

개인 미디어 이용자의 친구 맺기 유사성에 대한 연구:

네트워크, 사이버 커뮤니티, 클릭 패턴*

장덕진**, 천명규***, 김기훈****, 배 영*****

요약

본 연구는 사이버 공간에서 이루어지는 관계의 형성과정과 특성을 ITWAREHOUSE의 인택서비스에서 추출된 트래이스 자료와 회원들에 대한 설문조사를 이용하여 분석하였다. 이를 위해 분석단위를 개별 행위자가 아닌 행위자들 간의 이차관계로 설정하여 관심인맥 지경의 유사성에 대한 LSDV(Least Squares with Dummy Variables) 회귀분석을 시도하였다.

분석결과는 관심인맥의 연결망에서 비슷한 위치에 있는 사람들일수록, 사이버 커뮤니티에서의 활동경력이 비슷할수록, 그리고 사이트 내에서의 클릭 패턴이 비슷할수록 비슷한 상대를 선택하는 경향이 높다는 점을 보여준다. 분석결과는 개인미디어 서비스 기업의 입장에서 보면 회원들의 관계데이터를 이용한 새로운 마케팅 전략을 수립가능하게 하고, 학문적으로는 인터넷에서 활동성이 높은 이용자들을 대상으로 타겟 분석하여 기존의 무작위 표본추출이 가진 한계를 극복할 수 있는 가능성을 제시하고 있다.

키워드: 개인 미디어, 블로그, 네트워크, 사이버 커뮤니티, 클릭 패턴

I. 서론

한 사회 속에서 개인들의 삶은 자연적 혹은 의도적인 타인들과의 관계와 함께 이루어지게 된다. 정보통신 기술의 발전으로 인해 사회적 관계의 방식과 폭이 커지면서 전통적인 면대면(face to face)관계 뿐만 아니라 인터넷에 기반한 새로운 형식의 관계에 대해 활발한 논의가 이루어지고 있다(Wellman, 2001a; 2001b; Wellman & Gulia, 1999). 인터넷을 통한 새로운 관계의 형성과 유지는 오프라인보다 행위자에게 여러 가지 편의성을 제공하고, 관계 자체에 대한 사회적 제약이나 제재가 약할 수 있다는 점에서

훨씬 더 자유롭게 이루어진다. 더욱이 인터넷에서는 다양한 배경을 가진 사람들의 자유로운 관계형성 뿐만 아니라 자신의 관심과 목적에 따라 집단을 형성하게 되었는데, 상대방에 대한 오프라인에서의 경험을 공유하지 않은 상황에서도 자신과 조직에 대한 정체성의 확인과 수립이 가능하게 되는 양상에까지 나아가게 되었다(배 영, 2003a).

한국의 경우 초고속 통신망의 확산과 함께 사이버 공간이 보편화됨에 따라 인터넷에서의 새로운 관계 맺음은 이제 일상 속에 매우 중요한 요소로 작용하고 있다. ‘모교사랑(www.iloveschool.co.kr)’과 같은 관계형 커뮤니티 서비스가 90년대 말에서 2000년대

2006년 2월 20일 접수, 2006년 3월 13일 심사, 2006년 5월 22일 심사완료

* 본 연구는 한국학술진흥재단의 협동연구지원(2003-042-B00105)에 의해 이루어졌다. 그 동안 부분적인 연구결과는 여러 차례의 학술대회 및 워크숍 등을 통해 발표되었으며, 이 과정에서 많은 분들의 조언이 큰 도움이 되었다. 이 자리를 빌어 감사의 말씀을 드린다.

** 제1저자, 서울대학교 사회학과 교수 (dukjin@snu.ac.kr)

*** (주)잇이즈컴 대표이사 (chunmk@itiscom.co.kr)

**** (주)사이람 대표이사 (ghghim@cyram.com)

***** 교신저자, 숭실대학교 정보사회학과 교수 (ybae@ssu.ac.kr : 02-820-0047)

초까지 사이버 트렌드를 대표하면서 인터넷의 확산에 한 요인으로 작용했다면 2004년부터 현재까지 인터넷 이용자들에게 가장 관심을 끌고 있는 ‘미니홈피’ 서비스나 각 포털을 중심으로 한 블로그 서비스도 자신의 관심사를 타인과 공유할 수 있는 장으로서의 역할과 함께 사이버 공간에서의 관계 확장과 유지의 중요한 도구로 이용되고 있다. 자신과 관련한 정보를 선택적으로 제시하면서 관계 형성을 도모하고 자신이 필요로 하는 정서적 또는 도구적 만족을 위해 다양한 종류의 관계를 모색하게 된 것이다.

오프라인에서의 사회적 관계는 직접적인 접촉을 매개로 개인이 가진 특질과 속성에 기반한 경우가 많다는 점에서 사이버 공간에 비해 한정된 자원을 대상으로 한다. 때문에 특정한 관계의 형성에 있어 개인의 구체적인 의도도 중요한 요소이지만 개인을 둘러싸고 있는 사회적 배경이 미치는 영향도 크다. 이에 비해 사이버 공간에 있어서는 위의 경험적인 분석에서도 나타나듯이 상대적으로 개인의 의도와 자유로운 선택이 관계 형성에 미치는 영향이 보다 크다고 할 것이다. 다시 말해 추구하는 관계에 대한 선택의 폭이 현실공간에 비해 큰 특성을 갖고 있는 것이고, 이는 가능한 관계의 스펙트럼이 넓게 분포되어 있다는 것이다.

관계의 스펙트럼이 넓게 존재하는 사이버 공간에서도 현실세계와 마찬가지로 관계의 허브(Hub)가 존재한다. 관계에 있어 허브는 상대적으로 많은 연결을 가진 개인이라고 할 수 있다. 개인의 의도와 선택에 따라 관계의 형성이 자유롭지만 가수 ‘비’의 경우처럼 특정한 누군가는 타인에 비해 매우 많은 관계망을 갖게 된다는 것이다. 그렇다면 다음과 같은 질문들이 가능해진다. 어떠한 요인 때문에 특정한 개인은 허브의 위치를 차지하게 되는가? 바꿔 말하자면 특정한 개인을 선택한 많은 사람들은 제한된 정보와 자원 속에서, 그리고 상대방에 대한 인지나 신뢰가 제대로 이루어져 있지 않은 상황에서도 그 사람과 관계를 형성하는 이유는 무엇인가? 그리고 서로 다른 배

경과 사람들이 특정한 개인과의 관계를 공통적으로 선택한 이유는 무엇이고, 이들이 가진 특성은 유사함인가, 상이함인가?

이러한 질문들에 응답하기 위해 본 연구는 현재 사이버 공간에서 이루어지고 있는 사회적 관계의 양상과 동인을 IT 정보관련 업체인 ITWAREHOUSE에서 제공하고 있는 IT인택서비스를 중심으로 살펴보고자 한다. 이를 위해 우리는 IT인택서비스 회원들에 대한 설문조사 결과와 함께 2004년에 시작된 인택서비스의 트레이스 데이터도 관계 분석을 위한 자료로 활용하였다.

II. 사이버 공간의 사회적 관계

인터넷을 통한 커뮤니케이션의 발달로 기존의 관계망에 대한 유지비용이 절감되고 아울러 새로운 관계 형성의 기회 또한 증가하고 있다. 이렇듯 인터넷은 기존의 인간관계 방식이나 사회적 관계망의 형태에 있어서 다양한 변화를 추동하고 있는데, 인터넷을 통한 관계의 형성과 유지는 기존의 방식과 비교할 때 어떤 특성을 나타내고 있으며 향후의 전반적인 변화는 어떻게 나타날 것인가?

이에 대한 연구는 크게 두 가지 상반된 입장으로 요약할 수 있다. 먼저 역사적으로 봤을 때 새로운 기술이나 미디어의 출현과 마찬가지로 인터넷의 이용은 개인들의 물리적으로 제한된 시간을 차지하게 되는 것이고, 이에 따라 기존의 인간관계에 할애하던 시간이 줄어들 수밖에 없기 때문에 기존의 인간관계에 대해선 소홀해진다는 입장과 이와는 반대로 인터넷을 통하여 보다 많은 사람과의 관계 형성이 가능해지고 현재 유지하고 있는 기존 관계도 보다 활성화될 것이라는 낙관적 전망이 그것이다.(Nie, 2001; Orleans and Laney 2000; Howard et al. 2001; Wellman, 2001b; Pruijt, 2002).

올리언과 레이니(Orleans and Laney 2000: 57)는 이러한 상반된 시각을 제로섬(zero-sum)과 비

제로섬(non zero-sum) 사고방식으로 규정한다. 제로섬 시각은 새로운 미디어나 기술을 이용한다는 것은 기존의 시간과 인간관계를 줄여야만 가능하다고 보는 입장이고, 비 제로섬 시각은 오히려 새로운 미디어가 새로운 인간관계의 기회로 작용하여 기존의 시간과 관계를 상쇄시키는 역할을 하는 것이 아니라 새로운 수요의 창출로 연결된다고 보는 것이다. 이들은 가상공간과 현실세계는 서로 보완적인 관계에 있다고 주장하면서 인터넷을 이용한 커뮤니케이션이 기존의 현실 세계의 사회자본(social capital)을 보완하고 강화하는 기능을 하고 있다고 본다. 인터넷을 통한 개인들 간의 상호작용은 그 자체가 네트워크이고, 협동이나 자발성, 공유 등이 기본 속성이며 상대방의 존재를 기반으로 하는 인터넷의 기본적 특성이 사회자본의 특성과 부합하고 있다고 말한다.

인터넷에 기반한 사회적 관계가 현실세계의 관계와 어떤 연관 속에 있는지에 대한 연구는 활발히 진행되어 왔지만 사이버 공간에서 형성되는 관계 자체의 동학에 대한 연구는 상대적으로 매우 제한되어왔다. 인터넷 이용이 일상화되면서 사이버 공간에서의 관계 자체가 차지하는 비중이 커지고 있다는 점을 고려할 때 이에 대한 논의의 필요성 또한 증가하고 있다. 반즈(Barnes, 2001)는 인터넷에서 이루어지는 관계의 양상도 면대면 관계가 크게 다르지 않다고 말하며 상대방과 경험의 공유 여부, 자신을 잘 표현할 수 있는 글쓰기, 서로에 대한 적절한 역할 부여, 그리고 상대방에게 안정감을 주고 교류를 촉진 시킬 수 있는 대화의 존재와 같은 요소들이 관계 형성에 있어 중요한 자원이 된다고 밝힌바 있다. 아울러 심리학적 차원에서 오프라인과 마찬가지로 근접성(proximity)과 이로 인한 친밀성이 가져오는 접촉 빈도(intersection frequency)의 증가, 그리고 자신이 가지지 못한 전문적인 지식의 존재와 같은 요소들이 관계를 유인한다는 논의(Wallace, 1999)도 있지만 관계 전반에 대한 일반화를 시도하기에는 부족함이 많아 보인다. 또한 기존의 현실세계에서 이루어

지는 관계의 특성이나 속성의 연장선상에서 분석이 이루어지고 있다는 점에서 사이버 공간에서의 관계에 대한 실질적인 분석에는 큰 도움을 주지 못하고 있다.

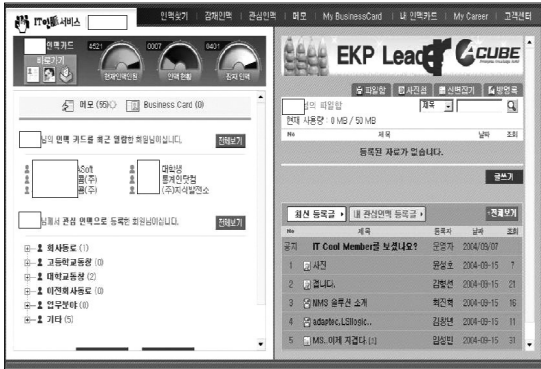
이처럼 사이버 공간의 관계에 대한 본격적인 연구가 어려웠던 것은 대상 자체가 가진 모호함에 기인한 바가 크다. 존스(Jones, 1998)가 지적한 것처럼 온라인에서의 개인과 집단들은 판독하기 어려운 방식으로 상호작용하고, 상호작용의 범위와 대상에 대해 시작과 끝을 설정하기 어렵다는 생각이 지배적이었기 때문이다. 이러한 이유로 본 연구가 갖는 의의는 커진다. 사이버 공간에서 활동하는 행위자들의 행태를 인맥서비스라는 매개를 통해 직접 관찰하고, 그들이 느끼고 있는 인식의 측면 또한 설문 조사 결과를 통해 파악할 수 있다는 점에서 그간 다른 연구들에서도 출해내지 못했던 다양한 사실들에 대한 접근이 가능해진 것이다. 앞으로의 분석을 통해 인지관계가 형성되어있지 않은 개인들 간에 어떠한 요소가 작용하여 사이버 공간에서의 네트워크의 형성과 특정한 대상과의 관계 맺기가 이루어지게 되는지를 실제적이고 다각적인 분석을 통해 알아본다.

