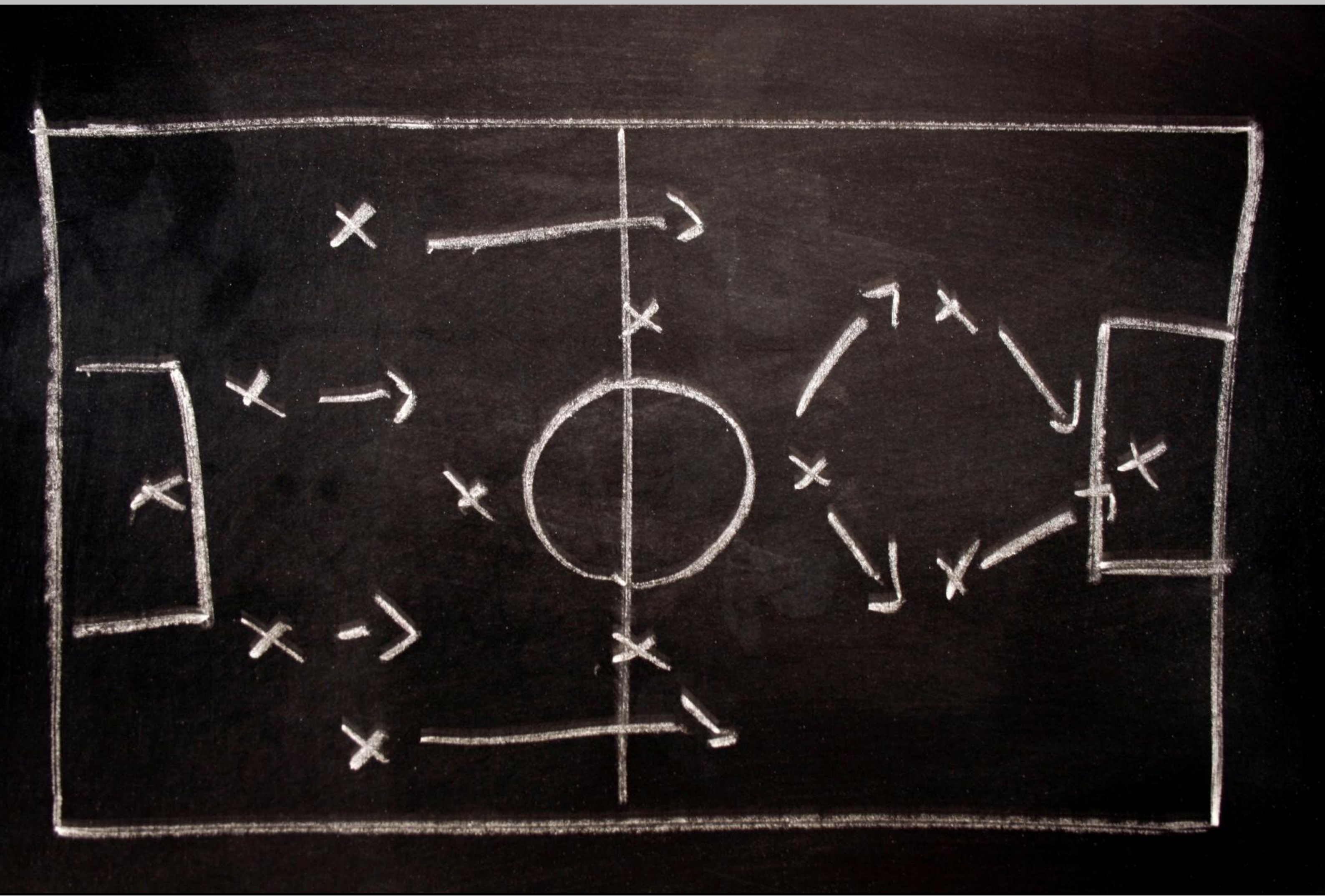
MASTER THESIS ECONOMETRICS & MANAGEMENT SCIENCE
OPERATIONS RESEARCH & QUANTITATIVE LOGISTICS

OPTIMIZING THE LINE-UP

IN SOCCER WITH REAL-LIFE DATA



Author:	Bayram KAVI [294688]
Academic supervisor:	Dr. Dennis HUISMAN
External supervisor:	Drs. Niels Calis
Co-reader:	Dr. Michel van de VELDEN

ERASMUS UNIVERSITY ROTTERDAM

Erasmus University Rotterdam
Erasmus School of Economics
Econometric Institute

1 August 2012

PREFACE

In the summer of 2011 I found an interesting thesis topic, optimization in soccer. First I doubted if this topic was suitable for my master Operations Research & Quantitative Logistics. After discussing this with the coordinator of the thesis, the topic was approved. Writing a thesis based on soccer was the ideal topic for me. This is because I am a huge fan of soccer, applying my knowledge of Econometrics on soccer was the best possible way to finish my study.

First I would like to thank my parents for supporting me and showing a lot of patience during the long nights I worked on my thesis. Especially in the final months I worked during the nights while they were sleeping.

Furthermore, I would like to thank my academic supervisor Dr. Dennis Huisman, external supervisor Drs. Niels Calis and former external supervisor Drs. Stefan Schuurman for providing me with interesting ideas, sharing their knowledge and continuous support during the writing of my thesis.

I would also like to thank my colleagues at Ortec TSS for sharing their knowledge of the database I used in my thesis. Finally I would like to thank Samad Bashir for taking time to read my thesis and correct me.

Bayram Kavi

ABSTRACT

This thesis treats the optimization of the line-up in soccer with real life data. The real life data used in this thesis consists of 40 matches played by Ajax in 2011-2012. The data are provided by Ortec TSS. With the data we created three profiles that will be matched to determine the optimal line-up for Ajax. A player profile containing the grades for each skill, position profile containing the grades of the skills required to play at a position and a team profile which adds additional sets of constraints to the Integer Programming model based on the team statistics. The team profile in particular is unique in the optimization of the line-up. These three profiles are matched with an IP model to optimize the line-up. The results obtained with the model are almost equal to the preferred line-up of the coach in 2011-2012. The only disadvantage of the methods used is that some players are overvalued or undervalued at some positions. This can be caused by the fact that the data we possess does not contain the physical skills of the players, only technical skills are included.

keywords: optimization, soccer, profiles, grading skills, matching problem.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

Soccer or football are the most commonly used names for association football. To distinguish the game from several other sports played with the foot in the 18th century the name association football was coined. The modern rules we are familiar with nowadays dates back on the mid- 19th century. The rules are determined in order to standardize the widely varying forms of football played at the public schools of England. Soccer can be described as a sport that is played between two teams of eleven players with a ball. It is played on a rectangular field of grass, with a goal in the middle of each of the short ends. The purpose of the game is to score a goal by driving the ball into the opposing goal. One of the eleven players on the field is called the goalkeeper, the only player that is allowed to touch the ball with their hands or arms. The other 10 players use their feet to kick the ball. The team that scores the most goals at the end of the match wins. The game is played in total 90 minutes; after 45 minutes there is a break of 15 minutes. It is also possible that the game ends with a tie and depending on the format of the competition it is possible that a draw is declared or the game goes into extra time. If after extra time the score is still equal a penalty shootout will determine the winner.

Nowadays, soccer is played at a professional level all over the world. It is the world's most popular sport which is played by over 250 million players in over 200 countries. A lot of people go regularly to football stadiums to follow their favourite team or just to enjoy the game. Football has also the highest global television audience in sports; billions watch the game on television or on the internet. This happens especially during international tournaments when national teams are playing like the FIFA world cup.

Because of the global popularity of soccer, huge sums of money are involved in soccer, from player wages to transfer fees. Soccer clubs are companies; they invest, have employees and even own properties. The players and the technical staff are a small but important part of the employees for a soccer club. One way to increase profits significantly is to win competitions. To win a competition the players have to perform better than their opponents. Besides the individual performance of players, teamwork is an important aspect in soccer. The performance of a team is dependent of the line-up of the team, which consists of 11 team members with different skills on each position. The fact that a manager is allowed to change only 3 players of the 7 available substitutions during a match makes the decision of the line-up even more crucial.

Nowadays computers are used to support people with various decisions. In soccer computers are generally used to keep up statistics during matches. These statistics could be used to support the technical staff. This is the part where management science is needed. One way to support the technical staff is determining the optimal line-up. This can be seen as matching the supply of players with the demand on the positions of the team. In soccer these are the starting eleven which must be decided out of all available players in a team. In general a team consists of 25 main players and in case of a lot of suspension or injuries, players from the youth squad could be added to complete the starting eleven.

The goal of this thesis is to find the optimal line-up for a soccer team; the questions stated below have to be answered to find the optimal line-up.

- *What kind of historical data is available?*
- *How do we relate the qualities of a player with the data?*
- *Which qualities are needed on a specific position?*
- *How do we describe teamwork with data?*
- *How do we determine the optimal line-up?*

First we should take a look how the technical staff determines the line-up in general. Then a description of ORTEC TSS will be given, the company that provides us the data. This chapter will end with the thesis outline.

1.1 SELECTION OF LINE-UP

The head of the technical staff in soccer is called the coach of the team. The coach is the one responsible for the team and the one that will be judged on the performance of the team. One of the tasks of the coach is determining the line-up before a match. The decision of the line-up depends on the formation, tactics and the available players. A formation is the composition of the positions of the eleven players. Because every formation needs one goalkeeper, there are 10 variable positions. Widely used formations are 4-4-2 and 4-3-3; the first number corresponds to the number of defenders, the second to the number of midfielders and the last to the number of attackers. Tactics of a team can be described as the game play of a team; this can be an attacking one or a defensive one for example. The coach also provides individual tactics to the players depending on their qualities. To decide the line-up, the coach needs a formation which he prefers. Depending on the strength of the team or the opponent the formation can change. This can happen before or during the match. The same strategy holds for the tactics of the team and players. It is mentioned before that the combination of the tactics and the formation is essential in team sports. Finally, the coach needs to judge the qualities of the players in order to assign them to positions where he thinks the player fits best. The players are judged in general during training sessions and during matches. During training sessions and friendly games it is usual that the coach experiments with different tactics, formations and the positions of players. This happens in general before the start of a season, if the technical staff is new to the club or new players have been purchased.

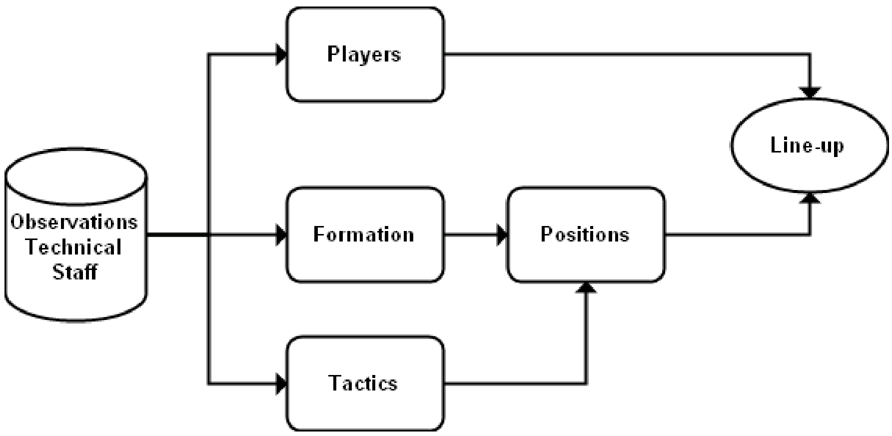


Figure 1.1 - Schematic view of the selection of the line-up by the coach

The procedure described above is the selection of the line-up in practice; it is what happens in reality so it is important for us to approach this procedure. In figure 1.1 a schematic view of this procedure is shown. This figure shows us what information we need to support the technical staff with the selection of the best line-up. From the figure we see that we need the qualities of the players and the qualities needed at the positions to decide the line-up. To determine the qualities at the position a formation and team/individual tactics are needed. These qualities are currently obtained from the observations of the technical staff during training sessions and matches.

1.2 ORTEC TSS

Ortec was established on April 1, 1981 in the Netherlands. In the past 30 years the ORTEC group has grown to one of the largest providers of advanced planning and optimization software solutions and consulting services. ORTEC Team Support Systems (ORTEC TSS) a part of the ORTEC group is responsible for developing decision support software for sports based on operational research methods.

The software used for soccer is called Effectivity in Action (EiA) and it is used by analysts to measure team and player effectivity. Since 2007 ORTEC TSS analyzes every match played in the Dutch Eredivisie and the Belgian Jupiler Pro League. Besides these two, several other competitions in Europe and international tournaments are analyzed as well. Thus many data are available at ORTEC TSS for a number of soccer competitions in Europe and the rest of the world. The information available at ORTEC TSS consists of the time, type, location and the grade of the actions performed by the players. An example of the action would be a pass of player X at location (x, y) in the 6^{th} minute with grade 4. The grade is based on the likert-scale which will be explained further in chapter 4.3. This information is currently used to create detailed match report, effectivity graphs that are shown in a Belgian newspaper and online for a Dutch newspaper which is called the live widget. The goal of ORTEC TSS is to make (team)sports more intelligible and more attractive for the athlete, club/trainer and staff, sponsor, media, scouts/agencies and the fans. This thesis will discuss the determination of the optimal line-up in soccer by using the data from EiA.

1.3 THESIS OUTLINE

The thesis is set up as follows. First a detailed problem definition will be given. In this chapter we will discuss what we need to build the decision support program and our main research question will be formulated. Literature related to our study will be presented in chapter 3. In chapter 4, we describe the dataset provided by ORTEC TSS. Chapter 5 will discuss and give a clear description of the methods used in the thesis. The results obtained by the methods will be shown in chapter 6. The last chapter contains the conclusion of this thesis with the strengths and weaknesses of the model and suggestions for further research on this topic.

2 PROBLEM DEFINITION

A problem every coach of a team faces is to decide the line-up. The line-up in soccer can be described as the eleven players starting the match on the field. In management science terms, this problem is matching the supply of players with the demand on the positions in the team. A more common description will be assigning the players to the right positions on the field where they can perform at their best. This thesis will treat the optimization of the line-up, the starting eleven before a match. The optimization of the squad; which is the starting eleven and the 7 substitutions is the next step of the coach after determining the line-up. With the optimization of the squad it is possible to anticipate on events during a match, this is out of our research's scope. In order to optimize the line-up we need to distinguish the players skills and the quality of those skills. To do this we first need to answer the questions stated below:

- *What is a skill?*
- *What is quality?*

In this chapter we will first give a clear description of the terms skill and quality. In the next part, we will describe how we can obtain the required skills from the given data. This is an important part of the thesis because it will serve as input for the optimization procedure. The remainder of the chapter will describe how the procedure of the selection of the line-up shown in figure 1.1 can be approached.

2.1 WHAT IS SKILL AND QUALITY?

The title of this thesis is 'Optimizing the line-up in soccer with real-life data'. The first question that arises after reading the title is 'What is optimal in sense of soccer?'. The optimal line-up in this thesis will be the best possible line-up with respect to the soccer skills of the players. In order to optimize we need to know the quality of a soccer skill. In a game like soccer there is not only one soccer skill, but many skills that are needed. Depending on the position of the player and tactics given by the coach, several different combination are needed at different positions on the field. This is why we need a clear description of the terms skill and quality.

2.1.1 SKILL

In Collins [2006] skill is defined as '*a special ability in sports, the ability acquired by training*' and as '*a trade or technique, requiring special training or manual proficiency*'. The classical definition of skill is '*the learned ability to bring about pre-determined results with maximum certainty often with the minimum outlay of time or energy or both*', Knapp [1977].

These three definitions of skill have in common that a skill is an ability that can be trained, so it can change over time in positive or negative sense. The definition of Knapp speaks about pre-determined results which is not always the case in sports. This is because in a free-flowing team sport like soccer, the results of a pass for example depends on the reactions of the teammates and the opponent. The right action has to be performed at the right time. One good pass does not say anything about the skill of the player. We want to know the quality of the skill.

2.1.2 QUALITY

According to Collins [2006] quality can be defined as '*a distinguishing characteristic or attribute*', '*the basic character of nature of something*' and '*the degree or standard of excellence*'. We want to know the quality of a certain soccer skill like passing. To distinguish a skill between players it is also important that we know when a player is not suitable in performing a skill. Therefore we want to measure the quality in a number; grade it. An example would be a player that has mastered a skill and can get grade 10 out of 10.

2.1.3 QUALITY OF A SOCCER SKILL

With the definitions of quality and skill given in the previous parts we can now describe what the quality of a soccer skill is. First we will describe what a soccer skill is. A soccer skill can be described as a ability that is used in soccer which can be acquired by training. Not all abilities can be mastered by everyone just by training, some people might have a talent for some skills and are better in performing them. As Knapp [1997] cited: '*Soccer is a game that is categorized as a free-flowing game requiring the execution of many aspects of skill in a dynamic context*'. In other words, a player might have a good technique but if he does not perform the right action at the right time (skill) then he becomes an 'almost' useless player.

We will use historical data to measure the quality of soccer skill, not by just looking at one action performed in one game but by looking at each action separately over several matches. By doing this we will be able to distinguish the quality of soccer skill for any player.

2.2 OBTAINING RELEVANT DATA

The next step in our analysis is to obtain the necessary information from the data. In sports like soccer, data is stored per match which cannot be used directly for our purpose. This is because the performance of a player in one match is not informative enough to reflect the players' skills. In figure 1.1 we see that with the observations of the technical staff the skills of the players, formation and tactics are decided. To support the technical staff this information must be obtained from the database. First the actions stored in the database must be translated into soccer skills for players. This will be called the player data we need; it must contain all skills in order to judge the capabilities, which can be used later used to match it with a position. Our first sub-question will be:

- *How do we translate the actions in the database into soccer skills of players?*

The tactics and formation can also be obtained from the database. The formation will not change a lot, so it can be assumed fixed for a specific team. As mentioned before in soccer we have individual and team tactics. So we need team data which describes the tactics of a team and tactics given to the players on specific positions. The individual tactics obtained from the team data are needed to determine which qualities are necessary to play at a given position. The next sub-question we need to answer is:

- *How do we reveal descriptive statistics related to individual and team tactics with historical data per match?*

2.3 CREATING PROFILES

With the skills and the team data obtained from the database, profiles can be created. Figure 1.1 shows us that at least two different profiles are needed, one expressing the qualities of players skills and one expressing the minimum needs of the skill at different positions. The player and position profiles are essential, but in order to approach reality with the optimization a team component should be included. An additional profile is needed in our approach and this will be called the team profile, the profile expressing the team tactics.

A disadvantage of the database is that we only possess information of the matches and not the training sessions. This means that the profiles will be based only on the information obtained from the matches. Figure 2.1 shows a soccer field with all possible positions a player can be assigned to. With these positions any soccer formation can be created. The abbreviations are listed in table 2.1. If we exclude the goalkeeper in figure 2.1 we see that we have five lines on the field. The locations of the lines shown in the figure below are not the exact positions because the exact positions are dependent of the tactics. The defensive line could be closer to the midfield if an attacking tactic is used for example.

Abbreviations of the positions			
GK	Goalkeeper	L	Left
D	Defender	CL	Central-left
DM	Defensive Midfielder	C	Central
M	Midfielder	CR	Central-Right
AM	Attacking Midfielder	R	Right
F	Forward		

Table 2.1 - Abbreviations' of the positions given in figure 2.1

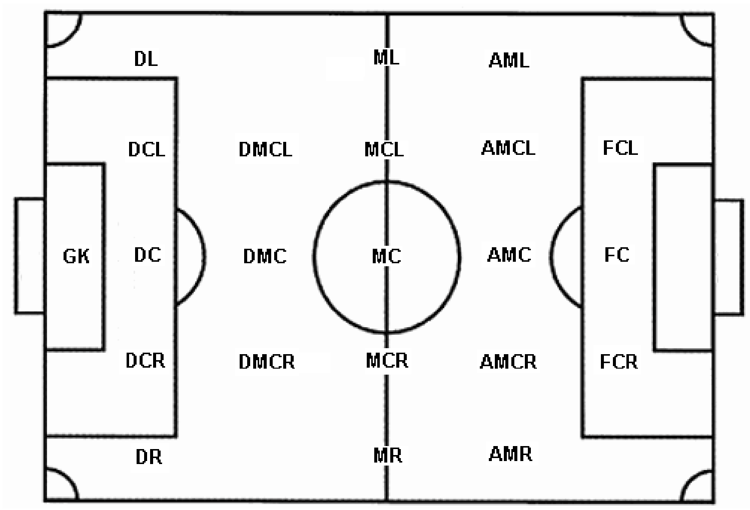


Figure 2.1 - Positions shown in soccer field

2.3.1 PLAYER PROFILE

From the player data containing the skills of each player, profiles can be created for each player. The profile must contain grades for all the skills of the player. Some examples of soccer skills are short-passing, long-passing, crossing and heading. All these skills should be graded separately in order to determine the quality of each skill. Besides a list of skills it is possible to also add the position to the profile where the player has played in many matches. The sub-question we obtain from this sub-section is:

- *How do we grade the skills of the players?*

2.3.2 POSITION PROFILE

The position profile is similar to the player profile; it is the same list with the soccer skills. The difference with the player profile is that the team data will be used to determine the minimum qualities of the skills needed to play at the position. A profile for a position also depends on the tactics (obtained from team data) and the formation (given). For this reason different profiles could be created for similar positions like MCL and MCR. One of the midfielders could have a defensive task of intercepting the opponent as much as possible and one can be the playmaker responsible of leading his teammates. This is why similar positions could have different tasks which are depending on the tactics and formation. Besides the grades the position profile should distinguish between the important skills needed at the position and the less important skills. This can be done by weighting the skills, the important skills will get a higher weight than less important skills. It is still possible that the data does not reflect the importance of some skills for some positions, this is because the data consists of actions and there is a lack of physical and mental skills. To create the position profiles the next sub-question must be answered:

- *How do we weight the skills for the positions in order to distinguish different tasks for the positions given the formation and tactics?*

2.3.3 TEAM PROFILE

Besides the player and position profiles, a team profile is needed to approach reality with the model. This is because the determination of the optimal line-up is not only dependent of the individual qualities of the players, teamwork is essential in sports. The team profile will express the cooperation between players depending on the tactics and the formation of the team. This will be done by examining the statistics that occur at the different lines separately. In soccer we can distinguish between the horizontal lines which are the defenders, midfielders , attackers and vertical lines left, right flank and centre. Checking the statistics at the different lines could provide us more information about the tactics in order to determine the teamwork. As mentioned before the formation is given and the team tactics are obtained from the team data. The sub-question of this section is defined as:

- *How do we expose and grade the quality needed for every line/flank based on the team tactics?*

2.4 MATCHING PROFILES

The optimal line-up of a soccer team could be obtained in several ways. A simulation model could be used to decide the optimal team. The database provided is not sufficient to use a simulation. This is because the database only contains actions of players. In a simulation we would need more than actions; mental and physical skills are necessary to do a realistic simulation of a soccer match.

