashions in more sports-related topics.

2.2.1 Mathematical Applications in Sports

While Operations research is useful in answering hypothetical questions regarding optimality in some sense for a team, most of what a team uses in decision-making analysis is outside of the OR discipline. The most basic of decision-making criteria comes from gaining information for evaluation amateur talent. Most leagues have a draft in which teams rotate a set order and draft amateur athletes to their team. Prior to a draft, the league usually hosts a combine, which is a combination of events over a few days, in which prospective and hopeful athletes participate in strength, speed, quickness, agility, and other drills. The purpose of these drills is to showcase their athletic prowess to the scouts and general managers of teams in the hopes of proving they have the talent to be drafted. The scouts in attendance take the performance numbers back to their team and assess many things. They look at the athletic numbers and determine a proper fit in the sport; quite a few collegiate football athletes make the team in a different position than the one they played in college. Additionally, scouts look at how those numbers stack up against the veterans already in the league. In the NFL, if a lineman was average in college, he might be way too small for the pros and not worth drafting. Or a wide receiver was moderately fast in college, but in the NFL, usually only the fastest athletes play wide receiver. By no means do these decisions require advanced modeling or intense mathematical formulation, but they are basic statistics used to determine potential talent.

One of the pioneers who made the idea of using advanced statistics to draft, sign, and trade players was Billy Beane of the Oakland Athletics (Lewis). Under general manager Billy Beane, the Oakland Athletics became a baseball sensation in the late 1990's and into the 2000's thanks to the statistical analysis Beane and his front office brought to the team. Beane put more stock in numbers and advanced statistics that told a truer story about a player than most general managers were using. With these tools in hand, Beane crafted playoff teams at a fraction of the cost of the big-money teams. Beane's teams consistently ranked in the bottom one-third of the league in terms of payroll, but were consistently at the top of the wins category. This largely went against the "at the time" idea of great teams needing to spend over $100 million easily if they wanted to compete. At the time, and still today to some extent, it is truly astounding to see a low payroll team compete with the top payroll teams.

The NBA has only recently begun to embrace advanced statistics. The once rare plus/minus is now commonplace in box scores. Usage rate and points per possession are quickly catching on as ways to measure a player's efficiency on offense. Most recently, a model has been created by Cervone and D'Amour (2014) from Harvard that shows the expected possession value (EPV) at every instant of a possession. This EPV model uses the vast amount of data collected by SportsVU cameras in every NBA arena to build the model. The EPV, for the first time ever, can measure the expected value or points of that possession based on every player's movement on the court. Every screen, every roll, every pass, every move affects the EPV. The results and insights from this knowledge are largely untapped, but there exists great potential to create incredibly efficient basketball possessions.

Statistics play a large role in the day-to-day, minute-to-minute, and even second-to-second analysis of understanding sports and what is happening. But there is little in terms of answering unknown questions that need optimal or for sure answers. Looking at another aspect of sports, scheduling a sports league, creating hundreds of match-ups with innumerable constraints at play, or determining a schedule by hand is nearly impossible. Thanks to breakthroughs in the OR field with the theory of scheduling sports leagues, OR has become a very popular way to continue to schedule and improve upon existing scheduling methods.

2.2.2 Operations Research in Sports Scheduling

In solving scheduling problems, most assume a round robin schedule. A round robin is where every team plays every other team exactly once. A more balanced schedule that is also used is a double round robin, where each team plays every other team exactly twice, once at home, and once away. The initial problem in sports scheduling dealt with breaks in a schedule. As defined by Schreuder (1980), a break is when a team has two consecutive home games or two consecutive away games. The goal was to minimize the number of breaks in a schedule. Using graph theory and the idea of hard and soft constraints, Schreuder proved that the minimum number of breaks for a league of n teams with n-1 rounds is n-2 breaks. Once Schreuder proved the minimum number of breaks, attention turned toward efficiently and quickly finding feasible and hopefully optimal schedules for different leagues of different sizes.

Regin (1998) took the same round robin model used by Schreuder and attempted to solve the model for different sizes of n (n=4,6,8,…) by using constraint programming to prove the n-2 breaks in an efficient time. For teams up to 60, Regin successfully proved that n-2 is the minimum number of breaks in only seconds. However, for fully constrained problems that more accurately reflect real-world scenarios, for n>12, there is no efficient method to create a schedule using constraint programming.

Nemhauser and Trick (1998) were tasked with finding a feasible and acceptable schedule for ACC basketball. Nemhauser and Trick proposed a 3-step method to finding the best feasible schedules. First, pattern sets of home, away, and byes are created that fit the number of teams in the conference. Second, these pattern sets are assigned games that correspond to the pattern sets created. Once the games are assigned, teams are assigned to the each set of games. One key difference that is present in this model as opposed to other models is that no constraint or combinatorial optimization is used. Instead, enumeration and integer programming is used to search for the best feasible solutions. What makes this, or any specific conference schedule especially difficult is the detailed constraints that come with the schedule. Given these special constraints and the 3 step process used, Nemhauser and Trick successfully showed that this process is a major improvement over the meticulous strategy of constraint programming. While no optimal schedule exists, a handful of very good and diverse schedules exist in an efficient manner. A very similar problem to the ACC scheduling problem is the more generalized traveling tournament problem where team travel is considered into a feasible, as balanced as possible schedule.

Schaerf (1999) took a different approach to working with the round robin tournament. Schaerf first uses *2n* teams and *2n-1* periods (double round robin, not mirrored). Instead of setting up constraint programming, Schaerf proposes first setting up a feasible schedule of dummy teams, where no specific team is in any period, but a pattern is created that satisfies the complementary constraints of a feasible pattern. A modified pattern is used that exists with 6n-6 breaks. Given this pattern, a bipartite graph is created with special attention given to the hard constraints that cannot be violated and the soft constraints that are trying to be minimized. Schaerf found that although specific cases could be solved this way, there is no way to ensure an optimal solution exists in the general case. Further, the general case and the two-step solution are both *NP*-complete problems. Similar to the round robin and double round robin tournaments are the round robin with no set venues or multiple venues (no permanent home stadiums for teams). This added wrinkle creates an entirely new type of problem to solve using the same types of tools.

A further approach to the round robin problem was undertaken by Hamiez and Hao (2000), and involved tabu search heuristic. Hamiez and Hao's idea was to use tabu search on constraint programming solutions as opposed to filtering algorithms and the inefficient branch-and-bound technique. These authors again look at the same round robin tournament from earlier. An initial solution is constructed using graph coloring techniques. From here, two out of the three hard constraints have been satisfied for the round robin tournament, the constraint that every match is different (*ALLDIFF*), and the constraint that every team must play every week, and exactly once every week (*WEEK*). The tabu search aims to satisfy the third constraint, which is that no team can appear more than twice in the same period throughout the tournament (*PERIOD*). Through their specific tabu search, solutions were instanced for n=40 and below, although it was noted that for n>28 the search time and instanced solutions are not efficient and no optimal solutions were approached.

The traveling tournament problem was researched extensively by Easton, Nemhauser, and Trick (2001). In the traveling tournament problem, a feasible acceptable schedule is sought out that satisfies the necessary constraints while also paying special attention to the travel patterns of teams. It is inherently unfair for a team to continuously fly across the country to play home and away games. As opposed to previous tournament problems, this problem is not concerned with minimizing breaks, but travel. What makes this problem truly unique is that it tries to balance a well-studied problem in balancing home and away patterns with another well-studied problem in the traveling salesman, which seeks to minimize travel distance. However, the combination of these two problems creates quite the complex scenario. A main tactic used in this process is to find the minimum lower bound each team must travel, regardless of any other constraints. Given these lower bounds, the next step is to find subsets of the minimum costs that include every city for each team. By taking each team's minimum path and adding them, the overall lower bound on the traveling tournament is established. From here, constraint programming is used to find feasible schedules based on the number of trips in a schedule. Now, a good set of feasible schedules has been created. It is also noted that integer programming can be used. The integer program requires full generation of all travel possibilities, but given the additional constraints that go along with the integer program, the run time is drastically improved from constraint programming.

Urban and Russell (2004) tackled the issue of multiple venues using a two phase constraint program. The multiple venue criterion adds the fact that where teams play is added to the decision making process in choosing a schedule. As such, the problem becomes more difficult. Initially, Urban and Russell used integer programming to solve optimally the problem, but did not find optimal solutions for $n>10$. Further, with the addition of the necessity to designate a home and an away team, integer programming became an inefficient model to use. Constraint programming is used to solve this problem more efficiently. The constraint program takes two phases. Phase one is assigning teams to pairs (games) and those pairs to venues. Phase two attempts to balance the home and away constraint to create an optimal schedule. Using these two phases, optimal solutions for $n=8$, 10, 12, 14, and 16 were all found in an efficient manner, which was a major improvement over the integer programming method. Along the lines of assigning home and away teams, scheduling for specific leagues can prove to be an incredibly difficult challenge when looking at all of the constraints that exist.

2.2.3 Clinching playoffs

Soon after research began to surface on applied OR to sports schedules, another large topic was addressed, which was when a team is eliminated from the playoffs or when a team has made the playoffs. This research and their subsequent findings are closely related to the research and results proposed in this thesis. This type of research uses existing sports schedules and leagues, as well as each sport's unique playoff qualification system to determine when a team has either qualified for the playoffs or been eliminated from the playoffs. OR comes into play because these problems determine the earliest qualification or elimination and in some they even give the scenarios that exist for the desired outcome. Along with the standard playoff qualification, many leagues have tiebreakers in place for when two or more teams tie for a playoff qualification spot. Certain OR playoff models even include one or more tie scenarios where tiebreakers are included in determining optimal playoff conditions.

To start, Gusfield and Martel (1999) analyzed many different sports scenarios and their mathematical playoff models with the idea of a minimum threshold (idea from Kevin Wayne (1999)) being used. They show that a generalized threshold idea exists for every playoff sport where after a certain point, a team cannot attain the minimum threshold required to qualify for the playoffs. At that point, the team is mathematically eliminated from playoff contention. Gusfield and Martel go on to show that this generalized threshold has no immediate connection to any sort of network flow or cut. Additionally, this research shows how network flows can find an undisputed champion in a group by finding the minimum number of games necessary to win outright by winning exactly the minimum number threshold.

Another conclusion drawn is that determining if a given team can be a wildcard team is *NP*-Complete in the general sense. Given a finite number of teams and divisions, a wildcard threshold is determined, and division-eligible teams are not included in the determination of the wildcard threshold.

The research by Gusfield and Martel shares a few common elements to the research presented here, but there are major differences. Gusfield and Martel dealt mainly with one large division and for each team determined its minimum threshold. They also discuss a one-division undisputed winner. Finally, they touch on the possibility of solving for a wildcard threshold in certain situations. The research presented in this paper deals with a multi-division, many-team league. The NFL also has a complicated tiebreaker system involved in determining the possibility of an undisputed champion within a division. This research does not implicitly determine a minimum overall threshold for each team, but instead determines a minimum points needed by each team at that particular point in time. Each team has its own threshold that it is striving for, and at some point they will not be able to attain the minimum points to clinch the division and are thusly eliminated. So, while the elements of threshold and elimination and undisputed winner exist in both types of models, the ideas of those objects are used very differently.

Another approach was taken by Delle Donne and Marenco (2011), who used mixed-integer programming to determine when a racecar driver had qualified for the playoffs. The authors focused their attention on the Argentine TC racing series. In the Argentine TC series, 42 racers compete over 11 races to qualify for 12 playoff spots. The playoff spots are awarded to the 12 drivers with the most cumulative points at the end of the racing season. The points are awarded at the end of every race with first place getting the most, second getting the second most, and so on and so forth. After tenth place, racers are grouped together and awarded points based on finishing position. Over the course of 11 races, each racer's point totals add up and slowly the playoff picture becomes clearer.

Delle Donne and Marenco used mixed integer programming to determine the minimum points a driver needed to secure a position in the playoffs, regardless of the other drivers' positions. The first MIP determines if a racer can win the maximum points possible in the remaining races and not qualify, meaning the racer is fixed at position 13 or worse. The conclusion was that in order for that to occur, the other 12 drivers must all be qualified. The second MIP determines the worst possible position for a driver given a fixed amount of points.

A binary search procedure was used to determine the possible positions for a driver with a fixed amount of points, and through those possibilities, the minimum was found.

