

12-2009

The ranking of football teams using concepts from the analytic hierarchy process.

Yepeng Sun 1976-
University of Louisville

Follow this and additional works at: <https://ir.library.louisville.edu/etd>

Recommended Citation

Sun, Yepeng 1976-, "The ranking of football teams using concepts from the analytic hierarchy process." (2009). *Electronic Theses and Dissertations*. Paper 1408.
<https://doi.org/10.18297/etd/1408>

This Master's Thesis is brought to you for free and open access by ThinkIR: The University of Louisville's Institutional Repository. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of ThinkIR: The University of Louisville's Institutional Repository. This title appears here courtesy of the author, who has retained all other copyrights. For more information, please contact thinkir@louisville.edu.

**THE RANKING OF FOOTBALL TEAMS USING
CONCEPTS FROM THE ANALYTIC HIERARCHY
PROCESS**

BY

Yepeng Sun

Speed Engineering School, 2007

A Thesis

Submitted to the Faculty of the
Graduate School of the University of Louisville
in Partial Fulfillment of the Requirements
for the Degree of
Master of Engineering

Department of Industrial Engineering

University of Louisville

Louisville, Kentucky

December 2009

**THE RANKING OF FOOTBALL TEAMS USING CONCEPTS FROM
THE ANALYTIC HIERARCHY PROCESS**

By

Yepeng Sun

Speed Engineering School, 2007

A Thesis Approved on

November 3, 2009

By the following Thesis Committee

X

Thesis Director

X

X

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Gerald Evans, for his guidance and patience. I attribute the level of my Masters degree to his encouragement and effort and without him this thesis, too, would not have been completed or written. One simply could not wish for a better or friendlier supervisor.

I would also like to thank the other committee members, Dr. Sunderesh Heragu and Dr. Darjen Chang, for their comments and assistance on my study and writing on Master thesis.

ABSTRACT

THE RANKING OF FOOTBALL TEAMS USING CONCEPTS FROM THE ANALYTIC HIERARCHY PROCESS

Yepeng Sun

October 10, 2009

A new method to rank football teams with the concepts associated with the Analytic Hierarchy Process(AHP) and Graph theory is developed in this thesis. A set of games is interpreted as a graph, in which every vertex represents a team, and every edge represents a team quality comparison evaluated directly from the score difference of a game. A weighted quality comparison value between every two teams is computed from the graph, based on the transitivity relationship of comparison values in the Analytic Hierarchy Process(AHP). All the quality comparison values are input into the software package Expert Choice as the pair-wise comparison values in term of team quality, then Expert Choice can give the rankings of the teams. The method is illustrated through the use of scores associated with the 2007 Big East Conference football season.

This method accounts for the margin of victory(score difference) in the rankings, which is the main difference as compared to the computer methods in the BCS system. Another characteristic of our method is that our ranking method considers indirect comparison as well as direct comparison between teams, using Graph theory to integrate and weight all these comparisons.

LIST OF TABLES

Table 1 Numerical Rating in Pair-wise Comparison	11
Table 2 The Numerical Rating in Our Problem.....	15
Table 3 The Weighting Vectors for Different Paths Set.....	30
Table 4 The Outcomes of Games in 2007 Big East Conference Season	34
Table 5 The IDs of Teams in Big East Conference.....	35
Table 6 The Outcomes of the Optimization.....	46
Table 7 The Outcomes of the Weighting Vector of the Optimization	47

LIST OF FIGURES

<u>Figure 1.</u> The Hierarchy of AHP	12
<u>Figure 2.</u> The Hierarchy of Team Ranking	14
<u>Figure 3.</u> The Conversion Function from Score Difference to Quality Difference	16
<u>Figure 4.</u> The Curve of the Conversion Function when $C = 0.04$	22
<u>Figure 5.</u> The Linear Function from D_{ij} to a_{ij}	23
<u>Figure 6.</u> The Graph of the Composed Function	23
<u>Figure 7.</u> The Graph of Composed Conversion Function when $C = 0.095$	25
<u>Figure 9.</u> The Searching Tree from Team 1 to Team 2 when $k=4$	28
<u>Figure 8.</u> A Graph showing a Set of Games.....	28
<u>Figure 10.</u> The Parameter in Excel	36
<u>Figure 11.</u> The Calculation of DQ and A_{ij} in Excel	36
<u>Figure 12.</u> The Graph Formed by the Set of Games	37
<u>Figure 13.</u> The Setup of Parameter C	38
<u>Figure 14.</u> Calculate the Weighted Quality Comparison Value between Teams.....	40
<u>Figure 15.</u> The Measurement of Optimization.....	41
<u>Figure 16.</u> Determine Parameters in Excel.....	42
<u>Figure 17.</u> The Structure of Our Problem in Expert Choice	43
<u>Figure 18.</u> The Matrix Input of Pair-wise Comparison	44
<u>Figure 19.</u> The Result of the Team Ranking	45

TABLE OF CONTENTS

ACKNOWLEDGMENTS	iii
ABSTRACT	iv
LIST OF TABLES	v
LIST OF FIGURES	vi
CHAPTER 1 INTRODUCTION.....	1
CHAPTER 2 LITERATURE REVIEW	3
2.1 The BCS System	3
2.1.1 Harris Interactive College Football Poll	3
2.1.2 USA Today/ESPN Coach's Poll	4
2.1.3 Anderson & Hester	4
2.1.4 Richard Billingsley	5
2.1.5 Kenneth Massey	5
2.1.6 Jeff Sagarin's USA Today	5
2.1.7 Peter Wolfe	5
2.1.8 Colley Matrix.....	6
2.1.9 The BCS Formula.....	6
2.2 Related Work.....	6

2.3 Description of the Analytic Hierarchy Process	9
CHAPTER 3 METHODOLOGY.....	14
3.1 Basic Idea	14
3.2 Formalization of Methodology.....	19
3.2.1 Input Processing	19
3.2.2 The Conversion Function	21
3.2.3 Path Search Algorithm.....	26
3.2.4 Compute the Weighted Pair-wise Comparison.....	29
3.2.5 Minimize the Error between the Actual Output and the Estimated Output	31
3.2.6 Rank the Teams on Quality.....	32
CHAPTER 4 RESULTS AND DISCUSSION	33
4.1 Implementation.....	33
4.1.1 Data Collection.....	33
4.1.2 Data Preprocessing.....	35
4.1.3 Setup for the Parameter C	38
4.1.4 Computation of the Weighted Quality Comparison Values between Teams ...	38
4.1.5 Optimizing the Prediction.....	40
4.1.6 Rank the Teams with Expert Choice	42
4.2 Results.....	45
CHAPTER 5 CONCLUSION.....	48

REFERENCE	50
APPENDIX A	52
CURRICULUM VITAE	53

CHAPTER 1 INTRODUCTION

The NCAA Football Bowl Subdivision(FBS) (formerly known as **Division I-A**) is the only NCAA-sponsored sport without an organized tournament to determine its champion. Schools in Division I FBS compete in post-season, with the champions of six conferences receiving automatic bids to the highly lucrative bowl games.

The **Bowl Championship Series (BCS)** is a selection system designed to give the top two teams in the FBS an opportunity to compete in a "national championship game". The BCS relies on a combination of polls and computer selection methods to determine relative team rankings, and to narrow the number of teams involved to two teams to play in the national champion bowl games held after the other college bowl games. The winner of this game is crowned as the BCS national champion, and is guaranteed at least a share of the national championship.

Before the BCS was introduced in 1998, Division I-A was based solely on two opinion polls, commonly referred to as the Coaches' and the AP(Associated Press) polls. The rankings of the teams in these two polls is determined by coaches and sports journalists respectively, who vote weekly for the top 25 teams based on the team's performances in all games played.

However, when the Coaches' poll and the AP poll could not make an agreement on the choice of the top team, a shared title of national champion could be assigned to two teams, which has occurred in 1997 with Michigan and Nebraska. In order to overcome this problem, from 1998 the new system was introduced, under which the top two teams at the end of season would play one final game for the championship. This system is called BCS, and it combines the AP poll, the Coaches' poll with several computer-based polls to neutralize the effects of the AP and Coaches'

polls, which represent expert opinion criticized by the public for the following two reasons: First the human pollsters are not objective observers and may have biases on certain schools based on their own perception; second it is impossible for a human pollsters to recall all outcomes of all games in the season, even if he or she watched all games.

Although the computer polls can overcome the above two weaknesses, they also have already been proven to be controversial in practice. Sometimes the computer polls could produce the ranking which differs significantly from the two human polls. For example, in 2001 University of Oregon got the second place in both the AP and Coaches' polls, but it got the 8th place in one of the computer polls and 7th place in two of the other computer polls. The low ranking of Oregon in these three computer polls is analyzed to be contributed to Oregon's many victories with narrow point margins(score difference), since the computer models in the BCS do not account for margin of victory to rank the teams involved. These kinds of instances happened many times due to the ignorance on margin of victory of games in the computer polls. The reason to ignore margin of victory in the computer polls is that the ranking result could be more consistent with the human polls if the computer polls only account a team's winning percentage and quality of opposition than consider the margin of victory between teams. As a negative result, this could lower the incentive for a team to earn more scores in games.

In this thesis, we propose a new computer model to weight the margin of victory in a holistic view, based on Graph theory and the Analytic Hierarchy Process(AHP). We consider a set of games as a graph, in which the teams are connected with games, and the margin of victory of games can indicate the quality comparison between teams; we compare the teams holistically based on the transitivity along the paths in the graph, and finally make pair-wise comparison between teams to rank the teams with AHP.

In Chapter 2, we will review the related literature concerning computer models including a detailed explanation of the BCS, and introduce some necessary knowledge on AHP. In Chapter 3, we introduce our methodology. In Chapter 4, the implementation of our methodology with an instance is given. Finally there are discussions and conclusions in Chapter 5.

CHAPTER 2 LITERATURE REVIEW

There are a variety of methods that have been developed for the rankings of athlete teams. In this chapter, first we introduced the currently used BCS ranking system; second the related other methods are introduced; finally AHP is introduced as a basis for our method.

2.1 The BCS System

The Bowl Championship Series (BCS)[1] is a ranking system to decide which two teams in the NCAA Football Bowl Subdivision (FBS) have the opportunity to compete in a national championship game. According to the BCS Formula, the BCS ranking score is calculated from the combination of 2 polls involving human elements (including Harris Interactive College Football Poll and USA Today/ESPN Coach's Poll) and 6 polls of computer elements (including Anderson & Hester, Richard Billingsley, Colley Matrix, Kenneth Massey, Jeff Sagarin's USA Today and Peter Wolfe). In the subsections of this section, each poll is introduced, finally the BCS Formula is described.

2.1.1 Harris Interactive College Football Poll

The Harris Interactive College Football Poll (Harris Poll)[2] is a weekly rankings of the top 25 NCAA Division I Football Bowl Subdivision (FBS) college football teams, and this poll is compiled by Harris Interactive, a market research company that specializes in Internet research.

To form the panel of Harris Poll composed by former players, coaches, administrators and current and former media, Division I-Bowl Subdivision college football conferences and

independent institutions firstly nominate 300 prospective panelists; then Harris Interactive selects 114 members from the nominees to participate in the actual panel.

The 114 members vote on the rankings of top 25 teams from each week in September to the end of the regular college football season, and Harris Interactive releases the ranking results weekly. There is no published information about the mechanism of voting inside Harris Interactive in Internet.

2.1.2 USA Today/ESPN Coach's Poll

The USA Today/ESPN Coach's Poll (Coach's Poll)[3] is made up of 61 head coaches at Division I FBS institutions, and all of them are members of the American Football Coaches Association. Similarly, the Coach's Poll also ranks the top 25 teams in FBS weekly, and the detailed mechanism of voting is not published. As a voting system made by a fewer of the coaches compared to Harris Poll, the Coach's Polls are under criticism due to its inaccuracy, subjectivity and bias[8].

2.1.3 Anderson & Hester

Jeff Anderson and Chris Hester are the owners of this computer ranking system, and the system has been a part of the BCS since its inception. The algorithm of the ranking system is unpublished as a secret. The Anderson & Hester Rankings are distinct in four ways according to its declaration in its official website [4]:

1. Unlike the polls, **these rankings do not reward teams for running up scores**. Teams are rewarded for beating quality opponents, which is the object of the game. Margin of victory, which is not the object of the game, is not considered.
2. Unlike the polls, **these rankings do not prejudge teams**. These rankings first appear after the season's fifth week, and each team's ranking reflects its actual accomplishments on the field, not its perceived potential.

3. These rankings compute the most accurate strength of schedule ratings. Each team's opponents and opponents' opponents are judged not only by their won-lost records but also, uniquely, by their conferences' strength (see #4).

