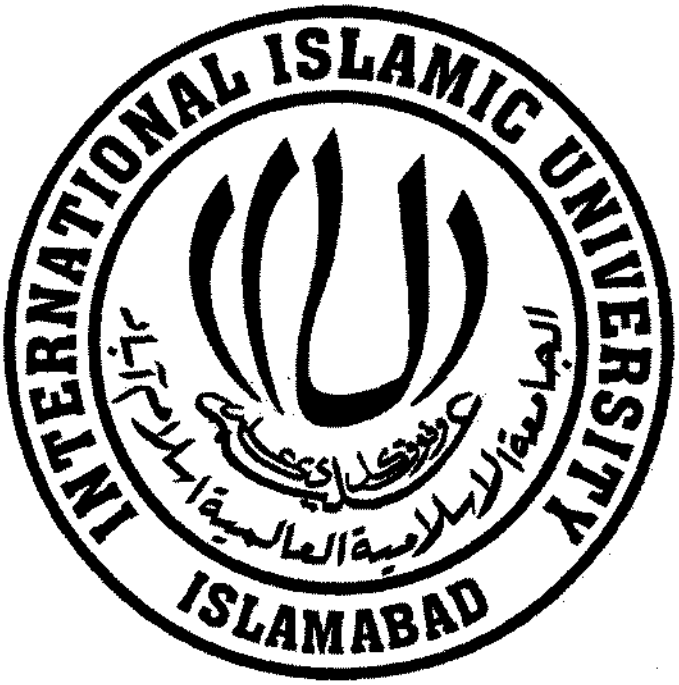
Link Fusion: A Unified Cricketer Ranking Algorithm



Department of Computer Science & Software Engineering

Faculty of Basic and Applied Sciences

International Islamic University Islamabad

2015

MS
003.54
AKL

- Measurement of quantities
- Link Fusion Algorithm
- Thesis

MS Thesis

Link Fusion: A Unified Cricketer Ranking Algorithm

Submitted in partial fulfilment of the requirements for the
MS Degree in Computer Science & Software Engineering Department

Supervisor:

Dr. Ali Daud

Assistant Professor, at the Department of Computer
Science & Software Engineering,
International Islamic University Islamabad, Pakistan

Akbar Hussain

687-FBAS/MSCS/F12

Department of Computer Science & Software Engineering
Faculty of Basic and Applied Sciences International
Islamic University Islamabad
2015

رَبِّ زِدْنِیْ عِلْمًا ۞

(طٰہٰ: 114)

اے میرے رب! میرے علم میں اضافہ فرما۔

In the name of

Allah,

The most Merciful and Compassionate the most Gracious and the Beneficent whose help and Guidance we always solicit at every step, and every moment.

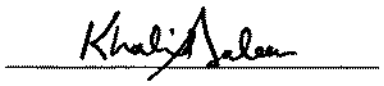
Date: 29/07/2015

Final Approval

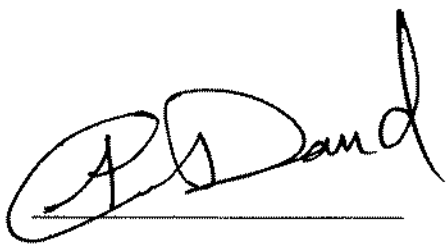
This is to certify that we have read and evaluated the thesis titled **"Link Fusion: A Unified Cricketer Ranking Algorithm"** submitted by **Akbar Hussain** under **Reg. No. 687-FBAS/MSCS/F12**. It is our judgment that this thesis is of sufficient standard to warrant its acceptance by International Islamic University, Islamabad for the degree of MS in Computer Science.

COMMITTEE

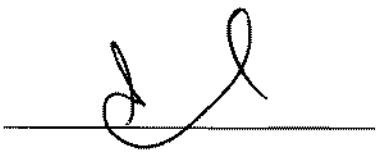
External Examiner
Dr. Khalid Saleem
Assistant Professor
Department of Computer Science
Quaid-e-Azam University, Islambad



Supervisor
Dr. Ali Daud
Assistant Professor
Department of Computer Science & Software Engineering
International Islamic University, Islamabad



Internal Examiner
Dr. Muhammad Adeel
Assistant Professor
Department of Computer Science & Software Engineering
International Islamic University, Islamabad



Dedication

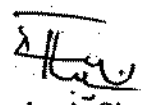
I dedicate this project to Allah, Hazrat Muhammad (PBUH), my beloved parents, respected teachers and all those who helped and prayed for my success.

DECLARATION

I **AKBAR HUSSAIN** S/O **SHER QUBAD**

Registration No. **687-FBAS/MSCS/F12**

Students of MS in Computer Science at International Islamic University do hereby solemnly declared that the thesis entitled **"Link Fusion: A Unified Cricketer Ranking Algorithm"**, submitted by me in partial fulfilment of MS degree in Computer Science, is my original work, except where otherwise acknowledged in the text, and has been submitted or published earlier and shall not, in future, be submitted by me for obtaining any degree from this or any other university or institution.



Student's Signature

Date: 29 July, 2015

DECLARATION

I Dr. **ALI DAUD**

Supervisor of **Mr. AKBAR HUSSAIN**

do hereby solemnly declare that the thesis entitled "**Link Fusion: A Unified Cricketer Ranking Algorithm**" being submitted as partial fulfillment of MS degree in the discipline of **Computer Science** has been completed under my guidance and supervision and is an original work of the student except where otherwise acknowledged in the text. It has not been submitted or published earlier for obtaining any degree from this or any other university or institution.

This thesis is completed in all respects and I am fully satisfied with the quality of student's research work. Now it is ready to be evaluated by external subject experts.

Date: 29 July, 2015

Signature: _____

Name in full: Dr. ALI DAUD

Address: Head of Data Mining and Information Retrieval Group, Assistant Professor at International Islamic University Islamabad, Pakistan

Phone: 051-9019509

Email: ali.daud@iiu.edu.pk

Acknowledgement

All praise to Almighty Allah, who gave me understanding, courage and patience to complete this research. I thank Him to be always there for us and to give me the world best parents who encouraged me at every stage of life and due to their efforts I am at this position today. I also thank him to give me a very cooperative teacher and supervisor in this research **Dr. Ali Daud** who helped and encouraged me to complete MS research thesis.

Akbar Hussain

687-FBAS/MSCS/F12

PROJECT IN BRIEF

PROJECT TITLE	:	Link Fusion: A Unified Cricketer Ranking Algorithm
UNIVERSITY	:	Department of Computer Science & Software Engineering International Islamic University, Islamabad
UNDERTAKEN BY	:	Akbar Hussain 687-FBAS/MSCS/F12
SUPERVISED BY	:	Dr. Ali Daud Assistant Professor, at the Department of CS & SE, International Islamic University Islamabad, Pakistan
TOOLS USED	:	Microsoft SQL Server 2012 MS Office 2013 for Documentation Wamp Server 2012 for Implementation
OPERATING SYSTEM:		Windows 8.1 Professional
SYSTEM USED	:	Haier Core i3 Intel ® CPU 2.00 GHs RAM 4 GB
START DATE	:	September 2014
COMPLETION DATE :		July 29, 2015

ABSTRACT

Ranking in cricket shows to quantify the importance of a player over other players. In cricket team selection, it is indispensable for a player to have good rank comparatively with other players. ICC consider overall cricket as one geographic to rank players based on different features like batting average, bowling average, strike rate, and economy rate etc. But we have divided overall cricket in five different regions (i.e. Africa, Americas, Asia Europe, and Oceania) and computed the ranking of the players (batsmen and bowlers) on each separate regions. There is a strong relationship among the players versus players and players versus teams, which is heterogeneous relationship. A player scores against a stronger bowler with good strike rate or takes wickets against stronger batsmen with lower strike rate and economy rate consider to be a good player. This study considers not only scoring the number of runs against a stronger bowler or getting wickets against a stronger batsman but also incorporates impact of the team's batting and bowling line-up against which he/she playing. Further we have calculated region based team strength which clearly influence player's ranking. In addition we have computed a Unified players ranking and compare with the current ICC ranking. We use the concept of link fusion algorithm as a model to rank the players. The results show that region based ranking can be effectively used to improve player's ranking for One Day International (ODI) cricket.

TABLE OF CONTENTS

LIST OF TABLES

No. **Description** **Page #**

LIST OF FIGURES

No. **Description** **Page #**

LIST OF ABBREVIATIONS

Avg	Average
Bat	Batsman
Bol	Bowler
ER	Economy Rate
ICC	International Cricket Council
IPL	Indian Premier League
ODI	One Day International
SR	Strike Rate
TS	Team Strength
W	Weightage
Wkt	Wicket

CHAPTER 1

INTRODUCTION

1. Introduction

In cricket player's ranking is very important. If a player perform well he will automatically represents the strength of the team. Each team have their good and best players in their own team and in overall teams. Player's importance is known on the technique of ranking.

1.1 Ranking Players in Cricket

Ranking is done in every field for the purpose to know about the orders and positions but in cricket ranking is very important. The tactic which are used to rank the players are the batting position, experience, home country advantage, match starting time and inning sequence, opposition team and its strength, performance and form of the players (S. Mukherjee, 2014).

1.2 Ranking Players based on Region Ranking

Region plays an important role in ranking the cricket players. Average scores achieves by a team depict the strength of a region. Finding ranking in cricket based on region is important but is very tricky task because you should check all the records of individual player in each region. Players ranking is the technique of accessing the worth of player according to their performance. Player's performance mostly depends upon batting average and batting strike rate for batsmen and bowling average, bowling strike rate, and economy rate for bowlers. In our proposed method we have find out players ranking for ODI matches on each region in order to see the strength of the teams on each region. We have ignored the overall ranking of the players and divided whole cricket jointly into five regions. We just considered the strength of the region individually for measuring the rank of the players. Players ranking need to be improved by considering region wise ranking.

1.3 Difference between overall and region based players ranking

ICC has ranked to each players based on total number of runs divided by its dismissals innings for batsmen and for bowlers it has ranked them total runs conceded in all matches (innings) divided by total number of wickets taken. But in region wise ranking it is quite different.

In region wise ranking, we have to know the importance of the players based on the specific regions/venues. We will classify the players that in which region the player is important and in which region they have bad performance, after that we will apply different technique to classify and find out the best players.

1.4 Cricket Sports

Cricket is a bat-and-ball game played between two teams of 11 players each. Each team takes it in turn to bat, attempting to score runs, while the other team fields. Each turn is known as an innings. The bowler delivers the ball to the batsman who attempts to hit the ball with his bat far enough for him to run to the other end of the pitch and score a run. Each batsman continues batting until he is out. The batting team continues batting until ten batsmen are out or the specific overs are completed and/or time is over, at which point the teams switch roles and the fielding team comes in to bat (https://en.wikipedia.org/wiki/Cricket).

1.4.1 Cricket Sports Formats

Internationally Cricket is playing throughout the world in three formats now-a-days for everyone, from which the five-day tactical tension of the Test match to the likeable and relish format of T-20 game. Basically all of the three formats have their own rules and technique of playing. These three formats are:

1. Test match Cricket
2. ODI (One Day International) Cricket and
3. Twenty20 Cricket

In test match both the team's will have permission to avail two innings each side and it starts for five continuous days, while ODI match is of 50 overs match for each side and T-20 is a limited over game of 20 overs per side (http://www.dummies.com/how-to/content/cricket-for-dummies-cheat-sheet.html).

1.4.2 Cricket Teams

Basically there are almost 124 member countries in the world in which cricket is played. ICC have divided all member countries in three different categories: these three categories are Full Member Countries, Associate Member Countries and Affiliate Member Countries. In out of 124 countries top 10 countries are Full Members countries in which the following countries are included Australia, Bangladesh, England, India, New Zealand, Pakistan, South Africa, Sri- Lanka, West Indies and Zimbabwe. Full member teams have the ICC permission to play all the three cricket format and follow ICC rules. In the list of 124 countries from number 11 to 47 countries are Associate Member countries. These member countries just play ODI and T-20 cricket games and

they also follow the ICC rules like Full Member countries. The rest of the countries are Affiliate Member countries, they can't play cricket internationally, and they just follow the ICC rules and regulations. (http://en.wikipedia.org/wiki/List_of_International_Cricket_Council_members).

1.5 Regions in Cricket Sports

We will rank players in cricket based on region wise, therefore we will categorize the regions. There are basically five regions in the world in which internationally cricket is playing. These regions are: (http://www.icc-cricket.com/about/67/icc-members/regions)

- **Region 1 (Europe Region)**: England
- **Region 2 (America Region):** West Indies
- **Region 3 (Asia Region)**: Pakistan, India, Sri Lanka and Bangladesh
- **Region 4 (Oceania Region):** Australia and New Zealand
- **Region 5 (Africa Region):** South Africa and Zimbabwe

1.6 Link Fusion Algorithm

In social network, hyperlink analysis and scientific citation in the web there are similarities among link analysis. These examples form single or multiple data spaces within the data objects which contain specific attribute related to data spaces within the data objects. There is homogeneous (inter-type links) or heterogeneous (intra-type links) relationship among the objects of data spaces in order to calculate a specific attribute which is called a unified link analysis framework called a "Link Fusion Algorithm". Most current web link analysis research only analyzes the hyperlinks within web pages, which can be considered as a homogeneous data space. But in the real world, the web pages will often interact with other types of objects, such as users and the queries. This work try to deal with these inter-relationships by expanding the link analysis to combine both inter-type link analysis and intra-type link analysis, and thereby improve web search performance. In Figure 1.1, an example of inter and intra type links by analyzing the relationship of three related data types in the web environment: user, web page, and query. Users and the queries they submit, plus the web pages they browse, form three homogeneous data spaces. They are correlated when a user submits queries, a user browses web pages, and a query references web pages. The three operations: submit, browse, and reference, involve inter-type links across these data spaces. The hyper-links within web pages, content-based similarity of queries, and social structure of users are

intra-type relationships within each space. It is obvious that when analyzing the attributes of web pages, not only the hyper-links between them, but also the users who browse them and the queries that reference them can play important roles.