The research of Delle Donne and Marenco is again somewhat similar to the research presented in this paper, though differences are apparent. Delle Donne and Marenco deal with playoff scenarios, analyzing the best a racer can be and miss the playoffs, and the worst a racer can be and make the playoffs. Also, they use MIPs to solve their research questions. However auto racing presents fundamental differences from the research presented here. First, auto racing cannot have ties within races; every driver gets a placement and rewarded points accordingly. Also, there is no winning, losing, or division/conferences, just a scale of points. Also, if two racers finish the season with the same number of points, there are no tiebreaker scenarios in the model to deal with assigning. While Delle Donne and Marenco's research touches on the topics of playoff scenarios using MIPs, the difference in sports creates vastly different research findings.

In 2008, Cheng and Steffy developed an integer program to determine if a team had clinched or been eliminated for the National Hockey League. The authors used MIPs to solve this playoff qualification/elimination query. First, for the interested team (they use team *k*), their upper bound on points is capped at the lowest possible, to determine the worst possible finish for *k* and still make the playoffs. Next, Cheng and Steffy also analyzed the maximum points team *k* could have and miss the playoffs; however, they found this question too computationally difficult to discern a solution of real value. Finally, the authors use a tiebreaker scenario and determine whether team *k* can qualify for the playoffs with a subset of tied teams.

Their research showed that in the entire 2004 season, only once could the IP determine that a team was eliminated and one team had qualified for the playoffs prior to the sportscaster's simple math of games remaining. No results were determined using the subset of tied teams although a formulation was given, and no results were determined using more than the simple tiebreaker of most points. Additionally, in any case but the extreme (when one team is doing great or poorly), it is difficult to determine the possibilities among the closely placed teams due to the large number of possible finishes available. Given these extreme and rare results, their research is in little way similar to the research presented in this thesis.

Another NHL paper, by Russell and van Beek, (2011), looks at the NHL Playoff picture from a constrained programming and network flow perspective, with the program solving multiple phases to build to a solution. Two main problems analyzed by the authors are determining the minimum number of points needed to guarantee a playoff spot, and determining the minimum number of points that could possibly qualify for a playoff spot. Russell and van Beek also incorporate the tiebreakers that the NHL has to determine the solutions to their two problems.

Russell and van Beek's research returned results that improved upon the current paper sports media's methods. Their research was the first complete NHL playoff qualification, and thus elimination, problem. Like the research presented in this thesis, Russell and van Beek found solutions for every team at any point in the season as to what their minimum points needed was. Also, the topic of "must win" games was addressed, as the same "critical" games idea is addressed here. However, Russell and van Beek avoided division leaders, did not give a complete schedule of how the results could play out, and used a 3-phase solving method to ultimately find a solution.

2.3 The NFL Playoff Problem

The NFL uses a complicated process involving 12 tiebreakers to determine playoff teams. Teams can qualify for the playoffs by winning their division or winning one of the wildcard spots. This complicated process eliminates the lowest or worst team from contention based on the NFL's qualification procedures. Usually, the team with the best record wins the qualification, but there have been many times where two or more teams are tied with the best record. Based on the NFL's rules for breaking these ties, all tied teams move through the qualification procedure until one team is eliminated. Then, if more than one team still remains, the process starts back over from the top until only one team remains. Given the possibility that multiple teams tie in multiple qualification categories, iterating through many tie-breaker procedures is an extensive process.

In the NFL, each division has a division champion (the team with the best record). Each of these 8 teams automatically qualifies for the playoffs. There are also two wildcard teams in each conference that qualify for the playoffs. The wildcard teams are two teams in the conference that have the best records among remaining teams (not division champions). Throughout the season, two or more teams in a division may have a similar record, which creates a problem with determining a division champion since there is no clear winner. However, if two or more teams are still tied in division, the NFL has created a tiebreaker system to determine which team is the division champion.

The NFL set of tiebreakers mainly uses the idea of Win-Loss Percentage (WLP), which was introduced earlier. The comprehensive list of all 12 NFL tiebreakers, in descending order are: best WLP head-to-head, best WLP in division games, best WLP in common games, best WLP in the conference, best average WLP among victories, best average WLP among all teams played, best combined ranking among conference teams in points scored and points allowed, best combined ranking among all teams in points scored and points allowed, best net points in common games, best net points in all games, best net touchdowns in all games, and finally a coin toss. If two teams are involved in the tiebreaker, once one team has an advantage that team is declared the division champion.

Different rules apply when three or four teams are tied for the division lead. When three or four teams are involved in a tiebreaker, the process is to drop the worst team and restart the process. One team must be completely dominated by the others to be dropped. Once dropped, the remaining teams begin back at the first tiebreaker and proceed until only one team is declared the division champion. For example, three teams are tied overall in WLP head-to-head (with team a>team b>team c>team a), and team a and team b dominate team c in division record, then team c is eliminated, and team a and team b revert back to the head-to-head tiebreaker, where team a would dominate team b and claim the division championship.

Starting with the first tiebreaker, for the best head-to-head WLP tie, the process to determine the winner is straightforward. For this and all subsequent descriptions of tiebreakers, team a is the desired team of the fans. If team a has a better WLP in games played against team b, then team a wins the tiebreaker. If three teams are involved, the scenario would assess how team a performed against team b and team c, b against a and c, and c against a and b. So, if team a lost twice to team b, but beat team c twice, its WLP is .500. If team b and team c both also finish with a .500 record, even though team a lost to team b head-to-head, since all three have equal WLP, then all three would advance to the next tiebreaker, which is best WLP in division games.

Best WLP in division games moves from strictly head-to-head matchups to all division games played by team a and team b. So now team a's WLP includes all division games and this record is compared to how team *b* did against the same division teams. For three teams, each team's WLP in division is compared. It is important to remember in a three or four way tie, one team must be completely dominated to be eliminated. For example, if team *a* has a .750 WLP in division against team *b* and *c*, and team *b* and *c* each have a .500 WLP, there is no way to determine if team *b* or team *c* should be eliminated. So those two teams move on to the next tiebreaker, WLP in common games.

Best WLP in common games looks at the opponents that the tie breaker teams have in common and compares how each team fared against those common opponents. So if team *b* and team *c* each reach this tiebreaker, and team *b* has a better WLP in common games, then team *b* would win this tiebreaker and team *c* would be eliminated. The tiebreaker process would start over between team *a* and team *b* with head-to-head WLP. However, if team *b* and team *c* have similar WLP in common games they would advance to the next tiebreaker: WLP in conference games.

The best WLP in conference compares the games each team played in the conference. These games include all games in division (*b* and *c* are equal), as well as all games outside of the division but in the conference, an additional 6 games. If *b* and *c* have a similar WLP in conference games, they advance to the next tiebreaker, which is WLP among victories.

WLP among victories (strength of victories) looks all the teams each tiebreaker team beat, and averages the WLP of those defeated teams. So if team *b* and team *c* each had 10 victories, then each 10 defeated teams' overall WLP would be averaged and compared for team *b* against team *c*. If team *b*'s defeated teams had a .750 WLP, then clearly team *b* played a very difficult schedule and beat very good teams. And if team *c*'s defeated teams only had a .650

WLP, then team *b* would win the tiebreaker. But, if each team has the same WLP, then the tied teams would move on to WLP among all teams played.

WLP among all games played (strength of schedule) now incorporates all teams played, not just teams that were defeated by a team. Now, if team *b* beat teams with a .750 WLP, but lost to teams with a .125 WLP, then team *b* had good wins but bad losses, and its WLP among all games played would even out to an overall average of all teams on the schedule. So, if team *b* and *c* were tied coming into this tiebreaker, then the deciding factor comes down to which team played a tougher schedule. If both teams had the same strength of schedule, both teams would advance to comparing offensive and defensive rankings within the conference.

After strength of schedule, WLP no longer matters. The next tiebreaker looks at each team's ranking within the conference in terms of points each team scored by over all games and points allowed by each team over all games. If team *a* was 15th in points scored (next to last in conference), and 1st in points allowed, team *a*'s combined ranking would be 16. Each team in the conference would need to have its combined ranking totaled, and those totals are ordered lowest (best) to highest. Comparing this with team *b*'s combined ranking would determine the winner. If both teams are tied in combined offensive and defensive rankings within the conference, the next tiebreaker expands this criterion to include the entire league.

Expanding the combined ranking to include the entire league allows for 32 possible rankings for offense and defense, making the opportunity for a tie even slimmer than in the conference. The same comparison rules and scenarios apply here as in the previous tiebreaker. So now instead of having 16 teams ordered and ranked, all 32 teams are ordered and ranked. Again, if both teams have the same combined ranking, they advance to the next tiebreaker.

The next tiebreaker is comparing the teams by best net points in common games. Net points in a game is defined as the desired team's score minus the opponent's score. For example, if team *a* played team *d*, and the score was 21-17 respectively, team *a* would have net points of 4. If team *d* was the desired team, team *d* would have net points of -4. This tiebreaker compares team *a* versus team *b*'s combined net points for all games in common with each other. If team *a* and team *b* are still tied up to this point, the tiebreaker of best combined net points in common games is expanded to include all games played.

Expanding the previous tiebreaker to include all games takes into account all wins and losses a team had throughout the season. The idea of determining net points remains, except now team *a*'s combined net points is compared to team *b*'s combined net points for all games played. If both teams have the same combined net points in all games, they advance to the next tiebreaker.

The next tiebreaker narrows the net points tiebreaker down to net touchdowns. Net touchdowns is defined as touchdowns scored in a game minus touchdowns allowed in a game. For example, if team *a* scores 3 touchdowns in a game and 24 points overall, and team *d* scores 2 touchdowns and 26 points overall, team *a* would have -2 net points, but +1 net touchdowns. And since this tiebreaker is only concerned with net touchdowns, the net points no longer matter (net points was previously used as a tiebreaker). This tiebreaker is compared for each team in the tiebreaker across all games played, so all of the touchdowns scored by and allowed by team *a* are compared against all of the touchdowns scored by and allowed by team *b*. If both teams are still tied after this tiebreaker, both teams advance to the final tiebreaker.

The final tiebreaker has involves no statistics or records. At this point, a coin flip is used to determine which team is eliminated from the playoffs. For two teams, after the coin flip, one team is eliminated and the other team automatically wins the division. For three or four teams, each team involved in the tiebreaker privately chooses heads or tails. The loser or losers are eliminated. The remaining teams revert back to the first tiebreaker, and if only one team remains, that team wins the division. If all teams choose the same side of the coin, the coin flip is performed again until one or more teams are eliminated.

Throughout the NFL's history, there have been no playoff scenarios that have gone past the strength of victory tiebreaker. This has occurred three times since the NFL adopted its current format of tiebreaker procedures in 2002, according to nfl.com. The strength of victory tiebreakers have also only ever been used in wildcard scenarios, never in a divisional championship scenario. The furthest divisional tiebreaker is the best WLP within the conference, which has also occurred only three times. Given that it is extremely rare that any tiebreaker would ever advance past the best strength of schedule tiebreaker, this research does not analyze any possible scenarios or outcomes beyond this tiebreaker.

Chapter 3 - The NFL True Fan Problem and SPORT Algorithm

This chapter describes the true fan problem and develops integer programs to determine the solutions. These integer programs are combined into a SPORT algorithm, which determines the minimum path to the division championship. Once the minimum path is determined, SPORT finds any critical games. These critical games are used to determine the games and teams that enable the easiest road to the playoffs.

3.1 The True Fan Problem

Every fan wants their team to be the champion and thus the fan's team must qualify for the playoffs. For this chapter, assume that the fan's team is team *a* and that the other teams in division are team *b,* team *c,* and team *d*. For a casual fan, rooting for teams *b*, *c,* and *d* are frowned upon as they are the main opposition to team *a* winning the division. However, a true fan sets aside that culture of unsupported distaste for the other teams when it is in the best interest of team *a*. For example, a fan of team *a* may root for a rivalry team to win a game (e.g. team *b*), if it helps team *a* qualify for the playoffs. Determining what teams have to win which games in order for team *a* to become the division champion is the NFL True Fan Problem

The concept of a minimum playoff path is vital to the NFL True Fan Problem. Given the outcomes for games already played, the minimum point division championship (MPDC) for team *a* is defined as the minimum number of points (2 points for a win, 1 point for a tie) required for team *a* to become the division champion. Any collection of outcomes to all the remaining games where team *a* earns MPDC points and is declared the division champion is called a minimum point path (MPP). If no such MPDC exists, then the team is eliminated from playoff contention.

Observe that MPP may not be the most likely path to become division champion. For instance, a team could make the playoffs by defeating a team that is undefeated or win two games versus teams that have not won. While the second scenario may be more likely, the first scenario is defined as the MPDC.

