

Networked Community: A connected Society

Soungwoong Yoon *, Sang-Hoon Lee**

Abstract

We are living in networks which are regarded as a society. However, it is difficult to designate a specific position or the impact over sociological relationships and virtual links. In this paper, we conceptualize two themes of the network as Physical Network and Virtual Network, and observe their cross-network effects. New concept called Networked Community (NC) is then introduced to walk through both PN and VN by using the element of connections say connectivity feature. Through modeling NC by the theme of network transposition and egocentric network, we try to comprehend all possible networks for detecting the problems and solutions by using both sides' idea. Experimental results show that we can model real-world problems and then analyze them through NC by measurable and structural manner.

▶Keyword: Physical Network, Virtual Network, Connectivity feature, Networked Community

1. Introduction

네트워크의 홍수로 이루어진 현대 사회에서 우리는 인지하거나 또는 인지하지 못하는 사이에 네트워크에 소속되고 이들을 이용하면서 살아간다. 생활 속에서 매일같이 이웃과 만나지만 웹 기반의 소셜네트워크 서비스 (SNS) 등을 통해 또 다른 '이웃'들과도 만나게 되는데, 이러한 현상은 그림 1과 같이 개인적인 연결들을 범세계적인 연결로 표기할 수 있는 개념의 근본이 된다.

이러한 관점에서 '사회(Society)'는 사회과학적 연결로서 가족, 친구, 이웃 등 물리적 또는 지리적으로 연결되어 있는 경우와 공학적 연결로서 가상의 연결을 지칭하게 된다. 본 연구에서는 이러한 연결의 개념들을 물리적 네트워크 (Physical Network; PN)와 가상 네트워크 (Virtual Network; VN)로 정의하는데, 이 정의를 통하여 이러한 네트워크들이 어떻게 생성되고

연결되며 상호관계를 가지는지, 또한 어떠한 요소들이 이러한 연결의 상호관계를 설명할 수 있는지 알아내고자 한다. 현재까지는 사회과학과 공학적 관점의 연결이 나뉘어 연구된 바 있으나, 이를 통합하고자 하는 시도는 이제 시작단계에 있다.

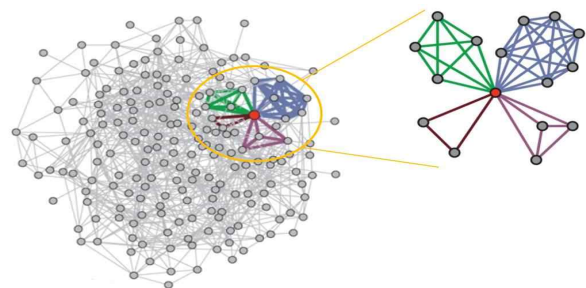


Fig. 1. A person's networks in global-scale network [1]

본 연구는 '사용자가 다양한 네트워크들 가운데 현재 위치를 파악할 수 있는가'라는 의문에서 시작되었다. PN과 VN 상의 친구 중에서 누가 더 가까운가, 두 네트워크 중 어디에서 문제가 시작되었는가와 같은 이런 의문은 'Six degrees of separation'[14]과 같이 모든 네트워크의 위계를 수평적으로 측정하여 해결할 수도 있겠으나, 이러한 접근으로는 개인이 가진 연결의 특성을 파악하는 데 한계가 있다. 때로 한 번도 보지 못한 SNS상의 친구가 바로 옆집의 이웃보다 가깝다고 느끼는 것은 어떻게 설명할 것인가? 이러한 현상에 따라 사회생활의 경계가 모호해지는 현상이 발생하는데 그 원인을 파악하지 못하고 있다. 이는 PN과 VN의 구분이 모호해지고 그 경계를 혼동하는

• First Author: Soungwoong Yoon, Corresponding Author: Sang-Hoon Lee
*Soungwoong Yoon (ysw1209@gmail.com), Dept. of Computer Science, Korea National Defense University
**Sang-Hoon Lee (hoony@kndu.ac.kr), Professor, Dept. of Computer Science, Korea National Defense University
• Received: 2017. 05. 22, Revised: 2017. 06. 11, Accepted: 2017. 06. 23.

것이 자연스러워졌으며, 개인의 연결에 관한 문제들이 두 네트워크를 망라하여 일어나기 때문이라고 할 수 있다. 즉 하나의 네트워크에서 문제를 해결하는데 다른 네트워크의 정보와 아이디어가 필요해졌다는 의미이다.

본 논문에서는 먼저 기존 연구의 방향에 따라 사회현상 전체의 네트워크를 PN과 VN이라는 두 가지 개념으로 분류하여 정의하고, 이 두 네트워크간의 상호영향성을 관찰을 통하여 평가하고자 한다. 다음으로 PN과 VN을 포괄하고 상호영향성을 측정 및 평가할 수 있는 새로운 네트워크 개념인 Networked Community (NC)를 제안하고자 하는데, 이 NC를 구성하는 요소로서 Connectivity feature (cf)도 개념화한다. 네트워크 이행 (Transition) 또는 자기화 네트워크 (Ego-centric network) 개념을 응용하여 NC의 연결을 정의함으로써 관찰되는 모든 네트워크를 통합, 각 네트워크의 문제점을 복합적으로 진단 및 해결하는 방법을 찾고자 하였으며, 실험을 통하여 NC 개념의 유용성을 보였다.

본 논문은 다음과 같이 구성된다. 2장에서는 관련연구를 살펴보고, 3장에서는 PN과 VN의 개념을 제시하고 이들 네트워크에서 일어나는 현상을 고찰한다. 4장에서는 NC 개념을 제시하고 5장에서 실험을 통해 그 유용성을 검증한다. 6장에서는 결론 및 향후 연구방향을 제시한다.

