Region-wise Ranking of Sports Players based on Link Fusion

Ali Daud [a,b]

[a] Faculty of Computing and Information
Technology,
King Abdulaziz University,
Jeddah, Saudi Arabia
adfmohamad@kau.edu.sa

Akbar Hussain [b]

[b] Department of Computer Science
and Software Engineering, International
Islamic University, Islamabad, Pakistan
akbar_hussain555@yahoo.com

Rabeeh Ayaz Abbasi [a,c]

[c] Department of Computer Science,
Quaid-i-Azam University,
Islamabad, Pakistan

rabbasi@kau.edu.sa

Naif Radi Aljohani [a]

[a] Faculty of Computing and Information
Technology, King Abdulaziz University,
Jeddah, Saudi Arabia
nraljohani@kau.edu.sa

Tehmina Amjad [b]

[b] Department of Computer Science
and Software Engineering, International
Islamic University, Islamabad, Pakistan
tehminaamjad@iiu.edu.pk

Hassan Dawood [d]

[d] Department of Software
Engineering, UET Texila,
Pakistan
hassan.dawood@uettaxila.edu.pk

ABSTRACT

Players are ranked in various sports to show their importance over other players. Existing methods only consider intra-type links (e.g., player to player and team to team), but ignore inter-type links (e.g., one type of player to another type of player, such as batsman to bowler and player to team) based on cognitive aspects. They also ignore the spatiality of the players. There is a strong relationship among players and their teams, which can be represented as a network consisting of multi-type interrelated objects. In this paper, we propose a players' ranking method, called Region-wise Players Link Fusion (RPLF) which is applied to the sport of cricket. RPLF considers players' region-wise intra-type and inter-type relation-based features to rank the players. Considering multi-type interrelated objects is based on the intuition that a batsman scoring high against top bowlers of a strong team or a bowler taking wickets against top batsmen of a strong team is considered as a good player. The experimental results show that RPLF provides promising insights of players' rankings. RLFP is a generic method and can be applied to different sports for ranking players.

1 INTRODUCTION

Teams consist of players in sports and better teams have better players. Players ranking in sports quantifies the importance of a player in a team. A higher rank is achieved by players through performing well and consistent for their teams. The higher the ranking, the more chances that the player will be selected in the playing side. Players are ranked on the basis of their performances in the sports, such as cricket, football, tennis, hockey, baseball, etc.

Most of the researcher considers both players past and recent performances. Previously, cricket sports objects' (batsman, bowler, team) ranking is performed using different features, such as, batting position, batting experience, home country advantage, region advantage, match starting time, inning sequence, opposition team strength, form of the players, batting average and strike rate for a batsman, bowling average, strike rate and economy rate for a bowler, session-wise player performance, Win / Loss ratio of a team, runs and wickets from which the matches are won by a team, or by applying PageRank [1, 3, 4, 12, 17]. PageRank based algorithm is also applied in football [2, 7] and baseball [16] to rank players by exploiting intra-type links. Tennis players' ranking is performed by applying PageRank for rating points based on the position of player and tournament value, time-dependency [6, 13] and by exploring intra-type links [14]. Intra-type links are utilized mostly in the past for different sports, while inter-type and region-wise information are not considered simultaneously. Cognitively, the success rate of team increases with "home advantage phenomenon" is widely expected and documented for different sports [15]. Intuitively, it has a great impact on players' performances as well.

In sports, multi-type interrelated objects exist, such as, the relationship between players and their teams, that can be presented in the form of a network. So, there is a homogeneous or heterogeneous relationship among the players of teams in order to calculate a specific feature. Cricket sports resembles baseball in its playing style where the pitcher is called bowler, the striker is called batsman, catcher is called wicket keeper, etc. Players ranking the game of cricket is mostly done based on their recent performances. Due to different batting and bowling conditions, cognitively region also plays an important role in the ranking of cricket players. Region-based ranking is important and challenging for selecting the best team for each region. In this work, a unified player ranking method RLFP is proposed to overcome the shortcomings of existing methods. RPLF considers both intra-type and inter-type links by exploiting spatiality of players along with other features. The detailed experimental study shows the effectiveness of the proposed algorithm. RLFP is quite general and requires no specific domain knowledge sit o can be applied in various sports.

The significant contributions of this work are as follows.
(1) Consideration of both intra-type and inter-type links
(2) Consideration of spatiality

(3) Region-wise Players Link Fusion algorithm applicable to different sports

The rest of the paper is organized as follows. Section 2 provides literature review for the ranking method in different sports. The ICC players current ranking criteria and problem definition are provided in section 3. The problem statement is discussed in Section 4. Section 5 elaborates our proposed RPLF method for cricket sports. Experimental results and discussion are provided in section 6. Section 7 finally concludes this work.

2   RELATED WORK

Ranking of players in sports is performed extensively in the literature. Different methods are proposed for different sports based on mathematical operations on features. Several methods explore network features by applying PageRank [20]. Salient efforts made in the sports of cricket, football, tennis, and baseball are discussed in the remaining section.