Ⅲ. 연구의 대상과 분석 변수 및 방법

1. 연구의 대상

본 연구에서 주요 연구대상이 되고 있는 ITWAREHOUSE사는 데이터웨어 하우스(DWH), 데이터베이스 마케팅(DBM), 고객관계경영(CRM) 분야의 솔루션과 컨설팅 및 마케팅, 그리고 각종 e-비즈니스 관련 솔루션을 제공하는 회사로서, 이 과정에서 얻게 되는 많은 IT관련 정보들을 체계적으로 분류 및 가공하여 회원들에게 제공함으로써 높은 평가와 함께 다수의 회원을 확보하고 있는 기업이다.

2004년 4월에 시작된 IT인맥서비스는 사이트에 등록된 이용자들이 인맥 형성과 관련하여 기초 연결



〈그림 1〉 사 인맥서비스의 실제 예

고리가 될 수 있는 학맥(고교, 대학)과 회사이름, 그리고 관심분야와 같은 기타 사항을 등록하면 자신과 관련이 있는 구성원들을 관심인맥으로 등록할 수 있도록 만들어져있다. 또한 일단 관심인맥으로 등록하면 폴더에 관심인맥의 종류별로 자신의 인맥이 정리되어 있어 관리하기에 편리할 뿐만 아니라 상호 교류할 수 있는 공간이 제공된다. 즉, 인맥서비스는 인맥 관련 정보와 함께 명함보내기과 게시판, 방명록, 자료실, 그리고 사진첩 등의 구성요소를 통해 회원들 상호 간의 대인 커뮤니케이션은 물론 매개 커뮤니케이션의 기능을 가지고 있다. 인맥서비스의 전반적인 구조나 인터페이스는 〈그림 1〉에서 나타나듯이 미니홈피의 변형된 형태로 볼 수 있다.

앞으로의 분석을 위해 이용되는 데이터는 IT인맥 서비스와 관련한 개별 회원 자료이다. 회원관련 자료에는 '개별 행동자료'와 '설문자료'가 포함된다. 회원 자료는 회원들의 성(sex)·직장 및 직무, 인터넷 경력과 접속환경 등에 대한 기본적인 배경 변수를 포함하고 있음은 물론, 각 회원 상호간의 인맥 형성의 기반이 되는 잠재인맥과 관심인맥의 현황 및 커뮤니케이션 정도와 상대방의 인맥카드에 방문한 횟수, 시점 등에 대한 모든 정보가 담겨져 있다. 회원들이 이 사이트에 접속하여 한번 클릭을 할 때마다 데이터로 형성되기 때문에 이러한 자료를 트래이스 데이터(Trace Data) 혹은 클릭 스트림 데이터(Click

Stream Data)라고도 부른다.

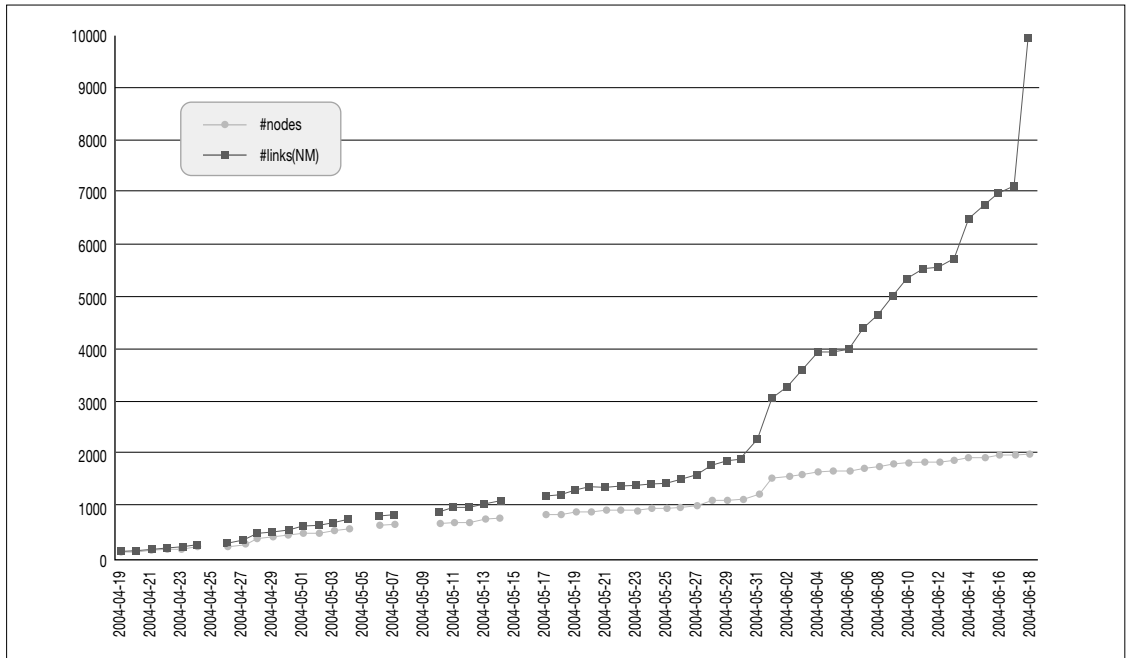
또한 온라인 상에서의 개인 행동 자료와 함께 포럼 구성원들의 개별적이고 주관적인 의식과 의도성을 파악해보기 위해 2004년 8월에 I사 홈페이지에서 설문조사를 실시하였다. 이 조사에는 총 586명이 참여하였으나 데이터 클리닝을 거쳐 분석에 이용할 수 있는 사례수는 572명의 응답이었다. 하지만 응답자들이 모두 인맥서비스를 이용하고 있는 것은 아니었기에 이 중에서 인맥서비스 이용자만을 추출한 결과 214명을 본 연구의 분석대상으로 하였다.

2. 인맥서비스에 나타난 관계의 주요 트렌드

본격적인 논의에 앞서 사이버 공간에서 이루어지는 관계의 양상과 실질적 과정을 살펴보기 위해 인맥서비스의 서비스 시작 시점부터 두 달간의 자료를 세부적으로 살펴보았다. 여기에는 행위자들의 가입을 통해 나타난 노드의 변화와 노드들 간의 관계 양상이 시계열적으로 나타나 있어 전반적인 관계 추이 및 네트워크의 구조를 파악하는데 도움을 준다.

〈그림 2〉는 서비스 개시 후 두 달 동안 이루어진 행위자의 증가와 행위자들 간 관계 형성의 추이, 즉 노드(node) 수와 링크(link) 수의 증가추세를 요약해 주고 있다. 처음에는 노드의 수가 더 많았으나, 불과 4일 만에 이 추세는 역전되어 링크의 수가 더 많아지는 것을 볼 수 있다. 즉 처음에는 소수의 적극적인 회원들에 의해 관심인맥으로 지정 당함으로써 데이터에 포함되는 사람들이 많았으나, 이들도 일단 이 공간에 들어오고 나면 적극적으로 다른 사람을 관심인맥으로 지정하고 관계를 맺어나간다는 것을 보여 준다. 이후 꾸준히 증가하다가 링크 수의 경우 5월 말을 기점으로 급속히 증가하는 양상을 나타낸다.

〈그림 3〉은 연결망 지름과 컴포넌트의 변화 추이를 나타내고 있다. 연결망 지름이란 연결망에서 가장 멀리 떨어져 있는 두 개의 노드들이 몇 단계를 거쳐서 연결되는지를 말하는 것이며, 컴포넌트는 몇 단계



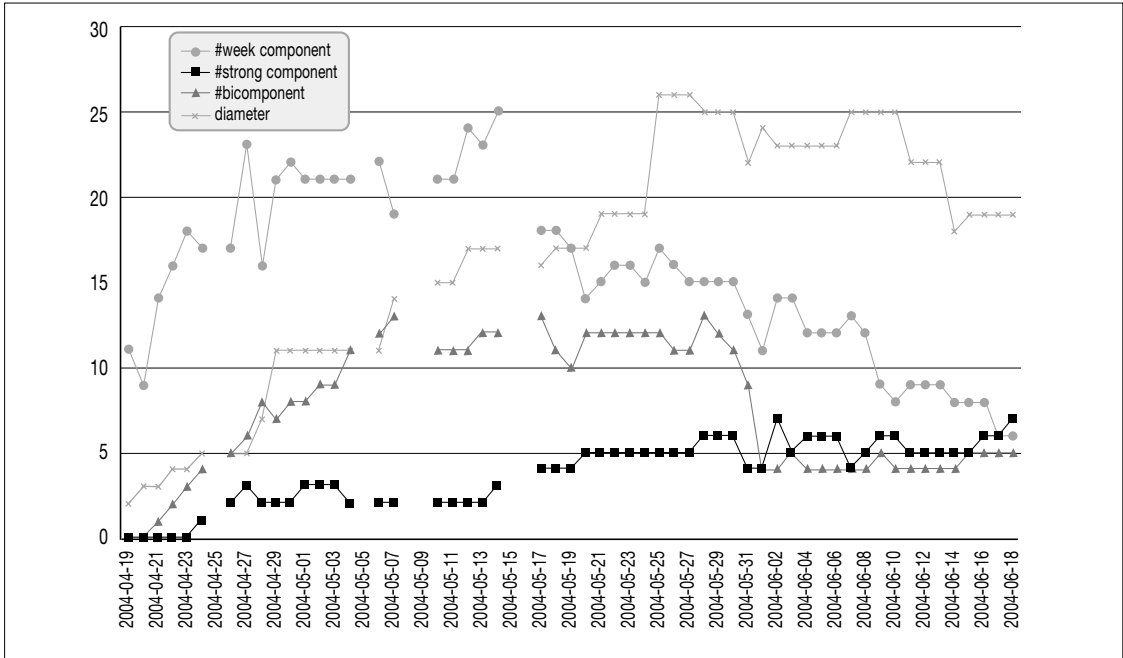
〈그림 2〉 노드와 링크수의 변화

를 거치든 상관없이 서로에게 연결될 수 있는 노드들로 이루어진 하위집단을 말한다. 이 중에서 강한 컴포넌트(strong component)란 두 개의 노드간에 양방향 모두 통로가 있어야 연결된다고 정의할 때의 컴포넌트이고, 약한 컴포넌트(weak component)란 어느 한 방향으로만 통로가 있어도 연결된다고 정의할 때의 컴포넌트를 말한다. 바이컴포넌트(bicomponent)란 컴포넌트에서 하나의 노드를 분리하기 위해서는 최소한 두 개의 링크를 제거해야 할 때의 컴포넌트를 말한다. 연결성이 좋고 성숙한 연결망은 대부분의 사람들이 어떻게든 서로 연결되기 때문에 컴포넌트의 수가 줄어드는 경향이 있다.