Another way is to match the profiles described in the previous section. The team profile will give the necessary information depending on the tactics for a given formation. The position profiles contain information of the essential skills needed to perform at a given position and finally the player profile provides information of the capabilities of a player. The position and player profiles can be matched to evaluate the performance of the players at a given position. The team profile will be used to judge the team; by adding the team profile the optimal line-up would not depend only the player and position profile.

A useful model to match the three different profiles is using a integer programming model to obtain the optimal line-up. With the optimal line-up we mean the starting eleven before a game. Our final sub-question is:

- *Which IP model do we use to match the profiles in order to optimize the line-up?*

2.5 GOAL

The goal of the thesis is to build a decision support program which will be able to optimize the line-up of a soccer team with real-life data as input upon request by ORTEC TSS. A large database from the system EiA is available at ORTEC TSS and they wish to use the database for more purposes. The requirements of ORTEC TSS are that the program is able to use the EiA as input and has a team-component. This was not possible at their previous program named Coach & Scout Assistant. Our idea to fulfil the requirements of ORTEC TSS is to use the database to obtain relevant data in order to create profiles. These profiles must be able to reflect qualities of players and the tactics of a team. Finally the profiles must be matched to optimize the line-up of a team. To reach our goal the sub-questions defined in the previous sections must be answered. The sub-questions are summarized below.

- *How do we translate the actions in the database into soccer skills of players?*
- *How do we reveal descriptive statistics related to individual and team tactics with historical data per match?*
- *How do we grade the skills of the players?*
- *How do we weight the skills for the positions in order to distinguish different tasks for the positions given the formation and tactics?*
- *How do we expose and grade the quality needed for every line/flank based on the team tactics ?*
- *Which IP model do we use to match the profiles in order to optimize the line-up?*

We can split the procedure in the thesis into 3 parts. The first part is obtaining relevant data of the players and the team. The first two sub-questions are related to the first part. The second part of the thesis can be described as using the data obtained from the first part to create the player, position and team profile. The 3rd till 5th sub-questions are related to the profiles. The last part consists of matching the profiles to come with an optimal line-up; our final sub-question covers this part. Our main research question covering all three parts with the sub-question will be:

- *How do we optimize the line-up in soccer with a real-life database containing only actions during a match?*

The procedure described in this chapter is summarized below and in figure 2.2 a schematic view is given. The data of the players and teams will be obtained from the TSS database. With the team data we can define the team tactics and for a given formation the team profile can be created. To create the position profile we need the formation and the team tactics again. The player profile will be created using the data of the players. These three profiles will be the input for the Integer Programming model. A detailed description of the procedure depicted below can be found in chapter 5.

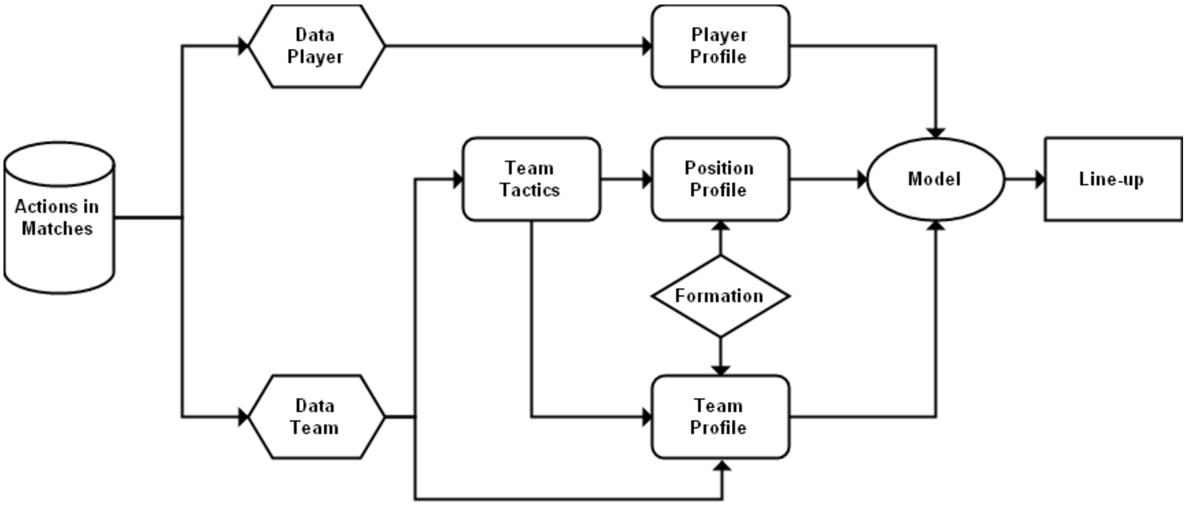


Figure 2.2 - Schematic view of procedure in thesis

3 LITERATURE REVIEW

In this chapter we will discuss articles related to our thesis. A lot of articles are written about soccer; but the majority of those articles are about betting and predicting match scores. The research done on the optimization of the line-up is limited. First we will discuss some interesting articles concerning team sports. Most of the articles will be about Coach & Scout Assistant or are based on it. The second part will cover similar studies in other areas.

3.1 SIMILAR STUDIES IN TEAM SPORTS

3.1.1 SOCCER

Comparable optimization results have been obtained by Boon & Sierksma (2003). This paper discusses a method to obtain the optimal team formation for soccer and volleyball. A basic Integer Programming Model is used to match quality supply and quality demand for soccer. In case of volleyball things are more complicated because a player does not have a fixed position. The input for the integer programming model consists of two quality lists, a position-score table and a player-score table. In the position-score table the relative importance of all qualities are determined for each position, and players are valued for each quality in the player-score table. The technical staff grades every player with a mark between 1 and 10 and grades also the quality list in the position-score table. In practice it could be that several position-score tables are needed to anticipate on changes between ‘offensive’ and ‘defensive’ tactics for example. A different tactic should be applied between playing against the previous champion or a relegation candidate in the competition. Both tables are based only on expert opinion so a human error factor is possible, demand and supply is subjective and it is time consuming for the coach to obtain the needed information.

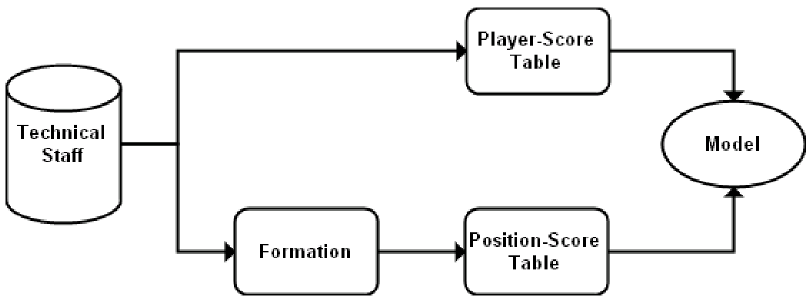


Figure 3.1 - Schematic view of Coach & Scout Assistant procedure

The player-score and position-score table form the input data for the program called Coach & Scout Assistant (CSA). The input data will be transformed into a list of weights, reflecting the performance of the players on any of the positions. With these weights the optimal formation can be determined. The main added value of the model is that it can be used for scouting. The difference between CSA and our procedure can be explained with the use of a schematic view. In figure 3.1 a schematic view of the procedure in CSA is depicted. If we compare the figure with figure 2.2 we see that the main difference is that CSA is based on expert opinion. We will use real-life data in order to optimize the line-up. Another major difference is that CSA does not contain a team component in the procedure.

Sierksma (2006) introduces two computer support systems; namely Coach and Scout Assistant and Effectivity in Action. The first system is used to calculate value-added and best positions of scouted players and the second one is currently used to generate graphs that reflect the effectivity of the teams during the match. A description of CSA is given in Boon & Sierksma (2003). In EiA all actions of both teams are registered using touch screen computers. For every action a score between 1-5 are given (the Likert scale). The main computer uses these actions and scores to calculate the current effectivity of both teams in order to draw a graph. As the match ends the teams receive an overall effectivity score, it is also possible to present the scores for individual players. These graphs are used in newspapers and one club uses it on a big screen in the stadium. The added value of EiA is that it uses analytical formulas that apply correction on facts like that goalkeepers are fewer times in ball possession than other players, also smoothing techniques are applied to obtain the effectivity graphs. The data that will be used in this thesis is also obtained from EiA.

With the information provided by CSA it is possible to do more research. Siesling (2004) used the calculations in CSA to determine the salaries of the players. The value-added of each player on any position are generalized to one core-value in order to determine the salary. A simulation is used to determine the salaries, injuries and suspension are taken into account in the simulation. Further CSA is used as input for a support system to determine the optimal training schedules for team sports named Soccer Training Support (STS) in Kleef (2005). Given the strongest team composition from CSA; the STS system can apply two visions. According to the first vision a player must train for his strongest position and the second vision lets the player train for the position which improves the overall team performance.

In practice it is possible that the optimal line-up from CSA does not match with the composition of the coach. In this case the information provided by CSA of scouted players is not correct anymore because it is compared with a different line-up than the line-up of the coach. In Kok (2005) this problem is tackled. Inverse optimizing techniques have been used to match the coach's team composition and the composition from CSA, with the coach's composition as input.

A different approach is applied in Duch, Waitzman and Amaral (2010). A network approach is developed that provides a powerful quantification of the contributions of individual players and of overall team performance. The performance of players in the European Cup 2008 soccer tournament are analyzed with the method. The method used is interesting but not useful for our research because the team performance is analyzed per match. Oberstone (2009) developed six independent variable multiple regression models to distinguish the performance of English Premier League soccer clubs. With the method it was possible to separate the top clubs from the middle and bottom clubs. The method applied in this paper is interesting but not useful for our study; this is because the identification of the team rating is based on only team statistics and not individual player performances. In Hirotsu & Wright (2003) a dynamic programming approach is used to find the optimal substitution strategy for a football match. Again the English Premier League data is used in order to maximize the probability of winning or the expected number of league points. This is an interesting topic in soccer but again not useful because it is out of the scope of this research. Our goal is to optimize the line-up which can be described as starting with the best possible line-up to the game, the anticipation on unexpected events like an injury or suspension of an player is out the scope of this thesis.

3.1.2 OTHER SPORTS

Baseball is a sport which provides many opportunities for the collection and evaluation of numbers. This is because the discrete nature of baseball. It is one of the first sports in which the analysis of statistics had been applied. The challenge in baseball was finding new measures of baseball performance in order to evaluate the performance of baseball players. The link between operational research and baseball is explained in Richard (2002). The paper shows how the use of statistical analysis of player performances has evolved in the last 30 years. Which has resulted in the introduction of linear regression analysis to measure player performance.

3.2 SIMILAR STUDIES IN OTHER AREAS

Human resources management is an area with similarities with our study. It is the same problem of matching supply and demand. The supply of the qualities of the employees should be optimally matched with the demand requirements of the available functions. Also in human resources management teamwork is essential for success. In Barrick & Stewart (1998) different methods of operationalizing team composition variables are examined. The purpose of the study was to evaluate how member characteristics of work teams relate to differences in team effectiveness. In terms of sports we should think how we can assess which individual qualities in soccer have a positive effect on teamwork and which a possible negative effect. By applying this, the quality of team would not be simply the sum of all individual qualities.

4 DATA DESCRIPTION

This chapter will describe the database that will be used in the thesis. As mentioned before the database is property of ORTEC TSS and it is obtained by analysts using the software Effectivity in Action (EiA). First some general information about the software EiA will be explained; how the data is gathered in particular. The second part will cover the available data for our research with some basic statistics. We will end this chapter with the explanation of the formulas used in the software EiA.

4.1 EFFECTIVITY IN ACTION

Effectivity in Action is an application owned by ORTEC TSS to gather data from various sports. It is currently applied in soccer, hockey and rugby. A soccer match is analyzed by 4 analysts, 2 per team. One analyst observes the actions during the match and registers the locations on a tablet. The second analyst uses a computer to store the actions in the database. For soccer there are 17 predefined actions which are summarized in table 4.1. Most actions have additional attributes which give detailed information about the action. An action can have at most 3 additional attributes and some actions does not have any additional attribute. A goal kick is an action without any attribute and a penalty kick can have 3 attributes. A penalty kick scored in the high right corner for example has three attributes; right corner, high and goal. EiA contains in total 104 different actions with the attributes included. All actions and their corresponding attributes are listed in the table A.1.

Actions	
Reception	Free kick awarded
Attacking action	Foul
Offside	Pass
Goal attempt	Save on goal attempt
Goal kick	Penalty
Move	Defending action
Corner	Direct free kick
Throw in	Indirect free kick
Interception	

Table 4.1 - List of all actions in EiA

During a soccer game the analysts decide how effective the actions performed by the players are. The actions are then graded based on their effectivity. We refer to Koerkamp [2004] for detailed information about what effectivity is. The grading of the action will be explained in chapter 4.3.

4.2 AVAILABLE DATA

The data that is available for the research are all actions with the ball for every player on the field from 2009 till 2011. In chapter 1.2 we mentioned that ORTEC started in 2007 with the use of EiA in sports. In the older versions of the software it was not possible to store the exact locations of the actions. The exact coordinates are available since November 2010, before November 2010 the field was split in 14 parts and it is only known in which part of the field the action was performed. So we cannot use all of the available data to research this topic. The model will be tested on Ajax a major Dutch soccer club. The reason that we have chosen for this club is that it has played the most matched in the season 2010-2011 in several competitions. Ajax played in total 55 matches in 5 different competitions and from these 55 matches 40 are useable for our research. The remaining 15 matches are not useable because often the exact coordinates are not stored or the match contains missing variables. Missing variables in the data simply means that the analysis of the match is not completed. The number of matches played per competitions and the useable matches can be found in table 4.2. The first three competitions are national competitions and the last two are continental competitions.

Competitions (2010-2011)	Matches played	Useable matches
Eredivisie	34	23
Dutch Cup	6	5
Dutch Super Cup	1	0
Champions League	10	8
Euro League	4	4
	55	40

Table 4.2 - Number of matches played per competition and the useable ones for Ajax

These 40 matches contains in total 60274 actions, 33958 of these actions are performed by Ajax. The remaining 26316 actions are from the opponents in the 40 different matches. The actions of the opponents are included because without them it is not possible to make calculations for the length and angle of the passes, we also need the passes which are interrupted by the opponent.

In table 4.3 we summarized the count and the percentage of the actions occurred in the 40 matches for Ajax. The actions in the table contain for every action all attributes, no distinction is made for the attributes of the actions. If we take a detailed look at the actions in the dataset we see that more than half of the actions are passes. This is indeed not strange because passing is the fastest way to move the ball in a field. In chapter 5 we will describe how the soccer skills are determined with the use of these actions summarized below.

Action	Count	Percentage
Interception	4863	14,32%
Goal kick	229	0,67%
Save on goal attempt	176	0,52%
Foul	565	1,66%
Throw in	690	2,03%
Corner	220	0,65%
Penalty	4	0,01%
Indirect free kick	452	1,33%
Direct free kick	38	0,11%
Offside	87	0,26%
Pass	18190	53,57%
Defending action	2647	7,79%
Attacking action	2691	7,92%
Free kick awarded	572	1,68%
Move	1111	3,27%
Reception	735	2,16%
Goal attempt	688	2,03%

Table 4.3 - Statistics per action for Ajax

4.3 GRADING IN EIA

The grading part of the software EiA is one of its main added values. The grading part consists of analytic formulas to calculate the grades of a player. The formulas are also capable of correcting the grade of a player if it has performed less actions than average. Finally smoothing techniques are used to provide effectivity graphs with EiA. This is all possible because all the actions registered by the analysts have been judged while registering. The judgement of the actions is based on their effectivity during that moment in the game. All 17 actions listed in table 4.1 are judged in the same way by using a 5 point likert-scale. The 5 point likert-scale is divided as:

1 - Very bad action leading to a goal against or a miss of a goal chance
2 - Bad action, loss of possession
3 - Neutral action with no improvement of the situation on the field
4 - Good action which improves the situation of the team.
5 - Very good action leading to a chance to score or stopping of a goal attempt of the opponent

From the division of the 5 point likert-scale above it is clear that grading in EiA is based on the effectivity of the action. Only the result of the action is of importance, if a players performs the right action at the right time but the opponent interrupts the action it is a bad action. The judgement of the actions are used to aggregate the actions to a grade on scale 1-10 for different levels. Currently there are 4 different levels used:

- *action*
- *action per player*
- *player*
- *team*

The first level is based on one action performed by all players, the second is based on one action per player, the third level is based on all actions of a player and the last level is based on all actions performed by the team. With a non-linear transformation these actions are aggregated to a grade on scale 1-10. To prevent that the grade converges to the mean when the number of actions grows the non-linear transformation has a s-shape with reference point (3, 5.5).

The choice of the steepness of the S-shape depends on the number of actions on which the mean is calculated. The function must be steeper if we have more observations. The S-shape function is defined as

$$f(mean|a) = 5.5 + (-1 + 2 \cdot sig) \cdot a \cdot \log(1 + 2 \cdot e^{\log\left(\frac{1}{4}e^{\frac{4.5}{a}} - 1\right)} \cdot |mean - 3|)$$

with

$$sig = \begin{cases} 1 & \text{if } mean \geq 3 \\ 0 & \text{if } mean < 3 \end{cases}$$

This part of the formula ensures that with a mean of less than 3 the final grade will be below 5.5 and with a mean larger or equal to 3 the final grade will be above 5.5.

The function $f(mean|a)$ has a domain of [1 5] and a scope of [1 10] so it can be used as a transformation function. The parameter a determines the steepness of the function in point (3, 5.5) and it is dependent of the mean number of actions of the desired level. In figure 4.1 the s-shape function is depicted for different values of a. Here we see that the line with a = 7 is almost linear, this line is used to grade when less actions are performed. As the number of actions increases, the shape of the line will change, a more curving line will be selected. The different curves in the s-shape function will ensure that the variation of the [1 10] scale grades will remain high.

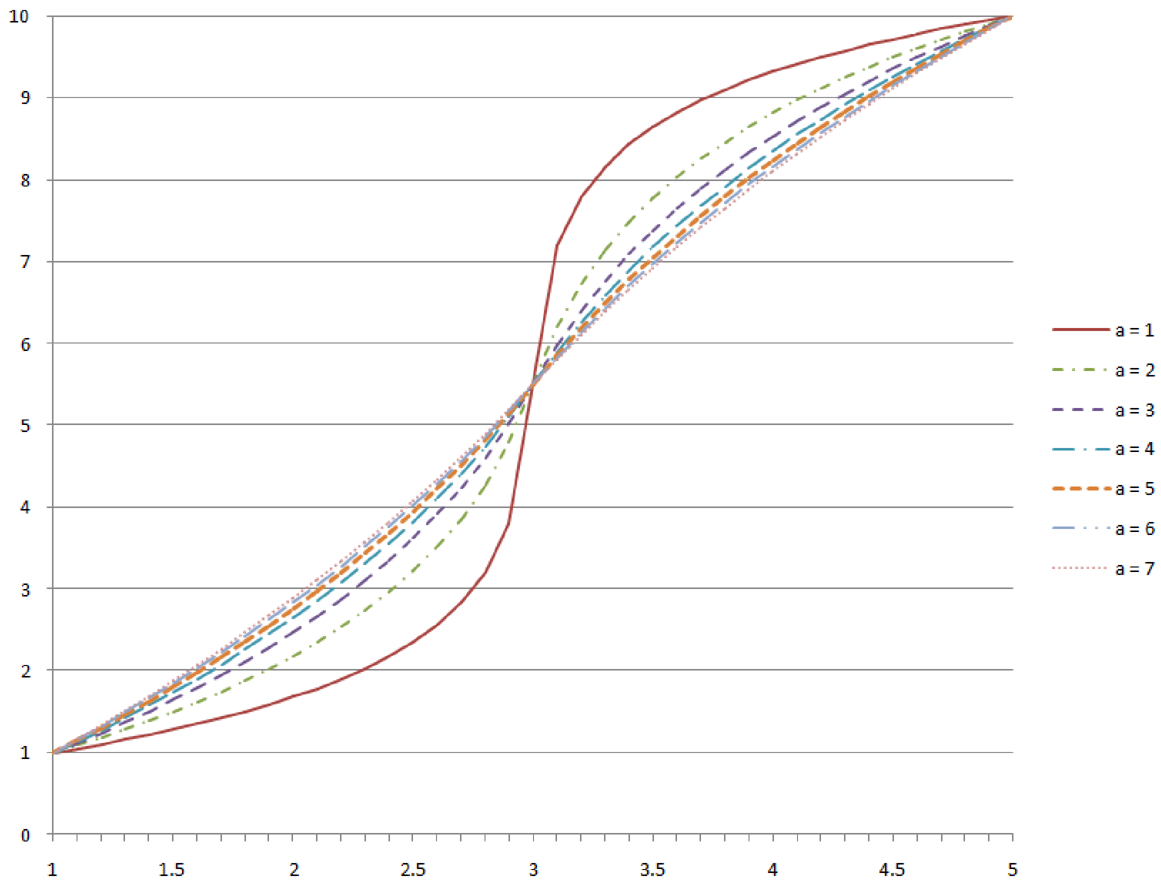


Figure 4.1 - S-shape transformation function

The input of the function f is the mean of the likert-scale grades. For the calculation of the mean the likert judgements l, the weights z_l, and the sum of the likert judgements Y_{ijl} are used. The likert judgement l = 4 is included as 4,4; this is done to increase the effect of it on the grade. The mean of the actions can be calculated with the formula below:

$$mean = \frac{\sum_{l=1}^{5} z_l \cdot Y_{ijl} \cdot l}{\sum_{l=1}^{5} Y_{ijl} \cdot z_l}$$

i	:	indices for player
j	:	indices for actions
l	:	likert-judgement, l = 1,..,5
Y_{ijl}	:	sum of grade l for action j of player i
z_l	:	weight of grade l

The weights corresponding to the likert judgements are:

$$\begin{aligned} z_1 &= 2 \\ z_2 &= 1 \\ z_3 &= 1 \\ z_4 &= 1 \\ z_5 &= 4 \end{aligned}$$

Here we see that the likert judgements 1 and 5 are given a higher weight z_l, this is done because the frequency of the grades 1 and 5 is lower than the other grades. We refer to Koerkamp [2004] for a detailed sensibility analyses on the effect of the weights to grades. We will end this chapter with a small example.