4. These rankings provide the most accurate conference ratings. Each conference is rated according to its non-conference won-lost record and the difficulty of its non-conference schedule.

2.1.4 Richard Billingsley

Richard Billingsley developed this method[5], the main components of which are: Won-Loss Records, Opponent Strength (based on the opponent's record, rating, and rank), with a strong emphasis on the most recent performance. However, it does not consider the margin of victory.

2.1.5 Kenneth Massey

Kenneth Massey described his method in his official website[6]. Only the score, venue, and date of each game are inputs to the Massey ratings algorithm. Every pair of scores of one game is converted as a value between 0 and 1 that estimates the probability that the winner of the game would win a rematch under the same conditions. The power rating representing the potential of a team is calculated from the probabilities through a Bayesian win-loss correction.

2.1.6 Jeff Sagarin's USA Today

The computer system published in USA Today was developed by Jeff Sagarin[7]. This system takes winning and losing as the sole factors, which makes it very "politically correct". However, this also makes it less accurate in predictions for upcoming games.

2.1.7 Peter Wolfe

Peter Wolfe[8] developed his ranking method called a maximum likelihood estimate. In it, each team i is assigned a rating value π_i that is used in predicting the expected result between it and its opponent j , with the likelihood of i beating j given by:

$$\pi_i / (\pi_i + \pi_j)$$

The probability P of all the results happening as they actually did is simply the product of multiplying together all the individual probabilities derived from each game. The rating values are chosen in such a way that the number P is as large as possible.

2.1.8 Colley Matrix

The method Colley Matrix [9] only considers wins and losses as input, and adjusts strength of schedule in a way that is free of bias toward conference, tradition, or region. The rating of a team is calculated from the numbers of winning games and number of losing games, and the strength of its components.

2.1.9 The BCS Formula

Finally, the BCS ranking score of a team is calculated according to the following formula:

$$BCS \text{ Ranking} = (Harris \text{ Poll } \% + Coach's \text{ Poll } \% + Computer \text{ Poll } \%)$$

“Harris Poll%” is the percentage that is calculated from the score of the team in the Harris Poll divided by the perfect score of the Harris Poll. Similarly, “Coach’s Poll%” is the percentage that is calculated from the score of the team in the Coach’s Poll divided by the perfect score of the Coach’s Poll. For example, in the Harris Poll, if Penn State has a score of 2631, and the perfect(full) score a team can earn in the Harris Poll is 2850, then $2631 \div 2850 = 0.923$ is Penn State’s “Harris Poll%”

In order to compute “Computer Poll%” of a team, among the team’s six scores of the six computer polls, the lowest score and the highest score should be taken off, the remaining scores are summed up; then the “Computer Poll %” is the sum of the four scores divided by 100.

2.2 Related Work

Many mathematicians and computer scientists have proposed alternative methods for improving team rankings.

James P. Keener[10] developed a method based on the Perron-Frobenius theorem, in which the strength of teams is expressed as a vector r , with the positive components r_j indicating the strength of the j th team. The score of team i is defined as

$$s_i = \frac{1}{n_i} \sum_{j=1}^N a_{ij} r_j$$

where a_{ij} is a nonnegative evaluation on the outcome of the game between team i and team j , N is the total number of teams in the competition, and n_i is the number of games played by team i . The matrix A with the entries $\frac{a_{ij}}{n_i}$ is called a preference matrix based on the paired comparison between teams, with the assumption that the strength of a team should be proportional to its score, that is,

$$Ar = \lambda r$$

where the ranking vector r is a positive eigenvector of the matrix A , and according to the Perron-Frobenius theorem, there exists an eigenvector r with nonnegative entries with a corresponding eigenvalue λ . Then the different methods of extracting the preference matrix from uneven and incomplete paired comparison between teams are discussed and developed.

Rick L. Wilson[11] developed a ranking method with a neural network approach: the teams and the games between teams form a graph, in which a node represents a team and an edge represents a game between the two teams represented by two nodes; and the graph is understood as a neural network, in which every node is taken as a neuron, every edge connecting with the neuron is taken as an input to it, and the value of the neuron representing the team strength is calculated with the transfer function through running an iterative program.

David Harville[12] developed a statistical regressive model for predicting the score difference of the game between two teams, by which the characteristic level of performance of

the teams over a long sequence of games is extracted from the random variation of performance shown in games, then the teams are ranked with the characteristic level of performance.

Similarly as [12], [13] takes the outcome of games between two teams as a normal distribution based on the strength of the teams over time, and developed a predictive model.

In the paper [14], the ranking problem is taken as a quadratic assignment problem: every team is assigned to a position in the ranking list, which is supposed to maximize the multiplication of the degree of victory in scores and the relative distance in the ranking list between any two teams.

R.T.Stefani[15] developed a predictive approach using least squares based on the following model:

$$w = Ar + v$$

Where w is the victory margin of vectors ($M \times 1$) of M games, A is the selection matrix with entries a_{ki} or a_{kj} (for a game k between team i and j , if team i defeat team j , $a_{ki} = +1$ and $a_{kj} = -1$; if the result is tie, $a_{ki} = 0$ and $a_{kj} = 0$), r is the rating vector on N teams, and v is the random variable vector on victory margins. David A. Harville[16] and Gilbert W.Bassett[17] gave modified versions of ranking method using least square with considering margin of victory.

David Mease[18] proposed a model using a penalized likelihood approach which results in a ranking process that attempts to mimic the thought processes of the human pollsters of the Coaches' polls. The underlying assumption is that every team keeps a reliable performance level, and its actual performance is its performance level plus the performance variation. Based on the model, the algorithm try to calculate the reliable performance level to fit the game outcomes the best, and use performance levels to rank the teams.

[19] proposed a network-based ranking system for US college football: similar to our work, it forms a network graph in which every team is a vertex and every game is an directed edge from

the winner to the loser; differently from our work, it only counts the direct and indirect win/loss to get scores of teams without considering margins of victory.

[20] proposed a method based on agent-based simulation: a set of voters are simulated as walkers in a network structure representing win/loss relationship among teams, the output of a game will change the walker's favor on the participant teams with some probability. Finally, the team who get more favors ranks high.

[21] proposed a ranking method with AHP. However, the evaluation on pair-wised comparison is very coarse and even unreasonable compared to our work.

[22] takes the ranking problem as an inverse equal paths problem, and developed a network based ranking system similar to [19].

2.3 Description of the Analytic Hierarchy Process

The Analytic Hierarchy Process (AHP)[23-25] is a structured approach to solve complex decisions in the economic, social and management sciences with inconsistencies and subjectivity. Instead of outputting an absolutely correct solution, AHP can help the decision makers identify their needs and objectives, and understand the problems more deeply. Based on some mathematical and psychological facts, it was developed by Thomas L. Satty in 1970s, and was implemented by the software package "Expert Choice"; and since then it has been studied extensively and applied widely. According to the past practical experience, it can be applied in the following situations:

- Choice: select one preferred alternative from a given set of alternatives, in which the alternatives are evaluated with multi-criteria.
- Ranking: evaluate a set of alternatives with multi-criteria, and sort out an order of alternatives following the general preference from most to least.
- Prioritization: determine the relative merit of a set of alternatives according to one or several criteria, in which the relative merits of alternatives in terms of the criteria are given as values.

- Resource Allocation: apportion resources to a set of alternatives according to some criteria such as emergency, importance and so on.
- Benchmarking: compare one alternative such as one solution, or operational performance of a company with those excellent alternatives, to identify the gap between the existing situation and the ideal situation.
- Quality management: integrate diversified quality measurements and quality control technique together to deal with the multidimensional aspects of quality and quality improvements.

AHP is a systematic approach, which decomposes the problem into a hierarchy and a series of pair-wise comparisons, finally the hierarchy and pair-wise comparisons are integrated with quantitative techniques. The application of AHP is supposed to follow the following basic steps:

- (1) Develop a hierarchy of objectives which consists of factors such as a general goal, criteria, sub-criteria and a set of alternatives, as shown in Figure 1. In this hierarchy, the top level defines the general goal of the problem or the system; the bottom level lists the alternatives which are to be evaluated, selected or ranked; in between the top level and the bottom level, there are several levels of criteria, sub-criteria, ..., and so on, which are obtained from the decomposition of the general goal and criteria, and every level is exactly non-overlapping partitions of the higher level.
- (2) Make pair-wise comparisons at each level. The comparisons have to do with how much more important one criterion/sub-criterion is than another criterion/sub-criterion with respect to their parent criteria, when the criteria are compared; and the comparisons have to do with how much better one alternative is than another, with respect to the criterion by which the alternatives are evaluated. In Figure 1, for the first level of criteria, the pair-wise comparisons should be made between criterion 1, criterion 2 and criterion 3 with respect to the goal; the comparisons also should be made between sub-criterion 1 and sub-criterion 2 with respect to criterion 1, but no comparisons should be made

between sub-criterion 2 and sub-criterion 3, because they have no common parent criteria.

Table 1 Numerical Rating in Pair-wise Comparison

Importance	Numerical Rating
Equal Importance	1
Weak Importance	3
Strong Importance	5
Very Strong Importance	7
Absolute Importance	9
2, 4, 6, 8 are intermediate values	

Performance	Numerical Rating
Equal Performance	1
Weakly Better	3
Strongly Better	5
Very Strongly Better	7
Absolutely Better	9
2, 4, 6, 8 are intermediate values	

And the answers given to the questions for the pair-wise comparisons are converted to numerical code, as shown in Table 1.

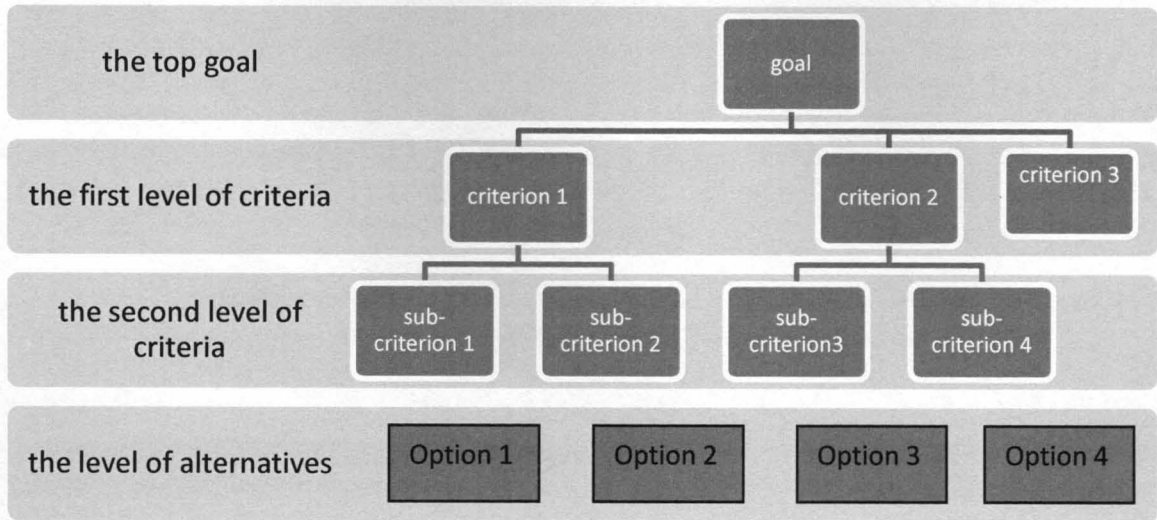


Figure 1. The Hierarchy of AHP

- (3) The pair-wise comparisons with respect to every goal, criterion or sub-criterion form an “influence matrix”, denoted as A , a square matrix with the number of rows and the number of columns equal to the number of criteria in the particular section of the hierarchy with respect to the goal, criterion and sub-criterion. The A matrix is used to determine the “local weight” for each factor, through using the “eigenvalue method” as following:

$$Aw = \lambda_{\max} w$$

$$\sum w_i = 1$$

where w is the n -dimensional eigenvector associated with the largest eigenvalue, λ_{\max} , of the influence matrix, A . λ_{\max} is determined by selecting the largest value of λ that satisfies the equation given by:

$$|A - \lambda I| = 0$$

The influence matrix $A = (a_{ij})$ is given by:

- The numerical rating provided by step (2), if i is more important or performs better than j ;
- $a_{ij} = 1$, if $i = j$;
- $a_{ij} = 1/a_{ji}$, if j is more important or performs better than i ;

The larger the “local weight” of a factor is, the more important or the better of performance the factor is, with respect to its parent factor. And if the numerical rating is perfectly consistent in the pair-wise comparisons, the influence matrix A will satisfy:

$$a_{ij} = \frac{w_i}{w_j}$$

Therefore, it can satisfy a transitive relationship as the following:

$$a_{ik} = a_{ij} * a_{jk} \text{ if there exist the comparison between } i \text{ and } j, j \text{ and } k.$$

But this is not the requirement AHP can work; on the contrary, AHP can deal with inconsistency due to human’s judgments or exceptions in the system.

