

## Historical ranking of vault players in artistic gymnastics using PageRank algorithm

Jiwun Yoon<sup>1</sup> & Jea-Hyeon Park<sup>2\*</sup>

<sup>1</sup>Kyungpook National University & <sup>2</sup>Korea National Sport University

**[Purpose]** The purpose of this study is to calculate the ranking of vault players in artistic gymnastics by individual and by country using the PageRank algorithm. The purpose of this study is to provide basic data that can be used in gymnastics events by comparing the performances of historical vault players with those of previous Olympic competitions. **[Methods]** The data collected for this purpose is a score of 117 vault players based on the results of the Olympic final event published in the International Gymnastics Federation (FIG). For data analysis, PageRank algorithm was used for calculating the ranking of vault players using MS-Excel and NetMiner. **[Results]** The results are as follows that. First, the PageRank algorithm is possible to calculate for historical vault players' rankings. Specifically, the ranking of vault players for historical Olympic calculated by PageRank ranked as Gervasio Deferr from Spain (ESP) at 1st, Alexei Nemov from Russia (RUS) at 2nd, and Klaus Koste from Germany (GDR) at 3rd. Second, Network of vault players' ranking is separated by each generation of Olympic games. **[Conclusions]** As a conclusions, it is possible to calculate national ranking of vault games of gymnastics by using PageRank algorithm.

**Key words:** Vault of Gymnastic, Players' Ranking, PageRank, Olympic

### 서론

“어떤 선수가 최고 선수인가?” 스포츠에서 선수평가는 선수, 감독, 팬, 스포츠사업가 등에게 중요한 관심사이다. 스포츠에서 선수를 평가하는 방법은 다양하다. 경기단체에서 사전에 정한 원칙에 따라 산정하는 랭킹은 선수평가를 위해 사용하는 가장 흔한 방법 중 하나이다. 스포츠 랭킹은 국가대표 선수 선발 준거, 프로선수 연봉 산정 기준(Soren, 1999), 스포츠판에게 흥미를 유도하는 도구로 활용되기도 한다. 특히 스포츠 랭킹은 스포츠 비즈니스 분야 마케팅 도구(Motegi & Masuda,

2012)로 활용되어 스포츠 산업 발전에 기여하는 바가 적지 않다. 즉, 경기단체에서 만들어내는 정보로서 중요한 가치를 담고 있는 도구가 바로 랭킹이라는 것이다.

스포츠 랭킹을 산출하기 위해 다양한 방법들이 소개되어 왔다. 스포츠 유형에 따라 상이한 모형을 적용하고 있다. 전통적인 스포츠 랭킹 산출방법은 선수가 획득한 승패 차이를 근거로 순위(Park & Newman, 2005)를 결정하고 있다. 현재 다수 프로리그에서 순위를 결정하는데 사용하는 방법이다. 출전한 경기 중 승리경기 횟수와 패한 횟수를 비교하여 랭킹을 매기는 승패모형을 적용한다. 승패모형은 경기에 출전한 모든 선수가 동일한 수로 경기를 수행하는 리그방식 경기운영에 적합하다. 예를 들어 10경기 중 5승한 선수와 6경기 중 5승한 선수를 동일한 랭킹으로 산정할 수 없기 때문이다.

논문 투고일 : 2017. 02. 28.

논문 수정일 : 2017. 04. 11.

게재 확정일 : 2017. 04. 28.

\* 교신저자 : 박재현(jhpark@knsu.ac.kr).

모든 선수가 동일한 경기 횟수를 수행하지 않았을 경우 Bradley-Terry(BT)가 랭킹 산출을 위해 제안한 승률 모형(Bradley & Terry, 1952)을 적용할 수 있다. BT 승률모형은 경기에 참가한 선수들이 상이한 횟수로 경기를 수행했다 하더라도 이전 시합 경기전적을 가중치로 반영하여 승률을 계산함으로써 승패모형이 갖는 문제를 해결하였다. BT 승률모형은 가중치를 합리적으로 적용할 수 있다는 이유 때문에 최근 스포츠랭킹 산출을 위한 통계모형으로 각광받고 있다(Nor & Mahmud, 2016; Shev et al., 2014). 예를 들면, 상대선수가 과거 획득한 승률을 반영하여 실력이 우수한 선수에게 승리한 경우와 실력이 열등한 선수에게 승리한 경우 최종 승률을 다르게 반영한다는 내용이 합리적이다. 그러나 BT 승률모형은 일대일 대응 경기(탁구, 테니스, 레슬링 등)에 한하여 승률계산이 가능하다는 단점을 가지고 있다. 상대 선수의 과거 승률을 반영하려면 대상 선수가 있어야 하기 때문이다. 사이클 경주, 쇼트트랙, 체조와 같은 종목은 여러 선수가 경쟁은 하지만 일대일로 대응하여 경기를 수행하지는 않는다. '다중경쟁자'가 출전하는 종목이다. '다중경쟁자' 출전 종목에서 선수 랭킹 산출에 BT 승률모형은 적합하지 않다(Gakis et al., 2016)는 주장이 일반적이다.

Gakis et al. (2016)은 다중경쟁자가 출전하는 사이클 경기에서 네트워크 이론에 기반을 둔 PageRank 알고리즘을 언급한 바 있다. PageRank는 BT 모형이 갖고 있는 장점인 이전 시합에서 획득한 선수 간 전적을 가중치로 반영한다는 동일한 논리를 가지고 있기 때문에 (Motegi & Masuda, 2012) 합리적이다. 특히 네트워크 이론에 기반을 두고 있기 때문에 무한정 많은 선수를 대상으로 랭킹을 산출한다 할지라도 이론적인 제한이 없다. 이에 PageRank는 다중경쟁자가 출전하는 스포츠에서 현실을 반영한 선수랭킹 산출 모형이라고 판단하여 다수 연구자가 스포츠분야에 활용 가능성을 주장하였다(Lazova & Basnarkov, 2015; Motegi & Masuda, 2012; Park & Newman, 2005).