연결망 지름도 이와 비슷한 패턴으로 계속해서 늘어나다가 5월말부터 줄어들기 시작하여 최종일에는 19에서 안정화되고 있다. 즉 이 연결망에서 가장 멀리 떨어져 있는 사람들도 19단계를 거치면 연결된다는 뜻이다. 약한 컴포넌트와 바이컴포넌트의 수 및 연결망 지름이 같은 시기에 줄어들고, 이와 때를 같

이 하여 강한 컴포넌트의 수가 늘어나는 것은 5월말 이후 링크의 수가 급격하게 증가하는 〈그림 2〉의 결과와 관련되어 있다. 사람들 사이의 관계가 본격적으로 늘어나기 시작하면서 과거에는 연결되지 않던 사람들이 연결되기 때문에 약한 컴포넌트와 바이컴포넌트의 수가 줄어드는 것이고, 이렇게 연결성이 좋아지면서 과거에는 여러 단계 떨어져 있던 사람들 사이에 더 빨리 연결될 수 있는 새로운 통로가 생기기 때문에 연결망 지름이 줄어드는 것이며, 과거에는 일방적인 관계였던 것이 이제는 쌍방향적으로 되는 경우들이 늘어나고 있기 때문에 강한 컴포넌트의 수가 늘어나는 것이다.

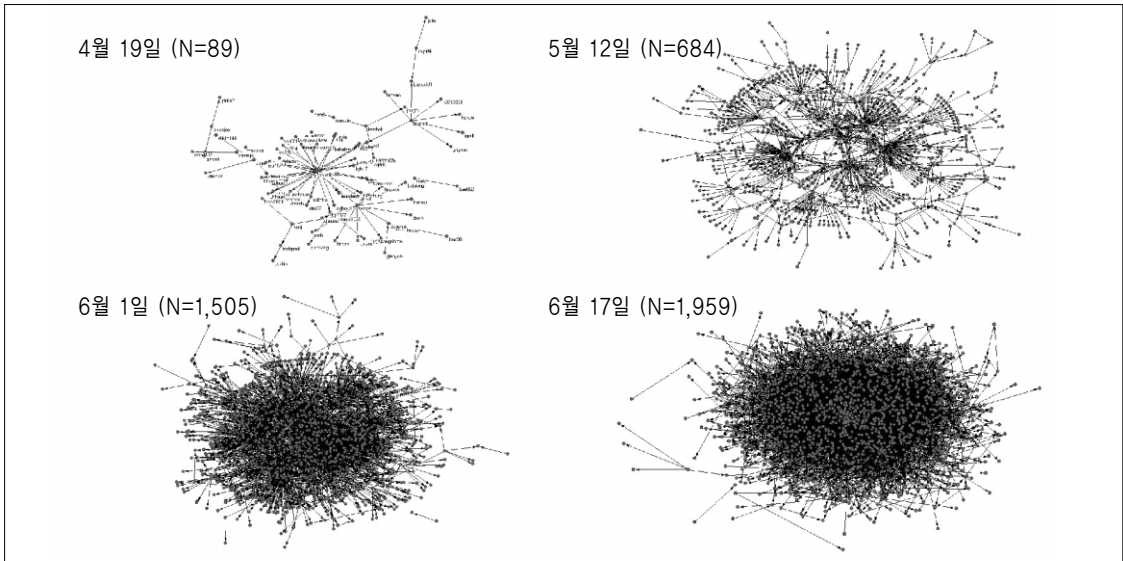
지금까지 살펴본 네트워크 특성의 시간적 추이 변화를 보면 4개 정도의 주요한 시점을 찾을 수 있고, 이에 따라 전체 60일간의 변화를 크게 3개의 시기로 구분하는 것이 가능하다. 첫 번째 시점은 인덱서비스가 시작된 4월 19일이고, 두 번째 시점은 컴포넌트와 연결망 지름의 변화가 나타나기 시작하는 5월 12일,



〈그림 3〉 연결망 지름(Diameter)과 컴포넌트(Component)의 변화

세 번째 시점은 링크의 수가 급격히 늘어나기 시작하면서 연결망에 큰 변화가 나타나기 시작하는 6월 1일, 네 번째 시점은 세 가지 종류의 컴포넌트가 거의

비슷하게 수렴하는 6월 17일이다. 〈그림 4〉는 이 네 개의 시점에서 각각 전체 네트워크가 어떻게 생겼는지를 보여주고 있다. 4월 19일에는 소수의 행위자



〈그림 4〉 4개 시점의 네트워크 변화

(n=89)들이 타인을 관심인맥으로 지정하고 있는 양상이 나타나고, 5월 12일에는 행위자의 수도 증가하고 관계의 양상도 매우 복잡해지고 있는데, 전체 구조 속에서 몇 개의 중심적인 허브가 존재하고 있는 것을 관찰할 수 있다. 6월 1일이 되면 시각적으로는 허브를 구분할 수 없을 정도로 복잡해졌지만 아직도 네트워크의 주변부에 빠져나와 있는 일부 행위자들이 눈에 띄고, 6월 17일에는 1,959명의 행위자 간에 거대한 네트워크 구조가 형성됨을 알 수 있다. 위에 제시한 그림들은 이제까지 추상적으로 막연하게 생각하고 있던 개인들 간의 관계 맺음 구조가 어떻게 형성, 성숙되어 가는지를 실제적으로 가시화하고 있다는 측면에서 그 의미가 크다고 본다.

3. 분석자료의 구성과 분석방법

(1) 분석시기의 선택

지금까지 살펴본 바와 같이 IT인맥서비스가 시작된 이후 60일 동안 많은 변화가 일어났으며, 60일 이후에도 변화는 또 일어날 것이다. 이러한 성격의 자료를 분석하는 데에는 시계열적인 분석과 횡단면 분석을 모두 생각할 수 있을 터이나, 이 자료의 경우 날짜에 따라 회원의 구성 자체가 계속해서 변하고 있기 때문에 같은 사람들에 대해 여러 시점에서 관찰하는 일반적인 패널 자료와는 차이가 있다. 따라서 이 논문에서는 한 시점을 택한 횡단면 분석을 시도하고자 한다. 그렇다면 어느 시점을 택할 것인가의 문제가 발생하는데, 시점 선택과 관련하여 우리는 다음의 조건을 충족시키고자 하였다. 하나는 신규 서비스인 IT인맥서비스가 그 자체의 역학을 가지기 시작한 이후가 좋겠다는 것이고, 다른 하나는 연결망 특성이 매일 같이 급변하기 보다는 어느 정도 안정성을 보이는 시기가 좋겠다는 것이다. 이러한 기준을 가지고 볼 때, 전체 60일 분량의 자료 중에서 6월 17일이 가장 적절하다고 판단하였다. 앞서 설명한 바와 같이 6월 17일에는 급격히 늘어난 링크들로 인해 전체적인

연결성이 높아지고 쌍방향 상호작용도 늘어난 상태이며, 5월말 이후 시작된 이러한 추세가 어느 정도 안정화 단계에 접어들었다고 보여지기 때문이다.

(2) 분석단위

앞에서도 설명하였지만, 이 논문의 주요 연구 질문은 사람들이 왜 같은 사람을 관심인맥으로 지정하는 것이다. 오프라인 세상에서 나타나는 ‘작은 세상’은 이미 잘 알려져 있고 그러한 현상은 온라인 상에서도 나타난다는 점이 밝혀져 있으나, 이러한 작은 세상을 가능하게 하는 원인은 결국 많은 사람들과 연결된 허브의 존재이다. 관계의 방향성을 고려하면 허브도 두 종류로 나눌 수 있는데, 자신이 다른 사람에게 관계를 요청하는(즉 타인을 관심인맥으로 지정하는) 종류의 허브가 있고, 많은 사람으로부터 관계를 요청받는(즉 타인으로부터 관심인맥으로 지정당하는) 종류의 허브가 있다. 이 중에서 주된 관심사는 관계를 요청받는 방향의 허브이다. 도구적인 인맥관리나 영업 등을 위해서 관계 요청을 남발하는 사람의 경우를 생각해 보면, 타인을 관심인맥으로 지정하는 허브는 본인의 노력만으로 쉽게 될 수 있지만 관심인맥으로 지정받는 허브는 본인이 의도적으로 통제할 수 없는 것이다. 지정받는 방향의 허브가 되기 위해서는 복수의 사람으로부터 동시에 관계를 요청받아야 하고, 이것은 집합적인 사회현상이기 때문에 “무엇이 사람들을 공통의 인물에 대한 집합적 관심으로 이끄는가”라는 고전적인 사회학적 질문과 맞닿아 있기도 하다.

이 질문에 답하기 위해서는 분석단위를 개인이 아니라 이자관계(dyad)로 바꾸어야 할 필요가 있다. A와 B라는 두 사람이 있다고 할 때, 우리의 관심은 A가 누구를 선택하였고 B가 누구를 선택하였는가라는 개별적 선택에 있는 것이 아니라 왜 A와 B가 공통의 인물인 C를 선택하였는가에 있기 때문이다. 6월 17일의 관심인맥 연결망에는 1,959명의 노드가 있기 때문에 이들의 이자관계 전체를 분석대상으로 한다

면 1,917,861개의 이자관계를 가지게 된다. 하지만 여러 가지 이론적이고 현실적인 이유들로 인해 우리는 분석대상을 대폭 축소해야만 했다. 첫째로, 우리는 이들의 선택에 영향을 미치는 여러 종류의 변수들을 포함하기를 원했는데, 그중에서도 특히 이들의 트레이스 자료와 2004년 8월에 이루어진 설문응답 자료를 포함하고자 하였다. 트레이스 자료는 두 종류가 있는데, 하나는 IT인택서비스의 모(母)사이트라고 할 수 있는 ITWAREHOUSE에서의 트레이스 자료이고, 다른 하나는 IT인택서비스 자체에서의 트레이스 자료이다. 이 자료들은 회원들이 두 개의 사이트를 이용하는 과정에서의 클릭 패턴이 공통인물에 대한 관심으로 유도하는지를 밝혀줄 것이다. 설문조사결과는 IT인택서비스에 대한 회원들의 주관적 평가가 어떤 영향을 주는지를 밝혀줄 것이다. 결과적으로 관심인택 연결망 자료와 두 사이트에서의 트레이스 자료, 그리고 설문응답자료를 모두 결합할 수 있는 회원은 231명으로 줄어들었다. 둘째로, 우리는 이들 중에서도 활동성이 활발한 사람들을 우선적으로 분석하고자 하였다. 그 이유는 대개의 인터넷 사이트는 가입만 해놓고 활동하지 않는 사람들이 워낙 많아서 이들이 분석결과를 흐려놓을 가능성이 높기 때문이다. 즉 인터넷 자료의 경우에는 무작위 표본추출보다는 활동성이 높은 사람들에 대한 전략적 표본추출(strategic sampling)이 더 적절하다고 판단한 것이다.¹⁾

이러한 판단에 따라 231명 중 외향중앙성(outdegree centrality)과 내향중앙성(indegree centrality)이 모두 상위 30% 이내에 드는 사람들을 추출하니 22명으로 나타났다. 이 22명 사이에 존재하는 231개의 이자관계가 이 논문의 최종적인 분석자료가 되었다. 여찌 보면 22명 사이의 관계란 분석대상으로 너무 적은 것이 아닌가라는 문제제기가 있을 수 있으나, 이 논문과 같이 분석단위를 이자관계로 전환하는 경

우에는 선행연구를 보면 불과 14개 산업 사이에 존재하는 91개의 이자관계를 분석대상으로 삼는 경우도 있으므로 22명 사이의 231개 이자관계는 그다지 적은 것이라고 할 수 없다(Mizruchi and Koenig 1986). 오히려 이슈는 자료의 크기보다 우리가 채택한 전략적 표본추출이 적절한 것인지의 여부라고 할 수 있으며, 인터넷 자료에서의 표본추출에 대한 논의는 앞으로도 계속되어야 할 것이다.