Suppose there are n different data objects within all data spaces (A1, A2, A3… An). Each data objects Ai contains a specific attribute Fi. Data objects within the same data space are interrelated $R_i \subseteq A_i \times A_i$, while data objects from different data spaces are intra-related $R_{ij} \subseteq A_i \times A_j$ ($i \neq j$). Suppose data objects of different data spaces are similar in nature. The specific attribute of a data objects of any data space is equal the summation of the attributes of the same data space that links to it and the addition of the other attribute of other data space that are related to it. Mathematically we can define as:

$$Fi = FiRi + \sum_{j \neq i} Fj\ Rji \quad \text{---------- (1.1)}$$

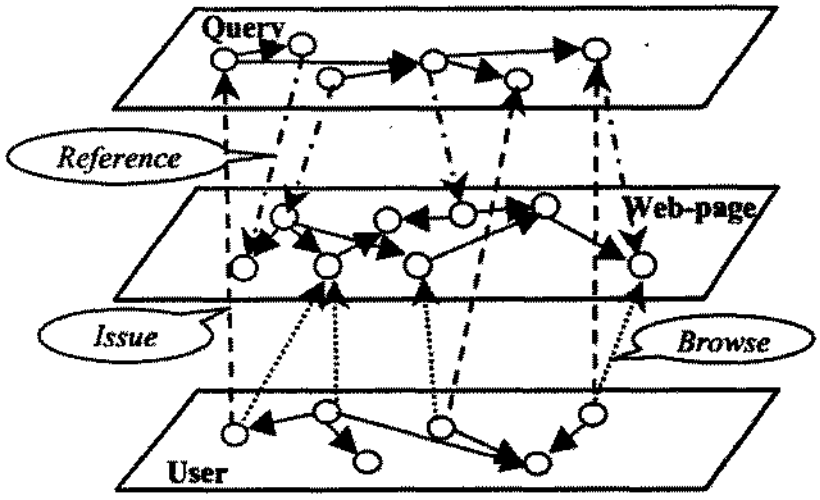


Figure 1.1 An illustration of multi type interrelated data spaces

For the simplicity of the equation of (1.1), we explain it with two data spaces, in which some attributes are related with each other. Suppose A & B are two data spaces in which A = (a1, a2, a3… a_n) and B = (b1, b2, b3 …b_n). There are four relationship between data space A and B. Relationships are Ra, Rb, Rab and Rba. In equations (1.2) and (1.3) L_A, L_B are the adjacency matrices within set A and B respectively. L_{AB} and L_{BA} are the adjacency matrices from the data objects of data space A to the data object of data space B respectively. L_{AB} (i, j) = 1, if the link

from object a_i to b_j exists, otherwise $L_{AB}(i, j) = 0$. Suppose W_a and W_b are the attribute vectors of objects in A, and B respectively. Then we can write equation 1 and 2 mathematically as:

$$W_a = L_a^T W_a + L_{ab}^T W_b \quad \text{---------- (1.2)}$$

$$W_b = L_b^T W_b + L_{ba}^T W_a \quad \text{---------- (1.3)}$$

Equation 2 and 3 can be generalized as:

$$W_M = L_M^T W_M + \sum_{\forall N \neq M} L_{NM}^T W_N \quad \text{---------- (1.4)}$$

On the basis of equation (1.4), a Unified Square Matrix U is created, where n is the total number of data objects from all the data spaces, in practice we can always ignore undesired intra and intertype relationships by setting the corresponding values of α and β to 0, L_M objects are on the diagonal direction, and all other data objects are L_{NM}. The matrix is given below.

$$U = \begin{bmatrix} \alpha_1 L_{11} & \beta_{12} L_{12} \dots & \beta_{1n} L_{1n} \\ \alpha_1 L_{11} & \beta_{12} L_{12} \dots & \beta_{2n} L_{2n} \\ \alpha_1 L_{11} & \beta_{12} L_{12} \dots & \alpha_{nn} L_{nn} \end{bmatrix} \quad \text{---------- (1.5)}$$

Suppose W is the attribute vector of all the data objects in different data spaces. The iterative approach proposed by the author is actually transforming the vector W using matrix U, which works like $W = U^T W$ that works iteratively up to the convergence of the calculation. In the Unified matrix U using all the adjacency matrices, we usually create a unified data space, which include different type of attributes of data objects in data spaces. In this case we say the Link Fusion Algorithm a link analysis (Wensi et al., 2004).

1.7 Research Contributions

The main objectives and contribution of our research are:

1. A unified link analysis framework, called "Link Fusion", to analyze inter and intra type links and to bring order to data objects in different data spaces at the same time. Next the proposed link fusion algorithm is applied in to a real world scenario of three data spaces: Hub-page data space, Authority space and User space. The concept of Link Fusion algorithm is used in our research for ranking cricket player based on each region.
2. We have calculated each team's batting strength and bowling strength on separate regions, in addition combining both the strengths conjointly.

3. We have calculated the performances of each batsmen against each bowler on each region and performances of each batsmen against each team on each region and similarly for the bowlers as well.
4. Finally a unified players ranking is computed and associate with ICC ranking.

1.8 Chapter Summery

In this chapter we have briefly discussed general idea of cricket sports, including ranking in ODI format, difference between region-based ranking and overall ranking in cricket, teams in cricket, cricket sports formats, regions in cricket sports, formats in cricket and finally link fusion algorithm which is the fundamental section of our research work.

CHAPTER 2

LITERATURE REVIEW

&

PROBLEM STATEMENT

2. Literature Review and Problem Statement

In cricket there is no research have been done before based on each region individually, even in other sports as well. Our proposed research is related to ranking cricket players on each region using the concept of Link Fusion Algorithm. The literature review given blow about cricket and brief of the Link Fusion Algorithm is discussed.

2.1 Cricket Game

Ranking is done in every field for the purpose to know about the order and position. In cricket Ranking shows the importance of the players which ultimately represents the strength of the team. Each team have their good and best players in their own team and in overall teams. Those players which have highest points is usually considered to be in top rank. Traditionally in sports, especially in cricket ratings are based on the number of winning, losing or tie matches which is measure by win/loss ratio, which is basically the subjective rating, like ICC rankings based on this specific rules (Daud and Muhammad, 2013). The success of a team is determined by the 'quality' of wins and not only the number of wins. Author has used the diffusion based PageRank algorithm on the networks in order to measure the important of winning a match by which rank the teams and its captains respectively (Duch et al., 2010).

Traditionally both batsmen and bowlers are graded on their batting average and bowling averages respectively. However in a Cricket game it plays significant role which one scores runs or claims a wicket against a strong bowling line-up or carrying a wonderful performance against a strong batting line-up team rated higher credit as compare to weak bowling line-up. The author has discovered the Social Network Analysis (**SNA**) applications to give credit to players on the basis of team's performance (Mukherjee, 2014). A team which wins a match through runs and wickets play significant role in team's ranking, similarly weighting factors are also important in ranking teams. The author has used four methods, t-index, team rank, weighted team rank and unified weighted team rank to rank teams in test cricket (Daud and Muhammad, 2013).

International Cricket Council (ICC) ranks cricket teams jointly they have not considered region/venue strength. In region wise ranking, we have to know the importance of the teams based on the specific region/venue. We will classify the teams that in which region the team is important and in which region they have bad performance (our proposed method).

Here the author has created the weighted and directed networks between the batsmen and batsmen when thy make batting partnership with each other. Also the author has quantify the performance of each players within a team by finding out its centrality score, pagerank score, in-strength, closeness centrality and betweeness centrality. Some players which have high average's and strike rates are not necessary but the important thing is its centrality score and betweeness (Mukherjee, 2013) In this paper the author has presented network based approach in order to examine the interaction between batsmen and bowlers. In this network if the same bowler faced by two batsmen then they are connected by unweighted link between them. Similarly for bowlers, if a batsman faced by two bowlers, then they are also connected by an unweighted link. In this way the network of batsmen to batsmen and bowler to bowler is created. In this way the exact value of shortest average path, average degree and clustering coefficient is determined and then compare them with Erdos-R'enye model and configuration model (Mukherjee, 2012).

The author have tried to combine bowling rate and measure their performance whom have participated in 4th IPL. A set of predictors which are affected on the bowlers performance are determined. For that purpose different linear regression techniques identified that were used that are responsible for the bowler's performance empirically. (Bhattacharjee and Pahinka, 2012). The number of runs and wickets from which team wins are important and affect teams ranking. The weightage factor is also important when two teams' wins similar number of matched from similar kind of opponents. The hybridization of h-index and PageRank based methods for ranking cricket teams is also effective as it considers graph, non-graph weightage as well as number of runs and wickets for both the in-strength of each player is computed through weighted and directed network of batsmen and bowlers. A gradient network is generated for computing the relative importance of a player over other, then PageRank is applied on the gradient network in order to compute the importance of each player (Satyam Mukherjee, 2013).

Author defined two kinds of links among data objects within different data types: intra-type links, which represent the relationship of data objects within a homogeneous data type, and inter-type links, which represent the relationship of data objects between different heterogonous data types. Then, the author proposed a unified link analysis framework, called "link fusion", to analyze inter- and intra-type links and to bring order to data objects in different data spaces at the same time. Next, author evaluated the effectiveness of the proposed link fusion algorithm by applying it into

a real world scenario of three data spaces: Hub-page space, Authority-page space, and User space (Wensi et al., 2004).

2.2 Other Sports Ranking

There are other sports other than cricket like football/Soccer and hockey etc. in which different models and technique have been used to rank the players. Below in literature review have been discussed.

2.2.1 Soccer Ranking

The author have used a network base approach which is applied measuring the performance of each soccer players individually. The author have also hypothesize that the approach we have generalize could be beneficial in other sports environment in which the quantification of other team member's contribution is important (Radicchi, 2011).

The author have presented the clearest and most straightforward approach "Colley Matrix Method." to rank teams in soccer. In this method they have discussed that the number of matches played by a team and the number of wins and lost, assuming no tie matches (Anjela at el., 2008).

2.2.2 Hockey

The most prominent method "plus/minus" method statistic have a lot of limitations, when scoring events are relatively common and player substitution are relatively uncommon, like in basketball, standard linear regression method works very well, but for hockey neither of this condition exists for this purpose the author have used the unique of the properties of the sports. They have model for each team the scoring rate and its own semi-Markov process (Thomas at el., 2013).

Author have used different features such as face-offs, hits and some other statistics features which has to be predicted. In addition of all of them the author have to predict the goals, shots, blocked shots and missed shots as well. Other than above statistics author have estimated the expected goals in 60 minutes by the player's contribution in the team, independent of his teammates, opposite team's players and the area in which he is appointed in which he playing the game (Macdonald, 2011).

2.3 Problem Statement

We define the problem of region based ranking of cricket players as an information retrieval problem. Given $\mathbf{P} = \{P_1, \ldots, P_n\}$ be an $n \times r$ matrix representing n dimensional feature vectors of n objects, where n is total numbers of players *(P)* and r is number of regions for each team T_i. Each row of $n \times r$ matrix corresponds to one Player P_i and each column corresponds to its standing in that specific region. Our, goal is to order n players according to their points within a region and find a vector representing its ranking in that specific region.

2.4 Research Objectives

The following are the main objectives of our research.

1. Collection of data for different features of batsmen, bowlers and teams on each region separately.
2. Calculating team's batting strength and bowling strength and then combined team's strength on all five different regions separately.
3. Computed a unified features of each of the players then combining with ICC current ranking.
4. Finally using the concept of "**Link Fusion Algorithm**" and applying as a model for ranking players in cricket.
5. All the data is based on ODI matches from 2001 to 2015 recently.

2.5 Chapter Summary

In this chapter we have described general overview of literature review related to cricket sports ranking, other than cricket sports ranking (including soccer, hockey and tennis ranking), link fusion algorithm related idea and concept, formal problem statement and research objectives.

CHAPTER 3

PROPOSED SOLUTION

3. Proposed Solution

In cricket ranking, batsmen always perform against bowlers and those bowlers are playing in some team, similarly bowlers also perform against batsmen which are playing in team. So there is a strong relationship among batsmen, bowlers and teams. For the sake of that relationship we have three data spaces that are batsmen, bowlers and Teams data spaces respectively, like link fusion algorithm, in our case in each data space there are a lot of data objects. When the data object of one data space related to other data space's object, we say that relationship is heterogeneous (intertype-relationship). Here data objects means batsmen, bowlers and teams. Below diagram shows the relationship of batsmen, bowlers and teams data objects respectively.

