

소셜네트워크 서비스 사용자 패턴 발견을 위한 사회 네트워크 분석 활용에 관한 연구: 페이스북을 중심으로*

하 병 국** · 장 용 수*** · 조 재 희****

목 차

요약	3.3 분석 프로세스
1. 서론	4. 분석 및 해석
2. 문헌 연구	4.1 네트워크 분석 시각화
2.1 소셜네트워크 서비스	4.2 시각적 네트워크 분석 내용
2.2 사회네트워크분석	4.3 네트워크 계량적 분석
3. 분석을 위한 데이터 및 분석 도구	5. 연구의 한계와 기대효과
3.1 네트워크 분석도구 NodeXL	참고문헌
3.2 페이스북 데이터 수집	Abstract

요약

개인의 소셜네트워크 서비스 활용의 증가는 기업의 비즈니스 활동에 새로운 기회로 주목 받고 있으며 소셜네트워크 서비스 관련연구 또한 많은 관심을 받고 있다.

본 연구는 사회네트워크분석을 통하여 소셜네트워크 서비스 사용자패턴을 발견하고자 한다. 그리고 사용자의 사용목적에 따른 네트워크의 패턴을 구분하고자 한다. 이를 위해 네트워크분석 도구인 NodeXL을 이용하여 소셜 네트워크 서비스 중 페이스북(Facebook.com)의 사용자를 분석하였다. 그 결과 단일 네트워크로 인식되었던 사용자의 네트워크를 여러 개의 하위그룹으로 구분할 수 있었다.

그리고 개인사용자의 페이스북 친구와 국내 K대학의 페이스북 친구를 비교하여 소셜네트워크 서비스의 사용목적에 따른 네트워크 구조의 차이를 발견하고 노력 하였다.

표제어: 소셜네트워크 서비스, 사회네트워크 분석, NodeXL

접수일(2012년 3월 10일), 수정일(1차: 2012년 3월 23일), 게재확정일(2012년 3월 29일)

* 이 논문은 2011년도 광운대학교 교내학술연구비 지원에 의해 연구되었음.

** 광운대학교 경영정보학과 대학원, djkoogy@gmail.com

*** 광운대학교 경영정보학과 대학원, jamitan@chol.com

**** 광운대학교 경영학부 교수, 교신저자, mis1@kw.ac.kr

1. 서론

요즘 들어 우리를 가장 많이 변화시킨 것을 말하라면 누구나 T.G.I.F.를 이야기할 것이다. 이는 개인의 삶뿐만 아니라 사회전반의 문화에 지대한 영향을 주었다. 우리는 Twitter로 약속을 정하고 Google에서 정보를 찾아서 iPhone으로 사진을 찍어 Facebook에 업로드하여 친구들과 공유를 한다. 이것이 T.G.I.F.-Life이다[3].

인터넷환경의 발전과 개인 미디어의 비약적인 성장에 힘입어 이런 삶이 가능해졌다. 그러나 이를 자세히 드러다 보면 단순한 기술의 발전만이 전부는 아니다. 그 중심에는 늘 사람이 있었다. 결국 사람과 사람을 연결하는 소통기술이 발전한 것이다. 우리는 이런 기술을 소셜네트워크 서비스(Social Network Services 이하: SNS)라 부른다. 앞에서 언급한 Twitter와 Facebook이 대표적인 SNS인 것이다.

SNS를 말하고자 할 때 가장 많이 언급되는 것이 페이스북(Facebook)일 것이다. 현재 6억 명이 넘는 가입자 수를 자랑하는 페이스북은 모바일 사용자만 1억 명이 넘는다[12]. 매일 6천만건의 업데이트가 일어나고 가입자들은 하루 평균 55분을 이용한다. 이런 페이스북의 사용자들은 자신의 정보, 프로필을 공개하여 친구맺음을 통하여 상대방의 글과 사진을 볼 수 있다. 그렇다고 트위터(Twitter)나 미투데이(Me2Day)처럼 완전히 개방된 서비스는 아니며 다소 폐쇄적인 SNS라 볼 수 있다.

기업에서도 이런 SNS를 활용한 움직임이 활발해지고 있다. 많은 기업들이 기업의 이미지, 상품 정보 제공에서부터 고객과의 소통의 장으로 활용하고 있다. 최근 소셜CRM과 같은 SNS데이터를 분석하여 기업의 의사결정에 반영하고자 하는 움직임 또한 일어나고 있다[14]. 하지만 고객정보공개, 활용인력부재, 명확하지 못한 ROI 등의 제약사항으로 활용에 어려움이 있다.

학계 또한 SNS의 활용(수익모델, 마케팅, 문화),

SNS 이용행태(재이용, 만족), SNS 특성(사회자본, 미디어, 보안) 등을 중심으로 활발한 연구들이 이루어지고 있다[2]. 그러나 많은 SNS 연구들이 설문위주의 개인 특성만을 반영하는 아쉬움을 가지고 있다.

SNS를 잘 사용하기 위해서는 본질을 이해할 필요가 있다. 가장 근본적인 것은 관계 형성이다. 서두에서 언급하였지만 결국 다른 사람과 관계를 형성하는 것이 가장 먼저일 것이다. 이런 관계를 계속하여 형성하다보면 관계들 사이에 공동체(네트워크)가 형성되고 그 공동체를 유지하거나 참여하기 위하여 사람들은 무엇인가 지속적으로 행동한다. 이런 행동을 사회구성론적 측면에서 “지속적인 상호작용” 혹은 “소속감”이라고 한다[6].

본 연구는 SNS 사용자의 형성된 네트워크를 분석하여 네트워크의 지속적인 상호작용을 위한 특성과 관계형성에 관한 패턴을 발견하고자 한다. 형성된 네트워크를 분석을 위해 사회네트워크분석(Social Network Analysis 이하: SNA)을 실시한다. 페이스북을 사용하는 개인(일반사용자)과 조직(국내 K대학)을 비교 분석 대상으로 한다. 분석 내용은 사회 네트워크 분석을 통한 시각적 분석과 측정지표를 활용한 통계적 분석을 실시한다.

2. 문헌 연구

2.1 소셜네트워크 서비스

소셜네트워크 서비스(Social Network Services)는 21세기 초반 1인 미디어, 1인 커뮤니티를 중심으로 하는 새로운 인적 네트워크 형성서비스이다. 이는 인터넷을 기반으로 정보를 공유하고 의사소통을 도와 사용자들이 서로서로 친구를 소개하고 대인관계를 넓히는 것을 목적으로 개설된 커뮤니티형 웹사이트이다. 싸이월드, 마이스페이스, 페이스북의 탄생은 SNS의 시작이라고 할 수 있다. SNS는 블로그, 개인 커뮤니티를 비롯한 매우 다양한 서비스 형태를 가지

고 있기 때문에 SNS를 정의하는 것은 아주 어렵다. 위키피디아에 의하면, “SNS란 사회적 관계 개념을 인터넷 공간으로 가져온 것으로 사람과 사람간의 관계 맺기를 통한 네트워크 형성을 지원하는 서비스이다.” SNS는 사이버 공간에 기초하여 사이버상의 인맥구축을 통해 사용자들이 새로운 네트워킹 가치를 찾아가는 참여형 서비스로 볼 수 있다.