Knowing that an MPDC exists provides fans with a glimmer of hope that their team could still qualify for the playoffs. An excellent example happened in 2002 in the AFC East. The New York Jets, New England Patriots, and Miami Dolphins all finished with a 9-7 record. With four games left in the season, the Dolphins were at 6-6, the Patriots at 7-5, and the Dolphins at 7-5. The Jets finished by winning two common games with the Patriots and Dolphins, and also beat the Patriots. The Patriots finished the season by winning two division games, including one against Miami, but lost to the Jets. The Dolphins finished the season by winning two common games with the Jets and Patriots, but losing to the Patriots. The final week of the season included the Patriots over the Dolphins in overtime and the Jets over the Green Bay Packers. Again, all three teams finished with a 9-7 record, but the Jets ended up winning the division. All three teams tied in the head-to-head category, Miami was eliminated by the division record category, and New York beat out New England in common games record. All of these tiebreak outcomes were decided in the final couple of weeks, with there being a very limited number of ways each team could win the division.

A key component of the above example was that the Jets had won the "appropriate" games to hold the tie breaker over New England and Miami. These "appropriate" games are called critical games. A game between any two teams is critical if changing the outcome of the game in every MPP increases the required points for team a to become division champion. More formally, a game g between team i and j is said to be critical for team a if one of the two scenarios happen. If team i beats team j and team a's MDCP increases, then the game is critical and the true fan must root for team j. Alternately, if team j beats team i and team a's MDCP increases, then the game is critical and the true fan must root for team i.

The concept of a critical game is vital to the NFL True Fan Problem. A great case of making the playoffs with a low number of wins is the NFL season of 2010, where the Seattle Seahawks made the playoffs by winning the NFC West division with a record of 7 wins and 9 losses. Meanwhile, the New York Giants and Tampa Bay Buccaneers, both in the NFC finished with 10 wins and 6 losses, but did not make the playoffs. Additionally, the St. Louis Rams, also of the NFC West, also finished with 7 wins and 9 losses. However, due to the NFL tiebreaker rules in place, Seattle won the division based on a better division record. Seattle only needed to win 7 games to win the division, and 4 of those 7 came in the division. Other division winners needed 10, 11, and 13 wins for the NFC East, North, and South respectively. Clearly, Seattle had the easiest path to the playoffs by only needing 7 wins. For good measure, in the playoffs Seattle beat an 11 win New Orleans team in the first round. Had Seattle won 10 or 11 games, they would have played the Green Bay Packers, the eventual Super Bowl® champions that year. So, for teams and fans alike, knowing the easiest path to the playoffs can help spur added excitement to the NFL season.

To solve for MPDC and to find MPP, several integer programs are created along with the SPORT algorithm, which identifies the critical games. The SPORT algorithm should be applied once game results are known. Thus, as results of games are obtained, new critical games may appear in a team's schedule. The next section describes the integer programs associated with SPORT.

3.2 IPs Associated with MPDC

To solve for MPDC, 31 integer programs are used. These integer programs look at every possible scenario for team *a* to become the division champion. These 31 IPs exhaustively evaluate all possible scenarios where team *a* could be the division champion. The following arguments describe these exhaustive cases.

An assumption is now made that in the final week, a divisional could score an arbitrarily large number of points. Thus, if team *a* were tied with another team through the first 7 tiebreakers, the other team could hold the 8^{th} tiebreaker (point differential). Consequently, team *a* must win the division championship prior to the 8^{th} tiebreaker and these models reflect this assumption.

Team *a* is declared the division champion if team *a* has a higher tie break against every single divisional team. The IP in Section 3.2.1 describes this IP. All other scenarios must have team *a* not having a tie breaker advantage over every other team in division.

In a three way tie, team *a* can be division champion and still lose a head to head tie breaker with one of the three teams. For this scenario to occur, each of the three teams holds the tie breaker, but their overall record among these three teams is equal. As tie breakers continue, the team that holds the head to head victory over team *a* must be the first team eliminated. Then team *a* holds the head to head advantage over the other team. This scenario can happen six times and is discussed in Section 3.2.2

The final 24 integer programs deal with the 4 types of 4-way head-to-head ties that can occur. Each of the 4 ways again has 6 different scenarios (4x6=24) of which teams among *a*, *b*, *c*, and *d* advance or are eliminated. These cases are exhaustive and more discussion is provided with each specific scenario.

The integer programs seek to minimize the number of points team *a* needs to earn for the season in order to secure the division championship. These points are derived from wins and ties throughout the season. The MPDC often involves team *a* beating one or more teams in one of the NFL division championship tiebreaker scenarios. The outcomes of these scenarios are determined by the integer programs, and reported as part of the MPP. Along with MPP, these integer programs are utilized in the SPORT algorithm to determine the critical games for each week.

The first of the 31 integer programs deals with team *a* defeating teams *b*, *c*, and *d* separately in one or more tiebreakers to become the division champion. For this scenario, *b*, *c*, and *d* must not have WLP greater than *a*. Team *a* can win on overall record and forgo dealing with the tiebreaker scenarios, but this is rarely the case since the goal is the minimum path to the division championship. The integer program takes care of which tiebreakers team *a* beats each of the other teams *j* on, and makes sure team *a* is the division champion.

3.2.1 Dominant Performance IP for Fan's Team

One integer program is the strictly dominant performance by team *a* versus each of the other teams in the division. Team *a* is strictly dominant in this integer program because team *a* is never worse in any tiebreaker scenario to every other team. This allows team *a* to be tied with more than one team in overall record, head-to-head, in division, etc. This means that even though all four teams may be tied in many tiebreakers, team *a* is better than every other team. Thus, team *a* is never eliminated. Once each team has been eliminated, team *a* is then the de facto champion since it dominates each of the other teams in at least one way. This integer program forces team *a* into this exact situation thereby guaranteeing that the minimum number of wins to be division champion is correct. This model also serves as the basis for the integer programs in the next sections.

Single Team Dominance Integer Program

To model the NFL true fan problem, the following decision variables are used.

Decision Variables

$p_{i,j} \epsilon \{0,1,2,3,4\}$ = the points that team i earns against team j for all $i,j \epsilon T$

$q_{i,j,w} \epsilon \{0,1,2\}$= the points that team i earns against team j in week w for all $w \epsilon W$, $i,j \epsilon T$

$r_{i,j,k}$ $\epsilon\{0,1\}$= 1if team i is equal or better than team j for tietype $k \epsilon K$ for all $i,j \epsilon T$

$v_{i,j} \epsilon \{0,1\}$ = 1 if team i beat team j for all $i,j \epsilon T$

$s_{i,j}$ = the points that team i earned towards is strength of victory by beating team j for all $i,j \epsilon T$

Parameters

TV_k is the value assigned to tiebreaker k. Its value is $10^{8-k.}$

$g_{i,j,w} \epsilon \{0,1\}$ = 1 if team i plays team j in week w; 0 if not, for all $w \epsilon S$, $i,j \epsilon T$

$NC_{i,j,m} \epsilon \{0,1\}$ = 1 if team i and team j both play team m; 0 if not, for all $i,j,m \epsilon T$

$C_{i,j,m} \epsilon \{0,1,2\}$ = 2 if team i and team j both play team m twice; 1 if team i and team j both play team m once; and 0 if not, for all $i,j,m \epsilon T$

$O_{i,j} \epsilon \{0,1\}$ = 1 if team i plays team j; 0 if not, for all $i,j \epsilon T$

NFL MPDC Two-Team Dominance Integer Program

Min $\Sigma_{j \epsilon T: j \neq a} p_{a,j}$

Subject to

$\Sigma_{k \epsilon K} TV_k r_{a,j,k} \geq \Sigma_{k \epsilon K} TV_k r_{j,a,k}+1$ for all $j \epsilon D_a$ and $j \neq a$ (2)

$\Sigma_{w \epsilon W} q_{i,j,w}=p_{i,j}$ for all $i,j \epsilon T$ (3)

$q_{i,j,w} + q_{j,i,w} = WO_{i,j,w} + WO_{j,i,w}$ for all $i,j \epsilon T$, $w \epsilon W$ (4)

$p_{i,j}-1 \leq M * v_{i,j}$ for all $i,j \epsilon T$ (5)

$2*v_{i,j} \leq p_{i,j}$ for all $i,j \epsilon T$ (6)

$s_{i,j} \leq M * v_{i,j}$ for all $i,j \in T$ (7)

$s_{i,j} \leq \Sigma_{k \in T}\ p_{j,k}$ for all $i,j \in T$ (8)

$s_{i,j} \geq \Sigma_{k \in T}\ p_{j,k} - M * (1-v_{i,j})$ for all $i,j \in T$ (9)

$\Sigma_{i \in T}\ \Sigma_{w \in W}\ q_{i,i,w} = 0$ (10)

$\Sigma_{i \in T}\ p_{i,j} - \Sigma_{i \in T}\ p_{j,i} + 1 \leq M\ r_{i,j,1}$ for all $i,j \in D_a$ (11a)

$\Sigma_{i \in T}\ p_{j,i} - \Sigma_{i \in T}\ p_{i,j} + 1 \leq M\ r_{j,i,1}$ for all $i,j \in D_a$ (11b)

$\Sigma_{i \in T}\ p_{i,j} - \Sigma_{i \in T}\ p_{j,i} + 1 \leq M\ (1-r_{j,i,1})$ for all $i,j \in D_a$ (11c)

$\Sigma_{i \in T}\ p_{j,i} - \Sigma_{i \in T}\ p_{i,j} + 1 \leq M\ (1-r_{i,j,1})$ for all $i,j \in D_a$ (11d)

$p_{i,j} - p_{j,i} + 1 \leq M\ r_{i,j,2}$ for all $i,j \in D_a$ and $i \neq j$ (12a)

$p_{j,i} - p_{i,j} + 1 \leq M\ r_{j,i,2}$ for all $i,j \in D_a$ and $i \neq j$ (12b)

$p_{i,j} - p_{j,i} \leq M\ (1- r_{j,i,2})$ for all $i,j \in D_a$ and $i \neq j$ (12c)

$p_{j,i} - p_{i,j} \leq M\ (1- r_{i,j,2})$ for all $i,j \in D_a$ and $i \neq j$ (12d)

$\Sigma_{i \in D_a}\ p_{i,j} - \Sigma_{i \in D_a}\ p_{j,i} + 1 \leq M\ r_{i,j,3}$ for all $i,j \in D_a$ and $i \neq j$ (13a)

$\Sigma_{i \in D_a}\ p_{j,i} - \Sigma_{i \in D_a} p_{i,j} + 1 \leq M\ r_{j,i,3}$ for all $i,j \in D_a$ and $i \neq j$ (13b)

$\Sigma_{i \in D_a}\ p_{i,j} - \Sigma_{i \in D_a}\ p_{j,i} \leq M\ (1- r_{j,i,3})$ for all $i,j \in D_a$ and $i \neq j$ (13c)

$\Sigma_{i \in D_a}\ p_{j,i} - \Sigma_{i \in D_a}\ p_{i,j} \leq M\ (1- r_{i,j,3})$ for all $i,j \in D_a$ and $i \neq j$ (13d)

$\Sigma_{m \in T}\ (NC_{i,j,m}\, p_{i,m}) - \Sigma_{m \in T}\ (NC_{j,i,m}\, p_{j,m}) + 1 \leq M\ r_{i,j,5}$ for all $i,j \in D_a$ and $i \neq j$ (14a)

$\Sigma_{m \in T}(NC_{j,i,m}\, p_{j,m}) - \Sigma_{m \in T}(NC_{i,j,m}\, p_{i,m}) + 1 \leq M\ r_{j,i,5}$ for all $i,j \in D_a$ and $i \neq j$ (14b)

$\Sigma_{m \in T}(NC_{i,j,m}\, p_{i,m}) - \Sigma_{m \in T}(NC_{j,i,m}\, p_{j,m}) \leq M\ (1- r_{j,i,5})$ for all $i,j \in D_a$ and $i \neq j$ (14c)

$\Sigma_{m \in T}\ (NC_{j,i,m}\, p_{j,m}) - \Sigma_{m \in T}(NC_{i,j,m}\, p_{i,m}) \leq M\ (1- r_{i,j,5})$ for all $i,j \in D_a$ and $i \neq j$ (14d)

$\Sigma_{i \in C_a}\ p_{i,j} - \Sigma_{i \in C_a}\ p_{j,i} + 1 \leq M\ r_{i,j,4}$ for all $i,j \in D_a$ and $i \neq j$ (15a)

$\Sigma_{i \in C_a}\ p_{j,i} - \Sigma_{i \in C_a}\ p_{i,j} + 1 \leq M\ r_{j,i,4}$ for all $i,j \in D_a$ and $i \neq j$ (15b)

