

사회네트워크분석을 통한 축구경기 분석

최승배¹ · 강창완² · 최형준³ · 강병욱⁴

¹²⁴동의대학교 데이터정보학과 · ³명지대학교 스포츠기록분석연구센터

접수 2011년 9월 25일, 수정 2011년 10월 19일, 게재확정 2011년 10월 25일

요약

사회네트워크분석은 개체들 간의 상호 정보 흐름인 사회적 구조에 대한 사회과학적·통계학적 분석이다. 본 연구에서는 한 축구경기에 대해서 사회네트워크분석을 이용하여 축구선수들 간의 패스 정보를 이용하여 (1) 어떤 선수가 팀에서 리더로서 얼마만큼 역할을 했으며, (2) 패스를 많이 해 준 선수 또는 패스를 많이 받은 선수들이 누구인지를 알아내어 경기에서 중요한 역할을 한 선수들을 탐지한다. 본 연구의 목적은 팀 내에서 각 선수들의 역할에 대한 수행 실태를 평가하고, 향후 경기에서 팀의 경기 전략을 수립하는데 기초자료로 활용하는데 있다. 본 연구에서는 먼저 포지션을 구분함이 없이 사회네트워크분석을 실시하였고, 각 포지션별 (수비수와 비수비수)로 각각 사회네트워크분석을 수행하였다. 본 연구에 대한 분석 결과는 다음과 같다. 첫째, 전체데이터를 이용하였을 때, 리더의 역할을 수행한 선수들은 이청용, 김정우, 기성용 선수이고, 서브리더는 이정수 선수였음을 알 수 있었다. 그리고 포지션별로는 수비수의 경우 리더의 역할을 수행한 선수는 이정수 선수임을 알 수 있었다. 비수비수 (미드필더와 공격수)의 경우에는 각 선수들이 부여된 포지션에서 최선을 다하였기 때문에 누구를 리더와 서브리더라고 할 것 없이 모두 훌륭한 경기를 소화하였다고 평가할 수 있다.

주요용어: 리더, 사회네트워크분석, 사회네트워크서비스, 서브리더.

1. 서론

요즘 인터넷은 대부분 가정에 보급되어 있을 정도로 보편화되어 있기 때문에 더 이상 특별한 것이 되지 못한다. 인터넷의 급속한 발전으로 정보를 얻기 위한 공간상에 제약은 없어졌으며, 사용자들은 사이버 공간상에서 서로간의 정보를 공유하고 있다. 인터넷의 발전은 사회미디어 (social media)의 등장을 가져왔으며, 이를 활용한 사회네트워크서비스 (social network service)가 활성화되고 있다. 이러한 환경들로 인하여 사회네트워크의 중요성과 활용성은 더욱 고조되고 있다. 또한 핸드폰과 통신 네트워크의 통합으로 스마트 폰이 등장함으로써 유비쿼터스 (ubiquitous) 시대에 접어들었다고 해도 과언이 아니다. 즉, 시간과 공간의 제약을 받지 않고 실시간으로 사용자들 간의 정보를 공유함으로써 일방적인 정보획득에서 양방향성 정보교류의 시대로 접어들었다.

사회네트워크분석 (social network analysis)은 사회네트워크서비스를 통하여 연결된 또는 연결되지 않은 개인 (또는 기관)으로 구성된 사회적 구조 (점과 선으로 구성된 망)에 대한 사회과학적·통계학적 분석방법이다. 여기서 연결여부는 친구/친족 관계, 금융거래, 친근감, 성관계, 신뢰도 등 다양하게 정의

¹ 교신저자: (614-714) 부산광역시 부산진구 가야동 산 24번지, 동의대학교 데이터정보학과, 부교수.
E-mail: csb4851@deu.ac.kr

² (614-714) 부산광역시 부산진구 가야동 산 24번지, 동의대학교 데이터정보학과, 교수.

³ (449-728) 경기도 용인시 처인구 남동 산 38-2번지, 명지대학교 스포츠기록분석연구센터, 팀장.

⁴ (614-714) 부산광역시 부산진구 가야동 산 24번지, 동의대학교 데이터정보학과, 대학원 석사과정.

된다 ('social network', Wikipedia, retrieved 2010/04/19). 즉, 사회네트워크분석은 사회네트워크에서 개체들 (사람, 회사, 기관, 홈페이지 등)의 상호작용을 통하여 얻어진 각종 데이터에 대한 개체 간의 관계를 분석하는 방법이다.