사회구성론적 측면에서 볼 때 SNS는 사이버 공동체에 포함된다고 볼 수 있다. 사회구성론은 인간의 관계는 사회적 관계 속에서 영향을 받고, 또 영향을 끼친다는 이론이고, 사이버 공동체란 사이버 공간에서 공통의 관심을 기반으로 정서적 유대를 가지고 지속적으로 상호작용하는 네트워크라 할 수 있다[6]. 사이버공동체의 개념을 충족시키는 필수조건은 무엇보다도 구성원 간의 “지속적인 상호작용”인데, 사이버공동체 구성원 간의 지속적인 상호작용이 가능하려면 공동체에 대한 “소속감”이 유지되어야 한다. 급변하는 현대사회에서 사람들은 심리적 안정과 소속감을 추구하게 되며, 오프라인에서는 쉽게 만날 수 없었던 공통의 관심사를 가진 사람들과 상호작용하는 과정에서 사이버공동체라는 장소에 의미를 부여한다. 이러한 소속감은 공동체에 대한 애정과 구성원 간의 유대감으로 나타나며, 사이버공동체가 오프라인 공동체로 확장되어 대면적이고 전인격적인 관계로 발전하기도 한다[6]. 이상의 논의를 고려하면 SNS를 얼마나 효과적으로 잘 활용하는지는 SNS 사용자들 간에 얼마나 더 정서적 유대를 갖느냐에 달려 있다고 볼 수 있다.

미디어적 측면에서 SNS를 대중매체와 구별하여 말하면 다음과 같은 특성으로 볼 수 있다.

첫째, 접근성으로 기존 미디어 즉 대중매체는 전형적으로 개인 또는 국가가 소유하는 형태를 취하지만, SNS는 일반적으로 적은 비용으로 혹은 무비용으로 누구나 이용 가능하다.

둘째, 유용성으로 대중매체 생산은 전문화된 기술과 훈련을 요구하지만 SNS는 누구나 제작수단을 쉽게

쓸 수 있도록 소프트웨어 기술을 재발명하기도 한다.

셋째, 최신성으로 대중매체가 생산하는 커뮤니케이션은 며칠, 몇 주에서 몇 달이 걸려야 생산되지만 SNS는 거의 즉각적인 반응을 가져올 수 있다.

마지막은 연속성으로 대중매체는 예를 들어 잡지는 한번 발간이 되고 나면 수정을 할 수 없지만, SNS는 코멘트나 편집을 통해 거의 즉각적으로 변경할 수 있다. 미디어적인 측면에서 SNS가 가진 이러한 특징들 탓에 더 많은 사람들이 더 많은 커뮤니케이션을 하는 것이 가능해지고 점차 증대되는 커뮤니케이션으로 인해 많은 사람들이 소비나 구매 행동 등에 영향을 받을 가능성이 커진다[5].

최근 들어 SNS의 관심과 관련 연구들이 활발하게 이루어지고 있다. 나종연(2010)은 소셜 네트워크 활용에 미치는 영향을 Shih and Venkatesh(2004)의 사용확산 모형으로 설명하였다. 고상민(2010)은 사회적 자본이론을 기반으로 SNS사용자의 사회적 자본과 관계를 설문 분석을 통하여 설명하고자 하였다. 서우중(2010)은 SNS를 웹사이트의 품질을 기반으로 지속의도와 구전의도의 실증을 실시하였다. 변창(2010)은 SNS사용만족이 지속적 참여와 같은 몰입, 효용에 미치는 연구를 실시하였다. 기존의 연구에서는 SNS의 가장 중요한 관계에 관한 연구가 부족하다는 것을 이야기 하고자 한다. 이를 위한 관계분석 연구가 필요하다고 생각된다.

2.2 사회네트워크분석

사회네트워크분석(Social Network Analysis)은 개체 간의 관계를 시각적으로 표현하여 개체 간 연결상태 및 연결구조의 특성을 계량적으로 분석하는 기법으로 [2], 사회네트워크는 일반적으로 친구, 직장동료 또는 정보를 교환하는 관계 등 다양한 사회적 관계를 통해 연결된 사람이나 조직 혹은 다른 사회적 개체의 집합을 의미한다[7].

사회 네트워크는 arnes(1954)에 의해 처음 사용된

용어로, 개인적인 인간관계가 확산되어 형성된 사람들 사이의 연결된 네트워크이다[1]. ‘사회적 관계’의 유형은 현실 세계에서 매우 다양하지만 Knoke and Kuklinski(1988)는 사회적 관계유형을 거래, 의사소통, 상호침투, 도구적, 감정적, 권위/권력, 친족관계로 보편적인 7가지 사회적 관계유형을 정의하고 있다. 사회적 관계유형을 이해하는 관점은 분석초점, 관계모양에 따라 구분될 수 있다. 먼저 분석초점에 따른 구분은 Human와 Provan(2000)이 ‘에고 네트워크(ego-centric network)’, ‘양자 네트워크(dyadic network)’, ‘전체 네트워크(total network)’ 3가지로 구분하였다[4].

에고 네트워크는 한 개인(Ego)을 중심에 위치시키고, 그 개인과 다른 노드(node)와의 연결을 표현한 네트워크를 의미한다. 이 에고 네트워크는 개인이라는 하나의 중심에서 밖으로 퍼져나가는 스타(Star) 모양을 가진다. 에고 네트워크 분석은 샘플 사이즈가 지나치게 커지기 쉽다는 단점을 가지고 있다. 이론적으로 샘플크기가 N이면 에고 네트워크도 N개만큼 필요하다는 것이다. 양자 네트워크는 두 사람 사이의 네트워크이며 두 사람을 쌍(pair) 단위로 분석한다. 만약 N명의 행위자가 있으면 모두 $N(N-1)/2$ 만큼의 쌍이 분석대상이 된다. 전체 네트워크는 N명 전체 행위자로 구성된 것을 의미하며 보통 네트워크라고 말할 때 가장 보편적으로 지칭하는 네트워크이다.

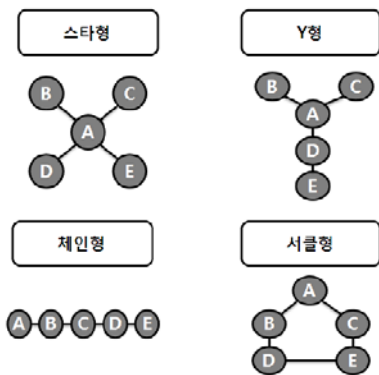


그림 1. 4가지 네트워크 유형

Fig. 1. The 4 Types of Communication Network

관계모양에 따른 구분은 스타형(Star), Y형(Y), 체인형(Chain), 서클형(Circle) 형태의 4가지로 구분된다. 4가지 형태는 다음의 그림 1과 같다.