체조경기는 다중경쟁자가 출전하는 종목에 해당한다. 여러 선수가 출전하여 경쟁하지만 선수들 간 일대일로 경쟁하여 승패를 만들어내지 않기 때문이다. 현재 국제 체조연맹(FIG)에서는 '포인트 시스템'으로 랭킹을 산출

하고 있다(FIG, 2016). 포인트 시스템은 대회에 포인트를 부여한 후, 대회에 출전한 선수들이 포인트를 획득하는 방식이다. 많은 포인트를 획득한 선수가 상위에 랭크 한다. 예컨대 XX대회 1위 선수에게는 50점, 2위 선수에게는 30점, 3위 선수에게는 10점 그리고 YY대회 1위 선수에게는 100점, 2위 선수에게는 70점, 3위 선수에게는 30점 등 포인트를 부여했을 때, 어떤 선수가 XX대회에 출전하여 1위 그리고 YY대회 출전하여 3위 하였다면 이 선수의 랭킹 포인트는 80점이 된다.

포인트 시스템은 대회 포인트를 결정하는데 있어 주관성과 임의성을 배제하지 못하고 있다. 대회에 출전하는 선수의 경기력을 임의로 조정할 수 없기 때문에 대회 포인트에는 동간성이 없다. 특히 FIG는 종목별 결승 경기에 진출한 선수 12명만을 대상으로 랭킹 포인트를 부여하기 때문에 종목별 결승에 진출하지 못한 대다수 선수들은 모두 0점 포인트를 갖게 된다. 13위 이하 선수 능력을 모두 0점으로 평가하는 불합리성을 갖고 있다. 포인트 시스템으로 운영하는 FIG 랭킹은 1년에 2회 개최되는 월드컵 경기 순위에 따라 차등 포인트를 부여한 후 선수들이 4년 동안 획득한 포인트를 합산하여 랭킹을 매긴다(FIG, 2016). 따라서 FIG 랭킹은 4년 동안 경기기록을 합산하여 순위를 매기기 때문에 즉각적으로 정보를 제공하지 못한다는 단점이 있다.

네트워크 이론에 기반을 두고 있는 PageRank는 FIG 랭킹 시스템이 가지고 있는 문제를 체계적으로 보완할 수 있다. PageRank는 네트워크 이론에 기반을 두고 있기 때문에 랭킹을 산출하는 선수 수에 제한이 없다. 예컨대 한번이라도 대회에 출전한 기록이 있는 선수라면 랭킹 산출이 가능하다는 뜻이다. 그 이유는 PageRank 알고리즘이 직접 경기를 수행하지 않았던 선수라고 할지라도 다른 선수와 수행한 경기의 상대결과를 근거로 서열을 매길 수 있기 때문이다. 또한 PageRank는 대회 등급에 따른 포인트를 임의적으로 매겨야 할 이유가 없다. 어떤 대회는 100점을 부여하고 다른 대회에는 20점을 주관적으로 부여해야 하는 문제를 해결할 수 있다. PageRank는 경기 결과만을 반영하여 순위를 산출하기 때문에 대회에 임의적 포인트를 부여할 이유가 없다.

이 연구는 PageRank 알고리즘이 갖고 있는 네트워크 분석 특성을 활용한다면 과거 선수와 현재 선수 간 실

력을 가늠해 볼 수 있을까라는 의문에서 계획되었다. 예컨대, “1980년대와 2010년대 도마 선수 중 누가 더 우수할까?” 혹은 “어떤 나라 도마 경기력이 가장 우수할까?”의 질문에 PageRank 알고리즘은 답을 줄 수 있을 것으로 판단(Radicchi, 2011)하였다.

이 연구는 PageRank 알고리즘을 활용하여 체조 도마경기 역대 순위를 산출하는 것이 목적이다. 구체적으로 이 연구는 PageRank 알고리즘을 이용하여 체조도마 경기에 적용하여 개인별 그리고 국가별 순위를 산출하여 체조 랭킹의 산출방법 개선을 위한 기초자료를 제공할 것이다.

## 랭킹모형

### PageRank 랭킹 모형

PageRank는 구글에서 링크 연결구조를 갖고 있는 웹 페이지의 상대적 중요도에 따라 가중치를 부여하여 웹 페이지 순위를 매기는 방법으로 처음 소개되었다 (Brin et al., 1998). PageRank 알고리즘은 중요한 웹페이지는 중요하지 않은 웹페이지보다 다른 페이지에 미치는 영향력이 더 크다고 전제한다. 예컨대 다른 웹페이지로부터 링크 받는 빈도가 높을수록 중요한 정보를 담은 웹페이지로 보기 때문에(Gyöngyi et al., 2004), 중요도가 높은 웹페이지는 그렇지 않은 페이지보다 PageRank 순위가 높다.

PageRank 알고리즘은 단지 검색결과에 대한 순위를 부여하는데 한정되어 있지 않다. 최근에는 축구리그 랭킹, 학술지 인용평가, 텍스트 의미 분석 등 전 학문분야에서 PageRank를 응용하여 사용하고 있다. 기본적으로 PageRank는 네트워크 이론에 근거를 두고 있다. 네트워크는 개체 간 거래관계로 표현할 수 있다. 거래관계를 화살표로 표현하면 화살표를 주는 개체와 화살표를 받는 개체로 구분해 볼 수 있다. 링크를 받는 웹페이지와 링크를 주는 웹페이지, 인용을 하는 학술논문과 인용을 받는 학술논문, 승리하는(승리를 받는) 선수와 승리하지 못하는(승리를 주는) 선수 등 대부분 개체관계는

종속적이다.