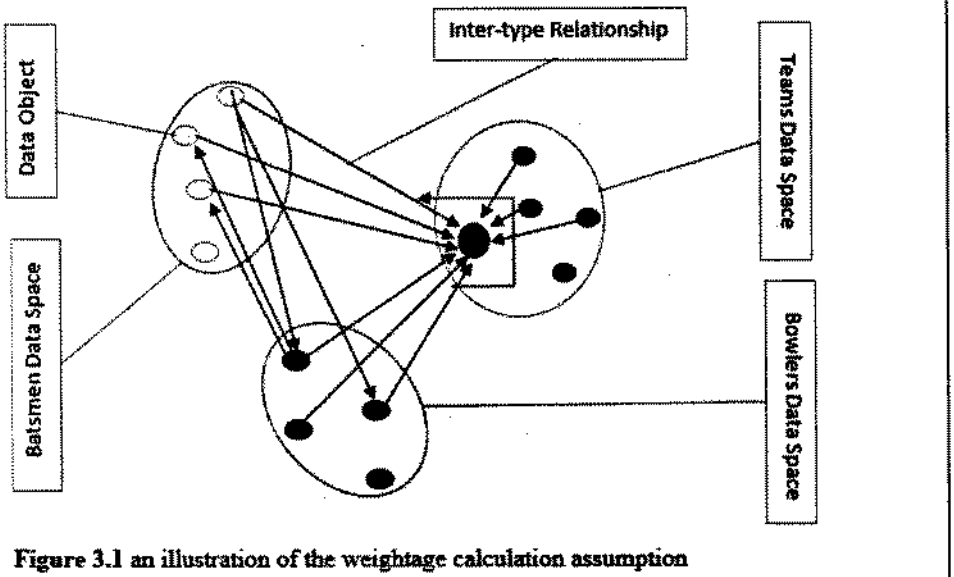


Figure 3.1 an illustration of the weightage calculation assumption

Mathematically the equations for batsmen and bowlers are given below:

$$W_{bat} = L^T_{Bat\,Bol}\, W_{Bat\,Bol} + L^T_{Bat\,Team} W_{Bat\,Team} \qquad (3.1)$$

$$W_{Bol} = L^T_{Bol\,Bat}\, W_{Bol\,Bat} + L^T_{Bol\,Team} W_{Bol\,Team} \qquad (3.2)$$

We have "$L^T_{Bat\,Bol}\, W_{Bat\,Bol}$" **and** "$L^T_{Bat\,Team} W_{Bat\,Team}$" **Matrices** for Batsmen

$$\begin{bmatrix} L'_{Bat\,Bol11} & L'_{Bat\,Bol12} \cdots & L'_{Bat\,Bol1n} \\ L'_{Bat\,Bol21} & L'_{Bat\,Bol22} \cdots & L'_{Bat\,Bol2n} \\ L'_{Bat\,Boln1} & L'_{Bat\,Boln2} \cdots & L'_{Bat\,Bolnn} \end{bmatrix} \begin{bmatrix} W'_{Bat\,Bol11} & W'_{Bat\,Bol12} \cdots & W'_{Bat\,Bol1n} \\ W'_{Bat\,Bol21} & W'_{Bat\,Bol22} \cdots & W'_{Bat\,Bol2n} \\ W'_{Bat\,Boln1} & W'_{Bat\,Boln2} \cdots & W'_{Bat\,Bolnn} \end{bmatrix}$$

$$+$$

$$\begin{bmatrix} L'_{Bat\,Team11} & L'_{Bat\,Team12} \cdots & L'_{Bat\,Team1n} \\ L'_{Bat\,Team21} & L'_{Bat\,Team22} \cdots & L'_{Bat\,Team2n} \\ L'_{Bat\,Teamn1} & L'_{Bat\,Teamn2} \cdots & L'_{Bat\,Teamnn} \end{bmatrix} \begin{bmatrix} W'_{Bat\,Team11} & W'_{Bat\,Team12} \cdots & W'_{Bat\,Team1n} \\ W'_{Bat\,Team21} & W'_{Bat\,Team22} \cdots & W'_{Bat\,Team2n} \\ W'_{Bat\,Teamn1} & W'_{Bat\,Teamn2} \cdots & W'_{Bat\,Teamnn} \end{bmatrix}$$

Similarly for bowlers we have these matrices "$L^T{}_{Bol\,Bat}\ W_{Bol\,Bat}$" **and**"$L^T{}_{Bol\,Team} W_{Bol\,Team}$"

$$\begin{bmatrix} L'_{Bol\,Bat11} & L'_{Bol\,Bat12} \cdots & L'_{Bol\,Bat1n} \\ L'_{Bol\,Bat21} & L'_{Bol\,Bat22} \cdots & L'_{Bol\,Bat2n} \\ L'_{Bol\,Batn1} & L'_{Bol\,Batn2} \cdots & L'_{Bol\,Batnn} \end{bmatrix} \begin{bmatrix} W'_{Bol\,Bat11} & W'_{Bol\,Bat12} \cdots & W'_{Bol\,Bat1n} \\ W'_{Bol\,Bat21} & W'_{Bol\,Bat22} \cdots & W'_{Bol\,Bat2n} \\ W'_{Bol\,Batn1} & W'_{Bol\,Batn2} \cdots & W'_{Bol\,Batnn} \end{bmatrix}$$

$$+$$

$$\begin{bmatrix} L'_{Team\,Bol11} & L'_{Team\,Bol12} \cdots & L'_{Team\,Bol1n} \\ L'_{Team\,Bol21} & L'_{Team\,Bol22} \cdots & L'_{Team\,Bol2n} \\ L'_{Team\,Boln1} & L'_{Team\,Boln2} \cdots & L'_{Team\,Bolnn} \end{bmatrix} \begin{bmatrix} W'_{Team\,Bol11} & W'_{Team\,Bol12} \cdots & W'_{Team\,Bol1n} \\ W'_{Team\,Bol21} & W'_{Team\,Bol22} \cdots & W'_{Team\,Bol2n} \\ W'_{Team\,Boln1} & W'_{Team\,Boln2} \cdots & W'_{Team\,Bolnn} \end{bmatrix}$$

Equation (3.1) above is used to calculate the combined performance of the batsmen against both the bowlers and teams against which he played on a specific region. Where W $_{bat}$ and W_{Bol} are the weights of a batsmen and bowlers respectively, $L^T{}_{Bat\,Bol}$ and $L^T{}_{Bol\,Bat}$ are the adjacency matrices from batsmen to bowlers and bowlers to batsmen respectively, if a batsman played against a bowler $L^T{}_{Bat\,Bol} = 1$ in the adjacency matrix from batsman to bowlers space else $L^T{}_{Bat\,Bol} = 0$, similarly in bowler to batsmen data space, batsman to teams and bowler to teams data spaces. $W_{Bat\,Team}$ And $W_{Bol\,Team}$ are the weights from batsman to teams and bowler to teams respectively. Similarly in equation (3.2), we have to calculate the combined performance of the bowler against both the batsman and teams against he played on a specific region. Where W $_{Bol}$ and W_{Bat} are the weights

of a bowlers and batsmen respectively, $L^T{}_{Bat\,Bol}$ and $L^T{}_{Bol\,Bat}$ are the adjacency matrices from bowlers to batsmen and batsmen to bowlers respectively, if a bowler is played against a batsman $L^T{}_{Bol\,Bat} = 1$ in the adjacency matrix from bowlers to batsmen space else $L^T{}_{Bol\,Bat} = 0$. $W_{Bol\,Team}$ And $W_{Bat\,Team}$ are the weights from bowler to teams and batsman to teams respectively. The equations for batsmen and bowlers are calculated below in a simple example for Asia Region (Region 3) with detail.

Suppose we have five batsmen (Bat), five bowlers (Bol) and five teams (T) respectively, (B1, B2, B3, B4, B5, b1, b2, b3, b4, b5, T1, T2, T3, T4, and T5). We compute team's batting and bowling strength and players versus player's performances in order to rank the players on Asia region.

For batting strength we first calculate eq.3.3 for each team then we will divide each score of eq. (3.3) on the maximum value out of them that will be the batting strength of each the team. The more the score the more will be the batting strength because batsmen always trying to get more average runs and strike rate against bowler he faced. Similarly for bowling strength we first calculate eq. (3.4) for each teams then we will divide each the maximum value out of them on each Team's score that will be the bowling strength of each the team. The more the score the more will be the bowling strength because bowler always trying to get wickets by conceding less number of average runs, strike rate and economy rate against batsmen he faced.

3.1 Region based Batsman, Bowlers and Teams Features in ODI Cricket

Since we have to rank players on each region that is why we have calculated different features of batsman, bowler and team on each region. Region based features of players (batsmen, bowlers) and teams are discussed below one by one.

3.1.1 Region based Team Features

We have calculated five different features related to team on each region. These features are Team Win/Loss ratio, Team's average runs score, Team Batting Strength, and Team Bowling Strength etc. on each of the five regions.

3.1.2. Region based Team's Win/Loss Ratio

We can define win/loss ratio by dividing total number of matches win on a specific region by the number of lost matches.

$$Region\ Based\ \frac{Win}{Loss}\ Ratio = \frac{Region\ Based\ Win\ Matches}{Lost\ Matchs} \quad \text{------- (3.3)}$$

e.g. Pakistan W/L Ratio on Asia Region = $\frac{130}{87}$ = 1.49

3.1.3 Region based Team's average Runs

Region based team's average runs can be defined as total number of runs scored by a team divided number of innings.

$$\text{Region based batting average runs} = \frac{Total\ No\ of\ Runs\ Scored\ of\ a\ Team\ on\ a\ Specific\ Region}{Number\ of\ Innings\ Played\ on\ a\ Specific\ Region} \quad \text{------- (3.4)}$$

Pak Batting Average Runs on Asia Region = $\frac{47964}{215}$ = 223.0883

3.1.4. Region based Team's batting Strength

Team's Batting Strength on Each Region:

$$\frac{((\text{Team player's Avg Runs} * 25)+(\text{Team's ER} * 25)+(\text{Region-wise Team's Batting avg} * 25)+(\frac{W}{L}*25))}{4} \quad (3.5)$$

Where "team player's average runs" is the average runs of a specific team on a specific region One can try different weights for team's player's average runs, team's economy rate, team's batting average and team win loss ratio (Player's Average Runs (30%), Economy Rate (20%), Team's Batting Average (20%) and Win/Loss (30%)) but in our case we set 25% to each features because all features are important for a team. By applying equation. (3.5), we have calculated all features of Team's Batting Strength.

Table 3.1 Teams Batting Strength on Asia Region

Teams	Teams W/L	Team Bat Avg	Player Bat avg	Runs/Over	Batting Strength	Team Bat Strength on Asia Region
T1	1.49	223.09	32.46	5.15	1638.70	0.92

T2	1.73	238.92	38.12	5.55	1777.01	1
T3	1.48	212.01	31.61	5.1	1563.76	0.88
T4	0.54	185.18	25.43	4.53	1347.99	0.76
T5	2.30	226.95	38.76	5.32	1708.32	0.96

By calculating team strength we have column "Batting Strength" values, then we will check the greater values in all of the values. T2 batting strength is maximum in all of the five values, so we have divided T2 batting strength on each of the Team's Batting Strength and the final column will shows the results of the batting strengths of all the teams. T5 have the best Team Strength and T4 have the lower Team Strength in all of the five Teams.

3.1.5. Region based Team's Bowling Strength

And similarly Team's Bowling Strength on Each Region:

$$\frac{((\text{Team Bol Avg Runs Concedded}*25)+(\text{Team's Bowling ER}*25)+(\text{Team's SR}*25)+(\frac{W}{L}*25))}{4} \quad (3.6)$$

Similarly for region based team's bowling strength One can try different weights for team's bowling average runs, team's economy rate, team's strike rate and team win loss ratio (team's bowling average runs (30%), team's bowling economy rate (20%), team's strike rate (20%) and team win/loss (30%)) but in our case we set 25% to each features because all features are important for a team. By applying equation. (3.5), we have calculated all features of Team's Batting Strength.

Table 3.2 Team's Bowling Strength

Teams	Teams W/L	Team Bol Avg	Team Bol ER	Team's Bol ST	Bowling Strength	Team Bol Strength on Asia Region
T1	1.49	25.86	36	4.30	422.84	1.20
T2	1.73	31.16	38.7	4.82	477.58	1.07
T3	1.48	27.02	35.8	4.52	430.15	1.18
T4	0.54	33.71	42.4	4.76	508.83	1
T5	2.30	25.10	32.2	4.66	401.65	1.27

Similarly for bowling strength we first calculate eq. (3.6) for each teams then we will divide the maximum value out of them on each team's score that will be the bowling strength of each the team. The more is the number of score the more will be the bowling strength because bowler always trying to get wickets by conceding less number of average runs, strike rate and economy rate against the batsmen he faced.

Here in table (3.2) T5 have best bowling strength and T4 have the lowest team's bowling strength in out of five teams.

3.2 Region based Batsman Features

Here we have discussed three basic features related to batsmen on each region. These are:

3.2.1 Region based Performance Index of Batsman (PIB)

We will calculate here the performance of batsmen against all bowlers in all regions

3.2.2 Region based Average of a Batsman

The batting average is defined as the number of runs scored by a batsmen in all innings divided by the number of completed innings.

3.2.3 Region based Strike Rate of a Batsman

Strike rate of a batsman is defined as the number of runs scored divided by the number of ball faced by a batsman

3.3 Region based Bowlers Features

Here we will discussed four basic features of a bowler on each region. These features are discussed below one by one.

3.3.1 Region based Quality Index of a Bowler (QIB)

We will calculate here the quality of bowler, that which bowler has quality bowling against which batsman in a specific region.

3.3.2 Region based Average of a Bowler

We will define region based average of a bawler the total number of runs conceded by a bowler divide by number of wickets taken.

3.3.3 Region based Strike Rate of a Bowler

We will define strike rate of a bowler by the number of balls bowled b a bowler on a specific region divided by the number of wickets taken.

3.3.4 Region based Economy Rate of a Bowler

Region based economy rate of a bowler can be define by number of runs conceded on a particular by the number of balls bowled with the multiplication of 6.

After calculating team's batting and bowling strength we will apply those strengths to calculate players ranking.

Batsman B1 to B5 have the runs scored, batting average and batting strike rate against bowler b1 to b5 in Asia Region.

Table 3.3 Batsmen Performance against Bowlers on Asia Region

Batsmen on Asia	**Performance Against b1**		**Performance Against b2**		**Performance Against b3**		**Performance Against b4**		**Performance Against b5**	
	Avg	**SR**	**Avg**	**SR**	**Avg**	**SR**	**Avg**	**SR**	**Avg**	**SR**
B1	27	65	53	125	36	77	34	107	39	86
B2	50	73	35	88	44	67	46	85	36	91
B3	18	98	41	153	37	133	25	77	19	78
B4	61	72	44	84	47	83	40	73	27	89
B5	37	102	26	136	56	91	31	125	51	73

Next, we set the weightages of all the batsman against all bowlers and then net batting weightages on Asia Region by using equation (3.1).