Cricket is one of the widely played and watched sport in the Asian region. Features such as batting position, batting experience, home country advantage, region advantage, match starting time, inning sequence, opposition team strength, and form of the players are used to rank players [3]. Batting average and batting strike rate for batsmen and bowling average, bowling strike rate and economy rate for bowlers have also been explored to rank players [17]. Captains of teams are ranked based on the winning performances of their teams. A directed and weighted network of teams and their captains is formed, and success of the team is measured by the quality of wins and not only by the number of wins using PageRank [13]. Players' performances evaluation is based on batting, bowling and fielding contributions in the match, session by session. Nominal multinomial logistic regression models are used to model match outcome probabilities for hypothetical and actual positions. Player performance is then calculated using changes in such probabilities session by session. The advantage of looking at the contribution at the end of each session of the match is evaluated in which wickets and runs are taken and conceded respectively [1]. Recently, Daud et al. presents a weighted page rank method to rank the cricket teams. It argues that only considering the number of Wins / Losses is not enough. There is also need to consider the strength of a team and also the number of runs and wickets from which the team has won/lost a match [4]. A bi-directional weighted network of players is generated for T-20 cricket matches by applying social network concepts for measuring performance and ranking players. All intra-country networks of players are investigated by treating players as nodes and the interaction between team members is denoted by edges. The selected team based on proposed ranking is compared with their Indian Premier League performances of 2016 [5]. Precedence of a team over the other teams in the game of cricket for one-day matches is explored recently. The batting and bowling productivity at the team level is emphasized in addition to previously used features [21]. A relatively new task of rising stars prediction for the game of cricket for both the batsmen and the bowlers is investigated. Features are derived from the co-players, team and opposite teams' entities. Generative and discriminative models are explored for prediction on real-world dataset [22].

Football is unanimously the most played and watched sports all over the world. Anjelya et al. apply PageRank based on "Colley Matrix Method" for ranking national football league teams. It first creates NFL graph and constructs a matrix for NFL graph and then computes the vector that contains the ranks of teams based on their seasonal performance. NFL graph is created in a way that each node in the graph represents some team of the league. When two teams play each other, a weighed directed edge is created among them and the direction is from the loser to the winner, assuming that there are no tie matches [2]. Duchet al. has highlighted the importance of teamwork leading towards greater creativity in European Cup soccer tournament. A network approach is developed which dominantly quantifies contributions of individual soccer players and overall team performance [7].

Tennis is one of the widely watched game all over the world. Intra-type links are applied for measuring the performance of each tennis player individually. It also hypothesizes that the generalized method can be beneficial in other sports environment in which the quantification of the opponent team members' contribution is important [15]. Another approach is used [6] to improve the ranking method of [14] by using PageRank. In its proposed scheme the data from the traditional ranking system is used where the rating points are given according to the position of the player and the value of the tournament. For example, Grand Slam tournament has the highest 2000 score for the winner. The PageRank is also used in [14] for ranking for tennis and table tennis players. A time depended approach is used with PageRank and named as Time-Dependent PageRank (tdPR).

Baseball is very much similar to cricket and is one of the most played and watched sports in North America. Shan et al. present GameRank method, which is a variation of HITS and PageRank to evaluate the pitching and batting ability for players in Major League Baseball (MLB), with a network perspective. It evaluates its proposed algorithm by comparing its results to ESPN Ratings, a popular baseball rating method [16].

Initially, PageRank [20] was used for ranking web pages and later applied to rank different objects. It computes a single vector for the whole graph to rank web pages, which produces biased results. In order to get more realistic results, topic-sensitive PageRank was proposed [8] to compute several PageRank vectors based on the topic relevance of the pages. Based on this topic-sensitive concept that the ranked results for each query should have their own vectors intuitively players should be ranked region-wise as well. Existing PageRank based methods for ranking sports objects also ignore inter-type links, which are considered important while ranking multi-type related web objects [18].

3   ICC Players Current Ranking

International Cricket Council (ICC) ranks players based on a pre-defined method. Exact formulae used in the method are not published, but factors used are provided on their website [10]. Missing a One Day International (ODI) match reduces 0.5% of their points for their country. All ODI matches are considered equal, except for ICC Cricket World Cup matches, where good performances add extra credit to the team and the players. More number of runs/wickets against weaker team contributes less credit than same performances against stronger teams [10]. Factors that are used for ranking of batsmen and bowlers are discussed below.

3.1 Batsmen Ranking Factors

ICC assigns points to each batsman based on the following factors [10]:

- Number of runs scored by a batsman
- Rating of the opposing bowling line-up
- A 100 runs scored by a player in a match where both teams scored 500 has less value than that of 100 runs scored in a match where both teams were bowled out under 200.

- A batsman receives bonus points when his team wins. That bonus will be higher for highly rated opposition teams.
- Batsmen gain significant credit for high strike rate.
- Batsmen get a small amount of credit for being not out.