Example 1

Suppose we have 2 players A and B with likert judgements:

l	1	2	3	4	5
Player A	2	4	10	6	4
Player B	1	2	5	3	2

The mean of the likert judgements of player A is:

(2·2·1)+(4·1·2)+(10·1·3)+(6·1·4.4)+(4·4·5) / (2·2)+(1·4)+(1·10)+(1·6)+(4·4) = 3.71

and of player B

(1·2·1)+(2·1·2)+(5·1·3)+(3·1·4.4)+(2·4·5) / (2·1)+(1·2)+(1·5)+(1·3)+(4·2) = 3.71

We see that both players have the same mean while player A has performed more action. With the S-shape function it is possible to distinguish for this fact, by using a different parameter for the steepness for the players. The grades corresponding for mean 3.71 with different values of parameter *a* are listed in the table below.

	Grades f(*mean*|*a*)						
	a = 1	*a* = 2	*a* = 3	*a* = 4	*a* = 5	*a* = 6	*a* = 7
grade	9.0	8.3	7.9	7.7	7.6	7.5	7.4

Because player A has performed more actions than player B, the grade of player A will be higher because the parameter will be lower than the parameter for player B.

With the functions described in this chapter the grades of the 5-point likert-scale are aggregated into a grade on scale 1-10 in the software Effectivity in Action. In chapter 5 we will discuss the usability of these function for grading the soccer skills.

5 METHODOLOGY

This chapter will describe all methods used in this thesis and will motivate the choices behind the methods. All sub-questions mentioned in chapter 2 will be answered. The first part will give a description of the changes made in the raw data obtained from EiA in order to use them in our thesis. In the second part we will describe how we used data obtained from the first part to create the profiles. The final part will explain how these profiles are matched to obtain the optimal line-up for a team.

5.1 DATA

The raw data we obtained from EiA contains all actions performed with the ball as explained in chapter 4.2. In this chapter we will explain how we used the data to create the soccer skills. The second part of this chapter will describe how the player and team tactics are obtained from the data. The first two sub-questions related to the data adjustments are stated below and will be answered in this chapter.

- *How do we translate the actions in the database into soccer skills of players?*
- *How do we reveal descriptive statistics related to individual and team tactics with historical data per match?*

5.1.1 PLAYER DATA

With the 17 actions and their attributes defined in EiA it is possible to distinguish between 104 different actions. As described before these are actions with the ball and needed to be translated into soccer skills. An action like goal kick does not need to be translated into a skill but the passes in the data could be split into several types of passes. The data also contains several heading attributes for different actions. They can be added in one skill heading. The soccer skills will be defined by splitting and aggregating the actions.

In table 4.3 we have seen that more than half of the actions are passes as expected. The passing action in EiA contains also passes with the head and crosses. Passing is essential in soccer because it is the fastest way to move the ball in the field. With the locations of the passes we are able to calculate the angle and the length of the pass, so we can split the passes. These calculations will be explained first. The locations consists of the x and y coordinates of the actions, but a soccer field is not a square so the length and the width are not the same. The coordinates for the x and y coordinate are stored as a number between 0 and 100. The left upper corner is stored as point 0,0. To calculate the length and the angle of a pass we need to know the width and length of the soccer field. The length and the width varies because the official rules for them differ per competition. The rule for the size of the field in the Dutch competitions is set by the KNVB; the width must be between 64 and 69m and the length between 100 and 105m. The UEFA has fixed sizes for the continental matches and these are 68m by 105m. Because the data we use contains mostly matches played in the Dutch competition we have chosen to take the average of the allowed size for the width and length of the field. For the length this means that it will be set to 66.5m and the width to 102.5m.

With the locations of pass and its destination given from the data the length of a pass can be calculated easily by using the Euclidean distance formula and adjusting for the length and width. After adjusting for the length and width the Euclidean distance formula will be:

$$length = \sqrt{((x_2 - x_1) \cdot 1.025)^2 + ((y_2 - y_1) \cdot 0.665)^2}$$

To calculate the angle of a pass the inverse tangent function is used, again we have to adjust for the length and width. Because the range of the inverse tangent function is restricted to (-90°,90°) we have to add 180° to get the correct angles between 0° and 360°. The formula used to calculate the angle is:

$$angle = \tan^{-1}\left(\frac{(y_2 - y_1) \cdot 0.665}{(x_2 - x_1) \cdot 1.025}\right) + 180$$

With the lengths and the angles of the passes we can split the passes into different soccer skills. First we will make a distinction between short and long passes. The threshold value between short and long passes is set to 25 meters. This value is obtained by rounding down the quarter of the length of a soccer field, which is in our case 102.5 meters. The second distinction will be the angle of the pass. With the angle we can determine the direction of the pass and we will use 3 directions for the passes; passes forward, backwards and wide. We will treat passes with an angle between 0° and 60° or 300° and 360° as passes forward, passes between 60° and 120° or 240° and 300° as passes wide and passes between 120° and 240° as passes backwards. Because all matches are analyzed from left to right the angles for forward and backward passes does not change in the second half of the game. Figure 5.1 shows the directions of the passes with the angles. We have not made a distinction between passes left and right because this is not of importance to the passing skill of a player, so a pass wide could be to the left or right side.

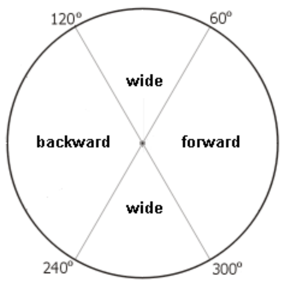


Figure 5.1 - Directions of passes

With the length and the angle of the passes we have split the passes into 6 different passing skills and we treat the crossing attribute as a separate skill. The coordinates have also been used to determine the length of a shot at goal to distinguish between long shots and the finishing skill of a player. The finishing skill is determined by checking if the shot is taken inside the penalty box. Further a set pieces skill is created which contains the off-ball situations corners and free kicks. These skills are not treated separately because they occur not frequently during a soccer game. With all the actions performed with the head a heading skill is created; this skill includes interceptions, passes, attacking and defending actions with the head. As a defensive skill we created interceptions; this skill contains all defending actions and interceptions except the ones made with the head. To distinguish between the defensive skill of a defender and attacker we also created a pressing skill. The pressing skill will contain the sliding and body attributes of the attacking actions. The last skill created is the goalkeeping skill, this skill includes the saves on goal attempts and interceptions of the goalkeeper. For three actions we have not made any changes, these actions could be translated directly into soccer skills. The three unchanged skills are goal kicks, reception and dribbling skills. With the soccer skills defined with the actions as described above we have answered our first sub-question:

- *How do we translate the actions in the database into soccer skills of players?*

In table 5.1 we summarized how many times the soccer skills are performed and their percentage in all matches of Ajax. We see in this table that it was indeed a good idea to split the pass actions into different kinds of passes. The pass action in the previous table contributed for more than 50% of the action which is now reduced by splitting them. Besides the 6 different passing skills we also included the short and long passing skills. This is done to have a general passing skills for the players.

Soccer skill	Count	Percentage
Short pass back	3288	10.44%
Short pass forward	5074	16.10%
Short pass wide	4795	15.22%
Long pass back	926	2.94%
Long pass forward	2092	6.64%
Long pass wide	1128	3.58%
Crossing	656	2.08%
Heading	1614	5.12%
Finishing	276	0.88%
Long shots	354	1.12%
Dribbling	1111	3.53%
Reception	735	2.33%
Interceptions	5768	18.31%
Goalkeeping	462	1.47%
Goal kicks	229	0.73%
Set pieces	710	2.25%
Short Passes	13157	41.76%
Long Passes	4146	13.16%
Pressing	2290	7.27%

Table 5.1 - Statistics per skill for Ajax

5.1.2 TEAM DATA

The 19 soccer skills we defined in the previous part are also useable to obtain the team data. These data will be used to create the position and team profile as explained in chapter 2.3. The soccer skills performed at the positions will be used for the position profile. For the team profile we will use all the soccer skills performed per match. To obtain more information about the tactics of a team the skills are used to create more detailed team based statistics. These statistics can be categorized into four groups, each group represents a different aspect of the soccer tactics. The four groups are:

- *Attacking statistics*
- *Defending statistics*
- *Passing statistics*
- *Discipline statistics*

The difference of the statistics in these groups is that they contain more detailed information of the skills performed. Besides the number of times a skill is performed we will take a closer look at the locations. For the locations we will distinguish between skills performed on the left-, right side or centre and skills performed at the own half or opponents half of the field. For the length of the passes we distinguish like before between short and long passes. Besides the locations we will check if a skill is performed successfully. The separation between successful and unsuccessful will be made with the Likert-scale defined in chapter 4.3. A skill performed with grade 3 or higher will be marked as successful. Another useful statistic for soccer is the percentage of the skills performed compared with similar skills; an example would be the percentage of the long and short passes.

Soccer statistic	Count	Success	%	Length	Direction	Half	Location
Passes	x	x	x	x	x	x	x
Possession	x		x				
Red cards	x						
Yellow cards	x						
Fouls	x					x	x
Interceptions	x	x	x			x	x
Goals concerned	x						
Crosses	x	x	x				x
Goals scored	x						
Shots	x	x	x	x			x

Table 5.2 - List of additional team statistics

In table 5.2 we summarized which additional statistics will be calculated for the soccer skills to create detailed team statistics. For the length of the soccer statistics we distinguish between short and long, the directions are forward, backward and wide, the half distinguishes between own and opponents half and the location between the left- right flank and centre of the field. The order of the categories of the statistics is: passing, discipline, defending and attacking statistics. The table contains some statistics which are not mentioned before; this is because these statistics could not be categorized as a soccer skill a player can perform or are simply not informative enough for individual players. The new statistics are, possession, goals scored and concerned, red and yellow cards and fouls committed.

To reveal the individual and team tactics of a team we will use the soccer skills defined in the previous chapter and they will be used to create more detailed statistics. This will be the team data and with this data we have answered our second sub-question:

- *How do we reveal descriptive statistics related to individual and team tactics with historical data per match?*

5.2 PROFILES

With the player and team data defined in chapter 5.1 we are now able to create the profiles we mentioned before. Three profiles will be created; a player, position and team profile. The player data will be used to create the player profile and the team data to create the position and team profile. Each of the sub-question stated below corresponds to one profile and will be answered in this chapter.

- *How do we grade the skills of the players?*

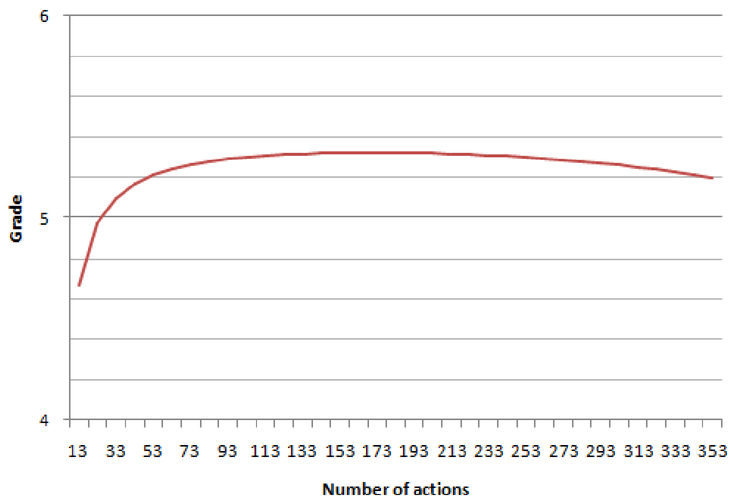
- *How do we weight the skills for the positions in order to distinguish different tasks for the positions given the formation and tactics?*
- *How do we expose and grade the quality needed for every line/flank based on the team tactics?*

5.2.1 PLAYER PROFILE

To create the player profiles the soccer skills defined in chapter 5.1 must be graded in order to distinguish the quality of the skills for different players. In chapter 2.1 we explained the terms quality and skill, and we described what the quality of a soccer skill is. To grade the skills we will use the grading formulas of the software EiA. However these formulas are based on the effectivity of the actions and our purpose is to grade the quality. Therefore some changes have to be made in the formulas to grade the quality. The following small example shows with a graph that the formula is based on the effectivity of the skill.

Example 2

Assume we grade a player for the passing skill. The player has played 40 matches and has given 3 bad passes. We will increase the number of 'neutral' passes to analyse the impact of it on the grade. The graph below shows the grade of the player with an increase of the number of 'neutral' passes. The y-axis shows the grade of the player and the x-axis shows the number of passes for each player. The graph starts with 13 passes (3 bad and 10 neutral passes) and we increase in each step the number of neutral passes by 10.



We see that the grade in the begin of the graph slightly increases from 4.7 to 5.3 and after that, it remains steady till the number of passes reaches 200. After 200 passes it starts to decrease, despite the fact that the number of 'neutral' passes increase. This is because the formula is based on the effectivity and not quality.

The grading formula can be adjusted to judge the quality of a skill. Currently the likert grades are divided into 3 categories; non-effective, neutral and effective. But to judge the quality of a skill we need 2 groups; unsuccessfully and successfully performed skills. A requirement for the two groups we need is that the previous 'neutral' actions must have a positive contribution to the grade because it is performed successfully. This can be done by slightly increasing the likert value of the 'neutral' action. This is the answer for the sub-question related to the player profiles is:

- *How do we grade the skills of the players?*

To determine the value of this increase we will compare the results of several test scenario's. In the previous example we examined the changes when 3 bad passes have occurred; now we will also consider the changes when 3 good passes, very good, very bad passes and finally only 'neutral' passes have occurred. By doing this we will be able to select the best increase in order to grade the quality of the several skills performed by players. The results of the several test scenarios will be discussed in chapter 6.

5.2.2 POSITION PROFILE

For the creation of the position profile the player data will be used. As explained before the player data contains for each soccer skill the grades and the count of the skills performed. In chapter 4.3 we mentioned 4 different grading levels on which the formula is based on. The grading level needed must be based on the actions performed per position. This is almost similar to the actions per player level, because when a player is substituted and the formation does not change he plays on the same position. The total amount of actions performed at the position where the substitution has occurred will be the sum of the actions of both players. This is why we can use the action per player level to grade the skills performed on the positions. The only difference is that when a player is substituted the grade of the position will be the grade of both players played on the position. This is how the grades for the position profile will be determined. Because for each position different skills are of importance we also need to give weights to the skills. These weights will be different for each position in the formation. The sub-question associated to the position profile is:

- *How do we weight the skills for the positions in order to distinguish different tasks for the positions given the formation and tactics?*

To determine the weights per position we will use the number of actions performed at the positions. First the fraction of the skills performed per position will be calculated, with this we will know which skills are used more often at the positions. Besides the fraction at the positions the fraction per skill at the different positions is also of importance. This will show us the distribution of a certain skill between the positions. The next step will be combining these two fractions into one fraction. This will be done by giving the two fractions a weight; this is needed because of the distribution of the skills. We have seen in table 5.1 that the interception skill is the largest group of all skills. Without the weights for the fraction the interception skill will have a too large weight for all positions. Several weights will be tested in order to determine the most suitable weights for the fraction. The exact value of the weights will be shown in the results chapter. The final step for calculating the weights will be scaling the weight so that the combined fractions will sum up to 1. All steps performed to determine the fractions for the positions are summarized below.

Calculation of the fractions

1. Calculate the fraction of skills performed per position, for all positions.
2. Calculate the fraction per skill at the different positions, for all skills.
3. Combine the fractions from step 1 and step 2 by giving the 2nd fraction a higher weight.
4. Scale the fraction obtained by step 3 per position so that they sum to 1.
5. Take only the grades above x%

To make the procedure of determining the fractions of the skills per position clear we will give a small example. To make the example simple we will assume that there are only 2 positions and 3 skills.

Example 3

Suppose that there are only 2 positions and 3 skills. The table below gives the number of occurrences of the skills per position. With the table below we can perform each step of the procedure. For each step one example of the calculation will be given.

	Skill 1	Skill 2	Skill 3
Position 1	200	75	30
Position 2	75	100	175

Step 1

Here the fraction of skills performed per position will be calculated. The first skill of position 1 is calculated as follows: 200 / (200 + 75 + 30) = 0.65. After repeating step 1 for all skills and positions the fractions will be:

	Skill 1	Skill 2	Skill 3
Position 1	0.65	0.25	0.10
Position 2	0.21	0.29	0.50

Step 2

To calculate the fraction per skill at the different positions for skill 1 and position 1 is calculated as follows: 200 / (200 + 75) = 0.73. Again we repeat this step for all skills and positions. The results are shown below.

	Skill 1	Skill 2	Skill 3
Position 1	0.73	0.43	0.15
Position 2	0.27	0.57	0.85

Step 3

With the 2 fractions obtained from step 1 and 2 we can now combine these 2 fractions into 1 fraction. As explained before we will give weights to the fractions. In this example the weight of the first fraction will be 1 and the weight of the second fraction will be 3. The calculation of the fraction will be (0.66 + (3 x 0.73)) / 4 = 0.71. After repeating this step for all skills and position the fractions will be:

	Skill 1	Skill 2	Skill 3
Position 1	0.71	0.38	0.13
Position 2	0.26	0.50	0.77

Step 4

The final step will be scaling the fractions from step 3 so that they sum to 1 per position. If we perform this step for the first position and skill this will be: 0.71 /(0.71 + 0.38 + 0.13) = 0.58.

	Skill 1	Skill 2	Skill 3
Position 1	0.58	0.31	0.11
Position 2	0.17	0.33	0.50

The calculation explained above is the fraction of the skills and positions. But it does not make a distinction between the different lines on the field. With the different lines we mean the defence, midfield and attack. Other lines which are not used in the method above are the flanks (left, right) and centre of the field. This can be done by calculating the count of the actions which occur on these lines and rank them. For the defence, midfield and attack the weights will be [w_3, w_2, w_1]. The highest ranked line will get weight w_3 and the lowest rank will get weight w_1. The flanks and centre will also be ranked and a weight will be added depending on the rank. The weight for the flanks and centre will be [v_3, v_2, v_1] so that the maximum weight will be w_3+v_3 and minimum weight w_1+v_1. These weights will be multiplied with the fraction calculated with the previous method. By doing this the importance of the different lines on the field will be combined with the positions. In table 5.3 the positions corresponding to lines are listed.

Line 1	Positions	Line 2	Positions
Defence	DR, DCR, DCL, DL	Right	DR, MCR, FR
Midfield	MCR, MC, MCL	Centre	DCR, DCL, MC, FC
Attack	FR, FC, FL	Left	DL, MCL, FL

Table 5.3 - Positions corresponding to the horizontal and vertical lines

The method for the calculation of the weights with which the fractions will be multiplied with is summarized below.

Calculation of the weights

1. Calculate for each skill the average count at the Defence, Midfield and Attack.
2. Rank the average count at the horizontal lines in decreasing order.
3. Set the ranked average count of the skill at the corresponding position to w_3, w_2, w_1.
4. Calculate for each skill the average count at the right side, left side and centre.
5. Rank the ranked average count at the vertical lines in decreasing order.
6. Set the average count of the skill at the corresponding position to v_3, v_2, v_1.
7. Sum the weights given at the horizontal and vertical lines.

5.2.3 TEAM PROFILE

Because the player and position profiles does not contain any information about team play and team tactics an additional profile is needed as explained before. The team profile will provide the information based on the team statistics. The team data and partly the player data will be used for the creation of the profile. The purpose of the team profile is to expose and grade the additional quality needed for the different lines and flanks based on the team tactics.

- *How do we expose and grade the quality needed for every line/flank based on the team tactics?*

The team data will be used first to determine the tactics of the team. These tactics will be based on the aggregated information obtained from the 40 matches we possess. The attacking, defending and passing statistics will be used to determine the tactics. Because the team data only contains team based statistics and does not contain any information about the positions, the player data will be used too (see table 5.2). The player data will be aggregated to position bases statistics.

With the team and position data pre-defined decision trees will be used to expose and grade the quality needed for every line and flank given the team tactics. For each one of the three statistics groups the determination of the general tactics will be explained first. The next step is to use position based statistics to determine the tactics per line and flank. In each node of the tree a statistic will be compared, this can be skills performed at the flanks with the centre or the midfielders with the attackers. The comparison will be based on the following rule: A is strictly larger than B if the difference between A and B is at least 10%; otherwise A is almost equal to B. The same rule holds for A is strictly smaller than B. The reason we use this rule is that the small differences in the exact values are not relevant in our case. If A is 15 and B is 17, then B is larger than A but the difference is 2. For simplification we will use the following notation in the remainder of this thesis: strictly larger than (>>), strictly smaller than (<<) and almost equal to (≈).

From the attacking statistics the crosses and the shots at goal will be used. First the method of the shots taken at goal will be explained. By analyzing the shots at goal the distinction between long shots and shots taken inside the penalty area will be made. Here we will check which types of shots are taken strictly more by the team. With strictly more we mean that the difference between the 2 types of shots must be at least 10%, otherwise we assume that the difference is too small and the types of shots are equally performed by the team. The next step will be the analysis of the shots for the different lines of the field. Here we will consider only the midfielders and the attackers because the majority of the shots are taken in general by midfielders or attackers. The defenders are not included in this method. This is the first assumption made in the team profile.