$\Sigma_{i \in C_a}\ p_{i,j} - \Sigma_{i \in C_a}\ p_{j,i} \leq M\ (1- r_{j,i,4})$ for all $i,j \in D_a$ and $i \neq j$ (15c)

$\Sigma_{i \in C_a}\ p_{j,i} - \Sigma_{i \in C_a}\ p_{i,j} \leq M\ (1- r_{i,j,4})$ for all $i,j \in D_a$ and $i \neq j$ (15d)

$\Sigma_{m \in T}\ s_{i,m} - \Sigma_{m \in T}\ s_{j,m} + 1 \leq M\ r_{i,j,6}$ for all $i,j \in D_a$ and $i \neq j$ (16a)

$\Sigma_{m \in T}\ s_{j,m} - \Sigma_{m \in T}\ s_{i,m} + 1 \leq M\ r_{j,i,6}$ for all $i,j \in D_a$ and $i \neq j$ (16b)

$\Sigma_{m \in T}\ s_{i,m} - \Sigma_{m \in T}\ s_{j,m} \leq M\ (1- r_{j,i,6})$ for all $i,j \in D_a$ and $i \neq j$ (16c)

$\Sigma_{m \in T}\ s_{j,m} - \Sigma_{m \in T}\ s_{i,m} \leq M\ (1- r_{i,j,6})$ for all $i,j \in D_a$ and $i \neq j$ (16d)

$\Sigma_{m \in T}\ O_{i,m}\, \Sigma_{n \in T}\ p_{m,n} - \Sigma_{m \in T}\ O_{j,m}\, \Sigma_{n \in T}\ p_{m,n} + 1 \leq M\ r_{i,j,7}$ for all $i,j \in D_a$ and $i \neq j$ (17a)

$\Sigma_{m \in T}\ O_{j,m}\, \Sigma_{n \in T}\ p_{m,n} - \Sigma_{m \in T}\ O_{i,m}\, \Sigma_{n \in T}\ p_{m,n} + 1 \leq M\ r_{j,i,7}$ for all $i,j \in D_a$ and $i \neq j$ (17b)

$\Sigma_{m\epsilon T}\ O_{i,m}\ \Sigma_{n\epsilon T}\ p_{m,n} - \Sigma_{m\epsilon T}\ O_{j,m}\ \Sigma_{n\epsilon T}\ p_{m,n} \ \leq M\ r_{j,i,7}$ for all $i,j \ \epsilon\ D_a$ and $i{\neq}j$ (17c)

$\Sigma_{m\epsilon T}\ O_{j,m}\ \Sigma_{n\epsilon T}\ p_{m,n} - \Sigma_{m\epsilon T}\ O_{i,m}\ \Sigma_{n\epsilon T}\ p_{m,n} \ \leq M\ r_{i,j,7}$ for all $i,j \ \epsilon\ D_a$ and $i{\neq}j$ (17d)

$r_{i,j,k}\ \epsilon\{0,1\}$ for all i,j,k

$q_{i,j,w}\ \epsilon\ \{0,1,2\}$ for all i,j,w

$p_{i,j}\ \epsilon\ \{0,1,2,3,4\}$ for all i,j

$v_{i,j}\ \epsilon\{0,1\}$ for all i,j

$s_{i,j}\ \epsilon\mathbf{Z}^+$ for all i,j

where M is an arbitrarily large value.

The objective function minimizes the total number of points that team i needs to gain from all other j teams. The constraints force that team i is the division champion. Thus, this IP seeks the minimum number of points that enable a team to be the division champion by dominating each team in its division on a tiebreaker.

The team must win the division over a set of tiebreakers (2), but does not have to win on any specific tiebreaker. Therefore, the team's tiebreaker score must be larger than every other team in its division. This is shown through the idea that the sum of the tie rankings multiplied by each ranking's value for the team must be greater than or equal to one plus each opponent's sum of its tie rankings multiplied by each ranking's value, thereby ensuring that the team's overall tie score will be larger, meaning it will be the division winner. This constraint does not force any tiebreaker score to happen, it simply forces one (or more) of the tiebreakers to occur in favor of team *a* over each team in the division.

Critical to this formulation is the value assigned to tie rankings. The highest tie ranking value must have a value that is strictly larger than the sum of all other tie ranking value. Thus, assigning a tie ranking value that doubles each time is sufficient. That is, tie ranking 6 has a value of 1, but tie ranking 1 has a value of 64. Thus, having all lesser tie breakers can never overcome the highest ranked tiebreaker.

Constraint set (3) shows the relationship between points (simply points *i* earns from *j*) and weekpoints (points *i* earns from *j* in week *w*). This constraint simply states that the sum of the weekpoints of *i* from *j* across all weeks should equal the total points *i* earns from *j*. While these variables could be removed, having this variable enables a simpler explanation of the model.

This fourth constraint shows that if two teams play each other, the most points they can earn for *weekpoints* is equal to 2. So, for each week *w*, if team *i* and *j* play, then there are exactly 2 *weekpoints* available; otherwise, each slot *w* in the $q_{i,j,w}$ array gets assigned a 0 since *i* and *j* did not play. Constraints (5) and (6) ensure that $v_{i,j}$ only equals 1 when team *i* beats team *j*. Constraints (7), (8), and (9) tie together $v_{i,j}$ and $s_{i,j}$. These constraints force $s_{i,j}$ to be exactly equal to the sum of the number of points of each team that *i* beat. Constraint (10) simply states that team *i* cannot earn any *weekpoints* from itself, because it does not play itself.

To replicate the tie breakers in the NFL into and IP format, a collection of four constraints are used. The equations consist of a left hand side that finds the difference between two teams based on the particular tie breaker, such as total points, points in division, or points in conference. The combination of the four right hand sides force the specific $r_{i,j,k}$ value based on the difference value on the left hand side. For each set of tie breaker constraints, three possible outcomes are created. First, if team *i* value – team *j* value is positive, team *i* gets a 1 for the corresponding $r_{i,j,k}$ and team *j* gets 0. Conversely, if team *i* value - team *j* value is negative, team *j* gets a 1 for the corresponding $r_{i,j,k}$ and team *i* gets 0. Lastly, if team *i* value - team *j* value is 0, both teams get a 1 for the corresponding $r_{i,j,k}$.

This set of constraints (11) sets up the first available tie-breaker between two teams, which is that team *i* has a better record than team *j*. To create this, four constraints are used.

Assume team i has more points than team j. The left hand side of the first constraint is positive and $r_{i,j,1}$ must be 1. The left hand side of the second constraint is positive and $r_{j,i,1}$ must be 0. The left hand sides of the third and fourth constraint are 0 or negative, thus $r_{j,i,1}$ and $r_{i,j,1}$ can be either 0 or 1. Based on this scenario, $r_{i,j,1}$ is 1 and $r_{j,i,1}$ is 0, which is correct.

Now, assume team j has more points than team i. The left hand side of the first constraint is 0 or negative and the $r_{i,j,1}$ can be 0 or 1. The left hand side of the second constraint is negative, and the $r_{j,i,1}$ can be 0 or 1. The left hand sides of the third and fourth constraints are strictly positive, and thus the $r_{j,i,1}$ is 1 and the $r_{i,j,1}$ is 0. For this scenario, $r_{i,j,1}$ is 0 and $r_{j,i,1}$ is 1.

Finally, assume team i and team j have equal points. Now, the left hand side of the first equation is positive, and the $r_{i,j,1}$ must be 1. The left hand side of the second constraint is 0, and the $r_{j,i,1}$ can be 0 or 1. The left hand side of the third constraint is positive, and the $r_{j,i,1}$ must be 1. The left hand side of the fourth constraint is 0 and the $r_{i,j,1}$ can be either 0 or 1. For this scenario, both $r_{i,j,1}$ and $r_{j,i,1}$ are 1.

The second set of tie breaker constraints (12) follows the same logic as the first set except this set deals with points only earned between team i and team j (head-to-head points). Thus, the left hand side of these constraints involve $p_{i,j}$ - $p_{j,i}$, the points that team i earned from team j. There are no teams being summed over here as this constraint only deals with team i versus each team j.

The third set of tie breaker constraints (13) follows previous logic and deals with points team i earned in the division versus team j's points in the division. For these constraints, the left hand side is concerned with points team i has earned from all teams in its division against the points team j earned against all teams in the division. So, the constraint still has $p_{i,j} - p_{j,i}$ except team j is summed over all teams in the division (D_i). Now, the difference in points between teams i and j reflects the difference in points earned in the division.

The fourth set of tie breaker constraints (14) follows previous logic and this set deals with points team i earned in games in common with team j, and points team j earned in games in common with team i. For this set of constraints, the left hand side compares $p_{i,j} - p_{j,i}$ with j being summed across games in common with i. A game in common is from the $NC_{i,j,m}$ data. For team i, if both team i and team j both have a game in common with team m, then the $NC_{i,j,m}$ becomes a 1, and this is multiplied by the points team i earned from team m. Thus, if team i played and beat team m, but team j did not play team m, then despite team i earning 2 points from team m, the $NC_{i,j,m}$ factor would be 0, and the overall effect would be a 0 for team i's left hand side.

The fifth set of tie breaker constraints (15) follows previous logic and this set of four constraints deals with points team i earned in the conference versus each team j's points in the conference. For these constraints, the left hand side is extended to include all teams played in the entire conference, not just games in the division and games in common. The left hand side shows this by having the points team i earned from all teams j in the conference and comparing the points earned to team j's points earned against all teams in the conference as well. The comparison is now $p_{i,j} - p_{j,i}$ with the j being summed across all teams in the conference.

The sixth set of tie breaker constraints (16) follows previous logic and this set deals with points earned from every team that team i defeated (strength of victory) against points earned by every team that team j defeated. For this set of constraints, the left hand side compares $s_{i,k} - s_{j,k}$ with k being summed across all teams. Victory points are the total points team k ends the season with, and are added to team i's strength of victory only if team i beat team k during the season.

If team i has more victory points than team j, and thus a better strength of victory, team i is rewarded with the tiebreaker points for this tiebreaker.

The seventh set of tie breaker constraints (17) follows previous logic and this set deals with strength of schedule. Strength of schedule is similar to strength of victory, except these constraints take into account all teams played, not just teams beat. The left hand side, $p_{m,n}$, looks at every opponent m that team i plays, $O_{i,m}$, and sums the points earned, $\Sigma_{n \epsilon T}$, by every opponent of team i, regardless if team i beat that opponent. This is compared to the points that every opponent of team j plays, $O_{j,m}$. If team i has a better strength of schedule (team i's opponents have more combined points than team j's), then team i is rewarded the tiebreaker points for this tiebreaker.

3.2.2 Three-Way Tie Scenario IP

Once three teams are allowed to tie in division, the tiebreaker determination process becomes quite complicated. The NFL employs a last team standing approach where one by one each team is eliminated until only one team is left. This method is different than, say, finding the first team to dominate the other teams in a tiebreaker. For the 3-way integer programs used, team a can no longer dominate each team directly, as this was covered in the previous case.

Table 3.2.2-1. Three Way Tiebreaker Case

3 Way Tie				
	a	b	c	d
a	x	4	0	3
b	0	x	4	3
c	4	0	x	3
d	1	1	1	x

For symmetry, assume teams *b* and *c* are tied with team *a*. This case has team *a* losing a head to head tiebreaker to a tied team and yet still being declared the division champion. This occurs if teams *a*, *b*, and *c* are all tied in win-loss-tie record and head to head. Team *a* dominates team *b* head-to-head, team *b* dominates team *c* head-to-head, and team *c* dominates team *a* head-to-head. If team c is eliminated first, then team *a* dominates team *b* on head-to-head. Thus, team *c* must be eliminated first, requiring both team *a* and team b have tiebreaker advantage over team c after the head-to-head tiebreaker. The following IP presents this case.