3.2 Bowlers Ranking Factors

ICC assigns points to each bowler based on the following factors [10]:

- Wickets taken and runs conceded by a bowler (bowling average) have a significant role.
- The level of runs scored in a match; e.g. bowling figures of 3 for 50 in a high-scoring match increases a bowler's rating more than the same figures in a low-scoring match.
- Bowlers who bowl a large number of overs in a match get credit, even if they take no wickets.
- Bowlers who get the most number of wickets in a victory receive bonus points.
- Bowlers gain significant credit for bowling at a low economy rate (bowling average).

4 PROBLEM DEFINITION

We define the problem of region-wise ranking of cricket players as an information retrieval problem. Given $P = \{P_1^1, P_2^1, P_3^2, ..., P_i^r\}$ be an $n \times r$ matrix representing n dimensional feature vectors of n players and r 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.

5 REGION-WISE PLAYERS LINK FUSION (RPLF)

Region plays a key role in cricket, as different regions have different playing conditions and have an impact on players' performances. Players with best performances on specific regions should be selected for that specific region. In this article, the players are ranked based on separate regions. There are five regions categorized in the world in which international cricket is played [19]. These are: **Europe** (England), **America** (West Indies), **Asia** (Pakistan, India, Sri Lanka and Bangladesh), **Oceania** (Australia and New Zealand), and **Africa** (South Africa and Zimbabwe) [9,10].

In cricket ranking, batsmen always bat against bowlers of the opposite team, similarly, bowlers bowl against batsmen of the opposite team. Hence, there is a strong relationship among batsmen, bowlers, and teams. Motivated from [18] to investigate this relationship, we consider three data spaces which are batsmen, bowlers, and teams. Each data space has many data objects. When the data object of one data space is related to another data space's object, we say that the relationship is heterogeneous (inter-type relationship). Here, data objects mean batsmen, bowlers, and teams. Figure 1 shows the relationship of batsmen, bowlers, and teams, respectively.

Suppose, there are n different data objects within all data spaces A= (b_1, b_2... b_n, l_1, l_2... l_m, T_1, T_2 ... T_p), where 'b_x' represents a batsman, 'l_x' represents a bowler and 'T_x' represents a team. Each data object A_i contains a specific attribute F_i. Data objects within the same data space are interrelated with intra-type associations $R_i \subseteq A_i \times A_i$, while data objects from different data spaces are inter-

related with inter-type associations $R_{ij} \subseteq A_i \times A_j$ ($i \neq j$). The specific attribute of a data object in one data type is equal to the sum of the attributes of other data objects in the same data space that link to it, plus the sum of the other related attributes of data objects in other data spaces and links to it.
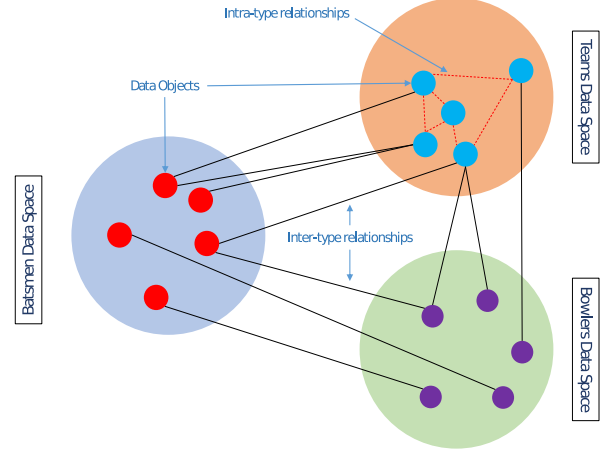

Figure 1: An illustration of weightage calculation assumption.

Mathematically, we can define an attribute as:
$$F_i = F_i R_i + \sum_{j=1}^{N} F_j R_{ji} \qquad (1)$$
Where, i is a data object whose attributes are to be calculated, $j=1$ to N are the data objects from which these attributes are to be calculated and R is the region on which they performed.

Equation (2) shows the region-wise performance of a batsman against both the bowlers and teams. Similarly, equation (3) shows the region-wise performance of a bowler against both the batsmen and teams.
$$RW_{bat} = L^T_{Bat\ Bol} W_{Bat\ Bol} + L^T_{Team\ Bat} W_{Team\ Bat} \qquad (2)$$
$$RW_{Bol} = L^T_{Bol\ Bat} W_{Bol\ Bat} + L^T_{Team\ Bol} W_{Team\ Bol} \qquad (3)$$
Where, RW_{Bat} and RW_{Bol} are the weights of batsmen and bowlers, respectively. $L^T_{Bat\ Bol}$ and $L^T_{Bol\ Bat}$ are the adjacency matrices from batsmen to bowlers and bowlers to batsman, respectively. If a batsman played against a bowler $L^T_{Bat\ Bol} = 1$ in the adjacency matrix from batsmen to bowler's space else $L^T_{Bat\ Bol} = 0$, similarly, in bowler to batsmen data space, batsman to teams and bowler to team's data spaces. $W_{Bat\ Team}$ and $W_{Bol\ Team}$ are the weights from batsmen to teams and bowlers to teams, 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 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. For batsmen's data space equation (2) can be written in matrix form as follows:

$$\begin{bmatrix} L^T_{Bat\ Bol11} & L^T_{Bat\ Bol12} & \cdots & L^T_{Bat\ Bol1n} \\ L^T_{Bat\ Bol21} & L^T_{Bat\ Bol22} & \cdots & L^T_{Bat\ Bol2n} \\ L^T_{Bat\ Boln1} & L^T_{Bat\ Boln2} & \cdots & L^T_{Bat\ Bolnn} \end{bmatrix} \begin{bmatrix} W_{Bat\ Bol11} & W_{Bat\ Bol12} & \cdots & W_{Bat\ Bol1n} \\ W_{Bat\ Bol21} & W_{Bat\ Bol22} & \cdots & W_{Bat\ Bol2n} \\ W_{Bat\ Boln1} & W_{Bat\ Boln2} & \cdots & W_{Bat\ Bolnn} \end{bmatrix}$$
$$+$$
$$\begin{bmatrix} L^T_{Team\ Bat11} & L^T_{Team\ Bat12} & \cdots & L^T_{Team\ Bat1n} \\ L^T_{Team\ Bat21} & L^T_{Team\ Bat22} & \cdots & L^T_{Team\ Bat2n} \\ L^T_{Team\ Batn1} & L^T_{Team\ Batn2} & \cdots & L^T_{Team\ Batnn} \end{bmatrix} \begin{bmatrix} W_{Team\ Bat11} & W_{Team\ Bat12} & \cdots & W_{Team\ Bat1n} \\ W_{Team\ Bat21} & W_{Team\ Bat22} & \cdots & W_{Team\ Bat2n} \\ W_{Team\ Batn1} & W_{Team\ Batn2} & \cdots & W_{Team\ Batnn} \end{bmatrix}$$

In the above matrices, batsmen perform against bowlers and teams, which shows their region-wise performance.

Similarly, for bowlers' data space equation (3), can be written in matrix form as follows:

$$\begin{bmatrix} L^T_{Bol\ Bat11} & L^T_{Bol\ Bat12} & \cdots & L^T_{Bol\ Bat1n} \\ L^T_{Bol\ Bat21} & L^T_{Bol\ Bat22} & \cdots & L^T_{Bol\ Bat2n} \\ L^T_{Bol\ Batn1} & L^T_{Bol\ Batn2} & \cdots & L^T_{Bol\ Batnn} \end{bmatrix} \begin{bmatrix} W_{Bol\ Bat11} & W_{Bol\ Bat12} & \cdots & W_{Bol\ Bat1n} \\ W_{Bol\ Bat21} & W_{Bol\ Bat22} & \cdots & W_{Bol\ Bat2n} \\ W_{Bol\ Batn1} & W_{Bol\ Batn2} & \cdots & W_{Bol\ Batnn} \end{bmatrix}$$
$$+$$
$$\begin{bmatrix} L^T_{Team\ Bol11} & L^T_{Team\ Bol12} & \cdots & L^T_{Team\ Bol1n} \\ L^T_{Team\ Bol21} & L^T_{Team\ Bol22} & \cdots & L^T_{Team\ Bol2n} \\ L^T_{Team\ Boln1} & L^T_{Team\ Boln2} & \cdots & L^T_{Team\ Bolnn} \end{bmatrix} \begin{bmatrix} W_{Team\ Bol11} & W_{Team\ Bol12} & \cdots & W_{Team\ Bol1n} \\ W_{Team\ Bol21} & W_{Team\ Bol22} & \cdots & W_{Team\ Bol2n} \\ W_{Team\ Boln1} & W_{Team\ Boln2} & \cdots & W_{Team\ Bolnn} \end{bmatrix}$$

5.1 Region-based Team's batting Strength

Similarly, in the above matrices bowlers perform against batsmen and teams, which shows their region-wise performance. We formulate various features related to a team for each region. These features are discussed below.

5.1.1 Region-based Team's Win/Loss Ratio

We can define win/loss ratio ($RTWLR^r_t$) of a team t in a region r by dividing total number matches won by the number of matches lost on a region.

$$RTWLR^r_t = \frac{Won^r_t}{Lost^r_t} \qquad (4)$$

Where, Won^r_t represents the number of matches won by team t in region r and $Lost^r_t$ represents the number of matches lost by a team t at region r. For example, if Pakistan has won 130 matches and has lost 87 matches in Asia region, then $RTWLR^{Asia}_{Pak} = \frac{130}{87} = 1.49$.

5.1.2 Region-based Team's average Runs

Region-based team's average runs ($RTAR^r_t$) of a team t for a region r is defined as the total number of runs (R^r_t) scored by the team t in the region r divided by number of innings played (I^r_t) by the team t in the region r.

$$RTAR^r_t = \frac{R^r_t}{I^r_t} \qquad (5)$$

For example, if Pakistan scores 47964 runs in 215 matches in Asia, then $RAR^{Asia}_{Pak} = \frac{47964}{215} = 223.0883$.

