

# 단백질 상호작용 네트워크에서 연결노드 추출과 그 중요도 측정†

## (Identifying Bridging Nodes and Their Essentiality in the Protein-Protein Interaction Networks)

안명상\*, 고정환\*\*, 유재수\*\*\*, 조완섭\*\*\*\*

(Myoung-Sang Ahn, Jeong-Hwan Ko, Jae-Soo Yoo, Wan-Sup Cho)

**요 약** 본 연구에서는 단백질 상호작용 네트워크에서 네트워크 견고성의 핵심노드는 허브노드이지만 연결노드 또한 허브노드와 같이 매우 중요한 역할을 하고 있음을 밝혀냈다. 네트워크 견고성에 가장 큰 영향을 미치는 핵심노드는 차수가 매우 큰 허브이다. 이 연구에서 새로이 제안된 연결노드는 단백질 상호 작용 네트워크에서 고밀도로 연결된 모듈들 사이에 위치하여 네트워크 견고성에 중요한 영향을 미치는 노드이다. 실제로 척도 없는 네트워크는 무작위 공격에 매우 강한 반면, 허브노드만을 제거하는 목표 공격에는 매우 취약하다. 기존에 핵심노드로 노드의 연결성을 연구한 방식과 사이 중앙성(centrality)을 기반으로 사이 중앙성(betweenness centrality)이 큰 노드가 전체 네트워크에서 핵심노드임을 연구한 방식이 있다. 그러나 노드의 연결성 즉, 차수와 사이 중앙성은 일정한 비례관계가 있어서 차수가 클수록 사이 중앙성 값이 커지기 때문에 기존의 두 연구는 큰 차이점을 갖지 못한다. 본 연구에서는 단백질 상호작용 네트워크도 척도 없는 네트워크이므로 노드의 차수와 사이 중앙성 간에 비례관계가 형성되지만, 차수가 작은 노드일수록 사이 중앙성 값이 매우 넓게 분포하고 있다는 특징에 착안하여 연결노드를 제안한다. 이 연구에서는 차수가 매우 작지만 사이 중앙성 값이 매우 큰 노드를 연결노드로 간주하여 실제 인터넷 상에서 공개된 생물 네트워크를 대상으로 견고성 분석 실험을 한다. 실험 결과 연결노드들이 생물 네트워크에서 제거되면 허브노드가 제거되는 것보다 더욱 큰 네트워크 분화가 발생하였으며, 이는 연결노드가 네트워크 견고성 측면에서 매우 중요한 요소임을 입증하는 것이다.

**핵심주제어 :** 네트워크 견고성, 단백질 상호작용 네트워크, 연결노드, 허브노드

**Abstract** In this research, we found out that bridging nodes have great effect on the robustness of protein-protein interaction networks. Until now, many researchers have focused on node's degree as node's essentiality. Hub nodes in the scale-free network are very essential in the network robustness. Some researchers have tried to relate node's essentiality with node's betweenness centrality. These approaches with betweenness centrality are reasonable but there is a positive relation between node's degree and betweenness centrality value. So, there are no differences between two approaches. We first define a bridging node as the node with low connectivity and high betweenness value, we then verify that such a bridging node is a primary factor in the network robustness. For a biological network database from Internet, we demonstrate that the removal of bridging nodes defragment an entire network severally and the importance of the bridging nodes in the network robustness.

**Key Words :** Bridging Node, Protein-Protein Interaction Network

† 이 논문은 2007년 교육인적자원부의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임 (지방연구중심대학육성사업/충북BIT연구중심대학육성사업단)

\* 충북대학교 경영정보학과 박사수료

\*\* 안동대학교 통계학과 교수

\*\*\* 충북대학교 전기전자컴퓨터공학부 교수

\*\*\*\* 교신저자 : 충북대학교 경영정보학과/BK21 u-Biz팀 부 교수 (wscho@cbnu.ac.kr)

## 1. 서 론

지난 세기 생물학의 주된 연구 주제였던 환원주의(Reductionism)는 세포의 구성 물질에 대한 지식과 그 기능에 대한 많은 지식을 알게 해주었다. 이러한 발전에 불구하고, 환원주의적 지식은 분자수준의 생물학적 기능에 한정되어 있었다. "하나의 유전자가 하나의 단백질을 만들고, 하나의 단백질이 하나의 기능을 수행 한다"는 환원주의의 일차원적인 시각은 복잡한 생물학적 기능을 이해하는데 한계가 있다. 즉, 특정 생물학적 기능은 하나의 특정 단백질로부터 발현되지 않으며, 대부분의 생물학적 특징은 다양한 세포 구성요소들의 복잡한 연관관계로부터 발생된다[13].

21세기의 생물학의 중요한 연구 주제는 복잡한 세포 구성 요소로 이루어진 세포 네트워크에 대한 동적 관계와 구조를 이해하는 것이다[3]. 최근 Yeast 및 여러 생물체에 대한 지놈(Genome) 수준의 단백질 쌍 데이터베이스가 구축됨에 따라 단백질 상호작용 네트워크의 구조 및 특징을 그래프 및 네트워크 이론[1]으로 분석이 가능하게 되었다. 이러한 네트워크들은 다양한 환경 속에서 기능을 수행하기 위하여 환경의 변화에 적응할 수 있는 견고한 네트워크 구조를 지니게 된다. 따라서 생물 시스템의 네트워크를 이해하기 위해서 생물 네트워크의 구조적 견고성을 측정하고, 네트워크의 견고성에 가장 핵심적인 요소를 밝혀내는 것은 생물 시스템의 진화 과정뿐만 아니라 진화의 목적을 이해할 수 있는 중요한 연구이다[8].

생물 네트워크의 구조적 견고성을 분석하는 연구로는 중앙성과 연결성을 이용하는 방식이 연구되어 왔다[8,9]. 노드의 연결성에 근거한 방식[8]은 네트워크의 견고성에 큰 영향을 미치는 핵심노드로 노드의 차수만을 고려한다는데 문제점이 있다. 즉, 연결노드가 많은 허브 노드만을 핵심노드로 간주한다. 그러나 허브노드는 전체 네트워크에서 지역적 중요성만 고려하기 때문에 차수가 적으면서 네트워크 견고성에 중요한 영향을 미치는 노드를 발견할 수 없는 문제점이 있다.

본 연구에서는 생물 네트워크의 구조적 견고성을 측정하기 위한 새로운 네트워크 견고성 측정 방식인 연결노드 개념을 제안한다. 연결노드란 차

수가 작으면서 네트워크의 모듈간의 연결통로 역할을 하는 노드이다. 연결노드를 이용한 견고성 측정 방식은 사이 중앙성을 이용한 방식[13]과 연결성을 이용한 방식을 결합한 방식으로서 기존의 차수만을 고려한 방식에서 발생했던 허브노드만을 핵심노드로 간주하는 문제점을 해결하게 된다. 제안된 방식은 네트워크에서 노드의 차수가 작아도 네트워크 견고성에 큰 영향을 미치는 노드들을 새롭게 찾아냄으로써 더욱 정확한 견고성 측정이 가능하게 된다.

제안된 방법의 유효성을 입증하기 위하여 인터넷 상에 공개된 실제 생명 네트워크 데이터에 대하여 제안된 분석 기법을 적용하였다. 실험결과 제안된 방식에서 찾아낸 연결노드들을 제거하면 네트워크가 급격히 분화되어 노드들의 유기적인 연관관계가 붕괴됨을 확인할 수 있었다.