(3) 종속변수의 구성

일단 분석단위를 이자관계로 전환하고나면 두 사람 사이의 관심인택 지정의 유사성을 어떻게 측정할 것인가가 문제가 된다. 이자관계에서의 유사성을 연구할 때에 종종 사용되는 방법은 다음 식에 의해서 얻어진다(Mariolis 1975; Mizruchi 1989).

$$S_{ij} = n_{ij}/(n_i * n_j)1/2$$

S_{ij}: i와 j 사이의 유사성

n_{ij}: i와 j가 공통으로 선택한 사람의 수

n_i, n_j: i와 j가 선택한 사람의 수

이 방법은 그 나름의 유용성을 가지고 있으나, 공식에서 보듯이 i와 j가 공통으로 선택한 사람에 대한 정보는 유사성 지수의 계산에 포함되지만 둘 중 한 사람만이 선택한 사람이 누구인지는 전혀 반영되지 않는다. 분모에는 i와 j가 각각 몇 명을 선택하였는가에 대한 정보만이 있기 때문이다. 이러한 한계에도 불구하고 자료가 이원연결망(two-mode network)으로 구성되어 있을 때에는 다른 선택의 여지가 없다. 이 유사성 지수는 기업의 정치자금 기부 행위 등을 분석할 때에 많이 사용되었는데, 이 경우 자료의 행에는 기부를 하는 기업의 명단이, 그리고 열에는 기부를 받는 정치인의 명단이 들어가는 전형적 이원 연결망이 된다. 하지만 우리의 자료는 누가 누구를

1) 이와 관련한 보다 상세한 논의로는 장덕진(2005)의 각주 2 참조.

관심인맥으로 지정하였는가라는 일원연결망(one-mode network)이기 때문에 굳이 한계를 무릅쓰고 이 지수를 사용해야 할 이유는 없다.

따라서 우리는 이자관계의 유사성(혹은 상이성)을 보여주는 변수로서 외향관계(outgoing ties)에서 계산된 구조적 등위성 거리(structural equivalence distance)를 사용하기로 하였다. 일반적인 구조적 등위성 거리는 다음 공식에 의해 얻어진다.

$$d_{ij} = [\sum q(z_{iq} - z_{jq})^2 + \sum q(z_{qi} - z_{qj})^2]^{1/2}$$

d_{ij} : i와 j 사이의 구조적 등위성 거리

z_{ij} : i가 j에게 보내는 관계의 강도

공식에서 보듯이, 유사성 지수와 비교할 때 구조적 등위성 거리는 이자관계를 구성하는 두 사람에 의해 공통으로 지정된 사람이건 아니건 상관없이 연결망에 포함된 모든 제3자와의 관계를 하나하나 고려할 수 있다는 장점이 있다.²⁾ 즉 구조적 등위성 거리는 유사성 지수에 비해 더 구조적이고 더 입체적인 개념인 것이다. 이원연결망의 경우에는 선택의 여지가 없이 유사성 지수를 사용할 수밖에 없지만, 일원연결망에서는 구조적 등위성을 계산하는 것이 가능하므로 우리는 구조적 등위성 거리를 사용하기로 하였다. 유사성 지수는 두 사람의 선택이 비슷할수록 그 값이 커지지만, 등위성 거리는 두 사람의 선택이 비슷해짐에 따라 그 값이 작아지므로 엄밀하게 말하면 상이성 지수인 셈이다. 구조적 등위성 거리는 22명 사이의 관계에 한정해서 계산한 것이 아니라 1,959명 전체

에 대해 계산하여 그 중 22명 사이의 231개 이자관계에 대한 자료를 사용하였다. 따라서 비록 분석에 사용된 이자관계는 231개로 한정되지만, 적어도 이 변수에 있어서는 1,959명과의 관계가 모두 고려되었다고 볼 수 있다. 또한 두 사람이 다른 사람으로부터 관심인맥으로 지정받는 것이 아니라 두 사람이 공통의 인물을 관심인맥으로 지정하는 것이 우리의 관심이므로 일반적인 경우와는 달리 외향 방향만을 고려하여 구조적 등위성을 계산하였다. 분석에 사용된 등위성 변수와 유사성 지수 사이의 단순상관관계는 $-0.281(p=.00)$ 로 나타났는데, 이는 종속변수의 선택이 분석결과에 큰 영향을 미칠 수 있음을 보여준다. 각 주 2에 설명한 이유들로 인해 이 차이는 연결망의 크기가 커질수록 더욱 크게 나타날 것이며, 이러한 경우에는 유사성 지수보다 구조적 등위성 거리를 사용하는 것이 훨씬 더 정확한 결과를 가져올 것이라고 생각된다.

(4) 독립변수

분석에 사용된 독립변수는 크게 보아 4개의 그룹으로 나뉜다. 첫째는 인구학적 변수로서, 두 사람의 연령차이(로그값)를 사용하였다. 표집된 22명이 모두 남성이었기 때문에 성별 차이는 사용할 수 없었고, IT분야 종사자들의 인맥 사이트라는 특성상 대부분의 사람들이 비슷한 업무분야에 종사하고 지역적으로도 서울 거주자가 압도적으로 많아서 업무분야의 일치여부나 지역적인 차이도 사용할 수 없었다. 둘째는 모(母)사이트인 ITWAREHOUSE에서의 활

2) 유사성 지수와 구조적 등위성 거리의 차이점을 더 구체적으로 설명하면 다음과 같다. 크게 보아 차이는 두 가지이다. 첫째는 유사성 지수의 경우 이자관계를 구성하는 두 사람이 각각 맺고 있는 관계를 나타내는 두 개의 벡터의 방향만이 중요할 뿐, 길이는 반영되지 않는다는 것이다. 예를 들어 A와 B가 각각 {1, 1, 0, 2, 0, 0}과 {1, 1, 2, 3, 4, 5, 6}으로 나타낼 수 있는 이웃구조를 가지고 있다고 하자. 이때 만약 A의 벡터가 2배가 되어서 {2, 2, 0, 4, 0, 0}이 된다 하더라도 유사성 지수에는 변함없다. 하지만 구조적 등위성은 벡터의 길이가 달라지는 것에 따라 변화하게 된다. 가중치를 가지는 데이터의 경우 이것은 큰 차이를 낳게 되는데, 우리의 자료는 관심인맥으로 지정했느냐 아니냐를 보여주는 1, 0으로 구성된 자료이므로 이에 따른 차이는 없게 된다. 둘째는 본문에서 설명한 바와 같이 A와 B가 공유하는 이웃 이외의 다른 이웃들이 반영되느냐의 여부이다. 설명을 단순화하기 위해 A와 B가 공유하는 이웃이 하나도 없는 상황을 가정하자. 이 경우 A만의 이웃 혹은 B만의 이웃이 몇 명이건 누구이건 상관없이 A와 B 사이의 유사성 지수는 0이 된다. 예를 들어 A의 이웃은 {0, 0, 0, 0, 1, 1}이고 B의 이웃은 {1, 0, 1, 0, 0, 0}이라면 유사성 지수는 0이 되지만 구조적 등위성 거리는 $(4)^{1/2} = 2$ 가 된다. 이러한 차이로 인해 나타나게 되는 결과 중 하나는 유사성 지수의 행렬은 0이 굉장히 많은 밀도 낮은 행렬(sparse matrix)이 되지만, 구조적 등위성 거리는 0이 거의 없는 짙은 행렬(dense matrix)이 된다는 점이다. 따라서 이것을 이 논문에서와 같이 이자관계 단위의 분석에 사용하게 되면 많은 0과 소수의 양수(positive number)를 가지는 우편포(positive skew)를 보이게 되지만, 구조적 등위성 거리는 상대적으로 정규분포에 가까운 모습을 보이게 되는 것이 일반적이다. 그렇다면 우리는 변수 구성 원리의 차이에서나 분포상의 특징에서나 모두 유사성 지수를 사용하는 것보다는 구조적 등위성 거리를 사용하는 것이 더 낫다고 판단하는 것이다. 두 변수의 차이를 분명하게 하는데 도움을 준 (주)사이람의 김세권씨에게 감사의 뜻을 전한다.

동경력 변수들이다. IT인택서비스는 원래 IT종사자들의 커뮤니티 사이트인 ITWAREHOUSE로부터 분리되어 나온 서비스이므로 IT인택서비스의 회원들 중에는 모사이트로부터 건너온 사람들이 많다. 이들은 IT인택서비스를 이용하기 이전에 상당기간 동안 ITWAREHOUSE에서 다양한 활동을 해온 사람들이기 때문에 모사이트의 서비스인 IT포럼 등을 통해서 서로를 알고 있는 경우들도 있을 수 있다. 따라서 모사이트에서의 활동경력이 신규 서비스인 IT인택서비스에서의 인택 지정의 유사성에 어떤 영향을 미칠지는 주요한 관심사이다. 모사이트 활동경력 변수들로서는 모사이트 가입기간(월 단위)의 차이, 포럼활동 더미, 그리고 클릭패턴 거리를 사용하였다. 앞서 언급한 바와 같이 모사이트에서의 활동경력이 긴 사람들은 IT인택서비스에서도 많은 사람들을 이미 알고 있을 가능성이 있으므로 이차관계를 구성하는 두 사람 사이의 가입기간 차이를 사용하였고, 모사이트로부터의 자료를 사용한 기존 연구들에서 포럼활동은 중요한 변수로 나타났기 때문에(배 영 2003a; 2003b; 2005; 장덕진 2005) 양쪽 다 포럼활동을 하는 사람들로 구성된 이차관계와 그렇지 않은 이차관계를 구분할 수 있는 더미 변수를 포함하기로 하였다. 클릭패턴 거리는 회원×페이지로 구성된 행렬로부터 위계적 집락분석을 행할 때 얻어지는 상이성(dissimilarity) 거리이다. 위계적 집락분석을 통한 클릭패턴 도출의 유용성은 선행연구에서 이미 증명된 바 있으므로 여기서는 상세한 논의는 생략한다(장덕진 2005 참조).

셋째는 IT인택서비스의 경험변수로서, 이는 다시 객관적 경험과 주관적 평가로 나뉜다. 객관적 경험변수는 앞에서와 같은 방식으로 얻어진 인택 사이트의 클릭패턴 거리를 사용하였고, 주관적 평가변수는 설문조사로부터 얻어진 만족도 차이를 사용하였다. 넷째는 관심인택 지정으로 만들어진 연결망 변수들이다. 여기에는 두 사람간 내향중앙성의 차이, 외향중앙성의 차이, 두 사람간 직접적인 관계의 존재 여부

더미, 그리고 두 사람간 제약도(constraint)의 차이 등이 포함되는데, 뒤의 두 변수에 대해서는 약간의 설명이 필요하다.

이 논문에서 분석하는 IT인택서비스도 결국은 블로그나 미니홈피와 같은 일종의 개인미디어인데, 개인미디어에서 자신의 관심을 끄는 누군가를 발견하는 주요한 방식은 자신과 직접 관계를 맺고 있는 사람의 개인미디어를 방문했다가 그 사람의 인택으로부터 발견하는 방식이 된다. 선풍적인 인기를 끌고 있는 싸이월드를 예로 들면 일촌의 미니홈피를 방문했다가 일촌의 일촌 중에서 누군가를 발견하고 그 사람의 미니홈피를 방문했다가 관심이 생기면 일촌신청을 하는 방식이다. 실제로 IT인택서비스의 경우에도 관심인택의 인택카드를 방문하면 그 사람의 관심인택 리스트를 볼 수 있고, 리스트에 나온 사람의 이름을 클릭하면 다시 그 사람의 인택카드로 이동할 수 있기 때문에 “관심인택의 관심인택”을 자신의 관심인택으로 지정할 가능성은 높은 편이라고 할 수 있다. 이러한 메커니즘이 실제로 존재한다면 두 사람 사이의 직접적 관계 여부 더미는 관심인택 지정의 유사성을 예측하는데 있어서 유의미한 역할을 수행할 것이다.