- (4) Determine the global weight (also called global priority) for each alternative in the hierarchy. The global weight of an alternative represents the overall score of that alternative, and this score indicates how better the alternative is compared to the other alternatives.

CHAPTER 3 METHODOLOGY

3.1 Basic Idea

The basic idea of our methodology is to make pair-wise comparisons between teams in terms of team quality according to their game performances, give numerical rating based on the score difference of the game between the two compared teams, form an influence matrix and calculate the overall weight using AHP. As shown in Figure 2, the AHP hierarchy for ranking football teams has only two levels: the top level “team quality” is not only the general goal but also the unique criterion; the bottom level consists of the teams to be ranked in term of team quality as the alternatives. The AHP gives the global weight of every team; the more weight a team has, the better the quality of the team.

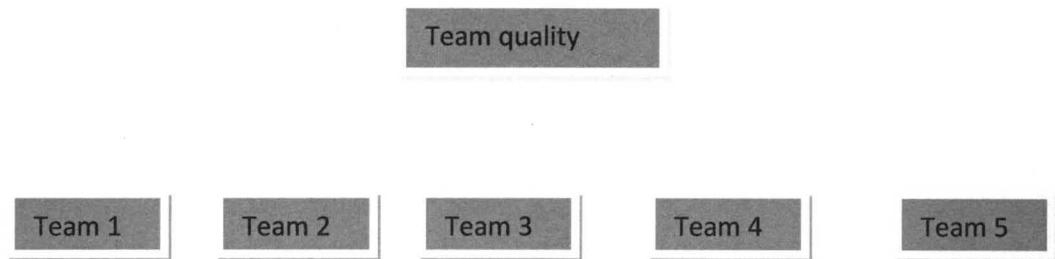


Figure 2. The Hierarchy of Team Ranking

In order to fulfill our idea, several problems must be discussed and solved:

- (1) Convert score difference to quality difference

If there exists a game between team A and team B, and if team A has beaten team B by a score of x to y (A wins over B , and $x > y$), how can we convert the score difference

$x - y$ to the quality difference which indicates the numerical rating representing the comparison of quality of two teams? We should construct a conversion function: the input is the score difference of one game between two teams, and the output is a number between 1 and 9, in which 1 indicates team A and team B perform equally well on team quality, and 9 indicate team A performs absolutely better than team B. Table 2 gives a more detailed description on the output:

Table 2 The Numerical Rating in Our Problem

Evaluation	Numerical Rating
1	A and B Perform Equally Well
3	A Performs Marginally Better than B
5	A Performs Markedly Better than B
7	A Performs Very Markedly Better than B
9	A Performs Absolutely Better than B

Any numbers between 1 and 9 represent intermediate values

The quality difference (DQ) of two teams is a non-decreasing, nonlinear function of the score difference (DS) as Figure 3 shows, when the two teams play a game. The function, $DQ = f(DS)$, has a non-increasing first derivative ($d(DQ)/d(DS)$) as DS increases in value. For example, if team A beats team B by 40 points, and team B beats team C by 40 points, then team A would be expected to beat team C not by 80 points, but by some number of points between 40 and 80.

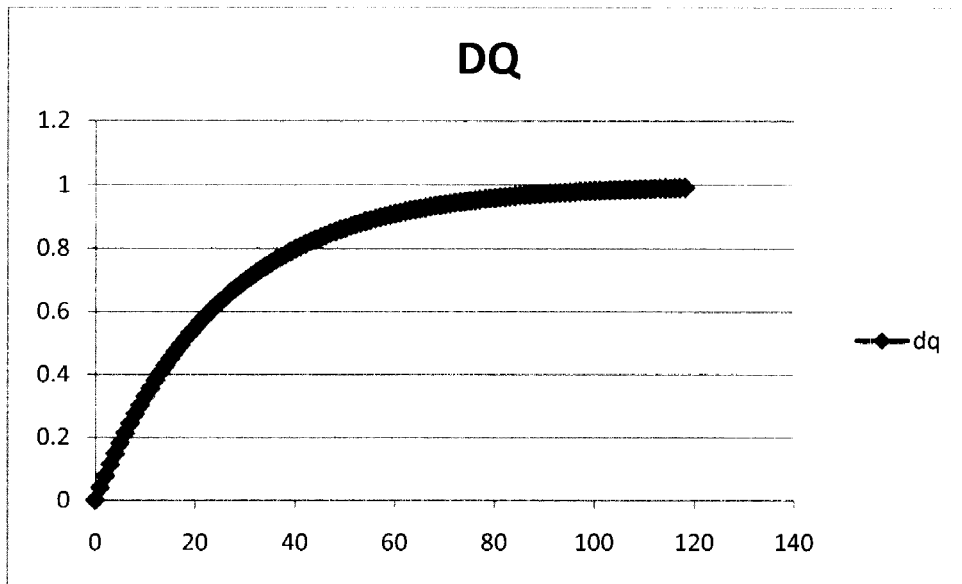


Figure 3. The Conversion Function from Score Difference to Quality Difference

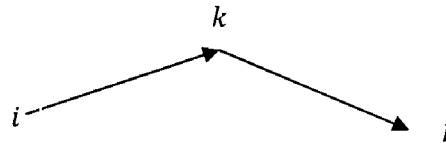
(2) An estimation method based on graph and transitivity

The ideal situation for team pair-wise comparisons is a Round-Robin Tournament in which each team plays a game with each of the other teams at least once. In a Round-Robin Tournament, we can obtain all the quality differences between the teams directly based on the score difference of the games they have played. However, in many schedules, every team only plays games with a subset of all team members; thus, there exist such situations that two teams have never played each other, and we cannot make a direct comparison on the team quality between these two teams.

A solution to deal with this problem is that we can make pair-wise comparisons indirectly with the transitivity relationship indicated in the AHP between two teams who have played no games with each other. For example, in the following diagram, team i and team j have no direct game played, but team i has beaten team k by a quality comparison value a_{ik} , and team k has beaten team j by a quality comparison value a_{kj} ; we can apply the

transitivity relationship indicated in AHP to get the indirect quality comparison between team i and j as following:

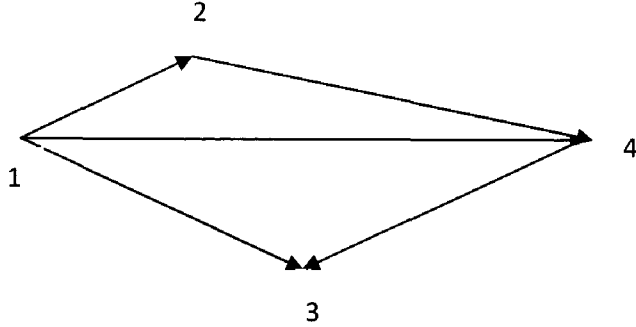
$$a_{ij} = a_{ik} * a_{kj}$$



However, transitivity relationship assumes “perfect consistency”, and unfortunately the perfect consistency cannot be guaranteed in AHP applications including ours. This inspired an estimation method based on graph theory: a set of games form a graph, in which every vertex represents a team, and every directed edge with a quality comparison value corresponds to a game, and the direction of the edge indicates the team of the start vertex wins over the team of the end vertex. The quality comparison value on the edge is calculated from the score difference with our conversion function ; in order to calculate the quality comparison value including direct comparison and indirect comparison between two teams, all paths between these two teams are found, the quality comparison for every path is calculated according to the transitivity relationship and is assigned a weight, and finally a value integrating these quality comparisons with these weights is calculated as a weighted quality comparison value.

For example, suppose that there are 4 teams denoted as vertex 1, 2, 3 and 4, and 5 games are played among these teams. This forms a graph shown in the following diagram; there exist 3 paths from team 1 to team 4: $1 \rightarrow 2 \rightarrow 4$, $1 \rightarrow 4$, and $1 \rightarrow 3 \rightarrow 4$; although the direction of the edge between 3 and 4 is from 4 to 3, we can convert the

edge with the quality comparison value a_{43} to an edge with the quality comparison value $1/a_{43}$ directed from 3 to 4.



In the following, NP_{ijm} is the m th quality comparison value along the m th path between team i and team j

For path $1 \rightarrow 4$, we can get the quality comparison value

$$NP_{141} = a_{14}$$

For path $1 \rightarrow 2 \rightarrow 4$, we can get the quality comparison value

$$NP_{142} = a_{12} * a_{24}$$

For path $1 \rightarrow 3 \rightarrow 4$, we can get the quality comparison value

$$NP_{143} = a_{13} * (1/a_{43})$$

Finally, the above quality comparison values are integrated as a weighted quality comparison value with given weights w_1 , w_2 and w_3 , and this weighted quality comparison value A_{14} would be taken as the corresponding entry of the influence matrix A in AHP.

$$A_{14} = w_1 * NP_{141} + w_2 * NP_{142} + w_3 * NP_{143}$$

And the condition is held at the same time:

$$w_1 + w_2 + w_3 = 1$$

(3) Home field advantage

The home field advantage must be considered when one team plays a game in its home city with another team. The home field advantage is measured in terms of expected extra points the home team can obtain. For example, if we set up the home field advantage as 3 points, team A has beaten team B by 38 points, and we suppose team A is at home, then team A won team B by 35 points after the home field advantage is considered; if team B is at home, the score difference should be adjusted as 41 points.

The general procedures to apply our methodology can be summarized as following:

- (1) Identify the conversion functions including the values of its corresponding parameters.
- (2) Seek the values for the parameters including the home field advantage and the weighting factors to weight the quality comparisons along different paths.
- (3) Compute all the weighted pair-wise comparisons on team quality between teams according to the conversion function and the corresponding parameters, form the influence matrix A and input it into AHP software package to make a ranking in terms of team quality.

3.2 Formalization of Methodology

In this section, we will refine our ideas, and give a detailed description of our methodology.

3.2.1 Input Processing

The input to our method is a set of tuples, every tuple specifies a game between two teams, and every tuple has the following format:

$\langle \text{winner}, \text{loser}, \text{winScore}, \text{losScore}, \text{homeTeam} \rangle$

winner : ID of the winning team in the game.

loser : ID of the losing team in the game.

winScore: score of the winning team.

losScore: score of the losing team.

homeTeam: ID of the home team.

And if the game result is a tie($winScore = losScore$), the *winner* or the *loser* just specify one of two teams in the game, and the order of them does not matter.

The input is simplified as the following format:

$\langle winner', loser', DS \rangle$

winner': ID of the winning team in the game.

loser': ID of the losing team in the game.

DS: the score difference after consideration of the home field advantage($DS \geq 0$).

The simplified tuples can be computed with the following algorithm:

Input: a set of tuples with the format $\langle winner, loser, winScore, losScore, homeTeam \rangle$.

Output: a set of simplified tuples with the format $\langle winner', loser', DS \rangle$.

Method: H is the number of points associated with the home field advantage.

if $homeTeam == winner$,

$diff = winScore - losScore - H$;

if $homeTeam == loser$,

$diff = winScore - losScore + H$.

if $diff \geq 0$,

$winner' = winner$,

$$\begin{aligned}
& \text{loser}' = \text{loser}; \\
& \text{if } \text{diff} < 0, \\
& \quad \text{winner}' = \text{loser}, \\
& \quad \text{loser}' = \text{winner}. \\
& DS = |\text{diff}|.
\end{aligned}$$

The set of simplified tuples can be easily interpreted or converted to a graph, in which every vertex is the ID of a team, and every edge weighted by the score difference DS (after consideration of home field advantage) represents a game between the two connected vertices of teams. We define the graph formally in the following:

$$G = \langle V, E \rangle$$

where V is the set of all vertices representing teams; E is the set of all edges representing games between teams.

3.2.2 The Conversion Function

We will explore the conversion functions with two steps: in the first step, we will devise a function which can convert every score difference to a corresponding value between 0 and 1; in the second step, we will devise a linear function which can map the value between 0 and 1 to a value between 1 and 9 representing the quality comparison between teams in terms of the AHP.

For the first step, we apply the following function on the score difference DS :

$$DQ_{ij} = 1 - e^{-C \cdot DS_{ij}} \quad (1)$$

DS_{ij} : the score difference of a game between team i and team j .

C : a parameter the value of which is to be determined.

DQ_{ij} : the quality difference as determined by a game between team i and team j ,

$0 \leq DQ_{ij} < 1$.

This function can satisfy all the requirements we need: the output value ranges from 0 to 1, we can easily use a linear function to map it onto a value between 1 and 9; when DS becomes larger, DQ approaches 1 and the derivative $d(DQ)/d(DS)$ becomes smaller and approaches 0. Figure 4 illustrates one possible graph.