스포츠 경기는 선수(개체) 간 승패 네트워크로 연결된 상호 종속적 관계에 있다. 예컨대 승리선수는 패배선수로부터 화살표를 받고 반대라면 패배선수는 승리선수에게 화살표를 보낸다. 사회연결망(social network)에서 사용하는 팔로워(follower)와 팔로잉(following) 개념과 동일하다. 네트워크 이론에서 이 관계를 노드(Node)와 링크(link)로 설명한다. 이때, 노드는 선수를 그리고 링크는 선수 간의 승패관계로 의미를 부여한다. 승자가 화살표를 하나 받을 때, 패자는 화살표를 하나 보낸다.

〈Fig. 1〉은 시합에 출전한 다중경쟁 선수(A, B, C, X, Y, Z)들이 네 경기(contest 1부터 contest 4)에 출전하여 획득한 순위이다. Contest 1에서 A는 B와 C보다 순위가 앞서 화살표를 두 개 받았고(2 follower를 가짐), C는 A와 B보다 순위가 뒤처져 받은 화살표가 없다. 특히, A는 Contest 1에서 2명에게 앞서 화살표 2개와 Contest 3에서 2명에게 앞서 화살표 2개를 받음으로써 전체 받은 화살표는 4개이다. Contest 1부터 Contest 4까지 다중경쟁 선수 간 네트워크 관계는 〈Fig. 1〉과 같다.

ranking Contest	1st	2nd	3rd
Contest 1	A	B	C
Contest 2	X	Y	Z
Contest 3	A	B	Y
Contest 4	C	X	Z

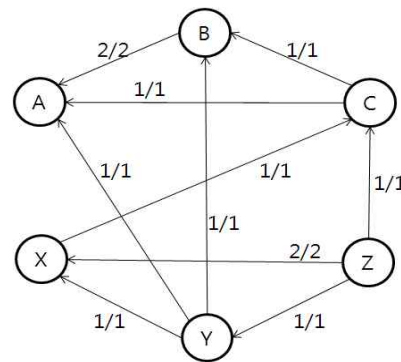


Fig. 1. Example of contest results

## 국제체조연맹 랭킹 모형

국제체조연맹(FIG)은 올림픽 주기 동안 개최되는 여덟 개 월드컵 경기를 종합하여 랭킹을 산출한다. 종목별 결승 진출한 1위부터 12위까지 선수를 대상으로 랭킹 포인트를 부여한다. 출전 횟수와 관계없이 선수들이 획득한 랭킹 포인트 중 상위 3개 점수를 합산하여 최종 랭킹을 산출한다. FIG 랭킹은 올림픽 출전권을 부여하는 목적으로 활용되고 있다(FIG, 2016).

FIG 랭킹 방법을 설명한 예는 <Table 1>에 제시하였다. 네 명 선수(A, B, C, D)들이 여러 차례 치러진 월드컵(Contest 1부터 Contest 8) 도마경기에 출전하여 획득한 포인트를 예로 설명하였다. A의 점수를 보자. A는 Contest 2와 Contest 5를 제외한 여섯 차례 경기에 출전하였다. FIG는 A가 획득한 포인트 중 상위 세 개 점수를 합하여 최종 랭킹점수를 산출하는 원칙을 갖고 있다. 따라서 A가 획득한 상위 세 개 점수는 30점(Contest 1), 25점(Contest 7), 18점(Contest 3)으로서 최종 랭킹 점수는 73점이다. 이러한 방법으로 최종 랭킹 점수를 계산하면 B는 57점, C는 37점, D는 51점으로 최종 랭킹은 A, B, D, C 순서로 매겨진다.

FIG 랭킹은 다음과 같은 제한을 가지고 있다. 첫 번째는 대회 중요도에 따라 부여된 랭킹 포인트는 일부 사람들로부터 결정된 주관적 의견에 불과할 수 있다. 대회에 부여한 랭킹 포인트에 대한 합리성이 의심되는 부분이다. 두 번째는 FIG에서 주관하는 국제대회는 올림픽, 세계선수권대회, 대륙 간 대회 등 다수 존재함에도 월드컵 대회 결승경기만으로 랭킹 포인트를 한정하고 있기 때문에 발생하는 정보 제한성이다. 사년에 한번 랭킹을 산출하여 발표하기 때문에 발표 당시에 이미 정보로서 가치는 상실되어 있다.

## 연구방법

### 분석대상자료

이 연구는 PageRank 알고리즘을 활용하여 체조 도마 선수를 대상으로 개인별 그리고 국가별 역대 순위를 산출하는 것이 목적이다. “체조 도마경기에서 역대 어떤 선수가 가장 우수한 선수일까?” 그리고 “어느 나라가 도마를 가장 잘할까?”에 대한 의문에서 시작하였다. 이 목적을 달성하기 위해 이 연구는 올림픽 경기로 수집자료를 한정하였다. 따라서 올림픽에 출전한 선수 중에서 “역대 최고 도마 선수는 누구인가?”가 보다 정확한 질문이다. 이를 위해 Fig. 홈페이지에 게시된 1896년 제1회 대회부터 2016 제31회 대회까지 도마 경기결과를 수집하였다. 이 연구에서 수집한 자료는 올림픽 도마 결승경기에 출전한 선수로 한정하였다. Fig. 홈페이지에 제공하고 있는 자료의 한계 때문에 수집할 수 있는 자료가 한정되었다. 하지만 이론적으로는 자료 수에 제한을 받지 않기 때문에 예선경기를 제외할 이유는 없다.

Fig. 홈페이지 게시 자료 중 선수이름, 국가, 출전대회 등과 같은 기본 정보와 함께 각 선수들이 도마 경기에서 획득하였던 원점수를 수집하였다. 올림픽 1회 대회와 2회 대회는 도마에서 획득한 원점수가 제공되지 않기 때문에 이 연구에서 분석에 실제 이용된 자료는 3회 대회부터 31회 대회를 범위로 하였다. 도마 경기 결승에 진출한 선수 117명이 획득한 기록을 수집하여 PageRank 알고리즘으로 분석하였다.