사회네트워크이론에서 ‘중심’의 중요성에 대한 인식은 비교적 뿌리가 깊다. 1940년대 미국 MIT대학의 알렉스 바베라스(Alex Bavelas) 교수와 그 제자인 해롤드 리 빗(Harold Leavitt)의 실험 연구결과는 ‘스타형’과 ‘Y형’에서 빠른 문제 해결을 얻게 된다는 점이다. 그 이유는 ‘스타형’과 ‘Y형’ 두 유형 모두 ‘중심’이 되는 사람을 통해 의사소통과 문제 해결을 하는 유형이라는 점이다. 말하자면, 확연하게 ‘중심’을 가진 집단은 문제 해결책을 찾는데 시간 소모가 적고 그 실행도 즉각적이었다는 발견을 제공한 것이다[1].

사회네트워크분석은 그래피론(Graph Theory)을 기반으로 한다. 그래프는 점(node)과 선(line)으로 표현하는 방법이며 점은 행위자를, 선은 행위자들 간의 관계를 표현하고 있다. 그래피론의 점과 선의 조합을 통해 사회적 관계(거래, 의사소통, 상호침투 등)를 표현하는 방법이 그래프이다. 이러한 그래프는 네트워크의 전체 윤곽을 이해하는데 시각적으로 도움을 준다. 그래프는 비방향 그래프(undirected graph), 방향 그래프(directed graph), 계량 그래프(valued graph)의 3가지 유형이 있다. 비방향 그래프는 관계의 유무에 초점을 맞춘 그래프이며 관계의 시작과 끝에 대한 방향은 없고 오직 관계의 존재여부에 대한 정보를 담고 있다. 방향 그래프는 관계의 유무뿐만 아니라 관계의 방향을 화살표를 이용해 동시에 표현하는 그래프로 관계의 시작(sender)과 끝(receiver)을 표현하기 때문에 방향 그래프는 비방향 그래프보다는 더 구체적인 정보를 제공한다. 계량 그래프는 관계의 강도(intensity)에 대한 정보를 나타내는 그래프이다. 관계의 강도는 보통 빈도 및 접촉의 빈도 등을 의미하는 것으로 나타난다. 그래프 이론에서는 점/라인, 궤적/경로/사이클, 컴포넌트/연결점/브릿지, 거리/최단거리/지름, 이웃점/연결정도/인디그리/아웃디그리, 연결정도의 평균과 분산에 대한 중요 용어들에 대해 이해해야 한다[1].

네트워크 구조를 파악하기 위해 사회 네트워크 분석에서 사용되는 측정지표로는 중심도(Degree Centrality), 사이도(Betweenness Centrality), 인접도(Closeness Centrality), 위세도(Eigen Centrality), 페이지랭크(Page-Rank)가 있다[1].

표 1. 네트워크 분석 측정지표
Tab. 1. Network Analysis Metric

측정치표	정의
중심도	연결정도중앙성값은 네트워크를 구성하는 하나의 노드가 다른 노드들과 가지는 직접적인 연결 관계의 정도를 나타낸다.
사이도	사이중앙성은 다른 노드들을 연결시키는 매개성을 나타내고, 높은 사이중앙성값을 갖는 노드는 네트워크의 흐름에 중요한 영향을 미치는 중개자 역할을 한다.
인접도	인접중앙성값은 노드가 네트워크 안에서 지리적으로 중심부에 위치하는지를 나타내고, 인접중앙성 값이 높은 노드는 네트워크에서 발생하는 상황을 잘 관찰할 수 있다
위세도	연결정도중앙성이 연결 수를 중시하고, 사이중앙성이 남들 사이에서 브로커의 역할을 중시한다면, 위세중앙성값은 연결된 상대방의 중요성에 가중치를 나타낸다.
페이지랭크	페이지랭크값은 각 결점이 자신의 나가는 링크(out-link)들에 골고루 자신의 중요성을 분배한다고 가정하고 반복적으로 각 결점의 중요성을 계산한 값을 나타낸다.

사회 네트워크분석에 대한 소프트웨어는 통계학 패키지인 SAS나 SPSS처럼 표준화된 프로그램은 없지만, 대표적으로 사용되는 소프트웨어로는 Pajek, PapperPuppy, Polyphonet, TeCFlow, NodeXL 등 70여 개 이상이 있다[4, 10].

본 연구에서의 사회네트워크분석의 소프트웨어로 자동화(SNS를 통한 데이터 추출의 자동화 여부)가 가능하고 엑셀을 기반으로 접근성(일반 사용자의 접근 가능 여부)이 높고, 사회네트워크의 다양한 시각화가 가능한 NodeXL을 분석 소프트웨어로 활용하고자 한다. 그리고 기본적으로 NodeXL에서는 트위터, 프리커, 유튜브 등의 SNS 클롤러(데이터수집기)를 제공한다. 사용자의 편의성을 한층 더 생각한다고 판단된다.

3. 분석을 위한 데이터 및 분석 도구

3.1 네트워크 분석도구 NodeXL

NodeXL은 엑셀의 Add-In 프로그램으로 작동하는 네트워크 분석 도구로서 SNS 서비스를 제공하는 웹사이트로부터 정보를 추출하거나 Outlook의 이메일 데이터를 이용하여 개인 네트워크 분석을 할 수 있도록 지원해주고 있다. 엑셀을 기반으로 하였기 때문에 다른 네트워크 분석 도구에서 사용한 데이터도 재사용할 수 있으며 Node-Link형태 외에 엑셀 기능을 활용한 다양한 다이어그램으로 표현이 가능하다[9].

NodeXL은 시스템 업데이트가 계속 진행되고 있으며 그 유용성이 증대되고 있다. NodeXL은 많이 사용되는 Excel 스프레드시트에 차트형식으로서 네트워크 그래프를 추가해주며, 초보자나 전문가 모두 쉽게 네트워크분석을 하도록 만들어진 소프트웨어이다[13].

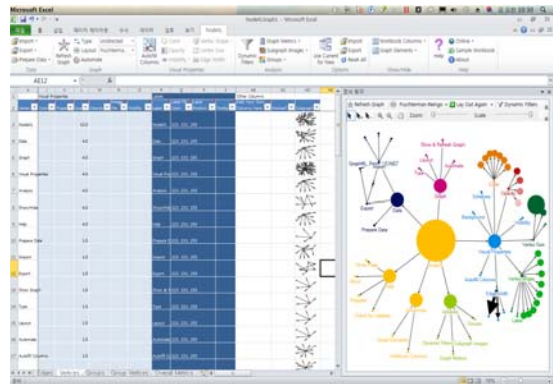


그림 2. NodeXL 화면
Fig. 2. User Interface of NodeXL

네트워크 분석에서 점으로 표시된 것이 개체(노드)이며 데이터에는 node로 표기되며 NodeXL에서는 Vertex로 표현된다. 이 노드의 표현 속성으로는 모양, 색, 크기, 사진으로 표현이 가능하며 값(value)은 측정지표 값을 사용할 수 있다.

점과 점을 이어주는 선을 연결선(링크)이라 한다.

edge로 데이터에 표시되며 NodeXL에서도 똑같이 표현된다. 링크는 연결, 방향, 굵기, 숫자 값으로 표현한다.