II. Related Works

네트워크의 연구는 사회과학과 공학 분야로 나누어진다. 두 분야에서 유사한 목표를 서로 다른 요소를 이용하여 분석하고자 한다고 볼 수 있는데, 특히 PN의 관점에서 VN을 해석하고자 하거나 그 반대의 연구가 주류를 이루었다.

1. Network Analysis

네트워크는 그 구성단위인 노드 (Node)의 물리적 연결을 뜻한다. 노드로 이용되는 것은 제한이 없고[8], 대표적인 네트워크의 분석은 노드간의 연결인 링크 (Link)를 이용한 최적화 방법이다. [3, 4] 최근 네트워크를 시스템을 이해하는 중요한 요소로 보고 노드의 의미를 확장하여 사회의 현상들을 해석하고자 시도한 연구도 있다. [1]

네트워크의 통합은 대상 네트워크의 동질적 연결성 (Homogeneous connectivity)을 기초로 하는 연구가 대표적인데, 관련 개념으로는 개념적 또는 최적화된 네트워크의 연결성을 정의하고자 하는 Network integration [17]과 유사한 연결성의 네트워크 상 공통요소를 이용하여 정의하고자 하는 Network fusion [18] 등이 있다. 이들 연구는 사회 현상을 이질적 네트워크의 연결성으로 정의하고자 하는 본 논문의 방향과는 차이가 있다.

현대의 대표적인 네트워크로 일컬어지는 인터넷은 개념적으로 연결된 광범위한 정보공간으로서, 사회와 연관하여 분석되는 소셜네트워크 (SN)의 연구는 주어진 SN에 대한 네트워크 관점의

분석과 그 문제점을 탐구하는데 중점을 두고 있다. [2, 15, 4, 11]

네트워크 구조를 실제 사회와 비교[7] 또는 그 반대의 연구 [11]도 있으나, 현재는 분석을 위한 요소를 추출하는 단계이다. 본 연구에서는 기존 연구에서 제시된 네트워크 개념과 분석방법을 효과적으로 결합할 수 있는 개념적인 방향을 제시하고자 한다.

2. Community

공동체는 지역적으로 제한된 사회의 일종[15]이다. 공동체 관련 연구는 구성원의 네트워크 구조와 상호작용을 통해 공동체 자체를 발견[1]하는 방법이 주류를 이룬다. 인터넷의 영역에서 공동체는 상대적으로 조밀하게 연결된 네트워크 구조로 정의[5]되기도 하는데, 이러한 공동체 발견을 위한 연구에는 지역적 이웃구조인 모듈성 (Modularity)을 이용한 연결구조 추정[16]과 군집을 이용한 공동체 성향 분석 (Weighted Community Clustering; WCC) [12] 등이 있다.

그러나 이러한 연구들은 주어진 네트워크 속에서 직관적으로 존재하는 공동체를 각자의 방법으로 검증하는 것이다.[6] 실제로 혼재된 네트워크에서 공동체를 정의하는 연구는 연결 공동체 (Link community) 개념을 통해 내부적으로 긴밀히 연결된 링크를 찾아 공동체를 판별하고자 하였다.[1] 본 연구에서는 공동체의 내부 및 외부적 연결요소를 포괄할 수 있는 인자에 대하여 정의하고, 이를 사회현상을 설명하는 방법으로 사용하고자 한다.

III. Concepts for Networks

네트워크 분석을 위하여 먼저 네트워크의 개념을 분류하고자 한다. 네트워크의 요소를 밝히고자 하는 기존 연구 [1, 14]와 관찰을 통하여 두 가지 개념으로 나누도록 한다.

1. Physical Network (PN)

사회과학적 의미의 네트워크는 연결 관계의 강도와 시계열적 측정이 곤란한 네트워크이다. 이 네트워크의 구성원인 개인은 자연스럽게 가족, 이웃, 친구라는 네트워크에 소속되게 되며 이는 자신의 의도와 무관한 경우가 많다. 이러한 형태의 네트워크를 PN으로 정의하며, 그 성격은 다음과 같다.

- PN은 의도되지 않은 연결을 통해 구성된다. 이에겐 출생, 이사, 입학 등의 현상을 들 수 있으며, 따라서 PN의 연결은 인식할 필요가 없는 연결로서 내재적인 연결로 불리기도 한다. 단일 PN의 연결성은 쉽게 연결되나 다수 PN의 연결성을 교차 분석하는 것은 매우 어렵다.

- PN은 매우 견고하게 연결되어 있으며, 이 연결은 의도적인 단절에도 잘 끊어지지 않는다. 단절의 예로 이사, 거절 또는 사망 등이 있는데, 때로는 사망 이후에도 제사 등 연결성이 어

는 정도 존재하기도 한다.

- PN의 구성원이 되기 위하여 개인정보를 직접적으로 요구하지는 않는다. 국적과 같이 PN이 구성될 때 필수적으로 파악되는 예외적인 경우도 있으나, 대체적으로 명시적으로 구성원이 되기 이전에 이미 구성이 완료되기도 한다.

PN의 연결은 공학적인 접근으로 밝혀내기 어려운 경우가 많다. 최근 생명공학의 연구는 인체라는 PN을 DNA라는 연결로 밝힌 것이라 할 수 있다. 만일 연결을 이루는 요소를 언어라 한다면 언어 기반의 PN은 국가 또는 지역사회라 지칭할 수 있게 된다.

대표적인 PN은 아래와 같다.

