

척도 없는 네트워크를 위한 그래프 레이아웃 알고리즘

(A Graph Layout Algorithm for Scale-free Network)

조 오 만 [†] 강 태 원 ^{**}
(Yong-Man Cho) (Tae-Won Kang)

요 약 네트워크는 공학이나 자연과학은 물론이고 사회과학의 여러 분야를 연구하는데 중요하게 사용되는 모델이다. 이런 네트워크를 좀 더 쉽게 분석하기 위해서는 시각적으로 네트워크의 특징을 잘 나타내는 것이 필요하다. 이러한 그래프 레이아웃 연구는 컴퓨터 기술이 발달함에 따라 많이 연구되고 있다. 그 중에서 요즘 새롭게 부각되고 있는 척도 없는(Scale-free) 네트워크는 다양한 분야에서 복잡한 현상들을 분석하고 이해하는데 유용하게 쓰이고 있다. 이 네트워크의 특징은 링크의 수(Degree)가 멱함수(power law) 분포를 보이고, 다수의 링크를 가지는 허브가 존재함이 알려졌다. 따라서 척도 없는 네트워크에서는 허브를 시각적으로 잘 표현하는 것이 중요하지만 기존의 그래프 레이아웃 알고리즘은 클러스터를 잘 표현하는 정도이다. 그래서 본 논문에서는 척도 없는 네트워크를 잘 표현하는 그래프 레이아웃 알고리즘을 제안한다. 본 논문에서 제안한 알고리즘에서 허브들 간에 작용하는 허브성 척력이 거리에 반비례하고, 허브들의 degree가 a 배 증가하면, 허브사이에 작용하는 척력의 크기는 a^γ (γ 는 연결선 지수)배가 된다. 또한, 전체 노드수와 전체 링크수에 따라 적용되는 힘의 크기를 조정하는 계수를 두어서 네트워크의 규모에 관계없이 허브성 척력이 적용되는 특성이 있다. 제안한 알고리즘이 허브를 잘 표현하는 그래프 레이아웃 알고리즘인지를 기존의 방식과 실험을 통해서 비교하였다. 실험의 절차는 먼저 네트워크에 허브가 존재하는 지를 식별한다. 허브의 존재를 식별하기 위한 방법은 연결선 지수를 확인하고, 연결선 지수의 값이 2와 3 사이에 있으면 허브가 존재하는 척도 없는 네트워크로 판단한다. 다음은 이 네트워크의 레이아웃 작성에 제안한 알고리즘을 사용한다. 그 결과, 제안한 그래프 레이아웃 알고리즘이 기존의 Noack 등의 클러스터 중심의 알고리즘에 비해서 척도 없는 네트워크의 허브를 확실히 잘 보여주고 있음을 확인할 수 있었다.

키워드 : 그래프 레이아웃, 척도 없는 네트워크, 허브, 멱함수, 허브성 척력

Abstract A network is an important model widely used in natural and social science as well as engineering. To analyze these networks easily it is necessary that we should layout the features of networks visually. These Graph-Layout researches have been performed recently according to the development of the computer technology. Among them, the Scale-free Network that stands out in these days is widely used in analyzing and understanding the complicated situations in various fields. The Scale-free Network is featured in two points. The first, the number of link(Degree) shows the Power-function distribution. The second, the network has the hub that has multiple links. Consequently, it is important for us to represent the hub visually in Scale-free Network but the existing Graph-layout algorithms only represent clusters for the present. Therefor in this thesis we suggest Graph-layout algorithm that effectively presents the Scale-free network. The Hubity(hub+ity) repulsive force between hubs in suggested algorithm in this thesis is in inverse proportion to the distance, and if the degree of hubs increases in a times the Hubity repulsive force between hubs is a^γ times (γ is a connection line index). Also, if the algorithm has the counter that controls the force in proportion to the total node number and the total link number, The Hubity repulsive force is independent of the scale of a network. The proposed algorithm is compared with Graph-layout algorithm through an experiment. The experimental process is as follows: First of all, make out the hub that exists in the network or not. Check out the connection line index to recognize the existence

· 이 논문은 2005년도 강릉대학교 학술연구조성비 지원에 의하여 수행되었음

† 정 회 원 : 강릉대학교 컴퓨터공학과

ymcho@kangnung.ac.kr

** 정 회 원 : 강릉대학교 컴퓨터공학과 교수

twkang@kangnung.ac.kr

논문접수 : 2006년 9월 18일

심사완료 : 2007년 3월 16일

of hub, and then if the value of connection line index is between 2 and 3, then conclude the Scale-free network that has a hub. And then use the suggested algorithm. In result, We validated that the proposed Graph-layout algorithm showed the Scale-free network more effectively than the existing cluster-centered algorithms[Noack, etc.].