3.2 페이스북 데이터 수집

페이스북은 제한적 오픈 소스이다. 일정조건을 충족하면 누구나 페이스북을 기반으로 앱, 서비스를 개발할 수가 있다. 간단한 버튼, 자신의 위젯은 물론이고 Social Plugins형태의 대화형 서비스 그리고 모바일 앱까지 다양하게 제공하고 있다.

트위터, 플리커, 유튜브 등과 달리 NodeXL에서 제공하는 페이스북 클롤러가 없어서 Bernie Hogan이 만든 페이스북 어플리케이션을 이용한다[11]. 개인의 네트워크를 얻기 위한 것임으로 로그인 이 필요하다. 본 연구에서는 개인의 네트워크와 조직(학교)의 네트워크를 비교하고자 하였다. 그리하여 국내 K대학의 페이스북 담당자의 협조를 받아 데이터를 수집할 수 있었다.

표 2. GraphML 소스 일부
Tab. 2. GraphML Source Code(selected)

```
<?xml version="1.0" encoding="UTF-8" ?>
- <graphml xmlns="http://graphml.graphdrawing.org/xmlns"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://graphml.graphdrawing.org/xmlnshttp://
graphml.graphdrawing.org/xmlns/1.0/graphml.xsd">
<key id="sex" for="node" attr.name="sex" attr.type="string" />
<key id="pic" for="node" attr.name="pic" attr.type="string" />
- <graph id="G" edgedefault="undirected">
- <node id="JaeXXXCho">
  <data key="sex">male</data>
  <data key="pic">http://profile.ak.fbcdn.net/hprofile-
ak-snc4/186959_5172538960000000.jpg</data>
</node>
- <node id="EmilyHwakyungKwon">
  <data key="sex">female</data>
  <data key="pic"> http://profile.ak.fbcdn.net/hprofile
-ak-snc4/203322_535411400000000jpg</data>
</node>
<edge source="김O중" target="WooXXXChoi" />
<edge source="김O중" target="KimXXX" />
<edge source="박O우" target="김O중" />
<edge source="WooXXXChoi" target="JeonXXXXHee" />
<edge source="WooXXXChoi" target="SonXXXwan" />
</graph>
</graphml>
```

페이스북 어플리케이션을 활용하여 자신의 친구 네트워크를 GraphML혹은 XML의 형태로 수집한다. 표 2는 페이스북 어플리케이션에서 제공받은 파일의 소스 일부분이다.

내용은 개체(노드)가 되는 <node id = "****">와 연결선 (링크)이 되는 <edge source = "****" target = "000"/> 내용으로 구성 되어있다.

3.3 분석 프로세스

본 연구에서는 다음과 같이 네트워크 분석을 하고자 한다. 페이스북 어플리케이션 Name Get Web 10을 활용하여 분석 대상의 페이스북에서 친구관계를 GraphML 혹은 XML을 수집한다. 수집된 데이터를 NodeXL을 활용하여 분석을 실시한다. 분석은 시각적 분석과 계량적 분석으로 구분한다. 시각적 분석에는 네트워크의 시각적 패턴을 발견하고자 한다. 시각적 패턴 발견의 가독성을 높이기 위한 분석 옵션을 변경하며 진행한다. 계량적 분석은 네트워크 분석에서 산출되는 측정지표 중 Degree, Betweenness, Closeness, Page-Rank를 정리 활용한다. 마지막으로 분석대상의 비교 및 종합적인 내용을 기술하고자 한다.

표 3. 분석 프로세스
Tab. 3. Analysis process

데이터수집	<ul style="list-style-type: none"> • 페이스북 앱을 활용
시각적분석	<ul style="list-style-type: none"> • 기본적인 네트워크 패턴발견을 위한 분석 • Fruchterman-Reingold Option(non)
통계적분석	<ul style="list-style-type: none"> • 네트워크 측정지표 활용 • Degree, Betweenness, Closeness, PageRank
해석	<ul style="list-style-type: none"> • 네트워크 내의 패턴발견 • 개인과 조직 네트워크 비교를 통한 패턴발견

NodeXL 사용을 위한 내용은 박한우 교수의 강의 자료내용과 Derek Hansen의 책을 참고하여 실시하였다[8].

4. 분석 및 해석

4.1 네트워크 분석 시각화

먼저 본 연구는 특정 시점의 페이스북 친구 데이터를 사용하였으며, 네트워크 비교를 위한 대상으로 일 반일(개인)과 국내 K대학교(조직)를 각각 수집하였다.

총 친구의 수는 개인은 192명, 학교는 2,519명의 친구와 관계를 맺고 있으며, 총 링크의 수는 개인은 956, 학교는 36,145개의 링크를 형성하고 있었다. 표 4는 NodeXL을 사용하여 분석한 기본 결과를 개인과 학교를 비교 정리하였다.

표 4. Overall Metrics 개인과 학교

Tab. 4. Overall Metrics(Individual and “K” university)

Graph Metric	개인	학교
Graph Type	Undirected	Undirected
Vertices	197	2519
Unique Edges	956	36145
Edges With Duplicates	0	0
Total Edges	956	36145
Self-Loops	0	0
Connected Components	19	89
Single-Vertex Connected Components	16	86
Maximum Vertices in a Connected Component	171	2429
Maximum Edges in a Connected Component	945	36143
Maximum Geodesic Distance (Diameter)	8	8
Average Geodesic Distance	3.378553	3.32168
Graph Density	0.04951828	0.01139710
NodeXL Version	1.0.1.172	1.0.1.172

네트워크 분석에서는 일반적인 네트워크 분석에서는 중심에 있는 것이 중요하다. 그러나 분석의 특성상 분석의 중심에는 항상 분석대상이 있으므로 그래프에는 자기 자신(ego)을 빼고 분석을 한다. 이미 모든 친구와 연결되어 있는 중심(ego)을 빼면 친구들의 관계를 더 잘 나타내 주기 때문이다[8].

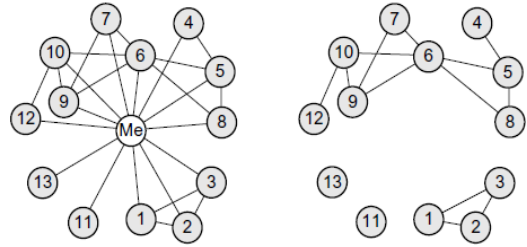


그림 3. 관계를 명확하기 위한 나 빼기

Fig. 3. The same ego network with and without ego. The second visualization (excluding ego) show the structure of the network more clearly

기본적인 분석은 Degree, Betweenness, Closeness, Eigenvector, PageRank와 Group by Cluster를 실시 하였다. 그래프 레이아웃 선택은 Harel-Koren Fast Multiscaling와 Fruchterman-Reingold를 사용하며 Lay-out options을 Iteration(반복) 그리고 Repulsion(노드 사이 저항값)을 수정(100, 3.3)하여 설정한다[8].

