

## An Estimation Model for Defence Ability Using Big Data Analysis in Korea Baseball

Ju-Han Heo\*, Yong-Tae Woo\*

\*Student, Dept. of Computer Engineering, Changwon National University, Gyeongsangnamdo, Korea

\*Professor, Dept. of Computer Engineering, Changwon National University, Gyeongsangnamdo, Korea

### [Abstract]

In this paper, a new model was presented to objectively evaluate the defense ability of defenders in Korean professional baseball. In the proposed model, using Korean professional baseball game data from 2016 to 2019, a representative defender was selected for each team and defensive position to evaluate defensive ability. In order to evaluate the defense ability, a method of calculating the defense range for each position and dividing the calculated defense area was proposed. The defensive range for each position was calculated using the Convex Hull algorithm based on the point at which the defenders in the same position threw out the ball. The out conversion score and victory contribution score for both infielders and outfielders were calculated as basic scores using the defensive range for each position. In addition, double kill points for infielders and extra base points for outfielders were calculated separately and added together.

▶ **Key words:** Big Data Analysis for Baseball, Defence Ability Evaluation Model, Baseball Data Analysis, Baseball Defence Strategy

### [요 약]

본 연구에서는 한국프로야구에서 수비수의 수비 능력을 객관적으로 평가하기 위한 새로운 모델을 제시하였다. 제안 모델에서는 2016년부터 2019년까지 한국프로야구 경기 데이터를 이용하여 구단, 수비 포지션별로 대표 수비수를 선정하여 수비 능력을 평가하였다. 수비력을 평가하기 위하여 포지션별로 수비 범위를 계산하고 계산된 수비지역을 분할하는 방법은 제안하였다. 포지션별 수비 범위는 같은 포지션의 수비수들이 타구를 아웃시킨 지점을 기준으로 컨벡스 헐(Convex Hull) 알고리즘을 이용하여 계산하였다. 포지션별 수비 범위를 이용하여 내야수, 외야수 모두에 대한 아웃전환점수, 승리기여점수를 기본 점수로 계산하였다. 또한, 내야수는 병살 점수, 외야수는 추가 진루 점수를 따로 계산하여 합산하였다.

▶ **주제어:** 야구 빅데이터 분석, 수비수 평가 모델, 야구 데이터 분석, 야구 수비 전략

- 
- First Author: Ju-Han Heo, Corresponding Author: Yong-Tae Woo
  - Ju-Han Heo (herojoon1378@gmail.com), Dept. of Computer Engineering, Changwon National University
  - Yong-Tae Woo (ytwoo@changwon.ac.kr), Dept. of Computer Engineering, Changwon National University
  - Received: 2023. 07. 19, Revised: 2023. 08. 01, Accepted: 2023. 08. 08.

## I. Introduction

미국프로야구에서는 효율적인 야구단 운영을 위하여 머니볼(Money Ball) 이론을 활발하게 활용하고 있다. 머니볼 이론은 최소의 비용으로 최대의 효과를 얻기 위한 경제학적 원칙을 데이터 분석을 통해 야구에 적용하는 방법이다. 미국프로야구 오클랜드야구단의 빌리 빈 단장은 스카우트의 주관적인 의견보다 데이터 분석을 통한 통계적 지표를 이용한 머니볼 이론을 야구단에 적용하여 성공적으로 구단을 이끌었다[1-2]. 최근에 국내외 프로야구에서 데이터 분석을 통해 선수의 타격과 같은 공격 능력뿐만 아니라 수비 능력을 객관적으로 평가하기 위한 연구도 활발하게 진행되고 있다[3-11].

하지만 수비수의 수비 능력을 효율적으로 평가하기 위한 기존 연구는 다음과 같은 문제점들이 존재한다. 먼저, 대부분의 기존 연구에서는 경기장을 일정한 크기로 나누어 수비 능력을 평가한다. 하지만 내야수와 외야수별로 수비 범위가 다르고, 타구의 분포가 다양한 관개로 포지션별로 수비 능력은 차이가 날 수 있다. 그리고 기존 연구에서는 대부분 수비 범위와 수비 능력을 중심으로 평가하는 관개로 수비 시점의 경기 상황을 고려하기 어렵다.