Key words : Graph Layout, Scale-free Network, Hub, Power Law, Hubity repulsive force

1. 서론

다양한 분야의 과학자들이 자연계와 인간 사회에서 일어나는 다양한 현상들이 어떤 공통점을 가지고 있는지를 연구하는 “복잡성의 과학”에 흥미를 가지고 연구하고 있다. 복잡성의 과학은 환원주의에 바탕을 둔 현대 과학의 한계점을 전일주의 입장에서 이해해 보려는 새로운 시도이다[1]. 한편 네트워크 연구는 복잡성의 과학을 연구하는 강력한 수단으로 새롭게 부각되면서 전 세계적으로 연구가 활발하다[2-5]. 우리 주위에서 흔히 볼 수 있는 사회 및 자연의 시스템은 다수 요소들의 네트워크로 구성되어 있으며, 그러한 시스템의 동작은 그 요소 자체가 아니라 그것의 구성이 중요하게 작용한다는 것이 알려졌다[6].

네트워크를 분석하기 위해서는 연구자가 복잡한 연결 관계를 쉽게 추적하고, 중요한 특징을 편리하게 시각화할 수 있어야 한다. 복잡한 네트워크를 정돈하여 나타내고 조작하기 위한 기반 기술은 자동 그래프 레이아웃 알고리즘(automatic graph layout algorithm)[7]이다. 그래프 레이아웃은 수학과 컴퓨터과학의 많은 분야에 사용되는 것으로 최근에는 생명공학이나, 유전자공학 같은 분야의 시각화에 많이 사용되고 있다[8].

네트워크를 정교하게 시각화하는 것은 네트워크의 중요한 특징을 빠르고 효율적으로 분석할 수 있도록 도와준다. 그래프 레이아웃 알고리즘은 차트, 계층구조, 흐름도 등을 표시하려는 목적에서 출발한 것으로 소프트웨어나 프로세스의 흐름, 컴파일러 체제 등을 나타내는 분야에 응용되었다[9,10]. 이후 그래픽스 기술이 발달하면서 그 사용범위가 크게 확대되어 복잡한 연결 관계를 시각화 하는 수단으로 크게 발전하여 이제는 고급 3차원 시스템으로 발전하고 있다[11,12]. 대표적인 그래프 레이아웃 알고리즘으로는 force-directed method[13], hierarchical method[14]가 있으며, 각각의 경우에도 여러 가지 변형이 있다[15]. 최근에는 유전 알고리즘을 적용하여 주어진 요구를 반영하는 - 예를 들어 링크의 교차 수를 최소화 - 그래프 레이아웃 방법이 개발되었다[16].

요즘 통계물리학이나 인공지능 분야에서 연구되기 시작하는 좁은 세상 네트워크(Small-world Network) [17-19]와 척도 없는 네트워크(Scale-free Network) [18-20]는 서로 다른 특징을 가지고 있다. 이러한 두 네

트워크의 차이로 인하여 네트워크를 표현하는 레이아웃 방법이 달라져야 한다. 기존의 레이아웃 방법들이 좁은 세상 네트워크의 중요한 특징인 클러스터를 잘 표현하는 알고리즘 중심으로 연구[21,22]되었다면, 본 논문에서는 척도 없는 네트워크에서 중요한 구성요소로 다루는 허브를 잘 표현하는 그래프 레이아웃 알고리즘에 대한 연구를 하였다. 본 논문에서 제안한 알고리즘의 특성은 허브성 척력이 거리에 반비례하고, 각 노드의 degree가 각각 a 배 증가하면 허브사이에 작용하는 척력의 크기는 a^γ (γ 는 연결선 지수)배가 된다. 또한, 전체 노드 수와 전체 링크 수에 따라 적용되는 힘의 크기를 조정하는 계수를 두어서 네트워크의 규모에 관계없이 적용되는 특성이 있다.