4.2 시각적 네트워크 분석 내용

분산된 노드들은 친구를 나타내고 있으며 링크는 친구관계를 나타내고 있다.

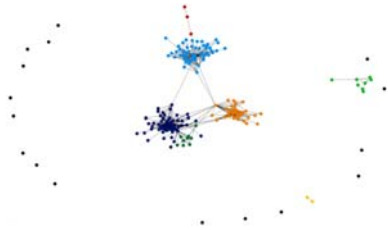


Layout Options:
Style; Layout the entire graph in the entire graph pane
Fruchetman-Reingold Layout: Repulsive:3.3, Iterations:10

그림 4. 개인의 네트워크 분석

Fig. 4. Network Analysis of the Individual

그림 4는 개인의 페이스북 친구 관계를 네트워크 그래프로 표현하였다. 일정한 패턴 없이 무질서하게 있는 것을 확인할 수 있다.



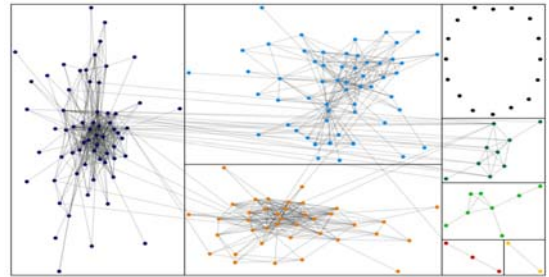
Layout Options:
Style: Layout the entire graph in the entire graph pane
Fruchetman-Reingold Layout: Repulsive:3.3, Iterations:100

그림 5. 개인의 네트워크 분석(그룹)
Fig. 5. Network analysis of the individual(group)

그림 5의 가장자리에 있는 검은 점들은 네트워크 내에서 다른 노드들과 관계를 하지 않은 유일한 존재들이다. 그림의 중심을 보면 세 개의 큰 하위그룹을 발견할 수 있다. 실제 각 그룹의 특징을 설명하면 상단의 파란색그룹은 대학 동아리와 관련 있는 친구들이다. 왼쪽 하단의 남색은 현재의 직장과 관련 있는 그룹이었다. 오른쪽 황색그룹은 최근 교육을 함께 받은 사람들이다. 그리고 그림의 오른쪽 녹색그룹은 지방에 있는 고등학교 동창들이다.

그림 6은 각 하위그룹을 명확히 분석하기 위해 각 그룹을 박스형태의 구역으로 구분하였다. 이런 각 하위그룹을 개별적으로 구분하여 분석하는 것을 컴포넌트 분석(Component Analysis)이라고 한다[1]. 이를 위해 그래프 옵션의 설정을 변경하였다. 변경된 옵션값은 그림의 하단에서 확인할 수 있다.

먼저 그림 6에서 확인할 수 있는 것은 7개의 그룹이 존재하는 것을 파악 확인할 수 있다. 먼저 대학(파란), 직장(남색), 교육(황색)의 네트워크 유형은 스타형을 나타내고 있다. 이는 네트워크 내에 중심역할자가 존재하는 것을 알 수 있다. 그리고 고교(녹색)는 Y형 네트워크의 형태를 보이고 있다. 이는 스타형과 동일하게 중심이 존재하는 것을 알 수 있다. 또한, 씨클형과 체인형의 유형도 함께 있음을 알 수 있다.

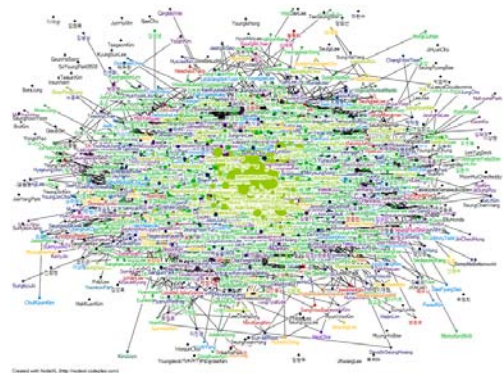


Layout Options:
Style: Layout each of the graph's groups in its own box and sort the boxes by group size, Box outline: 1
Fruchetman-Reingold Layout: Repulsive:3.3, Iterations:100

그림 6. 개인의 컴포넌트 분석
Fig. 6. Network analysis of the individual(component)

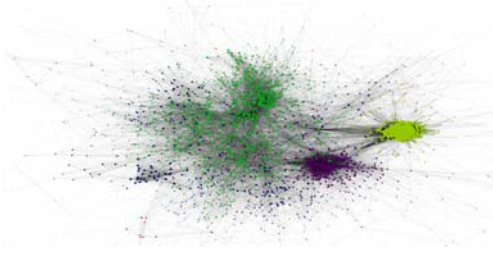
그림 7은 국내 K대학교에서 운영하는 페이스북에 등록된 친구의 네트워크 분석 화면이다. 개인 사용자에 비하여 노드의 수는 약 12배 링크의 수는 35배 많다는 것이 한눈에 보인다.

그림 8은 그림 5와 같이 그래프 옵션을 수정하여 하위그룹의 표현 하고자 하였다. 왼쪽의 연두색과 보라색의 두 하위그룹이 밀도 있게 발견되었다. 하지만 다른 하위 그룹을 확인하는 것은 어려워 보인다. 이를 극복하고자 컴포넌트 분석을 그림 9와 같이 실시하였다.



Layout Options:
Style: Layout the entire graph in the entire graph pane
Fruchetman-Reingold Layout: Repulsive:3.3, Iterations:10

그림 7. 학교의 네트워크 분석
Fig. 7. Network analysis of the "K" university

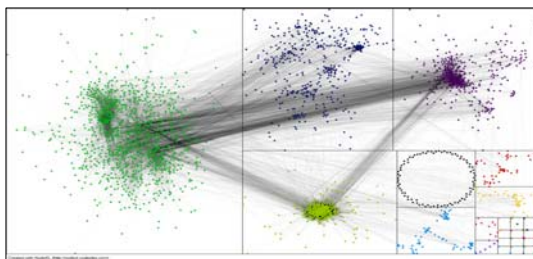


Layout Options:
 Style; LayOut the graph's connected components in Boxes at the bottom of the graph pane.
 Maximum size: 10, Box size: 16
 Fruchetman-Reingold Layout: Repulsive:3.3, Iterations:100

그림 8. 학교의 네트워크 분석(그룹)

Fig. 8. Network analysis of the “K” university(group)

그림 9에서 기대한 하위그룹의 유형을 발견하고자 하였으나 하위그룹 간의 링크가 많아 확인이 어렵다. 단지 22개의 하위그룹만 확인이 가능하다. 그리하여 각 하위집단 내의 링크만 유지하고 하위그룹 간의 링크를 생략하여 그림 10과 같이 수정된 컴포넌트 분석을 실시하였다.