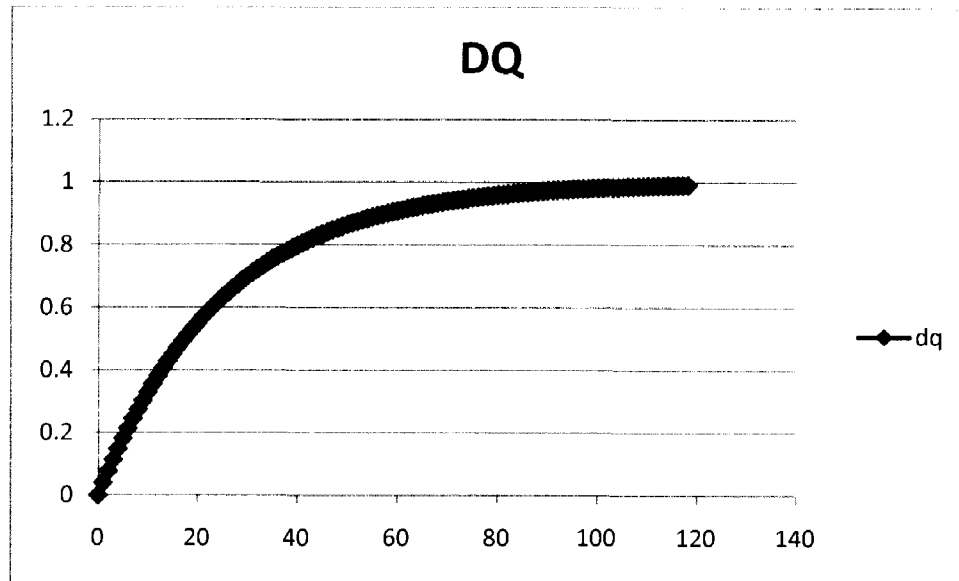


Figure 4. The Curve of the Conversion Function when $C = 0.04$

For the second step, we use the following linear function to map DQ to a value between 1 and 9 as the quality comparison:

$$a_{ij} = 1 + 8DQ_{ij} \quad (2)$$

Figure 5 shows the graph of this function.

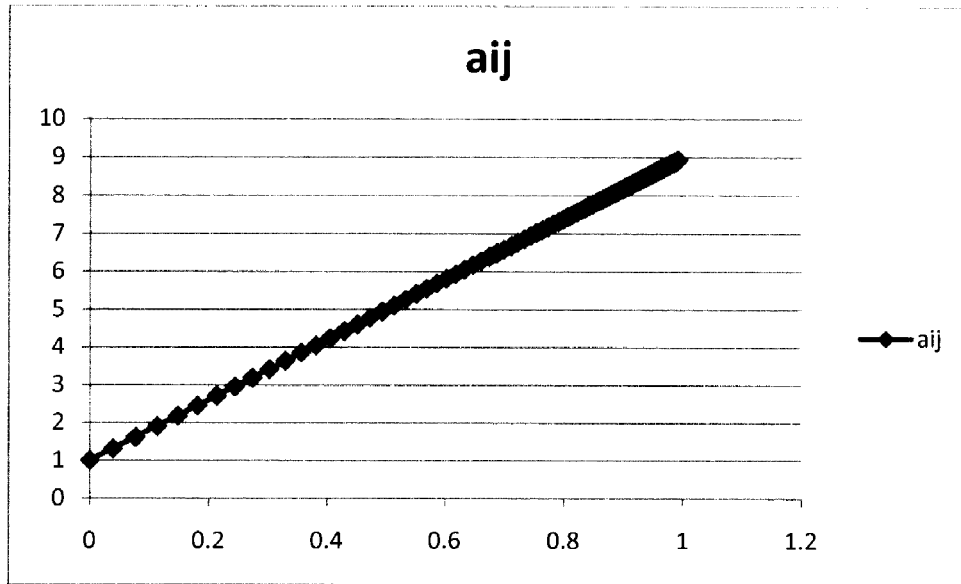


Figure 5. The Linear Function from D_{ij} to a_{ij}

As the composition of Equation (1) and (2), we can have:

$$a_{ij} = 9 - 8e^{-t DS_{ij}} \quad (3)$$

Figure 6 shows the graph of the composed function.

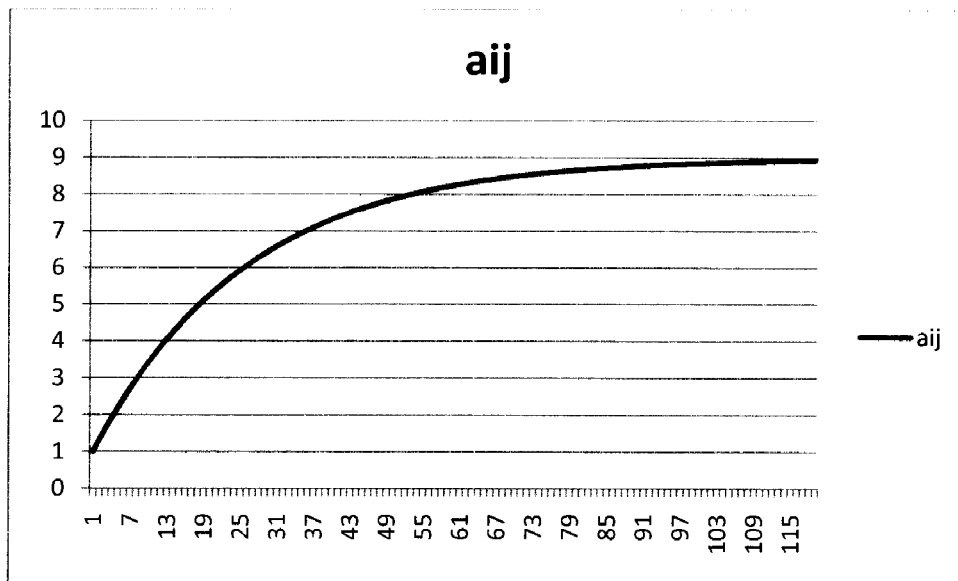


Figure 6. The Graph of the Composed Function

In order to use this function, we must determine a value for C . It is reasonable to assume that the value of C is determined by how many score difference corresponds to the statement “A Performs Absolutely Better than B”. The minimal score difference considered as “A Performs Absolutely Better than B” is a Threshold value, denoted as T . We can determine the value of C with the following formula:

$$9 - a_{ij} = 8e^{-C \cdot T} = P \quad (4)$$

Then we can have

$$C = -\frac{1}{T} \ln \frac{P}{8} \quad (5)$$

P is the precision which is a very small value. If $P = 0.01$, and we assume 70 is the threshold value, then $C \approx 0.095$. Figure 7 shows the curve of the composed conversion function when $C = 0.095$. And once the threshold for “A Performs Absolutely Better than B” is determined, the other threshold for others numerical rating will be determined automatically.

After all DS are converted to DQ , the graph G is also converted to a new graph consisting of (V, E') , in which every vertex still represents a team, and every edge represents a game between two connected teams indicating a quality comparison in terms of the AHP.

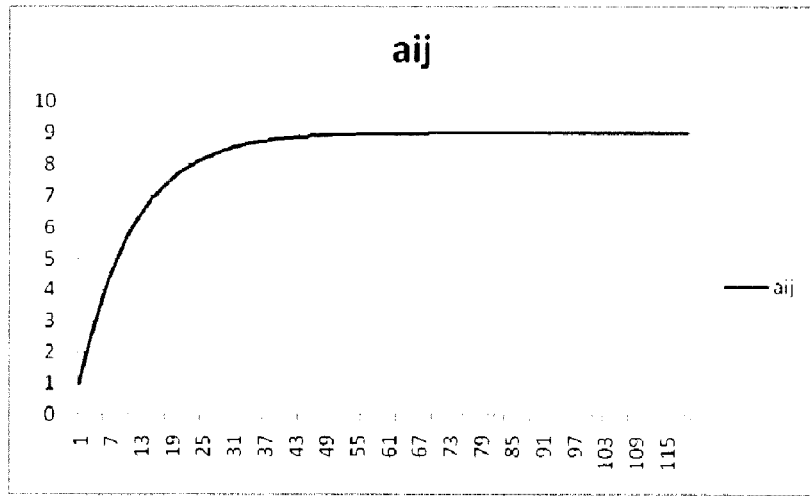


Figure 7. The Graph of Composed Conversion Function when

3.2.3 Path Search Algorithm

In order to compute the weighted pair-wise comparison on quality between two teams i and j , we must find all paths between i and j . A path in our problem is defined as: a finite sequence of vertices without repeated vertices (non-cyclic) such that from each of its vertices there is an edge to the next vertex in the sequence. In our problem, the direction of an edge does not matter, because it can be transformed to a reverse-directed edge with the inverted weight. The length of a path is defined as the number of edges the path has, and we call a path whose length is k as a k -segment path.

To guarantee that we can make pair-wise comparison between any two teams, the graph G must be connected, in which there is a path from one vertex to any other vertex. The following algorithm written as pseudo-code can validate the connectivity of a graph:

Input: a graph $G = \langle V, E \rangle$

Output: if the graph is connected.

Method:

1. *Begin at any arbitrary node of the graph, G .*
2. *Proceed from that node using either depth-first or breadth-first search, counting all nodes reached.*
3. *Once the graph has been entirely traversed, if the number of nodes counted is equal to the number of nodes of G , the graph is connected; otherwise it is disconnected.*

Given two teams i and j , in order to compute the weighted pair-wise comparison, we must find all the paths from vertex i to vertex j . Practically the comparison along the longer path have weaker weight in the weighted comparison, therefore we only count those paths within length k . For example, if $k = 4$, we only count those paths with 1, 2, 3 and 4 edges. This can simplify the path searching algorithm: with a graph $G = \langle V, E \rangle$, form a searching tree which starts from vertex i as the root node and travels along the edges in E without cycles (a vertex can only be visited once) until the length of the path from the root to the leaf nodes reaches to k or vertex j is visited; any path from the root node to the leaf nodes of vertex j forms a path for which we are looking. Figure 8 shows a set of games among 8 teams in the first half of the season; Figure 9

show a searching tree for comparison between team 1 and team 2 when $k = 4$, the paths from the root node to the green leaf are the paths which meet the requirements for this pair-wise comparison. The searching algorithm can be described as following:

Input: Graph $G = \langle V, E \rangle$, the starting vertex i , the end vertex j , the length limitation k .

Output: a set of paths from vertex i to vertex j without any cycle whose length is no more than k .

Method:

- 1. Proceed from vertex i , visit the adjacent vertices of vertex i , the adjacent vertices of the adjacent vertices of vertex i , ..., form a traversal tree from vertex i to the visited vertices, until the length of the path from the root to a leaf reach to k or the visitation meets vertex j .*
- 2. Collect the qualified paths as the path set for the pair-wise comparison between i and j .*

In Figure 9, the path set for pair-wise comparison between 1 and 2 consists of the following paths:

1→2

1→4→3→2

1→4→3→5→2

1→4→8→2

1→7→8→2

1→7→6→3→2

1→7→6→5→2

1→8→4→3→2

1→8→2

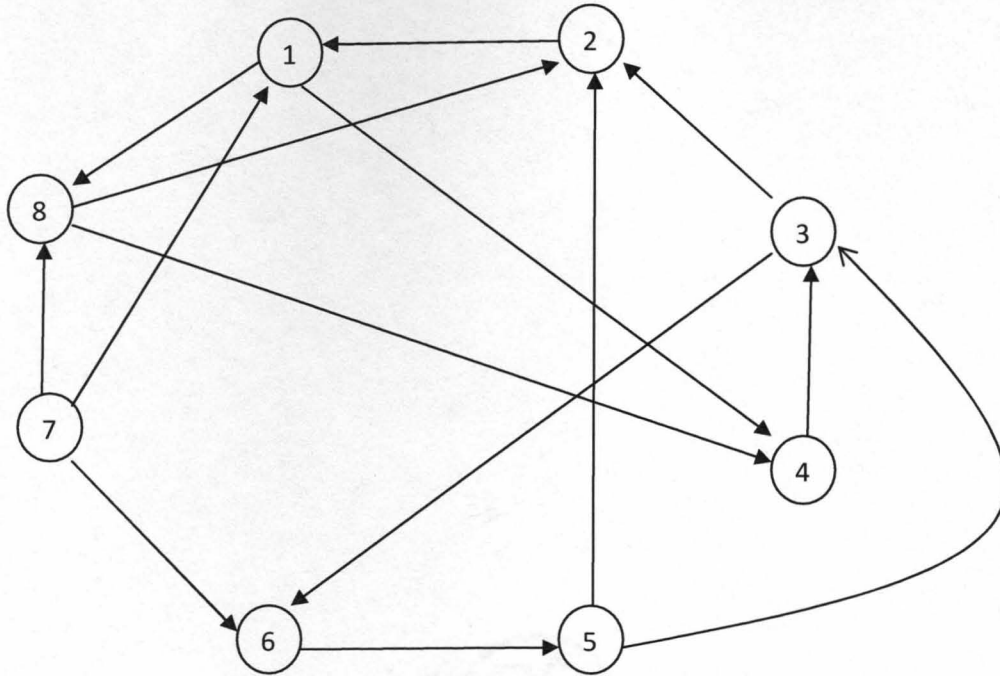


Figure 8. A Graph showing a Set of Games

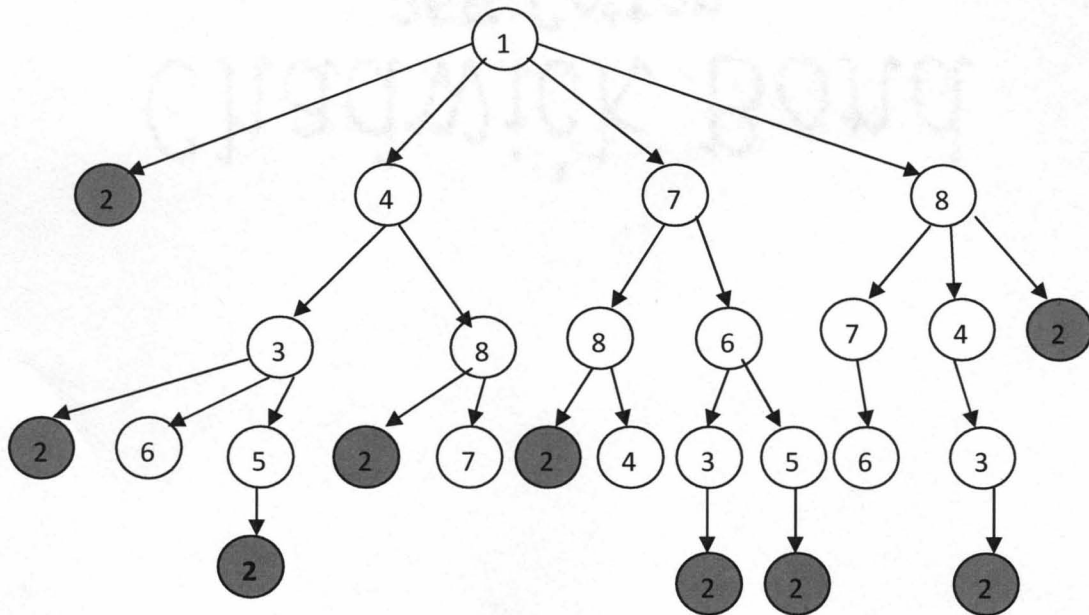


Figure 9. The Searching Tree from Team 1 to Team 2 when $k=4$

3.2.4 Compute the Weighted Pair-wise Comparison

For every qualified path computed from the above algorithm, according to the approaches described in the AHP, we can calculate a comparison value NP_{ijs}^l on team quality based on transitivity, i and j are IDs of the two teams, l is the length of the path ($l = 1, 2, 3, 4, \text{if } k = 4$), s is the index of the path (there might exist more than one path which have the same length), NP_{ijs}^l is the pair-wise comparison value on quality of comparing team i to team j along path s with l segments.

For pair-wise comparison between team i and j , we can form a set of values $\{NP_{ijs}^l\}$ based on the path set for pair-wise comparison between team i and j . In order to integrate and neutralize the set of values into one value A_{ij} , we must assign weights v_{ijs}^l to different NP_{ijs}^l , and they satisfy the following relationships:

$$A_{ij} = \sum v_{ijs}^l * NP_{ijs}^l$$

And $\sum v_{ijs}^l = 1$

There might exist different configurations of paths for one pair-wise comparison. For example, for one pair-wise comparison, the path set can contain several 1-segment paths, 2-segment paths, 3-segment paths and 4-segment paths; but for another pair-wise comparison, the path set can possibly contain several 2-segment paths and 3-segment paths; ... ; and so on. Table 3 shows the different configurations for path sets when $k = 4$, in which 1_2_3 means there only exist 1-segment, 2-segment, 3-segment paths in the path set.

For simplification, we assign the same weight to the paths with the same length in a path set. As shown in Table 2, in a path set which has the configuration 1_2_3, we assign the weight V_{22} to all paths with length 2; and every row of weights sum up to 1.

Table 3 The Weighting Vectors for Different Paths Set

Configuration(\no seg)	1	2	3	4
1_2_3_4	V11	V12	V13	V14
1_2_3	V21	V22	V23	0
2_3_4	0	V32	V33	V34
1_2_4	V41	V42	0	V44
1_3_4	V51	0	V53	V54
1_4	V61	0	0	V64
1_2	V71	V72	0	0
1_3	V81	0	V83	0
2_3	0	V92	V93	0
2_4	0	V102	0	V104
3_4	0	0	V113	V114
1	V121	0	0	0
2	0	V132	0	0
3	0	0	V143	0
4	0	0	0	V154

We can sort out the path set we obtained from Figure 9 and suppose the computed comparison values for every path as following:

NP_{ijm}^s is the team quality comparison value for the m th path connecting vertex i and vertex j with length s .

1-segment paths:

$$1 \rightarrow 2 \quad NP_{121}^1$$

2-segment paths:

$$1 \rightarrow 8 \rightarrow 2 \quad NP_{122}^2$$

3-segment paths:

$$1 \rightarrow 4 \rightarrow 3 \rightarrow 2 \quad NP_{123}^3$$

$$1 \rightarrow 4 \rightarrow 8 \rightarrow 2 \quad NP_{124}^3$$

$$1 \rightarrow 7 \rightarrow 8 \rightarrow 2 \quad NP_{125}^3$$

4-segment paths:

$$1 \rightarrow 4 \rightarrow 3 \rightarrow 5 \rightarrow 2 \quad NP_{126}^4$$

$$1 \rightarrow 7 \rightarrow 6 \rightarrow 3 \rightarrow 2 \quad NP_{127}^4$$

$$1 \rightarrow 7 \rightarrow 6 \rightarrow 5 \rightarrow 2 \quad NP_{128}^4$$

$$1 \rightarrow 8 \rightarrow 4 \rightarrow 3 \rightarrow 2 \quad NP_{129}^4$$

This path set has a configuration of 1_2_3_4, and the path set is divided into 4 subsets of paths according to their number of segments. First we average the comparison value for every subset of paths, since we assign the same weight to the paths with the same length in one pairwise comparison; second we assign the weights V_{11} , V_{12} , V_{13} and V_{14} to the 4 averaged comparison values respectively. We have the following equations to calculate A_{12} :

$$A_{12} = V_{11} * NP_{121}^1 + V_{12} * NP_{122}^2 + V_{13} * \frac{NP_{123}^3 + NP_{124}^3 + NP_{125}^3}{3} + V_{14} * \frac{NP_{126}^4 + NP_{127}^4 + NP_{128}^4 + NP_{129}^4}{4}$$

where $V_{11} + V_{12} + V_{13} + V_{14} = 1$.

3.2.5 Minimize the Error between the Actual Output and the Estimated Output

Since we have already validate the graph G is connected, even though the game set does not conform a Round-Robin Tournament, we can compute A_{ij} for any pair $i, j \in V$, which represents a weighted theoretical pair-wise comparison value on team quality. More importantly, the theoretical comparison values A_{ij} can be used to predict the output of the incoming games in the future (a game will happen between i and j). If team i and j have not played a game, we call A_{ij} as an estimated output; if they played a game, hence they have a direct comparison value a_{ij} computed from the output of the game with the conversion functions. Of course, we hope that $|A_{ij} - a_{ij}|$ is very small, if it is not zero.

Remember that we have not determined the value of C in the conversion functions and the weights shown in Table 3 yet. As an approach to decide these values, we can transform it as an optimization problem:

$$\min \sum (A_{ij} - a_{ij})^2$$

We determine the value of C and the weighting vectors such that it can minimize the error between the estimated output and the actual output of the games after the games are played.

3.2.6 Rank the Teams on Quality

Once we determine all the parameters including the value of C and the weighting vectors, and we know the actual output of all games we tried to predict, we can re-compute a_{ij} with the decided conversion functions, re-apply the path searching algorithms on the new graph with more games, and re-compute the weighted quality comparison values A_{ij} ; then we can use all A_{ij} to form the influence matrix (A_{ij}) as the input to the AHP approach, and compute the global weight of quality of every team to rank these teams.

CHAPTER 4 RESULTS AND DISCUSSION

Based on the formal model depicted in the previous chapter, we use Excel to implement the computation and analysis. Excel[26] is the most commonly used spreadsheet for PCs. It provides functions and tool packages for data analysis, computing and optimization that can satisfy the requirements of our implementation.

In this chapter, we will show our step-by-step implementation, give the final results, and finally provide some discussion.

4.1 Implementation

4.1.1 Data Collection

We collected the data from Scoreboard of Big East conference season in 2007[1], in which we only included the conference games. Table 4 shows the data, name of the home team, name of the away team, location, and score of each game in the 2007 Big East conference season.

We take the first 14 games as the first half of the season and the last 14 games as the second half of the season. And we suppose, at the time in between the two half seasons we use the results of the first half of the season to predict the results of the second half of the season, finally in the end of the full season we will compare the predicted results to the actual results to measure and optimize the accuracy of the prediction.

Table 4 The Outcomes of Games in 2007 Big East Conference Season

Date	Home Team	Away Team	Location	home score	away score
* Sat, Sep 22	Louisville	Syracuse	Louisville, Ky.	35	38
* Sat, Sep 22	Pittsburgh	Connecticut	Pittsburgh, Pa.	14	34
* Fri, Sep 28	USF	West Virginia	Tampa, Fla.	21	13
* Sat, Oct 06	Syracuse	West Virginia	Syracuse, N.Y.	14	55
* Sat, Oct 06	Rutgers	Cincinnati	Piscataway, N.J.	23	28
* Sat, Oct 13	Syracuse	Rutgers	Syracuse, N.Y.	14	38
* Sat, Oct 13	Cincinnati	Louisville	Cincinnati, Ohio	24	28
* Thu, Oct 18	Rutgers	USF	Piscataway, N.J.	30	27
* Fri, Oct 19	Connecticut	Louisville	East Hartford, Conn.	21	17
* Sat, Oct 20	Pittsburgh	Cincinnati	Pittsburgh, Pa.	24	17
* Sat, Oct 27	Louisville	Pittsburgh	Louisville, KY	24	17
* Sat, Oct 27	Rutgers	West Virginia	Piscataway, N.J.	3	31
* Sat, Oct 27	Connecticut	USF	East Hartford, Conn.	22	15
* Sat, Nov 03	Pittsburgh	Syracuse	Pittsburgh, Pa.	20	17
----- half season -----					
* Sat, Nov 03	USF	Cincinnati	Tampa, Fla.	33	38
* Sat, Nov 03	Connecticut	Rutgers	East Hartford, Conn.	38	19
* Thu, Nov 08	West Virginia	Louisville	Morgantown, W.Va.	38	31
* Sat, Nov 10	Syracuse	USF	Syracuse, N.Y.	10	41
* Sat, Nov 10	Cincinnati	Connecticut	Cincinnati, Ohio	27	3
* Sat, Nov 17	Connecticut	Syracuse	East Hartford, Conn.	30	7
* Sat, Nov 17	Rutgers	Pittsburgh	Piscataway, N.J.	20	16
* Sat, Nov 17	Cincinnati	West Virginia	Cincinnati, Ohio	23	28
* Sat, Nov 17	USF	Louisville	Tampa, Fla.	55	17
* Sat, Nov 24	Pittsburgh	USF	Pittsburgh, Pa.	37	48
* Sat, Nov 24	West Virginia	Connecticut	Morgantown, W.Va.	66	21
* Sat, Nov 24	Syracuse	Cincinnati	Syracuse, N.Y.	31	52
* Thu, Nov 29	Louisville	Rutgers	Louisville, Ky.	41	38
* Sat, Dec 01	West Virginia	Pittsburgh	Morgantown, W.Va.	9	13

4.1.2 Data Preprocessing

In order to simplify the expression and give convenience for the later computing, we assign a unique ID number to every team as shown in Table 5:

Table 5 The IDs of Teams in Big East Conference

Team name	Team No
Louisville	1
Syracuse	2
Rutgers	3
Cincinnati	4
West Virginia	5
USF	6
Connecticut	7
Pittsburgh	8

The next task is to calculate DQ (quality difference) and a_{ij} (quality comparison) according to Equations (1) and (2).

Figure 10 shows the definitions of the parameters in Excel: cell $S2$ defines the home advantage score, which is initialized as 3 points; cell $S3$ defines the parameter C in equation (1).