- **Assumption 1:** Shots at goal are taken by midfielders or attackers.

Now we will analyze the line where the shots are taken mostly by applying the same rule as for the shots only. The outcome of the method will be the line which is responsible for the majority of the type of shots taken, which can be also both lines in this case. If one of the lines is responsible for the majority of the shots the team will need at least one player who is capable of taking these shots. Some possible outcomes of the method for the crosses are: one midfielder with long shot skill, one attacker with finishing skill or one forward with finishing or long shot skill. So the method checks first which type of shot is preferred by the team and then compares the shots taken at specific lines on the field. The exact method for the shots at goal is summarized below. The id of the different scenarios are placed within brackets in the method below, the first character represents the skill and the number the corresponding scenario.

Method : Shots at goal

IF long shots ≈ short shots
  IF shots taken by Midfielders >> shots taken by Attackers
    [S.1] at least 1 *Midfielder* with *finishing* or *long shots* skill
  ELSEIF shots taken by Midfielders << shots taken by Attackers
    [S.2] at least 1 *Attacker* with *finishing* or *long shot* skill
  ELSEIF shots taken by Midfielders ≈ shots taken by Attackers
    [S.3] at least 1 *Midfielder* or *Attacker* with *finishing* or *long shot* skill

ELSEIF long shots << short shots
  IF short shots taken by Midfielders >> short shots taken by Attackers
    [S.4] at least 1 *Midfielder* with *finishing* skill
  ELSEIF short shots taken by Midfielders << short shots taken by Attackers
    [S.5] at least 1 *Attacker* with *finishing* skill
  ELSEIF short shots taken by Midfielders ≈ short shots taken by Attackers
    [S.6] at least 1 *Midfielder* or *Attacker* with *finishing* skill

ELSEIF long shots >> short shots
  IF long shots taken by Midfielders >> long shots taken by Attackers
    [S.7] at least 1 *Midfielder* with *long shot* skill
  ELSEIF long shots taken by Midfielders << long shots taken by Attackers
    [S.8] at least 1 *Attacker* with *long shot* skill
  ELSEIF long shots taken by Midfielders ≈ long shots taken by Attackers
    [S.9] at least 1 *Midfielder* or *Attacker* with *long shot* skill
END

The method for the crosses is a bit different than the method for the shots at goal. The first difference is that we will first check if the crossing skill is a part of the team tactics. This is because not every team plays with crosses. This will be determined by comparing the average crosses per match of the team with a value *x*. If the average crosses are above this value then we assume that crosses are part of the team tactics, otherwise crosses are not part of the team tactics given by the coach. The next step in the method in case that the crosses are part of the team tactics is comparing the crosses made at the flanks and the centre of the field. The same rule with strictly more, strictly less and equal to as for the shots at goal will be used to compare the crosses. The final step will be identifying the location of the line where a player is needed with the crossing skill. In case of the crosses we will consider all three lines on the field; the defence, midfield and attacking line. The line where the majority of the crosses are given will be the line where a player is needed with the crossing skill. The possible outcomes of the method are based on the location of the crosses (flanks or centre) and the line where they occur mostly (defence, midfield and attack). The method for the crosses with the possible outcomes is listed in the figure below.

Method : Crosses

```
IF      average crosses per match > x
        IF      crosses given at flanks ≈ crosses given from centre
                IF      most crosses given by Defenders
                                [C.1] at least 1 Defender with crossing skill
                ELSEIF  most crosses given by Midfielders
                                [C.2] at least 1 Midfielder with crossing skill
                ELSEIF  most crosses given by Attackers
                                [C.3] at least 1 Attacker with crossing skill
                ELSE
                                [C.4] at least 1 player in team with crossing skill

        ELSEIF  crosses given at flanks >> crosses given from centre
                IF      most crosses given by Defenders at flank
                                [C.5] at least 1 Defender at flank with crossing skill
                ELSEIF  most crosses given by Midfielders at flank
                                [C.6] at least 1 Midfielder at flank with crossing skill
                ELSEIF  most crosses given by Attackers at flank
                                [C.7] at least 1 Attacker at flank with crossing skill
                ELSE
                                [C.8] at least 1 player at flank with crossing skill

        ELSEIF  crosses given at flanks << crosses given from centre
                IF      most crosses given by Defenders at centre
                                [C.9] at least 1 Defender at centre with crossing skill
                ELSEIF  most crosses given by Midfielders at centre
                                [C.10] at least 1 Midfielder at centre with crossing skill
                ELSEIF  most crosses given by Attackers at centre
                                [C.11] at least 1 Attacker at centre with crossing skill
                ELSE
                                [C.12] at least 1 player at centre with crossing skill

ELSE
                [C.13] Crosses are not part of team tactics

END
```

The defending statistics are the second group of statistics which will be used. Because interceptions are the most important aspect of defending, they will be used to determine the defending tactics of a team. To determine if the team is playing with defensive tactics the number of interceptions performed at the own half will be compared with the interceptions performed at the opponents half using the rule described before. We can assume that a lot of interceptions at the opponents half in soccer means that the team is pressing to dominate the game. While a lot of interceptions at the own half means that the opponent is dominating in general. The assumptions we use here are:

- **Assumption 2:** Team is playing with defensive tactics if the interceptions at the own half are strictly larger than the interceptions at the opponents half; otherwise team is playing with offensive tactics

After determining the defensive or offensive tactics with the help of the interceptions it is important to know where these interceptions occur. We will compare the interception made at the flanks with the interceptions made at the centre of the field. Despite we expect that the majority of the interceptions will occur at the centre instead of the flanks, it is informative to know the distribution of the interceptions. If a team is pressing to dominate, it is usual that the players positioned at the flank must press also to capture the ball from the opponents. To include this scenario the distribution between the flanks and the centre is important. The final step of the method is determining the line where the majority of the interceptions occur. In case of defensive tactics we consider only the defenders and midfielders and for the offensive scenario the midfielders and attackers. The method for determining the scenarios related to interceptions is:

Method : Interceptions

```
IF        interceptions at own half >> interceptions at opponents half
          Defensive tactics
          IF        interceptions at flanks  ≈ interceptions at centre
                    IF        interceptions by Defenders >> interceptions by Midfielders
                              [I.1] at least 1 Defender with interception skill
                    ELSE      interceptions by Defenders ≈ interceptions by Midfielders
                              [I.2] at least 1 Defender or Midfielder with interception skill

          ELSEIF    interceptions at flanks  << interceptions at centre
                    IF        interceptions by Defenders >> interceptions by Midfielders
                              [I.3] at least 1 Defender at centre with interception skill
                    ELSE      interceptions by Defenders ≈ interceptions by Midfielders
                              [I.4] at least 1 Defender or Midfielder at centre with interception skill

          ELSEIF    interceptions at flanks  >> interceptions at centre
                    IF        interceptions by Defenders >> interceptions by Midfielders
                              [I.5] at least 1 Defender at flank with interception skill
                    ELSE      interceptions by Defenders ≈ interceptions by Midfielders
                              [I.6] at least 1 Defender or Midfielder at flank with interception skill

ELSE
          Attacking tactics
          IF        interceptions at flanks  ≈ interceptions at centre
                    IF        interceptions by Midfielders >> interceptions by Attackers
                              [I.7] at least 1 Midfielder with interception skill
                    ELSE      interceptions by Midfielders ≈ interceptions by Attackers
                              [I.8] at least 1 Midfielder or Attacker with interception skill

          ELSEIF    interceptions at flanks  << interceptions at centre
                    IF        interceptions by Midfielders >> interceptions by Attackers
                              [I.9] at least 1 Midfielder at centre with interception skill
                    ELSE      interceptions by Midfielders ≈ interceptions by Attackers
                              [I.10] at least 1 Midfielder or Attacker at centre with interception skill

          ELSEIF    interceptions at flanks  >> interceptions at centre
                    IF        interceptions by Midfielders >> interceptions by Attackers
                              [I.11] at least 1 Midfielder at flank with interception skill
                    ELSE      interceptions by Midfielders ≈ interceptions by Attackers
                              [I.12] at least 1 Midfielder or Attacker at flank with interception skill
END
```

The last statistics group that will be examined are the passing statistics. As with the attacking statistics the distribution between long and short passes will be made. Only for the passes a distinction between short and long passes is based on the length of the passes, for the shots it depends on the location at the field. This distinction is made because teams which are giving a lot of long passes play in general defensively. The opposite is also true in general, teams playing with a lot of short passes try to dominate the game, which is offensive. Another indicator for offensive and defensive play style would be the distribution of the passes given at the own half and the opponents half. If the majority of the passes are given at the own half the team is playing defensively. The assumption we use here will be:

- **Assumption 3:** Team is playing with defensive tactics if the passes given at the own half is strictly larger than the passes given at the opponents half; otherwise team is playing with offensive tactics.

Besides an attacking or defending play style the build up location of the passes is of importance. With the build up location we mean the side of the field were the build up, the passes are given. This could be through the centre of the field or by using the flanks. To determine the side of the build up the distribution of the fraction of the passes performed at the flanks and centre will be compared using the same rule as in the previous methods. The final step is comparing the fraction of the long and short passes, in case of defensive tactics only the defenders will be considered and for offensive tactics only the midfielders. The attackers are not included in this method. No comparison is made between the different lines on the field. The final method to determine the scenarios based on the passing statistics is:

Method : Passes

```
IF          passes at own half >> passes at opponents half
            Defensive tactics
            IF          passes at flanks >> passes at centre
                        IF          long passes >> short passes
                                    [P.1] at least 1 Defender at flank with long pass skill
                        ELSEIF      long passes << short passes
                                    [P.2] at least 1 Defender at flank with short pass skill
                        ELSEIF      long passes ≈ short passes
                                    [P.3] at least 1 defender at flank with short or long pass skill
            ELSEIF      passes at flanks << passes at centre
                        IF          long passes >> short passes
                                    [P.4] at least 1 Defender at centre with long pass skill
                        ELSEIF      long passes << short passes
                                    [P.5] at least 1 Defender at centre with short pass skill
                        ELSEIF      long passes ≈ short passes
                                    [P.6] at least 1 Defender at centre with short or long pass skill
            ELSEIF      passes at flanks ≈ passes at centre
                        IF          long passes >> short passes
                                    [P.7] at least 1 Defender with long pass skill
                        ELSEIF      long passes << short passes
                                    [P.8] at least 1 Defender with short pass skill
                        ELSEIF      long passes ≈ short passes
                                    [P.9] at least 1 Defender with short or long pass skill
ELSE
            Attacking tactics
            IF          passes at flanks >> passes at centre
                        IF          long passes >> short passes
                                    [P.10] at least 1 Midfielder at flank with long pass skill
                        ELSEIF      long passes << short passes
                                    [P.11] at least 1 Midfielder at flank with short pass skill
                        ELSEIF      long passes ≈ short passes
                                    [P.12] at least 1 Midfielder at flank with short or long pass skill
```

```
        ELSEIF  passes at flanks << passes at centre
                IF      long passes >> short passes
                                [P.13] at least 1 Midfielder at centre with long pass skill
                ELSEIF  long passes << short passes
                                [P.14] at least 1 Midfielder at centre with short pass skill
                ELSEIF  long passes ≈ short passes
                                [P.15] at least 1 Midfielder at centre with short or long pass skill
        ELSEIF  passes at flanks ≈ passes at centre
                IF      long passes >> short passes
                                [P.16] at least 1 Midfielder with long pass skill
                ELSEIF  long passes << short passes
                                [P.17] at least 1 Midfielder with short pass skill
                ELSEIF  long passes ≈ short passes
                                [P.18] at least 1 Midfielder with short or long pass skill
END
```

The 4 methods described above will provide us additional information based on the team tactics. With the outcomes of the 4 methods the team profile could be created based on attacking, defending and passing statistics.

5.3 MATCHING PROFILES

In the final part of this chapter we will explain how the player, position and team profile will be used to optimize the line-up. Before using the profiles to optimize the line-up the grading of the players on different positions will be explained. After the grading part has been explained the final sub-question stated below will be answered.

- *Which IP model do we use to match the profiles in order to optimize the line-up?*

5.3.1 GRADING PLAYERS ON POSITIONS

The function that will be used to grade the players on the position will be the grading function used in the Coach & Scout Assistant. The reason that this function will be used is that the rules the CSA grading function satisfy are exactly the same rules needed in this thesis. Before describing the rules we will explain the variables used in the function.

I	:	set of players
J	:	set of positions
Q	:	set of skills
i	:	indices for player
j	:	indices for actions
q	:	indices for skill
S_{iq}	:	score of player i for skill q
N_{jq}	:	demanded score at position j for skill q
N_{jq}^{α}	:	weight of skill q at position j

The rules the function $f(S_{iq}, N_{jq}) = \; ln\left(\frac{S_{iq}}{N_{jq}}\right)$ is satisfying are:

1. A shortage to a quality of a player results in a negative score
   if $S_{iq} < \; N_{jq}$ then $f(S_{iq}, N_{jq}) < 0$
2. A surplus to a quality of a player results in a positive score
   if $S_{iq} > \; N_{jq}$ then $f(S_{iq}, N_{jq}) > 0$
3. If the score of a player for a quality is equal to the demand for the quality at the position the score results in 0.
   if $S_{iq} = \; N_{jq}$ then $f(S_{iq}, N_{jq}) = 0$
4. Shortages of qualities are punished increasingly.
   $\frac{\delta^2 f(S_{iq}, N_{jq})}{\delta S_{iq}^2} < 0$
5. Surpluses of qualities are rewarded decreasingly.
   $\frac{\delta^2 f(S_{iq}, N_{jq})}{\delta S_{iq}^2} < 0$

The 4th and 5th rule are based on the second derivative of the function $f(S_{iq}, N_{jq})$, the function is concave because the second derivative is less than zero. This means that the shortages of the qualities are punished increasingly while the surpluses are rewarded decreasingly.

The function $ln\left(\frac{S_{iq}}{N_{jq}}\right)$ is used to calculate the scores for each skill of a player at a given position. To calculate the position score of player *i* at position *j* (*PSij*) the weight of skill *q* at position *j* (N_{jq}^{α}) is needed. Now the position score can be calculated by taking first the sum of all scores of the player for the position multiplied with the corresponding weight N_{jq}^{α}

$$\sum_{q \in Q} N_{jq}^{\alpha} \cdot ln\left(\frac{S_{iq}}{N_{jq}}\right)$$

and dividing the sum with the sum of all weights for the skills

$$\sum\nolimits_{q \in Q} N_{jq}^{\alpha}$$

This will be the weighted average over all skills of a player *i* for position *j,* the position score. The final formula for the calculation of the position scores would be:

$$PS_{ij} \; = \frac{1}{\sum_{q \in Q} N_{jq}^{\alpha}} \sum_{q \in Q} N_{jq}^{\alpha} \cdot ln\left(\frac{S_{iq}}{N_{jq}}\right)$$

To interpret the position scores calculated easier, the scores will be represented in percentages. This can be done by taking the exponent of the score and multiplying it with 100. The new score will be called the New Player Score (NPS_{ij}) of player *i* at position *j*.

$$NPS_{ij} = e^{PS_{ij}} \cdot 100 = exp\left[\frac{1}{\sum_{q\in Q} N_{jq}^{\alpha}} \sum_{q\in Q} N_{jq}^{\alpha} \cdot ln\left(\frac{S_{iq}}{N_{jq}}\right)\right] \cdot 100$$

5.3.2 INTEGER PROGRAMMING MODEL

The determination of the line-up in soccer can be seen as a matching problem, matching the supply of players with the demand at the positions. With a basic integer programming model the optimal teams can be calculated. The application Coach & Scout Assistant uses such a model to determine the optimal line-up. This model is based on the player and position tables which are similar to the player and position profiles defined in this thesis. We have an additional profile, the team profile which is included to add the teamwork aspect to the determination of the line-up. The team profile is created in such way that by adding additional restrictions to the basic integer programming model it will be able to calculate the optimal line-up with the team profile included. The variables used in the model are:

I	:	set of players
J	:	set of positions
i	:	indices for player
j	:	indices for actions
PS_{ij}	:	position score of player *i* at position *j*
$grade_j$	:	grade required at position *j*

and decision variables

$$x_{ij} = \begin{cases} 1 & \text{if player } i \text{ is at position } j \\ 0 & \text{if player } i \text{ is not at position } j \end{cases}$$

$$y_n = \begin{cases} 1 & \text{if scenario } n \text{ holds} \\ 0 & \text{if scenario } n \text{ does not hold} \end{cases}$$

The basic integer programming model is:

$$\max \sum_{i\in I} \sum_{j\in J} \left(PS_{ij} \cdot x_{ij}\right) \qquad (1)$$

subject to

$$\sum_{i\in I} x_{ij} = 1 \quad \forall j \in J \qquad (2)$$

$$\sum_{j\in J} x_{ij} \leq 1 \quad \forall i \in I \qquad (3)$$

$$x_{ij} \in \{0, 1\} \quad \forall i \in I, j \in J$$

The objective function (1) chooses from all feasible combination the line-up with the highest player scores. Constraint (2) ensures that for each position exactly one player is assigned. While in constraint (3) each player is assigned to at most 1 position.

This is how the model looks like when the team profile is not included. The team profile described in chapter 5.2.3 contains different scenarios for the shots, crosses, interceptions and passes. This means that for each group of statistics additional constraints are needed for the model. These constraints are dependent of the tactics of the team. For each group several scenarios are possible; 9 scenarios for the shots, 13 for the crosses, 12 for the interceptions and 18 for the passes. Each scenario results in a different set of constraint so these constraints must be defined first. Because we need the skills q of the player i in the model the variable S_{iq} will be included in the model.

The purpose of the constraints is that it must ensure that at least one player with the required skill is selected in the line-up at one of the positions that are of importance for the scenario. The number of constraints we need for a scenario depends on the number of skills q and positions j; this is because for a scenario at least 1 of the $N = (j \cdot q)$ constraints must hold. Depending on the number of skills times the positions, we need 3 different sets of constraints. If N is equal to 1 then the general form of the constraints is:

$$\sum_{i \in I} \sum_{q \in Skill} S_{iq} \cdot x_{ij} \geq grade_j \qquad \forall\, j \in Position$$

If N is equal to the 2 then we need that at least 1 of the 2 constraints holds, this can be done by adding an extra binary variably y to the constraint. The following 2 constraints ensure that at least one of the 2 constraints holds.

$$\sum_{i \in I} \sum_{q \in Skill} S_{iq} \cdot x_{ij} \geq grade_j \cdot y \qquad \forall\, j \in Position_1$$

$$\sum_{i \in I} \sum_{q \in Skill} S_{iq} \cdot x_{ij} \geq grade_j \cdot (1 - y) \qquad \forall\, j \in Position_2$$

The final possibility is that N is larger than 2, then we need $N + 1$ constraints so that 1 out of the N constraints hold. In this case we need to add N binary variables to the set of constraints which will ensure that at least one of the constraints holds. Then the sets of constraints needed are:

$$\sum_{i \in I} \sum_{q \in Skill} S_{iq} \cdot x_{ij} \geq grade_j \cdot y_1 \qquad \forall\, j \in Position_1$$

$$\cdots$$

$$\sum_{i \in I} \sum_{q \in Skill} S_{iq} \cdot x_{ij} \geq grade_j \cdot y_N \qquad \forall\, j \in Position_n$$

$$\sum_{n=1}^{N} y_n = 1$$

The final constraint here ensures that 1 of the constraints holds, the grades of the other N-1 will be set to 0 because y_n is a binary variable.

Further it is important to notice that the index of the skills q and positions j in these constraints is different than before. It depends on the scenario. This can be explained with a small example. Suppose that the crossing statistics of the team results in scenario C.2; which needs at least 1 midfield player with 7 as grade for the crossing skill. Here j = 3 (MCR, MC and MCL) q = 1 so we need 4 constraints for this scenario. The constraint corresponding to scenario C.1 will be:

$$\sum_{i \in I} \sum_{q \in Crossing} S_{iq} \cdot x_{ij} \geq 7 \cdot y_1 \qquad \forall\, j \in MCR$$

$$\sum_{i \in I} \sum_{q \in Crossing} S_{iq} \cdot x_{ij} \geq 7 \cdot y_2 \qquad \forall\, j \in MC$$

$$\sum_{i \in I} \sum_{q \in Crossing} S_{iq} \cdot x_{ij} \geq 7 \cdot y_3 \qquad \forall\, j \in MCL$$

$$\sum_{n=1}^{3} y_n = 1$$

With the general form of the constraints for the team profile all possible scenarios could be included to the model. As explained before the set constraints per scenario that will be added to the model depend on the number of skills and positions included in the scenario. The basic form of all constraints per scenario are listed in Appendix B. Based on the basic form the set of constraints could be easily created. If for all scenario's the number of skills and positions is equal to 1 (N = 1) the integer programming model will be:

$$\max \sum_{i \in I} \sum_{j \in J} \left(PS_{ij} \cdot x_{ij}\right) \qquad (1)$$

subject to

$$\textstyle\sum_{i \in I} x_{ij} = 1 \qquad \forall\, j \in J \qquad (2)$$

$$\textstyle\sum_{j \in J} x_{ij} \leq 1 \qquad \forall\, i \in I \qquad (3)$$

$$\textstyle\sum_{i \in I} \sum_{q \,\in\, \text{Attacking Skill 1}} S_{iq} \cdot x_{ij} \geq grade_j \qquad \forall\, j \in Positions \qquad (4)$$

$$\textstyle\sum_{i \in I} \sum_{q \,\in\, \text{Attacking Skill 2}} S_{iq} \cdot x_{ij} \geq grade_j \qquad \forall\, j \in Positions \qquad (5)$$

$$\textstyle\sum_{i \in I} \sum_{q \,\in\, \text{Defending Skill}} S_{iq} \cdot x_{ij} \geq grade_j \qquad \forall\, j \in Positions \qquad (6)$$

$$\textstyle\sum_{i \in I} \sum_{q \,\in\, \text{Passing Skill}} S_{iq} \cdot x_{ij} \geq grade_j \qquad \forall\, j \in Positions \qquad (7)$$

$$x_{ij} \in \{0, 1\} \qquad \forall\, i \in I, j \in J$$

6 RESULTS

In this chapter the result of the methods described in the previous chapter will be shown. The team that will be used is the Dutch soccer club Ajax in season 2010-2011. First the results of the sensibility analysis related to the player grades will be shown. The second part of this chapter will show the player, position and team profile of Ajax which will be used as input for the integer programming model. The final part will show the optimal line-up obtained by our model. We will also compare our line-up with the line-up obtained from basic IP model without team profile and with the line-up preferred by the coach.