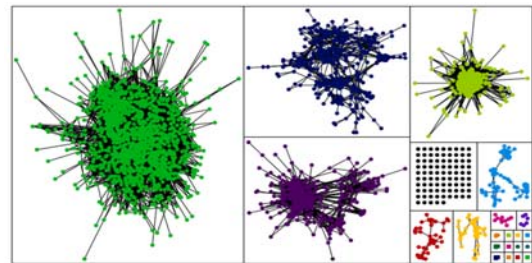
Layout Options:
 Style; Layout each of the graph's groups in its own box and sort the boxes by group size, Box outline: 1
 Fruchetman-Reingold Layout: Repulsive:3.3, Iterations:100

그림 9. 학교의 컴포넌트 분석

Fig. 9. Network analysis of the “K” university(component)

그림 10에서는 각 하위그룹의 뚜렷하게 네트워크 유형을 확인할 수 있다. 순서대로 확인하자면 앞의 4개의 네트워크는 스타형이라고 말할 수 있겠다. 그러나 오른쪽 하단에 있는 나머지 하위 그룹은 Y 유형의 모습을 하고 있다. 그러나 아쉽게도 개인 네트워크와는 달리 각 하위 그룹의 속성을 정의하기가

어렵다. 이유인즉, 각 하위 그룹은 단순한 인구 통계학적인 내용으로 설명하기 어렵기 때문이다. 하지만 본 연구에서 확인할 수 있는 내용은 분명히 각 하위 그룹의 특징은 발견할 수 있었다.

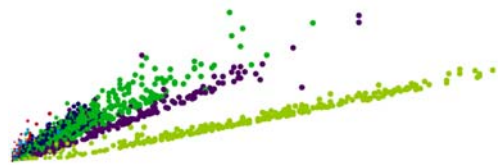


Layout Options:
 Style; Layout each of the graph's groups in its own box and sort the boxes by group size, Box outline: 1
 (Hide the intergroup edges, Use the Grid Layout for groups that don't have many edges)
 Fruchetman-Reingold Layout: Repulsive:3.3, Iterations:100

그림 10. 학교의 수정된 컴포넌트 분석

Fig. 10. Network analysis of the “K” university(component2)

그 특징을 발견하기 위해 네트워크 분석을 그림 11과 같이 실시하였다. 기존의 노드를 가로(X축)에는 밀도(연결정도), 세로(Y축)에는 페이지랭크 값으로 분산형 그래프를 만들었다. 이를 위해 별도의 그래프 작업을 실시한 것이 아니라 그래프 옵션을 아래와 같이 수정 변경하여 사용하였다.



Visual Properties Columns options:
 Vertex X: Degree, Vertex Y: PageRank ,
 Layout style: None
 Fruchetman-Reingold Layout: Repulsive:3.3, Iterations:100

그림 11. 밀도와 페이지랭크 분석

Fig. 11. A Comparative Analysis of Degree and PageRank.

그림 11에서 오른쪽 하단의 연두색 하위그룹은 페이지랭크가 낮은 것을 알 수 있으며, 길게 연결된

모양을 하는 것을 알 수 있다. 이를 다른 하위 그룹과 비교하자면 밀도가 높다는 것은 많은 링크를 가진다는 것이고 페이지랭크가 낮다는 것은 그 링크의 다른 노드가 또 다른 노드로 링크가 낮음을 말하는 것이다. 이는 연두색 하위 그룹은 그룹 내의 관계가 강한 폐쇄형 혹은 서클유형의 하위 그룹일 가능성이 크다.

4.3 네트워크 계량적 분석

시각적 네트워크 분석은 네트워크 전체의 모습을 조망하는 것에는 적절하나 세부적인 내용이나 개별 개체의 특성을 파악하기란 쉽지 않다. 그리하여 네트워크 측정지표를 활용한 분석을 하고자 한다.

표 5는 개인과 학교의 페이스북 사용자 네트워크 분석 측정지표를 요약 비교한 내용이다. 표 4에서

확인했듯이 노드의 수적차이가 중심도, 사이도의 값에 나타나는 것을 확인할 수 있다. 수치의 차이를 그대로 반영하기 어려워 이하 표에서는 상대적 값을 이용하여 표기 하였다.

표 5. 측정지표 요약 비교
Tab. 5. Metrics(Individual and “K” university)

측정지표		개인	조직
중심도	Minimum	0	0
	Maximum	42	510
	Average	9.705	28.698
	Median	7	15.000
사이도	Minimum	0.000	0.000
	Maximum	4,599.006	913,907.415
	Average	177.396	2719.424
	Median	8.821	795.143
인접도	Minimum	0.000	0.000
	Maximum	1.000	1.000
	Average	0.015	0.002
	Median	0.002	0.000

표 6. 개인의 네트워크 분석 각 측정지표 상위 10

Tab. 6. Metrics of the Individual(Top 10)

순위	ID	연결 중심도	그룹	ID	사이도	그룹	ID	인접도	그룹
1	102	0.214	G1	19	0.241	G6	17	1.000	G7
2	104	0.204	G1	84	0.231	G2	18	1.000	G7
3	103	0.184	G1	127	0.174	G1	48	0.111	G4
4	84	0.173	G2	35	0.141	G2	131	0.091	G4
5	98	0.163	G1	77	0.059	G1	129	0.083	G4
6	121	0.158	G1	103	0.054	G1	130	0.083	G4
7	110	0.153	G1	42	0.052	G3	46	0.077	G4
8	115	0.153	G1	104	0.050	G1	163	0.067	G4
9	43	0.148	G1	43	0.041	G1	156	0.059	G4
10	62	0.143	G6	79	0.040	G1	47	0.053	G4

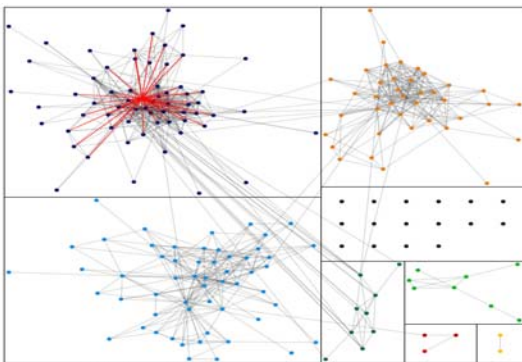
표 7. 조직의 네트워크 분석 각 측정지표 상위 10

Tab. 7. Metrics of the “K” university(Top 10)

순위	ID	연결 중심도	그룹	ID	사이도	그룹	ID	인접도	그룹
1	544	0.203	G4	544	0.288	G4	990	1.00000	G19
2	1076	0.092	G4	160	0.033	G4	989	1.00000	G19
3	2055	0.088	G8	1076	0.029	G4	2434	1.00000	G3
4	1989	0.086	G8	48	0.025	G4	2433	1.00000	G3
5	1836	0.086	G8	45	0.016	G4	544	0.00021	G4
6	1845	0.085	G8	524	0.015	G4	160	0.00017	G4
7	1852	0.085	G8	233	0.014	G4	1076	0.00017	G4
8	1861	0.082	G8	748	0.014	G4	45	0.00017	G4
9	785	0.082	G8	1237	0.013	G4	1237	0.00017	G4
10	2000	0.080	G8	676	0.011	G8	748	0.00017	G4