공동체 (Community) : PN에 경계를 둔다면 사람의 집단 단위는 가족, 친구, 동료 등의 개념으로 파악할 수 있다. 공동체란 일정 지역에 함께 거주하거나 취미가 비슷한 사람들의 집단 정도로 일컬어지는데, PN의 의미에서 공동체란 같은 목표를 가지고 삶을 공유하는 집단으로 정의하기도 한다. 정의하기는 어렵지만 모종의 요건이 공동체를 공고하게 결속하는데, 엄격하게 PN의 요건을 정의한다면 생활공동체(Gated community) 또는 공산주의 등으로 이해된다.

사회 (Society) : 사회는 본래 사람 자체를 의미하는데, 표현하자면 높은 단계의 공동체라 하겠다. PN으로서의 사회는 지역적 제한 등의 공동체의 요소를 느슨하게 적용한 경우로서, 예로 사회의 요소인 애국을 평소 느끼기는 어렵지만 세금관계 등 법적 분쟁요소나 국난의 위기 등에서는 비로소 작동함을 체감하게 된다. 그러므로 사회라는 PN의 요는 견고하게 연결되어 있기는 하나 측정이 어려운 요소도 있음을 알 수 있다.

2. Virtual Network (VN)

한편 PN과 같이 자연발생적이거나 지역적 제한을 받는 네트워크 이외의 네트워크도 존재하는데, 이를 VN이라 정의한다. VN은 통상적인 '네트워크'의 의미에 가까운 개념으로서, 느슨한 외연적 연결로 이루어진 오래전부터 존재해왔던 네트워크이다. VN의 예로 우편시스템을 살펴보면, 사람들은 우편을 통하여 연결을 유지하여 왔으며 시대적인 기술개발의 흐름과 함께 전화, 전보, 인터넷 및 휴대폰 등으로 그 범위가 확장된 것을 알 수 있다.

VN의 성격은 다음과 같이 규정될 수 있다.

- VN의 구성원이 되는 것은 PN에 비하여 제한이 적다. 요구되는 개인정보 항목만 입력하면 VN에 등록할 수 있으며, 이후 자유롭게 다른 VN 구성원들과 연결을 유지할 수 있다. 구성원이 된 이후의 VN 연결관계는 PN의 경우와 유사하다고 볼 수 있다.

- VN의 연결은 PN에 비해 쉽게 끊어질 수 있는데, 이는 접속 종료 등을 통하여 가능하다. 이때 개인정보는 실제 VN상에 다른 형태의 연결정보로 남아있게 되는데, 원하지 않을 경우 영구적인 삭제가 이론적으로 가능하다. 이 분야는 현재 활발하게 연구되고 있으나 아직 완전한 해결방법을 찾지 못했다.

- VN 구성원들은 대체로 개인정보의 일부분을 의도적으로

공개한다. 이와 같이 공개된 개인정보는 VN의 연결성을 정의하는 조건이 되는데, VN간의 연결에도 사용될 수 있다.

VN에서 사용되는 주요한 개념들은 다음과 같다.

네트워크 : 네트워크는 통상 VN과 동일한 개념으로 사용되는데, 이는 사회가 공동체의 집합과 같은 개념으로 사용되는 것과 같다. WWW와 같은 광대역 네트워크에서 개인은 상시 연결된 노드와 같은데, 이는 WWW상의 연결을 제외하면 일상생활이 매우 어렵기 때문이다. 경계가 없는 VN을 네트워크라고 말하기도 하는데, 이는 네트워크상 연결의 정도를 판단하는 기준이 사용자에게 따라 달라지기 때문이다. 이러한 개념에서 VN은 때로 PN과 혼동되어 사용되며, 나비효과와 같은 극단적인 연결의 VN 개념도 등장하게 된다.

소셜네트워크 (SN) : VN을 구성하는 노드수를 제한하면 이는 SN과 같이 측정가능한 VN이 된다. SN은 가장 성공적인 VN 모델중 하나인데, SN의 구성원이 되기 위해서는 일부 개인정보를 반드시 입력해야 하며, 이에 대한 보상으로 다른 SN 구성원들의 연결이 부여된다. 최근 연구동향은 SN이 사회현상을 대별하는 네트워크 분석의 대표적인 형태로 가정되었다.

PN과 VN의 구분에서 사용된 용어와 그 특징은 표 1과 같다. PN에서 사용된 연결의 성질은 가족과 같이 대체로 간단한 하나의 개념으로 되어있는 반면, VN에서의 연결의 성질은 나이, 성별, 지역 등 다양한 형태로 나타나는 것을 알 수 있다. 이러한 연결의 특징에 따라 PN의 연결 형태는 VN의 경우보다 그 경계를 정하기 어려운 특징이 있다.

Table 1. Terms and characters of PN and VN

Terms	PN	VN
Unit	People	Node
Connection	Relationship	Link, Edge
Degree	Intimacy, Friendship	Closeness, Weight
Characters	PN	VN
Connection theme	Singleton	Multiple
Connection topology	Hierarchical	Equivalent
Measuring degrees	Variant	Binary
Network boundary	Concrete	Sparse

3. Agglomerative Effects of PN's and VN's

PN과 VN은 서로 다른 네트워크를 비슷한 용어를 사용해서 정의하고 있으며, 우리 주변에 혼재되어 존재하므로 개인이 네트워크 내에서 위치를 정하는 데 있어 혼동을 준다. 이러한 네트워크를 망라하는 효과는 다음과 같이 관찰된다.

관찰 1. PN → VN : PN에서의 현상은 VN의 현상으로 재해

석될 수 있다. 정의적으로 PN의 한 사건 (event)은 VN에서 일어날 사건의 특정한 확률로 나타낼 수 있다.