본 논문의 구성은 다음과 같다. 제 2 장에서는 관련연구를 살펴본다. 제 3장에서는 이 논문에서 제안하는 연결노드의 성질에 대하여 설명하고, 제 4 장에서는 실제 데이터에 적용하여 견고성을 분석한다. 제 5장에서 결론을 맺는다.

## 2. 관련연구

이 장에서는 네트워크의 기본 특징을 살펴본 후에 생물 네트워크의 견고성에 관한 기존 연구들을 살펴본다

### 2.1 네트워크의 기본 특징

네트워크를 구성하는 노드의 가장 기본적인 성질은 노드의 차수(연결선)  $K$  이다. 노드의 차수는 그 노드가 얼마나 많은 이웃 노드와 연결되어 있는지에 대한 측정 단위이다.  $N$ 개의 노드와  $L$ 개의 링크를 가진 무방향성 네트워크에서 각 노드의 평균 차수  $\langle K \rangle$ 는 무방향성 네트워크이므로  $2L$ 의 링크를  $N$ 개의 노드로 나누는 아래 식으로 계산할 수 있다.

$$k \geq \frac{2L}{N} \quad (\text{수식1})$$

기본적으로, 다양한 네트워크들의 위상구조는 네트워크의 차수 분포(Degree Distribution)  $P(K)$ 를 기준으로 분류된다. 차수 분포는 임의로 선택된 노드가 정확히  $K$ 의 차수를 가질 확률이다.  $P(K)$ 는  $K=1, 2, \dots, n$ 에 대하여  $K$ 차수를 가지는 노드의 개수  $N(K)$ 를 전체 노드수로 나누어 얻는다.

차수 분포는 각 네트워크의 위상구조를 구별할 수 있는 중요한 측정 단위이다. 차수 분포가 종모양 또는 포아송(Poisson) 분포에서는 대부분의 노드가 평균차수  $\langle K \rangle$ 개의 링크를 가지며, 평균차수  $\langle K \rangle$ 보다 링크수가 굉장히 많거나 역으로 굉장히 적은 노드는 그래프가 지수적으로 감소하기 때문에 거의 존재하지 않는다. 반면, 멱함수 분포는 차수가 굉장히 큰 허브(Hub)노드와 차수가 적은 대다수의 일반 노드가 공존하고 있으며, 허브노드는 일반 노드에 대한 링크를 갖는 구조를 가진다. 따라서 노드의 차수분포가 멱함수를 따르는 네트워크는 평균적인 차수가 존재하지 않기 때문에 척도 없는(Scale-Free)네트워크[2]라고 칭한다.

네트워크에서 임의의 두 노드( $i, j$ )간 경로 길이(Path Length)  $L_{ij}$ 는  $i$ 노드에서 출발해서  $j$ 노드에 도착하는데 몇 개의 링크 거쳐야 하는가를 나타낸다. 임의의 두 노드 사이에는 여러 경로가 존재하게 되고, 그 중 가장 짧은 경로가 최단경로(Shortest Path)이다. 네트워크의 항해가능성(Navigability)을 측정하는 평균 최단경로길이(Average Shortest Path Length)  $ASL$ 은 네트워크의 모든 노드쌍( $i, j$ )의 최단경로의 평균값이다. 아래 식(2)로 평균 최단경로를 계산할 수 있다.

$$\langle ASL \rangle = \frac{2}{N(N-1)} \sum_{i < j} l_{i,j} \quad (\text{수식2})$$

상대적으로 짧은 경로를 거쳐 항해 가능한 네트워크를 작은 세상 네트워크(Small World Network)[3]로 부른다. 작은 세상 효과는 현실 세계에 존재하는 복잡계 네트워크의 공통된 특징이다. 네트워크에서 임의의 두 노드의 경로는 몇 개의 소수 링크만 거치면 연결된다. 작은 세상 효과는 처음 사회 네트워크에서 발견되었으며[4], 뉴런 네트워크에서 인터넷 및 세포 네트워크까지 모든 복잡한 네트워크가 가지는 특성이다. 척도 없는 네트워크는 작은 세상 효과를 더욱 크게 가진다. 랜

덤 네트워크[5]인 경우 두 노드가 연결되기 위해 거쳐야 되는 평균 최단 경로 길이는 약  $ASL \sim \log N$ 으로서 실제 노드의 개수  $N$ 에 로그리즘을 적용한 결과이다. 반면 척도 없는 네트워크인 경우 훨씬 더 큰 작은 세상 효과(약  $ASL \sim \log \log N$ )를 지니고 있다[6].

네트워크에서 모듈이나 클러스터는 고밀도로 연결된 노드들의 집합이다. 네트워크에 존재하는 노드들은 대부분 클러스터링(Clustering) 경향을 띠고 있다. 네트워크의 이런 특징은 클러스터링 계수(Clustering Coefficient)  $C_i$ 를 이용하여 정량화할 수 있다[3]. 아래 식(3)은 노드  $i$ 의 클러스터링 계수를 계산하는 식이다[3].

$$C_i = \frac{2n_i}{k_i(k_i-1)} \quad (\text{수식3})$$

클러스터링 계수는 한 노드와 연결선을 가진 이웃 노드들이 상호 연결하고자 하는 정도를 정량화한 것이다. 클러스터링 계수는 그 네트워크의 모듈화 정도를 가리키는 척도가 된다. 현실에 존재하는 대부분의 네트워크들의 클러스터링 계수 값은 동일한 크기의 무작위 네트워크의 클러스터링 계수보다 크다[7]. 즉, 대부분의 네트워크들은 모듈화 구조를 이루고 있음을 이를 통해 알 수 있다.

네트워크 위상구조를 결정하는 또 다른 측정도구는  $AC(k)$  함수이다. 이 함수는 차수가  $k$ 인 모든 노드에 대하여 클러스터링 계수 값의 평균을 구한 것이다. 만약  $AC(k)$ 가  $k$  값에 독립적이면 그 네트워크의 위상구조는 조밀하게 밀집된 작은 클러스터들이 계층적 관계를 가지지 않고 있음을 나타낸다. 반면  $AC(k) \sim k^{-\beta}$ 에 있다면 그 네트워크에 존재하는 클러스터들은 계층적 구조를 이루고 있다. 현실세계의 실제 네트워크에서  $\beta$ 값은 보통 1이다. 즉, 노드의 차수와 클러스터링 계수는 역관계를 가진다. 이 구조는 차수가 작은 노드들의 클러스터링 계수 값은 크고 차수가 커질수록 클러스터링 계수의 값이 작아지는 형태를 보인다. 즉, 링크 수가 적은 노드들은 지역적으로 밀집된 모듈을 형성하기 때문에 클러스터링 계수 값이 허브노드보다 크다. 계층적 구조를 가지는 네트워크에서 서로 다른 클러스터들은 허브 노드에 의해 상호 연관 관계를 유지하게 된다[7].

## 2.2 생물 네트워크의 견고성 (Biological Network Robustness )

이 절에서는 기존의 네트워크 견고성을 측정하기 위한 대표적인 방법들을 소개한다.

네트워크의 견고성을 측정하는 방법 중 가장 대표적인 것으로 위상 견고성(Topological Robustness)이 있다[8]. 위상 견고성은 네트워크가 공격을 당하거나, 일부분의 기능이 정지됨에도 불구하고 온전한 기능을 수행할 수 있는 능력을 의미한다.