이 연구에 참여한 연구대상 자료에 관한 일반적인 특성은 다음의 <Table 2>와 같다. 17개 올림픽 대회 결승경기에 진출한 도마 선수 117명이다.

Table 1. Example of FIG ranking system

2012-2016 Vault of Gymnastic World Championship										
player	Contest 1	Contest 2	Contest 3	Contest 4	Contest 5	Contest 6	Contest 7	Contest 8	rankin point	ranking
A	30	-	18	8	-	12	25	12	73	1
B	20	5	-	25	8	8	-	12	57	2
C	5	6	14	5	5	15	8	8	37	4
D	8	18	25	-	-	-	-	-	51	3

Table 2. Characteristics of collected data

no.	olympic	Number of player	no.	olympic	Number of player
1	3rd	3players	10	24th	8players
2	15th	6players	11	25th	8players
3	16th	6players	12	26th	8players
4	17th	6players	13	27th	8players
5	18th	6players	14	28th	8players
6	20th	6players	15	29th	8players
7	21st	6players	16	30th	8players
8	22nd	6players	17	31st	8players
9	23rd	8players	총	17 games	117players

## 자료분석

이 연구는 PageRank 알고리즘을 활용하여 체조 도마 경기에 출전한 선수를 대상으로 역대 순위를 산출하는 것이 목적이다. “역대 체조 도마 선수 중에서 누가 가장 우수한 선수일까?” 이 질문에 답하기 위해서 올림픽 대회 도마 결승경기에 진출한 도마선수로 자료를 한정하여 수집하였고, 수집된 자료를 MS-Excel과 NetMiner 프로그램을 사용하여 분석하였다. MS-Excel은 도마경기 출전한 선수를 대상으로 수집한 원천 기록을 행렬로 정리하는 목적으로 사용하였고 NetMiner는 행렬로 정리된 체조 도마 선수 기록을 활용하여 PageRank 값을 산출하는 목적으로 사용하였다.

PageRank 알고리즘을 적용하기 위한 구체적인 자료 처리 방법은 다음과 같다. 첫째, 역대 올림픽 경기기록을 수집한 후 각 경기 결과를 컴퓨터에 개별 입력하였다. 이때 선수들이 획득한 원점수를 기록하였고, 해당 선수 국가 등을 범주화하였다. 둘째, PageRank 알고리즘 적용을 위해 도마 경기 순위를 기초로 Source와 Target으로 구분된 1모드 행렬로 입력하였다. Source와 Target으로 구분된 1모드는 선수×선수 행렬에 해당한다. <Table 3>과 같이 Source는 승자(winner) 그리고 Target는 패자(loser)로 설정하였다. 셋째, PageRank 알고리즘은 선수×선수 1모드 행렬로부터 산출된 인접행렬을 정상화하여 변환된 행렬을 산출하게 된다. 이때 덤

핑팩터(damping factor)  $d$ 는 .85로 설정하였다. 덤핑팩터  $d$ 는 구글 엔진에서 검색 중 해당 페이지에 만족하지 못하고 다른 페이지로 가는 링크를 클릭할 확률로 해석할 수 있다. 덤핑팩터가  $1-d$ 라 함은 페이지를 잘못 찾아 다른 페이지로 넘어가는 비율로 볼 수 있다. 이 연구에서는 도마경기 실력이 낮음에도 앞선 순위에 잘못 배정되는 확률로 본다. PageRank는 일반적으로  $1-d$ 를 .15로 설정하고 있다(Boldi, Santini, & Vigna, 2005). PageRank 관련 보다 자세한 설명은 Page, Brin, Motwani, & Winograd(1997)를 참고 바란다.

이 연구에서 도마 선수들 간 실력 가중치는 선수들이 도마 경기에서 획득한 원점수 차이로 정의하였다. 세 명의 선수(A, x, D)가 출전한 도마 경기에서 A는 11.2점, x는 10.1점, D는 9.8점을 획득하였다. 이때, A와 x 간 관계에서 가중치는 0.1 그리고 A와 D 간 관계에서 가중치는 1.4이다. <Table 3>은 체조 도마경기에서 획득한 원점수를 근거로 산출한 NetMiner 입력 체계이다. 표에서 winner와 loser는 단순 개인 간 승패로 볼 수 없으며, 앞선 선수를 winner 그리고 뒤에 따라오는 선수를 loser로 정의하고 있다.

Table 3. Example of winner x loser matrix

contrast	winner	loser	weight*
1	A	x	0.100
2	B	y	0.050
3	C	z	0.250

\*weight is the difference between winner and loser score.

## 연구결과 및 논의

이 연구에서는 PageRank 알고리즘을 활용하여 체조 도마경기의 역대 순위를 개인별 국가별로 산출함으로써 역대 체조 도마선수들의 기량을 비교하고 체조 랭킹의 산출방법 개선을 위한 기초자료를 제공하는 것이 목적이다. 이를 위하여 PageRank 알고리즘에 기반 한 체조 랭킹을 개인별 국가별로 산출하였으며 체조 도마 경기의 역대 순위에 대한 결과는 다음과 같다.