예를 들어 생활공동체의 구조를 PN의 관점에서 파악하기는 매우 어려운데, 이는 공고한 내부 결속을 다지는 정보가 자체 기준으로 만들어져 있고 외부에 알려지지 않기 때문이다. 이를 VN의 관점에서 파악한다면 언론보도 등 간접적인 정보 (또는 이에 대한 반응)를 이용하여 알려진 분석방법을 이용하면 PN의 조사방법을 대체하는 효과가 있다는 것이다. 물론 생활공동체의 구조가 보편적인 네트워크의 규칙을 따른다는 가정이 필요하지만, 가장 복잡한 데이터 중심의 네트워크가 VN이라 한다면 생활공동체의 내부정보가 어느 정도 포함되고 분석 가능하다고 볼 수 있다.

이러한 착안을 확률적으로 정리하면 $Pr(PN \rightarrow VN)$ 이 된다. 현재까지의 PN 관련 연구에서의 네트워크 분석은 $Pr(PN \rightarrow VN)=1$ 로 가정한 것인데, PN 단독의 환경에서는 이와 같이 생각할 수 있지만 PN과 VN이 공존하는 환경에서는 확률화하기 어렵다. $Pr(PN \rightarrow VN)$ 은 PN과 VN의 연결확률로 이해될 있는데, 특정 관점을 기준으로 한다면 $Pr(PN \rightarrow VN)$ 은 0 또는 1로 단순화될 수 있음을 착안할 수 있다. 이 확률은 PN상의 연결의 유효성을 판단하는 기준에 따라 다르다.

관찰 2. $VN \rightarrow PN$: VN상의 문제는 PN상에서도 문제를 일으킨다. 정의적으로 VN의 한 사건은 PN에서 일어날 사건의 특정한 확률로 나타낼 수 있다.

VN에서 PN쪽으로는 연결은 그 결과에 치중되어 있는데, 이는 일반적으로 PN의 분석방법론으로 VN을 분석하기 어렵고 여겨지기 때문이다. 그러나 VN에서 일어나는 사건은 PN에도 명시적으로 영향을 미치게 되는데, 이는 최근 구체적으로 인과관계가 밝혀지는 왕따 문제 등에서도 관찰할 수 있다. 이때 VN의 사건은 PN상의 문제점의 전조로 받아들여지기도 한다.

이러한 내용을 수식으로 표현하면 $Pr(VN \rightarrow PN)$ 과 같은데, 직접적으로 $Pr(VN \rightarrow PN)$ 을 연구한 내용은 드물지만 $Pr(VN \rightarrow PN)=1$ 로 보고 사회현상을 해석하고자 하는 연구는 특히 생물학 및 병리학 분야에서 매우 활발하다. [9] 그러나, 이 연구들 또한 사람에게 영향을 미치는 주된 요소들에 대한 제한적인 VN을 모델화한 경우가 대부분이다.

위 두 관찰내용과 같이, PN과 VN의 연결관계는 필요충분관계가 아니다. 그럼에도 불구하고 일반적으로 $Pr(PN \rightarrow VN) \geq Pr(VN \rightarrow PN)$ 으로 받아들여지고 있으며, 기술 개발과 더불어 새로운 측정요소의 발견에 따라 $Pr(VN \rightarrow PN)$ 의 측정값이 갈수록 커지고 있다. 이는 VN의 규모가 PN 수준 이상으로 커지는 현상의 현상을 반영한다고 할 수 있다.

본 연구에서는 단순화를 위하여 네트워크의 범위를 사람의 네트워크로 한정하고, VN에서 사용한 용어를 PN에도 사용하기로 한다.

IV. Networked Community

혼재된 네트워크의 효과를 측정하기 위해서는 PN과 VN에 동시에 접근할 방법이 필요하다. 이를 위해 네트워크 내 노드의 연결뿐만 아니라 네트워크간의 연결에도 사용될 수 있는 요소들이 노드 내에 존재한다고 가정하고, 이를 연결요소(connectivity feature; cf)로 정의한다. cf는 네트워크 N의 노드 n_i 가 N'의 n_j 를 만났을 때 측정 가능하며, cf의 집합은 CF이다.

그림 2는 $|CF|=4$ 일 경우 cf의 개념을 나타낸다. 노드 'John'의 cf는 연결된 노드에 따라 다른 값으로 작용하고 때로는 무시해도 될 만큼 작아진다. 노드 '@john'과 '@mack'은 cf2에 의해 연결되어 있으나 그 값이 매우 작는데, 이는 cf2가 해당 SNS의 부도 확률로 측정되었기 때문이다. 하지만 'John'과 'Mother' 사이의 cf3=0.8은 큰데, 이는 모자간의 관계에서 매우 자연스러운 현상이다.

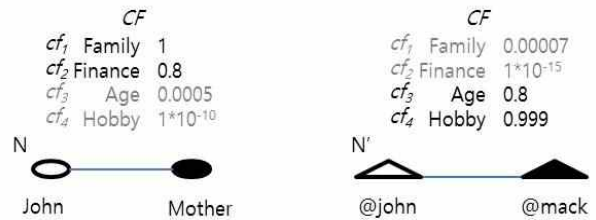


Fig. 2. Concept of connectivity feature

cf가 독립적이라 하면, cf의 활성화 확률(Pr_{CF})은 cf가 이용되는 연결을 설정하기 위한 최대값이며 아래와 같이 산출된다.

$$Pr_{CF}(n_i, n_j) = \underset{CF}{argmax} Pr_{cf}(n_i, n_j) \quad (1)$$

그림 2에서 $Pr_{CF}(John, Mother) = Pr_{cf1}(John, Mother) = 1$, $Pr_{CF}(@john, @mack) = Pr_{cf4}(@john, @mack) = 0.999$ 이다. Pr_{CF} 를 이용하면 네트워크의 경계 설정이 가능하며, 노드 $n_i, n_j \in N$ 의 관계를 설명할 수 있다. 가장 간단하게는 두 노드가 연결될 때 $Pr_{CF}(n_i, n_j) = 1$, 이외는 0이다. 모든 네트워크의 집합을 N이라 하고, 이를 구성하는 노드들의 조합에 따른 Pr_{CF} 에 의하여 연결이 결정된다고 가정하면, $N, N' \subset N$ 의 이행성(Transitivity)은 다음과 같이 추정할 수 있다.