Net batting performance is calculating by taking weightages of the performances of a batsmen

Table 3.4 Net batting weightages of Batsmen against all bowlers

Batsmen	Weightage vs b1	Weightage vs b2	Weightage vs b3	Weightage vs b4	Weightage vs b5	$W_{batsman\ bowler}$
B1	2110	4090	2620	3160	2890	2974
B2	2960	2810	2660	3080	2900	2882
B3	2500	4290	3770	2290	2130	2996
B4	3270	3000	3070	2660	2590	2918
B5	3150	3500	3500	3430	2990	3314

Table 3.4 shows the net batting weightages of each batsmen against all bowlers. It is obvious in the last column that B3 has best performance against all other batsmen.

Weightage of a batsman against each bowler on a particular region is calculated by the following formula:

$$\frac{((60*\text{batting avg against a specific bowler on a specific region})+(40*\text{batting SR on a specific region}))}{2} \quad (3.7)$$

One can try different weights for batting average and batting strike rate (50(batting average) and 50(batting strike rate) but in our case we assign weightages to an average and strike rate 60 % and 40 % respectively because batsman mostly focused on the number of runs scored instead of strike rate which is the important factor for a batsman.

After calculating the weightages of each batsman against all bowlers, we calculate its average weightage, in order to compare each batsman against all other batsmen.

Now coming towards the weightage of a 'batsman against teams' against which he played. It is obvious that a batsman score against the team's bowling line up. A batsman is more important which he scored more average runs with a high strike rate and the team's bowling strength is high on the other hand a team's bowlers will try to concede less average runs and strike rate. In first step we will find out the average score and SR of a batsman against each team then calculating its weightage using equation (3.7), just changing team instead of bowler then in the second step we will multiply team strength one by one by using equ. (3.6) to the batsman performance against that specific team on a particular region, then finally we added all the performances of a batsman against all the teams.

Table 3.5 Batsmen Average and SR against all Teams on Asia Region

Batsmen on Asia	Performance Against T1		Performance Against T2		Performance Against T3		Performance Against T4		Performance Against T5	
	Avg	SR	Avg	SR	Avg	SR	Avg	SR	Avg	SR
B1	27	81	51	105	36	79	41	101	38	98
B2	59	74	35	88	44	67	46	85	36	77
B3	29	91	41	153	37	133	25	77	19	78
B4	51	77	44	84	47	83	40	73	21	112
B5	39	111	26	136	56	91	35	108	51	76

Batsmen weights against all teams calculated using equation (3.7)

Table 3.6 Batsmen weights against all Teams on Asia Region

Batsmen	Weightage vs T1	Weightage vs T2	Weightage vs T3	Weightage vs T4	Weightage vs T5
B1	2430	3630	2660	3250	3100
B2	3250	2810	2660	3080	2620
B3	2690	4290	3770	2290	2130
B4	3070	3000	3070	2660	2870
B5	3390	3500	3500	3210	3050

Table 3.4 is the performances of the batsmen against the teams, but we have not considered here the strength of each team. So to calculate the accurate performance of each batsman against each team we will multiply the weighted score of a batsmen to the Team's bowling strength against a batsman scored. The final result of $W_{batsman}$ is given below in table 3.6 by finding the weighted average of both the $W_{batsman\ bowler}$ and the $W_{batsman\ Team}$ by equation 6 and 7 respectively.

$$\frac{((50*W\ batsman\ bowler)+(50*W\ batsman\ Team))}{2} \quad \text{---------- (3.8)}$$

W bats Team = Batsman Weighted Score against Each Team * Bowling Strength of tht Team --- (3.9)

Table 3.7 Batsmen weightages using Team's Bowling Strength

Batsmen	Weightage vs T1	Weightage vs T2	Weightage vs T3	Weightage vs T4	Weightage vs T5
B1	2430*1.20	3630*1.07	2660*1.18	3250*1	3100*1.27
B2	3250*1.20	2810*1.07	2660*1.18	3080*1	2620*1.27
B3	2690*1.20	4290*1.07	3770*1.18	2290*1	2130*1.27
B4	3070*1.20	3000*1.07	3070*1.18	2660*1	2870*1.27
B5	3390*1.20	3500*1.07	3500*1.18	3210*1	3050*1.27

Table 3.8 Batsmen net weightages using Team's Bowling Strength

Batsmen	Weightage vs T1	Weightage vs T2	Weightage vs T3	Weightage vs T4	Weightage vs T5	$W_{batsman\ Team}$
B1	2924.15	3867.50	3146.54	3250	3927.18	3423.08
B2	3910.90	2993.85	3146.54	3080	3319.10	3290.08
B3	3237.02	4570.69	4459.57	2290	2698.35	3451.13
B4	3694.30	3196.28	3631.54	2660	3635.80	3363.58
B5	4079.37	3729	4140.19	3210	3863.83	3804.48

Finally we add the weightages of batsmen against all bowlers and teams, which shows the net batting performances (attribute vector) of each batsmen. From the table 3.6, batsman B5 has the highest weighted score, so is the top batsman out of five batsmen.

Table 3.9 Batsmen weighted scores against bowlers and teams

Batsmen	$W_{batsman\ bowler}$	$W_{batsman\ Team}$	$W_{batsman}$
B1	2974	3423.07	159926.9
B2	2882	3290.08	154302
B3	2996	3451.12	161178.2
B4	2918	3363.58	157039.6
B5	3314	3804.48	177962

Table 3.10 Batsmen Ranks

Batsmen	$W_{batsman\ bowler}$	$W_{batsman\ Team}$	$W_{batsman}$	Ranks
B5	3314	3804.48	177962	1
B3	2996	3451.13	161178.2	2
B1	2974	3423.07	159926.9	3
B4	2918	3363.58	157039.6	4
B2	2882	3290.08	154302	5

Next all the weightages in table 3.10 above is sorted which shows ranks of all the batsmen in which B5 (Batsman) is top in the rank and B2 (Batsman) is at the end in the list of ranking on Asia Region.

Similarly for the bowlers the similar steps have been applied and ranks of all the bowlers are calculated.

Bowler b1 to b5 have conceded average runs, Strike rate and Econ Rate against Batsman B1 to B5 respectively on Asia Region.

Table 3.11 Bowlers Average, SR and ER conceded against all Batsmen on Asia Region

Bowlers on Asia Region	Performance vs B1			Performance vs B2			Performance vs B3			Performance vs B4			Performance vs B5		
	Avg	ER	SR	Avg	ER	SR	Avg	ER	SR	Avg	ER	SR	Avg	ER	SR
b1	27	6.5	31	53	3.9	28	36	7.7	29	34	7	24	39	6	34
b2	50	7.3	18	35	6.1	16	44	6.7	18	46	8.5	21	36	4	28
b3	18	4.8	25	41	5.9	39	37	5	36	51	3.7	51	19	7.8	41
b4	61	7.2	58	44	4.3	23	47	3	26	40	7.3	40	27	3	26
b5	37	3.9	21	26	7.0	21	56	4.2	17	27	5	27	51	5.2	16

Bowler weightages against all teams on Asia Region is calculated using equation 8.

$$\frac{((40*\text{bowler avg runs conceded})+(40*\text{bowling SR})+(20*\text{bowling ER}))}{3} \quad \text{---------- (3.10)}$$

One can try different weights for bowling average, bowling Strike Rate and bowling ER, but in our case we have given 40% weightage to average, 20% to SR and 40% to ER, because average and ER are most important for a bowler to control a batsman than strike rate.

After calculating the weightages of each bowler against all batsmen, we calculate its average weightage, in order to compare each bowler against all other bowlers as in table 3.12.

Table 3.12 Net bowling weightages of Bowlers against all batsmen

Bowler	Weightage vs B1	Weightage vs B2	Weightage vs B3	Weightage vs B4	Weightage vs B5	$W_{bowler\ batsmen}$
b1	653.33	945.33	776	706.67	826.67	781.6
b2	884	654.67	796	866.67	720	784.2667
b3	470.67	885.33	800	1069.33	630.67	771.2
b4	1296	797.33	840	897.33	573.33	880.8
b5	685.33	580	916	606.67	856	728.8

In table 3.12 column ($W_{bowler\ batsman}$) is the performance of bowlers against all batsmen. If the score of the bowler is low this means that bowler is best in his rank, because bowlers always trying to concede minimum average runs, with strike rate and economy rate to control match.

Now coming towards the weightage of a 'bowlers against teams' against which he played. It is obvious that a bowler takes wickets against the team's bating line-up with the average, SR and ER. A bowler is more important which he takes more number of wickets against a team having strong batting line-up with a low bowling average, SR and ER on the other hand a team's batsmen trying to score runs with high average and strike rate. In first step we will find out the average score conceded by a bowler, with SR and ER against each team's batsmen then calculating its weightage using equation. (3.10), just changing team instead of batsman then in the second step we will multiply team strength one by one by using equation. (3.5) to the bowler performance against that specific team on a particular region, then finally we take average of all the performances of a bowlers against all the teams. Then we will say that $W_{bowler\ team}$ (Bowler performance against teams).

Table 3.13 Bowlers Average, SR and ER conceded against all Teams on Asia Region

Bowlers on Asia Region	Performance vs T1			Performance vs T2			Performance vs T3			Performance vs T4			Performance T5		
	Avg	ER	SR	Avg	ER	SR	Avg	ER	SR	Avg	ER	SR	Avg	ER	SR
b1	40	6.5	29	43	4.1	31	29	6.2	27	34	7	34	39	5	24
b2	21	6.3	20	45	6.3	19	51	6.7	22	36	6.5	31	36	4	28
b3	34	5.8	35	51	5.1	31	33	5	36	51	5.7	41	39	6.8	43
b4	57	6.2	38	31	5.2	27	49	4.5	25	40	6.3	45	27	3	24
b5	37	4.9	41	29	7.0	25	58	4.2	20	37	6	32	31	5.2	26

Table 3.14 Net bowling weightages of Bowlers against all Teams

Bowler	Weightage vs T1	Weightage vs T2	Weightage vs T3	Weightage vs T4	Weightag e vs T5
b1	813.33	834.67	649.33	773.33	746.67
b2	497.33	810.67	916	773.33	720
b3	764	954.67	746.67	1029.33	897.33
b4	1096	662.67	880	917.33	560
b5	832	646.67	962.67	786.67	656

Table 3.14, shows the net bowling weightages of bowlers against all the teams on Asia region, but here the batting team strength of each team is not considered so for, in the table 3.15 we have considered the batting strength of each team with the bowler's weightage.

So to calculate the accurate performance of each bowlers against each team we will multiply the $W_{\text{bowler Team}}$ value by the Team's strength against a bowler bowled. The final result of W_{bowler} is given below in table 4.16 by finding the weighted average of both the $W_{\text{bowler batsman}}$ and the $W_{\text{bowler Team}}$ by this formula:

$$\frac{((50*W_{\text{bowler batsman}})+(50*W_{\text{bowler Team}}))}{2} \quad \text{---------- (3.11)}$$

Table 3.15 Bowlers weightages using Team's Batting Strength

Batsmen	Weightage vs T1	Weightage vs T2	Weightage vs T3	Weightage vs T4	Weightage vs T5
b1	813.33*0.93	834.67*1	649.33*0.88	773.33*0.76	746.67*0.96

b2	497.33*0.93	810.67*1	916*0.88	773.33*0.76	720*0.96
b3	764*0.93	954.67*1	746.6667*0.88	1029.33*0.76	897.3333*0.96
b4	1096*0.93	662.67*1	880*0.88	917.33*0.76	560*0.96
b5	832*0.93	646.67*1	962.6667*0.88	786.66*0.76	656*0.96

Table 3.16 Bowlers net weightages using Team's Batting Strength

Bowlers	**Weightage vs T1**	**Weightage vs T2**	**Weightage vs T3**	**Weightage vs T4**	**Weightage vs T5**	**$W_{bowlers\ Teams}$**
b1	752.35	834.67	571.41	586.63	715.68	692.15
b2	460.04	810.67	806.08	586.63	690.12	670.71
b3	706.71	954.67	657.06	780.82	860.09	791.87
b4	1013.82	662.67	774.40	695.86	536.76	736.70
b5	769.61	646.67	847.14	596.74	628.77	697.79

In table 3.17, we have combined the weights of bowlers against batsmen ($W_{bowler\ batsmen}$) and bowlers against teams ($W_{bowler\ Team}$). In the final column net bowlers weights ($W_{bowlers}$) is given.

Table 3.17 Bowlers weightage against Batsmen and Teams

Bowlers	**$W_{bowler\ batsmen}$**	**$W_{bowler\ Team}$**	**$W_{bowlers}$**
b1	781.6	692.15	36843.64
b2	784.27	670.71	36374.31
b3	771.2	791.87	39076.77
b4	880.8	736.70	40437.5
b5	728.8	697.79	35664.69

Next we have sorted all weightages of bowlers in ascending order then, clearly we will say that bowler b4 is the top ranked player out of five and b5 is 5^{th} in the list of five bowlers on Asia Region in our example. It should be noted that a bowler is more valuable and worthy that have low points out of all because when a bowler perform well he will try to concede low average runs, SR and ER against batsmen and teams so we have inversed all the ranking score of bowlers which clearly shows bowler's ranks.

Table 3.18 Bowlers Ranks

Bowlers	$W_{bowler\ batsmen}$	$W_{bowler\ Team}$	$W_{bowlers}$	Ranks
b4	880.8	736.70	40437.5	1
b3	771.2	791.87	39076.77	2
b1	781.6	692.15	36843.64	3
b2	784.27	670.71	36374.31	4
b5	728.8	697.79	35664.69	5

3.4 Chapter Summary

Here in this section we have studied proposed solution of our research having basic features of Teams (batting strength and bowling strength, team's win/loss ratio and team's batting average runs scored on etc.) on each region. We have also calculated Batsmen and Bowlers features in each region. Based on those features we have calculated ranking of players (i.e. Batsmen and Bowlers) on each separate regions.

