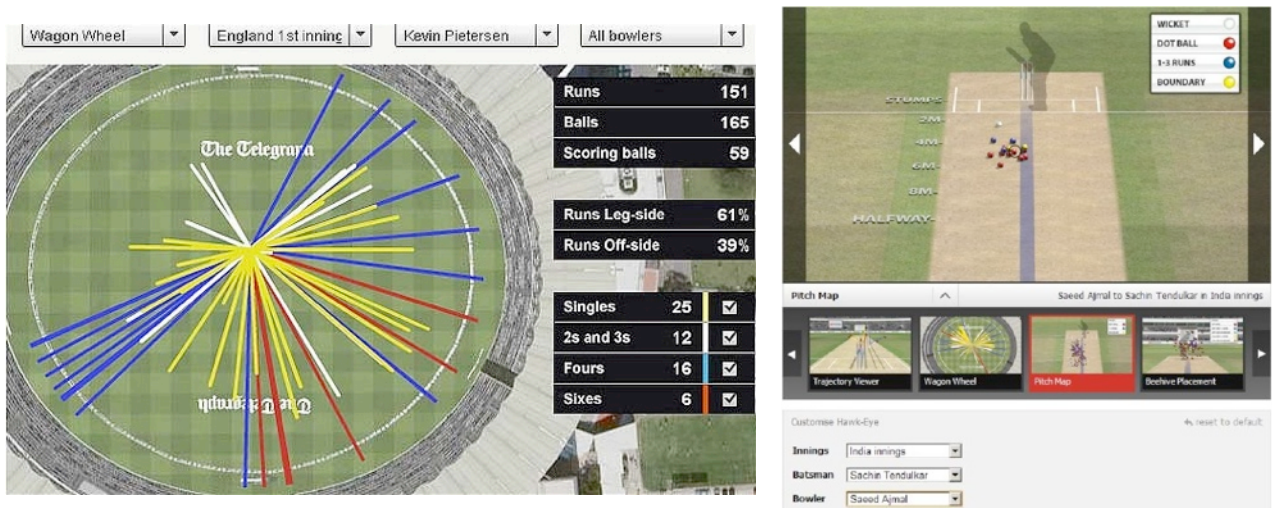*Department of Computer Engineering*
Faculty of Engineering
University of Peradeniya

Project Title : CRICT20 Database: Fully automated IPL-T20 Database

Project Advisor: Sampath Deegalla, **External Advisor**: Rajitha Navarathna
_____

ARE YOU INTERESTING IN CRICKET ? THIS PROJECT IS FOR YOU

With the introduction of T20 to Cricket, the game of Cricket is getting very exciting day by day in all over the world. From all the T20 tournaments, The Indian Premier League (IPL) is the most-watched Twenty20 Cricket league worldwide. IPL consist of different cities in India and each team consist of domestic players and well know international players. Teams such as *Mumbai Indians, Kolkata Knight Riders* and *Chennai Super Kings* have won twice while *Rajasthan Royals* and *Deccan Chargers* have won once. However some of the teams such as *Sunrisers Hyderabad, Delhi Daredevils* could not win a tournament or even to come to finals. Is it due to the unbalance of the teams ? Is it due to poor bowling for certain batsmen in other teams ? Or is it due to poor batting execution for certain bowlers in other teams ? and many more.

To discover these interesting data mining patterns, the project aims to build a T20 Cricket dataset using the previous IPL T20 matches.

PREREQUISITES

Enthusiastic, Excitement in Sports/Cricket

Things will Learn

Team work, Software programming (C/C++), DataBase Management, HTML , Web-based programming

Project Objective

The aim of this project is to create a CRICIPLT20 dataset using IPL T20 matches. Here we use the information from the commentary in Cricinfo website (http://www.espncricinfo.com/indian-premier-league-2014/engine/series/695871.html). The details of the matches are in text format in CricInfo website (eg: http://www.espncricinfo.com/indian-premier-league-2014/engine/match/729279.html?innings=1;view=commentary). The framework should have the ability to read the given webpage and extract ball-byball information related to the commentary in each bowl.

*Eg: **1.1** Malinga to Kallis, 2 runs, opens the face to dab this back of a length delivery behind point. Malinga manages some swing away from the bat as well.*
Information that can be extract from the text: Name of the Batsmen, Name of the bowler, Incident (wicket, runs, wide, no ball etc)
Eg:

BallerID	Batsmen ID	BallNO	Runs	Wickets	NOBalls	Wide	LB Byes	Fours	Sixes	Score
A_10	B_77	1.2	3	0	0	0	0	0	0	12
A_10	B_21	1.3	4	0	0	0	0	1	0	16
A_10	B_21	1.4	0	1	0	0	0	0	0	16

Project Requirements

I. IPL T20 2014 CricInfo matches
II. Programming Languages (Java, C/C++), Web-based programming

Project Phases:-

Phase 1:
- Familiar with IPL T20 commentary web pages
- Understanding the HTML, XML format in those web pages **(GOAL 1)**
- Discover different patterns to extract informations such as bowlers name, batsmen name, runs or wicket for that bowl. **(GOAL 2).** Refer the Table

Phase 2:
- Development of a framework to extract the information for all the overs for both team in every bowl **(GOAL 3). (**Refer figures)
- Integrate the system with MySQL (or any preferable database management) **(GOAL 4)**