본 연구에서는 한국프로야구에서 수비수의 능력을 효과적으로 평가하기 위한 새로운 모델을 제안한다. 본 연구에서는 수비 포지션별로 선수들이 타구를 아웃시킨 지점을 기준으로 컨벡스 쉘(Convex Hull) 알고리즘을 이용하여 수비 범위를 객관적으로 분할하기 위한 새로운 방법을 제안하였다.

## II. Preliminaries

### 1. Related works

#### 1.1 Study on existing model for evaluating the defenders

미국프로야구에서는 주로 스탯캐스트를 통해 수집된 경기 데이터를 이용하여 수비수를 평가한다. 하지만 미국프로야구에서 수비수 평가 지표로 사용되고 있는 UZR(Ultimate Zone Rating)의 수비지역은 크기가 고정되어 포지션별로 타구의 분포를 효과적으로 반영하기 어렵다[4]. 또한, UZR의 수비지역은 매년 변하는 타격이나 수비력을 적절하게 반영하기 어렵다.

한국프로야구에서는 수비수의 수비지역을 고려한 평가 지표가 미흡한 실정이다. 또한, 국내에서 외야수의 송구

정확도를 평가하는 기존 연구에서는 송구의 정확도만 고려하고 주루저지능력, 병살률 등과 같은 상세한 수비 능력을 고려하지 않았다[11].

### 1.2 UZR

UZR은 미국프로야구에서 가장 널리 사용되고 있는 수비수 평가를 위한 통계지표 중의 하나이다[4]. UZR은 자신의 능력으로 실점을 막았는지 자신의 실수로 실점하게 되었는지를 판단하기 위하여 야구 경기장을 일정하게 64개 구역으로 나누어 평가하는 지표이다. 이 지표는 각 항목 점수의 합으로 표현되며, 점수는 개인의 점수와 해당 포지션 전체 수비수 평균의 차로 계산된다[12]. Table 1은 UZR서 사용되는 4종류의 항목이다[13].

Table 1. Four evaluation items used in UZR

Item	Value
ARM	The sum of points fielded by outfielders through extra base stoppages
DPR	The sum of points fielded after an infielder drew a double play
RngR	Evaluate the degree to which a batted ball was converted into an out through the defense area
ErrR	Evaluating the difference in number of errors compared to average defenders

Table 1에서 ARM은 외야수가 추가 진루를 막았을 때 평균 외야수와 비교하여 수비한 점수의 총합을 나타낸다. DPR은 내야수가 병살을 이끌어냈을 경우 평균 내야수와 비교하여 수비한 점수의 총합을 나타낸다. RngR은 UZR의 분할된 수비 범위를 통해 평균 수비수보다 얼마나 많은 타구를 아웃으로 전환 시켰는지를 평가한다. ErrR은 각 수비수 포지션별 평균 수비수들과 비교하여 에러를 범한 정도를 평가하는 항목이다.

## III. The Proposed Scheme

### 1. Overview

본 연구에서는 한국프로야구에서 수비수의 수비력을 효과적으로 평가하기 위한 새로운 모델을 제안하였다. 제안 모델에서 수비 범위는 아웃 타구 좌표 데이터에서 같은 포지션의 수비수들이 타구를 아웃시킨 지점과 컨벡스 쉘(Convex Hull) 알고리즘을 이용하여 계산하였다[15]. 이후 해당 수비 범위에서 수비수가 서 있는 지점을 중심으로 8개의 범위로 분할한다. 2016년부터 2019년까지 연도별

한국프로야구 경기 데이터를 이용하여 내야수와 외야수의 수비지역과 수비 항목별로 평균 값을 구하였다. 평균 수비 지역과 항목별 평균 값을 이용하여 2019년도 한국프로야구 경기 데이터로부터 수비 항목별로 평가를 진행하였다.

$$\begin{aligned}
 \text{Infielder Evaluation Score} &= \text{Base Score} + \text{DPS} \\
 \text{Outfielder Evaluation Score} &= \text{Base Score} + \text{ARS} \\
 \text{Base Score} &= (\text{OCS} + \text{WCS})
 \end{aligned}$$

Equation 1. Evaluation Equation for infielders and outfielders

Equation 1에서처럼 내야수는 아웃전환점수, 승리기여 점수, 병살점수의 합으로 평가하였다. 외야수는 아웃전환 점수, 승리기여점수, 추가진루점수의 합으로 평가하였다. 본 연구에서 사용한 데이터는 (주)하이볼 사에서 제공한 2016년부터 2019년 연도별 한국프로야구 경기 데이터를 이용하였다.