다음 장들의 내용은 다음과 같다. 2장에서는 본 논문에서 기술하는 내용과 관련 있는 배경지식, 즉 좁은 세상 네트워크와 척도 없는 네트워크 및 그래프 레이아웃 알고리즘에 대해서 알아본다. 3장에서는 본 논문에서 제안한 그래프 레이아웃 알고리즘에 대한 설명이며, 4장은 제안한 알고리즘과 기존의 알고리즘들 간의 비교 실험 및 실험 결과를 분석한 내용이다. 마지막으로 5장에서는 본 논문의 결론부분으로서 이 연구의 성과 및 추후 연구방향을 제시한다.

2. 배경이론

이 논문에서는 좁은 세상 네트워크와 척도 없는 네트워크의 시각화를 네트워크의 구성요소, 즉 허브의 존재 유무에 따라 다른 그래프 레이아웃 알고리즘을 적용한다. 따라서 결집계수(Clustering coefficient)[3]가 높은 좁은 세상 네트워크와 허브가 존재하는 척도 없는 네트워크에 대해서 각 네트워크의 특징들을 살펴보고, 네트워크를 시각화 하는 기존의 그래프 레이아웃 알고리즘들에 대해서도 알아본다.

2.1 좁은 세상 네트워크(Small-World Network)

좁은 세상이라는 개념은 네트워크에 내재된 원리를 다루는 매우 중요한 것이다. 좁은 세상은 인터넷 link, 산불의 번짐, 전화 시스템, 먹이사슬, 기업의 지배구조, 부의 분배, 확산 및 지진 발생, 세포내 분자 네트워크, 두뇌 신경세포의 연결, 할리우드 배우 관계, 과학자들의 협동 네트워크, AIDS의 확산, 발생과정의 이해, 유전자

발현 등 자연에 보편적으로 내재된 것으로 알려져 있다 [24]. 좁은 세상 네트워크는 1998년 Watts와 Strogatz에 의해서 무작위 그래프(random graphs)의 한 종류로 정의 되었다. 그들은 그래프가 결집계수(clustering coefficient)와 평균 최단 거리(mean-shortest path length)에 의해서 구별될 수 있다고 말한다. 결집계수란 친구의 친구들 간에 얼마나 서로 친구인가를 나타내는 계수이다. 좁은 세상 네트워크는 같은 수의 노드와 링크를 가지는 무작위 그래프(random graph)와 비교해서 결집계수가 상당히 높고 평균 최단 거리가 생각보다 매우 낮은 특징을 가진다[3,25]. 다음은 Watts와 Strogatz가 정의한 좁은 세상 네트워크의 중요한 두 가지 특성이다.

- 평균 최단 거리(Characteristic path length : L)는 네트워크에 포함된 노드들 중에서 선택가능한 모든 두 노드들의 경로 길이(path length)들의 평균이다. 경로 길이, $d(i, j)$ 는 i 노드와 j 노드 사이의 최단 경로에 있는 링크 수(number of edges)이다.
- 결집계수(Clustering coefficient : C)는 지역적인 이웃들과의 당파성을 나타내는 척도이다. k 개의 이웃을 가지는 노드에 대해서, 이 k 개의 이웃들 사이에 최대 $kC_2 = k(k-1)/2$ 개의 링크가 존재할 수 있다. 노드의 결집계수는 이 최대 링크 수에 대하여 실제로 이웃들 간에 존재하는 링크의 비율로 구해진다. 네트워크의 결집계수(C)는 그래프에 있는 모든 노드들의 결집계수에 대한 평균이다.

그림 1은 Watts와 Strogatz에 의해서 좁은 세상 네트워크의 수학적 정의를 그림으로 표현한 것이다. 정규 네트워크(왼쪽)는 노드들을 원 형태로 놓고, 각 노드들이 바로 이웃 노드와 그 다음 노드들에 직접 연결 되도록 한다. 그러나 좁은 세상 네트워크(가운데)를 만들기 위해서는 무작위로 선택된 두 노드들을 연결하는 소수의 링크를 추가한다. 이런 소수의 장거리 링크는 멀리 떨어진 노드들 간에 지름길 역할을 하므로 네트워크내의 평균적 분리정도 즉, 평균 최단 거리를 급격히 줄여줄 수 있다. 하지만 이 소수의 장거리 링크는 결집계수에 거의

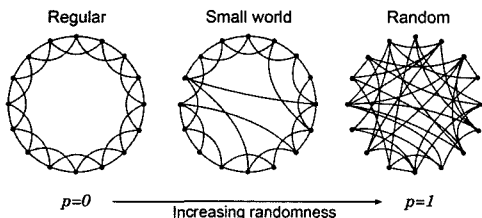


그림 1 확률 p 에 따라서 규칙 네트워크와 좁은 세상 네트워크와 무작위 네트워크가 만들어 진다[3].