6.1 SENSIBILITY ANALYSIS PLAYER GRADES

In the original data the likert grade with value 3 is used for neutral actions and the neutral actions do not increase the grade of a player's skill if the player has performed a lot of neutral actions. To grade the quality we treat the neutral actions as successful actions and therefore it must increase the grade. To determine the increase for value of the neutral likert grade the next 4 test cases are used. The values 3.05, 3.1, 3.15 and 3.2 will be compared with the usual value of 3. In each case we start with 10 neutral actions and increase them with 10 till the total number of neutral actions reaches 350. The difference in each case is that we add 3 very bad, bad, good and very good actions to the test case. This means that for a test case the neutral actions vary and the other actions remain the same. As comparison we have also a case were no other action has been added. Figure 6.1 shows us the graph were 3 very bad passes have been added to the neutral actions. Each line in the graph represents a different value for the neutral actions.

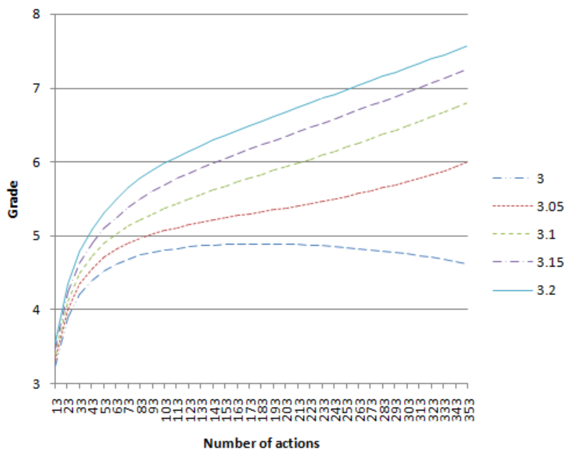


Figure 6.1 - 3 very bad passes given

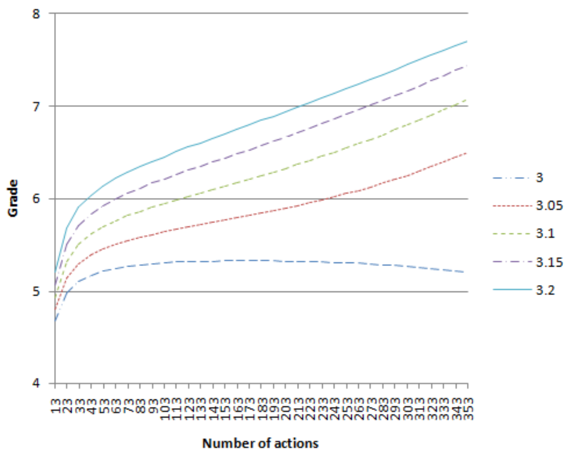


Figure 6.2 - 3 bad passes given

The figures above show us the unsuccessful actions, we see in both figures that the grade at the beginning are low for each value of the neutral actions and starts to increase rapidly with the neutral actions increasing. We see that increasing the neutral value of 3 contributes in a higher grade if the neutral actions are increasing, this does not happen for the original value of 3. By analyzing these 2 figures the most suitable value for the neutral actions will be 3.05. This is because the other values results in too high values, which will overestimate the qualities of the players.

The graphs of the successful action are placed in figure 6.3, 6.4 and 6.5. In figure 6.5 the graph with only neutral actions have been placed. Again we see that increasing the likert value gives us the expected result; a increase in neutral actions increases the grade. The bottom line in figure 6.5 makes it clear that for the original value this was not the case. In figure 6.3 and 6.4 the good and very good actions have been added to the neutral actions. Here we see that the grades are very high in the beginning and start to decrease rapidly. Because the weight of the grades for the good actions is higher than the other likert grades (see chapter 4.3) the grades with less neutral actions are very high. Especially in figure 6.4 we notice that the grades in the beginning of the graph are the highest in each line. This looks a bit strange but it is because we have compared neutral actions with 3 very good passes. It is very uncommon that a similar situation will occur in the data.

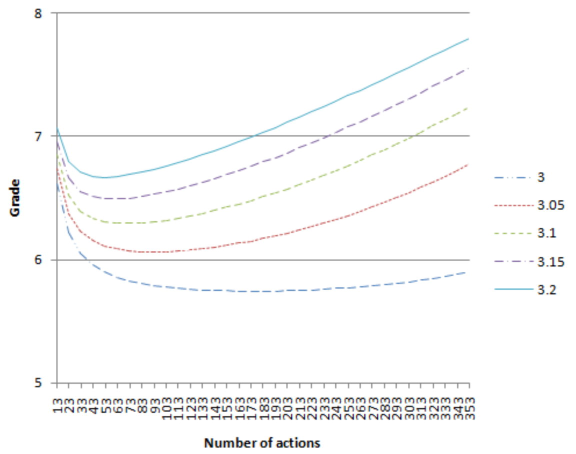


Figure 6.3 - 3 good passes given

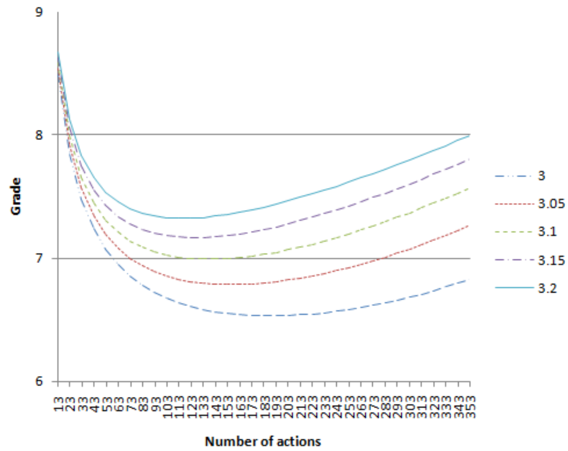


Figure 6.4 - 3 very good passes given

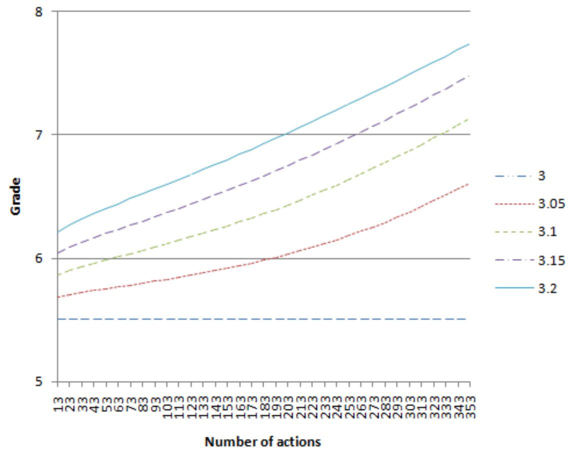


Figure 6.5 - only neutral passes given

With the use of the graphs above the value of the 'neutral' actions have been set to 3.05 and it will be used in the remainder of the thesis. Choosing a higher value would contribute in too high grades for the players skills with less actions. It would not be possible for us to distinguish player who have performed a lot of actions with players with less actions by the grade. Our purpose was to translate the neutral actions into successful action which would increase the grade slightly and this can be done with a value of 3.05.

6.2 PROFILES OF AJAX

With the likert value chosen in the previous part the profiles for Ajax can be graded given the data. First the player and position profile of Ajax will be shown and discussed, with these two profiles the player position scores can be calculated which will be used as input for the integer programming model. Also the impact of the choices made for the player and position profile at the position scores of the players will be discussed. The last part will show the team profile, this profile will decide which additional sets of constraints will be added to our model.

6.2.1 PLAYER PROFILES OF AJAX

Table 6.1 shows us the player skill grades for the 19 soccer skills we have defined. This table contains only 18 players of Ajax. In total we have grades 29 players of Ajax but 11 players played too less to receive a grade. The table containing all players can be found in Appendix C, table C.1. A lot of grades at table 6.1 have been set to 0.1. If a player has never performed the skill the grade is set to 0.1. This is obvious for a goal keeper but the data contains also players who have never played a match. Skills performed by the players which are less than the average of that skill have also been set to 0.1. This is done to prevent that the skill will be overestimated. If on average a player has given 20 short passes in 40 matches, all players with less than 20 short passes have been graded with 0.1. Otherwise a few good passes would lead to a high short passing grade which is not realistic if we want to grade the quality of a players skill.

Player Profile																			
Player	Short pass back	Short pass forward	Short pass wide	Long pass back	Long pass forward	Long pass wide	Cross	Heading	Finishing	Long shot	Dribbles	Reception	Interceptions	Goalkeeping	Goal kick	Set pieces	Short pass	Long Passes	Pressing
Vertonghen	4.9	8.6	7.9	4.2	8.3	5.7	0.1	7.3	7.5	5.0	6.9	0.1	8.7	0.1	0.1	5.6	8.3	8.5	0.1
Lindgren	0.1	0.1	0.1	0.1	0.1	0.1	0.1	6.8	0.1	6.6	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Emanuelson	3.8	0.1	0.1	4.0	0.1	4.6	6.5	0.1	0.1	0.1	7.6	0.1	8.4	0.1	0.1	6.0	0.1	4.2	6.0
Vermeer	0.1	0.1	0.1	0.1	3.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	9.0	5.5	0.1	0.1	0.1	0.1
de Jong	5.1	8.1	7.1	3.4	0.1	0.1	4.5	4.4	8.5	4.6	7.2	5.6	8.2	0.1	0.1	0.1	8.7	3.9	6.7
Zeeuw	4.5	8.5	7.0	2.9	7.8	5.6	4.0	6.1	7.6	5.8	0.1	0.1	8.3	0.1	0.1	6.8	8.4	5.6	6.7
El Hamdaoui	3.9	0.1	7.4	2.7	0.1	0.1	6.1	0.1	7.9	3.7	7.0	4.4	0.1	0.1	0.1	0.1	8.7	0.1	4.0
Verhoeven	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	8.4	5.3	0.1	0.1	0.1	0.1
Alderweireld	4.6	7.9	7.7	4.3	8.4	5.8	0.1	7.5	0.1	7.1	6.4	0.1	7.4	0.1	0.1	5.9	7.5	8.4	0.1
Blind	4.1	8.4	0.1	3.9	0.1	0.1	4.1	0.1	0.1	0.1	0.1	0.1	8.5	0.1	0.1	0.1	8.0	0.1	0.1
Van der Wiel	6.2	8.3	8.1	4.1	6.8	5.7	4.9	6.4	0.1	0.1	6.6	4.5	8.3	0.1	0.1	5.8	8.1	4.9	5.4
Anita	5.1	8.4	7.3	3.9	6.4	5.3	4.3	6.2	0.1	0.1	6.9	0.1	8.7	0.1	0.1	0.1	8.1	4.5	0.1
Enoh	5.0	6.9	7.2	3.7	7.2	5.3	0.1	6.4	0.1	0.1	0.1	0.1	8.8	0.1	0.1	0.1	7.5	6.3	0.1
Eriksen	5.2	8.6	7.2	3.1	0.1	6.1	4.1	6.7	8.4	4.0	7.3	6.1	8.0	0.1	0.1	5.0	8.6	5.4	6.3
Stekelenburg	0.1	4.8	0.1	0.1	8.4	5.4	0.1	0.1	0.1	0.1	0.1	0.1	0.1	8.8	6.4	5.1	0.1	8.3	0.1
Suarez	0.1	0.1	0.1	3.1	0.1	0.1	7.2	6.0	8.2	3.9	0.1	7.1	0.1	0.1	0.1	4.7	0.1	0.1	8.6
Sulejmani	3.8	7.6	6.6	3.3	0.1	0.1	5.7	5.7	7.7	5.2	7.5	6.5	8.0	0.1	0.1	6.0	8.3	3.9	7.3
Ebecilio	3.8	0.1	6.8	0.1	0.1	0.1	6.7	5.2	6.1	7.1	6.7	4.5	0.1	0.1	0.1	0.1	8.1	0.1	8.0

Table 6.1 - Player profile of AJAX with the scores of the players skills

The value of 0.1 is chosen for calculation reason (see section 5.3.1). Setting this value to 1 or 0 might look usual but it has an undesirable effect on the calculation of the position scores of the players. The effect of the values 0 and 1 will be discussed in the section 6.2.3.

6.2.2 POSITION PROFILES OF AJAX

The position profile consists of two parts, the grades required for the skills at a position and the weights for the skills. In table 6.2 the grades of the skills for the positions are listed. If a skill has not been performed at a position the grade has been set to 1. This is done again for calculation reasons. These grades are the denominators for the grading function so it is not possible to set them to 0. In the player profile skills which are performed less than the average count of the skill were set to a fixed number, this is not done for the position profile because we use also the weights for the position profile.

Grades Position Profile

Player	Short pass back	Short pass forward	Short pass wide	Long pass back	Long pass forward	Long pass wide	Cross	Heading	Finishing	Long shot	Dribbles	Reception	Interceptions	Goalkeeping	Goal kick	Set pieces	Short pass	Long Passes	Pressing
GK	4.4	4.3	4.6	5.8	8.2	5.5	1.0	9.2	1.0	1.0	5.8	1.0	7.5	8.9	6.2	5.2	3.1	8.1	7.9
DR	6.3	8.2	8.1	4.1	6.7	5.7	4.9	6.2	3.7	4.1	6.8	4.9	8.4	1.0	1.0	5.8	8.0	4.9	5.7
DCR	4.6	7.9	7.6	4.3	8.3	5.2	4.5	7.9	3.7	7.0	6.3	3.0	8.0	1.0	1.0	6.0	7.4	8.3	6.4
DCL	5.0	8.6	7.9	4.2	8.3	5.8	6.6	6.9	7.5	5.0	6.9	4.7	8.7	1.0	1.0	5.6	8.3	8.6	7.3
DL	4.0	8.4	7.0	3.7	7.4	4.4	4.8	6.5	7.5	5.3	7.5	7.2	8.3	1.0	1.0	7.0	7.8	5.0	6.3
MCR	4.4	8.5	7.7	3.1	7.6	5.3	4.6	6.6	8.2	4.5	6.4	4.4	8.8	1.0	2.5	6.8	8.3	4.6	5.5
MC	4.5	8.8	7.9	3.6	7.2	5.4	3.6	5.7	7.2	4.2	7.2	5.6	8.6	1.0	1.0	4.9	8.5	4.6	6.1
MCL	4.4	8.8	7.6	3.0	7.6	5.4	4.9	6.4	9.0	4.3	7.1	6.4	8.8	1.0	1.0	6.5	8.4	5.0	5.4
FR	3.0	7.9	6.5	3.1	6.9	4.8	5.2	5.8	7.4	5.0	7.2	4.9	8.3	1.0	1.0	3.9	8.2	3.6	7.5
FC	5.9	8.3	8.0	2.6	7.1	5.6	4.5	4.3	8.1	4.7	6.9	5.4	7.8	1.0	2.5	7.2	8.8	3.3	4.2
FL	3.6	8.0	7.2	3.2	7.0	5.0	6.6	5.0	7.2	6.5	6.9	5.4	8.3	1.0	1.0	5.3	8.3	4.1	8.4

Table 6.2 - Grades at the position for every skill

Weights Position Profile

Player	Short pass back	Short pass forward	Short pass wide	Long pass back	Long pass forward	Long pass wide	Cross	Heading	Finishing	Long shot	Dribbles	Reception	Interceptions	Goalkeeping	Goal kick	Set pieces	Short pass	Long Passes	Pressing
GK	0	.03	0	0	.11	0	0	0	0	0	0	0	0	1.69	1.64	0	0	.07	0
DR	.23	.32	.32	.21	.12	.16	.49	.21	0	.06	.18	.10	.48	0	0	.07	.46	.20	.04
DCR	.04	.45	.41	.04	.31	.64	0	.36	0	0	.11	0	.22	0	0	.05	.51	.55	0
DCL	0	.47	.43	0	.30	.45	0	.38	0	.04	.15	0	.26	0	0	.05	.51	.47	.05
DL	.12	.27	.20	.14	.23	.14	.22	.15	.05	.05	.24	0	.42	0	0	.09	.33	.18	.08
MCR	.26	.21	.21	.33	.12	.16	.09	.19	.19	.28	.15	.18	.36	0	0	.12	.31	.17	.10
MC	.14	.24	.28	.18	.05	.17	.03	.20	.14	.14	.16	.23	.11	0	0	.12	.39	.18	.25
MCL	.23	.13	.16	.23	.15	.10	.06	.11	.11	.18	.27	.08	.25	0	0	.19	.22	.11	.15
FR	.40	.07	.09	.21	0	.05	.42	.07	.30	.36	.11	.45	.12	0	0	.52	.15	.06	.29
FC	.22	.10	.15	.12	0	.06	.12	.23	.36	.19	.04	.53	0	0	0	.05	.24	.09	.58
FL	.29	.04	.05	.12	.05	0	.41	.04	.53	.30	.18	.24	.09	0	0	.17	.08	0	.46

Table 6.3 - Weights of the skills at the positions

The weights corresponding to the grades are listed in table 6.3. The weights used for determining the final weights of the skills are $[w_3, w_2, w_1, v_3, v_2, v_1] = [3, 2, 1, 2, 1, 0]$. The first 3 weights corresponds to ranked weights given for the horizontal lines and the last 3 to the ranked weights given to the vertical lines. By using these values a fraction will be multiplied with at most $w_3 + v_3 = 5$. The impact of using different weights will be discussed in chapter 6.3.2 were the player position scores will be shown. In table 6.3 we see some interesting differences between the weights for skills at the flanks. For the skills short passes back, long passes back, cross, finishing, reception and set pieces the difference of the weights at the left and right position differ. The expectation for these skills was that the weights at the flanks should be almost equal, but the differences are a result of the multiplication of the fractions we performed. In Appendix C, table C.2 the values with which fraction of a skill is multiplied is placed. Here we see that the fractions at the flanks are often not multiplied with the same value which resulted in the differences at the flanks in table 6.3. Further we see a lot of zeros, these weights are set to 0 because the fraction of the grades was less than 3% (see Chapter 5.2.2) so we assumed that they are not important for the position. In Appendix C, table C.3 and C.4 contains the weights of the position scores with the values $[w_3, w_2, w_1, v_3, v_2, v_1] = [5, 3, 1, 5, 2, 0]$ and $[w_3, w_2, w_1, v_3, v_2, v_1] = [10, 5, 1, 10, 5, 0]$.

6.2.3 PLAYER POSITION SCORES OF AJAX

With the player and position profiles we can calculate the player position scores. In chapter 5 we explained the function to calculate these scores. The function is:

$$NPS_{ij} = e^{PS_{ij}} \cdot 100 = exp\left[\frac{1}{\sum_{q\in Q} N_{jq}^{\alpha}} \sum_{q\in Q} N_{jq}^{\alpha} \cdot ln\left(\frac{S_{iq}}{N_{jq}}\right)\right] \cdot 100$$

with S_{iq} the grades of the players i for skill q, N_{jq} the grades needed at position j for skill q and N_{jq}^{α} the weight at position j for skill q. The player position scores of the 18 players of Ajax are listed in table 6.4; the complete table with all 29 players is listed in Appendix C, table C.5.

Player Position Scores											
Player	GK	DR	DCR	DCL	DL	MCR	MC	MCL	FR	FC	FL
Vertonghen	1.8	53.5	105.5	94.4	72.5	72.3	55.3	69.1	32.8	23.3	21.9
Lindgren	1.3	2.1	2.1	2.1	2.0	3.0	2.6	2.5	3.3	3.3	2.5
Emanuelson	1.4	15.2	8.1	6.4	15.2	12.8	9.4	17.1	21.4	9.4	15.5
Vermeer	81.3	1.7	1.9	1.7	2.0	1.9	1.7	2.0	2.0	1.8	1.7
de Jong	1.5	63.4	28.3	32.8	50.8	62.2	62.2	50.7	59.9	95.3	71.2
Zeeuw	1.7	69.9	82.7	76.1	70.9	69.0	60.6	59.2	61.5	53.0	53.3
El Hamdaoui	1.3	15.8	5.1	5.5	11.6	18.6	21.1	18.7	37.5	46.3	49.9
Verhoeven	69.8	1.5	1.4	1.3	1.5	1.7	1.6	1.6	2.0	1.8	1.6
Alderweireld	1.8	51.2	101.6	91.0	64.0	56.5	44.5	56.9	23.4	14.2	10.5
Blind	1.4	18.3	6.2	6.1	13.3	10.8	7.4	8.8	8.5	5.2	5.8
Van der Wiel	1.7	93.5	93.3	84.8	88.3	60.5	69.4	66.4	53.7	51.3	29.6
Anita	1.7	71.3	84.5	73.4	67.5	37.1	30.8	34.9	13.3	11.5	9.2
Enoh	1.7	34.9	76.1	62.7	35.9	27.7	23.6	21.1	7.6	9.4	4.3
Eriksen	1.5	83.6	66.8	63.6	71.5	88.4	95.5	78.9	109.5	113.9	86.4
Stekelenburg	101.1	4.1	13.2	9.8	5.6	4.3	4.6	4.5	4.2	2.8	2.2
Suarez	1.3	5.7	2.3	2.4	4.5	9.2	9.2	8.1	36.7	35.5	36.2
Sulejmani	1.5	69.3	29.9	34.8	57.9	71.6	74.1	66.8	109.2	105.1	91.6
Ebecilio	1.3	16.7	7.1	8.3	12.2	17.7	23.8	17.5	37.3	62.6	52.7

Table 6.4 - Position Scores of the players for every position

In table 6.4 we see some remarkable scores for 5 players. These are scores which are not expected based on the usual positions of the players. First we see that Vertonghen a defender of origin has high scores for the position MCR and MCL, Zeeuw which is a midfielder of origin has higher scores as defender, Van der Wiel a defender right of origin has high scores as a midfielder, Eriksen an attacking midfielder which has high scores as defender at the flanks and finally Sulejmani an attacker of origin with high scores as midfielder and defender right. All these strange scores for the players can be caused by the fact that the player position scores are based only on the soccer skills we defined. Besides the soccer skills we also need the physical skills of the players.