2. Defensive range measurement by position

본 연구에서는 수비수의 객관적인 평가를 위해 수비 포지션별 수비 범위를 측정하였다. 수비 범위는 포지션별 수비수의 능력으로 수비를 할 수 있는 평균 범위를 의미한다. Fig. 1은 수비 범위 측정 및 수비 범위를 분할하는 흐름도이다.

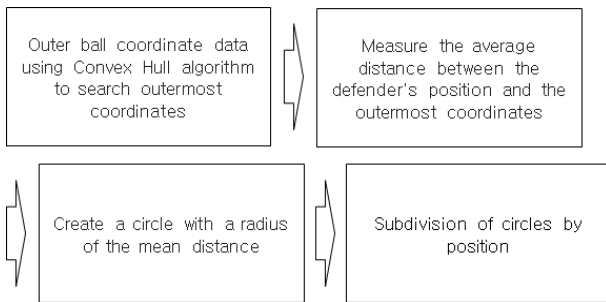


Fig. 1. Defensive range measurement and division flow chart

수비 범위 측정은 포지션별 아웃 타구 좌표 데이터를 컨벡스 헐 알고리즘을 이용하여 포지션별로 최외곽 좌표를 찾아낸다. 이후, 최외곽 좌표 사이의 평균 거리를 계산하여 포지션별 평균적인 수비 범위를 측정한다. 여기서 컨벡스 헐 알고리즘은 임의의 집합 X가 있을 때 집합 X를 모두 포함하는 가장 작은 Convex Set(볼록 집합)을 구하는 알고리즘이다[15]. 수비 포지션은 1, 2, 3 루수, 유격수, 좌익수, 중견수, 우익수 총 7개의 포지션을 대상으로 진행하였다. Fig. 2는 투수와 포수를 제외한 7개의 포지션을 대

상으로 컨벡스 헐 알고리즘을 이용하여 한국프로야구에서 수비수가 타구를 아웃시킨 범위를 나타낸다.

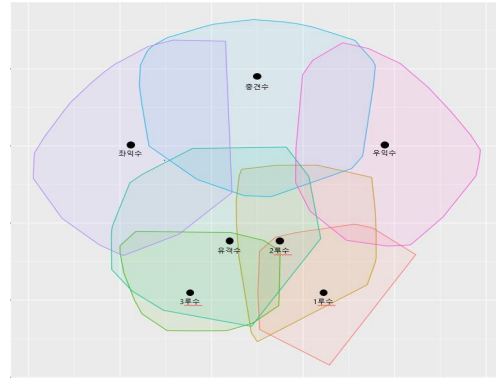


Fig. 2. The range at which a batted ball was thrown out by a defender in Korean professional baseball using the convex hull algorithm for seven positions excluding pitcher and catcher.

Equation 2는 포지션별 아웃 타구의 최외곽 좌표와 수비수 위치 사이 평균 거리를 구하기 위한 수식이다.

$$D_i = \frac{\sum_{n=0}^N \sqrt{(px_i - x_n^i)^2 - (py_i - y_n^i)^2}}{N_i}$$

$$i = \{1, 2, 3, 4, 5, 6, 7\}$$

$N_i$  = Number of outermost coordinates of  $i$ -position

$px_i$  =  $i$  position  $x$  coordinate

$py_i$  =  $i$  position  $y$  coordinate

$x_n^i$  = Outermost  $x$ -coordinate of the  $n$ th out batted ball at position  $i$

$y_n^i$  = Outermost  $y$ -coordinate of the  $n$ th out batted ball at position  $i$

Equation 2. Equation for the average distance between the outermost coordinates of an out hit ball by position and the position of the defender

Table 2는 Equation 2에서 사용되는  $i$ 가 의미하는 포지션에 대한 정의이다.