5.1.3 Region-based Team's Strike Rate

Region-based team's strike rate ($RTSR^r_t$) can be defined as total number of runs scored by a team (R^r_t) divided by the number of balls played (L^r_t). As the number of balls played in a match are much higher than the number of innings played (see Eq. 5), therefore it is multiplied with a factor k. Value of k is set to 6 as stated by [17].

$$RTSR^r_t = \frac{R^r_t}{L^r_t} \times k \qquad (6)$$

5.1.4 Region-based Team's Batting Average

Region-based team's batting average ($RTBA^r_t$) is defined as the average of batting averages of all the batsmen of a team playing in a region.

$$RTBA^r_t = \frac{\sum BA^r_t}{n_b} \qquad (7)$$

Where, BA^r_t is the batting average of team t in region r and n_b is the number of batsmen who played for the team t in region r.

Based on equations (4) to (7), a team's Region-based Batting Strength (RBS^r_t) can be computed as follows:

$$RTBS^r_t = RTAR^r_t \times \alpha + RTSR^r_t \times \beta + RTBA^r_t \times \gamma + RTWLR^r_t \times \delta \qquad (8)$$

Where, $\alpha + \beta + \gamma + \delta = 1$. In the absence of gold standard benchmark rankings, optimum combination of these variables is difficult to find. We use the values of $\alpha = \beta = \gamma = 0.2\ and\ \delta = 0.4$. δ is given a higher weight, because win-loss ratio of a team in a region is key factor [4].

Before applying eq. (8) all the feature values are normalized by the maximum value of the feature to get the teams' net batting strength.

5.2 Region-based Team's Bowling Strength

A team's region-based bowling strength can be calculated in a similar way to the batting strength. In bowling, *Average Runs Conceded (RTARC), Economy Rate (RTER), Bowling Strike Rate (RTLSR)* along with *Win/Loss Ratio (RTWLR)* plays a key role [10]. These values can be computed for regions (*RTARC, RTER, RTLSR, RTWLR*) in a similar way to equations (4) to (7). Once, individual values are computed, aggregated Region-based Team's Bowling Strength (RLS^r_t) can be calculated as follows:

$$RTLS^r_t = RTARC^r_t \times \alpha + RTER^r_t \times \beta + RTLSR^r_t \times \gamma + RTWLR^r_t \times \delta \qquad (9)$$

Values for $\alpha, \beta, \gamma,$ and δ are the same as in equation (8). Values are normalized by the maximum value. Values for bowling performance are also inverted, because a lower score in bowling means better performance.

5.3 Region-based Batsman Strength

Batsman region-based strength can be calculated using *Average of a Batsman (RBA)* against a specific bowler and *Strike Rate of a Batsman (RBSR)*, as both are key features of a batsman's strength.

Average of a batsman is defined as the number of runs scored by a batsman in all innings divided by the number of completed innings (innings in which the player is "not out" are excluded) on a region. Strike rate of a batsman is defined as the number of runs scored divided by the number of balls faced by that batsman in a region. Weightage of a batsman b against each bowler l on a region r is calculated as follows.

$$RBS^r_b = RBA^r_b \times \alpha + RBSR^r_b \times \beta \qquad (10)$$

Values for α and β are 0.6 and 0.4 respectively, because a team mostly needs high batting average of batsman as compared to strike rate of a batsman, although one can try different percentages.

Before applying eq. (10) all the feature values are normalized by the maximum value of the feature to get the batsman's net batting strength.

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

Result of RW_{bat} is calcualted by finding weighted average of both the $W_{Bat\ Bol}$ and $W_{Team\ Bat}$ by using following equation.

$$RW_{bat} = W_{Bat\ Bol} \times \alpha + W_{Team\ Bat} \times \beta \qquad (11)$$

Value for both α, β is 0.5.

5.4 Region-based Bowler Strength

Bowler region-based strength can be calculated using *Average of a Bowler (RLA), Strike Rate of a Bowler (RLSR)* and *Economy Rate of a Bowler (RLER)*, as all three are key features of a bowler strength.

Average of a bowler is defined as the total number of runs conceded by a bowler divide by the number of wickets taken on a region. Strike rate of a bowler is defined for a bowler as the average number of balls bowled per wicket taken on a region. Economy rate of a bowler is defined as the number of runs conceded per over on a region. Weightage of a bowler l against each batsman b on a region r is calculated as follows.

$$RLS^r_l = RLA^r_l \times \alpha + RLSR^r_b \times \beta + RLER^r_l \times \gamma \qquad (12)$$

Values for α, β and γ are 0.4, 0.4, 0.2, respectively. Because average and strike rate are considered relatively important than economy rate. Other weights can be used though.