Table 4. Historical ranking of vault players TOP 20

no.	name of player	country	PR	ranking	Number of olympic game & ranking
1	Gervasio Deferr	ESP	0.025	1	27thgame:1st / 28thgame:1st
2	Alexei Nemov	RUS	0.024	2	26thgame:1st / 27thgame:4th
3	Klaus Koste	GDR	0.017	3	20thgame:1st
4	Alexei Voropaev	RUS	0.016	4	26thgame:6th / 27thgame:2nd
5	Nikoai Andrianov	URS	0.015	5	20rdgame:3rd / 21stgame:1st / 22ndgame:1st
6	Lou Yun	CHN	0.015	5	23rdgame:1st / 24thgame:1st
7	Scherbo Vitaly	EUN	0.014	7	25thgame:1st / 26thgame:3rd
8	Leszek Blanik	POL	0.013	8	27thgame:3rd / 29thgame:1st
9	Yeo Hong-Chul	KOR	0.012	9	26thgame:2nd
10	Viktor Klimenko	URS	0.009	10	20thgame:2nd
11	Ri Se Gwang	PRK	0.008	11	31stgame:1st
12	Yang Hak-Sun	KOR	0.008	11	30thgame:1st
13	Denis Ablyazin	RUS	0.008	11	31stgame:2nd
14	Haruhiro Yamashita	JPN	0.007	14	18thgame:1st
15	Marian Dragulescu	ROM	0.007	14	28thgame:3rd / 29thgame:4th / 31sgame:4th
16	Li Xiaoshung	CHN	0.006	16	25thgame:4th / 26thgame:5th / 28thgame:7th
17	Kroll Sylvio	GDR	0.006	16	24thgame:2nd / 25thgame:6th
18	Viktor Chukarin	URS	0.006	16	15thgame:1st
19	Misutin Grigori	EUN	0.006	16	25thgame:2nd
20	Masao Takemoto	JPN	0.005	20	15thgame:2nd / 16thgame:6th

\*PR: PageRank value

### 체조도마 경기 역대 올림픽 선수랭킹

〈Table 4〉에 의하면 제 3회 올림픽부터 31회 올림픽까지 도마경기에 출전했던 선수들 중 최고의 선수는 27회 28회 올림픽에 출전하여 두 대회에서 모두 금메달을 획득한 스페인(ESP)의 Gervasio Deferr(PR=0.025) 선수인 것으로 확인되었다. 2위는 러시아(RUS)의 Alexei Nemov (PR=0.024) 선수, 3위는 독일(GDR)의 Klaus Koste(PR=0.017) 선수로 순위가 매겨졌다.

Gervasio Deferr 선수는 PageRank 랭킹모형에서 우수한 체조 선수로부터 상대적으로 승리한 횟수가 빈번하기 때문에 많은 화살표를 받고 있는 것이 1위에 등극한 이유로 설명할 수 있다. 체조 도마 경기의 올림픽 역대 최고 선수는 스페인의 Deferr 선수라고 말할 수 있다.

한편, 우리나라의 역대 도마 순위를 살펴보면, 26회 올림픽에서 2위를 획득했던 여홍철(PR=0.012) 선수가 9위에 랭크되었고 30회 올림픽에서 우리나라 역사상 첫 금메달을 획득했던 양학선(PR=0.008) 선수가 북한의 리세광(PR=0.008) 선수 그리고 러시아의 Denis Ablyazin (PR=0.008) 선수와 함께 11위에 랭크되는 것으로 나타났다.

흥미로운 것은 Gervasio Deferr 선수가 “왜 역대 최고의 도마 선수인가?”하는 질문이다. Deferr선수의 올림픽 출전횟수와 성적은 올림픽 2회 출전 및 각각의 대회에서 우승을 차지한 것을 확인할 수 있다. 최고의 도마 선수라는 사실은 분명함에도 Deferr 선수와 경기성적이 비슷한 Nikoai Andrianov(URS) 선수의 사례와 비교해보면 이러한 의문은 자연스러운 것이다. Andrainoy

선수는 올림픽 대회에 3번 출전하여 2번의 우승과 1번의 3위를 차지하였다. 성적으로만 놓고 보면 Deferr 선수에 뒤지지 않는다. 오히려 더 우수한 선수라고 판단하더라도 손색이 없을 것이다. 또한 Lou Yun (CHA) 선수 역시 Deferr 선수와 동일한 2회 출전에 2개의 금메달을 획득하였다. 올림픽 성적만으로 역대 최고 선수를 비교한다면 우열을 가리기 쉽지 않다.

PageRank는 이러한 문제를 합리적으로 해결할 수 있다. 수리적으로는 복잡한 모형과 분석절차를 가지고 있지만, Deferr 선수가 Andrianov와 Yun 선수와 직접 경쟁한 사례는 없더라도, 16위 Li Xiaoshung (CHA)과 17위 Kroll Sylvio (GDR) 선수가 출전하였던 24회와 25회 올림픽 대회를 매개로하여 상대적 순위를 매기게 된다. 특히 <Fig. 2>의 독립된 Andrianov 선수중심네트워크는 20회, 21회 그리고 22회 올림픽 결승진출자가 다른 회수 올림픽 대회와 링크하고 있지 않는 독립 세대이기 때문에 중심 네트워크에서 다소 밀려있는 것으로 해석할 수 있다. 결승경기 뿐 아니라 예선 경기를 모두 포함한다면 보다 타당한 역대 순위를 산출할 수 있을 것이다. 이 결과는 PageRank에서는 순위를 산정함에 있어 상대적 중요도에 따라 가중치를 부여하여 순위를 매기는 방식을 취하고 있다는 것을 확인한 것이다(Brin et al., 1998).

PageRank 시스템은 출전 횟수뿐만 아니라 잘하는 선수를 이겼을 때 더 많은 점수를 얻을 수 있다. 체조 경기에 많은 출전 횟수 보다는 기량이 좋은 선수에게 승리하는 것이 더 높은 순위에 랭크 될 수 있음을 시사하는 것으로 해석할 수 있다.

한편, 역대 체조도마 랭킹 상위 20위를 살펴보면 20회 올림픽 이후 올림픽에 출전한 선수들이 상위에 랭크 되는 것을 확인할 수 있었다. 과거에 비하여 선수들의 경기력이 향상된 것으로 추측해 볼 수 있다. 앞서 설명하였듯이 PageRank 알고리즘을 활용하여 랭킹을 산출하는데 있어 가중치를 활용한다. 경기기량이 높은 선수가 경기 기량이 낮은 선수보다 랭킹에 미치는 영향력이 더 크다는 가정 하에 기량이 높은 선수에게 가중치를 부여함으로써 더 많은 점수를 받게 한다는 논리이다 (Gyöngyi et al., 2004).