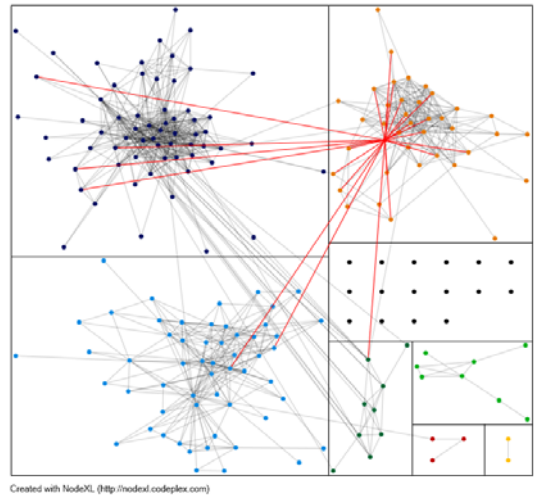
표 6은 개인 네트워크 분석 측정지표별 상위 10개를 비교하였다. 표 6에서 연결중심도(degree)가 가장 높은 사람은 “ID102”임을 알 수 있다. 중심도가 높다는 것은 많은 노드와 직접연결을 하고 있다는 것이다. 그리고 사이도가 가장 높은 사람은 “ID19”이다. 사이도(Betweenness)는 연결된 노드들이 다시 다른 노드들과 연결할 수 있는 것을 말한다. 단순히 생각하면 중심도가 높으면 네트워크의 가치가 높을 것이라고 생각할 수 있다. 하지만 네트워크의 가치를 단순 노드의 직접 연결만으로 판단하면 오류에 빠지기 쉽다. 이는 그림 12와 그림 13에서 잘 보여주고 있다. 중심도 1위인 “ID102”는 한 그룹 내에서만 링크를 가지는 것을 볼 수 있다. 앞 절에서 실시한 시각적 분석에서는 발견 할 수 없었던 사실을 통계적 기법과 함께 사용하여 새로운 사실을 쉽게 발견 할 수 있었다. 그리고 사이도 1위인 “ID19”는 다른 그룹과 링크를 가지고 있음을 확인할 수 있다.

표 6의 인접도(Closeness)순위에서 살펴보면 상위 2개의 값이 1임을 알 수 있다. 그림 13의 오른쪽 하단의 노란색 2노드간의 연결을 말하는 것이다. 단순 수치만으로 본다면 “ID17”, “ID18”이 네트워크에서의 친밀도가 높아 보이지만 그래프와 함께 분석 할 때 그 이유를 쉽게 알 수 있다.



Layout Options:
Style: Layout each of the graph's groups in its own box and sort the boxes by group size, Box outline: 1
Fruchetman-Reingold Layout: Repulsive:3.3, Iterations:100

그림 12. 중심도 1위 네트워크 그래프
Fig. 12. Network graph(Degree Top 1)



Layout Options:
Style: Layout each of the graph's groups in its own box and sort the boxes by group size, Box outline: 1
Fruchetman-Reingold Layout: Repulsive:3.3, Iterations:100

그림 13. 사이도 1위 네트워크 그래프
Fig. 13. Network graph(Betweenness Top 1)

표 7은 학교의 네트워크 분석 측정지표별 상위 10개를 비교하였다. 연결 중심도가 가장 높은 사람이 “ID544”임을 알 수 있고 사이도 또한 “ID544”임을 알 수 있다. 이를 표 6과 비교해본다면 특이한 점을 발견 할 수 있다. 측정지표와 함께 본 연구에서 구분된 하위 그룹을 함께 표기하였다. 학교의 네트워크에서는 사이도와 중심도에 특정 그룹이 집중되어 있는 것을 알 수 있다. 그리하여 표 8에서 각 네트워크에서 형성된 하위 네트워크의 예고와 노드를 비교분석 하였다. 개인은 총 7개의 그룹, 학교는 21개의 그룹으로 분류되었다. 그러나 학교에서 노드수로 정렬하면 보면 하위 14개 그룹은 10 이하의 노드를 가지고 있다. 여기서 주목할 점은 학교에서 G8 그룹으로 분류된 노드의 수는 310개이지만 링크의 수는 무려 14,007로서 G4 보다 월등히 많은 것을 알 수 있다. 이는 하나의 노드가 많은 링크를 가지고 있는 것을 의미한다. 이는 표 7에서 연결 중심도가 높은 노드들이 G8 그룹에 많이 있다는 것을 설명해주고 있다.

표 8. 하위 네트워크(컴포넌트)
Tab. 8. Sub-Network(component)

개인			학교		
그룹	Vertices	Edges	그룹	Vertices	Edges
G1	64	418	G4	1,124	10,837
G2	56	245	G1	412	2,111
G6	39	231	G10	393	5,218
G3	9	19	G8	310	14,007
G4	8	10	G2	71	166
G5	3	2	G5	43	94
G7	2	1	G7	37	73
			G9	10	10
			G11	6	5
			G6	3	3
			G13	3	3
			G15	3	2
			G3	2	1
			G12	2	1
			G14	2	1
			G16	2	1
			G17	2	1
			G18	2	1
			G19	2	1
			G20	2	1
			G21	2	1

5. 연구의 한계와 기대효과

본 연구는 많은 기업과 개인이 소셜네트워크 서비스를 사용함에 있어 단순한 정보 공유가 아닌 “지속적인 상호작용”을 위한 네트워크 모형과 그 특성을 확인 하고자 하였다.

먼저 소셜네트워크 서비스인 페이스북을 사용하는 사용자의 패턴 발견을 위한 사회네트워크분석을 실시하였다. 분석을 위한 분석도구는 SNS분석에 적합하고 엑셀 Add-In형식의 프로그램으로 쉬운 인터페이스와 분석을 위한 확장성이 뛰어난 NodeXL을 사용하였다. 분석을 위한 데이터 수집은 페이스북 어플리케이션을 활용하여 획득할 수 있었다. 하지만 개인 로그인을 설정해야 함으로 다양한 사례를 수집하지 못한 것이 한계점으로 작용한다. 그럼에도 개인 사용자의 페이스북 네트워크 분석과 학교에서 운영하는 페이스북의 친구 네트워크 분석은 큰 의미가 있다고 판단된다.

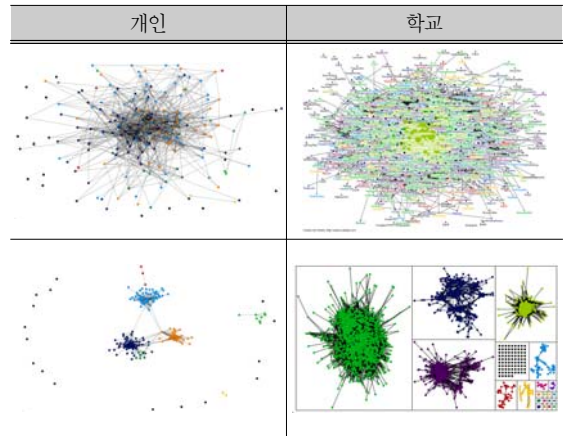


그림 14. 개인과 학교의 네트워크 분석
Fig. 14. Network graph Individual and “K” university)

위 그림 14에서도 나타나듯이 개인과 학교의 네트워크 유형과 하위 그룹의 구성을 한눈에 확인 할 수 있었다. 하지만 노드의 수가 증가하고 링크의 수가 증가함에 따라 네트워크의 판별을 위한 모형이 분명하지 않아 추가적인 분석이 필요하였다.