Before applying eq. (12) all the feature values are normalized by the maximum value of the feature to get the bowlers' net bowling strength.

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

Result of RW_{bol} is calcualted by finding weighted average of both the $W_{Bol\ Bat}$ and $W_{Team\ Bol}$ by using following equation.

$$RW_{bol} = W_{Bol\ Bat} \times \alpha + W_{Team\ Bol} \times \beta \qquad (13)$$

Value for both α, β is set to 0.5.

6   EXPERIMENTS

This section explains dataset, performance evaluation details and 2015 ICC cricket rankings. It also shows the detailed comparison of the players' (batsmen and bowlers) 2015 ICC rankings with the proposed unified players and region-wise rankings.

6.1 Dataset and Performance Evaluation

The dataset is taken from espncricinfo.com from January 1, 2001, to April 30, 2015 [10]. ICC puts cricket teams in three categories: (1) Full members, (2) Associate members and (3) Affiliate members. We have considered only full member ten countries, which are Australia, Bangladesh, England, India, New Zealand, Pakistan, South Africa, Sri Lanka, West Indies, and Zimbabwe. The region-wise dataset from January 1, 2001 to April 30, 2015 is taken for ODI matches only. Only those players are included which have played at least 10 matches in each region.

No gold standard dataset of cricket sports players exists to measure the quantitative effectiveness of any proposed method in terms of precision, recall or F-measure. Other ways like measuring Spearman's Rank Correlation and Kendall's Rank Correlation are not usable as we are providing a region-wise ranking of players, which is not provided by ICC or by any existing ranking work. ICC only provides general rankings of players. However, ODI ranking points system from ICC is used as a baseline method for comparing our unified players ranking method. We have discussed effectiveness of our proposed method with our university cricket team players for expert opinion and among sixteen most of them found it very useful, functional and time driven.

6.2 ICC Rankings of Players

ICC rates players on a scale of 0 to 1000 points. If a player's performance is improving based on his past record, his points increase; if his performance is declining, his points decrease. Tables 1 and 2 show ICC top ten rankings for batsmen and bowlers respectively.

6.3 Associating Current ICC Players Ranking with Unified Players Ranking

In this section, we discuss the proposed Unified Players' ranking and compare it with ICC ranking. Unified ranking aggregates the region-based rankings and shows performances of players across the regions.

6.3.1 Unified Batsmen Ranking

Unified batsmen ranking ($BatsmanUR_i$) of a batsman i can be computed as follows:

$$BatsmanUR_i = \frac{\sum_{r=1}^{n} BRW_{i,r}}{n} \qquad (14)$$

Where, $BRW_{i,r}$ is the weighted score of the batsman on region r. n is the number of regions on which the batsman has played.

Rankings based on the proposed unified batsmen ranking are shown in table 3 and ICC rankings are shown in table 1.

Table 1. ICC Batsmen Ranking

ICC Batsmen Ranking			
Rank	Name	Country	Rating
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

Table 2. ICC Bowlers Ranking in Cricket

ICC Batsmen Ranking			
Rank	Name	Country	Rating
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

Batsmen common in top-10 ranks of both the rankings are highlighted in table 1. AB de Villier is top in both rankings, H. M. Amla is 2nd in unified ranking and 3rd in ICC ranking, GJ Maxwell is 3rd in unified ranking and 9th in ICC ranking method, Verath Kohli is 5th in unified ranking and 4th in ICC ranking, MS Dhoni is 7th in unified ranking and 8th in ICC ranking, KC Sangakara is 8th in unified ranking and 2nd in ICC ranking, KS Williamson is 9th in unified ranking and 7th in ICC ranking and finally S Dhawan is 10th in unified ranking and 6th in ICC Ranking Method. Although AD Dilshan and GJ Bailey have high scores based on unified ranking but they are not being listed as top ten batsmen. Since ICC ranks players based on recent (2 or 3 years) performances that is why we see, ups and downs in both the methods. Eventually, Unified Batsmen Ranking and ICC Batsmen Ranking identify that both Ranking Methods have about 80% similarities, which clearly shows that the proposed Unified Batsmen Ranking method generates intuitive results.

Table 3. 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

6.3.2 Unified Bowlers Ranking

Unified bowler ranking ($BowlerUR_i$) of a bowler i can be computed as follows:

$$BowlerUR_i = \frac{\sum_{r=1}^{n} LRW_{i,r}}{n} \qquad (15)$$

Where, $LRW_{i,r}$ is the weighted score of the bowler on region r. n is the number of regions on which the bowler has played.

Table 4. 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
MShami	Ind	28788.47	10

Ranks of bowlers based on unified rankings are shown in table 4 and ICC rankings in table 2. Comparing the results, we observe that, Imran Tahir is top in unified ranking and 2nd in ICC ranking methods, MA Starc is 2nd in unified ranking and 1st in ICC Ranking, Saeed Ajmal is 4th in unified ranking and 7th in ICC Ranking, MG Johnson is 5th in unified ranking and 6th in ICC ranking. Although BAW Mendis, T Thushara, SL Malinga, JE Taylor, M Ntini and Muhammad Shami have high scores based on unified ranking but they are not listed as top ten ranking list of the bowlers. Since ICC ranks players based on recent (2 or 3 years) performances so, we see ups and downs in both the methods.