As Figure 11 shows, column E calculates the score difference between home team and away team, with $E_i = C_i - D_i$, in which E_i, C_i and D_i denote the corresponding cell, and i denotes the column index; column F calculates score difference adjusted by the home advantage score with the formula $F_i = E_i - \$S\2 ; column G indicates the winner team in this game with " $G_i = IF(F_i \geq 0, A_i, B_i)$ ", and column H indicates the loser team in this game with " $H_i = IF(F_i \geq 0, B_i, A_i)$ "; column I adjusts the score difference as a positive integer indicating by how many points the winner beats the loser, with $I_i = ABS(F_i)$; column J calculates the quality difference based on the current game with $J_i = 1 - EXP(-\$S\$3 * I_i)$; finally column K converts the quality difference to a quality comparison value in AHP with $K_i = 1 + 8 * J_i$.

	R	S	T	U
Parameters:				
Home Advantage		3		
C		0.112688	note: $dq=1-\exp(-C*DS)$	

Figure 10. The Parameter in Excel

	A	B	C	D	E	F	G	H	I	J	K
1	Home Team	Away Team	home score	away score	DS	Adj DS	winner	loser	win-lose	DQ	Aij
2	1	2	35	38	-3	-6	2	1	6	0.491418	4.931341
3	8	7	14	34	-20	-23	7	8	23	0.925117	8.400933
4	6	5	21	13	8	5	6	5	5	0.430753	4.44602
5	2	5	14	55	-41	-44	5	2	44	0.992975	8.9438
6	3	4	23	28	-5	-8	4	3	8	0.594041	5.752331
7	2	3	14	38	-24	-27	3	2	27	0.952288	8.618305
8	4	1	24	28	-4	-7	1	4	7	0.545618	5.36494
9	3	6	30	27	3	0	3	6	0	0	1
10	7	1	21	17	4	1	7	1	1	0.106571	1.852565
11	8	4	24	17	7	4	8	4	4	0.362851	3.902809
12	1	8	24	17	7	4	1	8	4	0.362851	3.902809
13	3	5	3	31	-28	-31	5	3	31	0.9696	8.756804
14	7	6	22	15	7	4	7	6	4	0.362851	3.902809
15	8	2	20	17	3	0	8	2	0	0	1
16	6	4	33	38	-5	-8	4	6	8	0.594041	5.752331
17	7	3	38	19	19	16	7	3	16	0.835198	7.681581
18	5	1	38	31	7	4	5	1	4	0.362851	3.902809
19	2	6	10	41	-31	-34	6	2	34	0.978321	8.826565
20	4	7	27	3	24	21	4	7	21	0.906187	8.249493
21	7	2	30	7	23	20	7	2	20	0.894996	8.159971
22	3	8	20	16	4	1	3	8	1	0.106571	1.852565
23	4	5	23	28	-5	-8	5	4	8	0.594041	5.752331
24	6	1	55	17	38	35	6	1	35	0.980631	8.845048
25	8	6	37	48	-11	-14	6	8	14	0.793537	7.348293
26	5	7	66	21	45	42	5	7	42	0.991199	8.929592
27	2	4	31	52	-21	-24	4	2	24	0.933097	8.464776

Figure 11. The Calculation of DQ and A_{ij} in Excel

In Figure 11, the red part is the games which are supposed to be played and ended in certain outputs, and the non-red part is the set of games which have not occurred yet and for which predictions are to be made. The two parts of data form a graph in Figure 12: every node represents a team, every solid arrow represents an "occurred" game, and every dotted arrow represents a game for which a prediction needs to be made. In addition, the a_{ij} values of the red part in Figure 11 should reside on the corresponding solid arrows in Figure 12.

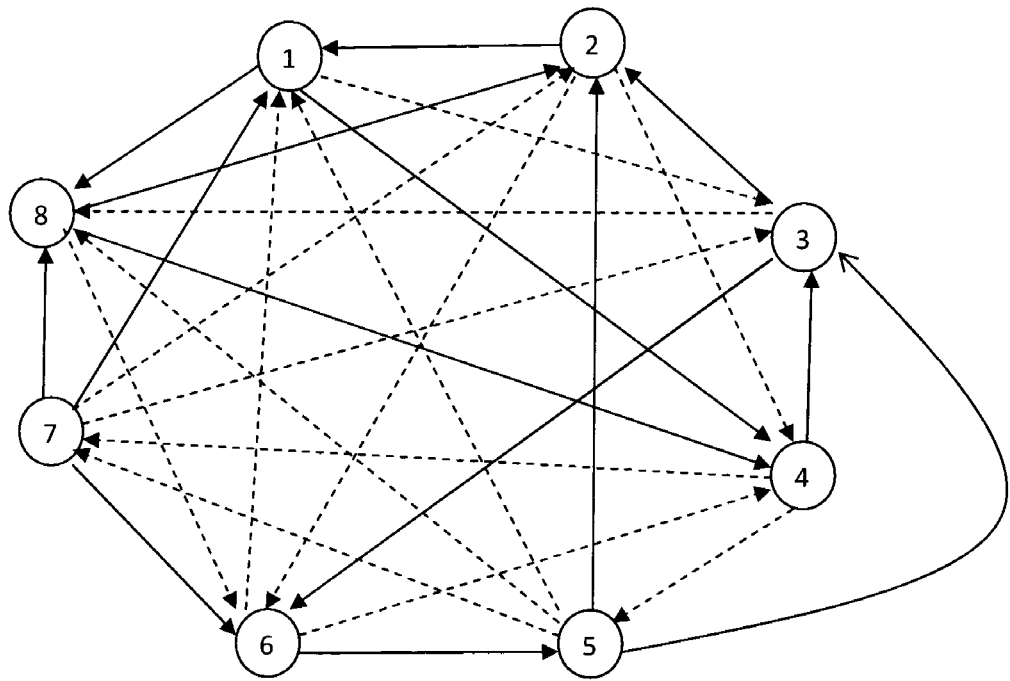


Figure 12. The Graph Formed by the Set of Games

4.1.3 Setup for the Parameter C

According to Equation 5, the parameter C is determined by the threshold value T and the precision value P . As Figure 13 shows, we define Cell S3(the C parameter) with “ $-\ln(U5/8)/S5$ ”, in which Cell U5 is the precision value, and Cell S5 is the threshold value. In the later experiment, we can adjust these two values to reflect the actual game scores.

	R	S	T	U	V
Parameters:					
Home Advantage		3			
C		0.095494453	note: $dq=1-\exp(-C*DS)$		
B		8	note: $A_{ij} = 1+B*dq$		
T(threshold):		70	P(precision):	0.01	

Figure 13. The Setup of Parameter C

4.1.4 Computation of the Weighted Quality Comparison Values between Teams

The next step is to calculate the weighted quality comparison values between teams according to the methodology previously presented.

First the weighting vectors in Table 3 are defined in the area of R8:V12. There exist 5 kinds of configurations listed in Figure 13 in our case. In every configuration, the summation of the weighting factors must be 1, and every weight factor is initialized to be equal to each other(the cell initialized as 0 is an ineffective cell).

Second the weighted quality comparison values are computed with the weighting vectors. Figure 14 shows one example of the calculation between team 1 and team 2. Column R(Paths) lists all the paths between 1 and 2 according to the graphic algorithm we specified in last chapter. A path is expressed by the sequence of the vertices it travels. For example, 1_8_2 means the path from 1 to 8 and from 8 to 2.

Column S(Calc Aij) lists all the quality comparison values along every path in terms of AHP. For example, for path 1_8_2, the corresponding cell is defined as “= K12 * K15” , which is actually the value of $A_{18} * A_{12}$.

Column T(Adj Calc Aij) adjusts the values that are greater than 9 in column S to 9, and adjusts the values that are smaller than 1/9 in column S to 1/9, with $T_i = IF(S_i > 9, 9, IF(S_i < 1/9, 1/9, S_i))$, since the comparison values in AHP are always in between 1/9 and 9.

Column U(Average) calculates the average value for every set of adjusted Aij whose paths have the same number of segments. For example, for the comparison between 1 and 2, there are only three 3-segment paths, for which the average value is defined in cell U21 as “=(T19+T20+T21)/3”.

Column V(weighted) calculates the final weighted quality comparison values highlighted with the red color. For the comparison between 1 and 2, there exist 1-segment, 2-segment, 3-segment and 4-segment paths, so that we apply the weighting vector for the configuration 1_2_3_4 on it, and the cell is defined as “=U17*S8+U18*T8+U21*U8+U27*V8” .

R	S	T	U	V	W	X
weight vector:						
configuration(no seg)	w1	w2	w3	w4	sum	
1_2_3_4	0.25	0.25	0.25	0.25	==	1
2_3_4	0	0.333333333	0.333333	0.333333	==	1
1_2_4	0.333333333	0.333333333	0	0.333333	==	1
1_3_4	0.333333333	0	0.333333	0.333333	==	1
1_4	0.5	0	0	0.5	==	1
Calculate the average Aij according to the paths:						
Paths	Calc Aij	Adj Calc Aij	Average	weighted		
1---2				5.256542		
1_2	4.931340672	4.931340672	4.931341			
1_8_2	3.902809316	3.902809316	3.902809			
1_7_8_2	4.534757522	4.534757522				
1_4_3_2	265.9687616	9				
1_4_8_2	1.37463551	1.37463551	4.969798			
1_4_3_5_2	31.51992626	9				
1_4_8_1_2	0.071424173	0.111111111				
1_7_6_5_2	83.77170345	9				
1_7_6_3_2	18.15623277	9				
1_8_4_3_2	755.127707	9	7.222222			
1---3				2.299247		
1_2_3	0.023529523	0.111111111				

Figure 14. Calculate the Weighted Quality Comparison Value between Teams

4.1.5 Optimizing the Prediction

It is reasonable to use the weighted comparison values as the predicted outputs. After we computed the weighted values for every pair of comparison between teams, we optimize the predicted outputs by changing the parameters(including the weighting vectors and the parameters in the pre-processing phase) with Solver Add-ins provided by Microsoft Excel .

	F	G	H	I	J	K	L	M	N	O	P	Q
1	Adj DS	winner	loser	win-lose	DQ	Aij	weighted Aij	error	error^2			
2	-6	2	1	6	0.49142	4.93134	0.1902391	4.7411	22.47804395			
3	-23	7	8	23	0.92512	8.40093	5.7875565	2.61338	6.829738938			
4	5	6	5	5	0.43075	4.44602	0.2896348	4.15639	17.27553959			
5	-44	5	2	44	0.99297	8.9438	0.7565914	8.18721	67.03037783			
6	-8	4	3	8	0.59404	5.75233	0.2032452	5.54909	30.79235232			
7	-27	3	2	27	0.95229	8.61831	0.2914359	8.32687	69.33675488			
8	-7	1	4	7	0.54562	5.36494	4.4654849	0.89946	0.809019946			
9	0	3	6	0	0	1	1.4248702	-0.42487	0.180514657			
10	1	7	1	1	0.10657	1.85257	1.0314995	0.82107	0.674148686			
11	4	8	4	4	0.36285	3.90281	0.3262411	3.57657	12.79184009			
12	4	1	8	4	0.36285	3.90281	3.4985858	0.40422	0.163396646			
13	-31	5	3	31	0.9696	8.7568	0.6422156	8.11459	65.84653991			
14	4	7	6	4	0.36285	3.90281	0.4702421	3.43257	11.78251794			
15	0	8	2	0	0	1	0.2590269	0.74097	0.549041091			306.539826
16	-8	4	6	8	0.59404	5.75233	2.041868	3.71046	13.76753459			
17	16	7	3	16	0.8352	7.68158	0.7079739	6.97361	48.63119071			
18	4	5	1	4	0.36285	3.90281	0.381095	3.52171	12.40247159			
19	-34	6	2	34	0.97832	8.82656	0.3366032	8.48996	72.07944489			
20	21	4	7	21	0.90619	8.24949	1.6856404	6.56385	43.08416189			
21	20	7	2	20	0.895	8.15997	1.2762512	6.88372	47.38559321			
22	1	3	8	1	0.10657	1.85257	5.49037	-3.6378	13.23362466			
23	-8	5	4	8	0.59404	5.75233	0.6248104	5.12752	26.29146628			
24	35	6	1	35	0.98063	8.84505	0.6470799	8.19797	67.20667808			
25	-14	6	8	14	0.79354	7.34829	3.6262443	3.72205	13.85364329			
26	42	5	7	42	0.9912	8.92959	2.7737048	6.15589	37.8949523			
27	-24	4	2	24	0.9331	8.46478	0.2428271	8.22195	67.60044902			
28	0	1	3	0	0	1	2.2992467	-1.29925	1.6880421			
29	-7	8	5	7	0.54562	5.36494	0.1730542	5.19189	26.95568119			492.074934
30									sum of error			798.61476
31									SD			5.34059507
32												

Figure 15. The Measurement of Optimization

We list all the weighted quality comparison values in column L(weighted Aij) as shown in Figure 15. Some of the values are kept the same as specified in Figure 14 synchronically, whereas some of the values may need an adjustment. For example, we only have the weighted quality comparison value from the loser 1 to the winner 2, but now we need the comparison value from the winner 2 to the loser 1 in row 2, as a solution we just use $A_{21} = 1/A_{12}$ to obtain the corresponding comparison value according to the computation in the AHP.