Table 2. Definition of the position meant by recipe used in Equation 2

$i$	1	2	3	4	5	6	7
	1B	2B	3B	SS	LF	CF	RF

Fig. 3은 Equation 3의 반지름으로 생성한 외야수의 수비 범위를 나타낸다.

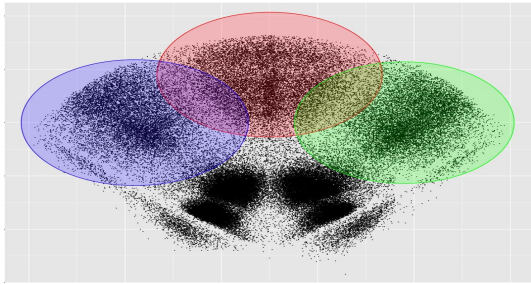


Fig. 3. The outfielder's fielding range generated by the radius in equation 3

### 3. Detailed division of defense range by position

수비 범위는 수비수 위치와 낙구 지점 사이의 거리, 타구가 날아오는 방향성을 고려하기 위해 세부 분할을 진행하였다. 수비수 위치와 낙구 지점 사이의 거리를 고려하기 위해서 포지션별 수비 범위보다 작은 1개의 원을 추가하여 세부 분할하였다. 또한, 타구가 날아오는 방향성을 고려하기 위해 타구가 날아오는 홈플레이트와 수비수의 위치를 이용하여 수비 범위를 세부 분할하였다. 포지션별 원은 세부 분할 결과 총 2개의 원으로 구성된다. 수비 범위 세부 분할 결과 포지션별 수비 범위는 총 8개의 수비 범위로 분할되었다.

### 4. Out conversion score

본 연구에서 제안한 아웃전환점수는 수비 범위에 떨어진 모든 타구 중 타구를 아웃으로 전환한 확률을 점수화한 것이다. Equation 3은 아웃전환점수를 계산하는 수식이다.

$$OCS_i = \sum_{n=0}^8 (IOCR_n - POCR_n^i) \times c_n^i$$

$i = \text{specific defender}$

$n = \{1, 2, 3, \dots, 8\}$

$POCR_n = \text{Average out conversion rate for } n \text{ region}$

$IOCR_n^i = i \text{ defender } n \text{ area out conversion rate}$

$c_n^i = \text{The number of out hits in area } n \text{ of defender } i$

Equation 3. Out conversion score calculation equation

Equation 3에서  $n$ 은 본 연구에서 제안한 포지션별 수비 범위 8개의 번호를 의미하고  $i$ 는 평가 대상이 되는 개인 수비수를 의미한다. 아웃전환점수는 포지션별 평균 아웃전환율과 개인 평균 아웃전환율의 차에 아웃 타구 발생 횟수를 곱하여 계산하였다.

### 5. Win contribution score

승리기여점수는 수비수의 수비 결과가 팀 승리에 기여한 정도를 의미한다. 본 연구에서 제안한 승리기여점수는 문형우가 제안한 마르코프 체인을 이용하여 계산한 한국 프로야구 기대 승리확률을 이용하여 산출하였다[10]. Equation 4는 승리기여점수를 계산하는 수식이다.

$$WCS_i = \begin{cases} \sum_{n=1}^S (1 - POCR_n) \times (WE_{after}^n - WE_{before}^n) & (\text{When defence successes}) \\ \sum_{n=1}^{F_i} (POCR_n) \times (WE_{after}^n - WE_{before}^n) & (\text{When defence fails}) \end{cases}$$

$i = \text{specific defender}$

$S_i = \text{The number of balls successfully fielded by defender } i$

$F_i = \text{The number of batted balls that defender } i \text{ failed to defend}$

$POCR_n = \text{Average out conversion rate in the defensive area where the } n\text{th batted ball landed}$

$WE_{after-before}^n = \text{Expected win probability before and after the } n\text{th hit defense result}$

Equation 4. Win contribution score calculation equation

승리기여점수는 수비를 성공한 경우와 실패한 경우를 나누어서 계산하였다. 수비 성공 시 가점을 받고 실패 시 감점을 받도록 계산하였다. 수비 성공 시 수비가 어려운 범위일수록 높은 점수를 받아야 하므로 아웃전환실패율인  $1-POCR$ 을 곱하여 계산하였다. 수비 실패 시 수비가 어려운 범위일수록 낮은 점수를 받아야 하므로 아웃전환율인  $POCR$ 을 곱하여 계산하였다.