다음으로 연결망에서의 제약도란 다음 공식에 의해 얻어진다(Burt 1992).

$$\text{constraint} = C = \sum_j c_{ij}$$

$$c_{ij} = (p_{ij} + [\sum_q p_{iq}p_{qj}]) / 2O_j, q \neq i, j$$

c_{ij} : j가 i에게 부여하는 제약의 정도

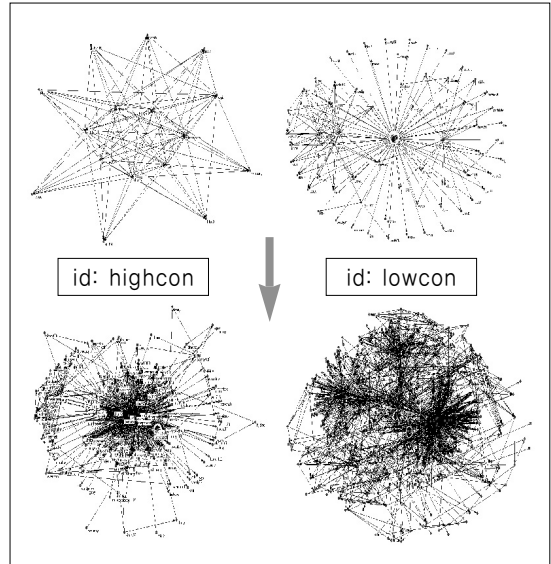
p_{ij} : i가 내보내는 전체 관계 중 j가 차지하는 비율

O_j : 해당 분야에서 j의 독과점 정도

개념적으로 설명하면, i의 개인 연결망에서 특정한 인물인 j가 차지하는 비중이 높을수록, j이외에 i가 관계를 맺고 있는 사람들이 또 다시 j에게로 모여들수록, 그리고 해당 분야에서 j가 차지하는 독과점의 정도가 높을수록 j로 인해 i가 받는 제약의 정도는 높

아지게 된다. 기업간 연결망 등에서 제약도를 계산할 때에는 흔히 j 의 시장내 매출액 비중 등이 독과점 정도로 사용되지만, 개인 연결망에서는 이와 같은 독과점 정도가 존재하지 않으므로 이 논문에서는 O_j 에 해당하는 항은 제외하고 계산하였다. i 는 자신의 연결망에 들어와 있는 모든 j 로부터 크고 작은 제약을 받게 되는데, 이것을 모든 j 에 대해 다 합친 것이 전체 연결망의 제약도이다. 이 분석에서는 전체 연결망 제약도의 차이를 사용하였다.

그렇다면 연결망 제약도의 차이가 관심인맥 지정의 유사성을 예측하는 데에 중요할 것이라고 예측하는 이론적인 이유는 무엇인가. 그것은 연결망 제약의 정도가 다른 사람들은 전체 연결망을 바라보는 시야가 다르기 때문이다. 보통 제약도가 높은 사람들은 규모가 작고 밀도가 높은 개인 연결망을 가지는 경향이 있는데, 이 경우 그의 시야에 들어오는 사람들은 수적으로도 몇 명 안 될뿐더러 그들 중 상당수는 자기들끼리 직접 관계를 맺고 있는 응집적 하위집단(cohesive subgroup)인 경우가 대부분이다. 응집적 하위집단에 속하게 되면 동질적인 사람들끼리 모이고, 한쪽으로 치우친 정보에 접할 가능성이 높으며, 편견을 증폭시키는 경향이 크다는 연구결과들이 존재한다(Burt 1992). 반면 제약도가 낮은 사람들은 연결망 전체를 조망하면서 밀도가 낮고 이질적인 사람들로 구성된 개인 연결망을 가지게 된다. 이들은 폭 넓고 균형 잡힌 정보를 접하기 때문에 똑같은 수의 관심인맥을 지정하더라도 누구를 자신의 관심인맥으로 지정하는 것이 가장 좋은지를 더 잘 알 수 있게 되고, 자신이 아니면 연결되지 않을 사람들을 연결해주는 브로커 역할을 하기 때문에 새로운 부가 가치 창출의 기회도 가장 먼저, 그리고 가장 많이 얻게 된다. 따라서 우리는 비슷한 정도의 제약도를 가진 사람들은 비슷한 사람들을 관심인맥으로 지정할 가능성이 높다고 보는 것이다.



〈그림 5〉 highcon과 lowcon의 단계별 개인 연결망 비교

연결망 제약도의 차이가 실제로 중요한 역할을 할 것이라는 우리의 이론적 예측을 뒷받침하기 위해 〈그림 5〉는 분석대상이 된 22명의 사람들 중에서 제약도가 가장 높은 사람과 제약도가 가장 낮은 사람을 예로 들어 이들의 1단계 개인 연결망과 2단계 개인 연결망을 비교해 보여주고 있다. 제약도가 .148로 가장 높은 highcon³⁾의 경우 1단계에서 16명하고만 연결되어 있고, 개인 연결망의 밀도는 .375이다. 반면 제약도가 가장 낮은 lowcon은 .038로 highcon의 4분의 1에 불과하며, 1단계에서 66명의 사람들과 연결되면서 밀도는 .037에 불과하다. 그림에서 시각적으로도 분명히 구분되듯이 highcon의 개인 연결망은 서로가 서로에게 연결되어 있는 응집적 하위집단인 반면, lowcon의 개인 연결망은 다양한 사람들과 집단들에게로 연결되는 부가 가치 창출형 연결망이다. 이러한 차이는 2단계로 가면 더욱 두드러진다. highcon의 2단계 개인 연결망은 157명과 연결되면서 .032의 밀도를 가지는데, 그림에서 보듯이 그중 대부분의 사람들이 가운데에 몰려서 서로가 서로에

3) 회원들의 프라이버시 보호를 위해 개인의 아이디는 변경하였음.

게 연결하고 있는 응집적 하위집단이다. 반면 lowcon의 2단계 개인 연결망은 2단계에서 무려 418명에게 연결되면서 밀도는 .009에 불과하다. 더욱 중요한 것은 그림에서도 뚜렷이 구분되듯이 lowcon은 2단계 개인 연결망을 통해 여러 명의 개인들뿐만 아니라 서로 분리되어 있는 몇 개의 주요한 집단들에게로 연결되고 있는 것이다. highcon 유형의 연결망을 가진 사람과 lowcon 유형의 연결망을 가진 사람은 상당히 다른 종류의 사람들을 관심인맥으로 지정할 것이라는 우리의 예측은 개인 연결망의 비교에서도 상당히 설득력 있는 것으로 나타난다.

(5) 분석방법

분석에 사용된 방법은 LSDV(Least Squares with Dummy Variables) 회귀분석이다. LSDV 회귀분석이란 22명 중 21명을 나타내는 더미변수를 포함한 선형회귀를 말하는데, 21개의 더미변수는 이차관계에서 각 변수가 나타내는 사람이 포함되었으면 1, 그렇지 않으면 0으로 코딩된다. 이차관계를 단위로 하는 회귀분석은 자료의 구성상 불가피하게 연결망 자기상관(network autocorrelation)을 포함하게 되

며, 따라서 각 관찰사례는 상호독립적이지 않다. LSDV 회귀분석은 이차관계를 구성하는 개인들에 대한 더미변수를 포함함으로써 오차항 속에 잠재적으로 존재하는 자기상관 요인이 모델 안으로 들어오게 되며, 오차항은 회귀분석의 가정을 충족시키게 된다(Hannan and Young 1977; Lincoln 1984).

IV. 분석결과

〈표 1〉은 분석에 사용된 변수들의 기술통계를 보여주고, 〈표 2〉와 〈표 3〉은 변수들간의 단순상관계수 및 부분상관계수(partial correlation coefficient)를 보여주고 있다. 기술통계에서 특기할만한 점은 별로 없으나, 자료구성에 사용된 22명 회원 중 일부는 IT인맥서비스에만 가입되어 있고 ITWAREHOUSE에는 가입되어 있지 않아서 이들에 대한 모사이트 클릭패턴 거리를 계산할 수 없기 때문에 이 변수의 관찰 사례는 166으로 줄어든다는 점을 알 수 있다. 나머지 변수들의 관찰수는 231이 아니라 225인데, 이는 전체 사례 중 회귀분석에서 극단치로 판명된 6개의 사례를 제외했기 때문이다.

〈표 1〉 분석에 사용된 변수들의 기술통계

	N	평균	표준편차	최소값	최대값
1. 연령차이(로그)	225	.66	.51	-1.00	1.36
2. 가입기간차이(로그)	225	2.43	.49	.70	3.24
3. 포럼 일치 더미	225	.24	.42	.00	1.00
4. 모사이트 클릭패턴 거리	166	2.78	.33	1.92	3.42
5. 인맥사이트 클릭패턴 거리	225	2.70	.33	1.80	3.64
6. 인맥사이트 만족도 차이 (절대값)	225	11.60	8.85	.00	40.00
7. 직접관계 더미 (로그)	225	-.94	.24	-1.00	.04
8. 내향중양성 차이 (절대값)	225	3.46	2.56	.00	10.00
9. 외향중양성 차이 (절대값)	225	31.32	23.77	.00	102.00
10. 제약도 차이 (절대값)	225	.03	.02	.00	.11

〈표 2〉 변수들 간의 단순상관계수

	1	2	3	4	5	6	7	8	9
2	.021								
3	-.181**	.080							
4	-.015	.030	.158*						
5	.088	-.034	-.136*	.427**					
6	.053	.282**	.058	.161*	.073				
7	-.066	.058	-.048	.158*	-.055	-.023			
8	.032	.110	.117	.016	.025	-.010	.060		
9	-.030	.042	.174**	.086	.020	-.017	.268**	.147*	
10	-.133*	-.086	-.053	-.026	.078	-.110	.125	-.090	.322*

* p<.05 ** p<.01

단순상관계수를 보면 일부 변수들 사이에서 상대적으로 높은 상관관계가 발견된다. 대표적으로 모사이트 클릭패턴 거리와 인맥사이트 클릭패턴 거리 사이의 단순상관계수가 .427(p<.01)이고, 외향중양성 차이와 제약도 차이의 단순상관계수가 .322(p<.05)이다. 하지만 이 논문에서 궁극적으로 사용할 분석은 각 회원을 나타내는 21개의 더미변수를 통제하는 LSDV 회귀분석이기 때문에 변수들 사이의 상관관

계 또한 단순상관관계가 아니라 더미변수들을 통제한 상태에서의 부분상관관계를 보아야 한다. 〈표 3〉에 보고된 부분상관계수를 보면 단순상관관계에서 존재하던 대부분의 관계는 사라지고 네 개의 상관관계만이 유의미하게 남는데, 이는 각각 인맥사이트 클릭패턴 거리와 제약도 차이, 내향중양성 차이와 외향중양성 차이, 모사이트 클릭패턴 거리와 인맥사이트 클릭패턴 거리, 그리고 외향중양성 차이와 제약도 차

〈표 3〉 변수들 간의 부분상관계수

	1	2	3	4	5	6	7	8	9
2	.060								
3	-.086	.072							
4	-.035	.041	.099						
5	.043	.108	.080	.322**					
6	.117	-.061	.002	.040	-.059				
7	-.056	.091	-.055	.123	-.089	.015			
8	.052	.126	.031	-.044	-.017	-.033	.070		
9	.022	.098	.058	.137	.125	-.070	.122	.195**	
10	-.009	-.010	.078	.029	.178*	-.074	-.071	.010	.411*

* p<.05 ** p<.01. 주) 더미변수를 통제한 부분상관계수임.

〈표 4〉 관심인맥 지정의 구조적 등위성 거리를 종속변수로 한 LSDV 회귀분석 결과

	모형 1	모형 2	모형 3	모형 4	모형 5
연령차이(로그)	.020 (.394)	.005 (.514)	-.006 (.505)	-.006 (.499)	-.002 (.830)
가입기간차이(로그)	.002 (.956)	.009 (.303)	.002 (.835)	.001 (.885)	-.003 (.717)
포럼활동 더미	.011 (.749)	-.013 (.278)	-.019 (.127)	-.020 (.120)	-.013 (.264)
모사이트 클릭패턴 거리	-.019 (.598)	-	-	-.003 (.839)	-.006 (.631)
인맥사이트 클릭패턴 거리	.076 (.036)	.009 (.439)	.016 (.197)	.017 (.194)	.019 (.118)
인맥사이트 만족도 차이 (절대값)	-.021 (.521)	.007 (.487)	.015 (.218)	.015 (.210)	.008 (.462)
직접관계 더미		-.032 (.000)	-.030 (.001)	-.030 (.001)	-.035 (.000)
내향중양성 차이(절대값)		.004 (.637)	.019 (.065)	.018 (.075)	.016 (.099)
외향중양성 차이 (절대값)		.297 (.000)	.304 (.000)	.303 (.000)	.257 (.000)
제약도 차이 (절대값)		.043 (.000)	.074 (.000)	.074 (.000)	.020 (.283)
외향중양성 차이* 제약도 차이					.088 (.000)
Total R2	.937	.991	.992	.992	.993
Residual R2	.090	.365	.369	.374	.374

주) 표준화 회귀계수이며, 괄호 속은 회귀계수의 p값임. 회원 개인을 나타내는 더미변수의 회귀계수는 생략되었음. Total R2는 더미변수를 모두 포함한 것이며, Residual R2는 더미의 효과를 제거하였을 때임.