Three-Way Tie Scenario Integer Program

Min $\Sigma_{j \in T:j \neq a} p_{a,j}$

Subject to

$\Sigma_{w \in W} q_{i,j,w} = p_{i,j}$ for all $i,j \in T$ (3)

$q_{i,j,w} + q_{j,i,w} = WO_{i,j,w} + WO_{j,i,w}$ for all $i,j \in T$, $w \in W$ (4)

$p_{i,j} - 1 \leq M * v_{i,j}$ *i* for all $i,j \in T$ (5)

$2 * v_{i,j} \leq p_{i,j}$ for all $i,j \in T$ (6)

$s_{i,j} \leq M * v_{i,j}$ for all $i,j \in T$ (7)

$s_{i,j} \leq \Sigma_{k \in T}\ p_{j,k}$ for all $i,j \in T$ (8)

$s_{i,j} \geq \Sigma_{k \in T}\ p_{j,k} - M * (1 - v_{i,j})$ for all $i,j \in T$ (9)

$\Sigma_{i \in T} \Sigma_{w \in W} q_{i,i,w} = 0$ (10)

$\Sigma_{i \in D_a} p_{i,j} - \Sigma_{i \in D_a} p_{j,i} + 1 \leq M\, r_{i,j,3}$ for all $i,j \in D_a$ and $i \neq j$ (13a)

$\Sigma_{i \in D_a} p_{j,i} - \Sigma_{i \in D_a} p_{i,j} + 1 \leq M\, r_{j,i,3}$ for all $i,j \in D_a$ and $i \neq j$ (13b)

$\Sigma_{i \in D_a} p_{i,j} - \Sigma_{i \in D_a} p_{j,i} \leq M\, (1 - r_{j,i,3})$ for all $i,j \in D_a$ and $i \neq j$ (13c)

$\Sigma_{i \in D_a} p_{j,i} - \Sigma_{i \in D_a} p_{i,j} \leq M\, (1 - r_{i,j,3})$ for all $i,j \in D_a$ and $i \neq j$ (13d)

$\Sigma_{m \in T} (NC_{i,j,m}\, p_{i,m}) - \Sigma_{m \in T} (NC_{j,i,m}\, p_{j,m}) + 1 \leq M\, r_{i,j,5}$ for all $i,j \in D_a$ and $i \neq j$ (14a)

$\Sigma_{m \in T}(NC_{j,i,m}\, p_{j,m}) - \Sigma_{m \in T}(NC_{i,j,m}\, p_{i,m}) + 1 \leq M\, r_{j,i,5}$ for all $i,j \in D_a$ and $i \neq j$ (14b)

$\Sigma_{m \in T}(NC_{i,j,m}\, p_{i,m}) - \Sigma_{m \in T}(NC_{j,i,m}\, p_{j,m}) \leq M\, (1 - r_{j,i,5})$ for all $i,j \in D_a$ and $i \neq j$ (14c)

$\Sigma_{m \in T} (NC_{j,i,m}\, p_{j,m}) - \Sigma_{m \in T}(NC_{i,j,m}\, p_{i,m}) \leq M\, (1 - r_{i,j,5})$ for all $i,j \in D_a$ and $i \neq j$ (14d)

$\Sigma_{i \in C_a} p_{i,j} - \Sigma_{i \in C_a} p_{j,i} + 1 \leq M\, r_{i,j,4}$ for all $i,j \in D_a$ and $i \neq j$ (15a)

$\Sigma_{i\epsilon C_a} p_{j,i}$ - $\Sigma_{i\epsilon C_a} p_{i,j}$ + $1 \leq M\ r_{j,i,4}$ for all $i,j\ _\epsilon\ D_a$ and $i \neq j$ (15b)

$\Sigma_{i\epsilon C_a} p_{i,j}$ - $\Sigma_{i\epsilon C_a} p_{j,i}$ $\leq M\ (1- r_{j,i,4})$ for all $i,j\ _\epsilon\ D_a$ and $i \neq j$ (15c)

$\Sigma_{i\epsilon C_a} p_{j,i}$ - $\Sigma_{i\epsilon C_a} p_{i,j}$ $\leq M\ (1- r_{i,j,4})$ for all $i,j\ _\epsilon\ D_a$ and $i \neq j$ (15d)

$\Sigma_{m\epsilon T}\ s_{i,m}$ - $\Sigma_{m\epsilon T}\ s_{j,m}$ + $1 \leq M\ r_{i,j,6}$ for all $i,j\ _\epsilon\ D_a$ and $i \neq j$ (16a)

$\Sigma_{m\epsilon T}\ s_{j,m}$ - $\Sigma_{m\epsilon T}\ s_{i,m}$ + $1 \leq M\ r_{j,i,6}$ for all $i,j\ _\epsilon\ D_a$ and $i \neq j$ (16b)

$\Sigma_{m\epsilon T}\ s_{i,m}$ - $\Sigma_{m\epsilon T}\ s_{j,m}$ $\leq M\ (1- r_{j,i,6})$ for all $i,j\ _\epsilon\ D_a$ and $i \neq j$ (16c)

$\Sigma_{m\epsilon T}\ s_{j,m}$ - $\Sigma_{m\epsilon T}\ s_{i,m}$ $\leq M\ (1- r_{i,j,6})$ for all $i,j\ _\epsilon\ D_a$ and $i \neq j$ (16d)

$\Sigma_{m\epsilon T}\ O_{i,m}\ \Sigma_{n\epsilon T}\ p_{m,n}$ - $\Sigma_{m\epsilon T}\ O_{j,m}\ \Sigma_{n\epsilon T}\ p_{m,n}$ + $1 \leq M\ r_{i,j,7}$ for all $i,j\ _\epsilon\ D_a$ and $i \neq j$ (17a)

$\Sigma_{m\epsilon T}\ O_{j,m}\ \Sigma_{n\epsilon T}\ p_{m,n}$ - $\Sigma_{m\epsilon T}\ O_{p,m}\ \Sigma_{n\epsilon T}\ p_{m,n}$ + $1 \leq M\ r_{j,i,7}$ for all $i,j\ _\epsilon\ D_a$ and $i \neq j$ (17b)

$\Sigma_{m\epsilon T}\ O_{i,m}\ \Sigma_{n\epsilon T}\ p_{m,n}$ - $\Sigma_{m\epsilon T}\ O_{j,m}\ \Sigma_{n\epsilon T}\ p_{m,n}$ $\leq M\ r_{j,i,7}$ for all $i,j\ _\epsilon\ D_a$ and $i \neq j$ (17c)

$\Sigma_{m\epsilon T}\ O_{j,m}\ \Sigma_{n\epsilon T}\ p_{m,n}$ - $\Sigma_{m\epsilon T}\ O_{i,m}\ \Sigma_{n\epsilon T}\ p_{m,n}$ $\leq M\ r_{i,j,7}$ for all $i,j\ _\epsilon\ D_a$ and $i \neq j$ (17d)

$\Sigma_{k\epsilon K}\ (TV_k\ r_{b,c,k}) \geq \Sigma_{k\epsilon K}\ (TV_k\ r_{c,b,k})$ +1 (18a)

$\Sigma_{k\epsilon K}\ (TV_k\ r_{a,c,k}) \geq \Sigma_{k\epsilon K}\ (TV_k\ r_{c,a,k})$ +1 (18b)

$\Sigma_{i\epsilon D_a}\ \Sigma_{j\epsilon D_a}\ r_{i,j,1}$ =1 for all $i,j \neq d$ (19a)

$\Sigma_{i\epsilon D_a}\ \Sigma_{j\epsilon D_a}\ r_{i,j,2}$ =1 for all $i,j \neq d$ (19b)

$\Sigma_{j\epsilon T}\ (p_{a,j} - p_{c,j}\ )$ =0 (20a)

$\Sigma_{j\epsilon T}\ (p_{b,j} - p_{c,j}\ )$ =0 (20b)

$\Sigma_{j\epsilon T}\ p_{d,j} \leq \Sigma_{j\epsilon T}\ p_{a,j}$ -1 (21)

$p_{a,b} \geq 3$ (22a)

$p_{b,c} \geq 3$ (22b)

$p_{c,a} \geq 3$ (22c)

$p_{a,b} - p_{b,c} = 0$ (23a)

$p_{c,a} - p_{a,b} = 0$ (23b)

$r_{i,j,k}\ \epsilon\{0,1\}$ for all i,j,k

$q_{i,j,w}\ \epsilon\ \{0,1,2\}$ for all i,j,w

$p_{i,j}\ \epsilon\ \{0,1,2,3,4\}$ for all i,j

$v_{i,j}\ \epsilon\{0,1\}$ for all i,j

$s_{i,j}\ \epsilon\mathbf{Z}^+$ for all i,j

where M is an arbitrarily large value.

This IP is similar to the two team dominant IP from the previous section. The constraints for the WLP tiebreaker (constraint set 10) and the head-to-head tiebreaker (constraint set 11) are removed, and their variables are set to 1 for the appropriate teams to help force the 3-way time. Additionally, constraint 2 has been removed as team *a* no longer needs to dominate every team in overall tiebreakers. Several other constraints are added as well to force this scenario.

Constraints 18a and 18b allow one of the two teams not team *a* to be dominant in one of the tiebreakers, allowing team *a* to then dominate that remaining team in head-to-head. Also, team *a* must dominate the least of the three teams as well, as team *a* cannot be eliminated first. This constraint exists so that team *a* has a team to eventually dominate in tiebreakers. If the other two teams remain tied throughout all tiebreakers, then team *a* would never dominate or win the division. The left side of the inequality is sum of the tiebreakers multiplied by each tiebreakers value for the tiebreaker score of team *b* versus team *c*. The right hand side is the same value, except for team *c* versus team *b*, plus 1. This inequality states that the total tie value of team *b* against team *c* must be at least 1 better than team *c*'s total tie value, ensuring team *b* is better in at least one tiebreaker category. Similar logic follows for team *a* also dominating team *c*.

The next set of constraints, 19a and 19b, force the tiebreaker variables for the first two tiebreakers to be equal to 1 for the three teams involved in the 3-way tie. This is due to the fact that constraint sets 10 and 11 from the two-team dominant IP were removed. Since these constraints were removed, nothing was forcing a value for these first two variables and they could then be a 1 or a 0 for no reason, which violates the validity of the model.

Following those constraints are constraints 20a, 20b. These constraints ensure that the three teams in the tie scenario have the same overall record. The left hand side compares the sum of all the points earned by two of the three teams in the tie scenario, and the difference in total points must be 0 to ensure that both overall records are equal. By the transitive property, having two such equalities ensures all three teams are equal.

Constraint 21 forces the only team not in the tie scenario to be dominated on overall points by team *a*, ensuring that a four-way tie scenario is not created. And by constraint set 20, the other two teams will also dominate the fourth team. The left hand side of the inequality is the total points that fourth team has earned, and the right hand side is the total points that team *a* has earned. The inequality states that the total points of the fourth team must be at best 1 point worse than team *a*'s total points, ensuring the fourth team is strictly worse on overall points.

The next set of constraints, 22a, 22b, and 22c force the head-to-head tiebreaker to be equal across all three teams in the three-way tie scenario. This set of constraints deals with each team earning at least 3 points from one of the other teams in the tie scenario. These constraints are needed so that one team is not eliminated in the head-to-head tiebreaker and no three-way tie exists. The left hand side of the inequality of the constraints is the points one team earned from another team, with each team dominating a different team (a team cannot both dominate and be dominated by the same team). The right hand side of the inequality is 3, since each team must outright dominate one of the others in head-to-head.

The next set of constraints, 23a and 23b ensure that all three teams gain the same amount from their non-dominated team in the three-way tie. The left hand side is the points difference between two teams earned from their non-dominated team, and the right hand side is zero to ensure that the points earned are equal between the two teams. Both constraints with the transitive property ensure all three teams are equal. One team cannot gain 3 points from one team and have that team gain 4 from the third team; this would violate the equality constraint as one team would have to finish with more points head-to-head.

3.2.3 Four-Way Tie Scenarios and IPs

This section provides an IP model that provides the division champion when all four teams have the same W-L-T record at the end of the regular season. This scenario is unlikely and has never occurred in the NFL. However, this scenario may still occur and could provide a glimmer of hope for a team and thus it is included here.

The complexity involved with 4-way ties in division is evident by the four times increase in the number of integer programs required to solve every situation from the 3-way tie scenarios. Similar to the 3-way tie IP, the case presented here is when team *d* is eliminated first, then team *c* and finally team *b*. Once such a scenario occurs, there are six ways to permute teams *b*, *c* and *d*.

In this scenario, all four teams have the same WLP record and have the same head to head, otherwise the scenario falls into either or the two previous cases. Since all four teams are tied head to head, they must also be tied in division. There are four cases to consider under these assumptions. Prior to giving these cases, the basic 4-way tie IP is given. Each case adds on a few special constraints.