### 6. Double play score

병살점수는 병살이 일어날 수 있는 상황에서 병살을 이끌어낼 확률을 점수화한 것이다. 병살이 일어날 수 있는 상황은 타구가 내야를 지나가는 땅볼이면서 주자 상황이 주자 1루와 주자 1, 2루와 주자 1, 3루 그리고 주자 만루인 상황으로 정의하였다. Equation 5는 병살확률을 이용한 수비수를 평가하기 위한 식이다.

$$DPS^{p_i} = IDPR_i - PDPR_p$$

$i = \text{specific defender}$

$p = \{1, 2, 3, 4\}$

$IDPR_i = \text{Defender } i\text{'s individual double play rate}$

$TDPR_p = \text{Overall average double play rate at position } p$

Equation 5. Defender evaluation formula using double kill probability

Equation 5의 병살점수는 내야수 포지션별 평균 병살률과 개인 평균 병살률의 차로 계산하였다. 평균 병살률의 차이가 양수이고 결괏값이 클수록 평균 내야수보다 병살을 많이 이끌어내는 것을 의미한다. Table 3은 Equation 5에서 이용한  $p$ 의 정의를 나타내는 표이다.

Table 3. The position that  $p$  used in Equation 6 means

$p$	1	2	3	4
	1B	2B	3B	SS

### 7. Additional runs score

추가 진루확률은 선행주자가 있는 상황에서 외야수가 타구를 수비했을 때 선행주자가 타구로 인해 추가적인 진루를 할 확률을 점수화한 것이다. 추가 진루확률은 타구가 떨어지는 범위에 따라 진루 결과에 영향을 미치기 때문에 같은 수비 범위에 떨어진 타구에 대해서 계산하였다. Table 4는 추가 진루 점수에서 평가 대상 주자 상황표이다.

Table 4. Conditions of runners to be evaluated in the additional runs score

Runner situation	Evaluated Runner
Bases 1	Runner on first base
Bases 2	Runner on second base
Bases 1, 2	Runner on second base
Bases 3	Runner on third base
Bases 1, 3	Runner on third base
Bases 23	Runner on third base
Bases full	Runner on third base

야구에서 주자 1로부터 만루까지 총 7가지의 주자 상황이 존재한다. 추가 진루점수 평가에 이용되는 주자는 주자 상황에서 제일 앞서 있는 주자이다. 예를 들어, 주자 1, 2루 상황에서는 제일 앞선 2루 주자가 주루를 시도하지 않으면 1루 주자는 주루를 시도할 수 없다. 따라서 제일 앞선 주자의 진루 결과를 평가에 이용하였다. Table 5는 추가 진루점수에서 이용되는 타격결과와 정상 진루에 대한 정의이다.

Table 5. Definition of normal run base used in the evaluation of additional runs score

Evaluation runner	Hitting result	Advance base
First Base	Single	Second Base
	Double	Third Base
	Fly Ball	First Base
Second Base	Single	Third Base
	Double	Home
	Fly Ball	Second Base
Third Base	Single	Home
	Double	Home
	Fly Ball	Third Base

야구에서 진루는 안타뿐만 아니라 직선타 혹은 플라이아웃을 시킨 후에도 진루할 수 있다. 따라서 평가에 이용하는 타격결과는 1, 2루타와 플라이아웃으로 구성하였다. Equation 6은 추가 진루점수를 계산하는 수식이다.

$$ARS_i = \sum_{z=1}^8 \sum_{b=1}^3 \sum_{h=1}^3 \sum_{r=1}^4 (PAVG_{z b h r}^i - IAVG_{z b h r}^i)$$