가정 1. 노드 $n \in N$ 은 다른 노드 $n' \in N'$ 으로 확률 $Pr_{CF}(n, n')$ 에 따라 이행될 수 있다.

그림 2에서 노드 'John'과 '@john'은 서로 다른 네트워크에 소속되어 있지만 실제로 같은 사람인데, 이는 두 네트워크의 '관점'이 다른 것을 의미한다. 가정 1에 따라 'John'과 '@john'은 $Pr_{CF}(John, @john)$ 로 연결될 수 있으며 이 값은 1에 가까워

야 하는데, 반대로 두 노드의 유사도를 Pr_{CF} 를 이용하여 측정할 수 있다고도 할 수 있다. 이와 같이 N 과 N' 은 Pr_{CF} 에 의하여 새로운 네트워크로 이행되는데, 이 네트워크를 Networked Community (NC)라고 정의한다.

기술적인 NC의 형성은 특정 노드 n_i 로부터 시작된다. n_i 와 연결된 노드들의 Pr_{CF} 를 산정하여 NC의 경계를 설정하는데, 이를 수식 1을 이용하여 표현하면 다음과 같다.

$$Pr_{CF}(n_i, n_j) = \underset{CF}{argmax} Pr_{cf}(n_j | n_i) \quad (2)$$

수식 2를 통한 NC의 구성은 시작노드로 선택된 노드 n_i 의 관점에서 네트워크를 정의하는 것이다. n_i 을 선택할 때는 빈도, 중앙성 등 다양한 기준을 통해 가장 영향력 있는 노드를 선택하는 것이 일반적이다. NC 구성방법은 아래와 같이 정의된다.

정의 1. NC는 시작노드 n_i 와 $Pr_{CF}(|n_i)$ 로 연결된 모든 노드의 집합이다

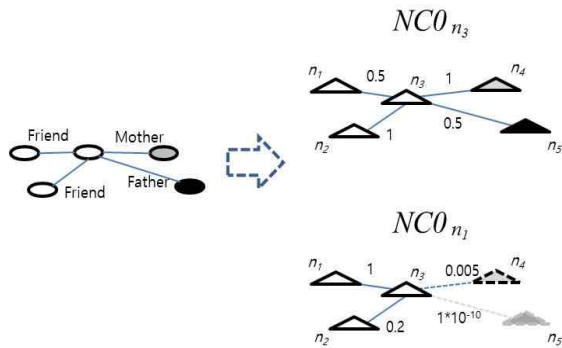


Fig. 3. Modeling NC

그림 3은 NC를 형성하는 과정을 나타낸 것으로서, 가상 이행(Pseudo-transition), 네트워크 흡수 또는 절대확률적 Pr_{CF} 조건이라고도 말할 수 있다. 이는 Pr_{CF} 가 NC의 모든 노드들에 대하여 상수로 정해질 수 있기 때문인데, 이런 네트워크는 실제 세계에서는 발견되기 어렵지만 NC를 통하여 PN과 VN을 연결하고 분석과정을 연구하기에 알맞은 형태라 할 수 있다.

V. Experiment

실험에서는 cf의 일반적인 효과를 측정하기 위해서 |CF|의 범위를 변화시키면서 적용하였다. 이를 위하여 보편적으로 PN이 구성되어 있다고 받아들여지는 네트워크 ($|CF| < \infty$)를 실험대상으로 선정하였으며, NC로 구성되었을 때 발견되는 네트워크의 특징을 분석하였다.

1. Diagnosis for Closed Organization [10]

일반적으로 조직은 목표 달성을 위해 수직적 관계가 강조되면서 업무 효율의 향상을 위하여 구성원의 업무강도나 중요성을 판단하고자 한다. 그러나 직원간의 연결은 업무적 관점과 개인적 관점이 혼재되어 있으며, 이러한 요소들을 판단하기 위한 조직진단은 구성원 개인 단위로 검사하고 이를 모아 분석해왔다.