4-Way Tie Core IP

Min $\Sigma_{j \in T: j \neq a} p_{a,j}$

Subject to

$\Sigma_{w \in W} q_{i,j,w} = p_{i,j}$ for all $i,j \in T$ (3)

$q_{i,j,w} + q_{j,i,w} = WO_{i,j,w} + WO_{j,i,w}$ for all $i,j \in T$, $w \in W$ (4)

$p_{i,j} - 1 \leq M * v_{i,j}$ i for all $i,j \in T$ (5)

$2 * v_{i,j} \leq p_{i,j}$ for all $i,j \in T$ (6)

$s_{i,j} \leq M * v_{i,j}$ for all $i,j \in T$ (7)

$s_{i,j} \leq \Sigma_{k \epsilon T}\ p_{j,k}$ for all $i,j\ _{\epsilon}\ T$ (8)

$s_{i,j} \geq \Sigma_{k \epsilon T}\ p_{j,k} - M * (1\text{-}v_{i,j})$ for all $i,j\ _{\epsilon}\ T$ (9)

$\Sigma_{i\epsilon T}\ \Sigma_{w\epsilon W}\ q_{i,i,w} = 0$ (10)

$\Sigma_{m\epsilon T}\ (NC_{i,j,m}\ p_{i,m})\text{-}\ \Sigma_{m\epsilon T}\ (NC_{j,i,m}\ p_{j,m}) + 1 \leq M\ r_{i,j,5}$ for all $i,j\ _{\epsilon}\ D_a$ and $i \neq j$ (14a)

$\Sigma_{m\epsilon T}(NC_{j,i,m}\ p_{j,m})\text{-}\Sigma_{m\epsilon T}(NC_{i,j,m}\ p_{i,m}) +\ 1 \leq M\ r_{j,i,5}$ for all $i,j\ _{\epsilon}\ D_a$ and $i \neq j$ (14b)

$\Sigma_{m\epsilon T}(NC_{i,j,m}\ p_{i,m})\text{-}\Sigma_{m\epsilon T}(NC_{j,i,m}\ p_{j,m}) \leq M\ (1\text{-}\ r_{j,i,5})$ for all $i,j\ _{\epsilon}\ D_a$ and $i \neq j$ (14c)

$\Sigma_{m\epsilon T}\ (NC_{j,i,m}\ p_{j,m})\text{-}\Sigma_{m\epsilon T}(NC_{i,j,m}\ p_{i,m}) \leq M\ (1\text{-}\ r_{i,j,5})$ for all $i,j\ _{\epsilon}\ D_a$ and $i \neq j$ (14d)

$\Sigma_{i\epsilon C_a}\ p_{i,j}\ -\ \Sigma_{i\epsilon C_a}\ p_{j,i}\ +\ 1 \leq M\ r_{i,j,4}$ for all $i,j\ _{\epsilon}\ D_a$ and $i \neq j$ (15a)

$\Sigma_{i\epsilon C_a}\ p_{j,i}\ -\ \Sigma_{i\epsilon C_a}\ p_{i,j}\ +\ 1 \leq M\ r_{j,i,4}$ for all $i,j\ _{\epsilon}\ D_a$ and $i \neq j$ (15b)

$\Sigma_{i\epsilon C_a}\ p_{i,j}\ -\ \Sigma_{i\epsilon C_a}\ p_{j,i}\ \leq M\ (1\text{-}\ r_{j,i,4})$ for all $i,j\ _{\epsilon}\ D_a$ and $i \neq j$ (15c)

$\Sigma_{i\epsilon C_a}\ p_{j,i}\ -\ \Sigma_{i\epsilon C_a}\ p_{i,j}\ \leq M\ (1\text{-}\ r_{i,j,4})$ for all $i,j\ _{\epsilon}\ D_a$ and $i \neq j$ (15d)

$\Sigma_{m\epsilon T}\ s_{i,m}\ -\ \Sigma_{m\epsilon T}\ s_{j,m}\ +\ 1 \leq M\ r_{i,j,6}$ for all $i,j\ _{\epsilon}\ D_a$ and $i \neq j$ (16a)

$\Sigma_{m\epsilon T}\ s_{j,m}\ -\ \Sigma_{m\epsilon T}\ s_{i,m}\ +\ 1 \leq M\ r_{j,i,6}$ for all $i,j\ _{\epsilon}\ D_a$ and $i \neq j$ (16b)

$\Sigma_{m\epsilon T}\ s_{i,m}\ -\ \Sigma_{m\epsilon T}\ s_{j,m}\ \leq M\ (1\text{-}\ r_{j,i,6})$ for all $i,j\ _{\epsilon}\ D_a$ and $i \neq j$ (16c)

$\Sigma_{m\epsilon T}\ s_{j,m}\ -\ \Sigma_{m\epsilon T}\ s_{i,m}\ \leq M\ (1\text{-}\ r_{i,j,6})$ for all $i,j\ _{\epsilon}\ D_a$ and $i \neq j$ (16d)

$\Sigma_{m\epsilon T}\ O_{i,m}\ \Sigma_{n\epsilon T}\ p_{i,n} - \Sigma_{m\epsilon T}\ O_{j,m}\ \Sigma_{n\epsilon T}\ p_{m,n} + 1 \leq M\ r_{i,j,6}$ for all $i,j\ _{\epsilon}\ D_a$ and $i \neq j$ (17a)

$\Sigma_{m\epsilon T}\ O_{j,m}\ \Sigma_{n\epsilon T}\ p_{j,n} - \Sigma_{m\epsilon T}\ O_{p,m}\ \Sigma_{n\epsilon T}\ p_{m,n} + 1 \leq M\ r_{j,i,6}$ for all $i,j\ _{\epsilon}\ D_a$ and $i \neq j$ (17b)

$\Sigma_{m\epsilon T}\ O_{i,m}\ \Sigma_{n\epsilon T}\ p_{i,n} - \Sigma_{m\epsilon T}\ O_{j,m}\ \Sigma_{n\epsilon T}\ p_{m,n}\ \leq M\ r_{j,i,6}$ for all $i,j\ _{\epsilon}\ D_a$ and $i \neq j$ (17c)

$\Sigma_{m\epsilon T}\ O_{j,m}\ \Sigma_{n\epsilon T}\ p_{j,n} - \Sigma_{m\epsilon T}\ O_{i,m}\ \Sigma_{n\epsilon T}\ p_{m,n}\ \leq M\ r_{i,j,6}$ for all $i,j\ _{\epsilon}\ D_a$ and $i \neq j$ (17d)

$\Sigma_{k\epsilon T}\ p_{i,k} = \Sigma_{k\epsilon T}\ p_{j,k}$ for all $i,j\ _{\epsilon}\ D_a$ and $i \neq j$ (24)

$\Sigma_{k\epsilon D_a}\ p_{i,k}\ =\ \Sigma_{k\epsilon D_a}\ p_{j,k}$ for all $i,j\ _{\epsilon}\ D_a$ and $i \neq j$ (25)

$r_{i,j,1} = 1$ for all $i,j\ _{\epsilon}\ D_a$ and $i \neq j$ (26a)

$r_{i,j,2} = 1$ for all $i,j\ _{\epsilon}\ D_a$ and $i \neq j$ (26b)

$r_{i,j,3} = 1$ for all $i,j\ _{\epsilon}\ D_a$ and $i \neq j$ (26c)

$r_{i,j,k}\ \epsilon\{0,1\}$ for all i,j,k

$q_{i,j,w}\ \epsilon\ \{0,1,2\}$ for all i,j,w

$p_{i,j}\ \epsilon\ \{0,1,2,3,4\}$ for all i,j

$v_{i,j}\ \epsilon\{0,1\}$ for all i,j

$s_{i,j}\ \epsilon\mathbf{Z}^{+}$ for all i,j

where M is an arbitrarily large value.

The only difference in the 4-Way Tie IP from the Dominant IP is that there is no tie breaker for W-L-T (constraint set 10), head-to-head (constraint set 11), or division record (constraint set 12). Rather constraint (24) requires all teams in division to have the same W-L-T record and (25) forces the same head to head and division record. The final added constraint set, 26a, 26b, and 26c, is so that all teams' first 3 tiebreaker variables are set to 1. Additional constraints are added to this IP to force certain scenarios. The different cases of eventual dominance are given in Table 3.2.2-2. For simplicity, the cases presented in this table do not consider ties. However, the IPs do allow for the tie scenarios by having points greater than or equal to 3. The cases below describe the exact order of how teams must be eliminated to assure that team *a* is the division champion. Any deviation from this order would result in an alternate division champion.

Table 3.2.2-2 Four-Way Tiebreaker Cases

Case 1.						Case 2.				
	a	b	c	d			a	b	c	d
a	x	4	2	0		a	0	4	0	2
b	0	x	4	2		b	0	x	4	2
c	2	0	x	4		c	4	0	x	2
d	4	2	0	x		d	2	2	2	x
Case 3.						**Case 4.**				
	a	b	c	d			a	b	c	d
a	x	4	0	2		a	x	2	2	2
b	0	x	4	2		b	2	x	4	0
c	4	0	x	2		c	2	0	x	4
d	2	2	2	x		d	2	4	0	x

Case 1 shows that each team dominates exactly one other team in head-to-head. In this case, team *a* wins the division if team *d* is eliminated first, because team *c* is then eliminated next due to only having 2 points in the head-to-head. Finally, team *b* is eliminated since it loses head-to-head. This scenario is modeled by adding the following constraints to the 4-way tie core IP.

$p_{a,b} \geq 3$ (27a)
$p_{b,c} \geq 3$ (27b)
$p_{c,d} \geq 3$ (27c)
$p_{d,a} \geq 3$ (27d)
$p_{a,b} - p_{b,c} = 0$ (27e)
$p_{b,c} - p_{c,d} = 0$ (27f)
$p_{c,d} - p_{d,a} = 0$ (27g)
$p_{a,c} - p_{b,d} = 0$ (27h)
$p_{b,c} - p_{a,d} = 0$ (27h)
$\Sigma_{k \epsilon K} (TV_k \, r_{i,d,k}) \geq \Sigma_{k \epsilon K} (TV_k \, r_{d,i,k}) + 1$ for all $i \,\epsilon\, D_a \backslash \{d\}$ (28)

The first 9 constraints force Case 1 according to table 3.2.2-2. Constraints set (23) forces team *d* to be worse than every other divisional team on tie breakers and thus, it is eliminated first. Team *a* is then the division champion since *d* is eliminated first.

Case 2 shows a scenario where each divisional team beat team *d* once, and dominated one other team in head-to-head. For this case, if team *d* is eliminated first, the remaining teams are all still tied in head-to-head with 4 points. Next, team *c* must be eliminated by a tiebreaker further down the line. Once team *c* is eliminated, team *a* dominates team *b* in head-to-head and is declared the champion. For Case 2, the following constraints are added to the 4-way Tie Core IP:

$p_{a,b} \geq 3$ (29a)

$p_{b,c} \geq 3$ (29b)

$p_{c,a} \geq 3$ (29c)

$p_{a,b} - p_{b,c} = 0$ (30a)

$p_{b,c} - p_{c,a} = 0$ (30b)
$p_{i,d} = 2$ for all $i \,\epsilon\, D_a$ and $i \neq d$ (31)
$\Sigma_{k \epsilon K} (TV_k \, r_{c,d,k}) \geq \Sigma_{k \epsilon K} (TV_k \, r_{d,c,k}) + 1$ (32a)
$\Sigma_{k \epsilon K} (TV_k \, r_{a,c,k}) \geq \Sigma_{k \epsilon K} (TV_k \, r_{c,a,k}) + 1$ (32b)
$\Sigma_{k \epsilon K} (TV_k \, r_{b,c,k}) \geq \Sigma_{k \epsilon K} (TV_k \, r_{c,b,k}) + 1$ (32c)

The first 6 constraints help force Case 2 as shown in Table 3.2.2-2. Constraints 32a, 32b, and 32c ensure that team *d* is eliminated first, and that team *c* is eliminated second so that team *a* will dominate team *b* in head-to-head tiebreaker once they are the only two teams left. This happens because team *c* dominates team *d* and team *a* and *b* both dominate team *c*, so they also will dominate team d. Thus, team a is the division champion since team d was eliminated first and team c was eliminated second.