In sections 6.2.1 and 6.2.2 we mentioned some choices made for the player and position profile. The first choice was that the empty skills for players are set to 0.1. Another possibility for this was setting the value to 0 which would be logical. The reason that we cannot set the empty skill scores to 0 is that it cannot be used in the grading formula. The other values we mentioned are 1 and 0.5. The reason we did not use these values is that it resulted in too high player position scores on all positions. The result of these values are listed in table C.6 and C.7. To compare the effect of the values we have selected 3 players for which we calculated the player position scores. In table 6.5 we placed the player position scores of 3 players for both values. Here we see setting the empty skills to 1 or 0.5 result in too high scores. With too high scores we mean that a player becomes suitable at almost any position.

Comparison Player Position Scores

Player	PS_{ij}	GK	DR	DCR	DCL	DL	MCR	MC	MCL	FR	FC	FL
Alderweireld	Original	1.8	51.2	101.6	91.0	64.0	56.5	44.5	56.9	23.4	14.2	10.5
	Empty skills 1	15.2	76.3	101.6	94.1	84.8	82.0	72.9	80.2	58.6	46.0	36.4
	Empty skills 0.5	7.9	67.6	101.6	93.1	77.9	73.3	62.8	72.3	44.5	32.3	25.0
	Weights 1	1.5	49.6	97.6	90.6	64.3	57.3	46.2	56.2	23.7	14.4	9.5
	Weights 2	1.4	52.8	98.1	89.0	64.9	58.6	46.0	56.3	23.0	13.0	8.9
Ebecilio	Original	1.3	16.7	7.1	8.3	12.2	17.7	23.8	17.5	37.3	62.6	52.7
	Empty skills 1	13.4	44.6	29.0	30.2	39.0	47.4	53.5	46.6	71.0	86.3	74.5
	Empty skills 0.5	6.7	33.2	18.9	20.5	27.5	35.2	41.9	34.7	58.5	78.4	67.2
	Weights 1	1.3	16.3	7.5	8.5	12.0	17.3	23.8	17.6	36.2	63.4	54.5
	Weights 2	1.3	15.8	7.4	8.4	11.6	17.2	23.8	17.4	37.0	65.7	55.2
Vertonghen	Original	1.8	53.5	105.5	94.4	72.5	72.3	55.3	69.1	32.8	23.3	21.9
	Empty skills 1	15.3	79.7	105.5	97.6	92.0	92.5	81.3	88.6	68.0	57.9	50.9
	Empty skills 0.5	8.0	70.7	105.5	96.7	85.7	85.9	72.4	82.2	54.6	44.0	39.5
	Weights 1	1.5	51.8	101.2	94.2	73.9	71.7	55.6	70.6	32.3	22.8	22.0
	Weights 2	1.4	55.2	101.8	92.6	74.1	73.8	54.6	70.5	32.2	20.7	20.5

Table 6.5 - Comparison of Player Position Scores

The second choice we made was the value for the weights [w_3, w_2, w_1, v_3, v_2, v_1]. The results of 2 different values for the weights have also be placed in table 6.5. Weight 1 corresponds to values [w_3, w_2, w_1, v_3, v_2, v_1] = [5, 3, 1, 5, 2, 0] and weight 2 to [10, 5, 1, 10, 5, 0]. For the different values for the weights we see that it lowers almost all position scores for the players, only some scores are a bit higher. The differences with the original values are negligible.

As a final comparison we calculated the mean and standard deviation of all player position scores for the different choices. These statistics are placed in table 6.6. With the original calculation of the player position scores we mean the choices we use in the remainder of the thesis. These choices are that the empty skill grades are set to 0.1 and the weights used for [w_3, w_2, w_1, v_3, v_2, v_1] are [3, 2, 1, 2, 1, 0]. We see here that the standard deviation does not vary a lot for the different choices. The mean however is higher if we set the empty skills to 1 or 0.5. The difference of the mean for the different values of the weights, compared with the mean of the original model is again negligible.

Calculation of PSij	Mean	Standard Deviation
Original	21.30	29.98
Empty skills set to 1	38.52	29.37
Empty skills set to 0.5	30.56	30.68
Weight set to [5, 3, 1, 5, 2, 0]	21.38	30.10
Weight set to [10, 5, 1, 10, 5, 0]	21.48	30.32

Table 6.6 - Mean and Standard Deviation of the PSij

6.2.4 TEAM PROFILE OF AJAX

With the results of the team profile we will be able to select the constraints corresponding to the scenario's defined. In table 6.7 the statistics needed to determine the 4 scenario's for Ajax are placed.

Team statistics of Ajax	Short Passing	Long Passing	Passing	Crossing	Finishing	Long shots	Shots	Interceptions
Mean	1175	370	1545	56	24	31	132	516
Defence	1442	495	1937	59	6	18	143	714
Midfield	1256	302	1558	34	26	41	134	548
Attack	977	179	1156	94	53	49	159	297
Defence centre	1539	683	2222	14	7	23	164	708
Midfield centre	1293	285	1578	36	36	47	157	422
Attack centre	879	174	1053	44	59	40	121	196
Defence flank	1346	306	1652	105	5	14	122	720
Midfield flank	1238	311	1548	33	21	39	123	612
Attack flank	1027	182	1208	120	51	54	178	348
Own half	-	-	209	-	-	-	-	113
Opponents half	-	-	223	-	-	-	-	89
Flanks	-	-	152	-	-	-	-	81
Centre	-	-	280	-	-	-	-	121

Table 6.7 - Team statistics of Ajax needed to determine scenario's

With the statistics listed in table 6.7 the decision trees created in chapter 5.2.3 have resulted in the next 4 scenario's.

- [S.3] at least 1 *Midfielder* or *Attacker* with *finishing* or *long shot* skill
- [C.7] at least 1 *Attacker at flank* crossing skill
- [I.3] at least 1 *Defender at centre* with *interception* skill
- [P.17] at least 1 *Midfielder at centre* with *short pass* skill

Before defining the sets of constraints needed per scenario we need to determine the grades of the skills per scenario. The grades for the skills finishing or long shot, crossing, interceptions and short passes are needed to complete the constraints. For each skill the average grade of the skill performed in all 40 matches of Ajax will be used. The average grades of the skills over the 40 matches are:

Finishing	:	6.7
Long shot	:	5.5
Crossing	:	5.5
Interceptions	:	8.0
Short passing	:	8.2

It is possible that the grades determined for the skills are too high, this can result in an infeasible solution of the IP model. If this occurs the grade that is too high will be set to 5.5. Another possibility is that the high grade will result into assigning players to positions where they do not fit in but have been assigned only because of the high grade in the constraint. If this occurs the grade will be also set to 5.5.

For the first scenario, [S.3] 'at least 1 *Midfielder* or *Attacker* with *finishing* or *long shot* skill' we have 6 possible positions and 2 skills. The number of constraints needed for this scenario is N =6·2 = 12, plus an additional constraints to ensure that one of the constraints holds so in total 13 constraints are needed.

$$\sum_{i\in I}\sum_{q\in Finishing} S_{iq}\cdot x_{ij} \geq 6.7\cdot y_1 \quad ,\forall j \in MCR \quad (1)$$

$$\sum_{i\in I}\sum_{q\in Long\ Shot} S_{iq}\cdot x_{ij} \geq 5.5\cdot y_2 \quad ,\forall j \in MCR \quad (2)$$

$$\sum_{i\in I}\sum_{q\in Finishing} S_{iq}\cdot x_{ij} \geq 6.7\cdot y_3 \quad ,\forall j \in MC \quad (3)$$

$$\sum_{i\in I}\sum_{q\in Long\ Shot} S_{iq}\cdot x_{ij} \geq 5.5\cdot y_4 \quad ,\forall j \in MC \quad (4)$$

$$\sum_{i\in I}\sum_{q\in Finishing} S_{iq}\cdot x_{ij} \geq 6.7\cdot y_5 \quad ,\forall j \in MCL \quad (5)$$

$$\sum_{i\in I}\sum_{q\in Long\ Shot} S_{iq}\cdot x_{ij} \geq 5.5\cdot y_6 \quad ,\forall j \in MCL \quad (6)$$

$$\sum_{i\in I}\sum_{q\in Finishing} S_{iq}\cdot x_{ij} \geq 6.7\cdot y_7 \quad ,\forall j \in FR \quad (7)$$

$$\sum_{i\in I}\sum_{q\in Long\ Shot} S_{iq}\cdot x_{ij} \geq 5.5\cdot y_8 \quad ,\forall j \in FR \quad (8)$$

$$\sum_{i\in I}\sum_{q\in Finishing} S_{iq}\cdot x_{ij} \geq 6.7\cdot y_9 \quad ,\forall j \in FC \quad (9)$$

$$\sum_{i\in I}\sum_{q\in Long\ Shot} S_{iq}\cdot x_{ij} \geq 5.5\cdot y_{10} \quad ,\forall j \in FC \quad (10)$$

$$\sum_{i\in I}\sum_{q\in Finishing} S_{iq}\cdot x_{ij} \geq 6.7\cdot y_{11} \quad ,\forall j \in FL \quad (11)$$

$$\sum_{i\in I}\sum_{q\in Long\ Shot} S_{iq}\cdot x_{ij} \geq 5.5\cdot y_{12} \quad ,\forall j \in FL \quad (12)$$

$$\sum_{n=1}^{12} y_n = 1 \quad (13)$$

For the second scenario, [C.7] 'at least 1 *Attacker at flank* crossing skill' the number of possible positions is 2 and we have 1 skill. In this case N = 2 so we need 2 constraints for this scenario.

$$\sum_{i\in I}\sum_{q\in Crossing} S_{iq}\cdot x_{ij} \geq 5.5\cdot \mathrm{y} \quad ,\forall j \in FR \quad (1)$$

$$\sum_{i\in I}\sum_{q\in Crossing} S_{iq}\cdot x_{ij} \geq 5.5\cdot (1-y) \quad ,\forall j \in FL \quad (2)$$

The third scenario, [I.3] 'at least 1 *Defender at centre* with *interception* skill' the number of possible position is again 2 and we have 1 skill. Here N = 2 and so 2 constraints are needed.

$$\sum_{i\in I}\sum_{q\in Interceptions} S_{iq}\cdot x_{ij} \geq 8.0\cdot \mathrm{y} \quad ,\forall j \in DCR \quad (1)$$

$$\sum_{i\in I}\sum_{q\in Interceptions} S_{iq}\cdot x_{ij} \geq 8.0\cdot (1-y) \quad ,\forall j \in DCR \quad (2)$$

The final scenario, [P.17] 'at least 1 *Midfielder at centre* with *short pass* skill' the number of possible positions and skills are both equal to 1. Only 1 constraint is needed for this scenario.

$$\sum_{i\in I}\sum_{q\in Short\ Pass} S_{iq}\cdot x_{ij} \geq 8.2 \quad ,\forall j \in MC \quad (1)$$

With all sets of constraints defined for the scenario's the final integer programming model that will be used to optimize the line-up of Ajax will be:

$$\max \sum_{i\in I} \sum_{j\in J} \left(PS_{ij} \cdot x_{ij}\right) \qquad (1)$$

subject to

$$\sum_{i\in I} x_{ij} = 1 \qquad (2)$$

$$\sum_{j\in J} x_{ij} \leq 1 \qquad (3)$$

$$\sum_{i\in I} \sum_{q\in Finishing} S_{iq} \cdot x_{ij} \geq 6.7 \cdot y_1 \qquad \forall j \in MCR \qquad (4)$$

$$\sum_{i\in I} \sum_{q\in Long\ Shot} S_{iq} \cdot x_{ij} \geq 5.5 \cdot y_2 \qquad \forall j \in MCR \qquad (5)$$

$$\sum_{i\in I} \sum_{q\in Finishing} S_{iq} \cdot x_{ij} \geq 6.7 \cdot y_3 \qquad \forall j \in MC \qquad (6)$$

$$\sum_{i\in I} \sum_{q\in Long\ Shot} S_{iq} \cdot x_{ij} \geq 5.5 \cdot y_4 \qquad \forall j \in MC \qquad (7)$$

$$\sum_{i\in I} \sum_{q\in Finishing} S_{iq} \cdot x_{ij} \geq 6.7 \cdot y_5 \qquad \forall j \in MCL \qquad (8)$$

$$\sum_{i\in I} \sum_{q\in Long\ Shot} S_{iq} \cdot x_{ij} \geq 5.5 \cdot y_6 \qquad \forall j \in MCL \qquad (9)$$

$$\sum_{i\in I} \sum_{q\in Finishing} S_{iq} \cdot x_{ij} \geq 6.7 \cdot y_7 \qquad \forall j \in FR \qquad (10)$$

$$\sum_{i\in I} \sum_{q\in Long\ Shot} S_{iq} \cdot x_{ij} \geq 5.5 \cdot y_8 \qquad \forall j \in FR \qquad (11)$$

$$\sum_{i\in I} \sum_{q\in Finishing} S_{iq} \cdot x_{ij} \geq 6.7 \cdot y_9 \qquad \forall j \in FC \qquad (12)$$

$$\sum_{i\in I} \sum_{q\in Long\ Shot} S_{iq} \cdot x_{ij} \geq 5.5 \cdot y_{10} \qquad \forall j \in FC \qquad (13)$$

$$\sum_{i\in I} \sum_{q\in Finishing} S_{iq} \cdot x_{ij} \geq 6.7 \cdot y_{11} \qquad \forall j \in FL \qquad (14)$$

$$\sum_{i\in I} \sum_{q\in Long\ Shot} S_{iq} \cdot x_{ij} \geq 5.5 \cdot y_{12} \qquad \forall j \in FL \qquad (15)$$

$$\sum_{n=1}^{12} y_n = 1 \qquad (16)$$

$$\sum_{i\in I} \sum_{q\in Crossing} S_{iq} \cdot x_{ij} \geq 5.5 \cdot y_{13} \qquad \forall j \in FR \qquad (18)$$

$$\sum_{i\in I} \sum_{q\in Crossing} S_{iq} \cdot x_{ij} \geq 5.5 \cdot (1 - y_{13}) \qquad \forall j \in FL \qquad (19)$$

$$\sum_{i\in I} \sum_{q\in Interceptions} S_{iq} \cdot x_{ij} \geq 8.0 \cdot y_{14} \qquad \forall j \in DCR \qquad (20)$$

$$\sum_{i\in I} \sum_{q\in Interceptions} S_{iq} \cdot x_{ij} \geq 8.0 \cdot (1 - y_{14}) \qquad \forall j \in DCR \qquad (21)$$

$$\sum_{i\in I} \sum_{q\in Short\ Pass} S_{iq} \cdot x_{ij} \geq 8.2 \qquad \forall j \in MC \qquad (22)$$

$$x_{ij} \in \{0, 1\} \qquad \forall i \in I, j \in J \qquad (23)$$

$$y_n \in \{0, 1\} \qquad \forall n \in N \qquad (24)$$

6.3 OPTIMAL LINE-UP OF AJAX

With the results of all 3 profiles it is now possible to optimize the line-up of Ajax. To compare the results of the IP model used in the thesis, 2 additional line-ups will be shown. The first extra line-up will be the optimal line-up based on the basic IP model. As mentioned before the basic IP model does not contain a team profile, it is a match of the player and the position profile. This makes it possible to see the effect of the team profile we included in the model. The second extra line-up will be based on the information obtained from the data without any optimization. The players in this line-up will be based on the matches they played in the 40 matches from the data, so it will reflect the choice of the coach in reality. For every position the player who has played the most matches will be assigned to the position.

To compare the 3 different line-ups the average team value will be used. The average team value is the sum of all individual players scores divided by the total number of players, which is 11 in case of soccer. Before comparing the line-ups, each line-up will be discussed first.

In table 6.8 the players selected at each position with their corresponding position scores are placed. The team value obtained by the optimization is 920, which corresponds to an average team value of 83.63. The first we notice in this table is that a player is included in the line-up with a score below 55. This player is Ebecilio and he is positioned as forward left (FL). This could be caused by the fact that the player has for some important skills a grade with 0.1. As explained in the player profile section, players who has performed a skill less that the average of the team are set to 0.1. In table 6.9 the detailed profile of Ebecilio is placed with the grades for the skills, the grades required to play as forward left and the corresponding weights. Here we see that Ebecilio has for 5 skills a grade of 0.1 and for 6 skills with a grade less than required. This explains why the grade of Ebecilio is low as forward left.

Position	Player	Score
GK	Stekelenburg	101.1
DR	Van der Wiel	93.5
DCR	Enoh	76.1
DCL	Alderweireld	91.0
DL	Anita	67.5
MCR	Zeeuw	69.0
MC	Eriksen	95.5
MCL	Vertonghen	69.1
FR	Sulejmani	109.2
FC	De Jong	95.3
FL	Ebecilio	52.7

Table 6.8 - Optimal line-up

Ebecilio	Short pass back	Short pass forward	Short pass wide	Long pass back	Long pass forward	Long pass wide	Cross	Heading	Finishing	Long shot	Dribbles	Reception	Interceptions	Goalkeeping	Goal kick	Set pieces	Short pass	Long Passes	Pressing
Grades	3.8	0.1	6.8	0.1	0.1	0.1	6.7	5.2	6.1	7.1	6.7	4.5	0.1	0.1	0.1	0.1	8.1	0.1	8.0
Grades FL	3.6	8.0	7.2	3.2	7.0	5.0	6.6	5.0	7.2	6.5	6.9	5.4	8.3	1.0	1.0	5.3	8.3	4.1	8.4
Weight	.29	.04	.05	.12	.05	0	.41	.04	.53	.30	.18	.24	.09	0	0	.17	.08	0	.46

Table 6.9 - Profile of Ebecilio as Forward Left (FL)

The line-up placed in table 6.10 is obtained by only matching the player and position profile, which is the basic IP model discussed earlier. The team value of the basic IP is equal to 920 with an average team value of 83.63. The line-up obtained with the basic IP model is exactly the same as in the IP model with the team profile. We expected a line-up with a higher team value because no additional constraints are involved in this model. This means that the scenarios already hold for the best possible line-up based on the position scores.

Position	Player	Score
GK	Stekelenburg	101.1
DR	Van der Wiel	93.5
DCR	Enoh	76.1
DCL	Alderweireld	91.0
DL	Anita	67.5
MCR	Zeeuw	69.0
MC	Eriksen	95.5
MCL	Vertonghen	69.1
FR	Sulejmani	109.2
FC	De Jong	95.3
FL	Ebecilio	52.7

Table 6.10 - Optimal line-up with basic IP model (without team profile)

The last line-up in table 6.11 contains the preferences of the coach. This is line-up contains the players who has played the most matches at the positions. The total team value of this line-up is equal to 851.9 with an average team value of 77.45. It is remarkable that the line-up of the coach contains 3 players with a score less than 55. This means that for these players the data was not able to judge them correctly based on their skills. In our line-up Enoh is positioned as DCR with score 76.1 while in the line-up of the coach Enoh is positioned as MCL with grade 21.1. Another major difference is that Vertonghen is positioned as MCL in our model while he is an defender of origin. Finally we see that El Hamdaoui is positioned as FC with a score of 46.3. All these remarkable results could be caused by the fact the player position scores we calculated are only based on the skills we defined. As mentioned before we miss an important aspect of soccer skills, namely the physical soccer skills.

Position	Player	Score
GK	Stekelenburg	101.1
DR	Van der Wiel	93.5
DCR	Alderweireld	101.6
DCL	Vertonghen	94.4
DL	Anita	67.5
MCR	Zeeuw	69.0
MC	Eriksen	95.5
MCL	Enoh	21.1
FR	Sulejmani	109.2
FC	El Hamdaoui	46.3
FL	Ebecilio	52.7

Table 6.11 - Line-up with 11 most positioned players

To analyze the effect of the grades used for the constraints in the team profile we will compare 3 additional line-up. In the first line-up we have set the grades in all 4 scenario's to 7 to see the effect of it. Table 6.12 shows the line-up with the grades equal to 7. The team value of the line-up is 903.5 with an average team value of 82.14. The only difference with the original model is that Suarez is now positioned as forward left instead of Ebecilio. This is because Suarez is the only player with a crossing skills higher than 7 (see table 6.1).

Position	Player	Score
GK	Stekelenburg	101.1
DR	Van der Wiel	93.5
DCR	Enoh	76.1
DCL	Alderweireld	91.0
DL	Anita	67.5
MCR	Zeeuw	69.0
MC	Eriksen	95.5
MCL	Vertonghen	69.1
FR	Sulejmani	109.2
FC	De Jong	95.3
FL	Suarez	36.2

Table 6.12 - Line-up with grades 7

In the second line-up we lowered the grades to 6. This resulted again in exactly the same line-up as in our original model. In the final line-up the grades will be set to 8. This model was infeasible, no solution is found. This is because none of the players has for the crossing skill a grade higher of equal to 8.

Line-up	Total Team value	Average Team value
With team profile	920	83.63
Without team profile	920	83.63
Grades set to 6	920	83.63
Grades set to 7	903.5	82.14
Coach	851.9	77.45

Table 6.13 - Comparison of different line-up values

We end this chapter with a final comparison between our line-up with the line-up obtained by the basic IP model and the line-up of the coach. Based on the total and average team value, there is no difference between the model with or without team profile. They both result in the highest total team value (see table 6.13). Further setting the grades of the team profile to 7 resulted in a lower team score. The effect of the crossing grade has a large impact on the selection of the line-up. Only a few players have a high grade for the crossing skill which explains the result. Further we expected that the line-up without team profile could result in a higher team value than our model because we added extra constraints to the model. But this was not the case.