최근 사회네트워크서비스가 이슈화됨으로써 사회네트워크와 관련된 많은 연구가 있다. 대표적인 연구로써 전희주와 안철경 (2010)은 사회네트워크서비스에 대한 사용실태와 유용성을 강조하면서 보험회사에 적용 가능성을 기술하고 있다. Kang (2010)은 사회연결망 연결정도 중심성을 활용한 신규고객 상품추천방법의 추천정확성 향상 방안에 대한 연구를 수행하였다. Bae 등 (2009)은 학생들 간의 친밀도와 관련된 설문조사를 수행하여 자료를 얻고, 얻어진 자료를 R 패키지를 이용한 사회네트워크분석을 통하여 학생 조직에서 중요한 역할을 담당하는 구성원들을 탐색하는 등의 내용을 다루고 있다. 사회네트워크분석과 관련하여 유용한 서적으로 시중에 나와 있는 것으로 김용학 (2007), 손동원 (2010), 허명희 (2010) 등이 있다. 그리고 스포츠와 관련하여 Howard (1993)는 스포츠에서 구성원들의 상호작용, 관계, 구조 변화 등을 이해하기 위해서 구조적인 사회네트워크의 접근을 시도하였다. 최근 건강에 대한 높은 인식으로 인해 다양한 스포츠와 관련한 많은 응용연구들이 행해지고 있다 (신상근 등, 2009; 홍중성 등, 2010; 이장택, 2010; 박철용과 이미숙, 2011). 스포츠분야에서 사회네트워크를 활용한 국내논문은 몇몇이 있지만 대표적인 연구로서 김혜진 (2007)은 국내외에서 사회네트워크분석을 스포츠에 적용한 사례를 제시하면서 패스와 관련된 경기에서는 사회네트워크분석을 통하여 팀 경기 분석방법의 대안이 될 수 있음을 기술하고 있다.

본 연구에서는 한 축구경기에 대해서 사회네트워크분석을 이용하여 축구선수들 간의 패스 패턴을 통하여 해당 경기에서 (1) 어떤 선수가 팀에서 리더로서 얼마만큼 역할을 했으며, (2) 패스를 많이 해 준 선수 또는 패스를 많이 받은 선수들이 누구인지를 알아내어 경기에서 중요한 역할을 한 선수들을 탐지한다. 본 연구에서는 전체 데이터를 이용한 분석을 먼저 수행하고, 각 포지션별 (수비수와 비수비수)로 분석을 수행한다. 축구 경기에서 경기력을 평가하는 많은 요인이 있지만 패스 역시 경기력을 판단할 수 있는 중요한 요인이기 때문에 본 연구에서는 패스 요인이 경기력을 가장 잘 평가할 수 있다는 가정을 한다. 본 연구의 목적은 축구결과 데이터를 이용하여 팀 내에서 각 선수들의 역할에 대한 수행 실태를 평가하고, 향후 경기에서 팀의 경기 전략을 수립하는데 기초자료로 활용하는데 있다.

본 연구의 2절에서 사회네트워크분석과 관련된 개념 및 사회네트워크분석을 위한 몇몇의 중심성 측정도 (centrality measurement)들에 대해서 간단히 소개한다. 3절에서는 분석데이터인 축구 경기에 대해서 사회네트워크분석을 수행하고, 전체 선수들의 측면과 포지션별 측면에서 분석한 결과를 제시한다. 마지막으로 4절에서는 결론으로서 연구결과를 요약하고, 연구결과에 대한 장단점 및 향후 연구 과제를 기술함으로써 본 연구를 맺는다.

2. 사회네트워크분석

2.1. 사회네트워크분석의 정의 및 개요

사회네트워크분석을 정의내리기 이전에 '사회네트워크', '사회미디어'와 '사회네트워크서비스'라는 개념을 먼저 알아야 한다.

사회네트워크

사회네트워크는 웹 사이언스 (web science)의 연구 분야 중 하나로, 웹상에서 개인 또는 집단이 하나의 개체가 되어 각 개체들 간의 상호의존적인 관계에 의해서 만들어지는 사회적 관계 구조를 말한다. 모든 개체들은 네트워크 안에 존재하는 개별적인 주체들로서 사람, 회사, 기관, 홈페이지 등이 될 수 있다. 사회네트워크에서 방향성은 누가 누구에게 정보를 제공하였는지에 대해서 알 수 있지만 비방향성은 이

러한 방향성을 알 수 없는 경우이다. 두 개체들 간에 교류한 자원의 종류로서 텍스트, 영상 등이 될 수 있다.

사회미디어

사회미디어의 용어를 처음으로 사용한 사람은 가이드와이어그룹 (guidewire group)을 창업한 크리스 쉬플리 (Chris Shipley)이다. 사회미디어란 사람들이 의견, 생각, 경험, 관점 등을 서로 공유하기 위해 사용하는 온라인 플랫폼 (platform)을 말한다. 사회미디어는 텍스트, 이미지, 오디오, 비디오 등의 다양한 형태를 가지고 있다. 대표적인 사회미디어로서 블로그 (Blog), 사회네트워크 (Social network), 메시지보드 (Message boards), 팟캐스트 (Podcasts), 위키스 (Wikis), 비디오블로그 (Vlog) 등이 있다 (전희주와 안철경, 2010).

사회네트워크서비스

사회네트워크서비스는 사회미디어의 한 종류로서 사회네트워크를 이용하는 사람들 간의 의사소통과 정보공유를 위한 온라인 서비스이다. 사회네트워크서비스는 사용자 상호간에 친구를 소개하여, 친구관계를 넓힐 것을 목적으로 개설된 웹기반 커뮤니티 웹사이트이다. 사회네트워크서비스의 예로서 미국의 페이스북 (Facebook), 유튜브 (Youtube), 트위터 (Twitter), 마이페이스 (Myspace) 등과 국내의 싸이월드 (Cyworld) 등이 있다.