Similar to Case 2, Case 3 shows that each team beat d once, and dominated one other team in head-to-head. For this case, team c is eliminated first. From here, team b would be eliminated next due to having the worst head-to-head record. Team a and team d would be left with similar head-to-head records, so team a would need to dominate team d in some other eventual tiebreaker to win the division. To create Case 3, the following constraints were added to the 4-way Tie Core IP:

$p_{a,b} \geq 3$ (33a)

$p_{b,c} \geq 3$ (33b)

$p_{c,a} \geq 3$ (33c)

$p_{a,b} - p_{b,c} = 0$ (34a)

$p_{b,c} - p_{c,a} = 0$ (34b)

$p_{i,d} = 2$ for all $i \in D_a$ and $i \neq d$ (35)

$\Sigma_{k \in K} (TV_k\, r_{a,c,k}) \geq \Sigma_{k \in K} (TV_k\, r_{c,a,k}) + 1$ (36a)

$\Sigma_{k \in K} (TV_k\, r_{b,c,k}) \geq \Sigma_{k \in K} (TV_k\, r_{c,b,k}) + 1$ (36b)

$\Sigma_{k \in K} (TV_k\, r_{d,c,k}) \geq \Sigma_{k \in K} (TV_k\, r_{c,d,k}) + 1$ (36c)

$\Sigma_{k \in K} (TV_k\, r_{a,d,k}) \geq \Sigma_{k \in K} (TV_k\, r_{d,a,k}) + 1$ (36d)

The first 6 constraints help force Case 3 as shown in table 3.2.2-2. Constraints 36a, 36b, 36c, and 36d ensure that team c is eliminated first, and that team b is eliminated second. Once team b is eliminated, team a will need to dominate team d so that team a is the only team left. This case differs from Case 2 in how the case plays out since team c is eliminated first, not team d. However, team a is still the division champion in this case since team c is eliminated first.

Case 4 differs from the previous two cases in that team a is now the team that has split with every other team. For this case, any team b, c, or d can be eliminated first. Once one of them has been eliminated, one of those remaining two will be eliminated by having the worst head-to-head record with only 2 points. Finally, the remaining team and team *a* will have identical head-to-head records, and team *a* will need to dominate that team in some other eventual tiebreaker to win the division. For this case, the following constraints were added to the 4-Way Tie Core IP:

$$p_{b,c} \geq 3 \tag{37a}$$

$$p_{c,d} \geq 3 \tag{37b}$$

$$p_{d,b} \geq 3 \tag{37c}$$

$$p_{b,c} - p_{c,d} = 0 \tag{38a}$$

$$p_{c,d} - p_{d,b} = 0 \tag{38b}$$

$$p_{i,a} = 2 \text{ for all } i \in D_a \text{ and } i \neq a \tag{39}$$

$$\Sigma_{k \epsilon K} (TV_k \, r_{a,d,k}) \geq \Sigma_{k \epsilon K} (TV_k \, r_{d,a,k}) + 1 \tag{40a}$$

$$\Sigma_{k \epsilon K} (TV_k \, r_{b,d,k}) \geq \Sigma_{k \epsilon K} (TV_k \, r_{d,b,k}) + 1 \tag{40b}$$

$$\Sigma_{k \epsilon K} (TV_k \, r_{c,d,k}) \geq \Sigma_{k \epsilon K} (TV_k \, r_{d,c,k}) + 1 \tag{40c}$$

$$\Sigma_{k \epsilon K} (TV_k \, r_{a,b,k}) \geq \Sigma_{k \epsilon K} (TV_k \, r_{b,a,k}) + 1 \tag{40d}$$

The first 6 constraints help force Case 4 as shown in table 3.2.2-2. Constraints 40a, 40b, 40c, and 40d ensure that team *d* is eliminated first, and that team *c* is eliminated second. Once team *c* is eliminated, team *a* will need to dominate team *b* so that team *a* is the only team left. Once team *a* dominates team *b* in tiebreakers, team *a* is then the division champion in this case, since team *d* is eliminated first.

Due to permutations, this section has described 31 different IPs that are solved to find solutions to the NFL True Fan problem. The next section describes the SPORT algorithm that manages these 31 IPs to identify answers.

3.3 The SPORT Algorithm

This section provides the pseudo code for the Shortest Path Of Remaining Teams (SPORT) Algorithm, which solves the NFL True Fan Problem. The SPORT algorithm is a new algorithm that solves each of the 31 integer programs and, using the resultant MPDC and MPP, determines the critical games for the upcoming week for team *p*. Based upon the solution of the IPs, SPORT adjusts and reports solutions MPDC and MPP. Part of the reporting is a list of critical games for each of the 32 NFL teams, which is a solution to the NFL True Fan Problem.

The SPORT Algorithm requires many inputs which allow it to run. One input is the week, *w*, in question. Another input is the result of all completed games. In putting the result of a game, the score is irrelevant and only a win, loss or tie must be recorded. Finally, SPORT requires the schedule for all teams.

For each team, SPORT seeks to determine the critical games. As each team becomes the desired team, the core component of SPORT solves numerous IPs. These models are loaded based on a counter, *count1*, with the initial model being loaded each time a new desired team is chosen. After 31 IPs are solved, SPORT identifies the MPDC for the desired team and identities which IPs provided an MPP.

For each unplayed game in week *w*, SPORT assigns one team as the winner and checks MPDC. This check is made by verifying whether or not MPDC has increased in all models that had an MPP. This process is repeated with the other team winning. A final case is tested where the teams tie. Thus, for any given week, assuming no byes, SPORT solves up to 1,519 IPs for a single team in a single week. Thus, SPORT could solve over 48,000 IPs per week.

Once the SPORT Algorithm has run to completion many results are reported. The most important result is the critical games for the upcoming week *w*. For this week and for each team *a*, the critical games are reported as which teams *i* need to beat which teams *j*. Additionally, the specific IP model scenario that forces the critical game is reported, which can be a two-way, three-way, or four-way tie scenario. Also, if no solution is found for a team *a*, then the algorithm reports "No Solution" for that team, as that team is eliminated from divisional championship contention.

Another helpful output that goes along with the IP model scenario is the solution that the IP model gives, the minimum points that team *a* needs to win the division. The current points for team *a* are also calculated during the SPORT Algorithm's run and are reported with the minimum points. Finally, an MPP is reported so that fans have a clear path to a division championship.

Formally, the pseudo code for SPORT is as follows.

The SPORT Algorithm

For each $a \epsilon T$

    $Z^{min} \leftarrow M$
    For each IP *model* from Chapter 3
        Solve IP *model*,
        if $Z^{MODEL} < Z^{min}$, then
            unflag all *models*
            flag this *model*
            $Z^{min} \leftarrow Z^{MODEL}$
            Solution ← *model*'s solution
        else
            if $Z^{MODEL} = Z^{min}$
                flag *model*
            End if
        End Else
    End if
    If Z^{min} < M, then
        Report Z^{min} as the MPDC and *Solution* as the MPP
        For each to be played game $g = (t_i, t_j, w, -1)$ in the upcoming week

test ← 0

For each flagged IP *model*

Create *model'* by adding the constraint $q_{i,j,w} = 2$ to *model*

Solve *model'*

If *model'* is feasible and $Z^{model'} = Z^{min}$, then

team *i* beating team *j* is not critical for team *p*.

test ← 1

exit for loop

End If

End for

If *test* = 0, then

report team *i* beating team *j* in slot *k* is critical for team *a*.

test ← 0

For each flagged IP *model*

Create *model''* by adding the constraint $q_{j,i,w} = 2$ to *model*

Solve *model''*

If *model''* is feasible and $Z^{model''} = Z^{min}$, then

team *j* beating team *i* is not critical for team *a*

test ← 1

exit for loop

End If

End for

If *test* = 0, then

report team *j* beating team *i* in slot *k* is critical for team *a*

test ← 0

For each flagged IP *model*

Create *model'''* by adding the constraint $q_{i,j,w} = 1$ to *model*

Solve *model'''*

If *model'''* is feasible and $Z^{model'''} = Z^{min}$, then

team *i* tying team *j* is not critical for team *a*

test ← 1

```
                        exit for loop
                    End If
                End for
                If test = 0, then
                    report team i tying with team j in slot k is critical for team a
            End For
        End If
    End For
End For
```

Although SPORT theoretically solves the NFL True Fan problem, there is no guarantee that SPORT is a computationally tractable algorithm. SPORT solves many IPs and all it takes is one truly complex IP for SPORT to report neither an MPDC, nor MPP, nor critical games. The next chapter provides computational results, which demonstrate that SPORT is indeed a practical method to solve the NFL True Fan problem in most instances.

Chapter 4 - Computational Results

This chapter gives the computational results obtained from running the SPORT algorithm for the 2012-2013 NFL season. These results were obtained from running OPL through CPLEX 12.4 (CPLEX) on a Dell XPS with a 3.40 GHz Intel Core i7-3770 processors with 8GB of RAM. The results obtained include run time, number of critical games for each week, and number of teams eliminated for each week. Finally, this chapter gives several interesting instances to show the use of MPDC and MPP.

4.1 Instances Run

For the 2012-2013 season, instances were run for weeks 2 through week 17. Week 1 was not included as every team has the exact same solution at that point. The solution is 3 wins, one over every team in the division, and one tie over some other team. The results do not start to change until week 2, so week 2 is where the instances began. For all instances ran, the results are summarized in Table 4.1-1. Weeks 2 through week 5 had no significant results in terms of possible three-way tie and four-way tie scenarios and teams eliminated from the division championship. Most tie scenarios either are not the easiest possible path this early or those scenarios are eliminated by the results of the games from those weeks.

There are no team eliminations early, because the earliest a team can be eliminated from the division championship is after week 8. This is due to the fact that one team can have 8 wins and 0 losses, and another team can have 0 wins and 8 losses. Both teams can finish with 8 wins and 8 losses, and once tiebreakers are factored in, either team could lose out on the division championship. Thus, week 8 is the earliest one team could be eliminated from the division championship.

Week 6 is the first week with interesting results. In this week, the Baltimore Ravens, have 10 models all with the minimum path possible. One way is the dominant model, one is a three-way tie scenario, one is a four-way tie, case 1 scenario, two are four-way tie case 2 scenarios, two are four-way tie case 3 scenarios, and three are four-way tie case 4 scenarios. This is by far the most models any team has throughout the season. Four of these ten models have the same critical game as well; however, all of these models drop out for week 7.

Week 7 through week 10 again have few results of significance. These results were the critical games for the weeks, with no teams being eliminated from the division championship race. Week 9 has the Baltimore Ravens, with a three-way tie scenario possible, which is interesting since this scenario is still possible this late in the season. This is also the last week where any scenario other than the dominant model scenario is possible for any team.

Critical games appear when a team is near elimination. In many scenarios, the team must win and have the current division leader lose to retain a glimmer of hope for the championship. In these scenarios, such games for team *a* are critical. However, team *a* may be behind by five games with six to play. Most likely, there are no critical games. If team *a* loses and the leader wins, then team a is eliminated even though there was not a critical game.

Week 11 is the first week where critical games start carrying important implications for teams since week 11 is where teams are first eliminated from division championship contention if they do not win. Week 11 features three such games, with the Cleveland Browns, Jacksonville Jaguars, and Carolina Panthers all needing to win their week 11 game. None of them do, however, and they are all eliminated from their division championship race. These eliminations show up only when running the instances for week 12, so once teams are identified as being eliminated, it is straightforward to know that their game in the previous week was clearly a critical game.

Beginning with week 12 there is at least one team eliminated from their division championship race every week. Week 14 has the most teams eliminated in one week with eight teams being eliminated. Week 13 has nine critical games, and three teams being eliminated. In addition to week 14 having eight teams eliminated, it also has three critical games, only one of which involves a team potentially being eliminated from its division championship race.

Week 15 has one team eliminated from its division championship race, and has 10 critical games, which is the most for any week in the season. These 10 games are due in part to week 16 having five teams eliminated from their division championship race. That leaves five games that are critical to teams that still have a chance to win their division even if they lose. Three of these games involve teams who are currently in the lead in their division. These games are critical because if they win, it forces their minimum path to increase. So these games are basically games that the division leader can afford to not win and its division lead will not be affected. The other two games involved teams that were not facing elimination from their division championship, but also had almost no slack left in terms of how much larger their minimum path could be before being eliminated. These games are critical as the result keeps these teams in the hunt for their division championship with the minimum possible points.

Week 16, in addition to its 5 eliminated teams, also has seven critical games for its week. The final week before the end of the season, week 17, has three critical games, each of which involves a team that will lose out on its division championship if they lose in week 17.

The 2012-2013 NFL season features one team, the Cincinnati Bengals, who ties with their division champion, the Baltimore Ravens. However, the Bengals lost to the Ravens on division record so the Bengals did not win their division.