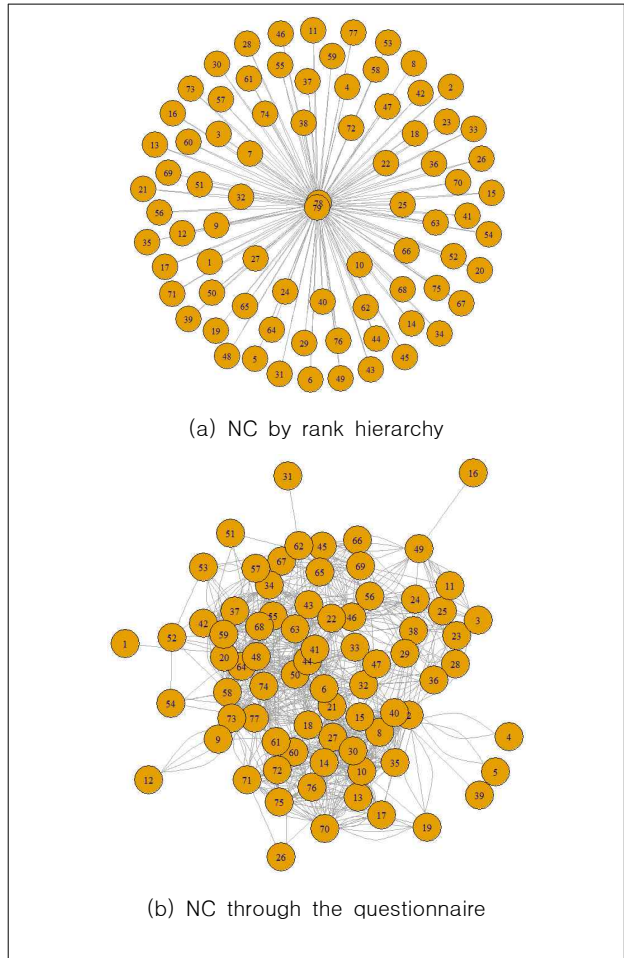


Fig. 4. Drawing NCs #1

실험을 위하여 군 단위부대를 대상 조직으로 선정하였으며, 한 조직의 인원수는 70~400명 수준이다. 이들을 계급에 따라 구성한 NC의 예는 그림 4(a)와 같은데, 통상적인 조직진단에 사용되는 이러한 네트워크 형태를 통해서 조직의 현상이나 특정한 인원을 발견하기 어렵다.

한편 구성원들에 대한 설문조사를 통하여 조직내 긴밀도를 측정했는데, 기존의 사회과학적 조사방법을 일부 변경하여 조직 내 본인 이외의 구성원에 대한 평가를 하도록 하였으며, 4가지 분야의 질문에 해당되는 구성원을 최대 7명까지 선정하도록 하였다. 이러한 조사를 통한 NC의 예는 그림 4(b)와 같다.

이러한 조사를 통하여 기존 개인단위에서는 측정하기 어려웠던 특정 조직원을 확인할 수 있었는데, 그림 4(b)에서 16, 31번과 같은 네트워크 외곽에 존재하는 구성원 등이다. 이는 조직

내 왕따와 같은 문제점들을 잠재적으로 가지고 있는 인원을 판별하는데 도움을 주었다.

군 조직에서는 심리적 문제가 있는 인원들을 A(심각), B(관찰요망), C등급(경미)으로 분류하여 관리하고 있으며, 이는 의사 및 조직 리더의 장기간의 관찰을 통하여 지속적으로 분류되고 관리된다. 많은 노력이 드는 이 분류를 그림 4(b)의 NC를 통해 추정할 수 있었으며, 그림 5는 이 추정 결과의 일치 정도를 보여주는데, 특히 실험대상 조직내 심각단계의 구성원들을 모두 알아내는 성과를 보였다.

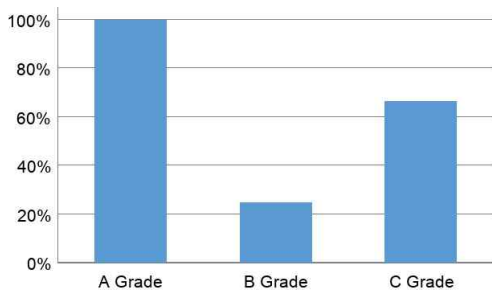


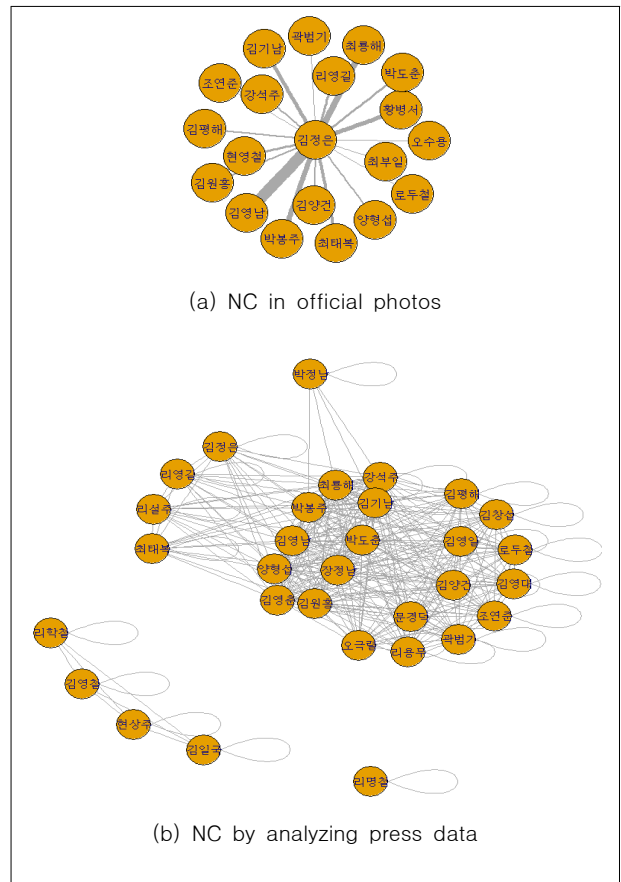
Fig. 5. Ratio of detecting at-risk members

2. Intra-network Changes [19]

이 절에서는 NC의 개념을 통해 북한의 권력 구조가 간접적인 정보를 통해 효과적으로 파악될 수 있음을 보였다. 실험의 대상으로 북한 권력층을 선택한 이유는 1) 북한은 PN으로 여겨질 수 있는 견고한 내부 네트워크가 존재한다고 추정되며, 특히 권력층은 상호 밀접하게 연결되어 있다고 알려졌고, 2) 북한은 내부적으로 정보가 철저히 통제되고 있으므로 PN 관련 자료를 직접 수집하기 어려워 NC를 통한 간접적인 구조 파악방법이 유용하게 적용될 수 있기 때문이다.