네트워크의 견고성을 측정하기 위하여 두 가지 시나리오가 필요하다. 첫 번째 시나리오는 매 단계마다 무작위로 선택된 노드들을 제거하여 그 네트워크의 견고성을 측정하는 방식이다. 두 번째 시나리오는 네트워크에서 선택된 노드만을 차례로 제거하는 방식으로서 특히 매 단계마다 차수가 가장 많은 허브 노드들을 제거해 나감으로서 견고성을 측정한다. 견고성 측정값으로는 평균 최단 경로길이(ASL)을 측정한다. 척도 없는 네트워크는 차수가 적은 수 많은 노드와 차수가 굉장히 큰 허브노드가 공존하는 구조이다. 이런 척도 없는 네트워크에서 무작위로 노드를 제거하면 대부분 선택되는 노드는 차수가 적은 노드가 되고 그 노드를 제거하더라도 ASL은 거의 변화가 없으므로 네트워크 견고성에 영향을 미치는 노드가 아니다. 반면 허브노드를 제거하면 ASL은 급격하게 증가한다. 즉, 허브노드 제거로 인하여 노드들이 상호작용하기 위한 경로길이가 급격히 늘어났으므로 허브 노드는 핵심노드이다.

생물학적 네트워크들은 노드의 차수분포가 멱함수 법칙을 따르는 척도 없는 네트워크이다[2]. 척도 없는 네트워크에서 허브 노드만을 제거 대상으로 하는 목표 공격(Target Attack)과 무작위로 선택된 노드를 제거하는 무작위 공격(Random Attack)에 대하여 각각 다른 대응력을 보여준다. 척도 없는 네트워크는 노드의 무작위 공격에 대하여 램덤 네트워크[5]보다 견고함을 지닌 반면, 목표 공격에는 매우 취약한 형태를 취하고 있다[8]. 척도 없는 네트워크는 차수가 적은 많은 노드와 차수가 매우 많은 허브 노드들이 공존하는 구조를 취하고 있다[2]. 이런 네트워크에서 무작위로 노드를 제거하면 대부분 차수가 적은 노드들이 제거될

것이며, 허브 노드들이 제거될 확률은 매우 낮아지게 된다. 따라서 무작위 노드 제거로 인한 네트워크의 견고성에 미치는 영향은 매우 적게 된다. 반면, 척도 없는 네트워크에서 매 단계마다 차수가 큰 허브들을 제거하면 네트워크에 견고성에 심각한 피해를 가져오게 된다. 결국 세포 네트워크는 척도 없는 네트워크이므로 무작위 공격, 무작위적 노드의 기능 정지, 무작위적 노드의 변형에 매우 견고함을 지니고 있는 반면, 대부분의 노드들을 조정하고, 노드간의 연결 통로 역할을 하는 허브 노드만을 의도적으로 제거하면 그 네트워크는 심각하게 파괴되어 제 기능을 수행하지 못하는 상태가 된다[8].

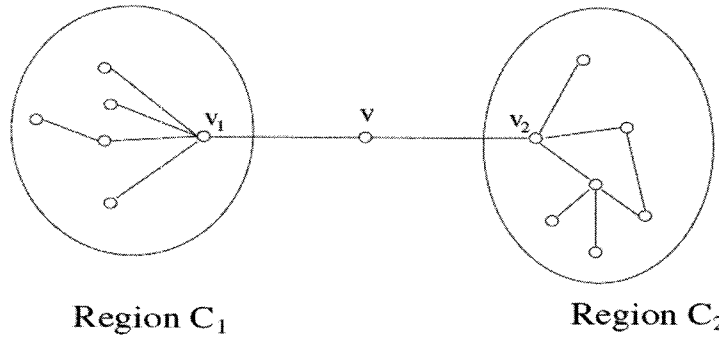
## 3. 연결 노드(Bridging Node)

이 장에서는 기존 연구에서 네트워크의 핵심노드를 결정하는 측정치인 연결성(Connectivity)과 사이 중앙성(Betweenness Centrality)[9]에 대하여 소개하고, 연결성과 사이 중앙성을 조합하여 이 논문에서 제안하는 연결노드를 설명한다.

### 3.1 네트워크의 핵심 노드(Essential Node in the Biological network )

복잡한 네트워크에서 모든 노드가 동등한 역할을 하지는 않는다. 예를 들면, 한 노드의 제거는 네트워크 견고성에 영향을 거의 미치지 않는 반면, 또 다른 노드의 제거는 네트워크에 심각한 영향, 특히 그 노드의 제거로 인하여 일부분의 연결이 끊겨 고립되는 현상을 유발할 수 있다. 공격에 대처하기 위한 네트워크 견고성 및 전염병에 대응하기 위한 면역시스템 고안 등에서 네트워크의 핵심노드에 대한 문제는 매우 중요한 주제이다[8].

노드의 중요성을 측정하기 위한 측정치는 중앙성(Centrality)과 연결성(Connectivity)으로 구별된다[9]. 중앙성은 전체 네트워크에서 그 노드의 위치를 기반으로 노드의 중요성을 측정한다. 연결성은 한 노드가 주변 이웃 노드와 얼마만큼 연결되어 있는가에 대한 정량화된 단위이다. 따라서 연결성은 지역적 중요성(Local Importance Quantity)을



[그림 1] 노드의 연결성과 중앙성

측정하며, 중앙성은 광역적 중요성(Global Importance Quantity)을 측정한다[9].

[그림 1]에서 노드 v1과 v2는 연결성이 가장 큰 노드들이다. 이 노드들은 다른 노드에 비해 이웃 노드들과 많은 연결 링크를 가지고 있기 때문이다(  $k(v1)=5$ ,  $k(v2)=4$ ). 반면, [그림 1]을 전체 네트워크로 볼 때 위치적으로 가장 중앙에 위치하고, 전체 노드들이 상호 연결을 위해 필수적으로 거쳐야 하는 노드 v는 중앙성이 가장 큰 노드이다. 노드 v의 연결성 또는 지역적 중요성은 이웃 노드와 오직 2개의 연결 링크를 가지므로 매우 낮다. 그러나 노드 v의 광역적 중요성은 매우 크다. 전체 네트워크에서 region c1과 region c2가 상호 작용을 하기위해 필수적으로 거쳐야하는 노드이기 때문이다. 따라서 노드 v는 연결선이 적은 노드지만 전체 네트워크의 견고성에 중요한 역할을 하고 있다. 즉, 이 노드의 제거는 전체 네트워크가 분화되는 심각한 영향을 끼치게 되기 때문이다.

두 개의 부분 네트워크를 연결하고 있는 중앙지점에 위치한 노드의 중요성을 측정하는 측정치로 사이 중앙성(Betweenness Centrality, BC)이 있다 [9]. 사이 중앙성은 네트워크에 존재하는 모든 노드 쌍의 최단경로와 그 최단경로가 특정 노드를 경유하는 비율을 측정한다. 아래 식 (4)는 노드 v에 대한 사이 중앙성 BC(v)를 계산하는 공식이다 [9].