Column M(error) lists all the differences between the actual outputs(Aij) and the predicted outputs(weighted Aij) as the error of the prediction, with $M_i = K_i - L_i$.

Column N(error^2) lists all the squares of the errors, with $N_i = M_i * M_i$.

Cell Q15 sums up all the squared errors for the first half of the season, with “=SUM(N2:N15)”; cell Q29 sums up all the squared errors for the second half of the season, with

"=SUM(N16:N29)"; cell Q30 specifies the total sum of the squared errors with "=Q30", and cell Q31 is the SD(Standard Deviation) between the actual output and the predicted output with "=SQRT(Q30/28)".

In order to optimize the prediction, we must minimize SD by changing the parameters; and by this way we can seek for the best parameters to set up our models for future predictions. We use Excel Solver Add-in to minimize SD, as shown in Figure 16. We set up SD(\$Q\$31) as the target cell to minimize it by changing the cells containing the parameters and the weighting vectors, with subjection to the constraints : the cells to be changed must be positive, and the summation of every set of weighting vector must be 1.

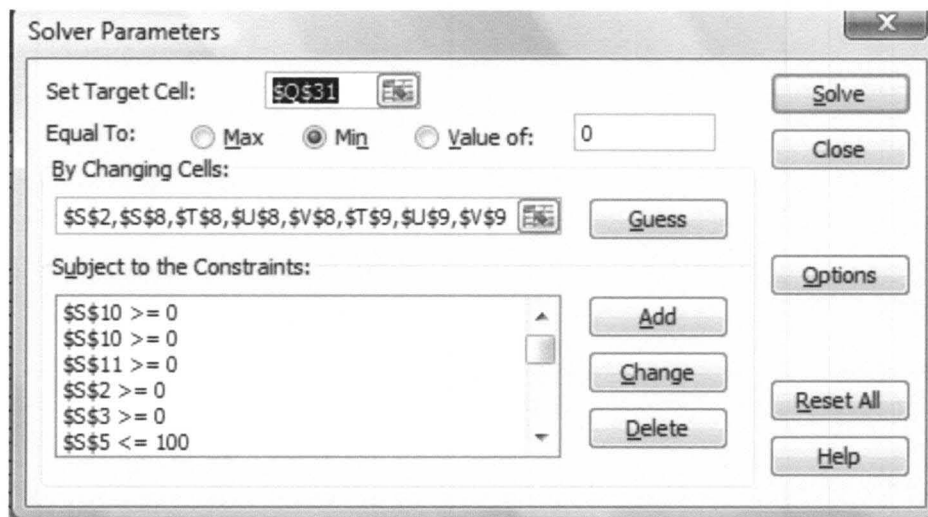


Figure 16. Determine Parameters in Excel

4.1.6 Rank the Teams with Expert Choice

Expert Choice[27] is a robust, desktop-based application that enables teams to prioritize objectives, evaluate alternatives and achieve alignment, buy-in, and confidence around important organizational decisions. With Expert Choice, you can combine the expertise and intuition of your team with quantitative information to provide valuable insights, explore what-if scenarios, and achieve stakeholder consensus and understanding.

As Figure 17 shows, the team ranking problem is structured as a two-hierarchy AHP problem: the top level is the unique criteria “Team Quality”, the low lever lists the 8 teams as alternatives.

As Figure 18 shows, the weighted A_{ij} are input into the matrix of pair-wise comparison. In addition, the red color number indicates an inverse comparison. For example, the cell crossed by team 3 and team 8 is 3.5 with red color, which means team 8 is better than team 3 with a quality comparison value of 3.5; if the color of that number is black, which means team 3 is better than team 8 with a quality comparison value of 3.5.

Figure 19 shows the ranking output, in which the teams are listed from high quality to low quality.

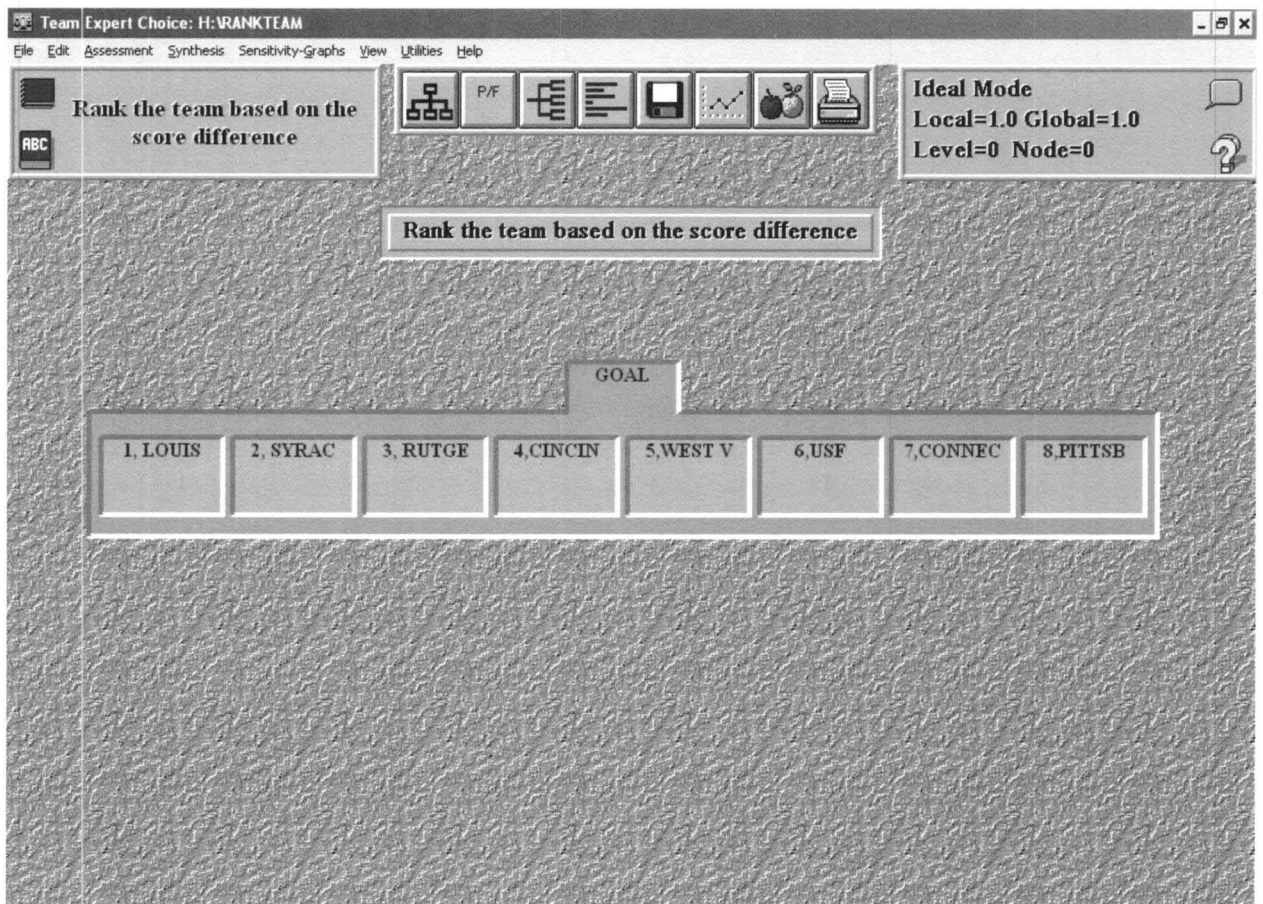


Figure 17. The Structure of Our Problem in Expert Choice

With respect to GOAL

8,PITTSB: 8,Pittsburgh

is 3.5 times (MODERATELY to STRONGLY) more IMPORTANT than

4,CINCIN: 4,Cincinnati

[Best Fit]	2. SYRAC	3. RUTGE	4.CINCIN	5.WEST V	6.USF	7.CONNEC	8,PITTSB
1. LOUIS	4.5	1.2	4.9	4.1	1.8	1.7	3.5
2. SYRAC		8.4	3.4	8.9	5.5	3.3	1.0
3. RUTGE			5.3	8.6	1.0	1.9	4.7
4.CINCIN				2.1	6.5	4.1	↑ 3.5
5.WEST V					4.0	4.1	4.4
6.USF						3.5	3.9
7.CONNEC							8.1

Figure 18. The Matrix Input of Pair-wise Comparison

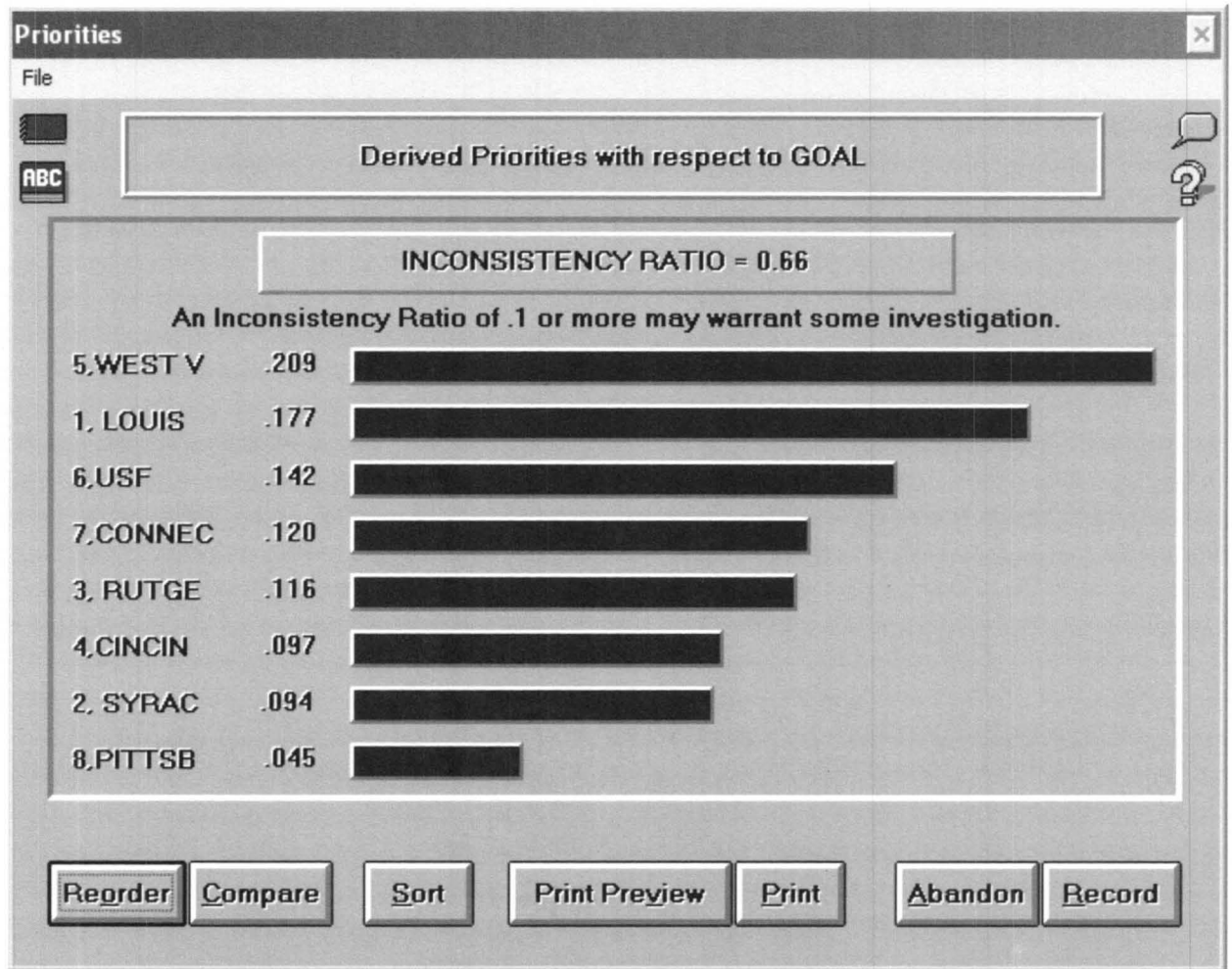


Figure 19. The Result of the Team Ranking

4.2 Results

We set up $T = 70$ and $P = 0.01$, which are considered to be reasonable values (since the pair-wise comparisons in AHP are subjective judgments), hence the conversion function is decided. We ran the optimization with Solver and obtained the following results:

Table 6 The Outcomes of the Optimization

winner	loser	win-lose	DQ	Aij	weighted Aij	error	error ²
2	1	6	0.43615	4.489199	0.222756885	4.266442	18.20253
7	8	23	0.888794	8.110353	8.110347947	5.37E-06	2.89E-11
6	5	5	0.379651	4.037204	4.037193523	1.06E-05	1.11E-10
5	2	44	0.985031	8.880246	8.88022263	2.32E-05	5.39E-10
4	3	8	0.53418	5.273439	5.273309773	0.000129	1.66E-08
3	2	27	0.924101	8.392807	8.392473902	0.000333	1.11E-07
1	4	7	0.487503	4.900027	4.900025801	1.28E-06	1.64E-12
3	6	0	0	1	0.999999388	6.12E-07	3.75E-13
7	1	1	0.091077	1.728613	1.728613748	-9.7E-07	9.49E-13
8	4	4	0.31749	3.539918	3.539887647	3.08E-05	9.48E-10
1	8	4	0.31749	3.539918	3.542970827	-0.00305	9.32E-06
5	3	31	0.948198	8.585585	8.572375112	0.01321	0.000174
7	6	4	0.31749	3.539918	3.539903351	1.51E-05	2.27E-10
8	2	0	0	1	0.999996716	3.28E-06	1.08E-11
4	6	8	0.53418	5.273439	0.154756685	5.118682	26.20091
7	3	16	0.783012	7.264093	0.517366002	6.746727	45.51832
5	1	4	0.31749	3.539918	0.24439493	3.295524	10.86048
6	2	34	0.961102	8.688816	0.182760903	8.506055	72.35297
4	7	21	0.865391	7.923131	4.136977452	3.786153	14.33496
7	2	20	0.851903	7.815225	3.313967156	4.501258	20.26133
3	8	1	0.091077	1.728613	4.747472949	-3.01886	9.113517
5	4	8	0.53418	5.273439	2.108716901	3.164722	10.01546
6	1	35	0.964645	8.717157	0.550081938	8.167075	66.70112
6	8	14	0.737347	6.898778	3.862945143	3.035833	9.21628
5	7	42	0.981881	8.855044	4.134460345	4.720584	22.28391
4	2	24	0.898922	8.191379	0.294195164	7.897184	62.36552
1	3	0	0	1	1.292283503	-0.29228	0.08543
8	5	7	0.487503	4.900027	0.22577849	4.674249	21.8486
sum of error							409.3615
SD							3.823618

Table 6 shows the weighted Aij in the column “weighted Aij”, and the difference between the actual quality comparison values based on the actual game results and the predicted quality comparison values computed with our method with the following weighting vectors in Table 7.

Table 7 The Outcomes of the Weighting Vector of the Optimization

configuration(no seg)	w1	w2	w3	w4
1_2_3_4	0.999999	0	7.99E-07	0
2_3_4	0	0	0.011556	0.988444
1_2_4	0.996874	5E-05	0	0.003076
1_3_4	0.999999	0	3.29E-07	3.29E-07
1_4	1	0	0	3.29E-07

And the ranking output has already been displayed in Figure 19 and in Appendix A.

CHAPTER 5 CONCLUSION

In this thesis, we have proposed a new method to rank football teams according to the outcomes of games with the techniques borrowed from AHP and graph theory. A set of outcomes of games between teams is interpreted as a graph, composed by the vertices representing teams and the edges representing games with score differences; the pair-wise comparisons on team quality are made, through integrating the quality comparisons based on direct games and the indirect quality comparisons derived from some specific paths in the graph; finally the pair-wise comparison values are input into Expert Choice to develop a ranking. Our method gives the following contributions compared to the computer models in the BCS:

1. Our method accounts for margins of victory, which is not considered by the computer models in the BCS. Instead the BCS models only count the winning percentage and the quality of opposite teams for the ranking. This can reduce the motivation for the teams to earn more score difference in games, and result in conservative game strategies. Our method takes margin of victory into account through pair-wise quality comparisons between teams based on outcomes of games. This will not only inspire the teams that are going to win to try their best to enlarge the score difference, but also inspire the teams that are going to lose to attempt to reduce the score difference.
2. Our method takes a holistic view to evaluate the quality of teams. Specifically the set of games is interpreted as a graph, a specific set of paths is searched for every pair-wise comparison between teams, and the quality comparison value is calculated from the integration of different quality comparison values for different paths; thus every pair-wise

comparison value can capture the holistic information in a set of games for the whole season.

There are still some improvements that we can make through the following further work:

1. In order to calculate the parameters which can reflect the model more accurately, we need more data about games to be involved in the implementation, and we probably need a mechanism to filter the outlier data out. In the implementation, we only used the outcomes of games of Big East conference in 2007, which consists of 8 teams and 28 games, to compute the fitting model; however, this cannot guarantee that the model we calculated out can reflect the fitting model statistically, therefore more data including more teams and games should be introduced into the computation to detect the most fitting model. When more data are introduced, we must pay attention to the outlier data, which is interpreted as an exception compared to most of the other data, because these outlier data could interrupt the computation to seek for more fitting parameters for the model.
2. We can try to detect some new conversion functions. Our conversion function expressed by Equations (1) and (2) might not be the best function to convert a score difference into a quality comparison value. There is one shortcoming in our conversion function, to make our model work we must setup Threshold T and Precision P , in which there are always some subjective judgments.

In general, although our method is not flawless, it approved the feasibility to develop a method to rank the teams based on margins of victory and pair-wise comparisons. Based on this feasibility, we can develop better methods along this way.

REFERENCE

1. website, b.e.c. 2007; big east conference website]. Available from:
http://www.bigeast.org/SportSelect.dbml?SPSID=103568&SPID=11215&DB_OEM_ID=19400&DB_MENU_ID=&Q_SEASON=2007.
2. Harris. *Harris Interactive Poll*. Available from: 1.
http://en.wikipedia.org/wiki/Harris_Interactive_College_Football_Poll>
2. <http://www.harrisinteractive.com/news/bcspoll.asp>>
3. Poll, C.s., *Web page of Coach's Poll*
4. Anderson and Hester. *web page of Anderson&Hester*. Available from: 1.
http://www.andersonsports.com/football/ACF_frnk.html>
5. **Billingsley, R.** *Web page of Richard Billingsley*. Available from: 1.
http://www.cfr.com/Archives/ISO_08.htm>
6. Massey, K. *Massey Ratings Theory*. 1997; Available from:
<http://www.mratings.com/theory/index.htm>.
7. Sagarin, J., *web page of Jeff Sagarin*.
8. **Wolfe, P.**, *web page of Peter Wolfe rating*.
9. Colley, W.N., *Colley's Bias Free College Football Ranking Method: The Colley Matrix Explained*. 2002, Princeton University.
10. Keener, J.P., *The Perron-Frobenius theorem and the ranking of football teams*. SIAM Rev., 1993. **35**(1): p. 80-93.
11. Wilson, R.L., *Ranking College Football Teams: A Neural Network Approach*. Interfaces, 1995. **25**(4): p. 44-59.
12. Harville, D., *The Use of Linear-Model Methodology to Rate High School or College Football Teams*. Journal of the American Statistical Association, 1977. **72**(358): p. 278-289.
13. Glickman, M.E. and H.S. Stern, *A State-Space Model for National Football League Scores*. Journal of the American Statistical Association, 1998. **93**(441): p. 25-35.
14. Cassidy, C.R., L.M. Maillart, and S. Salman, *Ranking Sports Teams: A Customizable Quadratic Assignment Approach*. Interfaces, 2005. **35**(6): p. 497-510.
15. Stefani, R.T., *Football and Basketball Predictions Using Least Squares*. Systems, Man and Cybernetics, IEEE Transactions on, 1977. **7**(2): p. 117-121.
16. Harville, D.A., *The Selection or Seeding of College Basketball or Football Teams for Postseason Competition*. Journal of the American Statistical Association, 2003. **98**(461): p. 17-27.

17. Bassett, G.W., Jr., *Robust Sports Ratings Based on Least Absolute Errors*. The American Statistician, 1997. **51**(2): p. 99-105.
18. Mease, D., *A Penalized Maximum Likelihood Approach for the Ranking of College Football Teams Independent of Victory Margins*. The American Statistician, 2003. **57**(4): p. 241-248.
19. Park, J. and M.E.J. Newman, *A network-based ranking system for US college football*. Journal of Statistical Mechanics: Theory and Experiment, 2005(10): p. P10014.
20. Callaghan, T., P.J. Mucha, and M.A. Porter, *Random Walker Ranking for NCAA Division I-A Football*. 2003.
21. Boginski, V., S. Butenko, and P. Pardalos, *Matrix-based methods for college football rankings*, in *Economics, Management and Optimization in Sports*. 2004.
22. Hochbaum, D.S., *Ranking Sports Teams and the Inverse Equal Paths Problem*, in *Internet and Network Economics*. 2006, Springer Berlin / Heidelberg.
23. Cambron, K.E. and G.W. Evans, *Layout design using the analytic hierarchy process*. Comput. Ind. Eng., 1991. **20**(2): p. 211-229.
24. Golden, B.L., E.A. Wasil, and P.T.H. (Eds.), *The Analytic Hierarchy Process-- Applications and Studies*. 1989: Springer-Verlag.
25. Saaty, T.L., *Decision making for leaders : the analytic hierarchy process for decisions in a complex world*. 1990, Belmont, Calif.: Lifetime Learning Publications. xi, 291 p.
26. Microsoft. *Excel Official Website*. 2007; Available from: <http://office.microsoft.com/en-us/excel/default.aspx>.
27. *Expert Choice* Available from: <http://www.expertchoice.com/products-services/expert-choice-115>>

APPENDIX A

Rank the team based on the score difference

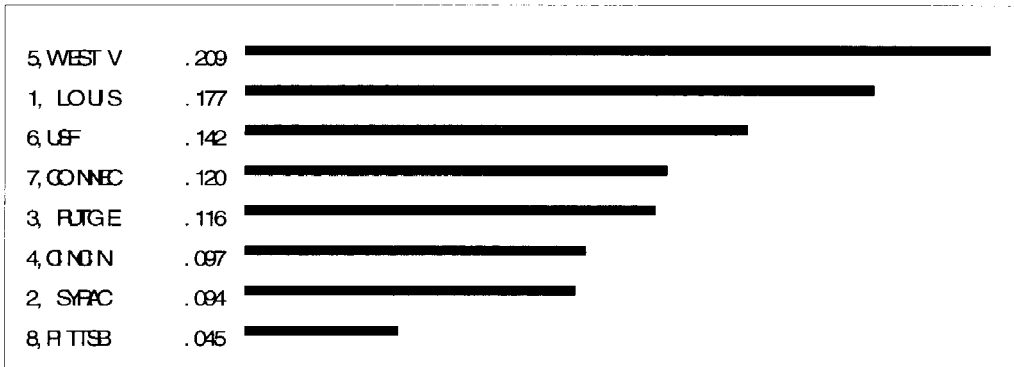
Node: 0

Compare the relative IMPORTANCE with respect to GOAL

	1, LOUIS	6, USF	7, CONN	3, RUTGE	4, CINN	2, SYRAC	8, PITTSB
5, WEST V	(4.1)	(4.0)	4.1	8.6	2.1	8.9	4.4
1, LOUIS		1.8	(1.7)	1.2	4.9	4.5	3.5
6, USF			(3.5)	1.0	6.5	(5.5)	3.9
7, CONN				(1.9)	(4.1)	3.3	8.1
3, RUTGE					(5.3)	8.4	4.7
4, CINN						(3.4)	(3.5)
2, SYRAC							1.0

Row element is ... times more than column element unless enclosed in ()

Abbreviation	Definition
Goal	Rank the team based on the score difference
5, WEST V	5, West Virginia
1, LOUIS	1, Louisville
6, USF	6, USF
7, CONN	7, Connecticut
3, RUTGE	3, Rutgers
4, CINN	4, Cincinnati
2, SYRAC	2, Syracuse
8, PITTSB	8, Pittsburgh



Inconsistency Ratio = 0.66

Educational Use Only

CURRICULUM VITAE

NAME: Yepeng Sun

ADDRESS: 789 Eastern PKWY APT 4

Louisville, KY 40217

DOB: Hebei, China – May 6, 1976

EDUCATION

&TRAINING: M.S.E., Industrial Engineering

University of Louisville

2007—2009

M.S., Computer Science

Aalborg University, Denmark

2004—2006

B.S., Computer Science and Technology

Northeastern University, China

1996—2000

WORK

EXPERIENCE: Software Developer

Advanced System Development Co.Ltd

2000-2003

Software Developer

Application Solutions & Technologies Inc

2003--2004