이다. 이중 앞의 두 개는 상관계수가 그다지 크지 않으므로 심각한 문제는 아니겠으나, 뒤의 두 개는 부분상관계수가 .322(p<.01)와 .411(p<.05)이기 때문에 이로 인한 다공선성 문제는 회귀분석에서 유의해서 보아야 할 부분이다.

LSDV 회귀분석의 결과는 〈표 4〉에 제시되어 있다. 무작위로 추출된 표본이 아니기 때문에 통계적 유의성 검정은 별 의미가 없으나, 이런 경우에도 발견적인 도구(heuristic device)로서의 역할은 수행

할 수 있고, 이 논문의 표집에 사용된 것과 같이 활동성이 높은 사람들에 한정해서 일반화할 수 있는 가능성이 있기 때문에 p값을 함께 제시하였다. 〈표 4〉에는 모두 4개의 모형이 제시되어 있다. 모형 1은 연결망 변수들을 제외하고 인구학적 변수, 모사이트 활동변수, 그리고 인맥사이트 경험변수만을 포함한 것이다. 이 모형에서는 인맥사이트 클릭패턴 거리가 커질수록 관심인맥 지정의 유사성이 낮아지는 것으로 나타나는 것을 제외하면 유의미한 변수가 발견되지

않는다. 모형 2는 여기에 연결망 변수들을 더하면서 모사이트 클릭패턴 거리를 제외한 것이다. <표 1>에 나타났듯이 모사이트 클릭패턴 거리는 225개 이자관계 중 166개에 대해서만 알 수 있기 때문에 이 변수를 사용하면 59개의 사례수를 잃어버리게 된다. 그렇다고 해서 이 변수를 제외하면 원래의 커뮤니티 사이트에서의 클릭패턴이 신규 인맥서비스에서의 관심인맥 유사성에 어떤 영향을 미치는가라는 흥미로운 질문에 답할 기회를 잃어버리게 된다. 이 문제를 해결하기 위해서 우리는 이 변수를 활용하되 모형 2와 모형 3을 비교함으로써 59개의 사례를 잃어버리는 것의 효과를 동시에 평가해보고자 하는 것이다.

두 개의 모형은 동일한 모형으로서 모두 모사이트 클릭패턴 거리를 제외한 것이고, 차이는 분석에 포함된 사례수에 있다. 모형 2에서는 225개의 사례가 모두 사용되었고, 모형 3에서는 이 변수를 사용할 경우 잃어버리게 될 59개의 사례를 제외한 166개의 사례에만 한정하여 분석한 것이다. 연령 차이 변수의 계수 부호가 바뀌긴 하였으나 여차피 부분회귀도표(partial regression plot)는 거의 수평에 가까운 것으로 나타났기 때문에 별 의미가 없다. 이것을 제외하면 두 개의 모형은 거의 같은 결과를 보여주는 데, 유일한 차이는 내향중양성의 차이의 p값이 모형 3에서는 대폭 낮아져서 통계적으로 유의미한 차이를 나타낸다는 것이다. 이 변수는 모형 4와 5에서도 약간의 p값 증가가 있기는 하지만 대체로 유의미하게 유지된다. 이러한 차이는 어디서 오는 것일까. 모형 2에서 모형 3으로 가면서 제외된 59개의 사례가 무엇인지를 생각해볼 필요가 있다. 이자관계 단위의 자료를 구성하기 위해 사용된 22명의 회원들 중 3명은 모사이트인 ITWAREHOUSE의 회원이 아니면서 IT인맥서비스에만 가입한 사람들이다. 따라서 이들에 대해서는 모사이트 클릭패턴 거리를 계산할 수 없었던 것인데, 이들이 자료에서 제외되자 갑자기 내향중양성 차이가 관심인맥의 유사성에 유의미한 영향을 미치는 것으로 나타난 것이다. 한 가지 가능한 해

석은 모사이트와 인맥 사이트를 모두 이용하는 사람과 인맥 사이트만 이용하는 사람들은 관심인맥을 알게 되는 경로가 다를 수 있다고 보는 것이다. 선행 연구들에서 반복적으로 증명되었듯이, 모사이트인 ITWAREHOUSE는 인터넷 커뮤니티인 “포럼” 활동이 매우 활발하였고, 또한 매우 중요한 역할을 하고 있는 것으로 나타났다. 다대다(多對多) 커뮤니케이션인 포럼의 특성상 모사이트에서의 활동경력을 가지고 있는 사람들은 사이트 내의 유명 인사들을 알고 있을 가능성이 매우 높고, 서로간에도 아는 사이일 가능성도 높다. 이들은 인맥 사이트에 가입해서도 모사이트에서 원래 알던 사람들을 우선적으로 관심인맥으로 지정했을 가능성이 높다. 그렇다면 모사이트에서 비슷한 정도로 알려졌던 사람들은 인맥 사이트에서도 비슷한 정도의 내향중양성을 가지게 되었을 것이고, 이들은 또 다시 비슷한 사람들을 자신의 관심인맥으로 지정했을 수 있는 것이다. 이러한 해석이 옳은 것인지 확증할 수는 없으나, 모사이트의 회원이 아닌 사람들과 관련된 59개의 이자관계를 제거했을 때 내향중양성 차이가 갑자기 유의미하게 변하는 현상과 같은 자료를 사용한 선행 연구로부터의 결과들을 종합하면 상당히 개연성이 높은 해석이라고는 말할 수 있을 것이다.

모형 4는 모사이트에서의 클릭패턴 거리와 함께 연결망 변수들을 포함한 것이다. 먼저 p값을 기준으로 보면 4개의 연결망 변수들이 모두 유의미한 것으로 나타난다. 앞서 예상한 바와 같이 둘 사이에 직접적인 관계가 있는 이자관계에서는 관심인맥 지정의 유사성이 현저히 높아지고, 이자관계를 구성하는 사람들 사이에 내향중양성과 외향중양성의 차이가 커질수록 관심인맥 지정의 유사성은 낮아진다. 제약도 차이도 예상했던 결과를 보여주는데, 두 사람 사이의 연결망 제약도가 서로 다르면 다를수록 관심인맥 지정의 유사성이 낮아지고 있는 것이다. 우리의 예상대로 제약도가 다른 사람들은 연결망을 조망하는 시야의 폭 자체가 다를 가능성이 많다. 연결망 변수들은

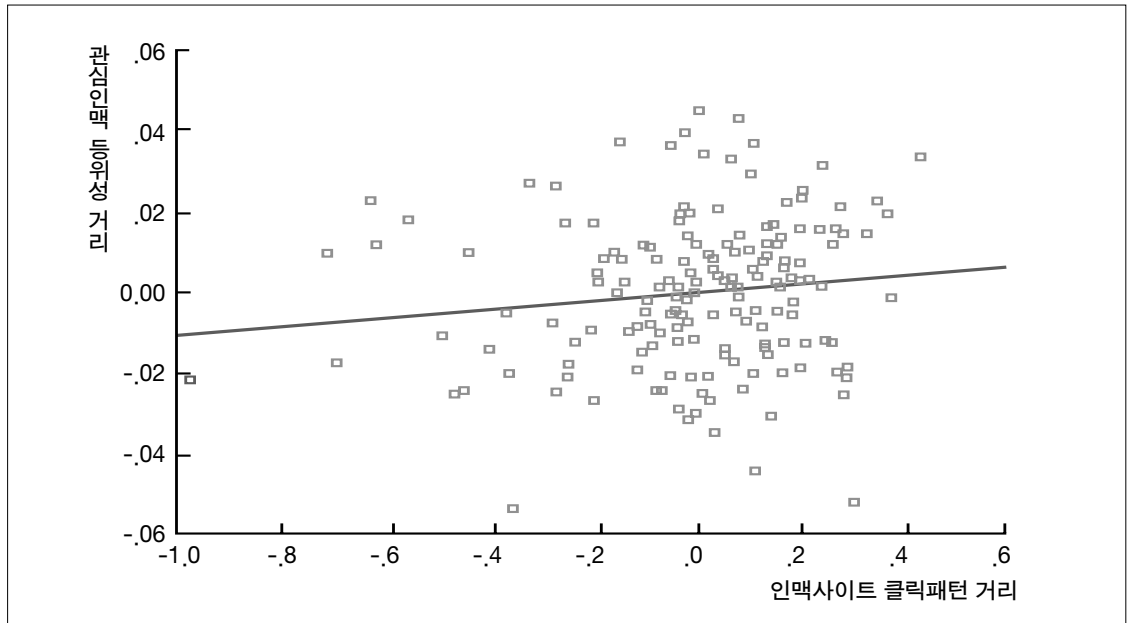
모두 유의미하게 나타났지만, 그중에서도 효과의 크기가 압도적인 것은 외향중양성의 차이이다. 이 변수가 이렇게 중요한 역할을 하는 이유를 다음과 같이 생각해볼 수 있다. 앞의 III절에서 논의한 60일간의 집중화 추이 변화 및 <표 1>의 기술통계에서도 나타나듯이, 이 연결망은 내향중양성에 비해 외향중양성이 매우 집중화 되어 있는 연결망이다. 즉 관심인맥 지정을 받는 사람들은 비교적 분산되어 있지만, 남들보다 훨씬 많은 사람들을 관심인맥으로 지정하는 소수의 사람들이 있다는 뜻이다. 이들은 모사이트에서의 활동경력 등을 통해 많은 사람들을 이미 알고 있는 사람일 수도 있고, 혹은 그야말로 인맥을 쌓기 위한 도구적인 이유로 상대방 회사의 규모나 직급 등 외형적인 기준이 충족되면 관심인맥 지정을 남발하는 사람들일 수도 있다. 둘 중의 어느 경우가 되었든 이들이 관심인맥으로 지정하는 사람들은 서로 겹치게 될 가능성이 많다. 앞의 경우라면 모사이트에서의 유명 인사들을 공유할 것이고, 뒤의 경우라면 대기업의 직급이 높은 사람들을 공유할 것이다. 반면 외향중양성이 낮은 사람들은 소수의 사람들만을 관심인맥으로 지정한다는 뜻인데, 이들은 인맥 서비스에 가입해 있는 사람들 중 자신의 친한 친구들만을 지정한다든가 하는 비교적 소극적인 사용자들이다. 따라서 이들이 관심인맥으로 지정하는 사람들은 모사이트에서의 유명 인사이거나 직급이 높은 사람들이기 보다는 자기 주변 사람들로 한정될 것이다. 그렇다면 외향중양성의 차이가 관심인맥 지정의 유사성에 큰 영향을 미치는 이유를 이해할 수 있을 것이다. 앞에서 외향중양성 차이와 제약도 차이, 그리고 모사이트 클릭패턴 거리와 인맥사이트 클릭패턴 거리 사이의 부분상관계수가 높게 나타났기 때문에 이로 인한 다공선성 문제가 발생할 수 있음을 지적한 바 있다. 하지만 회귀분석 결과 이 두 변수들의 분산중폭요인(VIF)

값은 순서대로 2.575, 3.336, 2.959, 2.986으로 심각한 문제는 아니라고 할 수 있다.⁴⁾

모형 5에서는 제약도 차이의 분산중폭요인이 가장 높았던 것을 감안하여 이 변수와 외향중양성 사이의 상호작용항을 추가하였다. 다른 결과들은 모두 동일하지만 흥미로운 것은 제약도 차이의 효과가 사라지고 상호작용항이 유의미한 것으로 나타났다는 점이다. 상호작용항의 회귀계수도 .088로 압도적 변수인 외향중양성 차이를 제외하고는 가장 높게 나타난다. 이 결과가 시사하는 바는 다음과 같다. 두 사람 사이의 관심인맥 지정의 유사성을 결정하는데 있어서 제약도 차이 그 자체가 중요한 것은 아닐 수 있다. 하지만 제약도의 차이가 외향중양성의 차이와 결합될 경우에는 급격히 관심인맥 지정의 유사성을 낮추어 놓을 수 있다는 뜻이다. 앞서 제시한 제약도 공식에서 나타나듯이, 연결망 제약도는 자기 연결망 크기에 상관없이 계산되고 0에서 1 사이로 변하는 값이다. 따라서 똑같은 정도의 제약도를 가지더라도 어떤 사람은 100명의 사람과 연결되어 있을 수 있고, 어떤 사람은 5명의 사람과 연결되어 있을 수도 있다. 5명과 연결된 사람이 매우 낮은 제약도를 가진다 하더라도 그 효과는 그다지 크지 않을 것이다. 하지만 100명과 연결된 사람이 낮은 제약도를 가진다면 그것의 파급효과는 엄청날 것이다. 모형 5는 우리가 예측한 제약도 차이의 효과가 정확히는 관심인맥 지정의 수와 결합되었을 때 본격적으로 나타난다는 것을 보여 준다.