$i = specific\ defender$

$z = \{1, 2, 3, \dots, 8\}$

$b = \{1, 2, 3\}$

$h = \{1, 2, 3\}$

$r = \{1, 2, 3, 4\}$

Equation 6. Additional runs score equation

Equation 6에서  $z$ 의 1부터 8까지는 본 연구에서 제안한 포지션별 8개의 수비 범위를 의미하고,  $b$ 는 Table 5에서 정의한 평가 대상 주자 1루, 2루, 3루를 의미한다.  $h$ 는 Table 6에서 정의한 타격결과 1루타, 2루타, 플라이아웃을 의미하고  $r$ 은 각 주자의 진루 결과인 1루, 2루, 3루, 홈을 의미한다. PAVG와 IAVG는 각각 수비 범위, 주자 상황, 타격결과, 진루 결과를 고려한 포지션별 평균 추가 진루확률과 개인 평균 추가 진루확률을 의미한다. 추가 진루점수는 수비 범위, 주자 상황, 타격결과, 진루 결과를 고려한 포지션별 평균 추가 진루확률과 개인 평균 추가 진루확률의 차를 이용하여 계산하였다.

### IV. Experimental Results

제안 모델에서는 2016년부터 2019년도까지 한국프로야구 경기 데이터 중에서 투수와 타자 간 발생한 타격 결과로 인해 이루어진 수비, 송구 기록 등의 데이터를 수비수를 평가하기 위한 데이터로 이용하였다. 수비수를 평가하기 위한 항목인 승리기여점수, 추가 진루점수, 병살점수를 계산하기 위해 아웃 상태, 주자 상황 등의 데이터를 이용하였다. 내야수에 대한 평가 항목은 아웃전환점수, 승리기여점수, 병살점수이고, 외야수의 평가 항목은 아웃전환점수, 승리기여점수, 추가 진루점수를 평가하였다. 수비수 평가는 각 항목의 합으로 평가하는 관계로 각 항목의 점수는 Z-점수로 표준화한 후 합산하였다[14].

#### 1. Outfielder Evaluation Results

##### 1.1 Right field evaluation items and final score

Table 6은 본 연구에서 제안한 수비수 평가 모델에서 제안한 Out Conversion Score, Win Contribution Score, Additional Runs Score 항목을 이용하여 2019년 한국프로야구 우익수들의 평가 결과이다.

Table 6. Korean professional baseball 2019 evaluation result of right fielders

Rank	Player	OCS	WCS	ARS	Total
1	Smolinski	0.6471	1.1132	1.2643	3.0246
2	Park Geonu	0.6489	0.3748	0.8735	1.8973
3	Kang Baekho	0.2110	0.8318	0.7295	1.7723
4	Che Eunseong	0.1756	0.5559	0.7751	1.5067
5	Lee Myeonggi	0.2182	1.1473	-1.0527	0.3129
6	Hoying	1.6529	-0.1925	-1.7544	-0.2941
7	Koo Jauk	0.2412	-0.3678	-0.6294	-0.7560
8	Han Dongmin	-0.8115	-0.2853	-0.2055	-1.3023
9	Sands	-1.7544	-1.6115	0.6634	-2.7024
10	Son Aseop	-1.2291	-1.5660	-0.6639	-3.4590

Table 6에서 스몰린스키 선수는 모든 항목에서 높은 점수를 획득하여 1위를 기록하였다. 반면 손아섭 선수는 모든 항목에서 평균 이하의 점수를 얻어 10위를 기록하였다.

##### 1.2 Analysis of outfielder evaluation results

2019년 한국프로야구 공식기록에서 수비수 평가 지표인 수비율은 우익수 구자욱 선수 99.1%, 우익수 스몰린스키 선수 97.9%로 구자욱 선수가 더 높은 수비율을 기록

하였다. 하지만 본 연구에서 제안한 아웃전환점수, 승리기여점수, 추가진루점수를 이용한 측정결과 우익수 스몰린스키 선수는 1위, 우익수 구자욱 선수는 7위를 기록하였다. Fig. 4는 우익수 수비수 수비 범위와 세부 수비 범위를 명명한 결과이다.

