한국컴퓨터정보학회 하계학술대회 논문집 제26권 제2호 (2018. 7)

빅데이터 분석과 게임이론을 활용한 야구선수 영입 모델

김윤후[O], 김상헌[*], 최형준[*], 정재은[*]
[O*]중앙대학교 컴퓨터공학과
e-mail: {thomas77777[O], comnamu1[*], sangpon[*]}@cau.ac.kr, j3ung@cau.ac.kr[*]

Baseball Player Scouting Model using Big Data Analysis and Game Theory

Yun-Hu Kim[O], Sang-Heon Kim[*], Hyung-June Choi[*], Jai-Eun Jung[*]
[O*]Dept. of Computer Engineering, Chung-Ang University

● 요 약 ●

빅데이터 분석은 스포츠에서 다양한 분야에서 사용되고 있다. 야구산업에서도 세이버 메트릭스를 활용하여 전술 훈련, 개인 훈련 등 다양한 방면으로 활용되고 있다. 본 논문에서는 기존의 연구인 빅데이터 분석과 게임이론을 활용한 축구선수 영입 모델을 야구에 적용킨 시뮬레이션을 진행하고 합리적인 의사결정 모델을 제안한다.

키워드: 산업구조분석(Industrial Structural Analysis), 경쟁분석(Competitive Analysis), 차별성(Differentiation)

Ⅰ. Introduction

2011년 개봉한 영화 '머니볼'은 메이저리그 최하위팀인 오클랜드 애슬레틱스(Oakland Athletics)가 야구 데이터 분석을 통해 기적을 이뤄내는 영화이다. 이와 같이 빅데이터 분석은 이미 스포츠 산업에서 다양한 방식으로 적용되고 있으며 이를 스포츠 데이터 분석이라 부른다.

KBO(Korea Baseball Organization League)에서도 선수별 타점, 타율, 희생번트, 출루율 등 다양한 데이터를 기록한 후 야구를 통계학적, 수학적으로 분석하는 방법론인 세이버 메트릭스(Sabermetrics)를 활용하여 전술 훈련이나 개인 훈련 등 여러 방면으로 활용하고 있다[1].

그렇지만 현재의 KBO 이적 시장은 스카우터가 경기를 관람하며 그들의 직감에 의해서 선수를 영입하는 경우가 대부분이다. 이러한 영입 방식은 논리성과 객관성을 만족하기에는 어려움이 있다.

본 논문에서는 기존에 연구된 기대 퍼포먼스 E를 야구에 적용시키기 위하여 새로운 야구 기대퍼포먼스 BE를 정의하였다. 게임이론을 대입하여 각 야구 구단이 최적화된 영입을 할 수 있도록 하는 것을 목표로 하며 축구뿐만이 아닌 야구에서도 이를 적용할 수 있음을 보여준다[2].

Ⅱ. Backgrounds

2.1 선수 역량 척도

기존의 연구에서는 공신력을 가지는 축구선수 평점 측정 사이트인 '후스코어드(Whoscored)'의 평균평점을 r로 정의하고 선수 가치를 예측한 사이트인 '트랜스퍼 마르크트(Transfer Markt)'의 선수가치데이터를 v로 정의하였다.

임의의 포지션 $m_k \in M$에 대하여 임의의 선수 $p_i \in P$의 기대 퍼포먼스 E는 다음의 수식(1)과 같다.

$$E_i^m = \frac{r_i^{m_k} - \mu(r^{m_k})}{\sigma(r^{m_k})} - \frac{v_i^{m_k} - \mu(v^{m_k})}{\sigma(v^{m_k})} \quad (1)$$

2.2 게임이론 대입

기존 연구에서는 구단과 구단의 경쟁구도가 아닌 내쉬 균형을 통한 합리적인 선수 영입 모델을 제안했다. 각 구단들은 $Min(v_i)$, $p_i \in P$로 $Max(r_i)$, $p_i \in P$를 기대한다. 하지만 E가 높은 선수들의 수요가 증가하게 되면 가격경쟁으로 인해 기존의 v보다 높은 값을 지불해야하기 때문에 임의의 팀 $t_a \in T$에 대한 합리적인 의사결정 D는 수식(2)와 같다.

$$D(T_a, T_b) = \arg max(P_i + P_j) \quad (2)$$
(단, $p_i, p_j \in P^{\forall}$ 이며 $p_i \neq p_j$)