지금까지는 통계적으로 유의미하게 나타난 연결망 변수들의 효과에 대해서만 논의하였지만, 다른 변수들의 효과에 대해서도 관심을 돌려볼 필요가 있다. p값만을 기준으로 보면 연결망 변수들만이 유의미하게 나타나지만, 앞서도 언급한 바와 같이 무작위 표본이 아닌 이 자료에서 p값은 그다지 중요한 것은 아

4) 회귀분석에서의 다공선성 판단기준으로 흔히 사용되는 VIF의 값이 얼마 이하여야 만족할만한 한 수준이라고 할 수 있는지에 대해서는 절대적인 기준은 없다. 그러나 일반적으로 사용되는 기준들을 보면 가장 관대하게 적용할 경우 10 이하, 가장 엄격하게 적용할 경우 2 이하의 VIF 값을 요구한다. 3.336은 가장 엄격한 기준보다는 높지만 가장 관대한 기준에 비하면 훨씬 낮기 때문에 다공선성 문제가 심각하지는 않다고 보는 것이다.



〈그림 6〉 인맥사이트 클릭패턴 거리의 부분회귀도표

니다. 또한 무작위 표본이라 하더라도 “영가설이 참인데도 불구하고 표본에서 특정한 통계치가 얻어질 확률”이라고 하는 p값의 원래 의미를 생각해보면 .05나 .1과 같은 기준은 하나의 관행일 뿐, 절대적인 의미는 없는 것이기도 하다. 그렇다면 우리는 모형 4와 모형 5에 나타난 몇몇 다른 변수들에도 주목할 필요가 있다고 생각한다. 예를 들면 모형 4에서 포럼활동 더미와 인맥 사이트 클릭패턴 거리, 인맥 사이트 만족도 차이 등을 보면, 이 세 변수의 p값은 각각 .120, .194, .210으로 관행적인 p값의 기준에는 미치지 못하지만, 그럼에도 불구하고 관행적 기준에 상당히 접근하고 있는 것도 사실이다. 포럼활동 더미의 경우를 예로 들면 정확한 해석은 “다른 변수들의 영향력을 통제했을 때, 이자관계를 구성하는 두 사람이 모두 포럼활동을 하고 있을 경우 이들의 관심인맥 지점의 유사성이 실제로는 없음에도 불구하고 모형에서와 같이 나타날 확률은 .120에 불과하다”는 것이다. 그렇다면 이것은 관행적 기준에서 벗어난다고 해서 쉽게 무시해버릴 수 있는 것은 아니라고 할 수 있

다. 포럼활동 더미와 인맥사이트 만족도 차이는 모형 3과 모형 4에서는 비슷하게 유지되지만, 상호작용항을 포함한 모형 5에서는 이 두 변수의 p값이 .264와 .462로 다시 높아지고 있기 때문에 확실하게 말하기는 어렵지만 그럼에도 불구하고 표집의 차이 등에 의해 유의미하게 나타날 가능성은 상당히 있다고 하겠다. 인맥사이트 클릭패턴 거리는 모형 3에서 모형 4, 모형 5로 감에 따라 미미하게나마 회귀계수는 높아지고 p값은 낮아지고 있어서 가장 견실한 모습을 보여준다. 최종모델에서의 p값은 .118로, 관행적 기준인 .1에 비해 1.8% 밖에는 차이가 나지 않는다. 〈그림 6〉은 최종 모델에서 이 변수의 부분회귀도표이다. 그림에서 나타나듯이 인맥 사이트 내에서의 클릭패턴이 달라짐에 따라 관심인맥 지점의 상이성도 높아지는 비교적 깨끗한 관계를 보여주고 있다.

끝으로, 우리의 기대와는 달리 모사이트에서의 클릭패턴 거리가 아무런 영향을 미치지 못하고 있음을 알 수 있다. 이 변수가 포함된 모형 1, 모형 4, 모형 5에서 모두 아무 효과가 없는 것으로 나타나고 있다.

이것은 IT인맥서비스의 회원 중 상당수가 ITWAREHOUSE에서 건너온 사람들이기 때문에 모사이트에서의 클릭패턴 거리가 관심인맥 지정의 유사성에 영향을 줄 것이라는 우리의 예상과는 정면으로 배치되는 결과여서 주목을 끈다. 같은 사이트에서의 클릭패턴을 사용한 선행연구(장덕진 2005)와 비교해보면 이러한 결과가 나타난 원인을 추론해볼 수 있는데, 선행연구에서는 모사이트가 가지고 있는 총 151개의 페이지에 대한 클릭 자료를 집락분석하여 5개의 집락을 얻은 후 이것을 다항로지스틱 회귀 분석의 종속변수로 활용하였다. 즉 클릭패턴을 범주형 변수로 취급한 것이다. 이때 얻어진 5개 집락의 내용을 보면, 각각 ① 포럼에서 주로 활동하는 ‘적극적/정서적 공동체’, ② 자료실을 주로 이용하는 ‘특화목적 집단’, ③ 기업경영 전반에 대한 정보와 메모 주고받기를 주로 활용하는 ‘개인 네트워크를 통한 일반적 정보교류 집단’, ④ 뉴스 일반과 상품소개 등에 치중하면서 사이트 내에서 자신의 클릭패턴을 아직 안정화시키지 못한 ‘익명적/일반적 범주’, 그리고

⑤ ‘신입회원 범주’로 나뉘었다. 이러한 집락구분은 모사이트에서의 클릭 패턴을 연속변수로 사용하는 것이 적절치 못할 수 있음을 시사한다. 클릭패턴 거리의 차이는 똑같다 하더라도 ①과 ③ 사이의 거리는 ①과 ④ 사이의 거리와는 의미가 다를 수 있는 것이다. 이것은 단순히 ‘서로 다르다’는 의미에서 집락을 구분하는 데에는 유용할 수 있지만, 같은 거리는 같은 의미로 취급하는 연속변수에서는 질적인 차이를 흐려놓을 가능성이 높은 것이다.

그렇다면 똑같은 클릭패턴 거리인데 왜 인맥서비스에서의 클릭패턴 거리는 나름의 유의미한 결과를 보여주는지에 대한 해명도 필요하다. 이것은 모사이트와 인맥 사이트에서의 페이지 구성의 내용이 큰 차이를 보인다는 점에서 그 이유를 찾을 수 있을 것 같다. 선행연구에서 나타나듯 모사이트는 비교적 이질적인 151개 페이지를 가지고 있는데 반해 비교적 단순한 구성을 가진 인맥 사이트는 모두 24개의 페이지로 구성되어 있는데, 이것을 분류해보면 <표 5>와 같이 정리된다.

<표 5> 인맥사이트 페이지 분류

대분류	중분류	세분류	대분류	중분류	세분류
회원등록 단계		이용안내	내 인맥카드 작성 단계 (계속)	파일함 관련	파일함 등록
		수정			파일함 상세
		인맥카드			파일함 리스트
		잠재인맥리스트	내 인맥카드 보기 단계	방문자 전체리스트	
내 인맥카드 작성 단계	Career 관련	My Career 등록	인맥 지정 단계		등록
		My Career 상세			등록인맥 전체보기
		My Career 수정			방명록 보기
	Photo 관련	My Photo 등록	방명록 상세보기		
		My Photo 상세	방명록 쓰기	방명록 등록하기	
		My Photo 수정		방명록 수정하기	
		My Photo 리스트	비동의자 파악	비동의자 인맥카드	
			검색		검색
					검색결과

이질적인 페이지들로 구성된 모사이트와는 달리, 인맥사이트는 회원등록 단계로부터 적극적인 회원으로 변신하기까지의 점진적인 과정이 그대로 반영된 페이지 구성을 보여준다. <표 5>에 제시된 ‘세분류’가 원래의 페이지 분류이고, 이것을 단계별로 묶어서 필요에 따라 ‘중분류’와 ‘대분류’로 제시하였다. 먼저 대분류 ‘회원등록 단계’에 해당하는 페이지들은 신입회원이면 누구나 처음에 클릭하게 되는 페이지들이다. 회원등록과 동시에 자신의 인맥카드가 저절로 생성되고, 기존 회원 중 학연이나 업무분야 등에 따라 자신과 관련이 있는 잠재인맥의 리스트를 제공받으며, 필요하면 이용안내를 참고하고, 인맥카드를 수정한다. 이 단계가 지나면 스스로를 소개하기 위해 자신의 인맥카드를 꾸며나가는 단계이다. 남들이 자신을 알 수 있도록 스스로의 경력을 등록하고(Career 관련), 더 적극적인 사람이라면 스스로를 소개할 사진도 올려놓으며(My Photo 관련), 더 나아가 남들과 공유하고 싶은 이야기나 업무자료 등도 올려놓는다(파일함 관련). ‘내 인맥카드 작성단계’를 충실히 거친 회원이라면 ‘회원등록 단계’에 머무른 회원보다 더 적극적이라고 할 수 있을 것이다. 이 단계가 끝날 즈음이면 다른 사람들이 내 인맥카드를 찾아오기 시작하는데, 어떤 사람들이 나를 찾았는지에 관심을 가지기 시작하는 것이 세 번째 단계인 ‘내 인맥카드 보기 단계’이다. 여기서 더 적극적이 되면 나도 다른 사람을 관심인맥으로 등록하고 내가 등록한 사람들을 관리하기 시작하는데, 이것이 네 번째의 ‘인맥 지정 단계’이다. 더 나아가면 내가 지정한 사람들의 인맥카드에 찾아가서 처음에는 구경만 할 것이고(방명록 보기), 더 적극적이 되면 스스로도 글을 남길 것이며(방명록 쓰기), 이보다 더 적극적이 되면 내가 인맥으로 지정했는데 동의하지 않은 사람들은 어떤 사람들인지까지도 찾아보게 될 것이다(비동의자 인맥카드). 인맥사이트 활용의 적극성에 있어서 마지막

단계는 ‘검색’이다. 검색을 사용하지 않을 경우에는 자동으로 제공되는 관심인맥리스트에서 선택하거나 혹은 나를 지정한 사람에 대해 나도 그를 지정하는 방식으로 인맥이 만들어진다. 하지만 여기에 만족하지 못하는 사람들은 회원 전체에 대해 특정 조건을 충족시키는 사람을 검색하여 찾아나설 것이기 때문이다. 따라서 검색을 사용하는 단계가 마지막 단계라고 할 수 있다. 이러한 페이지 구성에서 나타나듯이, 모사이트와는 달리 인맥사이트에서의 클릭패턴 거리는 집락구분을 위해서 유용할 뿐만 아니라 연속변수로서의 특성도 비교적 잘 갖추고 있다고 할 수 있다. 즉 <표 5>에서 구분한 여러 단계들 중 1단계와 3단계보다 1단계와 4단계가 실제로 더 멀다고 해석해도 무리가 없는 것이다. LSDV 회귀분석에서 인맥사이트 클릭패턴 거리가 나름의 유의미성을 보여준 것이나 선행연구에서 클릭패턴 거리에 근거한 집락구분이 유용하게 나타난 것, 그리고 한국의 대표적인 인터넷 사이트들 중 상당수가 일부 회원에 대해 트레이스 데이터를 구축하기 시작한 사실⁵⁾ 등은 향후 인터넷 공간에서 일어나는 일들을 이해하기 위해 이 변수가 중요하게 사용될 가능성이 높다는 점을 시사한다. 비록 기대와는 달리 모사이트 클릭패턴 거리가 관심인맥 유사성을 예측하는데 도움이 되지 않는다는 발견은 실망스럽지만, 이러한 결과는 페이지 구성의 내용에 따라 트레이스 자료의 측정을 달리해야 한다는 것을 알려준다는 점에서 앞으로의 연구를 위해 유용하게 쓰일 수 있을 것이다.