사회네트워크분석

사회네트워크분석은 수많은 개체들과 그 개체들 사이의 관계망 내에서 어떤 개체들이 얼마나 중요한 역할을 하는지, 하위 그룹이 없는지 또는 연결 관계가 무엇을 의미하는지 등의 개체들 간 사이에 어떤 패턴을 찾기 위한 사회과학적 또는 통계학적 방법론이라고 할 수 있다. 사회네트워크분석을 수행하기 위해서는 중심성과 밀도 등 몇몇의 측도를 필요로 한다. 본 연구에서는 중심성 (centrality)을 가지고 주로 분석을 수행한다. 이러한 사회네트워크분석에 사용되는 측도에 대해서는 2.2절에서 소개한다.

2.2. 사회네트워크의 측도

사회네트워크에서의 측도는 여러 가지가 있지만 중심성 측도는 이해하기 쉽고 사회네트워크분석에서 자주 사용되기 때문에 본 연구에서도 중심성 측도를 이용하여 분석결과를 해석한다. 중심성은 사용되는 사회네트워크에서 어느 개체가 중심적인 위치에 있는 지에 대한 측도로서 다음의 표 2.1과 같이 여러 가지로 측정할 수 있다 (Opsahl 등, 2010).

표 2.1 중심성 측도들

측도	실명
연결선	- 노드와 다른 노드가 연결된 강도. 즉, 인접 사용자수 - 이 양이 큰 노드들을 보통 허브라고 부른다.
근접성	- 커뮤니티 내 전체 노드와의 거리 - 얼마나 다른 노드들과 가까이 있는지를 측정하는 양.
중개성	- 서로 다른 커뮤니티를 연결시키는 영향도 - 어떤 노드가 다른 노드들이 서로 연결 될 때 얼마나 통로 역할을 많이 하는지를 측정하는 양.

연결선

연결선 (연결선의 수, degree)은 네트워크 내에서 어떤 사람이 다른 어떤 사람과 얼마나 많이 접해 있는가를 측정하는 개념으로서 연결중심성 (degree centrality)라고도 한다. 이 측도는 연결된 방향에 따라 입선 수 (in.degree)와 출선 수 (out.degree)로 나눌 수가 있는데, 입선 수는 다른 사람들로부터 연결

되는 수이고, 출신 수는 다른 사람들에게 연결되는 수를 의미한다. 따라서 연결선이 많은 사람일수록 해당 사회네트워크에서 중심적인 역할을 하는 사람이라고 할 수 있다.

근접성

근접성 (closeness)은 어떤 개체가 다른 개체들과 얼마나 친밀한 (가까운) 지를 측정하는 개념으로서 근접중심성 (closeness centrality)라고도 한다. 각 개체들의 근접성은 어떤 특정 개체가 다른 개체까지 도달할 때까지 거치는 선의 수 (최단경로의 길이)의 평균의 역수로 계산된다. 예로서 A-B-C-D-E로 연결되어 있는 라인 네트워크 (line network)가 있다고 할 때, A에 대한 근접성은 $4/10$ 이다. 즉, A에서 B로는 경로가 1개, C로는 2개, D로는 3개, 그리고 E로는 4개이기 때문에 A에 대한 근접성은 평균이 $(1+2+3+4)/4=10/4$ 이기 때문에 $4/10$ 가 된다. 마찬가지로 다른 개체들 B, C, D, E의 근접성은 각각 $4/7, 4/6, 4/7, 4/10$ 가 되고 이 중에서 근접성이 가장 큰 개체를 중심개체로 본다. 따라서 근접성이 가장 큰 개체 C가 가장 중요한 역할을 한다. i 번째 개체에 대한 근접성을 다음과 같이 일반화할 수 있다 (허명희, 2010).

$$C_c(i) = \frac{n-1}{\sum_{i \neq j} d(i,j)}, \quad i = 1, \dots, n. \quad (2.1)$$

여기서 $d(i,j)$ 는 개체 i 에서 j 까지의 최단 경로 길이이고, n 은 개체수이다.

중개성

중개성 (betweenness)은 어떤 개체가 다른 개체 쌍 사이에 위치하는 정도를 측정하는 개념으로서 두 개의 개체 간에 중간에 끼이는 경우의 수 (두 개체들을 연결할 경우 거쳐 가는 경우의 수)이며 중개중심성 (betweenness centrality)이라고도 한다. 2)에서와 같이 A-B-C-D-E로 구성된 네트워크로 예를 들면, 양 끝에 있는 개체 A와 E는 중개정도가 없기 때문에 가장 중개성이 낮으며, 가장 중간에 있는 개체 C가 중개의 정도가 가장 많기 때문에 중심성이 높다. 개체 B에 대한 중개성 (무방향성의 경우)은 A-B-C, A-B-D, C-B-A, D-B-A로 4이다. 이와 마찬가지로 나머지 개체들 A, B, C, D, E의 중개성은 각각 0, 4, 8, 4, 0이다. 따라서 개체 C가 중개성이 가장 크고 중개성이 클수록 중심개체로 본다. 따라서 앞의 간단한 네트워크 예에서 개체 C가 근접성 측면에서 가장 중요한 역할을 한다. 어떤 개체 i 에 대한 중개중심성은 다음과 같이 계산된다 (허명희, 2010).