간접정보를 이용한 북한 권력층의 PN구조를 파악하기 위한 기존 연구는 최고지도자가 권력층과 찍은 공식사진을 이용하였는데, 이 사진에서의 위치를 PN으로 가정하였다. 특히 매년 12월 공식행사에서 촬영되는 공식사진 상의 자리 배열이 권력 서열을 의미한다고 여겨지고 있으며, 이 정보를 이용한 NC 구성의 예는 그림 6(a)와 같다.

외부로 알려진 정보는 북한 매체들의 보도가 대표적인데, 한국 정부가 발행하는 '북한 동향'을 통해 제공되고 있으며 북한 국영 TV 보도 및 국영 신문 (노동신문)의 보도가 망라되어 있다. 이 '북한 동향' 자료에서 보도된 사건과 그 참가 인물들을 기반으로 NC를 구성하였는데, 인물을 노드로, 사건을 서브 노드로 한 2-모드 SN기법을 활용하였다. 이를 이용한 NC 구성의 예는 그림 6(b)와 같다.



(a) NC in official photos

(b) NC by analyzing press data

Fig. 6. Drawing NCs #2

인물들의 중요성을 산정하기 위하여 SNA의 Eigenvector centrality[13] 기법을 이용, 한 인물의 중요도가 다른 인물들에게 미치는 영향을 포함하도록 했다. 표 2는 NC 구성 방법에 따른 효과를 보여준다.

Table 2. Person match ratio on NCs

Year	2012	2013	2014
Precision@	@32	@30	@20
by Degree Centrality	0.813	0.700	0.600
by Eigenvector Centrality	0.906	0.867	0.633

Precision@의 차이는 각 연도별 공식사진 상의 인물 수에 따라 조정했으며, 실험결과 기존의 방법[13]보다 비교적 정확하게 권력층 인물들이 나타남을 알 수 있다.

특히 특정 인물의 순위 변화는 북한의 체제 변화를 가늠하는 중요한 척도인데, 3년간의 NC상 중요도 변화를 측정할 결과 그림 7(a)에서 보인 권력층에서 실제 순위 하락이 추정된 인물들은 그림 7(b)와 같이 NC 실험치 역시 동일한 하락경향을 보인 것으로 나타났다.

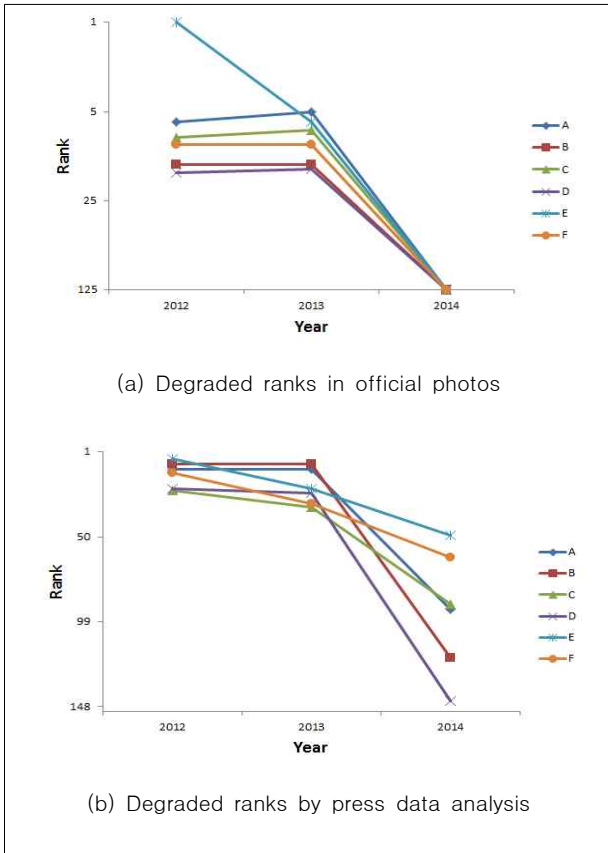


Fig. 7. Personal-level evaluation: Degraded ranks

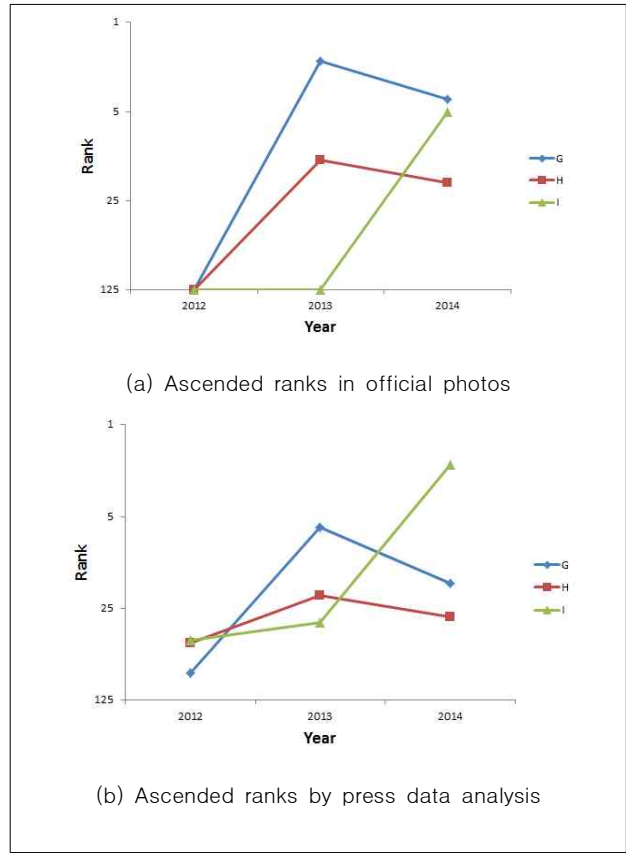


Fig. 8. Personal-level evaluation: Ascended ranks

더욱 흥미로운 사실은 권력층상 순위가 상승한 인물들에 대한 실험 결과인데, 이는 기존의 권력층에 포함되지 않았던 것으로 간주되던 인물들의 예측결과도 얻어낼 수 있기 때문이다. 그림 8(a)와 같이 기존에 밝혀지지 않아 공식 사진에 포함되지 않았던 인물들도 그림 8(b)와 같이 NC 분석을 통하여 미리 파악이 가능했으며, 그 변화추이 또한 실제와 비슷했다.