III. Modeling & Simulation

본 논문에서는 기존의 연구인 빅데이터 분석과 게임이론을 활용한 축구선수 영입 모델을 야구 도메인에 적용하여 시뮬레이션을 진행해보았다. 본 시뮬레이션은 KBO 데이터를 분석하는 공신력 있는 사이트인 '스테티즈(Statiz)'의 2017 KBO 선수 데이터를 기반으로 하였다. 좌익수와 우익수의 모집단을 추출한 후 수식(4)를 통해 BE를 계산하였다. 기존의 축구를 도메인으로 한 연구는 평점 데이터 r과 가치 데이터 v를 통하여 E를 계산하였지만 본 시뮬레이션에서는 우익수와 좌익수의 수비 능력을 측정해야하기 때문에 스테티즈에서 정의한 평균 선수 대비 득점 기여 수치인 sFR RAA(statiz Fielding Runs Runs Above Average)를 sr로, 포지션 조정이 포함된 평균 대비 수비 승리 기여 지표인 WAA with ADJ(Wins Average with position Adjust)를 wa로, 선수 연봉은 s로, 선수의 연차는 a라고 정의하였다[3]. KBO에서는 선수 가치측정 시 구단마다 서로 다른 가산점과 감산점을 이용하고 있기에 본 논문에서는 임의의 선수 포지션 $p_i \in P$에 대하여 임의의 포지션 $m_k \in M$의 출장 경기 수 f_m^p라 정의하면 선수의 가치 pv를 수식(3)과 같이 정의한다[4].

$$pv = (s/a) * f_m^p \tag{3}$$

위의 변수들을 기반으로 한 야구 기대 퍼포먼스 BE는 수식(4)와 같다.

$$BE_i^m = \frac{sr_i^{m_k} - \mu(sr^{m_k})}{\sigma(sr^{m_k})} + \frac{wa_i^{m_k} - \mu(wa^{m_k})}{\sigma(wa^{m_k})} - \frac{pv_i^{m_k} - \mu(pv^{m_k})}{\sigma(pv^{m_k})} \tag{4}$$

야구에서는 좌익수로 기용된 선수가 우익수로 기용되기도 하고 우익수로 기용된 선수가 좌익수로 기용되기도 한다. 즉, 각 구단의 전술 및 상황에 따라 한 선수가 좌익수 혹은 우익수에 배정되는 것이다.

예컨대 조수행(두산 베이스), 이명기(기아 타이거즈), 박정음(넥센 히어로즈)이 동시에 자유 이적 시장 리스트에 올랐고 A팀은 우익수를 B팀은 좌익수를 필요로 한다고 가정했다.

Table 1. KBO 외야수 시뮬레이션 결과

좌익 / 우익	조수행	이명기	박정음
조수행	2.64(2.25)	2.64(1.13)	2.64(0.85)
이명기	2.16(2.25)	2.16(1.13)	2.16(0.85)
박정음	0.86(2.25)	0.86(1.13)	0.86(0.85)

<표 1>의 값들은 각 선수들이 우익수로 기용되었을 경우와 좌익수로 기용되었을 경우를 구분하여 BE값을 계산한 표이다. 값이 클수록 투자 대비 수비 능력 효율이 높음을 의미한다.

합리적 의사결정을 위한 D값이 가장 높은 경우는 4.89로 A팀과 B팀 모두 조수행을 영입했을 때이다. 하지만 한 선수를 영입하기 위해 두 팀이 경쟁을 하면 더 높은 금액의 연봉을 지불해야하고 연봉이 상승하면 BE값은 하락한다. 즉, $p_i \neq p_j$의 조건과 내쉬 균형을 만족시키는 경우인 D=4.41을 선택하는 것이 가장 합리적인 의사결정이다. D=4.41의 경우는 B팀에서 조수행을 좌익수로 영입하고 A팀에서 이명기를 우익수로 영입하는 방법이다.

IV. Conclusions

본 논문은 야구 산업에서의 선수 이적이 객관성과 논리성이 결여된 감독의 직감으로만 이루어진다는 문제점에서 출발하였다. 이에 기존의 연구에서 제시한 빅데이터 분석과 게임이론을 통한 축구선수 영입 모델을 야구에 적용시켜보았다. sFR RAA, WAA with ADJ, 연봉, 연차, 포지션 별 출장 경기 수의 변수를 활용하여 야구 기대 퍼포먼스인 BE를 제시하여 시뮬레이션을 진행하였다.

본 연구에서 제안한 야구 기대 퍼포먼스는 수비에 국한된 자수이며 다양한 변수들을 포함시키지 못했다는 아쉬움을 지닌다.

향후에는 이 한계점들을 보완한 새로운 모델을 제시하여 구단별 영입 최적화를 연구할 예정이다. 이러한 데이터 분석이 스포츠 시장에 직접적으로 다양한 가치 창출에 기여할 수 것을 기대한다.

ACKNOWLEDGEMENT

본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 SW중심대학지원사업의 연구결과로 수행되었음(20170001000021001)

REFERENCES

[1] Hong, J. S., Kim, J. Y. and Sin, D. S., "Alternative hitting ability index for KBO," Journal of the Korean Data & Information Science Society, Vol. 27, No. 3, pp.677-687, May 2016.

[2] Kim, Y. H., Jung, J. E., "Soccer Player Scouting Model using Big Data Analysis and Game Theory," Convergent and Smart Media Systems, June 2018.

[3] Basco, D., Zimmerman, J., "Measuring defense: Entering the zones of fielding statistics," The Baseball Research Journal, Vol. 39, No. 1, 2010

[4] Park, S. H., "Performance Factors Affecting the High Annual Salary of Korean Professional Baseball Batters," The Korea Journal of Sports Science, Vol. 17, No. 2, pp.485~494, 2008.