$$C_b(v) = \sum_{i \neq v} \sum_{j \neq v} \frac{\sigma_{ivj}}{\sigma_{ij}}, \quad i = 1, \dots, n. \quad (2.2)$$

여기서 σ_{ij} 는 개체 i 에서 j ($i \neq j$)로 가는 최단경로 수이고, σ_{ivj} 는 개체 i 에서 개체 j ($i \neq j$)로 가는 경로 중에서 v 를 거치는 경로의 수이다.

커뮤니티

커뮤니티 (community)는 임의의 사회네트워크분석을 수행했을 경우 몇 개의 부 집단 (sub group)이다. 즉, 하나의 커뮤니티는 부 그룹들로 구성된다. 이러한 커뮤니티 내에서 개체들이 어떤 역할을 하느냐에 따라 리더 (leader), 서브리더 (subleader), 팔로우 (follower) 등으로 나눌 수 있으며, 이것들을 커뮤니티 역할 (community role)이라고 한다. 이들에 대한 정의는 표 2.2에 주어져 있다.

본 연구에서 사용한 커뮤니티 탐색을 위한 알고리즘으로 커뮤니티 간보다 커뮤니티 내의 결속력이 높은 커뮤니티를 하나의 그룹노드로 형성하는 Louvain 방법을 사용하였다 (Blondel 등, 2008). Louvain 방법의 알고리즘은 다음의 두 단계에 걸쳐 수행된다. 첫째, 각 노드들을 하나의 커뮤니티로서 초기화시킨다. 둘째, 각 노드가 속해있는 커뮤니티로부터 결속력이 높은 이웃하는 커뮤니티로 이동시킨다. 이러한 과정을 결속력이 더 이상 향상되지 않을 때까지 반복한다.

표 2.2 커뮤니티 내에서 개체들의 역할

역할	설 명
리더	- 그룹 혹은 커뮤니티를 이끌어 가는 개체. - 양방향 의사소통이 많은 개체.
서브리더	- 리더의 성격을 띠고 있지만 실제 리더보다 영향도가 적은 개체.
팔로우	- 리더와 서브리더를 제외한 모든 개체.
중개	- 이상점들 중 서로 다른 커뮤니티 연결시키는 다리역할을 하는 개체.

3. 데이터 분석

3.1. 분석데이터

본 연구에서 사용된 분석데이터는 2010년 6월 4일 대한민국과 스페인과의 친선경기에 대한 축구 경기결과에서 우리나라 대표 팀에 대한 경기 내용이다. 분석데이터에서 분석에 적용될 선수들은 교체선수를 포함하여 모두 13명의 선수들이다. 사회네트워크분석을 위한 원시데이터는 표 3.1에 주어져 있다.

표 3.1 원시데이터

일련번호	패스한 선수	패스 받은 선수	가중치
1	KIM Jung Woo	KI Sung Yueng	2
2	KIM Jung Woo	LEE Chung Yong	1
3	KIM Jung Woo	LEE Jung Soo	3
4	KIM Jung Woo	OH Beom Seok	3
5	KIM Jung Woo	CHO Yong Hyung	4
⋮	⋮	⋮	⋮
94	AHN Jung Hwan	LEE Young Pyo	4
95	AHN Jung Hwan	PARK Chu Young	1
96	CHA Du Ri	KIM Jung Woo	1
97	CHA Du Ri	KI Sung Yueng	2
98	CHA Du Ri	CHO Yong Hyung	1

사회네트워크분석을 위한 툴은 넷마이너 (netminer), 유시아이넷 (UCINET), 파이엑 (pajek), 크랙플롯 (Krackplot), R 등이 있다. 이러한 각 분석 툴들은 장단점이 있지만 본 연구에서는 통계학에서 많이 사용되고 있는 SAS의 모듈로서 사용자들이 접근하기가 용이한 Optgraph procedure for SNA를 사용하였다. 참고로 향후 분석결과에서 AJW는 안정환, KJW는 김정우, KNI는 김남일, CYH는 조용형, CDR은 차두리, OBS는 오범석, YKH는 염기훈, LCY는 이청용, KSY는 기성용, LJS는 이정수, KJS는 김재성, LYP는 이영표, PCY는 박주영 선수들을 의미한다. 그리고 한국은 전반에 4-4-1-1 (수비-미드필드-공격형미드필드-공격수)의 전술을 쓰다가 후반에 안정환을 투입하여 4-4-2 (수비-미드필드-공격수)의 전술로 바꾸었다. 이러한 이유로 공격수 간의 패스는 거의 전무하기 때문에 분석의 의미가 없다고 판단하여 포지션별 사회네트워크분석을 수비수와 수비수가 아닌 두 집단에 대해서 분석하였다. 그리고 교체 선수는 수비수에서 차두리와 오범석이 교체되었으며, 비 수비수에서는 염기훈과 안정환이 교체되었고, 김재성과 김남일이 각각 교체되었다. 따라서 본 연구에서는 교체된 선수들을 경우, 두 선수를 같은 선수로 인식하게끔 하여 분석을 수행하였다. 예를 들면, 교체된 염기훈과 안정환은 YKH_AJH으로 표현하였다.