이와 같은 논리에 따라 해석을 해보면 20회 올림픽

이전의 출전 선수들 간의 기량의 차이는 20회 이후 올림픽 출전 선수들 간의 기량차이보다 적다는 것으로 해석할 수 있다. 체조 경기의 기술 수준과 연기내용이 세대가 지남에 따라 빠르게 변화하고 있다는 선행연구(Kim & Kim, 2011)가 이를 뒷받침 해주고 있다.

한편, PageRank에 의한 체조 역대 순위의 네트워크는 어떠한가를 확인하기 위하여 PageRank 체조랭킹 연결망을 시각화 하였다. 체조랭킹 연결망은 <Fig. 2>와 같다.

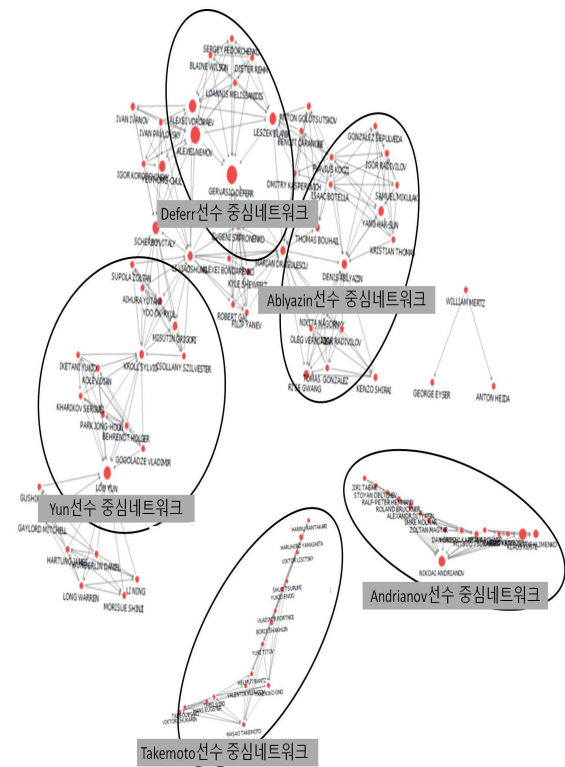


Fig. 2. Network of historical ranking for player's

<Fig. 2>에 의하면 역대 올림픽 체조도마 출전 선수 117명에 대한 PageRank 체조랭킹의 선수연결망은 크게 5개로 구분되는 것을 확인할 수 있다. 즉 1위를 중심으로 이루어지는 'Deferr선수 중심네트워크', 13위를 중심으로 이루어지는 'Abylazin선수 중심네트워크', 5위를 중심으로 이루어지는 'Yun선수 중심네트워크'와 'Andrianov선수 중심네트워크' 그리고 20위를 중심으



로 이루어지는 'Takemoto선수 중심네트워크'로 구분되고 있다. 이는 동일한 세대에 체조도마경기에 출전한 선수들 간 연결망이 구성되고 있는 것으로 해석할 수 있다. 특히 Deferr 선수, Ablyazin선수 그리고 Yun선수 중심네트워크는 도마경기에 출전한 선수들의 세대는 다르지만 서로 비슷한 세대로 묶여 있으며 Andrianov선수와 Takemoto선수 중심네트워크는 완전 고립된 네트워크의 세대로 해석할 수 있다. 네트워크는 선수 간 승패 결과를 반영하여 화살표의 연결이 이루어지는데, Andrianov선수 중심네트워크와 Takemoto선수 중심네트워크는 해당 시대 올림픽 결승진출자만으로 네트워크를 구성하고 있기 때문에 고립된 네트워크를 형성하게 되었다.

### 체조도마 경기 역대 올림픽 국가랭킹

이 연구에서는 PageRank 알고리즘을 활용하여 체조도마경기의 국가별 순위를 산출하였다. 결과는 <Table 5> 및 <Fig. 3>과 같다.

<Table 5>에 의하면 역대올림픽 도마경기의 국가랭킹 1위는 러시아(PR=0.104)로 나타났으며, 2위 스페인(PR=0.073). 3위 한국(PR=0.071) 순으로 나타났다. 이는 러시아 선수들이 역대 올림픽 체조도마 경기에서 많은 선수들이 올림픽 결승에서 승리하였음을 의미한다. 이와 같은 결과는 <Fig. 3>을 통해서도 확인가능하다.

<Fig. 3>을 보면 러시아의 노드가 가장 크며 많은 링크를 갖고 있는 것을 확인할 수 있다. 이때 네트워크는 노드(국가)와 링크(승, 패를 주고받는 결과)의 관계를 구조화 한 것을 의미하는데(Yoon & Park, 2015) 노드가 클수록 다른 선수로부터 승리를 통한 화살표를 많이 받은 것으로 해석할 수 있다.

한편, 우리나라 체조 도마종목은 24회 올림픽에서 3위, 25회 올림픽에서 3위, 26회 올림픽에서 2위, 30회 올림픽에서 1위의 성적을 거두었다. 여홍철 선수의 여1, 여2 기술과 양학선 선수의 양1, 양2 기술까지 다양한 신기술 개발로 세계 기계체조 도마 종목에서 강국임을 입증(Song et al, 2014; Han, 2013) 한 바 있다. 이 연구 결과는 우리나라 체조 도마종목이 강국이라는 그동안의 주관적 인식을 객관적 데이터로 증명하였다는 사실에

의미가 적지 않다. 우리나라 도마 경기력은 역대 올림픽 성적 기준으로 세계 3위로 평가할 수 있다. 이는 PageRank를 활용한 체조랭킹 산출을 통해 국가별 순위 비교를 객관적으로 평가할 수 있음을 시사한다.