이러한 실험결과는 예측을 위해서도 NC가 유용하며, 특히 cf의 우세도 측정 결과의 집합인 CF를 이용하여 개별 노드의 현재 상태와 그 변화의 측정도 가능함을 시사한다.

VI. Conclusions

본 연구에서는 네트워크를 PN과 VN으로 개념화하고 혼재된 네트워크의 현상을 관찰하였으며, 이에 기초한 모든 네트워크의 연결에 영향을 미치는 cf의 개념과 이를 이용한 새로운 네트워크의 개념인 NC를 제시하였다. NC는 PN과 VN을 포괄하여 사회현상을 분석하는 데 유용하며, 실험결과 사회현상과 그 변화를 추정할 수 있었다.

앞으로 NC의 구조와 범용성, 구성방법에 대한 추가적인 연구가 필요하며, NC 모델의 검증에 사회현상을 반영하여 발전

시켜야 한다. 또한 cf의 증가를 지속적으로 처리 가능한 NC 모델의 개발이 필요하다.

REFERENCES

- [1] Y.-Y. Ahn et al., "Link communities reveal multi-scale complexity in networks," *Nature*, 466:761-764, 2010.
- [2] M. V. Vieira et al., "Efficient search ranking in social networks." *CIKM '07*, pp. 563-572, 2007.
- [3] L. Xin-ru and G. Haijuan. "International human resource management research focus and cutting-edge analysis." *ISMOT 2012*, pp. 495-500, Nov 2012.
- [4] B. Cao et al., "Graph-based workflow recommendation: On improving business process modeling." *CIKM '12*, pp. 1527-1531, 2012.
- [5] T. Chakraborty et al., "On the permanence of vertices in network communities." *KDD '14*, pp. 1396-1405, 2014.
- [6] R. Devooght et al., "Random walks based modularity: Application to semi-supervised learning." *WWW '14*, pp. 213-224, 2014.
- [7] Y. Kim et al., "Trailing organizational knowledge paths

through social network lens: integrating the multiple industry cases." J. Knowledge Management, 18(1):38-51, 2014.

- [8] M. E. J. Newman. "The structure and function of complex networks." SIAM REVIEW, 45:167-256, 2003.
- [9] C. Poletto et al., "Quantifying spatiotemporal heterogeneity of mers-cov transmission in the middle east region: A combined modelling approach." Epidemics, 15:1-9, 2016.
- [10] D. Park and S. Lee, "Diagnosis model for closed organizations based on social network analysis." KIISE Trans. on Comp. Prac. Vol.21, No.6, pp. 393-402, 2015.
- [11] G. Park et al., "Influencerank: Trust-based influencers identification using social network analysis in q&a sites." IEICE Transactions, 95-D(9):2343-2346, 2012.
- [12] A. Prat-Perez et al., "High quality, scalable and parallel community detection for large real graphs." WWW '14, pp. 225-236, 2014.
- [13] J. Scott. Social network analysis. Sage, 2012.
- [14] J. Travers and S. Milgram. "An experimental study of the small world problem." Sociometry, 32:425- 443, 1969.
- [15] S. Wasserman. "*Social network analysis: Methods and applications*", volume 8. Cambridge university press, 1994.
- [16] J. Zhang, et al., "Who proposed the relationship?: Recovering the hidden directions of undirected social networks." WWW '14, pp. 807-818, 2014.
- [17] K. Xu et al., "A tutorial on the internet of things: from a heterogeneous network integration perspective." IEEE Network, 30(2), pp. 102-108, 2016.
- [18] B. Wang et al., "Similarity network fusion for aggregating data types on a genomic scale." Nature Methods 11, pp. 333-337, 2014.
- [19] W. Lee et al., "Evaluation of Structural Changes of a Controlled Group Using Time-Sequential SNA." Journal of KIISE Vol.43, No.10, pp. 1124-1130, 2016.

Authors



Soungwoong Yoon received the B.S. degree in Engineering from Hanyang University in 1992 and M.S. degree in Computer Science and Engineering from Korea National Defense University (KNDU), Korea, in 2004, respectively.

Mr. Yoon joined the Ph.D course of the Department of Computer Science at KNDU, Seoul, Korea, in 2015 and currently Ph.D candidate. He is interested in knowledge discovery, information retrieval, big data and social network analysis.



Sang-Hoon Lee received the B.S. degree in Computer Science from Sungkyunkwan University in 1978, M.S. degree in Computer Science from Yonsei University, Korea, in 1989 and Ph.D. degree in Computer Science from Kyoto University, Japan, in 1997, respectively.

Dr. Lee joined the faculty of the Department of Computer Science and Engineering at KNDU, Seoul, Korea, in 2000. He was the head of KNDU Computing Center, the director of KNDU library, the graduate school president and currently a Professor in the Department of Computer Science, KNDU. He is interested in database, information retrieval, big data and social network analysis.