The fact that the line-up based on the preferences of the coach have resulted in the lowest total team value is as expected. As mentioned before the data we possess does not contain all necessary information to grade the players and the required grades at the positions. The physical skills of the players such as speed, acceleration, condition and strength are not included in the data, which is important to grade the players and the requirements at the positions.

7 CONCLUSION & DISCUSSION

The aim of this thesis is to optimize the line-up in soccer by using real-life data. The line-up is very important in soccer, this is because of the rule that only 3 players can be substituted during a match. So it is important for the coach that he starts with the best possible line-up to a game. To accomplish the optimization of the line-up the selection of the line-up in practice is discussed first. The information the coach possesses to decide the line-up consists of his observations during trainings session and matches. These observations are based on the soccer skills of the players. Before defining these soccer skills the terms *skill* and *quality* have been described. The soccer skills created are based on the description of these two terms.

The next step of the coach consist of defining the formation and tactics he will use. With a formation and tactic in mind the coach judges the skills of the players and the skills required to play at a position on the field. To judge the skills of the players a player profile is created which grades all the soccer skills of all players. For the skills required at the positions a similar profile is created, the position profile. The position profiles also contain weights for each skill besides the grades required for the positions. For both profiles the formation is assumed to be fixed. To cover the team tactics given by the coach the team profile is created. In the team profile the 3 major aspects of soccer are analyzed to reveal the tactics of the team. These 3 aspects are attacking, defending and passing. For each one of these aspects several scenarios have been created and depending on the team statistics one scenario will be chosen.

The final choice of the coach is the determination of the line-up based on the observations. In our case we used a IP model to determine the best line-up with as input the player, position and team profile. The player and position profile will be matched and the team profile will add additional sets of constraints based on the team tactics. The main research question we answered in this thesis is:

- *How do we optimize the line-up in soccer with a real-life database containing only actions during a match?*

To compare the line-up found with our model we added an extra line-up which is based on the choice of the coach. The major difference between the two line-ups was that in our model a defender is placed as midfielder and a midfielder as defender. We have also found that the optimal line-up contained players with a position score of less than 55, which means that he is according to the grading function not suitable for that position. Also the impact of the team profile is compared in the result. The basic IP model is used to define the line-up without team profile, this resulted in exactly the same line-up as with the team profile included. This could be caused by the grades given to the skills required in the team profile. To analyze the effect of the grades we examined the line-up with different grades for the skills in the team profile. Setting all grades to 7 resulted in a lower objective value and a change at 1 position, while raising the grades to 8 resulted in a infeasible model. The reason of this is that none of the players has a crossing grade of 8. Lowering the grades to 6 resulted into the same line-up as with the original grades we used. The extra set of constraints we added to the model depends on the grades given to each aspect. The costs of including a team component to the model with extra constraints to the model can result indeed in a lower objective value as expected but it has the advantage that it adds a team component to the model based on the team statistics.

The main added value of the methods in the thesis is the translation of the actions from the database into soccer skills. By grading the soccer skills the progress of the players can be evaluated. This can be used also during training sessions of the coach, by combining his own observations with the grades for the skills more accurate individual training schedules could be created. An advantage of optimizing the line-up is that it can be used for scouting purposes, with the position scores and average team value the weakness and strength of the team can be revealed with data.

As mentioned before the weakness of the data used for this model is that it does not contain any physical skills of the players. Physical skills like speed, acceleration, strength and condition are essential in team sports like soccer. A suggestion for further research on this topic would be obtaining and including these physical skills. With the physical skills included it will be easier to distinguish between players.

8 BIBLIOGRAPHY

[1] Ali, A. [2011], Measuring soccer skill performance: A review, Scandinavian Journal of Medicine and Science in Sports 21, 170-183.

[2] Barrick, M.R. & Stewart, G.L. [1998], Relating Member Ability and Personality to Work-Team Process and Team Effectiveness, *Journal of Applied Psychology,* 83, 377-391.

[3] Boon, B.H. & Sierksma, G. [2003], Team Formation: Matching Quality Supply and Quality Demand, *European Journal of Operational Research* 148, 277-292.

[4] Collins [2006], *Collins English Dictionary & Thesaurus*, Glasgow, Fourth edition, Harper Collins Publishers .

[5] Duch, J. Waitzman J.S. & Amaral L.A.N. [2010], Quantifying the Performance of Individual Players in a Team Activity, *PloS ONE* 5(6), e10937.

[6] Hirotsu, N. & Wright, M. [2003], Determining the best strategy for changing the configuration of a football team, *The Journal of Operational Research Society*, Vol. 54, No. 8, 878-887.

[7] Kleef, M. van [2005], Optimale Trainingsschema's voor Teamsporten, PhD thesis, Rijksuniversiteit Groningen, Groningen, The Netherlands.

[8] Knapp, B. [1977], Skill in sport: the attainment of proficiency, London: Routledge.

[9] Koerkamp, S. Klein [2004], Effectiveness of actions & the weighted average grade calculation, PhD thesis, Rijksuniversiteit Groningen, Groningen, The Netherlands.

[10] Kok, B. [2005], Teamsamenstellingen als Inverse Optimaliseringsprobleem, PhD thesis, Rijksuniversiteit Groningen, Groningen, The Netherlands.

[11] Oberstone, J. [2009], Differentiating the Top English Premier League Football Clubs from the Rest of the Pack: Identifying the Keys to Success, *Journal of Quantitative Analysis in Sports*: Vol. 5: Iss. 3, Article 10.

[12] Richard, J.P. From Scientific Baseball to Sabermetrics, *NINE: A Journal of Baseball History and Culture*, Volume 11, Number 1, 34-48.

[13] Sierksma, G. [2006], Computer Support for Coaching and Scouting in Football, *The Engineering of Sports 6: Developments for Innovation*, Springer, 215-219.

[14] Siesling, D.S. [2004], Soccer Salary, PhD thesis, Rijksuniversiteit Groningen, Groningen, The Netherlands.

APPENDIX A

<table>
<tr><th>Action</th><th>Attributes</th><th>Action</th><th>Attributes</th></tr>
<tr><td>Reception</td><td>Right foot
Left foot
Right leg
Body</td><td>Attacking Action</td><td>Body
Sliding player
Sliding opponent
Fake pass
Head</td></tr>
<tr><td>Goal attempt</td><td>Right foot
Left foot
Head
Volley
*Besides target
*On the post
*Over
*Goal
*Saved by keeper</td><td>Penalty</td><td>Right corner
Left corner
Trough the centre
*High
*Over the ground
*Post
*Besides target
*Over
**Goal
**Saved by keeper</td></tr>
<tr><td>Direct free kick</td><td>Right foot
Left foot
*Saved by keeper
*Besides target
*Over
*Goal
*On the post</td><td>Corner</td><td>Right foot
Left foot
*Curved out
*Curved in
*Short corner</td></tr>
<tr><td>Indirect free kick</td><td>Right foot
Left foot
Assist</td><td>Save on goal attempt</td><td>Ground
In the air
*Opponent goal</td></tr>
<tr><td>Defending Action</td><td>Body
Sliding player
Sliding opponent
With head</td><td>Pass</td><td>Right foot
Left foot
Head
Cross pass</td></tr>
<tr><td rowspan="4">Interception</td><td rowspan="4">Head
On deep pass
On cross pass
*Sliding</td><td colspan="2">Goal kick</td></tr>
<tr><td colspan="2">Move</td></tr>
<tr><td colspan="2">Throw in</td></tr>
<tr><td colspan="2">Free kick awarded</td></tr>
<tr><td colspan="2">Offside</td><td colspan="2">Foul</td></tr>
</table>

Table A.1 - All actions in database with additional attributes

APPENDIX B

Skill	Index q	notation
Short pass	{1}	SPA
Long pass	{2}	LPA
Passing	{1,2}	PAS
Crossing	{7}	CRO
Finishing	{9}	FIN
Long shot	{10}	LSH
Interceptions	{13}	INT
Finishing or Long shot	{9,10}	FI_LS

Table B.1 - Skills used in constraint with indices and short notation

Position	Index j	notation
Defender	{2,3,4,5}	DEF
Defender at flank	{2,5}	DEF_F
Defender at centre	{3,4}	DEF_C
Midfielder	{6,7,8}	MID
Midfielder at flank	{6,8}	MID_F
Midfielder at centre	{7}	MID_C
Attacker	{9,10,11}	ATT
Attacker at flank	{9,11}	ATT_F
Attacker at centre	{10}	ATT_C
Player at flank	{2,5,6,8,9,11}	FLA
Player at centre	{3,4,7,10}	CEN
Midfielder or Attacker	{6,7,8,9,10,11}	M_A
Defender or Midfielder	{2,3,4,5,6,7,8}	D_M
Defender of Midfielder at centre	{3,4,7}	D_M_C
Defender of Midfielder at flank	{2,5,6,8}	D_M_F
Midfielder or Attacker at centre	{7,10}	M_A_C
Midfielder of Attacker at flank	{6,8,9,11}	M_A_F

Table B.2 - Positions used in constraint with indices and short notation

Scenario	Index j	Index q	SHOTS AT GOAL Constraint	∀	N
S.1	*MID*	*FI_LS*	$\sum_{i\in I}\sum_{q\in FI\_LS} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in MID$	6
S.2	*ATT*	*FI_LS*	$\sum_{i\in I}\sum_{q\in FI\_LS} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in ATT$	6
S.3	*M_A*	*FI_LS*	$\sum_{i\in I}\sum_{q\in FI\_LS} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in M\_A$	12
S.4	*MID*	*FIN*	$\sum_{i\in I}\sum_{q\in FIN} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in MID$	3
S.5	*ATT*	*FIN*	$\sum_{i\in I}\sum_{q\in FIN} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in ATT$	3
S.6	*M_A*	*FIN*	$\sum_{i\in I}\sum_{q\in FIN} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in M\_A$	6
S.7	*MID*	*LSH*	$\sum_{i\in I}\sum_{q\in LSH} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in MID$	3
S.8	*ATT*	*LSH*	$\sum_{i\in I}\sum_{q\in LSH} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in ATT$	3
S.9	*M_A*	*LSH*	$\sum_{i\in I}\sum_{q\in LSH} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in M\_A$	6

Table B.3 - Basic constraint per scenario for the shots at goal statistic

Scenario	Index j	Index q	CROSSES Constraint	$\forall$	N
C.1	DEF	CRO	$\sum_{i\in I}\sum_{q\in CRO} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in DEF$	4
C.2	MID	CRO	$\sum_{i\in I}\sum_{q\in CRO} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in MID$	3
C.3	ATT	CRO	$\sum_{i\in I}\sum_{q\in CRO} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in ATT$	3
C.4	J	CRO	$\sum_{i\in I}\sum_{q\in CRO} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in J$	11
C.5	DEF_F	CRO	$\sum_{i\in I}\sum_{q\in CRO} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in DEF\_F$	2
C.6	MID_F	CRO	$\sum_{i\in I}\sum_{q\in CRO} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in MID\_F$	2
C.7	ATT_F	CRO	$\sum_{i\in I}\sum_{q\in CRO} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in ATT\_F$	2
C.8	FLA	CRO	$\sum_{i\in I}\sum_{q\in CRO} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in FLA$	6
C.9	DEF_C	CRO	$\sum_{i\in I}\sum_{q\in CRO} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in DEF\_C$	2
C.10	MID_C	CRO	$\sum_{i\in I}\sum_{q\in CRO} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in MID\_C$	1
C.11	ATT_C	CRO	$\sum_{i\in I}\sum_{q\in CRO} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in ATT\_C$	1
C.12	CEN	CRO	$\sum_{i\in I}\sum_{q\in CRO} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in CEN$	4
C.13	-	-	-	-	-

Table B.4 - Basic constraint per scenario for the crosses statistic

Scenario	Index j	Index q	INTERCEPTIONS Constraint	$\forall$	N
I.1	DEF	INT	$\sum_{i\in I}\sum_{q\in INT} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in DEF$	4
I.2	D_M	INT	$\sum_{i\in I}\sum_{q\in INT} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in D\_M$	7
I.3	DEF_C	INT	$\sum_{i\in I}\sum_{q\in INT} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in DEF\_C$	2
I.4	D_M_C	INT	$\sum_{i\in I}\sum_{q\in INT} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in D\_M\_C$	3
I.5	DEF_F	INT	$\sum_{i\in I}\sum_{q\in INT} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in DEF\_F$	2
I.6	D_M_F	INT	$\sum_{i\in I}\sum_{q\in INT} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in D\_M\_F$	4
I.7	MID	INT	$\sum_{i\in I}\sum_{q\in INT} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in MID$	3
I.8	M_A	INT	$\sum_{i\in I}\sum_{q\in INT} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in M\_A$	6
I.9	MID_C	INT	$\sum_{i\in I}\sum_{q\in INT} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in MID\_C$	1
I.10	M_A_C	INT	$\sum_{i\in I}\sum_{q\in INT} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in M\_A\_C$	2
I.11	MID_F	INT	$\sum_{i\in I}\sum_{q\in INT} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in MID\_C$	2
I.12	M_A_F	INT	$\sum_{i\in I}\sum_{q\in INT} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in M\_A\_F$	4

Table B.5 - Basic constraint per scenario for the interceptions statistic

Scenario	Index *j*	Index *q*	PASSING Constraint	∀	*N*
P.1	*DEF_F*	*LPA*	$\sum_{i\in I}\sum_{q\in LPA} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in DEF\_F$	2
P.2	*DEF_F*	*SPA*	$\sum_{i\in I}\sum_{q\in SPA} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in DEF\_F$	2
P.3	*DEF_F*	*PAS*	$\sum_{i\in I}\sum_{q\in PAS} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in DEF\_F$	4
P.4	*DEF_C*	*LPA*	$\sum_{i\in I}\sum_{q\in LPA} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in DEF\_C$	2
P.5	*DEF_C*	*SPA*	$\sum_{i\in I}\sum_{q\in SPA} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in DEF\_C$	2
P.6	*DEF_C*	*PAS*	$\sum_{i\in I}\sum_{q\in PAS} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in DEF\_C$	4
P.7	*DEF*	*LPA*	$\sum_{i\in I}\sum_{q\in LPA} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in DEF$	4
P.8	*DEF*	*SPA*	$\sum_{i\in I}\sum_{q\in SPA} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in DEF$	4
P.9	*DEF*	*PAS*	$\sum_{i\in I}\sum_{q\in PAS} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in DEF$	8
P.10	*MID_F*	*LPA*	$\sum_{i\in I}\sum_{q\in LPA} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in MID\_F$	2
P.11	*MID_F*	*SPA*	$\sum_{i\in I}\sum_{q\in SPA} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in MID\_F$	2
P.12	*MID_F*	*PAS*	$\sum_{i\in I}\sum_{q\in PAS} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in MID\_F$	4
P.13	*MID_C*	*LPA*	$\sum_{i\in I}\sum_{q\in LPA} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in MID\_C$	1
P.14	*MID_C*	*SPA*	$\sum_{i\in I}\sum_{q\in SPA} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in MID\_C$	1
P.15	*MID_C*	*PAS*	$\sum_{i\in I}\sum_{q\in PAS} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in MID\_C$	1
P.16	*MID*	*LPA*	$\sum_{i\in I}\sum_{q\in LPA} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in MID$	3
P.17	*MID*	*SPA*	$\sum_{i\in I}\sum_{q\in SPA} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in MID$	3
P.18	*MID*	*PAS*	$\sum_{i\in I}\sum_{q\in PAS} S_{iq}\cdot x_{ij} \geq grade_q$	$j \in MID$	6

Table B.6 - Basic constraint per scenario for the passes statistic

APPENDIX C

Player Profile																			
Player	Short pass back	Short pass forward	Short pass wide	Long pass back	Long pass forward	Long pass wide	Cross	Heading	Finishing	Long shot	Dribbles	Reception	Interceptions	Goalkeeping	Goal kick	Set pieces	Short pass	Long Passes	Pressing
Oleguer	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Vertonghen	4.9	8.6	7.9	4.2	8.3	5.7	0.1	7.3	7.5	5.0	6.9	0.1	8.7	0.1	0.1	5.6	8.3	8.5	0.1
Lindgren	0.1	0.1	0.1	0.1	0.1	0.1	0.1	6.8	0.1	6.6	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Emanuelson	3.8	0.1	0.1	4.0	0.1	4.6	6.5	0.1	0.1	0.1	7.6	0.1	8.4	0.1	0.1	6.0	0.1	4.2	6.0
Vermeer	0.1	0.1	0.1	0.1	3.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	9.0	5.5	0.1	0.1	0.1	0.1
Cvitanich	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
de Jong	5.1	8.1	7.1	3.4	0.1	0.1	4.5	4.4	8.5	4.6	7.2	5.6	8.2	0.1	0.1	0.1	8.7	3.9	6.7
Zeeuw	4.5	8.5	7.0	2.9	7.8	5.6	4.0	6.1	7.6	5.8	0.1	0.1	8.3	0.1	0.1	6.8	8.4	5.6	6.7
El Hamdaoui	3.9	0.1	7.4	2.7	0.1	0.1	6.1	0.1	7.9	3.7	7.0	4.4	0.1	0.1	0.1	0.1	8.7	0.1	4.0
Verhoeven	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	8.4	5.3	0.1	0.1	0.1	0.1
Alderweireld	4.6	7.9	7.7	4.3	8.4	5.8	0.1	7.5	0.1	7.1	6.4	0.1	7.4	0.1	0.1	5.9	7.5	8.4	0.1
Blind	4.1	8.4	0.1	3.9	0.1	0.1	4.1	0.1	0.1	0.1	0.1	0.1	8.5	0.1	0.1	0.1	8.0	0.1	0.1
Zeegelaar	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Ooijer	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Wiel	6.2	8.3	8.1	4.1	6.8	5.7	4.9	6.4	0.1	0.1	6.6	4.5	8.3	0.1	0.1	5.8	8.1	4.9	5.4
Boilesen	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Lukoki	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Deckers	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Anita	5.1	8.4	7.3	3.9	6.4	5.3	4.3	6.2	0.1	0.1	6.9	0.1	8.7	0.1	0.1	0.1	8.1	4.5	0.1
Enoh	5.0	6.9	7.2	3.7	7.2	5.3	0.1	6.4	0.1	0.1	0.1	0.1	8.8	0.1	0.1	0.1	7.5	6.3	0.1
Eriksen	5.2	8.6	7.2	3.1	0.1	6.1	4.1	6.7	8.4	4.0	7.3	6.1	8.0	0.1	0.1	5.0	8.6	5.4	6.3
Ozbiliz	0.1	0.1	0.1	0.1	0.1	0.1	3.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	3.6
Stekelenburg	0.1	4.8	0.1	0.1	8.4	5.4	0.1	0.1	0.1	0.1	0.1	0.1	0.1	8.8	6.4	5.1	0.1	8.3	0.1
Suarez	0.1	0.1	0.1	3.1	0.1	0.1	7.2	6.0	8.2	3.9	0.1	7.1	0.1	0.1	0.1	4.7	0.1	0.1	8.6
Sulejmani	3.8	7.6	6.6	3.3	0.1	0.1	5.7	5.7	7.7	5.2	7.5	6.5	8.0	0.1	0.1	6.0	8.3	3.9	7.3
Tainio	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
H. Hussein	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Ebecilio	3.8	0.1	6.8	0.1	0.1	0.1	6.7	5.2	6.1	7.1	6.7	4.5	0.1	0.1	0.1	0.1	8.1	0.1	8.0
Castillion	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1

Table C.1 - Player profile of all 29 players of Ajax

Multiplication values of Fractions

Player	Short pass back	Short pass forward	Short pass wide	Long pass back	Long pass forward	Long pass wide	Cross	Heading	Finishing	Long shot	Dribbles	Reception	Interceptions	Goalkeeping	Goal kick	Set pieces	Short pass	Long Passes	Pressing
GK	1	1	1	1	1	1	1	1	1	1	1	1	1	5	5	1	1	1	1
DR	3	4	4	3	4	4	4	4	2	3	3	3	5	0	0	3	4	4	1
DCR	1	5	5	1	3	5	2	5	1	1	2	2	3	0	0	1	5	5	3
DCL	1	5	5	1	3	5	2	5	1	1	2	2	3	0	0	1	5	5	3
DL	2	3	3	2	5	3	3	3	3	2	4	1	4	0	0	2	3	3	2
MCR	4	3	3	5	3	3	3	3	3	4	4	4	4	0	0	4	3	3	2
MC	2	4	4	3	2	4	1	4	2	2	3	3	2	0	0	2	4	4	4
MCL	3	2	2	4	4	2	2	2	4	3	5	2	3	0	0	3	2	2	3
FR	5	2	2	4	2	2	5	2	4	5	2	5	3	0	0	5	2	2	3
FC	3	3	3	2	1	3	3	3	3	3	1	4	1	0	0	3	3	3	5
FL	4	1	1	3	3	1	4	1	5	4	3	3	2	0	0	4	1	1	4

Table C.2 - Weights with which the fractions are multiplied with $[w_3, w_2, w_1, v_3, v_2, v_1] = [3, 2, 1, 2, 1, 0]$