CHAPTER 4

EXPERIMENTS

4. Experiments

In this chapter we will discuss region based datasets, preprocessing steps, ICC current ranking, and a Unified Players Ranking to our experimentation and the results and discussion in detail disjointedly.

4.1 Data Set

We have taken the dataset for our experiments from espncricinfo website and archive from 01 January 2001 to 29 April 2015 and ODI ranking points system from ICC, as existing method [25]. ICC has divided all cricket teams in Full, Associate and Affiliate members, in which full members have the authority to play all type of matches that is T-20, ODI and Test cricket. We have just ranked players of the full member countries in ODI cricket format. Full members countries are ten countries Pakistan, India, Sri Lanka, Bangladesh, Australia, New Zealand, England, West Indies, South Africa and Zimbabwe, in which Bangladesh and Zimbabwe are considered weak teams and all others are strong teams. We used two categories in ODI cricket format that are batting and bowling. Both categories have their own data variables and features. There is no unique dataset exists from which we measure the results of our proposed and the existing method to rank cricket players in order to calculate the accuracy of precision, recall or f-measure. Instead, we have measured our proposed method subjectively with the existing ICC ranking to show the performance of each player on each region and then finally add all the performances of all the teams on all the regions, then compare with the ICC ranking method. Also we have discussed the subjective method with the help of our university players and our supervisor, which are most familiar and interested with the cricket.

4.2 Pre-processing steps

Before taking the dataset for calculation using our following features we use some preprocessing steps to get more related and remove unnecessary records.

1. For collecting our data from the dataset, we have included record of the players (batsmen & bowlers) from Jan 2001 to Apr 2015.
2. The data we have taken from the dataset is based on each region separately (Europe, Oceania, Asia, Africa, and Americas Regions).

3. The data we have taken is related to ODI cricket format.
4. We have included those players which have played at least 10 ODI matches and currently playing cricket on separate regions.
5. We have taken the data of 150 to 200 batsmen and similarly 150 to 200 bowlers on each regions, according to the record matches of the region.
6. We have calculated team strength of batting and bowling separately and then combined team strength on each region.
7. We have computed a Unified Players Ranking and compare with ICC current Ranking.
8. Finally we have applied the concept of Link Fusion Algorithm as a model for ranking.

4.3 Results and Discussion

This section shows the detail comparison of all the players (e.g. batsmen and bowlers) on each separate regions. We compare the current ICC ranking with our Region based ranking. The results shows in two forms **1)** in tabular form and in **2)** graphical form. We have applied the concept of Link Fusion Algorithm which ranks all the batsmen and bowlers based on each region separately. We have categorized ranking of the batsmen and bowlers on each region one by one in the next discussion.

ICC Current Ranking

ICC ranks players on the basis of pre-programed algorithm. No one knows about the algorithm else ICC members. Missing a match reduce only a half per cent (½%) of their points for their country. All ODI matches are considered equal, except for ICC Cricket World Cup matches, where good performances gain extra credit. Big scores or wicket hauls against very weak nations get much less credit than the same performances against the main ODI countries. Some of the factors for batsmen and bowlers have been used are discussed below.

4.3.1 Batsmen Ranking Factors

ICC gives points to batsmen based on some pre-defined algorithm to which a normal human being not familiar. Some of the principles have been given by which ratings take place:

- Number of runs scored by a batsmen.

- Ratings of the opposing bowling line-up; the more value is given to the batsman's innings if the combined ratings of the attack is higher.
- A player scores 100 runs in a match where all teams scored 500 is worth less than 100 runs in a match where all teams were bowled out for 200. And if a team scores 500 in the first innings and 200 in the second innings, a century in the second innings will get more credit than in the first innings (because the general level of run scoring was higher in the first innings)
- A Batsman scores in a victory receive a bonus points. That bonus will be higher for highly rated opposition teams (i.e. win bonus against the current Australia team is higher than the bonus against Zimbabwe.)
- Batsmen gain significant credit for rapid scoring. They only get a small amount of credit for being not out (because a not out batsman is, by definition, batting at the end of the innings when the value of his wicket is low)

4.3.2 Bowlers Ranking Factors

ICC also gives points to bowlers based on some pre-defined rules to which a normal human being not familiar. Some of the principles have been given by which ratings take place.

- Wickets taken and runs conceded by a bowler (bowling average) perform a significant role.
- The level of run-scoring in the match; bowling figures of 3-50 in a high-scoring match will boost a bowler's rating more than the same figures in a low-scoring match
- Bowlers who bowl a large number of overs in the match get some credit, even if they take no wickets.
- Bowlers who get most number of wickets in a victory receive a bonus points. That bonus will be higher for highly rated opposition teams.
- Bowlers gain significant credit for bowling low economy rate.

4.3.3 Comparison of Current ICC Players Ranking with Region based Ranking

ICC rated Players on a scale of 0 to 1000 points. If a player's performance is improving on the basis of his past record, his point's increases; if his performance is declining his points will go down.

Here we have given the ICC Current batsmen and bowler's top ten ranking list which we compare to our Region based players ranking. By comparing batsmen ranking in both methods, A.B. de Villiers is top in the list of ICC batsmen ranking and also he comes in most of the region in top ten list. AB ed Villiers is ranked 1st, 3rd, 25th, 1st, and 7th, on Asia, Africa, Europe, Oceania and America Regions respectively. H.M Amla, V. Kohli, T.M. Dilshan, S. Dhawan, M. S. Dhoni, GJ Maxwell and GJ Bailey are 3rd, 4th, 5th, 6th, 8th, 9th and 10th respectively in the current ICC batsmen ranking also they are in good position in region based top ten ranking. K. C. Sangakara and K.S. Williamson on the other hand have best position in current ICC batsmen ranking, although they both have good position in region based batsmen ranking but they are not included in region based top ten ranking list.

Table 4.1. ICC Batsmen Ranking

1	A.B. de Villiers	SA	902
2	K.C. Sangakkara	SL	860
3	H.M. Amla	SA	828
4	V. Kohli	IND	822
5	T.M. Dilshan	SL	793
6	S Dhawan	IND	756
7	K.S. Williamson	NZ	733
8	M.S. Dhoni	IND	731
9	G.J. Maxwell	AUS	720
10	G.J. Bailey	AUS	697

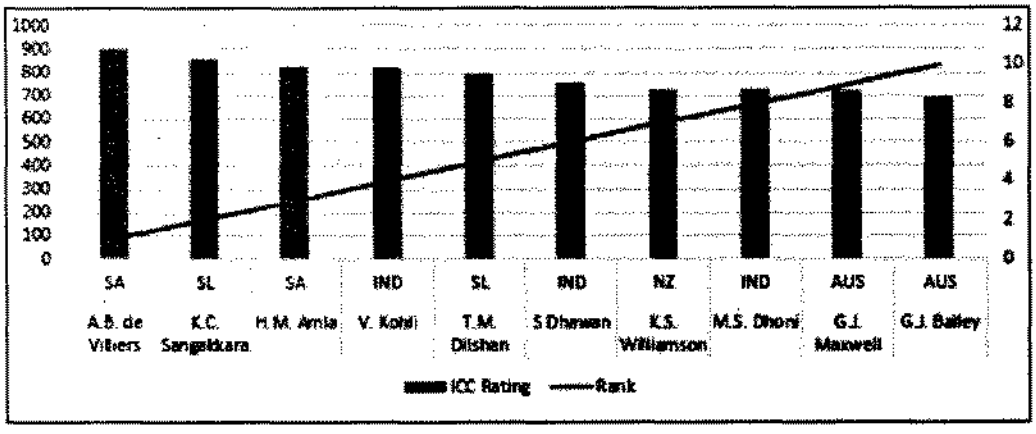


Figure 4.1: ICC Batsmen Ranking

By comparing ICC bowlers with Region Based ranking most of the ICC top ten bowlers also comes in region based top ten ranking. By comparing batsmen ranking in both methods, M. A. Starc is top in the list of ICC bowlers ranking and also he comes in most of the region in top ten list. M. A. Starc is ranked 1st, 2rd, and 1st in ICC, Asia, and on Oceania Regions respectively and in the rest of the regions he has good ranking but not comes in top ten ranking. Imran Tahir, D. W. Steyn, M. G. Johnson, Saeed Ajmal, and M. Morkel are 2rd, 3th, 6th, 7th, and 10th respectively in the current ICC bowlers ranking also they are in good position in region based top ten ranking. S. P. Narine, T. A. Boult, Shakib ul Hasan and J. M. Anderson on the other hand have best position in current ICC bowlers ranking, although they have good position in region based ranking but they are not included in region based top ten ranking list.

Table 4.2. ICC Bowlers Ranking in Cricket

1	M.A. Starc	AUS	783
2	Imran Tahir	SA	734
3	D.W. Steyn	SA	717
4	S.P. Narine	WI	709
5	T.A. Boult	NZ	693
6	M.G. Johnson	AUS	693
7	Saeed Ajmal	PAK	682

8	Shakib Al Hasan	BAN	676
9	J.M. Anderson	ENG	667
10	M. Morkel	SA	666

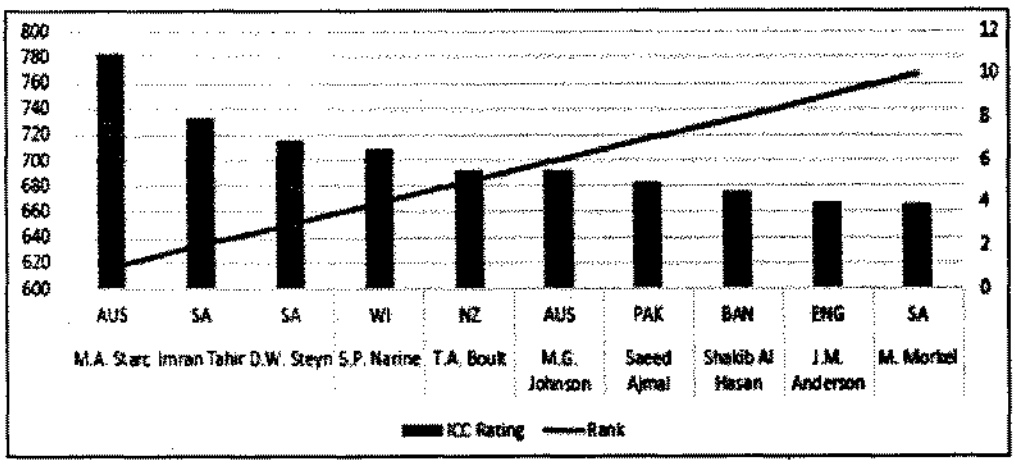


Figure 4.2: ICC Bowlers Ranking

4.3.4 Batsmen ranking based on Asia Region

On Asia region we have taken 161 top batsmen out of ten full member countries in which AB De Villiers (SA) is the top batsmen in all players, second is Verath Kohli.(Ind) 7th is Shahid Afridi (Pak) and 10th is H.M Amla (SA). In this region we have shown the top ten batsmen list in which 4 are Indian batsmen, 3 of them are Australian, 2 of them are South African and just a single batsmen is of Pakistan due to getting more number of average runs and strike rates against strong bowlers and stronger teams.

We have given weightages to each batting average and strike rate 60% and 40% respectively. In all of these batsmen even Shahid Afridi has very low batting average i.e. 25.25 but highest strike rate i.e. 128.81, also he perform well against strong bowlers and teams. On the other hand GJ Maxwell has both batting average and Strike Rate good as compare to Shahid Afridi that is why he is 6th but he has still low batting average as compare to all other batsmen in the list of ranking below on Asia Region.

Table 4.3 Batsmen Ranking Based on Asia Region

Players Name	Country	Players Rank Score on Asia Region	Rank
AB de Villiers	SA	204429.91	1
V Kohli	Ind	199627.85	2
GJ Bailey	Aus	198440.96	3
S Dhawan	Ind	194537.35	4
MS Dhoni	Ind	192048.21	5
GJ Maxwell	Aus	192012.16	6
Shahid Afridi	Pak	187042.22	7
SR Watson	Aus	186078.71	8
V Sehwag	Ind	184978.01	9
HM Amla	SA	178474.78	10

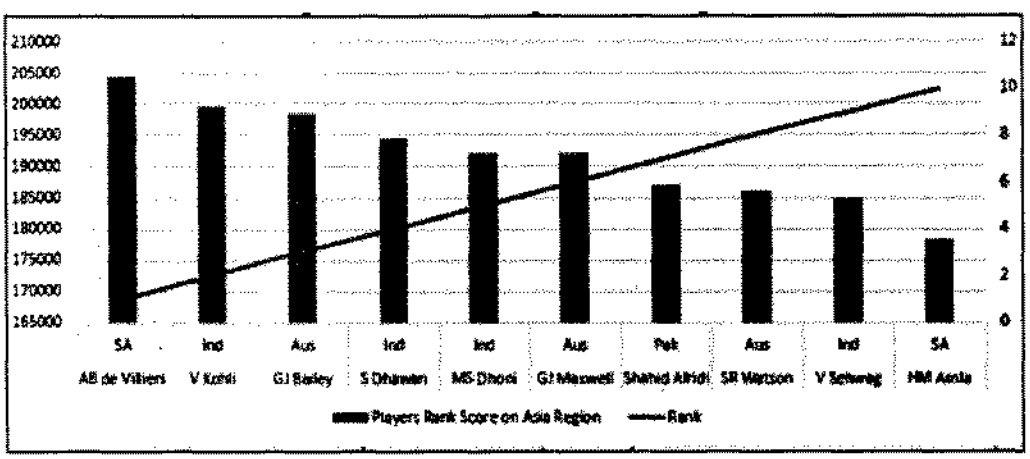


Figure 4.3: Batsmen Ranking Based on Asia Region

4.3.5 Bowlers ranking based on Asia Region

On Asia region we have taken 170 top bowlers out of ten full member countries in which Imran Tahir (SA) is the top performer in all players, second is MA Starc (Aus) and 7th is Saeed Ajmal (Pak). On this region we have shown the top ten bowlers in the list in which 3 of them are each of Australia and South Africa, and one of each England, New Zealand, Pakistan and Sri Lanka in the list of top ten bowlers.