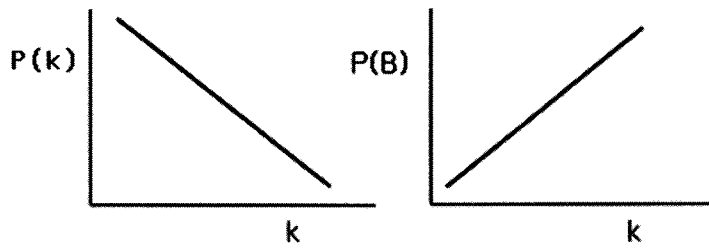
$$BC(v) = \sum_{s \neq v \neq t, s, t \in V} \frac{p_{st}(v)}{p_{st}} \quad (\text{수식4})$$

위 공식에서  $p_{st}$ 는 노드 s와 t간의 최단경로 수를 나타내며  $p_{st}(v)$ 는 노드 v를 경유하는 s와 t의

최단경로 개수이다.  $BC(v)$ 의 값이 크면 클수록 모든 노드 쌍들의 최단경로가 노드 v를 지나는 횟수가 많다. 즉, 노드 v는 노드 쌍의 최단경로 사이에 위치하게 됨을 의미한다. 또 다른 의미로서 노드 v를 통한 정보의 흐름이 많음을 의미한다. [그림 1]에서 노드 v는 두 개의 이웃 노드만을 가지고 있지만 region c1과 region c2에 위치한 모든 노드 쌍들의 최단경로는 노드 v를 거쳐야만 하므로 사이 중앙성이 매우 높다. 이 경우 노드 v가 제거되면 네트워크가 분할되는 결과를 가져오게 되어 전체 네트워크의 유기적인 기능을 상실하게 된다.

네트워크의 척도 없음(Scale-Free)은 무작위 공격에 대하여 네트워크에 견고성을 보장하는 중요한 특징이다[8]. 척도 없는 네트워크는 노드의 차수 분포  $P(k)$ 가 멱함수를 따른다. 멱함수 특징에 의하여 차수가 적은 많은 노드들과 차수가 매우 많은 허브 노드들이 존재하게 되며, 이 허브노드들이 네트워크를 조정 및 중재하는 중요한 노드로 판단된다. 따라서 허브 노드만을 목표로 하는 공격에 척도 없는 네트워크는 매우 취약하다. 이러한 측면에서 노드의 중요성은 노드의 연결성에 의해 결정된다고 볼 수 있다.

[그림 2]의 왼쪽 그래프는 차수 k와 차수분포  $P(k)$ 의 관계를 로그 스케일을 적용하여 일반화한 것이며, 오른쪽 그래프는 차수 k와 사이 중앙성 분포  $P(B)$ 의 관계를 표현한 것이다. [그림 2]의  $P(k)$ 분포에서 볼 수 있듯이 링크 수(k)가 많은 허브노드는 일반 노드보다 적지만 전체 네트워크에서 또 다른 작은 네트워크를 만드는 매우 중요한 노드이다. 또한 전체 네트워크에서 위치 중요성을



[그림 2] 척도 없는 네트워크의 차수분포  $P(k)$ 와 사이 중앙성 분포  $P(B)$

측정하는  $P(B)$ 분포에서 허브 노드는 매우 큰 사이 중앙성 값을 가지므로 결론적으로 허브노드는 지역적으로 광역적으로 모두 중요한 노드임을 알 수 있다.

### 3.2 연결 노드(Bridging Node)

이 절에서는 복잡계 네트워크의 견고성에 큰 영향을 미치는 핵심 노드로 연결노드를 제안하고 그 중요성을 검증한다.

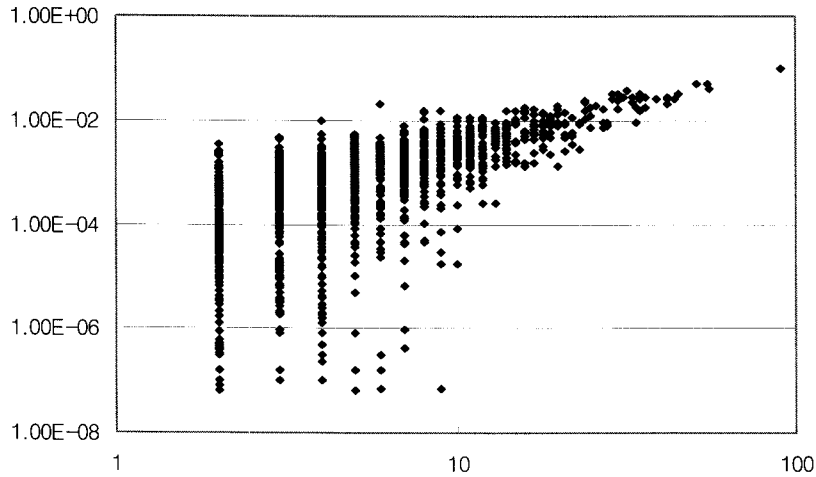
네트워크에서 모듈이나 클러스터는 고밀도로 연결된 노드들의 집합이다. 클러스터에 대한 연구는 사회 네트워크에 대한 연구에서 시작되었다. 사회 네트워크에는 친밀성, 안전성 등 여러 이유로 밀접한 클러스터를 만드는 경향이 있다. 한 클러스터에 존재하는 노드들은 서로 밀접하게 연관 관계를 가지고 있고, 외부로는 몇 개 안 되는 링크가 있어서 이 링크를 통해서 다른 클러스터와 커뮤니케이션을 할 수 있다[10]. 여기서 클러스터 내부에 존재하는 밀접한 링크는 ‘강한 유대관계(Strong Link)’이며, 클러스터와 클러스터를 연결하는 링크는 ‘약한 유대관계(Weak Link)’이다. 이런 네트워크에서 강한 연결을 제거할 때와 약한 연결을 제거할 때 네트워크에 미치는 영향을 예측할 수 있다. 즉, 강한 연결을 제거한다 할지라도 네트워크 노드간의 연결 경로에는 거의 영향을 미치지 않는다. 이 경우 밀집도 있게 연결되어 있는 클러스터 내에서의 노드가 제거된다 하더라도 다른 노드간의 상호연결에 영향을 주지 못하기 때문이다. 그러나 클러스터와 클러스터를 연결하고 있는 약한 고리를 제거하면 노드의 경로 길이는 매우 길어지게 되며

더욱 심각하면 클러스터들의 분리 현상까지 영향을 미칠 수 있다. 사회 네트워크에서 약한 연결은 강한 연결보다 훨씬 중요한 경우가 있다. 약한 연결이 다수의 사회 네트워크들을 하나로 이어 붙이는 결정적인 접착제 역할을 하기 때문이다[10]. 네트워크에서 이러한 “약한 연결”의 정도를 이 논문에서는 연결노드(Bridging Node)로 측정하게 된다.

클러스터링 정도에 대한 정량화된 값인 클러스터링 계수는 차수( $k$ )가 커지면 작아진다. 즉, 노드의 차수와 역관계를 형성하고 있는데 이는 네트워크가 모듈을 형성하며 이 모듈들이 계층 구조를 이루고 있다는 의미이다. 단, 한 노드가 연결노드가 되기 위해서 클러스터링 계수의 값이 작아야 한다. 클러스터링 계수의 값이 크면 이 노드와 연결된 이웃 노드들이 서로 밀접하게 연결되어 있다는 의미이므로 그 노드를 제거해도 이웃 노드들의 상호연결에는 큰 영향을 미치지 못하기 때문이다.