Weights Position Profile

Player	Short pass back	Short pass forward	Short pass wide	Long pass back	Long pass forward	Long pass wide	Cross	Heading	Finishing	Long shot	Dribbles	Reception	Interceptions	Goalkeeping	Goal kick	Set pieces	Short pass	Long Passes	Pressing
GK	0	.03	0	0	.11	0	0	0	0	0	0	0	0	3.39	3.28	0	0	.07	0
DR	.46	.56	.57	.42	.22	.28	.98	.37	0	.13	.29	.20	.97	0	0	.14	.81	.34	.04
DCR	.04	.90	.81	.04	.52	.28	.04	.71	0	0	.17	0	.37	0	0	.05	1.02	1.09	.03
DCL	0	.93	.87	.00	.50	.90	0	.76	0	.04	.22	0	.44	0	0	.05	1.02	.93	.10
DL	.19	.46	.34	.21	.47	.23	.37	.25	.11	.07	.48	0	.73	0	0	.14	.55	.30	.12
MCR	.51	.35	.36	.66	.19	.26	.17	.32	.31	.55	.27	.36	.71	0	0	.24	.51	.29	.15
MC	.21	.49	.55	.31	.08	.33	.03	.41	.21	.21	.27	.38	.16	0	0	.19	.79	.36	.50
MCL	.39	.20	.24	.40	.30	.16	.10	.16	.23	.29	.54	.12	.42	0	0	.32	.33	.16	.25
FR	.79	.10	.13	.42	.03	.08	.84	.11	.53	.72	.16	.90	.24	0	0	1.03	.22	.09	.48
FC	.36	.21	.31	.18	0	.13	.20	.47	.60	.31	.04	.93	0	0	0	.08	.49	.19	1.15
FL	.50	.04	.05	.19	.10	0	.72	.04	1.06	.53	.35	.40	.13	0	0	.30	.08	0	.81

Table C.3 - Weights of the skills at the positions with $[w_3, w_2, w_1, v_3, v_2, v_1] = [5, 3, 1, 5, 2, 0]$

Player	Short pass back	Short pass forward	Short pass wide	Long pass back	Long pass forward	Long pass wide	Cross	Heading	Finishing	Long shot	Dribbles	Reception	Interceptions	Goalkeeping	Goal kick	Set pieces	Short pass	Long Passes	Pressing
Weights Position Profile																			
GK	0	.03	0	0	.11	0	0	0	0	0	0	0	0	6.77	6.56	0	0	.07	0
DR	.84	1.21	1.22	.77	.47	.59	1.84	.79	0	.23	.59	.37	1.94	0	0	.25	1.74	.74	.04
DCR	.04	1.80	1.62	.04	1.04	2.55	.07	1.42	0	0	.29	0	.75	0	0	.05	2.03	2.18	.06
DCL	0	1.87	1.73	0	1.01	1.81	.04	1.52	0	.04	.37	.03	.88	0	0	.05	2.05	1.86	.19
DL	.37	.91	.67	.43	.93	.45	.74	.50	.20	.14	.90	0	1.57	0	0	.27	1.11	.61	.25
MCR	.96	.69	.72	1.33	.39	.53	.32	.64	.62	1.03	.58	.67	1.33	0	0	.45	1.03	.58	.25
MC	.35	.91	1.04	.61	.13	.62	.03	.76	.35	.35	.53	.76	.26	0	0	.31	1.48	.68	.93
MCL	.78	.33	.40	.86	.55	.26	.19	.27	.43	.58	1.08	.20	.84	0	0	.63	.55	.27	.50
FR	1.59	.20	.26	.79	.07	.16	1.69	.22	1.14	1.44	.32	1.79	.44	0	0	2.06	.44	.18	.95
FC	.73	.38	.56	.30	0	.23	.39	.86	1.20	.62	.04	1.99	0	0	0	.16	.89	.34	2.30
FL	1.08	.04	.05	.39	.17	0	1.54	.04	2.13	1.13	.65	.81	.26	0	0	.64	.08	0	1.74

Table C.4 - Weights of the skills at the positions with [w_3, w_2, w_1, v_3, v_2, v_1] = [10, 5, 1, 10, 5, 0]

Player	GK	DR	DCR	DCL	DL	MCR	MC	MCL	FR	FC	FL
Player Position Scores											
Oleguer	1.3	1.5	1.4	1.3	1.5	1.7	1.6	1.6	2.0	1.8	1.6
Vertonghen	1.8	53.5	105.5	94.4	72.5	72.3	55.3	69.1	32.8	23.3	21.9
Lindgren	1.3	2.1	2.1	2.1	2.0	3.0	2.6	2.5	3.3	3.3	2.5
Emanuelson	1.4	15.2	8.1	6.4	15.2	12.8	9.4	17.1	21.4	9.4	15.5
Vermeer	81.3	1.7	1.9	1.7	2.0	1.9	1.7	2.0	2.0	1.8	1.7
Cvitanich	1.3	1.5	1.4	1.3	1.5	1.7	1.6	1.6	2.0	1.8	1.6
de Jong	1.5	63.4	28.3	32.8	50.8	62.2	62.2	50.7	59.9	95.3	71.2
Zeeuw	1.7	69.9	82.7	76.1	70.9	69.0	60.6	59.2	61.5	53.0	53.3
El Hamdaoui	1.3	15.8	5.1	5.5	11.6	18.6	21.1	18.7	37.5	46.3	49.9
Verhoeven	69.8	1.5	1.4	1.3	1.5	1.7	1.6	1.6	2.0	1.8	1.6
Alderweireld	1.8	51.2	101.6	91.0	64.0	56.5	44.5	56.9	23.4	14.2	10.5
Blind	1.4	18.3	6.2	6.1	13.3	10.8	7.4	8.8	8.5	5.2	5.8
Zeegelaar	1.3	1.5	1.4	1.3	1.5	1.7	1.6	1.6	2.0	1.8	1.6
Ooijer	1.3	1.5	1.4	1.3	1.5	1.7	1.6	1.6	2.0	1.8	1.6
Van der Wiel	1.7	93.5	93.3	84.8	88.3	60.5	69.4	66.4	53.7	51.3	29.6
Boilesen	1.3	1.5	1.4	1.3	1.5	1.7	1.6	1.6	2.0	1.8	1.6
Lukoki	1.3	1.5	1.4	1.3	1.5	1.7	1.6	1.6	2.0	1.8	1.6
Deckers	1.3	1.5	1.4	1.3	1.5	1.7	1.6	1.6	2.0	1.8	1.6
Anita	1.7	71.3	84.5	73.4	67.5	37.1	30.8	34.9	13.3	11.5	9.2
Enoh	1.7	34.9	76.1	62.7	35.9	27.7	23.6	21.1	7.6	9.4	4.3
Eriksen	1.5	83.6	66.8	63.6	71.5	88.4	95.5	78.9	109.5	113.9	86.4
Ozbiliz	1.3	2.6	1.4	1.4	2.2	2.0	2.2	2.1	3.9	4.1	4.4
Stekelenburg	101.1	4.1	13.2	9.8	5.6	4.3	4.6	4.5	4.2	2.8	2.2
Suarez	1.3	5.7	2.3	2.4	4.5	9.2	9.2	8.1	36.7	35.5	36.2
Sulejmani	1.5	69.3	29.9	34.8	57.9	71.6	74.1	66.8	109.2	105.1	91.6
Tainio	1.3	1.5	1.4	1.3	1.5	1.7	1.6	1.6	2.0	1.8	1.6
H. Hussein	1.3	1.5	1.4	1.3	1.5	1.7	1.6	1.6	2.0	1.8	1.6
Ebecilio	1.3	16.7	7.1	8.3	12.2	17.7	23.8	17.5	37.3	62.6	52.7
Castillion	1.3	1.5	1.4	1.3	1.5	1.7	1.6	1.6	2.0	1.8	1.6

Table C.5 - Player Position Scores of all 29 players of Ajax

Player Position Scores											
Player	GK	DR	DCR	DCL	DL	MCR	MC	MCL	FR	FC	FL
Oleguer	13.4	15.5	13.9	13.0	15.3	16.9	16.1	16.3	19.9	18.4	15.9
Vertonghen	15.3	79.7	105.5	97.6	92.0	92.5	81.3	88.6	68.0	57.9	50.9
Lindgren	13.4	17.8	16.8	16.3	17.5	21.9	20.0	19.8	24.9	23.8	19.6
Emanuelson	13.8	41.7	28.8	25.5	41.9	39.7	33.8	44.6	54.7	36.4	42.5
Vermeer	87.2	16.0	15.3	14.3	16.7	17.5	16.4	17.3	19.9	18.4	16.1
Cvitanich	13.4	15.5	13.9	13.0	15.3	16.9	16.1	16.3	19.9	18.4	15.9
de Jong	14.1	79.0	52.9	55.1	72.9	81.2	80.8	73.5	85.7	103.5	83.9
Zeeuw	15.1	83.1	88.9	83.8	85.7	86.4	81.5	79.5	87.2	81.3	73.2
El Hamdaoui	13.4	42.3	25.5	25.4	37.1	45.4	48.0	44.8	65.6	69.4	66.5
Verhoeven	80.3	15.5	13.9	13.0	15.3	16.9	16.1	16.3	19.9	18.4	15.9
Alderweireld	15.2	76.3	101.6	94.1	84.8	82.0	72.9	80.2	58.6	46.0	36.4
Blind	13.7	46.0	28.2	27.2	40.6	38.2	31.9	34.0	35.8	28.6	26.7
Zeegelaar	13.4	15.5	13.9	13.0	15.3	16.9	16.1	16.3	19.9	18.4	15.9
Ooijer	13.4	15.5	13.9	13.0	15.3	16.9	16.1	16.3	19.9	18.4	15.9
Van der Wiel	15.0	97.3	93.3	87.2	95.7	82.7	85.8	84.7	81.6	77.0	55.7
Boilesen	13.4	15.5	13.9	13.0	15.3	16.9	16.1	16.3	19.9	18.4	15.9
Lukoki	13.4	15.5	13.9	13.0	15.3	16.9	16.1	16.3	19.9	18.4	15.9
Deckers	13.4	15.5	13.9	13.0	15.3	16.9	16.1	16.3	19.9	18.4	15.9
Anita	14.9	84.8	87.3	80.4	83.7	66.2	60.2	63.4	44.5	40.9	33.3
Enoh	15.1	63.1	84.5	75.7	64.1	58.3	53.2	50.6	35.3	37.6	24.3
Eriksen	14.2	90.4	81.2	77.4	85.8	95.6	99.5	89.3	109.5	113.9	89.6
Ozbiliz	13.4	18.3	13.9	13.3	17.3	18.0	18.1	18.0	25.1	24.3	22.5
Stekelenburg	101.1	23.7	37.8	31.9	27.3	25.5	25.3	25.6	27.3	21.9	18.3
Suarez	13.4	27.2	17.2	16.9	24.4	34.2	34.4	31.8	70.4	70.4	64.7
Sulejmani	14.1	82.8	54.1	56.7	77.3	86.2	87.6	82.6	112.8	110.1	94.9
Tainio	13.4	15.5	13.9	13.0	15.3	16.9	16.1	16.3	19.9	18.4	15.9
H. Hussein	13.4	15.5	13.9	13.0	15.3	16.9	16.1	16.3	19.9	18.4	15.9
Ebecilio	13.4	44.6	29.0	30.2	39.0	47.4	53.5	46.6	71.0	86.3	74.5
Castillion	13.4	15.5	13.9	13.0	15.3	16.9	16.1	16.3	19.9	18.4	15.9

Table C.6 - Player Position Scores (empty skill grades set to 1)

Player Position Scores											
Player	GK	DR	DCR	DCL	DL	MCR	MC	MCL	FR	FC	FL
Oleguer	6.7	7.7	7.0	6.5	7.7	8.4	8.1	8.1	10.0	9.2	7.9
Vertonghen	8.0	70.7	105.5	96.7	85.7	85.9	72.4	82.2	54.6	44.0	39.5
Lindgren	6.7	9.4	9.0	8.9	9.1	12.0	10.8	10.7	13.5	13.1	10.6
Emanuelson	7.0	30.8	19.6	16.8	30.9	28.2	23.0	33.4	41.3	24.3	31.4
Vermeer	85.4	8.2	8.1	7.6	8.8	9.0	8.3	9.0	10.0	9.2	8.2
Cvitanich	6.7	7.7	7.0	6.5	7.7	8.4	8.1	8.1	10.0	9.2	7.9
de Jong	7.2	74.0	43.8	47.1	65.4	74.9	74.7	65.7	77.0	100.9	79.9
Zeeuw	7.9	78.9	87.0	81.4	80.9	80.7	74.5	72.7	78.5	71.5	66.5
El Hamdaoui	6.7	31.5	15.7	16.0	26.1	34.7	37.5	34.4	55.5	61.4	61.0
Verhoeven	77.0	7.7	7.0	6.5	7.7	8.4	8.1	8.1	10.0	9.2	7.9
Alderweireld	7.9	67.6	101.6	93.1	77.9	73.3	62.8	72.3	44.5	32.3	25.0
Blind	6.9	34.8	17.9	17.3	29.1	26.1	20.6	22.6	23.2	17.1	16.8
Zeegelaar	6.7	7.7	7.0	6.5	7.7	8.4	8.1	8.1	10.0	9.2	7.9
Ooijer	6.7	7.7	7.0	6.5	7.7	8.4	8.1	8.1	10.0	9.2	7.9
Van der Wiel	7.8	96.1	93.3	86.5	93.4	75.2	80.5	78.7	72.0	68.1	46.1
Boilesen	6.7	7.7	7.0	6.5	7.7	8.4	8.1	8.1	10.0	9.2	7.9
Lukoki	6.7	7.7	7.0	6.5	7.7	8.4	8.1	8.1	10.0	9.2	7.9
Deckers	6.7	7.7	7.0	6.5	7.7	8.4	8.1	8.1	10.0	9.2	7.9
Anita	7.8	80.5	86.5	78.2	78.5	55.6	49.2	53.0	31.0	28.0	22.6
Enoh	7.9	52.8	81.9	71.5	53.8	46.6	41.6	38.9	22.2	24.8	14.4
Eriksen	7.2	88.3	76.5	73.0	81.2	93.4	98.3	86.0	109.5	113.9	88.6
Ozbiliz	6.7	10.1	7.0	6.7	9.3	9.4	9.7	9.5	14.3	14.2	13.7
Stekelenburg	101.1	14.0	27.5	22.4	16.9	14.9	15.1	15.2	15.6	11.7	9.7
Suarez	6.7	17.0	9.3	9.3	14.7	23.0	23.1	21.1	57.9	57.3	54.3
Sulejmani	7.2	78.5	45.3	48.9	70.8	81.5	83.3	77.5	111.7	108.6	93.9
Tainio	6.7	7.7	7.0	6.5	7.7	8.4	8.1	8.1	10.0	9.2	7.9
H. Hussein	6.7	7.7	7.0	6.5	7.7	8.4	8.1	8.1	10.0	9.2	7.9
Ebecilio	6.7	33.2	18.9	20.5	27.5	35.2	41.9	34.7	58.5	78.4	67.2
Castillion	6.7	7.7	7.0	6.5	7.7	8.4	8.1	8.1	10.0	9.2	7.9

Table C.7 - Player Position Scores (empty skill grades set to 0.5)

Player Position Scores											
Player	GK	DR	DCR	DCL	DL	MCR	MC	MCL	FR	FC	FL
Oleguer	1.3	1.6	1.4	1.3	1.5	1.7	1.6	1.6	2.0	1.8	1.6
Vertonghen	1.5	51.8	101.2	94.2	73.9	71.7	55.6	70.6	32.3	22.8	22.0
Lindgren	1.3	2.1	2.1	2.1	2.0	3.1	2.5	2.5	3.4	3.3	2.5
Emanuelson	1.4	16.4	7.8	6.1	14.8	13.9	8.7	17.8	22.5	9.4	15.8
Vermeer	87.7	1.7	1.8	1.7	2.1	1.9	1.7	2.0	2.1	1.8	1.7
Cvitanich	1.3	1.6	1.4	1.3	1.5	1.7	1.6	1.6	2.0	1.8	1.6
de Jong	1.4	64.0	29.5	34.0	49.0	63.4	62.9	49.6	57.7	95.7	70.8
Zeeuw	1.5	70.3	85.0	78.6	66.8	68.1	63.0	55.3	61.5	56.7	52.2
El Hamdaoui	1.3	15.9	5.2	5.5	11.3	19.1	20.1	19.3	37.4	44.0	52.6
Verhoeven	79.0	1.6	1.4	1.3	1.5	1.7	1.6	1.6	2.0	1.8	1.6
Alderweireld	1.5	49.6	97.6	90.6	64.3	57.3	46.2	56.2	23.7	14.4	9.5
Blind	1.4	19.8	6.2	6.2	12.5	11.7	7.3	8.3	8.8	5.1	5.2
Zeegelaar	1.3	1.6	1.4	1.3	1.5	1.7	1.6	1.6	2.0	1.8	1.6
Ooijer	1.3	1.6	1.4	1.3	1.5	1.7	1.6	1.6	2.0	1.8	1.6
Van der Wiel	1.5	93.0	93.1	86.6	87.4	60.2	73.7	64.2	54.2	55.0	26.5
Boilesen	1.3	1.6	1.4	1.3	1.5	1.7	1.6	1.6	2.0	1.8	1.6
Lukoki	1.3	1.6	1.4	1.3	1.5	1.7	1.6	1.6	2.0	1.8	1.6
Deckers	1.3	1.6	1.4	1.3	1.5	1.7	1.6	1.6	2.0	1.8	1.6
Anita	1.5	71.3	85.1	76.4	68.9	36.4	33.3	34.3	13.0	11.9	8.3
Enoh	1.5	34.2	77.1	67.7	34.9	27.2	26.3	19.3	7.4	10.1	3.7
Eriksen	1.4	83.7	70.2	66.6	67.5	89.2	96.8	75.4	107.4	115.2	85.7
Ozbiliz	1.3	2.6	1.4	1.4	2.1	2.1	2.2	2.1	4.0	4.3	4.4
Stekelenburg	101.0	3.9	13.1	9.9	5.7	4.1	4.7	4.5	4.4	2.8	2.2
Suarez	1.3	6.0	2.3	2.3	4.2	9.8	8.4	8.1	39.3	34.7	39.0
Sulejmani	1.4	70.4	30.5	35.2	54.9	73.9	72.8	65.1	109.0	105.3	90.9
Tainio	1.3	1.6	1.4	1.3	1.5	1.7	1.6	1.6	2.0	1.8	1.6
H. Hussein	1.3	1.6	1.4	1.3	1.5	1.7	1.6	1.6	2.0	1.8	1.6
Ebecilio	1.3	16.3	7.5	8.5	12.0	17.3	23.8	17.6	36.2	63.4	54.5
Castillion	1.3	1.6	1.4	1.3	1.5	1.7	1.6	1.6	2.0	1.8	1.6

Table C.8 - Player Position Scores (Weights set to [w_3, w_2, w_1, v_3, v_2, v_1] = [5, 3, 1, 5, 2, 0])

Player Position Scores											
Player	GK	DR	DCR	DCL	DL	MCR	MC	MCL	FR	FC	FL
Oleguer	1.3	1.5	1.4	1.3	1.5	1.7	1.6	1.7	2.0	1.8	1.6
Vertonghen	1.4	55.2	101.8	92.6	74.1	73.8	54.6	70.5	32.2	20.7	20.5
Lindgren	1.3	2.1	2.1	2.1	2.0	3.1	2.5	2.5	3.4	3.2	2.5
Emanuelson	1.4	15.0	7.5	6.0	15.0	13.7	8.7	20.1	21.9	9.3	17.1
Vermeer	91.2	1.7	1.8	1.7	2.1	1.9	1.7	2.0	2.1	1.8	1.7
Cvitanich	1.3	1.5	1.4	1.3	1.5	1.7	1.6	1.7	2.0	1.8	1.6
de Jong	1.4	63.1	29.4	34.2	49.0	63.1	63.8	50.1	57.6	96.2	70.0
Zeeuw	1.4	71.4	86.3	79.3	68.7	67.4	60.3	54.7	61.5	53.9	53.5
El Hamdaoui	1.3	15.0	5.1	5.4	10.8	19.1	20.4	20.0	37.9	45.6	53.4
Verhoeven	84.2	1.5	1.4	1.3	1.5	1.7	1.6	1.7	2.0	1.8	1.6
Alderweireld	1.4	52.8	98.1	89.0	64.9	58.6	46.0	56.3	23.0	13.0	8.9
Blind	1.4	19.2	6.1	6.3	12.9	11.6	7.3	8.6	8.6	4.9	5.3
Zeegelaar	1.3	1.5	1.4	1.3	1.5	1.7	1.6	1.7	2.0	1.8	1.6
Ooijer	1.3	1.5	1.4	1.3	1.5	1.7	1.6	1.7	2.0	1.8	1.6
Van der Wiel	1.4	93.6	93.0	87.6	88.2	60.6	75.7	63.9	52.8	53.8	26.0
Boilesen	1.3	1.5	1.4	1.3	1.5	1.7	1.6	1.7	2.0	1.8	1.6
Lukoki	1.3	1.5	1.4	1.3	1.5	1.7	1.6	1.7	2.0	1.8	1.6
Deckers	1.3	1.5	1.4	1.3	1.5	1.7	1.6	1.7	2.0	1.8	1.6
Anita	1.4	74.1	86.3	77.9	69.6	37.6	33.8	33.7	12.7	10.7	7.7
Enoh	1.4	36.9	79.7	69.9	36.2	27.8	26.5	18.4	7.2	9.2	3.5
Eriksen	1.4	83.4	69.9	66.4	67.5	89.0	97.3	75.7	107.5	115.2	85.5
Ozbiliz	1.3	2.5	1.4	1.4	2.1	2.0	2.2	2.2	4.0	4.4	4.7
Stekelenburg	100.9	4.1	13.3	9.9	5.6	4.2	4.6	4.3	4.4	2.8	2.2
Suarez	1.3	5.5	2.2	2.3	4.2	9.5	8.3	8.6	39.8	37.4	42.3
Sulejmani	1.4	68.9	29.9	34.9	55.0	73.2	72.9	66.8	109.0	106.2	91.2
Tainio	1.3	1.5	1.4	1.3	1.5	1.7	1.6	1.7	2.0	1.8	1.6
H. Hussein	1.3	1.5	1.4	1.3	1.5	1.7	1.6	1.7	2.0	1.8	1.6
Ebecilio	1.3	15.8	7.4	8.4	11.6	17.2	23.8	17.4	37.0	65.7	55.2
Castillion	1.3	1.5	1.4	1.3	1.5	1.7	1.6	1.7	2.0	1.8	1.6

Table C.9 - Player Position Scores (Weights set to [w_3, w_2, w_1, v_3, v_2, v_1] = [10, 5, 1, 10, 5, 0])