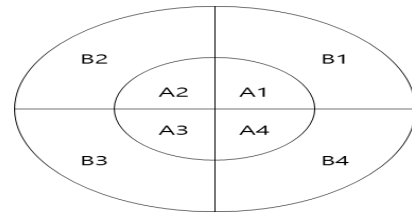


Fig. 4. Result of naming the right field defensive range and detailed defensive range

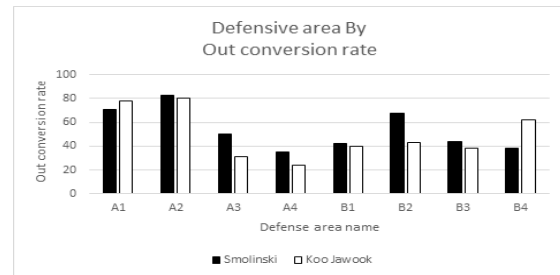


Fig. 5. Right fielder Smolinski and Koo Ja-wook's out conversion rate by detailed defensive range

Fig. 5에서 x축은 수비 범위 명으로 아웃전환율을 이용하여 오름차순으로 정렬하였다. y축은 수비 범위별 아웃전환율을 나타낸다. 1위를 기록한 스몰린스키 선수는 7위를 기록한 구자욱 선수보다 상대적으로 많은 수비 범위에서 구자욱 선수보다 높은 아웃전환율을 기록하였다. Table 7은 우익수 스몰린스키와 구자욱 선수의 수비 결과에 따른 승리기여도 기여/감점 횡수를 나타낸 Table이다.

Table 7. Right fielders Smolinski and Koo Ja-wook contributed to victory, number of points deducted, and probability of contribution

Player	Number of contribution	Number of deduction	Contribution probability
Smolinski	137	118	53.73
Koo Jauk	212	216	49.53

Table 7에서 1위를 기록한 스몰린스키 선수는 기여 횡수 137번 감점 횡수 118번으로 53.73%의 기여 확률을 기록하였다. 반면 7위를 기록한 구자욱 선수는 기여 횡수 212번 감점 횡수 216번으로 49.53%의 기여 확률을 기록하였다.

## V. Conclusions

본 연구에서 제안한 수비수 평가 모델은 한국프로야구 수비수의 수비력을 효과적으로 평가하기 위한 모델이다. 제안 모델은 아웃전환점수, 승리기여점수, 병살점수, 추가진루점수 등의 새로운 평가 항목을 이용하여 수비수의 수비력을 다양한 항목으로 평가할 수 있는 모델이다. 본 연구에서는 경기 상황, 타구의 난이도, 낙구 지점, 수비의 가치 등을 다양한 상황을 평가 항목으로 고려한 새로운 형태의 수비수 평가 모델이다.

현재 한국프로야구에서는 자살, 보살, 실책 횟수 등을 이용한 단순한 형태의 수비율에 의해 수비수를 평가하는 문제점이 있다. 본 연구에서 제안한 평가 모델은 구체적인 항목을 통한 효과적인 수비수 평가 모델이라 생각한다. 앞으로 현재 단계에서는 수집이 어려운 타구의 속도, 각도 등의 데이터를 수집하여 새로운 수비 범위 생성과 구장마다 다른 야구장 크기를 고려하여 수비 평가를 할 수 있는 추가적인 연구가 필요하다.

본 연구에서 제안한 수비수 평가 모델은 야구팬들에게 현재 한국프로야구에서는 제공하지 않는 데이터와 평가 결과를 제공하여 객관적인 정보를 제공할 수 있다. 또한, 야구 관계자에게는 상황에 맞게 필요한 수비수를 투입하고 경기 전력 수립에 사용할 수 있다. 그리고 수비수 평가 결과를 이용하여 선수 영입과 합리적인 연봉 협상에도 효과적으로 활용할 수 있으리라 기대한다.