사이 중앙성 값은 차수( $k$ )가 커지면 그에 따라 커지게 된다. 허브 노드의 사이 중앙성 값이 큰 이유는 허브 노드를 중심으로 직접적으로 연결된 수많은 이웃 노드들의 최단 경로는 이 허브노드를 통해서 연결되기 때문이다. 따라서 허브 노드는 사이 중앙성 값은 매우 크지만 클러스터를 구성하는 중심 역할을 하는 노드이므로 연결노드는 아니다.

결론적으로 노드들이 클러스터를 형성하고 클러스터의 고립을 막기 위해 클러스터간 연결고리 역할을 해주는 연결노드는 사이 중앙성 값이 커야만 한다. 또한 사이 중앙성 값이 큰 허브노드와 다르게 클러스터를 구성하는 역할이 아니라 클러스터간의 연결고리 역할을 수행하므로 링크 수가 많지 않으며 클러스터링 계수의 값은 작아야 한다.



[그림 3] DIP yeast\_core data의 사이 중앙성 분포  $BC(k)$

[그림 3]은 미국 UCLA에서 운영중인 DIP (Database of Interacting Proteins)[11] 데이터베이스의 yeast\_core 데이터를 대상으로 차수에 따른 사이 중앙성 분포인  $BC(k)$ 를 보여주고 있다. [그림 3]에서 x축은 노드의 차수이며 y축은 로그 스케일을 적용한 사이 중앙성 값을 표시한 것이다.

[그림 3]에서 허브 노드는 많은 이웃 노드와 직접적으로 연결되어 있고, 이웃 노드간 최단 경로는 허브 노드를 거쳐야 하므로 허브 노드의 사이 중앙성 값은 커진다. 차수가 증가하면 사이 중앙성 값도 증가함을 알 수 있지만,  $BC(k)$ 의 범위는  $k$ 가 작을수록 넓게 분포함을 보여주고 있다. 즉, 허브 노드들은 대부분 비슷한 사이 중앙성 값을 가지지만 링크 수가 적은 노드일수록  $BC(k)$  값의 범위는 매우 넓게 분포한다. 범위 분포를 세로로 보았을 때 가장 위쪽에 위치한 노드들은 링크수가 매우 적지만 매우 큰  $BC$  값(HBLC: High Betweenness and Low Connectivity)을 가지고 있다[12]. HBLC 노드들은 컴퓨터에 의해 인위적으로 생성된 척도 없는 네트워크에서는 발견되지 않는 생물 네트워크의 독특한 특성이다[12]. 이 HBLC 노드들은 링크 수는 아주 적지만 사이 중앙성 값은 허브 노드만큼의 큰 값을 가진다.

이러한 연결노드의 특성을 반영하여 노드  $v$ 에 대한 연결노드인자(Bridging Node Factor)  $BNF(v)$ 를 식(5)과 같이 정의할 수 있다. 제안된

모델에서 연결노드인자는 노드  $v$ 가 연결 노드가 될 수 있는 측정값으로서 이 값이 크면 연결 노드가 될 가능성이 높음을 의미한다.

식 (5)에서 연결노드는 동일 차수의 노드,  $N(v)$ 들에 비하여 상대적으로 사이 중앙성  $BC(v)$  값이 매우 큰 노드이다. 즉, 두 클러스터 간의 상호작용을 위해 필수적으로 거쳐야 하는 노드이기 때문이다. 또한, 연결노드는 클러스터링 계수 값이 작은 노드이다( $\frac{1}{C(v)}$ ). 만약, 클러스터링 계수 값이 큰 노드라면 이 노드가 제거 되더라도 이웃 노드들이 서로 밀접하게 연결되어 있어서 이웃 노드간의 상호작용에 영향을 미치지 못하기 때문이다. 단, 클러스터링 계수가 0인 노드인 경우 그 노드가 제거되면 이웃 노드들이 상호작용을 할 수 없으므로 연결노드가 될 가능성이 가장 높게 되며, 식에서 이 점을 반영하여  $C(v)$ 값을 0.1로 설정하였다. 이 값은 실제 네트워크 데이터에서  $C(v)$ 가 0이면 다른 노드에 비하여 연결노드가 될 가능성이 10배 정도 큼을 의미하는 경험적인 계수 값이다.

$$BNF(v) = \frac{1}{C(v)} \times \frac{BC(v)}{N(v)} \quad (\text{수식5})$$

여기서,  $C(v) = 0$  이면 0.1로 지정함.

단,  $C(v)$  : 노드  $v$ 의 클러스터링 계수값

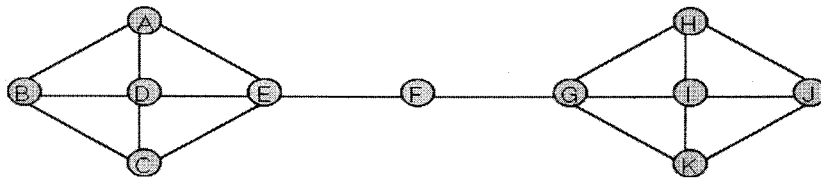
$N(v)$  : 노드  $v$ 의 차수

$BC(v)$  : 노드  $v$ 의 사이 중앙성 값

연결노드인자  $BNF(v)$ 는 노드  $v$ 가 연결 노드가 되기 위한 측정치이다. 이 정의에 의하여 노드  $v$ 가 연결노드가 되기 위해서 클러스터링 계수의 값  $C(v)$ 는 작아야 하며, 노드의 차수  $N(v)$ 에 대하여 상대적으로 사이 중앙성  $BC(v)$  값이 커야 한다.

[ Example 1 ]

[그림 4]의 그래프는 연결 노드를 측정하기 위한 예제 네트워크이다. 이 그래프는 두 개의 서브 네트워크가 노드 F에 의해 연결된 형태이다.



[그림 4] 예제 네트워크

[표 1]은 예제 네트워크의 각 노드에 대한 차수  $N(v)$ , 사이 중앙성 값  $BC(v)$ , 클러스터링 계수 값  $C(v)$ 과 연결 노드 인자 값( $BNF$ )을 보여주고 있다. 노드 F가 가장 큰 연결노드인자 값을 가짐을 알 수 있다. 위 예제 네트워크에서 노드 G, E, I, D는 각 링크가 4인 허브 노드이다. 단 노드 D와 I는 클러스터링 계수 값이 노드 G와 E보다 큼을 알 수 있다. 노드 D와 I는 이웃 노드들을 연결하여 모듈을 구성하는 허브노드이기 때문이다. 반면 노드 G, E는 사이 중앙성 값은 크고 클러스터링 계수 값이 상대적으로 작으므로 모듈의 일부분임을 알 수 있다. 노드 F는 네트워크에서 링크수가 가장 작지만 사이 중앙성 값이 매우 크다. 따라서 노드 F가 모듈과 모듈을 연결하는 연결노드가 될

[표 1] 그림 4 네트워크에서 각 노드의 측정치

Node	$N(v)$	$BC(v)$	$C(v)$	$BNF(v)$
F	2	0.55	0.1	2.75
G	4	0.54	0.33	0.409091
E	4	0.54	0.33	0.409091
I	4	0.05	0.66	0.018939
D	4	0.05	0.66	0.018939
K	3	0.05	0.66	0.025253
H	3	0.05	0.66	0.025253
C	3	0.05	0.66	0.025253
A	3	0.05	0.66	0.025253
B	3	0.007	0.66	0.003535
J	3	0.007	0.66	0.003535

[표 2] 노드 제거에 따른 네트워크의 직경 및 평균최단경로의 변화

	노드 수	직경	평균 최단경로
전체 네트워크	11	6.0	2.38
시나리오 1 F 노드 제거	10	2.0	0.48
시나리오 2 D or I 노드 제거	10	6.0	2.36



수 있음을 알 수 있다.