We have given 40%, 40%, and 20% weightages to each bowler's average, bowler's economy rate and bowlers strike rate respectively. Wickets taken ratio and runs conceded per over are very important that is why we have given 40% to each.

In all of them Imran Tahir has conceded 15.41, 4.27, and 21.6 average runs, economy rate and strike rate respectively and also he perform well against strong bowlers and teams. On the other hand Saeed Ajmal has conceded 21.06, 3.99, and 31.6 bowling average, economy rate and strike rate respectively. Saeed Ajmal is a little expensive in his strike rate that is why he in 7th in the list. KD Mills and SCJ Broad have both highest strike and economy rate due they are 9th and 10th in ranking list.

Table 4.4 Bowlers Ranking Based on Asia Region

Players Name	Country	Players Rank Score on Asia Region	Rank
Imran Tahir	SA	27226.44	1
MA Starc	Aus	26181.84	2
MG Johnson	Aus	25920.24	3
LL Tsotsobe	SA	25322.16	4
BAW Mendis	Sri	25220.08	5
R McLaren	SA	24627.80	6
Saeed Ajmal	Pak	24043.88	7
MJ Clarke	Aus	23703.95	8
KD Mills	NZ	22007.67	9
SCJ Broad	Eng	19035.40	10

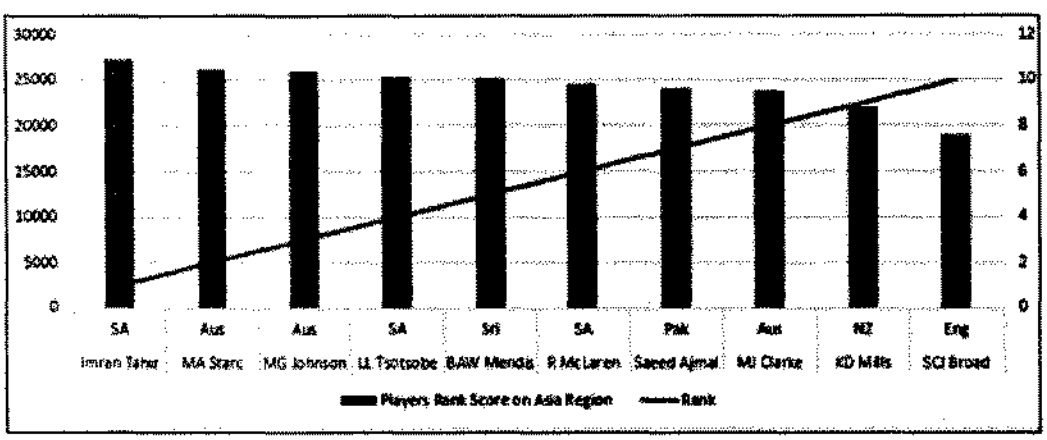


Figure 4.4: Bowlers Ranking Based on Asia Region

4.3.3 Batsmen ranking based on Africa Region

On Africa region we have taken 172 top batsmen out of ten full member countries in which AD Russell (WI) is the top batsmen in all players, second is H. M. Amla (SA), 3rd is AB de Villier

(SA), 9th and 10th are L. D. Chandimal and T. M. Dilshan (Sri) respectively. We have shown Africa region based top ten batsmen ranking results in which 2 of them of each Australia, South Africa, Sri Lanka and Pakistan and one of each India and West Indies.

We have given weightages to each batting average and strike rate 60% and 40% respectively. In all of these batsmen although KP Pieteresen has second best batting average i.e. 87.14 on Africa region but most of the score he made against weak bowlers and weak teams, that is why he is number seven in the list of ranking.

Table 4.5 Batsmen Ranking Based on Africa Region

Players Name	Country	Players Rank Score on Africa Region	Rank
AD Russell	WI	215203.11	1
HM Amla	SA	204159.87	2
AB de Villiers	SA	204048.76	3
MR Marsh	Aus	202165.34	4
SR Watson	Aus	195909.33	5
YK Pathan	Ind	195700	6
M Hafeez	Pak	189683.38	7
M Yousuf	Pak	189321.25	8
LD Chandimal	Sri	186900.36	9
TM Dilshan	Sri	183091.67	10

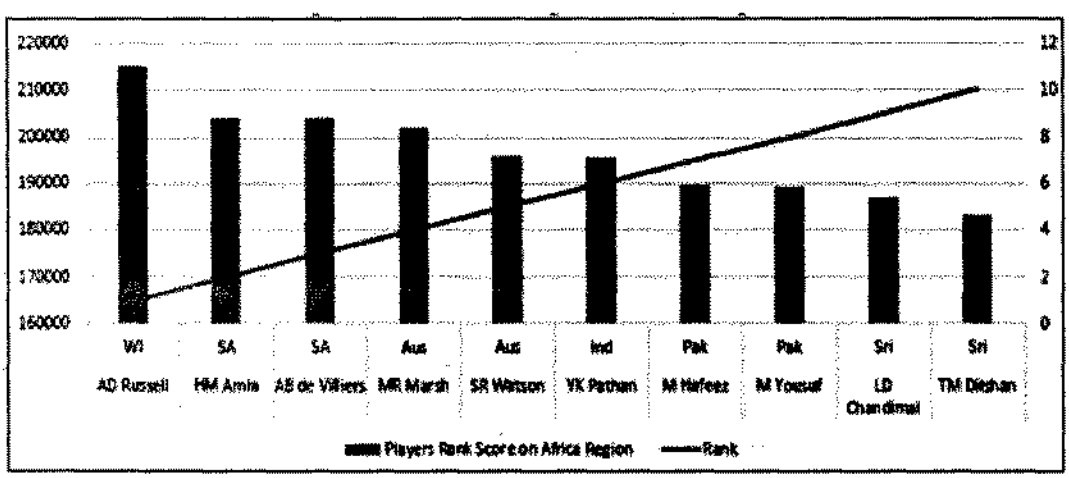


Figure 4.5: Batsmen Ranking Based on Africa Region

4.3.6 Bowlers ranking based on Africa Region

On Africa region we have taken 150 top bowlers out of ten full member countries and computed region based ranking in which Mashrafe Mortaza (Ban) is the top performer in all players, second is BAW Mendis (Sri), and Saeed Ajmal (Pak) is the 10th bowler in this list. In this region we have shown the top ten bowlers in the list in which 4 of them are to South African, 2 of them are each of Pakistan and India and 1 of each Sri Lanka and Bangladesh.

We have given 40%, 40%, and 20% weightages to each bowling average, economy rate and strike rate. Wickets taken ratio and runs conceded per over are very important that is why we have given 40% to each.

Table 4.6 Bowlers Ranking Based on Africa Region

Players Name	Country	Players Rank Score on Africa Region	Rank
Mashrafe Mortaza	Ban	32055	1
BAW Mendis	Sri	31801.67	2
A Mishra	Ind	31755	3
VD Philander	SA	31368.33	4
Mohammad Irfan	Pak	31068.33	5
Z Khan	Ind	30528.33	6
Imran Tahir	SA	30228.33	7
DW Steyn	SA	28821.67	8
R McLaren	SA	28688.33	9
Saeed Ajmal	Pak	28481.67	10

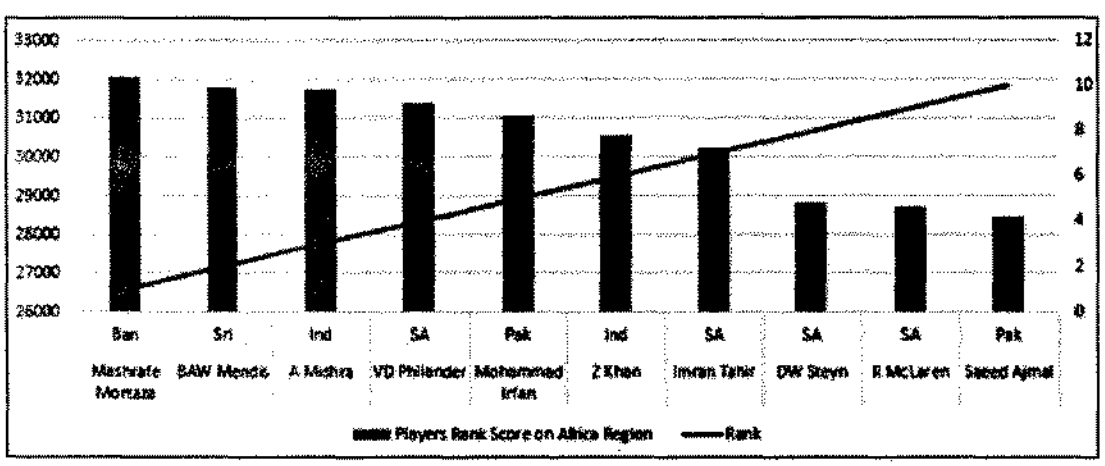


Figure 4.6: Bowlers Ranking Based on Africa Region

4.3.7 Batsmen ranking based on Europe Region

On Europe region we have taken 170 batsmen out of ten full member countries in which RA Jadeja (Ind) is the top batsmen in all players, second is MJ Guptill (NZ), 9th is SK Raina (Ind) and 10th is WW Tharanga (Sri). On this region top ten batsmen list have shown in which 3 of them are Indian, 2 are Australian and one of each New Zealand, Pakistan, South Africa, Sri Lanka and West Indies.

We have given weightages to each batting average and strike rate 60% and 40% respectively. In all of these batsmen although MJ Guptill has best Strike Rate but he has low batting average as compare to RA Jadeja, that is why he is on second in the list of ranking. On the other hand Misbah Ul Haq has very good batting average, but he has lowest strike rate in all other 9 batsmen. That is why he goes to 8th in the list of ranking.

Table 4.7 Batsmen Ranking Based on Europe Region

Batsmen Name	Country	Batsmen Rank Score on Europe Region	Rank
RA Jadeja	Ind	251890	1
MJ Guptill	NZ	246220	2
S Dhawan	Ind	213350	3
DJG Sammy	WI	209030	4
AJ Finch	Aus	200435	5
SE Marsh	Aus	188285	6
HM Amla	SA	180720	7
Misbah-ul-Haq	Pak	174640	8
SK Raina	Ind	174130	9
WU Tharanga	Sri	173590	10

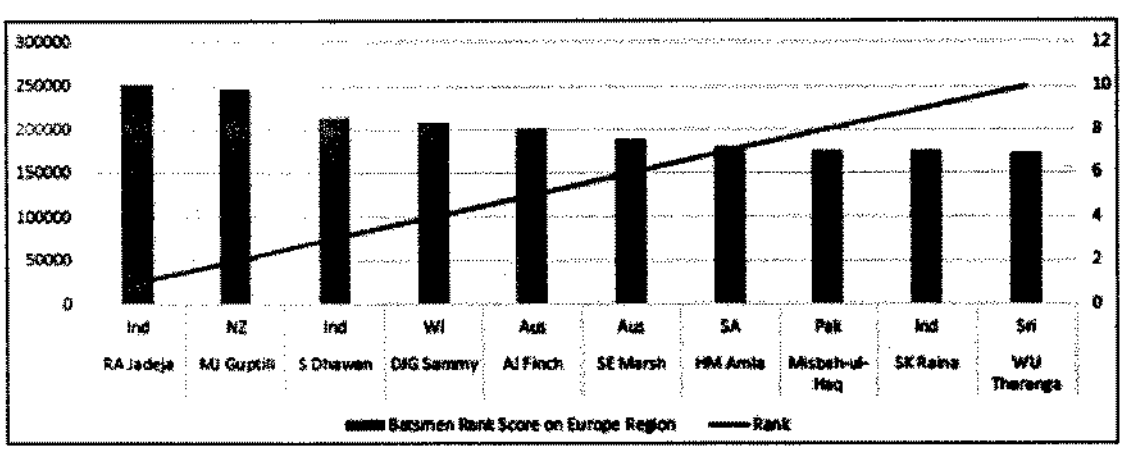


Figure 4.7: Batsmen Ranking Based on Europe Region

4.3.8 Bowlers ranking based on Europe Region

On Europe region we have taken 150 top bowlers out of ten full member countries and find out their ranking in which MJ McClenaghan (NZ) is the top performer in all bowlers, second is Saeed Ajmal (Pak), and 10th is Harbhajan Singh (Ind). On this region we have shown the top ten bowlers in the list in which 3 of them are Pakistani bowlers, 2 of them are each England, India and New Zealand bowlers and just a single bowler belong Sri Lanka.