## ACKNOWLEDGEMENT

This research is financially supported by Changwon National University in 2021~2022

## REFERENCES

- [1] J. H. Cho, "Utilization and Prospect of Sport Big Data," *The Korean Journal of Measurement and Evaluation in Physical Education and Sport Science*, Vol. 14, No. 3, pp. 1-11, Dec. 2012. DOI: 10.21797/ksme.2012.14.3.001
- [2] A. Utami, "Analysis of Decision Making Process in Moneyball: The Art of Winning an Unfair Game," *The Winners*, Vol. 16, No. 1, pp. 57-63, Mar. 2015. DOI: <https://doi.org/10.21512/tw.v16i1.1555>
- [3] N. M. Watanabe, S. Shapiro, and J. Drayer, "Big Data and Analytics in Sport Management," *Human Kinetics Journals*, Vol. 35, No. 3, pp. 197-202, Apr. 2021.
- [4] M. G. Litchman, "Ultimate zone rating," *The Baseball ThinkFactory*, 2013.
- [5] S. T. Jensen, K. E. Shirley, and A. J. Wyner, "Baseball: a Bayesian Hierarchical Model for Evaluating Fielding in Major League Baseball," *The Annals of Applied Statistics*, Vol. 3, No. 2, pp. 491-520, Jun. 2009. DOI: 10.1214/08-AOAS228
- [6] J. M. Kim, C. Dibrell, E. Kraft, and D. Marshall, "Data analytics and performance: The moderating role of intuition-based HR management in major league baseball," *Journal of Business Research*, Vol. 122, Jan. 2021. DOI: <https://doi.org/10.1016/j.jbusres.2020.08.057>
- [7] J. Y. Park, J. S. Kim, and Y. T. Woo, "A motion classification and retrieval system in baseball sports video using Convolutional Neural Network model," *Journal of Korean Data Information Science Society*, Vol. 26, No. 8 pp. 31-37, Aug. 2021. DOI: 10.9708/jksci.2021.26.08.031
- [8] J. T. Lee, "Pitching grade index in Korean pro-baseball," *Journal of Korean Data Information Science Society*, Vol. 25, No. 3 pp. 485-492, May. 2014. DOI: <https://doi.org/10.7465/jkdi.2017.28.2.317>
- [9] Y. J. Seo, H. W. Moon, and Y. T. Woo, "A Win/Lose prediction model of Korean professional baseball using machine learning technique," *Journal of The Korea Society of Computer and Information*, Vol. 24, No. 2, pp. 17-24, Nov. 2019. DOI: 10.9708/jksci.2019.24.02.017
- [10] H. W. Moon, Y. T. Woo, and Y. W. Shin "Run expectancy and win expectancy in the Korea Baseball Organization (KBO) League," *The Korean journal of applied statistics*, Vol. 29, No. 2, pp. 321-330, Feb. 2016. DOI: 10.5351/KJAS.2016.29.2.321
- [11] S. H. Lee, Y. R. Jung, H. W. Moon, and Y. T. Woo, "A Baseball Batter Evaluation Model using Genetic Algorithm," *Journal of The Korea Society of Computer and Information*, Vol. 24, No. 1, pp. 41-47, Nov. 2019. DOI: 10.9708/jksci.2019.24.01.041
- [12] FanGraphs, "The Fangraphs UZR Primer," <http://www.fangraphs.com/blogs/the-fangraphs-uzr-primer/#16>.
- [13] FanGraphs, "UZR Primer" <http://www.fangraphs.com/library/defense/uzr>
- [14] Naver Wikipedia, "Z Score" <https://terms.naver.com/entry.naver?docId=275110&cid=50315&categoryId=50315>
- [15] R. V. Chadnov, A. V. Skvortsov, "Convex hull algorithms review," *Proceedings. The 8th Russian-Korean International Symposium on Science and Technology, KORUS 2004, Tomsk, Russia*, Vol. 2, pp. 112-115, Jan. 2004.

## Authors



Ju-Han Heo received the B.S. and M.S. degrees in Computer Engineering from Changwon National University, Korea, in 2015 and 2017, respectively. Mr. Heo is a Developer in Hibrain.net Co.

He is interested in Data Modeling and Data Analysis.



Yong-Tae Woo received the B.S., M.S. and Ph.D. degrees in Computer Science and Engineering from Kyungpook National University, Korea, in 1982, 1984 and 1995, respectively.

Dr. Woo is a Professor in the Department of Computer Engineering, Changwon National University since 1987. He is also CEO of Hibrain.net Co. He is interested in Data Modeling, Internet Business, and Big Data Analysis areas.