[표 2]는 예제 네트워크에서 노드의 연결 노드 인자 값(BNF)이 가장 큰 노드 F를 제거하는 경우와 허브 노드로서 차수가 가장 크고 지역적 모듈을 구성하는 역할을 하는 D나 I 노드를 제거할 때 전체 네트워크에 미치는 영향을 표로 나타낸 것이다. [표 2]에서 보는 것과 같이 노드 제거 전 예제 네트워크의 전체 노드 개수는 11개이며, 지름은 6이고, 평균 최단경로는 2.38인 네트워크이다.

- 시나리오 1: 노드 F를 제거했을 때 직경이 2.0으로 작아지고 평균 최단경로 또한 0.48로 매우 작아지는 결과를 보였다. 이 결과는 노드 F의 제거로 전체 네트워크가 2개의 독립적인 모듈로 분리되어 각각의 두 모듈의 직경은 2이며 평균 최단경로가 0.48임을 보여준다.
- 시나리오 2: 노드 D나 I를 제거할 경우 전체 네트워크의 견고성에 큰 영향을 미치지 못하였다. 단, 평균 최단경로가 약간 줄어든 이유는 허브 노드 1개가 제거되었기 때문이지만 실제 현실세계의 복잡한 네트워크에서는 허브노드를 차례로 제거하면 평균 최단 경로 값은 네트워크의 분화가 심하게 일어나기 전까지 계속 증가한다. 위 단순한 실험에서 연결노드로 선정된 노드들

은 허브노드와 함께 네트워크의 견고성에 아주 큰 영향을 미침을 알 수 있다. 특히, 연결노드를 차례대로 제거하면 특이하게 평균 최단경로가 줄어드는 특징을 보임으로서 연결노드의 제거로 인하여 곧바로 전체 네트워크가 독립된 모듈로 분화되는 현상을 확인할 수 있다. 네트워크, 특히 생물 네트워크는 그 기능을 온전히 수행하기 위해서 모든 모듈들이 유기적으로 연결되어 전체적인 시스템을 유지하게 되는데, 이런 네트워크에서 모듈들의 분화는 네트워크의 심각한 기능 상실을 의미한다[6].

## 4. 실험

본 장에서는 제안된 연결노드 모델을 입증하기 위하여 미국 UCLA에서 운영하는 공개 생명공학 데이터베이스[11]를 이용하여 네트워크의 견고성에 관한 분석실험을 수행한다.

### 4.1 실험 환경

실험에 사용된 네트워크 데이터베이스는 미국 UCLA DIP (<http://dip.doe-mbi.ucla.edu>: Database

[표 3] yeast PPI 코어 데이터 [<http://dip.doe-mbi.ucla.edu>]

DIP45E	DIP512N S25365	Tup1p 283218	P16649	JN0133	83454	DIP696N	Cyc8p	P14922
DIP51E	DIP13N A40083	Pho2p 101447	P07269	A25872	82888	DIP142N	Bas1p	P22035
DIP52E	DIP760N S62015	Los1p 2133145	P33418	S38043	539302	DIP962N	Sol1p	P50278
DIP68E	DIP665N S25365	Mig1p 283218	P27705	S17246	101546	DIP696N	Cyc8p	P14922
DIP100E	DIP376N 72877	Sup35p P05453	EFBYS2	72877		DIP376N	Sup35p	P05453
DIP103E	DIP376N 626763	Sup35p P05453	EFBYS2	72877		DIP800N	Sup45p	P12385
DIP123E	DIP41N XUBYSO	Lpd1p 2144399	P09624	A30151	82983	DIP1102N	Kgd2p	P19262
DIP124E	DIP365N XUBYSO	Kgd1p 2144399	P20967	DEBY	1070439	DIP1102N	Kgd2p	P19262
DIP125E	DIP671N YCBYI	Ilv6p 68242	P25605	S19411	83137	DIP1104N	Ilv2p	P07342
DIP126E	DIP744N S48761	Tps1p 630358	Q00764	S34979	6319602	DIP823N	Tps2p	P31688
DIP127E	DIP744N S54473	Tps1p 1084868	Q00764	S34979	6319602	DIP891N	Tps3p	P38426
DIP128E	DIP744N S36868	Tps1p 6323537	Q00764	S34979	6319602	DIP753N	Tsl1p	P38427
DIP129E	DIP823N S54473	Tps2p 1084868	P31688	S48761	630358	DIP891N	Tps3p	P38426
DIP130E	DIP753N S48761	Tsl1p 630358	P38427	S36868	6323537	DIP823N	Tps2p	P31688

of Interacting Proteins)의 코어 데이터를 사용하였다. 이 데이터베이스에는 yeast PPI 네트워크 (Protein-Protein Interaction Network) 데이터들이 축적되어 인터넷상에 공개되고 있다(표 3 참고). [표 3]의 첫 번째 라인에서 DIP:512N은 512번 단백질 즉 노드(N:Node)를 나타내며, 이 노드는 DIP:696N 노드와 상호연결 관계가 있음을 표현하고 있다. DIP:45E는 두 노드간의 링크(E:Edge)를 표현한다. 데이터베이스에는 5952개의 단백질 노드간의 상호관계를 그래프 형태로 표시하고 있다.

PPI는 상호작용하는 단백질들로 구성된 복잡계 네트워크이다. 이 네트워크에서 단백질은 노드로, 단백질들 사이의 상호 관계는 링크로 표현된다. 하나의 네트워크는 매우 방대한 단백질들과 그들 사이의 복잡한 관계들로 구성되어 있다. 이런 네트워크를 단백질 상호 네트워크라고 부른다. 단백질 상호작용 네트워크는 인터넷이나 다른 네트워크가 가지는 특징인 척도 없는 네트워크 구조와 계층적 구조를 가지고 있다.

#### 4.2 실험 방법과 결과

이 연구에서는 네트워크에서 허브노드 뿐만 아니라 모듈간의 연결고리 역할을 하는 연결노드를 제안하고 네트워크 견고성에 중요한 역할을 담당함을 밝혀내는 것이다. 제안된 연결노드는 차수가 작으면서 상대적으로 사이 중앙성 값이 크며 또한 클러스터링 계수 값은 상대적으로 작은 노드들이

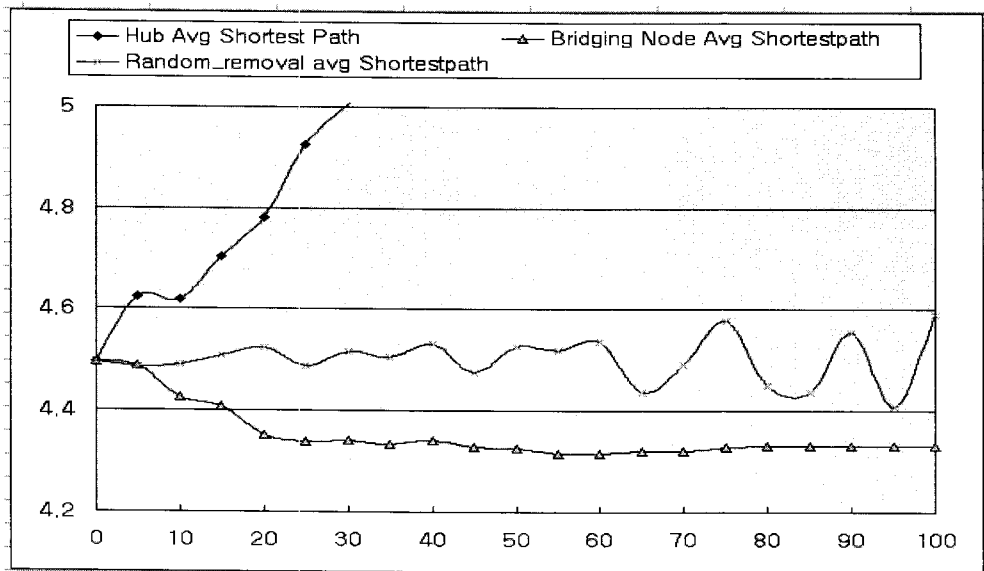
다.