Table 4.8 Bowlers Ranking Based on Europe Region

Bowlers Name	Country	Bowlers Rank Score on Europe Region	Rank
MJ McClenaghan	NZ	30821.67	1
Saeed Ajmal	Pak	30075	2
RA Jadeja	Ind	29768.33	3
Junaid Khan	Pak	29728.33	4
TG Southee	NZ	29595	5
RAS Lakmal	Sri	29241.67	6
JC Tredwell	Eng	28248.33	7
Umar Gul	Pak	27948.33	8
CJ Jordan	Eng	27821.67	9
Harbhajan Singh	Ind	21148.33	10

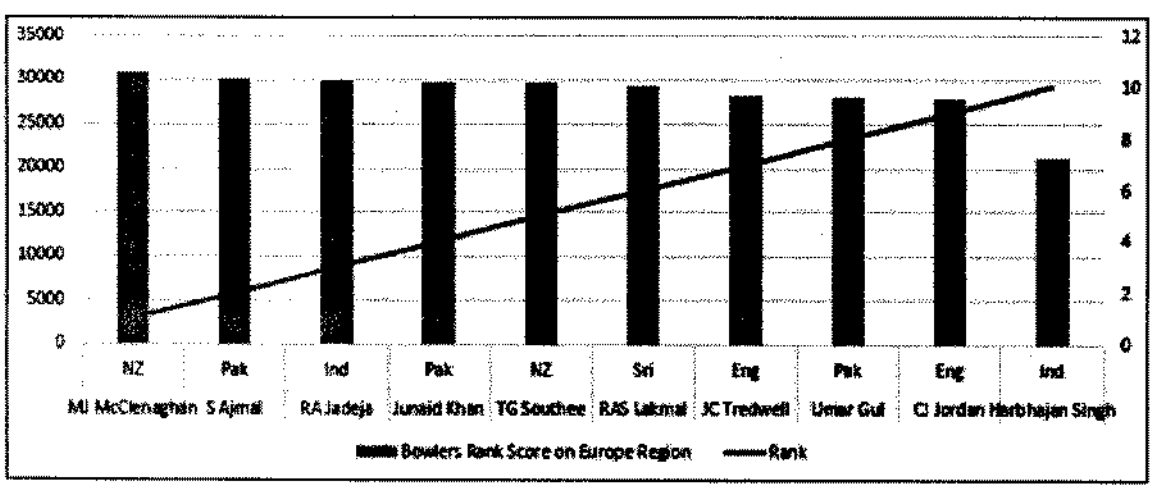


Figure 4.8: Bowlers Ranking Based on Europe Region

4.3.9 Batsmen ranking based on Oceania Region

On Oceania region we have taken 142 top batsmen out of ten full member countries in which AB De Villiers (SA) is the top batsmen in all players, second is SC Williams (Zim) and 3rd is AD Russell (WI) 10th is DA Miller (SA).

We have given weightages to each batting average and strike rate 60% and 40% respectively. In all of these batsmen even Shahid Afridi and AD Russel have very low batting average i.e. 28.6 and 28.5 but highest strike rate i.e. 145.9 and 159.81 respectively, and also they both perform well against strong bowlers and teams. On the other hand AB de Villiers has best batting average as compare to all other batsmen on that region and his strike rate is also well that is why he is top in the list of ranking on Oceania Region.

Table 4.9 Batsmen Ranking Based on Oceania Region

Batsmen Name	Country	Batsmen Rank Score on Oceania Region	Rank
AB de Villiers	SA	270820	1
SC Williams	Zim	222625	2
AD Russell	WI	214485	3
BRM Taylor	Zim	208385	4
Shahid Afridi	Pak	200725	5
CJ Anderson	NZ	196820	6
IJL Trott	Eng	196280	7
L Ronchi	NZ	194855	8
JP Faulkner	Aus	192860	9
DA Miller	SA	188865	10

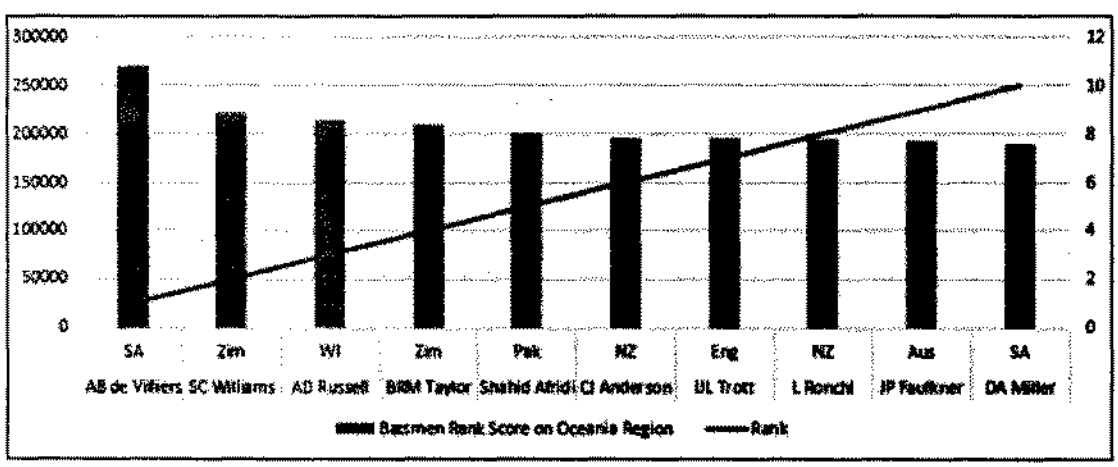


Figure 4.9: Batsmen Ranking Based on Oceania Region

4.3.10 Bowlers ranking based on Oceania Region

On Oceania region we have taken 141 top bowlers out of ten full member countries in which MA Starc (Aus) is the top performer in all players, second is LL Tsotsobe (SA), 6th is Wahab Riaz (Pak), and 10th Muhammad Shami (Ind). On this region we have shown the top ten bowlers in the list in which 3 of them are each Australian and South African bowlers, and one of each England, India, Pakistan and Zimbabwe bowlers. We have given 40%, 40%, and 20% weightages to each bowler's average, bowler's economy rate and bowlers strike rate. Wickets taken ratio and runs conceded per over are very important that is why we have given 40% to each.

Table 4.10 Bowlers Ranking Based on Oceania Region

Bowlers Name	Country	Bowlers Rank Score on Oceania Region	Rank
MA Starc	Aus	30003.33	1
LL Tsotsobe	SA	29603.33	2
CJ McKay	Aus	28160	3
RJ Harris	Aus	27481.67	4
M Morkel	SA	26025	5
Wahab Riaz	Pak	25165	6
KJ Abbott	SA	25138.33	7
CJ Anderson	Eng	23413.33	8
JE Taylor	Zim	22395	9
M Shami	Ind	20851.67	10

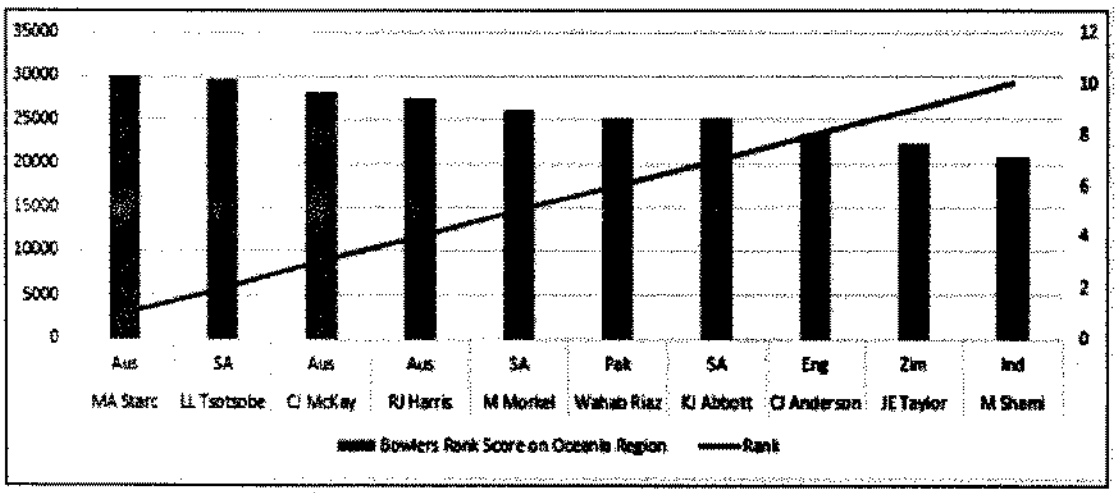


Figure 4.10: Bowlers Ranking Based on Oceania Region

4.3.11 Batsmen ranking based on America Region

On Americas region we have taken 131 top batsmen out of ten full member countries in which HM Amla (SA) is the top batsmen in all players, second is AD Russell (WI), 3rd and 4th respectively are Umar Akmal and Misbah ul Haq (Pak) and 10th is AD Mathews (Sri).

We have given weightages to each batting average and strike rate 60% and 40% respectively. In all of these batsmen even JH Kallis and Misbah Ul Haq both have very low strike rate i.e. 78.36 and 66.86 but highest batting averages i.e. 59 and 74 respectively, but still they are in the top ten list. On this region there are 3 batsmen in the list of South Africa, 2 of each Australia and Pakistan and One of each India, Sri Lanka and West Indies. Although West Indies have played a lot of matches on this region but still they have just a single batsmen in the list of top ten.

Table 4.11 Batsmen Ranking Based on America Region

Bowlers Name	Country	Bowlers Rank Score on America Region	Rank
HM Amla	SA	236381.25	1
AD Russell	WI	218850	2
Umar Akmal	Pak	204236.25	3
Misbah-ul-Haq	Pak	198767.5	4
MJ Clarke	Aus	196756.25	5
SR Watson	Aus	190962.5	6
AB de Villiers	SA	182080	7
JH Kallis	SA	181755	8
V Sehwag	Ind	180676.25	9
AD Mathews	Sri	179983.75	10

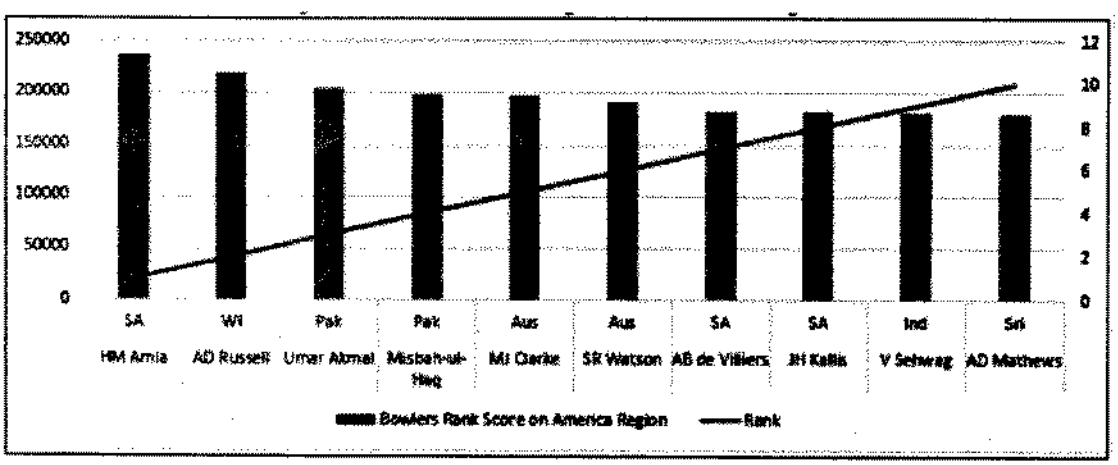


Figure 4.11: Batsmen Ranking Based on America Region

4.3.12 Bowlers ranking based on America Region

On Oceania region we have taken 145 top bowlers out of ten full member countries in which B Kumar (Ind) is the top performer in all players due to getting more wickets by conceding very few average runs, economy rate and strike rate in all 145 bowlers against strong teams, second is HMRB Herath (Sri), 3rd is Abdul Razzaq (Pak), and 10th is SL Malinga (Sri).

On this region we have shown the top ten bowlers in the list in which 3 of them are of India and Sri Lanka each, and one of each Bangladesh, New Zealand, Pakistan and South Africa.
We have given 40%, 40%, and 20% weightages to each bowler's average, bowler's economy rate and bowlers strike rate. Wickets taken ratio and runs conceded per over are very important that is why we have given 40% to each.

Table 4.12 Bowlers Ranking Based on America Region

Bowlers Name	Country	Bowlers Rank Score on Oceania Region	Rank
B Kumar	Ind	26251.67	1
HMRKB Herath	Sri	26005	2
Abdul Razzaq	Pak	25578.33	3
M Morkel (SA)	SA	25436.67	4
Al-Amin Hossain	Ban	25105	5
A Mishra	Ind	22945	6
Mohammad Sami	Ind	20560	7
AD Mathews	Sri	19636.67	8
TG Southee	NZ	18745	9
SL Malinga	Sri	17318.33	10

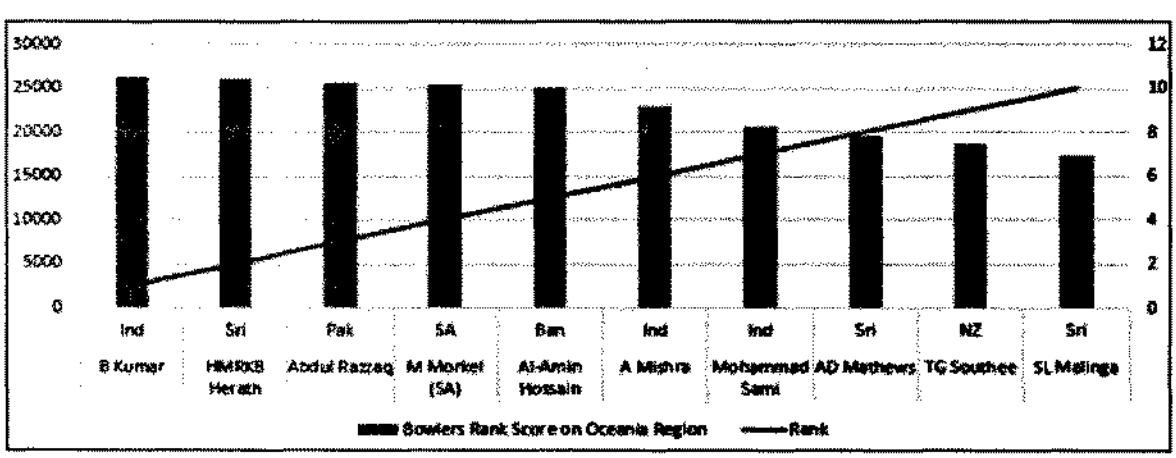


Figure 4.12: Bowlers Ranking Based on America Region

4.4 A Unified Players Ranking

In the above sections we have discussed regions based players related results and discussions and compare those results with ICC current ranking. In this section we have to calculate a Unified Players ranking and then associate with ICC current ranking.