3.2. 분석결과

3.2.1. 전체분석

표 3.2와 표 3.3은 분석데이터의 전체에 대한 사회네트워크분석 결과 중심성 측도 (degree, closeness, between)들에 대한 기술통계량과 커뮤니티이다. 분석결과 총 3개의 커뮤니티 수가 형성 되었고, 세 가지 측도들 중 근접성의 견지에서 1에 가까울수록 중심이 된다고 할 수 있다. 표 3.2에서 평균근접성이 0.91로 얻어졌기 때문에 평균적으로 ‘선수들은 대체적으로 가까운 관계에 있을 것이다’라고 유추할 수 있다. 표 3.3의 커뮤니티가 포지션별로 일치하지 않는 것을 볼 수 있다. 이 결과는 포지션 간에도 패스가 많이 이루어 질 수 있기 때문에 당연한 결과라고 볼 수 있다.

표 3.2 중심성 측도의 기술통계량 (전체선수)

측도	평균±표준오차	최소값	최대값
입선수	8.00±1.33	5.000	9.000
출선수	8.00±1.15	5.000	9.000
연결선	16.0±1.83	13.00	18.00
근접성	0.91±0.08	0.783	1.000
중개성	0.01±0.01	0.000	0.027

표 3.3 커뮤니티 (전체선수)

선수	포지션	커뮤니티
KJW	수비수	0
KSY	미드필더	0
OBS_CDR	수비수	0
LJS	수비수	1
CYH	수비수	1
KNLKJS	미드필더	1
LCY	미드필더	2
YKHAJH	미드필더	2
LYP	수비수	2
PCY	공격수	2

표 3.4는 중심성 측도들을 가지고 각 선수의 경기력을 평가한 결과이다. 분석결과, 연결선 (degree = in_degree + out_degree)과 근접성 그리고 중개성의 측면에서 1위부터 5위까지는 ‘미드필더’가 차지하고 있지만 4위와 5위는 교체선수들의 합이기 때문에 순수한 선수들의 기량의 결과라고 보기 힘들다. 표 3.4의 결과로서 김정우, 기성용, 이청용 선수가 세 가지 측도 모두에서 상위권에 있기 때문에 스페인 친선 경기에서 매우 활약을 많이 한 선수들이라고 할 수 있다. 참고로 축구경기에서의 상식에서 (1) 미드필더는 입선수 (in_degree)도 많을 것이고 출선수 (out_degree)도 많을 것이라고 판단할 수 있다. 그리고 (2) 수비수는 출선수가 많을 것이고, (3) 공격수는 입선수가 많을 것이라고 판단해 볼 수 있다. 표 3.4의 결과에서 각 포지션에 따른 입선수와 출선수의 결과는 전기한 내용을 반영하고 있음을 알 수 있다.

그림 3.1은 10명의 선수들 (교체선수들은 하나의 선수로 취급)의 중심성의 측도들과 선수들 간의 패스와 관련된 정보를 이용하여 작성된 것으로서 선수들 간의 관계를 흐름도로 표현한 것이다. 여기서 양방향 화살표는 두 선수들 간에 서로 패스를 주고받았음을 의미하고, 단방향 화살표는 어떤 선수가 다른 선수에게 패스를 주거나 받기만 한 경우를 의미한다. 예를 들면, 양방향 화살표는 선수 A가 선수 B에게 패스를 주고받은 경우이며, 단방향 화살표는 선수 A가 선수 B에게 패스를 주거나 선수 A가 선수 B에게 패스를 받은 경우를 의미한다. 그림 3.1의 결과에서 각 커뮤니티를 도형으로 표현하였다. 즉, 커뮤니티 0은 육각형, 커뮤니티 1은 육각형, 커뮤니티 2는 원으로 표현하였다. 각 커뮤니티에서의 리더 (leader)와 서브리더 (sub-leader)들을 찾는 것도 의미가 있지만 본 연구에서는 전체 선수들에서의 리더

표 3.4 중심성 측도의 결과 (전체선수)

선수	포지션	커뮤니티	입선수	출선수	연결선	근접성	중개성
KJW	미드필더	0	9	9	18	1.000	0.027
KSY	미드필더	0	9	9	18	1.000	0.027
LCY	미드필더	2	9	8	17	0.947	0.015
KNLKJS	미드필더	1	9	8	17	0.947	0.025
YKH_AJH	미드필더	2	9	8	17	0.947	0.021
LJS	수비수	1	7	9	16	0.900	0.012
OBS.CDR	수비수	0	8	8	16	0.900	0.006
CYH	수비수	1	7	8	15	0.857	0.004
LYP	수비수	2	5	8	13	0.783	0.004
PCY	공격수	2	8	5	13	0.783	0.000