6.4 Comparison of Current ICC Players Ranking with Region-based Ranking

In this section, we will compare region-based players' rankings with ICC ranking. By comparing batsmen ranking in both methods, A.B. de Villiers is top in the list of ICC batsmen ranking and he also appears among top ten in most of the regions. AB de 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 an advantageous position in region-based top ten ranking. K. C. Sangakara and K.S. Williamson on the other hand, have the best position in current ICC batsmen ranking, although they both have an advantageous position in region-based batsmen ranking but they are not included in region-based top ten ranking list because they perform almost equally in all the regions.

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

6.4.1 Batsmen ranking based on Asia Region

From region Asia, 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). Table 5 shows the top 10

batsmen for region Asia. In this region, there are 4 Indian, 3 Australian, 2 South African and 1 Pakistani batsmen.

Table 5. 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

One can try different weightages to batting average and strike rate (70%(average), 30% (strike rate) or (50%(average), 50% (strike rate)) [20], but we have given weightages to each batting average and strike rate 60% and 40%, respectively, because batting average is important than strike rate for a batsman and team. In these batsmen, even Shahid Afridi has very low batting average i.e. 25.25 but highest strike rate i.e. 128.81, in addition, he performs well against strong bowlers and teams. Due to good strike rate, Shahid Afridi is included in Asia region but in the rest of the regions, he has very low strike rate and average runs, so he is on 20th to 40th positions.

6.4.2 Bowlers ranking based on Asia Region

For region Asia, 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), as shown in table 6.

Table 6. Bowlers Ranking Based on Asia Region

Players Name	Country	Players Rank Score in 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

In all of the bowlers Imran Tahir has conceded 15.41, 4.27, and 21.6 average runs, economy rate, and strike rate respectively, in addition, he performs 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 at his strike rate that is why he is on 7th position on the list. KD Mills and SCJ Broad have both highest strike and economy rates, they are 9th and 10th in the ranking list.

6.4.3 Batsmen ranking based on Africa Region

For region Africa, we have taken 172 top batsmen out of ten full member countries in which AD Russell (WI) is the top batsman in all players, second is H. M. Amla (SA), 3rd is AB de Villier (SA), 9th and 10th are L. D. Chandimal (Sri) and T. M. Dilshan (Sri) respectively as shown in table 7. Although KP Pieteresen has second best batting average, i.e., 87.14 on Africa region but most of his runs in this region are against weak bowlers and teams, that is why he is 7th in the list.

Table 7. 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

6.4.4 Bowlers ranking based on Africa Region
Table 8 shows top 10 bowlers of region Africa. 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. For this region we have shown the top ten bowlers in the list in which 4 of them are South African, 2 of them belong to each Pakistan and India and 1 is from Sri Lanka and Bangladesh each.

Table 8. 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

6.4.5 Batsmen ranking based on Europe Region
For 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, the second is MJ Guptil (NZ), 9th is SK Raina (Ind) and 10th is WU Tharanga (Sri) as shown in table 9.
In all batsmen, although MJ Guptill has best strike rate but he has low batting average as compared to RA Jadeja, that is why he is second on the list. On the other hand, Misbahul Haq has a very good batting average, but he has lowest strike rate on the list. That is why he goes to 8thposition in the list of ranking.

Table 9. 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

6.4.6 Bowlers ranking based on Europe Region
For Europe region we have taken 150 top bowlers out of ten full member countries and find their ranking in which MJ Mc Clenaghan (NZ) is the top performer in all bowlers, second is Saeed Ajmal (Pak) and 10th is Harbhajan Singh (Ind) as shown in table 10.

6.4.7 Batsmen ranking based on Oceania Region
For region Oceania, 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). Rankings are shown in table 11. In these batsmen, even Shahid Afridi and AD Russel have very low batting average i.e. 28.6 and 28.5 but highest strike rates i.e. 145.9 and 159.81 respectively, and also, they both perform well against strong bowlers. On the other hand, AB de Villiers has best batting average as compared to all other batsmen on that region and his strike rate is also high that is why he is top on the list.

Table 10. 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

Table 11. 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

6.4.8 Bowlers ranking based on Oceania Region
For 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, 3 bowlers of them are each Australian and South African, and one of each England, India, Pakistan, and Zimbabwe as shown in table 12.

Table 12. 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
WahabRiaz	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

6.4.9 Batsmen ranking based on America Region
Table 13 shows results for region America. For this 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 are Umar Akmal and Misbah Ul Haq (Pak) respectively and 10th is AD Mathews (Sri).

In all batsmen, JH Kallis and Misbah Ul Haq both have very low strike rates, 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.

Although, West Indies have played a lot of matches in this region, but still they have just a single batsman in top ten.