실험에서는 DIP 데이터베이스의 yeast\_core 단백질 대상으로 연결노드를 밝혀내고 그 연결노드의 중요성을 실험하였다. 실험 방법은 허브노드와 연결노드를 추출하여 차례대로 가장 큰 노드를 일정 수(5개씩)의 노드를 제거하면서 네트워크의 평균 경로길이와 클러스터링 개수를 측정하였다. 여기서 허브 노드는 링크가 가장 많은 노드 순으로 추출하였으며 연결노드는 이 논문에서 제시한 연결노드인자 값을 계산하여 인자 값이 가장 큰 순으로 추출하였다. 무작위 노드는 매 순간 무작위로 추출된 노드를 제거하였다. 아래 [표 4]는 실험 결과를 보여주고 있다. [표 4]에서와 같이 노드 삭제 전 전체 네트워크의 평균 최단 경로는 4.495이며, 각 노드를 5개씩 제거 후의 평균 최단 경로를 계산하였다. [그림 5]는 이 [표 4]실험 결과를 그래프로 표현한 것이다.

[그림 5]는 [표 4]의 실험 결과를 허브노드, 연결노드 그리고 무작위로 선택된 노드들을 제거하면서 네트워크의 평균 최단경로를 그래프로 표시한 것이다. 여기서 x축은 제거된 노드 수를 의미하며 y축은 노드가 제거될 때의 평균 최단 경로를 표시한 것이다. [그림 5]에서 가장 위쪽 라인은 허브노드를 제거할 때의 평균 최단 경로 값을 표현한 것이다. [그림 5]에서 보는 것과 같이 허브노드를 제거하면 급격하게 경로 값이 증가함을 보여주고 있다. 가운데 라인은 무작위 제거로 인한 변화이다. [그림 5]에서 가장 아래 라인인 논문에서 제시

[표 4] 실험 결과

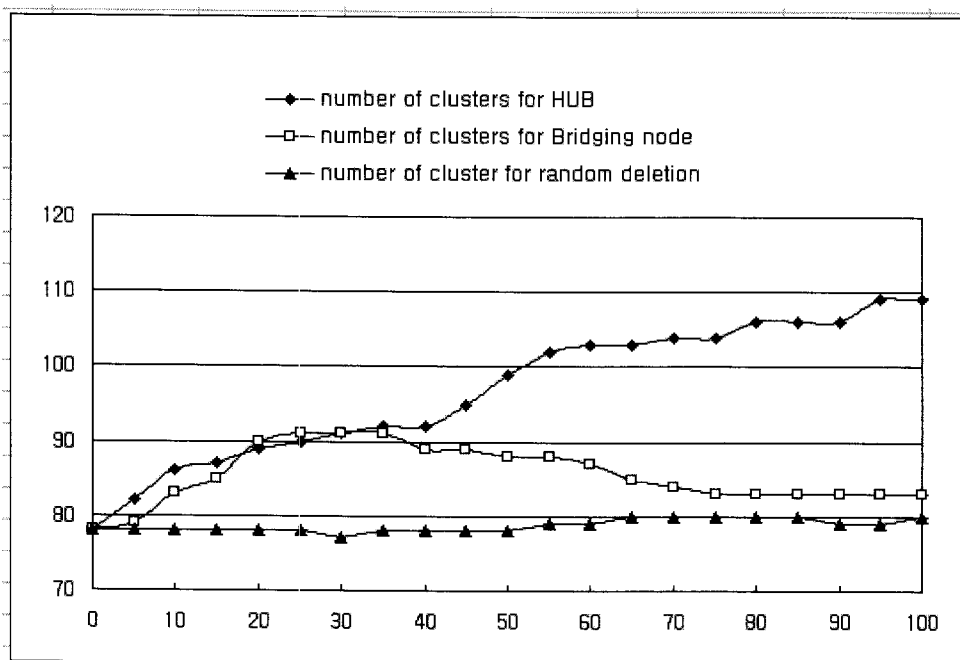
number of node deletion	Hub Avg Shortest Path	Bridging Node Avg Shortest Path	Random_removal Avg Shortest Path
0	4.495	4.495	4.495
5	4.62350753	4.48733514	4.48523288
10	4.61882454	4.42746648	4.49027369
15	4.70416671	4.40865794	4.50818824
20	4.78224901	4.35249275	4.52373314
25	4.92580228	4.33835386	4.48880282
30	5.00419916	4.34064305	4.51638469
35	5.09037024	4.33340494	4.50591367
40	5.16906092	4.34134409	4.53120144
...	....	...	...
440	0.07484792	4.29615948	4.39926175



[그림 5] 노드제거에 따른 평균 최단경로 변화

한 연결 노드를 제거할 때의 평균 최단경로의 변화 값을 보여주고 있다. [그림 5]에서 허브 노드를 제거해 나가면 네트워크가 분화가 심하게 발생하기 전까지 평균 최단 경로길이는 계속 늘어나게 됨을 보여주고 있다. 네트워크는 정상적인 기능을 수행하기 위해 노드들이 유기적으로 밀접하게 연

결되어 있어야 하다. 그러나 허브노드들이 제거되면 그 영향으로 노드간의 상호 연결 경로가 급격히 증가하여 네트워크의 기능 수행에 심각한 영향을 미치게 된다. 반면 무작위로 노드를 선택하여 제거하면 네트워크는 거의 영향을 받지 않고 있다. 제거되는 노드들이 대부분 연결선이 적은 일반노



[그림 6] 노드제거에 따른 클러스터 개수의 변화

드이기 때문이다. 그러나 이 연구에서 제안하는 연결노드를 제거하면 네트워크의 경로 길이가 일정 수준까지 계속 줄어드는 특징을 보이고 있다. 그 이유는 시나리오1 예에서 알 수 있듯이 중요한 연결경로가 제거되면 곧바로 네트워크가 여러 클러스터로 분화되기 때문이다. 실험에서 보는 것과 같이 연결노드 제거로 인하여 네트워크의 분화가 계속 진행되다가 25~30개의 연결노드 제거 후에는 더 이상 분화가 되지 않고 있다. 이 사실로 네트워크에 치명적인 분화를 일으키는 연결노드는 전체 노드의 0.05%정도이다. 단 이 노드들은 허브노드보다 네트워크 견고성에 더욱 치명적인 손상을 야기할 수 있음을 실험에서 보여주고 있다.