와 서브리더를 찾는 것에 의미를 둔다. 그리고 교체된 선수는 순수한 기량을 평가할 수 없기 때문에 선수들의 역할을 탐지하는 것에서는 제외하였다. 그림 3.1에서 색깔이 진하면서 테두리의 선이 있는 선수를 리더, 색깔이 진하고 이탤릭체로 된 선수를 서브리더로 표현하였고, 선수들의 역할은 세 가지 중심성 측도를 고려하여 결정하였다. 그림 3.1의 결과로부터 리더의 역할을 수행한 선수들은 김정우, 기성용, 이청용 선수이고, 리더의 성격을 띠고는 있지만 리더에 미치지 못한 서브리더는 이정수 선수임을 알 수 있다. 지금까지 전체 선수 10명에 대해서 사회네트워크분석을 수행한 결과를 보였다. 축구경기에서 승리하기 위해서는 선수들 간에 원활한 ‘패스’를 통하여 공격을 하고, 공격의 성공으로 해당 경기에서 승리를 할 수 있는 것이다. 선수들의 경기력을 평가하는데 있어서 포지션별로 임무가 주어져 있기 때문에 선수 전체에 대해서 분석하는 것보다 포지션별로 분석을 수행하는 것이 타당할 수 있다. 따라서 포지션별 (수비수와 비수비수)로 나누어 분석하는 것이 타당한 지를 알아보기 위해서 검정을 수행할 필요가 있다. 일반적으로 비수비수가 수비수보다 경기의 활동력이 높을 것으로 판단되기 때문에 대립가설을 ‘비수비수가 수비수보다 경기의 활동력이 높다’로 설정하여 가설을 검정하였다.

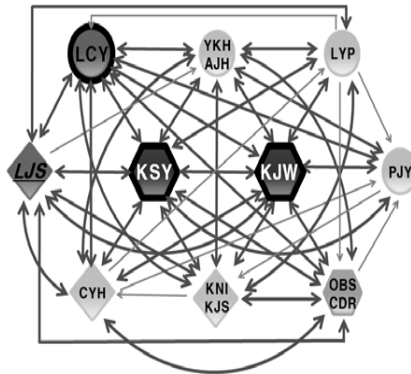


그림 3.1 선수들 간의 관계도 (전체선수)

여기서 중심성의 측도들을 검정변수로 하여 수비수 집단과 비수비수 집단의 정규성 검정 결과 정규성을 따르지 않고, 개체수도 적기 때문에 비모수검정법인 Mann-Whitney 방법을 사용하여 위의 가설을 검정하였다. 그 결과는 표 3.5에 주어져 있다.

표 3.5의 결과 모든 중심성 측도에서 유의수준 0.05하에서 유의한 결과를 보여 비수비수가 수비수보다 경기의 활동력이 높음을 알 수 있다. 따라서 포지션별로 세부적인 분석이 필요하다고 판단되어 포지

표 3.5 Mann-Whitney 검정 결과

중심성 측도	연결선	근접성	중개성
유의확률	0.032	0.032	0.043

선별 사회네트워크분석을 수행한다.

3.2.2. 포지션별 분석

본 절에서는 포지션별로 사회네트워크분석을 수행한다. 본 연구에서는 포지션을 나누는데 있어서 감독의 전략과 전술로 인하여 공격수가 전반에는 1명, 후반에는 2명으로 개체수가 너무 적기 때문에 공격수를 미드필더로 통합하였다. 따라서 본 절에서는 포지션을 (1) 수비수와 (2) 비 수비수 (공격수와 미드필더)로 나누어 사회네트워크분석을 수행한 결과를 제시한다.

(1) 수비수 분석

본 연구의 분석대상에서 수비수 선수들 내에서 어떤 관계가 있는지를 알아보기 위해서 사회네트워크 분석을 수행한 결과가 표 3.6과 표 3.7에 주어지 있다. 표 3.6은 수비수 선수들에 대한 세 가지 중심성 측도의 기술통계량에 대한 결과를 보여 주고 있다. 표 3.7은 연결선, 근접성 그리고 중개성의 중심성 측도들을 가지고 수비수 선수별로 경기력을 평가한 결과이다. 분석결과, 세 측도에 대해서 비슷한 값을 보이고 있지만 이정수가 중심적인 역할을 하고 있음을 알 수 있다.

그림 3.2의 선수들 간의 관계 흐름도에서도 수비수의 경우 이정수가 리더로서의 역할을 수행하였음을 알 수 있다.