Table 5. Historical ranking of vault of country TOP 10

no.	County	PR	Ranking
1	RUS	0.104	1
2	ESP	0.073	2
3	KOR	0.071	3
4	EUN	0.064	4
5	CHN	0.055	5
6	JPN	0.042	5
7	ROM	0.041	7
8	URS	0.039	8
9	GDR	0.037	9
10	POL	0.036	10

\*PR: PageRank value

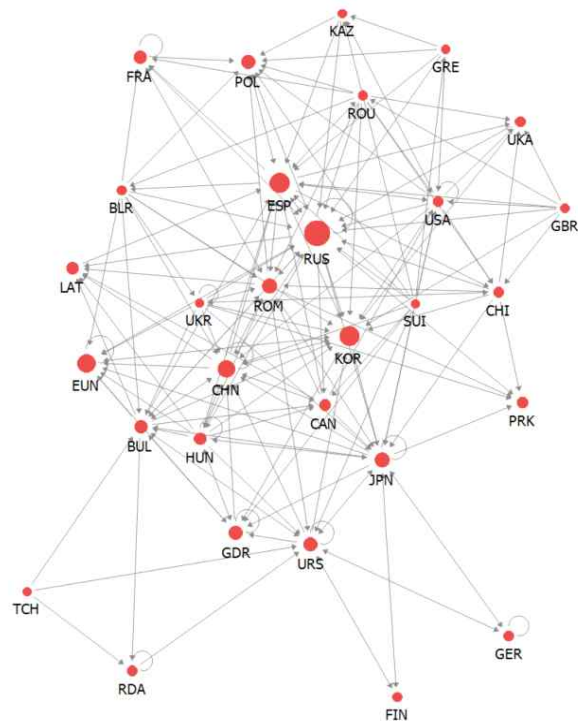


Fig. 3. Network of historical ranking for countries



## 결론 및 제언

체조경기에서 랭킹의 산출은 선수개인의 발전과 스포츠발전을 위해 중요한 요소 중 하나이다. 랭킹을 통해 선수의 경기력을 평가하여 각종 상황에 반영할 수 있으며(Soren, 1999) 랭킹을 통한 선수 간 국가 간 비교를 통해 스포츠팬으로부터 정보와 재미를 더해 줄 수 있기 때문이다. 그러나 그동안 국제체조의 랭킹시스템으로는 전체 선수를 대상으로 한 랭킹 산출의 어려움뿐만 아니라 랭킹관련 다양한 정보제공에 제한이 있어 왔다. 이에 이 연구에서는 PageRank 방법을 활용하여 체조랭킹시스템의 제한점을 극복해보고자 하였다.

따라서 이 연구에서는 PageRank 알고리즘을 활용하여 체조 도마경기의 역대 순위를 개인별 국가별로 산출함으로써 역대 체조 선수들의 기량을 비교하고 체조 랭킹의 산출방법 개선을 위한 기초자료를 제공하는 것이 목적이었다. 이 연구의 얻은 결론은 다음과 같다.

첫째, PageRank 알고리즘을 활용하여 체조 도마경기의 선수랭킹 산출이 가능하다. 구체적으로 PageRank로 산출한 역대올림픽 도마 선수 순위는 스페인(ESP)의 Gervasio Deferr 선수이며, 2위는 러시아(RUS)의 Alexei Nemov 선수, 3위는 독일(GDR)의 Klaus Koste 선수 순이다. 둘째, PageRank 알고리즘을 활용하여 산출한 체조 도마랭킹의 선수연결망은 올림픽 출전 세대별로 구성된다. 셋째, PageRank 알고리즘을 활용하여 체조 도마경기의 국가랭킹 산출이 가능하다. 구체적으로 PageRank로 산출한 역대올림픽 도마 국가순위는 1위 러시아, 2위 스페인, 3위 한국 순이다.

이 연구는 PageRank 알고리즘을 활용함으로써 체조 도마 경기의 역대올림픽 선수에 대한 랭킹 산출하여 역대 최고 도마 선수는 누구인가? 역대 최고 도마 국가랭킹 1위는 누구인가에 대한 궁금증을 확인할 수 있었다는 점에서 의의가 있다. 이 연구를 진행하면서 나타난 몇 가지 제한점을 토대로 다음과 같은 제언을 하고자 한다.

첫째, 이 연구가 체조도마 경기 역대 순위를 산출하면서 역대올림픽에 한정하였다는 점이다. 물론 올림픽은 체조뿐만 아니라 세계 스포츠인들의 최고 기량을 가리는 대회라는 점은 부정할 수 없다. 그러나 올림픽이라는 대회에 한정함으로써 모든 체조선수들을 포함하고 있지

는 못하다. 이에 후속연구에서는 올림픽 뿐 만 아니라 세계선수권대회, 아시아경기대회 등 다양한 주요대회 경기 결과를 확인함으로써 보다 흥미로운 정보를 제공할 것을 기대한다.

둘째, 이 연구에서는 체조 랭킹을 산출하면서 도마 종목에 한정하였다는 제한이 따른다. 체조 종목은 남자 6종목, 여자 4종목으로 구성되어 있으며 종목별대회, 단체전, 개인전을 포함하며 총 12경기를 치루는 종목이다. 따라서 후속 연구에서는 다양한 종목의 랭킹 산출을 통해 의미 있는 정보를 제공하기를 제안한다.

셋째, 이 연구는 체조 랭킹을 산출하는데 있어 PageRank 알고리즘을 활용하여 그 가능성을 확인하였으나 랭킹에 대한 타당성을 확인하지 못하였다는 제한이 따른다. 그러나 랭킹시스템 개선을 위한 기초자료를 제공했다는 점에서 그 의의가 크다. 향후 후속연구에서는 새로운 체조랭킹 시스템 개발과 타당도를 검증할 수 있는 연구들이 진행되기를 기대한다.