[그림 6]은 각 노드를 제거하면서 발생하는 클러스터 변화를 그래프로 표현한 것이다. [그림 6]에서 가장 위쪽 라인은 허브 노드를 제거할 때의 클러스터 변화량을 보여주고 있으며, 두 번째 라인은 연결 노드를 제거할 때의 클러스터 변화량을 보여준다. 실험에서 특정 노드의 제거로 인하여 네트워크의 클러스터 개수가 증가하는 단편화의 특성을 보이는 노드는 허브 노드와 연결노드이다. 특히 연결노드의 경우 35개 정도의 노드를 제거할 경우 허브노드를 제거할 때보다 단편화가 더욱 심하게 나타남을 볼 수 있다. 즉, 연결노드 제거로 인하여 발생된 클러스터로 인하여 [그림 5]에서와 같이 평균 최단경로가 줄어드는 것이다.

주목해야 할 또 다른 특징은 [그림 6]에서 허브노드와 연결노드의 제거로 인한 클러스터 개수의 증가 추세는 많은 차이가 나지 않는다. 그러나 [그림 5]에서와 같이 허브노드의 경우 평균 최단 경로가 급격히 늘어나고 연결노드의 경우는 급격히 줄어들고 있다. 그 이유는 허브노드와 연결노드 제거로 인해 생기는 클러스터 개수는 비슷하지만 허브노드 제거로 발생하는 클러스터는 대부분 독립적인 노드(singleton)들이기 때문이다. 따라서 연결노드 제거로 인해 최단경로가 급격히 줄어드는 것이다. 또 다른 특징으로는 연결노드는 40개 제거 후에는 오히려 클러스터 수가 줄어드는 특징을 보인다. 약 40개 이후부터 제거되는 클러스터들은 대부분 두 개의 노드와 한 개의 링크로 이루어진 작은 클러스터들이 제거되기 때문에 네트워크에 큰 영향을 미치지 못한다.

이 실험결과로부터 Yeast\_core에 포함된 전체 5,728개의 단백질 중 30~40개 정도는 네트워크의 견고성에 매우 핵심적인 노드이며, 이들이 모듈과 모듈을 연결하는 연결노드 역할을 한다는 것을 알 수 있다.

## 5. 결론

이 논문에서는 생물 네트워크의 생존에 매우 중요한 핵심노드로 연결노드를 제안하고, 실험을 통하여 그 중요성을 입증하였다. 또한, 거대한 생물 네트워크에서 연결노드를 쉽게 판별하여 찾아내는 알고리즘도 제시하였다. 연결노드는 차수가 매우 작지만 사이 중앙성 값이 매우 크며 상대적으로 클러스터링 계수 값이 작은 노드들이다. 네트워크 견고성에서 연결노드의 중요성을 입증하기 위하여 미국 UCLA에서 운영하는 생물 네트워크 데이터베이스에 대하여 견고성 측정 실험을 수행하였다. 실험에서 연결노드를 제거하면 네트워크의 경로 길이가 일정 수준까지 계속 줄어드는 특징을 보였다. 그 이유는 연결노드가 제거되면 곧바로 네트워크가 여러 클러스터로 분화되기 때문이다. 독립된 클러스터로 분리된다면 각 클러스터는 자신만의 고유한 기능을 수행할 수 있을지라도 분화 이전에 클러스터간의 협력에 의해 수행되어진 기능을 할 수 없게 된다. 특히 연결노드의 경우 실험에서 35개 정도의 노드를 제거할 경우 허브노드를 제거할 때보다 단편화가 더욱 심하게 나타남을 볼 수 있었다. 그러나 그 이상의 연결노드 제거는 무작위 제거와 같이 전체 네트워크에 거의 영향을 미치지 못하였다. 이 사실로 실험에 사용된 생물 네트워크에서 치명적인 분화를 일으키는 연결노드는 전체노드의 0.05% 정도로 판단된다. 이 노드들은 허브노드보다 네트워크 견고성 측면에서 더욱 치명적인 손상을 야기할 수 있음을 실험에서 확인할 수 있었다.

향후 연구에서는 연결노드로 인해 분화되는 기능적 모듈에 대한 체계적인 연구가 필요하다. 연결노드를 찾아 체계적으로 연결노드를 제거하면 전체 네트워크를 구성하는 기능적 모듈들을 찾을 수 있다.

## 참 고 문 헌

- [1] S. H. Strogatz, "Exploring complex networks," *Nature*, vol. 410, no. 6825, pp. 268-276, 2001.
- [2] A. L. Barabasi and E. Bonabeau, "Scale-free networks," *Scientific American*, vol. 288, pp. 60-69, 2003.
- [3] Duncan J. Watts and Strogatz H. Steven, "Collective Dynamics of 'Small-World' Networks," *Nature*, vol. 393, pp. 440-442, 1998.
- [4] Stanley Milgram, "The Small-World Problem," *Psychology Today*, vol. 1, pp. 60-67, 1967.
- [5] P. Erdos and A. Reny, "On random graphs," *Publicationes Mathematicae*, vol. 6, pp. 290-297, 1959.
- [6] R. Cohen and S. Havlin, "Scale-free networks are ultra small," *Phys. Rev. Lett*, 90, 058701, 2003.
- [7] E. Ravasz, et al., "Hierarchical organization of modularity in metabolic networks," *Science*, vol. 297, pp. 1551-1555, 2002.
- [8] R. Albert, H. Jeong and A. L. Barabasi, "Error and Attack tolerance of complex networks," *Nature*, vol. 409, pp. 378-382, 2000.
- [9] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, vol. 40, pp. 35-41, 1979.
- [10] M. S. Granovetter, "The Strength of Weak Ties," *American Journal of Sociology*, vol. 78, pp. 1360-1380, 1973.
- [11] DIP: Database of Interacting Proteins, "http://dip.doe-mbi.ucla.edu/".
- [12] J. Maliackal Poulo, et al., "High-Betweenness Proteins in the Yeast Protein Interaction Network," *J. biomed. Biotechnol.*, vol. 2, pp. 96-103, 2005.
- [13] Barabasi A. L., et al., "Network Biology: understanding the cell's functional organization," *Nature Review*, vol 5, Feb, 2004



안 명 상 (Myoung-Sang Ahn)

- 충북대학교 경영정보학과 (경영학사)
- 충북대학교 전자계산학과 (이학석사)
- 충북대학교 경영정보학과 (박사과정 수료)
- 관심분야: Data Warehouse, Data Mining, OLAP, Bioinformatics
- Email: epita55@cbnu.ac.kr



고 정 환 (Jeong-Hwan Ko)

- 경북대학교 통계학과 (이학박사)
- 오클랜드 대학교 방문교수
- 현 안동대학교 교수
- 관심분야: 신뢰성, 품질관리, 통계상담



유 재 수 (Jae-Soo Yoo)

- KAIST 전산학과 (공학박사)
- 현 충북대학교 BIT 연구중심 대학육성사업단 부단장
- 현 충북대학교 전전컴 교수
- 관심분야: Database, BIT, 정보검색, 멀티미디어 분산객체 시스템
- Email: yjs@chungbuk.ac.kr



조 완 섭 (Wan-Sup Cho)

- 한국과학기술원 전산학과 (공학박사)
- 미국 U. of Florida Post. Doc
- 한국전자통신연구원, 연구원 & 포닥
- 현 충북대학교 경영정보학과 부교수
- 관심분야: Data Warehouse & OLAP, CRM, DB, Bioinformatics, Web Service
- Email: wscho@cbnu.ac.kr