표 3.6 중심성 측도의 기술통계량 (수비수)

중심성	평균±표준오차	최소값	최대값
in_degree	2.25±0.96	1.000	3.000
out_degree	2.25±0.50	2.000	3.000
degree	4.50±1.29	3.000	6.000
Closeness	0.82±0.14	0.667	1.000
Betweenness	0.13±0.20	0.000	0.417

표 3.7 중심성 측도의 결과 (수비수)

선수	입선수	출선수	연결선	근접성	중개성
LJS	3	3	6	1.000	0.417
OBS_CDR	3	2	5	0.857	0.083
CYH	2	2	4	0.750	0.000
LYP	1	2	3	0.667	0.000

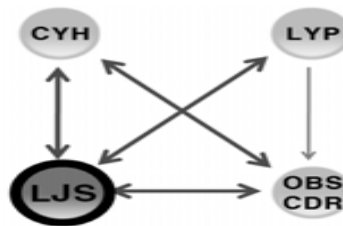


그림 3.2 선수들 간의 관계도 (수비수)

(2) 비수비수 (공격수와 미드필더)

본 연구의 분석대상에서 비수비수 선수들 내에서 어떤 관계가 있는지를 알아보기 위해서 사회네트워크 분석을 수행한 결과가 표 3.8과 표 3.9에 주어져 있다. 표 3.8은 비수비수 선수들에 대한 세 가지 중심성 측도의 기술통계량에 대한 결과를 보여 주고 있다. 표 3.9는 세 가지 중심성 측도들에 대한 비수비수 선수들의 경기력을 평가한 결과이다. 비수비수의 경우에는 기성용, 김정우, 이청용 선수들 모두는 세 가지 중심성 측도에서 모두 같은 값을 보이고 있어 비수비수로서의 역할을 잘 소화해 내었다고 할 수 있다. 박주영의 경우는 가장 낮은 선수로 얻어진 것으로 보이나 세 측도의 값으로 보면 다른 선수들과 별 차이를 보이지 않기 때문에 박주영 선수 역시 경기력 면에서 활약이 두드러졌다고 할 수 있다. 즉, 미드필더들과 공격수들은 스페인과의 친선 경기에서 모두 우수한 역할을 수행했다고 할 수 있다. 굳이 평가를 한다면 염기훈, 안정환, 김남일, 김재성 선수들은 교체된 선수들이기 때문에 선수들의 평가는 무의미하기 때문에 기성용, 김정우, 이청용 선수들이 비수비수 포지션에서 많은 활약을 했다고 할 수 있다.

그림 3.3의 선수들 간의 관계 흐름도를 보면, 비수비수 (미드필더, 공격수)의 경우 모든 선수들이 부여된 포지션에서 최선을 다하였고, 누구를 리더와 서브리더라고 할 것도 없이 모두 훌륭한 경기를 소화하였다고 평가할 수 있다.

표 3.8 중심성 측도의 기술통계량 (비수비수)

중심성	평균±표준오차	최소값	최대값
in.degree	4.83±0.41	4.000	5.000
out.degree	4.83±0.41	4.000	5.000
degree	9.67±0.52	9.000	10.00
Closeness	0.97±0.05	0.909	1.000
Betweenness	0.01±0.01	0.000	0.013

표 3.9 중심성 측도의 결과 (비수비수)

player	입선수	출선수	연결선	근접성	중개성
YKH_AJH	5	5	10	1.000	0.013
KSY	5	5	10	1.000	0.013
KJW	5	5	10	1.000	0.013
LCY	5	5	10	1.000	0.013
KNL_KJS	5	4	9	0.909	0.000
PCY	4	5	9	0.909	0.000

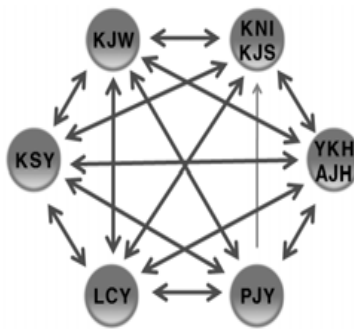


그림 3.3 선수들 간의 관계 흐름도 (비수비수)

4. 결론

본 연구에서는 스페인과의 친선경기 데이터를 이용하여 우리나라 선수들의 경기력을 분석하기 위하여 사회네트워크분석을 수행하였다. 이를 위하여 전체데이터를 이용하여 분석하였고, 포지션(수비수와 비수비수)별로 구분하여 분석을 수행하였다.

분석결과는 다음과 같다. 전체데이터를 이용하였을 때, 리더의 역할을 수행한 선수들은 김정우, 기성용, 이청용 선수이고, 리더의 성격을 띠고는 있지만 리더에 미치지 못한 서브리더는 이정수 선수였음을 알 수 있었다. 그리고 포지션별로는 수비수의 경우 리더의 역할을 수행한 선수는 이정수 선수였음을 알 수 있었다. 비 수비수(미드필더, 공격수)의 경우에는 모든 선수들이 세 가지 축도를 기준으로 평가하였을 때 비슷한 점수를 보여 누구를 리더와 서브리더라고 할 것도 없이 모두 훌륭한 경기를 소화하였다고 평가할 수 있다.

