

Baseball-AI

Simultaneous Winning Rate Computing

Ji Am Chung^{*}, Yeong Jae Byun[†], Seung Myeon Park[‡], Jun Jeon[§]

^{}Department of Information System
Hanyang University*

Goyang, Gyeonggi Province

Email: hummernk@gmail.com

*[†]Department of Information System
Hanyang University*

Goyang, Gyeonggi Province

Email: qusdudwowo@gmail.com

*[‡]Department of Information System
Hanyang University*

Yongin, Gyeonggi Province

Email: antimoto@nate.com

*[§]Department of Information System
Hanyang University*

Seoul, Korea

Email: jeonjun2@gmail.com

Abstract— Today baseball is becoming more and more popular and getting new fans. Those who just have put their first step into baseball - find it hard to follow the rules and to understand the changes in the situation. Unlike our baseball league KBO, MLB already has many meaningful methods for in-game analyzing with real time statistics. So we are going to adopt those ideas into KBO league. Winning Probability Added, which is also known as WPA, is one of good factors which allows easy understanding of the situation and the impacts that pitcher or batter makes each time. Our project aims to show Winning Rate and WPA factor in real-time, with the algorithm that fits KBO situation.

1. Introduction

Baseball, even though the myriad scandals, attracts more and more fans and is becoming more nationwide sports. There is a major problem that, though baseball fans are inflowing, its hard for beginners to understand the rules and catch situation of the match. Though baseball is much more a number-oriented sport than any others such as soccer or basketball, the classical stats do not fully reflect the match stream and evaluate the value of the players properly. However, in MLB, a lot of sabermetricians have already quantified vague situations and values in numerical way, and KBO tends to follow it. (i.e Babip, OPS, etc..). Korean baseball websites nowadays use the same WPA algorithm which had been used in earlier days of MLB, so the differences between two leagues are not considered. We wanted

to solve this problem by creating our own WPA algorithm based on KBO, which not only calculates the stats suitable for KBO, but also analyzes winning possibilities of each team and the impacts each players make. We also focus on helping beginners to understand baseball. We will make KBO winning rate DB and we are going to put that into our program. The program represents ongoing situation of the match and each teams winning possibilities simultaneously. Also, by showing quantified stats such as batters WPA or pitchers WAR, it shows how much powerful the player is (or has been) throughout the game, seasons or his entire career. Whether baseball match is underway or not, the program will show player rank categorized by players, teams, or positions according to users request. It is not just Baseball statistic calculator, but also somewhat like websites such as Statiz or KReport. The software indicates different types of statistical analysis, and shows them in visualized ways. We thought that current WPA algorithm used in Korean web sites does not fit Korean baseball situation. That is why we have started this project.

This project is composed of 4 steps.

- 1) Crawl the data of last 10 seasons of KBO from baseball statistics websites.
- 2) Compute the data to make some statistics.
- 3) Apply the statistics to WPA algorithm.
- 4) Real-time data capturing and showing the winning rate and WPA.

April 2, 2016

2. Requirements

2.1. Data Handling

2.1.1. Crawling.

- Get every single raw data of KBO from baseball webpage to construct the root database.
- Crawling source : <http://www.koreabaseball.com>.

2.1.2. Capturing.

- Get real-time data when the match is underway.

2.1.3. Real-Time Mirroring.

- Program should immediately renew the database according to the result of the match.

2.1.4. Computing.

- Calculate the numerical data to make some meaningful statistics.
- Every single data has different weight. e.g.) Hits at 1st inning have different value from those at 9th.
- Calculate numerical values including WPA.

2.1.5. Data Storage.

- Save every single stats data.
- Divide players into two tables. One table is for players who is in active service, the other for retired.
- Table for players who is in active service needs to be updated constantly, and the other table doesn't.
- User who wants to conceal the program from screen can do that by clicking window minimization button.

2.2. Function

2.2.1. EXCEL Compatibility.

- User can export datas of specific player or stats to MS Excel files.
- User can import fixed form of MS Excel file of specific game result to compute changes of KBO algorithm winning rate shown as image file which can also be exported as jpeg, gif, png, or bmp.
- User can import fixed form of MS Excel file of specific league(fantasy or amateur) data to compute WAR stats. This data can be exported as EXCEL file.

2.2.2. On-Board Posting.

- Someone who wants to post any idea or thoughts can share what they have.
- Make another Q&A board so as to help beginners solve their curiosity.

2.2.3. Board Log-in & Sign-out.

- Log-in to or Sign-out from Board.
- User who logged-in the board can upload their post or reply to other users

2.2.4. Stats Visualization.

- Show current state of game in a table.
- Show current winning average of each team
- Show current WPA stats of players
- Show player's photograph

2.3. User Interface

2.3.1. Window Minimization & Window Maximization.

- User who wants to see the program widely can do that by clicking window maximization button

2.3.2. Program Turn On & Turn Off.

- User can turn on the program by clicking desktop icon
- If user try to power on the program even if that is already turned on, terminate existing program and launch the program again
- User can turn off the program by clicking x button at the top-right corner of the program

2.3.3. Mouse Click Event.

- Provide user with three options [To Home, Window Minimization, Termination] when user right-click any area within program
- If user double-clicks search button, program shows last request

2.3.4. Player Stat Pop-Up.

- When user clicks certain player, program shows his profile by generating a new pop-up
- If player is a pitcher, pop up list of the first string who has not on the match yet
- If player is taking the field, pop up his profile as batter
- Pitcher pop-up profile stats list : ERA(Earned Run Average) for applicable season, WPA, WAR, WHIP for last 5 matches, (KBB 9), hyperlink connected to NAVER article about him
- Batter pop-up profile stats list : BA(Batting Average), WAR, WPA, OPS for last 5 matches, BABIP for applicable season, hyperlink connected to NAVER article about him
- The number of pop-up cannot be over two

2.3.5. Player Ranking.

- Sort players by team, position, date and game with WPA stats

2.3.6. Data Searching.

- Searching option constitutes of match schedule, player and stats and player
- If option match schedule is chosen, program shows match schedule as a calendar
- If user click one of date, there are three cases. First one is past match, so program shows match log. Second one is on-going match, so program directs user to the match. And the last one is coming match, so program shows every details of the match including players, referees, park, appointed first thrower, weather forecast
- If option player and stat is chosen, program shows the applicable stats separately by entire players, team, position, monthly
- If option player is chosen, program shows every single stat of applicable player

2.3.7. Get Information real-time.

- User can choose the way one gets some information(pop-up or push window)
- Pop-up is a kind of window, so when user have it on the screen, one cannot click main program
- Information could be as follows
- Agreed Decision : User can get information about agreed decision and its details
- Cancellation in case of rain : User can get information when the match is cancelled in case of rain by getting a pop-up or push window
- Player Substitute : In case of substituting player, User can get information why the player was substituted with other, and information about that other player
- When option is pop-up, user can have additional function which is Multi-View. By doing so, user can watch several matches simultaneously
- If there have multiple pop-ups, eliminate pop-up windows sequentially after checking them

2.4. Development Environment

2.4.1. IDE(Integrated Development Environment).

- Developing Language : Google-Go
- Developing Platform : GO
- Use NAVER weather API to get weather information
- Use NAVER time API to get time information

2.5. Execution Environment

2.5.1. Network Environment.

- Accessing to network needs basically
- If network is off-line, user cannot launch program

2.5.2. Least Background Memory.

- There should be enough available memory space to execute the software.