PageRank 알고리즘이 이론적으로 폭넓은 확장성을 가지고 있는 랭킹모형이라는 사실은 분명하지만, 스포츠 랭킹을 산출할 때 적용범위와 수리적 모형에 대한 논의는 이제 시작단계에 있다. 따라서 적지 않은 연구자들이 PageRank 알고리즘에 관심을 두고 있는 것은 사실이지만 스포츠현장에서 적용하기 위한 더 많은 증거가 요구되는 것도 사실이다. 앞으로 더 많은 후속연구를 통하여 보다 합리적인 스포츠랭킹 시스템이 자리잡을 수 있기를 기대한다.

## 참고문헌

- Boldi, P., Santini, M., & Vigna, S. (2005). PageRank as a function of the damping factor. In *Proceedings of the 14th international conference on World Wide Web* (pp. 557-566). ACM.
- Kim, Y. J. & Kim, H. Y. (2011). Strategic Analysis for Vault in 2010 Women's Gymnastics Artistic World Championships. *The Korean Journal of Sport*, 9(3), 193-203.
- Song, J. H., Park, J. C., Lee, Y. S., Back, J. H., Oh, S. B., Park, J. H., & Lee, G. S. (2014). Comparative Analysis of Determinant Factor of Yeo2 Motion in Vault. *Korean*

- Journal of Physical Education*, 53(1), 463-474.
- Yoon, J. & Park, J. H. (2015). Semantic Network Analysis for Content Analyzing of Qualitative Research in Adapted Physical Activity. *Korean Journal of Physical Education*, 54(5), 877-889.
- Han, Y. S. (2012). Elements, Line Deductions and Landing Success Rate of the Vault Final in the London Olympics 2012. *The Korea Journal of Sports Science*, 22(1), 1069-1076.
- Bradly, R. A., & Terry, M. E. (1952). Rank analysis of incomplete block designs. *The method of paired comparisons*, *Bipmetika*. 42, 450-470.
- FIG (2016). *Artistic Men's Rules*. <http://www.fig-gymnastics.com/site/rules/disciplines/art>. 2016년 10월 10일 검색.
- Lazova, V., & Basnarkov, L. (2015). *PageRank Approach to Ranking National Football Teams*. CoRR abs/1503.01331.
- Page, L., Brin, S., Motwani, R., & Winograd, T. (1997). *PageRank: Bringing order to the web* (Vol. 72). Stanford Digital Libraries Working Paper.
- Park, J. & Newman, M. E. J. (2005). A network-based ranking system for US college football. *Journal of Statistical Mechanics: Theory and Experiment*. 10014-10027.
- Motegi, S., & Masuda, N. (2012). A network-based dynamical ranking system for competitive sports. *Scientific Reports*, 2, 904.
- Nor, S. B. M., & Mahmud, Z. (2016, October). Teams ranking of Malaysia Super League using Bayesian expectation maximization for Generalized Bradley Terry Model. In N. Aziz, J. Zulkepli, N. Ahmad, & J. Jamil (Eds.), *AIP Conference Proceedings* (Vol. 1782, No. 1, p. 050011). AIP Publishing.
- Brin, S., Page L., Motwami, R. & Winograd, T.(1998). The PageRank Citation Ranking: Bringing Order to the Web, *Technical report, Computer Science Department*, Stanford University, Stanford, CA.
- Gakis, K., Pardalos, P., Choi, C.H., & Park, J.H. (2016). A probabilistic model for multi-contestant races. *Athens Journal of Sports*, 3(2), 111-118.
- Gyöngyi, Z., Garcia-Molina, H., and Pedersen J.(2004). Combating web spam with trustrank. *Proceedings of the Thirtieth international conference on Very large data bases*. 30, 576-587.
- Radicchi, F.(2011). Who Is the Best Player Ever? A Complex Network Analysis of the History of Professional Tennis. *PLoS ONE*, 6(2). e17249. doi:10.1371/journal.pone.0017249
- Shev, A., Fujii, K., Hsieh, F., & McCowan, B. (2014). Systemic testing on Bradley-Terry Model against nonlinear ranking hierarchy. *PloS one*, 9(12), e115367.
- Soren, P. S.(1999). *An overview of some methods for ranking sports*. Knoxville: TN.

## PageRank 알고리즘을 활용한 체조도마 선수 역대 랭킹

윤지운(경북대학교), 박재현(한국체육대학교)

**[목적]** 이 연구는 PageRank 알고리즘을 활용하여 체조 도마경기의 역대 순위를 개인별 그리고 국가별로 산출하여 산출방법 개선을 위한 기초자료를 제공하는 것이 목적이었다. **[방법]** 이 목적을 위하여 활용된 자료원은 국제체조연맹(FIG)에 게시되어 있는 올림픽 17회(3회부터 31회)의 도마 결승전 경기결과이며 117명의 선수들의 점수이다. 자료분석을 위하여 MS-Excel 프로그램과 NetMiner 4.0 프로그램을 활용하여 PageRank 체조랭킹을 산출하였다. **[결과]** 연구결과는 다음과 같다. 첫째, PageRank 알고리즘을 활용하여 체조 도마경기의 선수랭킹 산출이 가능하다. 구체적으로 PageRank로 산출한 역대올림픽 도마 선수순위는 스페인(ESP)의 Gervasio Deferr선수이며, 2위는 러시아(RUS)의 Alexei Nemov선수, 3위는 독일(GDR)의 Klaus Koste선수 순이다. 둘째, PageRank 알고리즘을 활용하여 산출한 체조 도마랭킹의 선수연결망은 올림픽 출전 세대별로 구성된다. 셋째, PageRank 알고리즘을 활용하여 체조 도마경기의 국가랭킹 산출이 가능하다. 구체적으로 PageRank로 산출한 역대올림픽 도마 국가순위는 1위 러시아, 2위 스페인, 3위 한국 순이다. **[결론]** 이 연구의 결과는 체조랭킹시스템 개선을 위한 기초자료로 활용할 수 있고 체조경기에서 다양한 경기정보를 생성하는데 기여할 것으로 기대한다.

**주요어:** PageRank, 역대올림픽, 체조선수, 도마랭킹