본 연구에서는 선수들의 경기력을 평가하는데 있어서 '패스'가 가장 중요하다는 가정 하에 '패스' 변수 하나만을 사용하여 사회네트워크분석을 수행하였다. 실제 축구 경기에서 경기력에 대한 평가를 하기 위해서는 '감독의 전략과 진술', '선수들의 컨디션', '운동장의 상황(응원자들의 수 등)', '날씨' 등을 고려해서 종합적으로 판단해야 한다. 이러한 점에서 본 연구의 결과는 단점이 있으며, 향후 연구과제로서 상기의 다양한 변수들을 이용하여 선수들의 경기력을 평가해야 할 것으로 판단된다. 또한 한국의 축구 경기력 향상을 위하여 한 경기만을 이용하는 것보다 여러 경기를 이용하여 종합된 데이터를 이용하여 분석을 수행해 보는 것도 의미가 있을 것이라고 판단된다. 그리고 우리나라의 선수들의 역할에 대해서 분석하는데 있어서 다른 팀이 영향을 미칠 수 있기 때문에 한 경기의 전체 선수들에 대한 분석과 상대팀의 선수들에 대한 분석을 수행하는 것도 의미가 있다. 그러나 본 연구의 목적이 우리나라의 선수들의 기량을 평가하여 향후 팀의 전략 및 전술을 세우기 위한 기초자료로 활용하는 것이기 때문에 우리나라 선수들만을 분석하였다. 다른 측면에서의 향후 연구과제로서 사회네트워크분석과 공간통계학적 분석을 결합한 분석을 통하여 보다 정밀한 정보를 얻을 수 있을 것으로 기대된다.

참고문헌

- 김용학 (2007). <사회 연결망 분석>, 박영사, 서울.
- 김혜진 (2007). 사회연결망(social network analysis: SNA를 이용한 스포츠 경기분석. <한국체육측정평가학회지>, **9**, 99-112.
- 박철용, 이미숙 (2011). 스포츠영재성 검사 항목과 코스타스 점수간의 연관성 분석. <한국데이터정보과학회지>, **22**, 57-64.
- 손동원 (2010). <사회 네트워크 분석>, 경문사, 서울.
- 신상근, 조용주, 조영석 (2009). 유럽 리그에서 득점과 실점을 이용한 승점 추정에 관한 연구. <한국데이터정보과학회지>, **20**, 837-844.
- 이장택 (2010). 엘리트 10종경기 선수들의 경기력 패턴에 관한 연구. <한국데이터정보과학회지>, **21**, 1071-1079.
- 전희준, 안철경 (2010). 사회네트워크서비스(Social Network Service)의 확산과 보험회사의 활용 가능성. <손해보험협회>, **504**, 18-31.
- 허명희 (2010). <R을 활용한 사회네트워크분석 입문>, 자유아카데미, 서울.
- 홍종성, 정민섭, 이재형 (2010). 2010 남아공 월드컵 축구 예측모형 분석. <한국데이터정보과학회지>, **21**, 1137-1146.
- Bae, J. D., Min, B. M. and Kim, J. S. (2009). Centrality analysis in students' friendship network and its strategic applications. *Journal of the Korean Data Analysis Society*, **11**, 1467-1478.
- Howard, L. N. II (1993). Social network analysis of sport: Emphasizing social structure in sport sociology. *Sociology of Sport Journal*, **10**, 315-321.
- Kang, B. S. (2010). Performance improvement methods for new customer recommendations using degree centrality of social network. *Journal of the Korean Data Analysis Society*, **12**, 1511-1522.
- Opsahl, T., Agneessens, F. and Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, **32**, 245-251.

Social network analysis for a soccer game

Seungbae Choi¹ · Changwan Kang² · Hyongjun Choi³ · Byungyuk Kang⁴

^{1,2,4}Department of Data Information Science, Dongeui University

³Invasion Game Analysis Team, Myongji University

Received 25 September 2011, revised 19 October 2011, accepted 25 October 2011

Abstract

Social network analysis is the social statistical analysis of any social structure involving a stream of mutual information between observations. In this study we used the results of passes between players in a soccer game. The analysis contents are as follows. (1) Players with important or leading roles are identified. (2) Players are assessed by pass frequency and the success rate of passes. The purpose of this study is for use as basic data for future team strategy, and achieves this by evaluating the role of each individual player within a team. In this study, social network analysis without separating positions is conducted, and is also performed for defensive and attacking positions respectively. The results of this study are as follows: First, when complete team data were available, the players performing leadership roles were Jung-woo Kim, Sung-yeung Ki and Chung-young Lee, whereas Jeong-su Lee acted as a sub-leader. In case of data for defensive positions Jeong-su Lee was a leading player, and in terms of attacking positions, all of the players excelled in the game and could be evaluated as playing lead roles.

Keywords: Leader, social network analysis, social network service, subleader.

¹ Corresponding author: Associate professor, Department of Data Information Science, Dongeui University, Busan 614-714, Korea. E-mail: csb4851@deu.ac.kr

² Professor, Department of Data Information Science, Dongeui University, Busan 614-714, Korea.

³ Director, Invasion Game Analysis Team, Myongji University, Gyeonggi-do 449-728, Korea.

⁴ Master course student, Department of Data Information Science, Dongeui University, Busan 614-714, Korea.