4.4.1 A Unified Batsmen Ranking

In a Unified batsmen ranking we have calculated unified ranking of the currently playing batsmen by using the following equation.

$$Batsman_i = \frac{\sum_{R=1}^{N} R_W}{n} \text{ --------- (4.1)}$$

Where i=1….N batsmen, R is the Regions (Asia Region, Africa Region etc.), R_w is weighted score of the batsman on a specific region and n is the number of Regions (i.e. a batsman perform on three regions n=3)

In a Unified Batsmen Ranking we have computed the collective ranking of all batsmen playing on each region and compare the net rating with ICC current batsmen ranking.

Comparing our Unified batsmen ranking with ICC ranking, AB de Villier is top in both ranking methods, H. M. Amla is second in a Unified Ranking Method and 3 in ICC ranking, GJ Maxwell is 3rd in our method and 9th in ICC ranking method, Verath Kohli is 5th in our ranking method and 4th in ICC ranking, MS Dhoni is 7th in our method and 8th in ICC ranking, KC Sangakara is 8th in our ranking and 2nd in ICC ranking, KS Williamson is 9th in our Unified Ranking Method and 7th in ICC ranking and finally S Dhawan is 10th in our Unified Batsmen Ranking Method and 6th in ICC Ranking Method. Although AD Dilshan and GJ Bailey have good ranking position in a Unified Batsmen Ranking but they are not coming in top ten ranking list batsmen even they best position in ICC ranking. Since ICC ranks players on the basis of recent (2 or 3 years) performance that is why we see up and down in both the method. Ultimately a Unified Batsmen Ranking and ICC Batsmen Ranking identifies that both Ranking Methods have about 70 % similarities which clearly shows that our Unified Batsmen Ranking method generates better results.

Table 4.13 A Unified Batsmen Ranking

Batsmen Name	Country Name	A Unified Batsmen Rank Score	Rank
AB de Villiers	SA	197977.73	1
HM Amla	SA	192411.18	2
GJ Maxwell	Aus	189598.58	3
AD Russell	WI	186179.37	4
V Kohli	Ind	170782.89	5
MJ Clarke	Aus	169177.95	6
MS Dhoni	Ind	169132.06	7
KC Sangakara	Sri	166403.11	8
KS Williamson	NZ	165938.10	9
S Dhawan	Ind	164614.97	10

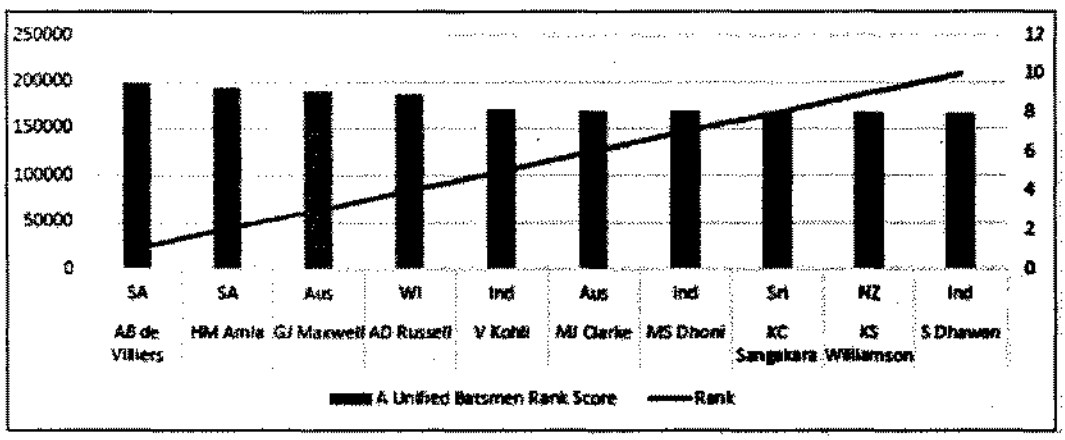


Figure 4.13: A Unified Batsmen Ranking

4.4.2 A Unified Bowlers Ranking

In a Unified bowlers ranking we have computed unified ranking of the currently playing bowlers by using the following equation.

$$Bowler_i = \frac{\sum_{R=1}^{N} R_w}{n} \quad \text{---------- (4.2)}$$

Where i=1….N bowlers, R is the Regions (Asia Region, Africa Region etc.), R_w is weighted score of the bowlers on a specific region and n is the number of Regions (i.e. a bowler perform on three regions n=3)

Comparing our Unified bowlers ranking with ICC ranking, Imran Tahir is top in our Unified Ranking Method and 2nd in ICC ranking methods, MA Starc is first in our Unified Ranking and 2nd in ICC Ranking, Saeed Ajmal is 4th in a Unified Ranking and 7th in ICC Ranking, MG Johnson is 5th in a Unified bowlers Ranking and 6th in ICC ranking. Although BAW Mendis, T Thushara, SL Malinga, JE Taylor, M Ntini and Muhammad Shami good ranking position in a Unified Batsmen Ranking but they are not coming in top ten ranking list of the bowlers even they got best position in ICC ranking. Since ICC ranks players on the basis of recent (2 or 3 years) performance that is why we see up and down in both the method. Ultimately a Unified Bowlers Ranking and ICC Bowlers Ranking identifies that both Ranking Methods have about 60 % similarities which clearly shows that our Unified Bowlers Ranking method generates better results.

Table 4.14 A Unified Bowlers Ranking

Bowlers Name	Country Name	A Unified Bowlers Rank Score	Rank
Imran Tahir	SA	32508.16	1
MA Starc	Aus	32462.89	2
BAW Mendis	Sri	32432.91	3
Saeed Ajmal	Pak	32385.27	4
MG Johnson	Aus	31808.34	5
T Thushara	Sri	31673.46	6
SL Malinga	Sri	29297.62	7
JE Taylor	WI	29082.37	8
M Ntini	SA	28972	9
M Shami	Ind	28788.47	10

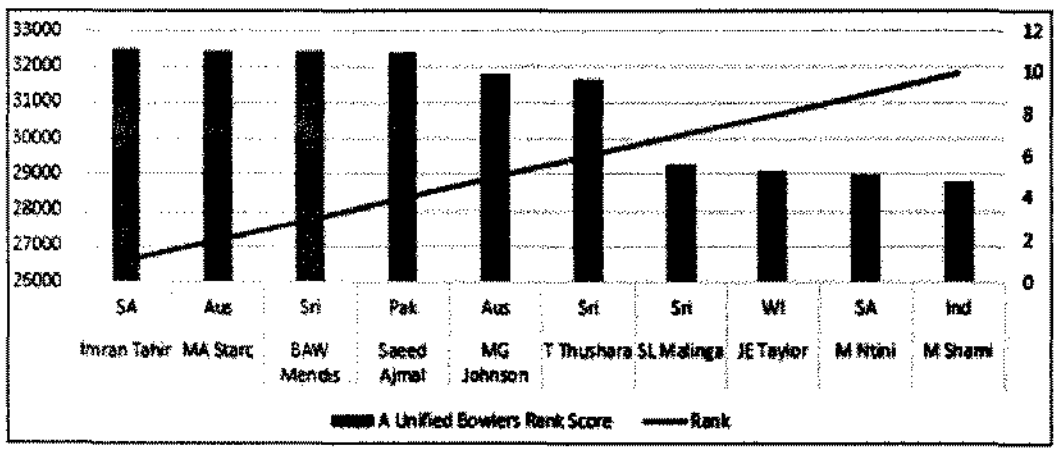


Figure 4.14: A Unified Bowlers Ranking

4.4 Chapter Summary

In this chapter we have studied experimental results of batsmen and bowlers on each separate region. We have qualitatively analyze overall record of the players on the basis of each region. The dataset from 2001 to 2015 (based on individual region) will be taken from ESPNcricinfo and archive website and based on the ranking features by using Link Fusion Algorithm as a model. We have included all those players whom have taken part in at least 10 matches in a specific region on ODI cricket matches.

CHAPTER 5

CONCLUSION
&
FUTURE WORK

5. Conclusion

To summarize our work, we studied the ranking of players on each region in One Day International (ODI) cricket. We have applied the concept of Link Fusion Algorithm on all five regions separately for ranking the players in both format i.e. batsmen and bowlers in cricket. Although Link Fusion Algorithm is used in many scenarios like by applying it into a real world scenario of three data spaces: user, query and webpage data spaces but it has not been used in cricket ranking.

We used here three main fields of cricket sports which are 1) team's batting and bowling strengths, 2) batsmen and 3) bowlers. For team's batting strength we have calculated four basic features Team's average runs scored on each of the five regions, team's win/loss ratio, team's player's average runs, team's strike rate. Similarly for team's bowling strength we have calculated team's average runs scored, team's win/loss ratio, team's bowler's average runs conceded and team's economy rate conceded. For batsmen we have investigated batsmen average runs and batsmen strike rate against each bowlers and teams respectively on each of the five regions and for bowlers we have considered bowler's average runs conceded, bowler's strike rate and bowler's economy rate conceded against each batsmen and teams respectively on each of the five region.

We explored that the team's batting and bowling strength is very important for finding the ranking of batsmen and bowlers respectively. The teams which have high batting strength are stronger in batting line-up as compare to those which have weak team's strength also the teams which have stronger team's bowling strength are stronger in his bowling line-up.

After computing region based ranking, we have calculated a unified players ranking which clearly signifies that our proposed ranking method generates better results.

This dissertation is the first attempt to find out ranking cricket players using Link Fusion Algorithm concept in ODI cricket format on each regions separately in order to see the strength of each team and players on each region. This method can be apply to rank region based test cricket and T-20 cricket players.

To conclude our work we have compared our results with ICC ad hoc ranking and with performances of the players based on each region. Almost 70 % of the ICC top ten players comes in our region based top ten ranking space. In addition there is no work done on region based cricket players ranking using Link Fusion Algorithm.

5.1 Future Work

In cricket ranking ICC ranks players based on their current performance after each series but there is no work done based on region based ranking in cricket. We have investigated the record of all the teams' vs teams, batsmen vs teams, bowlers versus teams, batsmen vs bowlers and bowlers versus batsmen based on each separate region. We have applied link fusion algorithm concept and find out ranking of batsmen and bowlers on each region and on the basis of region based ranking a unified players ranking is calculated. In the future, Link Fusion Algorithm can be applied to rank players based on each region in T-20 cricket and in test cricket format.

6. REFERENCES

[1] Anjela, Y. Govan. and Carl D. (2008). Generalizing Google's PageRank to Rank National Football League Teams. Biofinformatics, 5(4), 1-151.

[2] Boulier, B. & Stekler, H. (2003). Predicting the outcomes of National Football League games. International Journal of Forecasting, 19, 257–270

[3] Bhattacharjee, D., Pahinka, D., 2012. Analysis of Performance of Bowlers using Combined Bowling Rat. International Journal of Sports Science and Engineering, 6(3), pp 184-192.

[4] Brown, L. D. (2008). In-season Prediction of Batting Averages – A Field Test of Empirical Bayes and Bayes Methodologies. The Annals of Applied Statistics, 2, 113152

[5] Corral, J. and Prieto-Rodriguez, J. (2010). Are differences in ranks good predictors for Grand Slam Tennis matches?. International Journal of Forecasting, 2(6), 551 -563

[6] Mukherjee, S., 2014. Quantifying individual performance in Cricket—A network analysis of batsmen and bowlers. Phys. Stat. Mech. Its Appl. 393, 624–637.

[7] Croucher, J. S. (2000). Player ratings in one-day cricket. In Proceedings of the Fifth Australian Conference on Mathematics and Computers in Sport, 95-106

[8] Daud, A., Muhammad, F., 2013. Ranking Cricket Teams through Runs and Wickets, in: Active Media Technology. Springer, pp. 156–165.

[9] Dingle, N., Knottenbelt, W., & Spanias, D. (2013). On the (page) ranking of professional tennis players. In Computer Performance Engineering: Springer Berlin Heidelberg, pp. 237-247

[10] Duch, J. S. Waitzman. and Amaral, L. (2010). Quantifying the performance of individual players in a team activity. PloS One, 5(6), e10937

[11] Hoerl, A. E. and Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. Technometrics, 6(2), 55-67

[12] https://en.wikipedia.org/wiki/Cricket

[13] (http://www.dummies.com/how-to/content/cricket-for-dummies-cheat-sheet.html).

[14] http://www.icc-cricket.com/about/67/icc-members/regions.

[15] http://en.wikipedia.org/wiki/FIH _World _Rankings

[16] http://www.icc-cricket.com/player-rankings/about

[17] Kimber, A. C. and A. R. Hansford. (1993). A statistical analysis of batting in cricket. Journal of the Royal Statistical Society. Journal of the American Society for Information Science and Technology, 156(3), 443-445

[18] Macdonald, B. (2011). A Regression-Based Adjusted Plus-Minus Statistic for NHL Players. Journal of Quantitativ e Analysis in Sports, 7(3), 29

[19] Mukherjee, S., 2013. Complex Network Analysis in Cricket: Community structure, player's role and performance index. Advances in Complex Systems, 16(08), 1350031.

[20] Mukherjee, S., 2012. Identifying the greatest team and captain— A complex network approach to cricket matches. Phys. Stat. Mech. Its Appl. 391, 6066–6076.

21] Radicchi, F. (2011). Who Is the Best Player Ever? A Complex Network Analysis of the History of Professional Tennis. PLoS One, 6(1), e17249

22] Thomas, A. C., Ventura, S. L., Jensen, S. T., & Ma, S. (2013). Competing process hazard function models for player ratings in ice hockey. The Annals of Applied Statistics, 7(3), 1497-1524.

[23] Wikipedia.http://en.wikipedia.org/wiki/List_of_International_Cricket_Council_members.

[24] Xi, W., Zhang, B., Chen, Z., Lu, Y., Yan, S., Ma, W. Y., & Fox, E. A., 2004. Link Fusion: A Unified Link Analysis Framework for Multi-Type Interrelated Data Objects. In Proceedings of the Thirteenth International World Wide Web Conference, pp. 319- 32

5] http://www.espncricinfo.com/