V. 결론

전국적인 “싸이 열풍”에서 보듯이, 사람들은 인터넷 상에서 개인 미디어를 통한 친구 맺기에 빠져들고 있다. 그리고 이렇게 맺어진 친구관계들은 개인 수준에서는 온라인 친구일 뿐이지만 거시적인 시각에서

5) 논문에서 기업명을 밝힐 수는 없지만, 이미 상당수의 대형 포털 등은 회원 중 일부에 대해 장기간에 걸쳐 트레이스를 추출하는 패널 자료를 구성하여 운용하고 있다.

보면 인터넷 전체의 구조를 바꾸고 오프라인의 현실에까지 커다란 영향을 미치고 있다. 그렇다면 그들은 왜 특정한 사람과 친구를 맺기를 원하는 것일까. 여기서 질문의 핵심은 A가 C와 친구관계를 맺었고 B가 C와 친구관계를 맺었다는 개인적 사실에 있지 않다. 중요한 것은 왜 A와 B가 공통적으로 C와 친구관계를 맺었느냐는 질문이다. 왜냐하면 이러한 공통의 친구관계를 통해 온라인 공간의 허브가 탄생하고, 허브의 존재는 끝없이 넓어 보이는 온라인 공간을 '좁은 세상'으로 바꾸어 놓으며, 좁아진 세상에서는 정보의 유통 속도가 빠를 뿐만 아니라 특정한 방향으로 편중되고, 이렇게 빠르면서도 편중된 거대한 분량의 정보는 사회 전체에 영향을 미치기 때문이다. 한국 사회가 이미 여러 차례 경험하였듯이, 눈사태와도 같이 견잡을 수 없게 번지는 온라인의 영향력을 가져오는 한 가지 미시적 원천은 공통의 친구관계이다.

이러한 문제의식 하에 이 논문에서 IT인맥서비스의 첫날부터 60일간의 자료를 분석하여 위의 질문에 답하고자 시도하였다. 신규 서비스의 등장 이후 다양한 네트워크 특성 지표들의 변화를 시계열적으로 추적한 결과 한달 반 정도의 시간이 지난 이후 안정기에 접어들기 시작하였다고 판단할 수 있었고, 안정기에 접어든 다음인 6월 17일 자료를 택하여 활동성이 높은 22명 회원들 사이의 231개 이차관계에 대한 LSDV 회귀분석을 통해 공통의 친구 맺기에 영향을 미치는 요인들이 무엇인지를 파악하고자 하였다. 분석결과는 다음과 같다. 비슷한 정도의 외향중양성이거나 내향중양성을 보이는 사람들은 비슷한 상대를 관심인맥으로 지정하는 경우가 많았고, 비슷한 정도의 제약도를 보이는 사람들도 비슷한 상대를 지정하는 경우가 많았다. 연결망 관련 변수들이 일제히 유의미한 결과를 보이는 데에는 개인 미디어간의 연결망에서 비슷한 위치에 있는 사람들이 전체 연결망을 조망하는 비슷한 시야를 가지게 되는 것이 주요하게 작용한다고 보여진다. 특히 외향중양성 차이와 제약도 차이의 상호작용항이 매우 중요하게 나타나는 것은 개

인 연결망의 구조와 크기가 결합될 때 이러한 시야의 차이가 더욱 중요해짐을 보여준다.

관행적으로 사용되는 통계적 유의미성의 기준에 미치지지는 못하였으나, 모사이트에서의 포럼활동 더미나 인맥사이트에서의 클릭패턴 거리도 무시할 수 없는 요인인 것으로 드러났다. 하지만 기대와는 달리 모사이트에서의 클릭패턴 거리는 아무런 영향을 미치지 못한다는 사실도 알 수 있었다. 분석의 대상이 된 IT인맥서비스가 사이버 커뮤니티 위주였던 모사이트로부터 분리되어 나온 것이기 때문에 다대다(多對多) 커뮤니케이션의 장(場)인 포럼에서의 인간관계는 일대일 커뮤니케이션인 친구관계에서도 유효하게 이어지고 있으며, 인터넷 상에서 사람들이 취하는 행동 중 가장 미시적 행동인 "클릭"의 패턴도 그 사람의 사회적 관계를 특정 방향으로 유도하는 경향이 있다고 보여진다. 모사이트에서의 클릭패턴 거리가 유의미하게 나타나지 않은 것은 모사이트와 인맥사이트의 페이지 구성에 큰 차이가 있기 때문이라고 생각된다. 이 결과는 비록 우리의 기대에 어긋난 것이기는 했으나, 향후 중요하게 사용될 수 있는 클릭패턴 분석에서 페이지 구성에 따라 변수 구성을 어떤 방식으로 하는 것이 적절할 지에 대한 단초를 제공해주었다는 점에서 그 의의를 찾고자 한다.

연구결과는 학문적인 함의와 더불어 기업경영적인 측면에서의 함의도 함께 가진다고 생각된다. 예를 들어 다음 카페나 싸이월드 미니홈피 같은 서비스들은 한국이 만들어낸 세계적인 히트상품이라고 할 수 있을 것이다. 하지만 그들의 성공에 가려서 잘 드러나지 않을 뿐, 이와 유사한 서비스들이 수없이 많이 나타났다가 성공하지 못하였다. 성공한 서비스와 그렇지 못한 서비스의 차이는 무엇인가? 블로그 방문자에게 그 블로그 주인의 일촌을 볼 수 있게 하는 것과 볼 수 없게 하는 것은 서비스의 성패에 어떤 영향을 미칠까? 일촌만 보여주는 것과 이촌 이상까지 모두 볼 수 있게 하는 것은 또 어떤 차이를 낳을 것인가? 같은 사람에게 일촌신청을 한 사람들끼리 서로 연결

시켜 주는 것이 도움이 될까, 아니면 서로 모르게 하는 것이 도움이 될까? 이러한 질문들은 사소한 것 같지만 사실은 서비스 전체의 성패를 결정지을 수도 있는 중요한 선택이 될 수 있다. 이 논문의 분석결과는 지금까지 답하지 못했던 이러한 질문들에 대한 일부분의 해답을 제시한다는 점에서 기업경영적인 측면의 함의를 가진다고 할 수 있다. 학문적인 함의이자 동시에 이 논문의 가장 큰 한계라고 생각되는 것은 표본추출의 문제이다. 본문에서도 상세히 기술하였으나, 실질적으로 활동하지 않는 아이디가 훨씬 더 많은 인터넷 자료의 경우 무작위 표본추출은 적절하지 않다는 것이 우리의 판단이었고, 따라서 우리는 활동성이 높은 소수의 회원들만을 대상으로 하여 분석하였다. 하지만 이러한 판단이 옳은 것인지에 대한 합의된 정답은 아직까지 없는 것 같다. 현재로서는 이 문제에 대한 학문적 논의의 활성화를 제안하는 것으로 우리의 대답을 대신하고자 한다.

■ 참고문헌

- 배 영 (2003a). "사이버 커뮤니티의 내부동학과 관계구조: 28개 전문인 포럼의 비교분석." 『한국사회학』, 37(3): 109-134.
- 배 영 (2003b). "사이버 커뮤니티의 경험적 분석을 이용한 사회자본의 유형화." 『한국사회학』, 37(5) : 161-186.
- 배 영 (2005). "사이버 공간의 사회적 관계: 개인미디어를 이용한 관계의 형성과 유지를 중심으로." 『한국사회학』, 39(5) : 55-92.
- 장덕진 (2005). "웹사이트 이용자의 클릭패턴 결정요인에 대한 연구: I사 클릭스트림 데이터를 중심으로." 『한국사회학』, 39(4): 101-130.
- 중앙일보(2005). "탐사기획: 온라인사회해부-인터넷 친구맺기." 5월 9일.
- Barnes, S. (2001). *Online Connection*. New York: Hampton Press.
- Burt, Ronald S. (1992). *Structural Holes: The Social Structure of Competition*. Cambridge: Harvard University Press, edited by David R. Heise. San Francisco: Jossey-Bass.
- Hannan, Michael T. and Alice A. Young (1977). "Estimation in Panel Models: Results on Pooling Cross-Sections and Time Series." *Sociological Methodology* 1977 :52-83.
- Howard, P., Raine, L., and Jones, S. (2001). "Days and nights on the Internet: The impact of a diffusing technology." *The American Behavioral Scientist*, 45(3): 383-404.
- Jones, S. (2001). *Cyber Society 2.0*. Thousand Oaks: Sage Publications.
- Katz, J. E., Rice, R. E., & Aspden, P. (2001). "The Internet, 1995-2000: Access, civic involvement, and social interaction." *The American Behavioral Scientist*, 45(3): 405-419.
- Lincoln, James R. (1984). "Analyzing Relations in Dyads." *Sociological Methods and Research*, 13: 45-76.
- Mariolis, Peter (1975). "Interlocking Directorates and Control of Corporations." *Social Science Quarterly*, 56: 425-439.
- Mizruchi, Mark S. and Thomas Koenig (1986). "Economic Sources of Corporate Political Consensus: An Examination of Interindustry Relations." *American Sociological Review*, 51: 482-491.
- Mizruchi, Mark. S. (1989). "Similarity of Political Behavior Among Large American Corporations." *American Journal of Sociology*, 95(2): 401-424.
- Nie, N. H. (2001). "Sociability, interpersonal relations, and the Internet: Reconciling conflicting findings." *The American Behavioral Scientist*, 45(3): 420-435.
- Nie, N. H., & Erbring, L. (2000). "Internet and society: A preliminary report." online available: www.stanford.edu/group/siqss. (검색일: 2005. 2. 5).
- Orleans, M., Laney, M. C. (2000). "Children's computer use in the home: Isolation or sociation?" *Social Science Computer Review*, 18(1): 56-72.

Pruijt, H. (2002). Social capital and the equalizing potential of the internet. *Social Science Computer Review*, 20(2): 109-115.

Shenk, D. (1997). *Data smog: Surviving the information glut*. New York: Harper Collins Books.

Wallace, P. (1999). *The Psychology of the Internet*. Cambridge: Cambridge Univ. Press.

Wellman B. (2001a). Computer Networks As Social Networks, *Science*, 293: 2031-2034.

Wellman B. (2001b). "Does the internet increase, decrease, or supplement social capital? Social networks, participation, and community commitment." *American Behavioral Scientist*, 45(3): 436-455.

Wellman, B. & Gulia, M. (1999). "Net-Surfers Don't Ride Alone: Virtual Communities as Communities." Barry Wellman (ed.), *Networks in the Global Village: Life in Contemporary Communities*, 331-366. Boulder: West View Press.

■ 필자소개

- 장덕진(Duk-jin Chang)**
 1999, PhD in Sociology, The University of Chicago
 현재, 서울대학교 사회학과 조교수
 관심분야 : 네트워크 분석, 경제사회학, 정보사회학
- 천명규(Myoung-kyu Chun)**
 1985, 사회학사, 성균관대학교
 현재, (주)잇이즈컴 대표이사
 관심분야 : 정보사회학
- 김기훈(Gi-hoon Ghim)**
 1987, 사회학 석사, 서울대학교
 현재, (주)사이람 대표이사 및 송실대 정보사회학과 겸임교수
 관심분야 : 정보사회학
- 배영(Young Bae)**
 2003, 사회학 박사, 연세대학교
 현재, 송실대학교 정보사회학과 조교수
 관심분야 : 정보사회학, 사회자본론