Table 4.1-1 2012-2013 Results Summary

Week	Critical Games	Eliminated Teams	Run Time (in minutes)
2	4	0	60
3	3	0	68
4	3	0	68
5	2	0	65
6	5	0	85
7	1	0	54
8	1	0	69
9	4	0	76
10	3	0	74
11	4	0	65
12	5	3	87
13	9	3	86
14	3	8	69
15	10	1	76
16	7	5	61
17	3	3	58
Sum	67	23	1121
Average	**4.19**	1.44	**70.06**
Average Over Last Half of Season	5.33	**2.56**	72.44

4.2 Comparison to the 2013-2014 NFL Season

For comparison, results for week 9 through week 17 of the 2013-2014 NFL season were computed. These results are summarized below in Table 4.2-1. Run times for this season were not included as the same integer programs and SPORT algorithm were run on the same computers as the 2012-2013 season and had similar times. This NFL season featured almost as many critical games in its final nine weeks as the entire previous season had total. It was interesting to note that for both seasons, week 12 was the first week that a team was eliminated from its division championship race. Both seasons had similar results for the average number of teams eliminated over the last half of the season at about 2.5 teams.

The 2013-2014 season featured three weeks with critical game totals that totaled at least as much as the highest total in the 2012-2013 season. Week 10 of the 2013-2014 season was a very interesting week for two reasons. First, this week featured two critical games for teams that were not playing in those games. Further, these games were between division opponents, so the teams needing these critical games would have had to root for a division rival to win. The second interesting result was that the Tennessee Titans had two MPDC paths with a 4-way tie, case 4 scenario. For being 10 weeks into the season with two of these two paths still possible is rare as it was the only team across two seasons to have these paths occur.

Table 4.2-1 2013-2014 Results Summary

Week	Critical Games	Teams Eliminated
9	5	0
10	7	0
11	4	0
12	2	2
13	12	1
14	11	7
15	6	6
16	10	0
17	3	4
Sum	60	20
Average	6.67	2.22

In terms of flex scheduling, the critical games found for the 2012-2013 season were compared to what the NFL determined to be the best choice for their flex game. Two games were flexed into the Sunday Night slot, the 49ers at Seahawks in week 16 and the Cowboys at Redskins in Week 17.   Week 16 featured seven critical games, one of which was the 49ers-Seahawks game the NFL chose. Week 17 featured three critical games, one of which was the

Cowboys at Redskins game the NFL chose. Out of the seven weeks of flex scheduling possible (week 11 through week 17), the NFL only flexed two of those games. The other five featured games that were not a critical game for that week, despite each week having at least one critical game. The NFL could be considered fortunate since week 12 and week 15 happened to have critical games already scheduled, but that leaves three weeks (weeks 11, 13 and 14) in the last half of the season that could have featured critical games, but did not. Had the NFL utilized the SPORT algorithm on its schedule for those weeks, they could have noted the potential interest in those games and considered using its flex scheduling on those games.

Chapter 5 - Conclusions and Future Work

In the NFL, each team's desire is to win their division and qualify for the playoffs. Given the complex nature of the NFL's tiebreaker procedure for playoff qualification, there are numerous ways that a team can qualify for the playoffs. Ideally, a team would like to qualify for the playoffs by taking the easiest path to the playoffs. This easiest path is given as winning the least number of games possible. By using the NFL's tiebreaker procedure to find each team's easiest path, many games throughout the season arise that can potentially affect a team's easiest path. These games are identified as critical games as those games can possibly change the easiest path of a team, making the easiest path more difficult by requiring additional wins or ties later in the season. Identifying these games is called the NFL True Fan Problem and is formally defined in the introduction.

The task of identifying the critical games and the easiest path of each team was accomplished using integer programming. In total, 31 integer programs were used to encompass every two-team, three-team, and four-team division tie scenario. Once the easiest path was found from these scenarios, a new algorithm, Shortest Path of Remaining Teams (SPORT), is used to find the critical games for the upcoming week.

By solving the integer programs and implementing the SPORT algorithm, results for two NFL seasons were obtained. Both seasons showed that week 12 was the earliest a team was eliminated. The average number of critical games for both seasons was 65 games per season. Finally, the 2012-2013 season's critical games were compared to the NFL's flex schedule and found that three weeks could have featured a critical game in the flex slot, but the NFL missed out on its opportunity to feature the critical games in those slots.

5.1 Future Research

The primary extension of this research should consider the critical games for the NFL wildcard teams. A wildcard team is a team that qualifies for the playoffs, but does not win their division championship. These are the two teams in each conference with the best record after all the division champions have been decided. A portion of this research attempted this problem, but the dominant IP did not solve within a few days. This lack of a solution is believed to be due to the number of teams (12) that could tie and the complexity of having the team be the second or third worst team in the division. When coupled with all of the numerous tie scenarios, the number of IPs that would need to be solved would be excessively numerous. A new approach would have to be created to solve this interesting and important problem.

For the NFL, there is also more work to be done in the MPDC and MPP area. The integer programs presented do not perfectly represent the strength of victory since proper weight cannot be applied when an uneven number of unique opponents is applied. For example, one team beats seven teams while beating one of those seven teams twice (6 unique opponents, 7 victories), and another team beats seven unique opponents (7 unique opponents, 7 victories). Currently there is no way to illustrate the fact that the total points from beating those opponents should be averaged to create an accurate portrayal of strength of victory and still have a linear model.

Applying the SPORT Algorithm to other sports may create similar benefits for those sports. For the NBA, there are certain TV nights where major networks carry NBA games. Knowing which games are more important coming up would aid in creating a better TV schedule in terms of getting more important and exciting matchups on TV. The same idea can be applied to the MLB. Usually a single game in a series is picked up for a primetime game on TV. Being able to look forward into the schedule and determining which series' and which games in a series are most impactful and important would help put a better TV baseball schedule together. With the rise of Major League Soccer (MLS) in the United States, the MLS could look at this as a way to place marquee matchups on TV and attract a wider audience with an underlying rooting interest in the outcome.

References

"Appendix I: The Nature of Operations Research." *Operations Research* 19.5, Guidelines for the Practice of Operations Research (1971): 1138-48. Print.

"Ancient Egyptian Sports." State Information Services. www.sis.gov.eg/En/Templates/Articles/tmpArticles.aspx?ArtID=1694#.UwRYLGJdWSo

Burke, Martin J., Esq. "Lies, Darn Lies, and Statistics." *401 (k) Advisor* 20.4 (2013): 12. Print.

Cervone, Dan, Alexander D'Amour, Luke Bornn, and Kirk Goldsberry. "POINTWISE: Predicting Points and Valuing Decisions in Real Time with NBA Optical Tracking Data."

Cheng, Eddie, and D. Steffy. "Clinching and Elimination of Playoff Berth in the NHL." *International Journal of Operations Research* 5.187-192 (2008): 17.

Coleman, B. J. "Identifying the Players in Sports Analytics Research." *Interfaces* 42.2 (2012): 109-18. Print.

CPLEX Solver, IBM's Software Information Page, http://www-01.ibm.com/software/info/ilog/

Delle Donne, Diego, and Javier Marenco. "Studying Playoff Qualification in Motorsports via Mixed-Integer Programming Techniques." *Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology* 226.1 (2012): 32-41.

Duran, Guillermo, Mario Guajardo, and Rodrigo Wolf-Yadlin. "Operations Research Techniques for Scheduling Chile's Second Division Soccer League." *Interfaces* 42.3 (2012): 273-85. Print.

Easton, Kelly, George Nemhauser, and Michael Trick. "Solving the travelling tournament problem: a combined integer programming and constraint programming approach." Practice and Theory of Automated Timetabling IV. Springer Berlin Heidelberg, 2003. 100-109.

ESPN NFL Scorecenter. www.scores.espn.com/nfl/scoreboard

Fry, Michael and Jeffrey Ohlmann. "Introduction to the Special Issue on Analytics in Sports, Part II: Sports Scheduling Applications." *Interfaces* 42.3 (2012): 229-31. Print.

Gusfield, Dan and Chip Martel. "Thresholds for sports elimination numbers: Algorithms and complexity." *Algorithms and Data Structures*. Springer Berlin Heidelberg, 1999. 330-341.

Hamiez, Jean-Philippe and Jin-Kao Hao. "Solving the sports league scheduling problem with tabu search." *Local Search for Planning and Scheduling*. Springer Berlin Heidelberg, 2001. 24-36.

Hurley, W. J. "Special Issue on Operations Research in Sport." *Computers & Operations Research* 33.7 (2006): 1871-3. Print.

Jenkins, Mike J. "Advanced Materials and Sporting Performance." *Interdisciplinary Science Reviews* 27.1 (2002): 61-6. Print.

Kauppi, Katri, Claire Moxham, and David Bamford. "Should we Try Out for the Major Leagues? A Call for Research in Sport Operations Management." *International Journal of Operations & Production Management* 33.10 (2013): 1368-99. Print.

Klis, Mike. "NFL Fact Book: Denver is 17th Largest TV Market; Cleveland is 18th." The Denver Post First-And-Orange Blog. http://blogs.denverpost.com/broncos/2013/08/01/nfl-fact-book-denver-is-17th-largest-tv-market-cleveland-is-18th/20473/

Kolekar, Maheshkumar H. "Bayesian Belief Network Based Broadcast Sports Video Indexing." *Multimedia Tools and Applications* 54.1 (2011): 27-54. Print.

Lewis, Michael. *Moneyball: The Art of Winning an Unfair Game.* New York: W.W. Norton, 2003.

London 2012. The Official Website of the Olympic Movement. www.olympic.org/london-2012-summer-olympics

Maidenberg, Micah. "Tracking their Every Move." *Crain's Chicago Business* Jan 20, 2014 2014: 3. Print.

Martinez, J A Martinez, and Laura Martinez. "A Stakeholder Assessment of Basketball Player Evaluation Metrics." *Journal of Human Sport and Exercise* 6.1 (2011): 153-83. Print.

"Math Modelers Take on Sports Figures." *IIE Solutions* 31.6 (1999): 8. Print.

McClure, Brian, Richard Cassady, Chase Rainwater, and Justin Chimka. "Optimizing the Sunday Singles Lineup for a Ryder Cup Captain." *Interfaces* 42.2 (2012): 180-90. Print.

McMurray, Scott. "Basketball's New High-Tech Guru." *U.S. News & World Report* Dec 11, 1995 1995: 79-80. Print.

Nemhauser, George Nemhauser,Michael Trick. "Scheduling A Major College Basketball Conference." *Operations Research* 46.1 (1998): 1-8. Print.

NFL Schedules. www.nfl.com/schedules

"Nielsen Tops of 2012: Television." Nielsen Newswire. http://www.nielsen.com/us/en/newswire/2012/nielsen-tops-of-2012-television.html

"Professional Sports Average Salary/Revenue/Salary Cap." Statistic Brain. www.statisticbrain.com/professional-sports-average-salary-revenue-salary-cap/

Régin, Jean-Charles. "Minimization of the number of breaks in sports scheduling problems using constraint programming." *DIMACS Series in Discrete Mathematics and Theoretical Computer Science* 57 (2001): 115-130.

Russell, Robert, and Timothy Urban. "A Constraint Programming Approach to the Multiple-Venue, Sport-Scheduling Problem." *Computers & Operations Research* 33.7 (2006): 1895-906. Print.

Russell, Tyrel, and Peter van Beek. "A Hybrid Constraint Programming and Enumeration Approach for Solving NHL Playoff Qualification and Elimination Problems." *European Journal of Operational Research* 218.3 (2012): 819-28. Print.

Sathyan, Thuraiappah, Richard Shuttleworth, Mark Hedley, and Keith Davids. "Validity and Reliability of a Radio Positioning System for Tracking Athletes in Indoor and Outdoor Team Sports." *Behavior Research Methods (Online)* 44.4 (2012): 1108-14. Print.

Schaerf, Andrea. "Scheduling Sport Tournaments using Constraint Logic Programming." *Constraints* 4.1 (1999): 43-65. Print.

Schreuder, Jan A. M. "Combinatorial Aspects of Construction of Competition Dutch Professional Football Leagues." *Discrete Applied Mathematics* 35.3 (1992): 301-12. Print.

Schwartz, B. L. "Possible Winners in Partially Completed Tournaments." *SIAM Review* 8.3 (1966): 302-8. Print.

Smith, Kelly "Valuing Professional Sports Franchises: An Econometric Approach." About.com Economics. http://economics.about.com/library/weekly/aa043004g.htm

Sport Vu Player Tracking. Stats.com. http://www.stats.com/sportvu/sportvu.asp

Streib, Noah, Stepen Young, and Joel Sokol. "A Major League Baseball Team Uses Operations Research to Improve Draft Preparation." *Interfaces* 42.2 (2012): 119-30. Print.

Wayne, Kevin D. "A new property and a faster algorithm for baseball elimination." *SIAM Journal on Discrete Mathematics* 14.2 (2001): 223-229.

"What Is the Furthest Down the Line a Tie Breaker Has Ever Gone." Professional Football Researchers Association. http://www.pfraforum.org/index.php?showtopic=1212