분석결과 형성된 네트워크는 하위 그룹이 존재하며 그 하위그룹은 특정 연결 관계를 가지는 것을 알 수 있었다. 개인의 페이스북 네트워크는 직장, 대학 동기, 고교동창과 같은 개인의 일상 및 과거의 관계를 알 수 있는 단초를 기반으로 하위그룹을 해석할 수 있었다. 그러나 학교의 페이스북 네트워크에서는 세부적인 특성(개인정보)을 활용하여 네트워크를 해석할 수 없었다. 하지만 추가분석을 통하여 네트워크 특성을 구분 하고자 하였다. 그리하여 패쇄적이거나 서클형의 네트워크를 가지고 있음을 확인 할 수 있었다.

또한 본 연구에서는 네트워크 분석을 심도 있게 활용하기 위해서는 시각적인 분석과 함께 측정지표를 활용한 통계적 분석을 병행하는 것이 의미 있다고 주장한다. 이는 각 네트워크상에서 “누가 관심을 많이 받고 있는가?”를 알고 싶다면 중심도(degree)를 지표로 활용할 수 있을 것이다. 그러나 단순히 하나의 지표만으로 네트워크를 이해하는 것은 위험하다.

표 6에서 확인 했듯이 각 측정지표의 의미에 따라 네트워크의 가치매김이 달라 질 수 있다.

본 연구의 한계점으로 먼저 네트워크 분석을 통한 네트워크 유형의 구분은 가능하였으나 네트워크의 구체적인 특성을 찾아내지를 못하였다. 이를 극복하기 위해서는 추가적인 분석과 하위 그룹의 형성과정을 이해 할 수 있는 개인정보가 필요하다는 것을 확인 할 수 있었다. 이에 본 연구는 다차원적인 데이터 수집과 분석이 필요하다고 제안 한다. 예를 들어 시간차원을 추가한다면 과거부터 현재까지의 네트워크 형성 과정을 시뮬레이션 할 수 있을 것이다.

향후 연구에서는 SNS의 다차원적인 데이터수집 및 분석을 위한 통합적 분석 방법을 제시하고자 한다. 또한 활용 가능한 분석도구를 지속적으로 연구에 반영하여 분석의 가용성을 높이고자 한다. 이를 통하여 지속적인 상호작용이 가능한 네트워크를 제안하고자 한다.

참 고 문 헌

[국내 문헌]

- [1] 손동원 (2002), “사회 네트워크 분석”, 경문사.
- [2] 안정민 (2011), “SNS 사용자의 이용 행태와 성격이 정서적 유대감 형성에 미치는 영향에 대한 연구”, 아주대학교 경영대학원, 석사학위논문.
- [3] 양윤직 (2011), “TGIF 스토리”, 커뮤니케이션북스 03.
- [4] 이찬균 (2011), “사회네트워크분석을 통한 e-커머스 기업 고객서비스 트렌드 진화에 관한 연구”, 한국외국어대학교, 박사학위논문.
- [5] 정기한, 정지희, 신재익 (2010), “기업의 SNS 활동이 브랜드 및 구전, 구매의도에 미치는 영향”, 한국경

영학회, 통합학술대회.

- [6] 최고은 (2010), “사이버 공동체와 정치참여에 대한 기술철학적 접근: 사회구성론을 중심으로”, 한국교원대학교, 석사학위논문.

[국외 문헌]

- [7] Aleman-Meza, B., Nagarajan, M., Makrishnan, C., Ding, L., Kolari, P., Sheth, A. P., Arpinar, I. B., Joshi, A., and Finin, T. (2006), “Semantic Analytics on Social Networks: Experiences in dressing the Problem of Conflict of Interest Detection”, In Proceedings of the 15th International Conference on WWW, 407-416.
- [8] Hansen, Derek. Shneiderman, Ben. Smith, Marc (2010), “Analyzing Social Media Networks with Nodexl”, Morgan Kaufmann Publishers.
- [9] Marc A. Smith, Ben Shneiderman, Natasa Milic-Frayling, Eduarda Mendes Rodrigues, Vladimir Barash, Cody Dunne, Tony Capone, Adam Perer, Eric Gleave (2009), “Analyzing (Social Media) Networks with NodeXL,” In C&T '09: Proc. Fourth International Conference on Communities and Technologies, Lecture Notes in Computer Science.

[웹사이트]

- [10] http://en.wikipedia.org/wiki/Social_network_analysis_software.
- [11] <http://apps.facebook.com/namegenweb>.
- [12] <http://en.wikipedia.org/wiki/Facebook>.
- [13] <http://nodexl.codeplex.com>.
- [14] <http://www.etnews.com/201106130210>.



하 병 국 (ByungKook Ha)

동국대학교 경제학 학사, 광운대학교 경영대학원 경영학 석사학위를 취득하였다. 현재 광운대학교 대학원에서 경영정보학 전공으로 박사 과정에 재학 중이다. 주요 관심분야는 비즈니스 인텔리전스, CRM, SNA 등이다.



장 용 수 (Yongsoo Jang)

전남대학교 사범대학 학사, 고려대학교 교육대학원 전산교육 석사학위를 취득하였다. 현재 광운대학교 대학원에서 경영정보학 전공으로 박사 과정에 재학 중이다. 주요 관심분야는 CRM, OLAP, 교육 데이터 자산의 전략적 활용 모델 등이다.



조 재 희 (JaeHee Cho)

연세대학교에서 경영학 학사, Miami University (Ohio)에서 경영학 석사, University of Nebraska at Lincoln에서 경영정보학 박사를 취득하였다. 펜타시스템테크놀로지(STM(LG CNS의 전신)에서 컨설턴트로 근무하였고, 현재 광운대학교 경영학부 교수로 재직 중이다. 주요 관심분야는 데이터자산의 전략적 활용, 다차원모델링, OLAP, 데이터웨어하우스, 비즈니스 인텔리전스 등이다.

Users Pattern Discovery of Social Network Service by Social Network Analysis : Focusing on Facebook

ByungKook Ha* · Youngsoo Jang* · JaeHee Cho**

ABSTRACT

Companies see a new business opportunity in the increased popularity of social network services and the related studies are also gaining more attention. This study attempted to analyze the social networks and thereby find a pattern in the use of social network services. Network users' pattern has been categorized by their purpose of use. Among various social network services, we selected the Facebook and its users were analyzed by a network analysis tool called NodeXL. In the end, several subgroups have been identified in a seemingly homogeneous network.

Furthermore, the network shape differences according to the usage of social network services has been studied by comparing "friends" of an individual Facebook user with those of the K University Facebook page.

Keywords: Social Network Services, Social Network Analysis, NodeXL

* Graduate School, Kwangwoon University

** School of Business, Kwangwoon University