Table 13. 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

6.4.10 Bowlers ranking based on America Region

For region America, we have taken 145 top bowlers out in which B Kumar (Ind) is the top performer in all players due to getting more wickets by conceding very low 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) as shown in table 14.

Table 14. 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

7 CONCLUSIONS

This paper investigates the region-wise players link fusion based ranking in an intuitive way. One can see that region-wise cricket player rankings are quite applicable. It can be also concluded that multi-type interrelated objects both intra-type and inter-type relationships simultaneous modeling is necessary. One can conclude that teams batting, and bowling strength is very important for ranking players. The teams which have high batting strength are stronger in batting line-up as compared to those which have weak team's strength. Similarly, the teams which have stronger bowling strength are stronger in his bowling line-up. Unified players ranking based on regions wise ranking also clearly signifies the importance of our proposed method.

Situational assessment skills were examined in soccer by using video-based simulated task environment [11]. The study investigated both skill and gender-based differences to recognize the ability to predict the opponent's next move or strategy. Intuitively, who are the players which pre-read the mind of opponents? In future, this kind of consideration in addition to existing methods for players' ranking in sports can open a new area of research.

REFERENCES

[1] Akhtar, S., Scarf, P., and Rasool, Z. 2014. Rating players in test match cricket. Journal of the Operational Research Society, 66(4), 684-695.

[2] Govan, Y. A., Meyer, C. D., and Albright, R. 2008. Generalizing Google's PageRank to Rank National Football League Teams. SAS Global Forum Paper 151.

[3] Bailey, M., and Clarke, S. R. 2006. Predicting the match outcome in one-day international cricket matches, while the game is in progress. Journal of sports science & medicine, 5(4), 480.

[4] Daud, A., Muhammad, F., Dawood, H., and Dawood, H. 2015. Ranking Cricket Teams. Journal of Information Processing and Management, 51, 62-73.

[5] Dey, P., Ganguly, M., and Roy, S. 2017. Network centrality based team formation: A case study on T-20 cricket. Applied Computing and Informatics, In Press.

[6] Dingle, N., Knottenbelt, W., and Spanias, D. 2013. On the (page) ranking of professional tennis players. In Computer Performance Engineering, ed: Springer, 237-247.

[7] Duch, J., Waitzman, J. S. and Amaral, L. A. N. 2010. Quantifying the performance of individual players in a team activity. PLoS ONE, 5, e10937.

[8] Haveliwala, T. H. 2002. Topic-sensitive PageRank. In Proceedings of the International Conference on World Wide Web (WWW), 517-526.

[9] ICC. 2017. https://www.icc-cricket.com/about/members, last accessed, October 31, 2017.

[10] ICC.2017.http://www.icc-cricket.com/player-rankings/about, last accessed October 31, 2017

[11] Itay, B. 2013. Game Reading Skills in Soccer. Department of Educational Psychology and Learning Systems, degree granting department, Florida State University.

[12] London, A., Németh, J., and Németh, T. 2014. Time-dependent network algorithm for ranking in sports. Acta Cybernetica, 21, 495-506.

[13] Mukherjee, S. 2012. Identifying the greatest team and captain—A complex network approach to cricket matches. Physica A: Statistical Mechanics and its Applications, 391, 6066-6076.

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

[15] Ribeiro, H, V., Mukherjee, S., Zeng, X. H. T. 2016. The Advantage of Playing Home in NBA: Microscopic, Team-Specific and Evolving Features. PLOS One, http://dx.doi.org/10.1371/journal.pone.0152440.

[16] Shan, Z., Li, S., and Dai, Y. 2012. GameRank: ranking and analyzing baseball network. In International Conference on Social Informatics, 244-251.

[17] Van-Staden, P. J. 2008. Comparison of Bowlers, Batsmen and All-rounders in Cricket Using Graphical Display. Technical Report 08/01, Department of Statistics, University of Pretoria, South Africa.

[18] Xi, W., Zhang, B., Chen, Z., Lu, Y., Yan, S., Ma, W. Y., and Fox, E. A. 2004. Link Fusion: A Unified Link Analysis Framework for Multi-Type Interrelated Data Objects. In Proceedings of the 13th International World Wide Web Conference, 319- 32.

[19] Wikipedia.2017.http://en.wikipedia.org/wiki/List_of_International_Cricket_Council_members, last accessed October 31, 2017

[20] Page, L., Brin, S., Motwani, R. and Winograd, T. 1998. The pagerank citation ranking: Bringing order to the web. Technical report, Stanford Digital Libraries SIDL-WP-1999-0120.

[21] Ahmad, H., Daud, A., Wang, L., Ahmad, I., Hafeez, M., Yang, Y. 2017. Quantifying Team Precedence in the Game of Cricket, Cluster Computing.

[22] Ahmad, H., Daud, A., Wang, L., Hong, H., Dawood, H., Yang, Y. 2017. Prediction of Rising Stars in the Game of Cricket.