표 1 자연계에 존재하는 네트워크에 대한 평균 거리 및 결집계수[5]

	평균 최단 거리	무작위 그래프 평균 최단 거리	결집 계수	무작위 그래프의 결집계수
영화배우망	3.65	2.99	0.79	0.00027
전력망	18.7	12.4	0.080	0.00054 ¹⁾
C. elegans 신경망	2.65	2.25	0.28	0.05

영향을 주지 않는다. 무작위 그래프(오른쪽)는 임의의 두 노드를 선택해서 이 두 노드 간을 연결하는 링크를 추가한다.

표 1은 우리 생활 주변에 흔히 존재하는 3개의 네트워크 모두가 평균 최단 거리는 $L_{actual} \geq L_{random}$ 이 것이고, 결속계수는 $C_{actual} \gg C_{random}$ 이 것으로, 이 네트워크들이 좁은 세상 네트워크가 가지는 특성을 보여준다. 각 네트워크들의 노드 수(n)와 각 노드 당 평균 링크의 수(k)는 영화배우 네트워크는 $n=225,226$ 이고, $k=61$ 이고 전력망은 $n=4,941$ 이고, $k=2.67$ 이다. C. elegans 신경망은 $n=282$ 이고, $k=14$ 이다.

2.2 척도 없는 네트워크(Scale-free Network)

1960년 Erdoes-Renyi 이후 네트워크의 이론에서 중요한 발전을 이룬 계기는 1998년 Strogatz와 Watts에 의해 이루어 졌고, 이들의 연구에 영향을 받아 Huberman와 Barabasi는 실제의 웹 구조가 어떻게 되어있는지를 연구하였다[4,5]. 그 연구의 결과 웹에서 하이퍼링크의 연결선 수 분포가 멱함수 분포 곡선 즉, $P(k) \sim k^{-\gamma}$ 의 형태를 지니고 있다는 것이 밝혀졌다[5]. 이것의 의미는 홈페이지의 수가 매우 많다면 연결선 수가 굉장히 많은 홈페이지가 존재한다는 것이다. 또한 연결선 수가 적은 홈페이지는 매우 많이 존재한다. 이렇게 연결선 수가 멱함수 법칙에 따르는 네트워크를 “척도 없는 네트워크”이라고 부른다.

척도 없는 네트워크는 자연계나 인간 사회에서 존재하는 여러 현상에 나타나는데, 항공망, 영화배우 네트워크, 논문공저자 네트워크, 인터넷, 생물계의 네트워크(단백질 상호작용 네트워크, 신진대사 네트워크), 무역거래 네트워크, 주식회사간의 주식가격에 대한 상호 관계 네트워크 등에서 쉽게 찾아 볼 수 있다[23-26]. 여기서 주목할 만한 것은 척도 없는 네트워크에서는 연결선 수가 굉장히 많은 노드가 존재한다는 사실이다. 이를 흔히 “허브”라고 부른다[23]. 허브의 출현으로 두 점 사이의 평균거리는 무작위 네트워크에서의 평균거리보다 훨씬

1) 원 논문[5]에서는 0.005로 되어있지만, $C = k/n$ 을 적용해보면, $2.67/4941 \approx 0.00054$ 이다.

짧아진다[5]. 따라서 이런 네트워크에서는 허브의 존재를 파악해서 허브가 적으면 허브를 보강해서 네트워크의 견고성을 강화시킬 수 있다. 또는 이런 허브에는 좀더 강력한 보안과 자원을 투자함으로써 이런 네트워크를 무력화하려는 어떠한 침해로부터 사전에 허브를 안전하게 보호할 수 있다[24]. 웹의 구조가 척도 없는 네트워크로서 허브가 존재한다는 사실이 밝혀지면서 다른 많은 분야들에서도 허브를 찾아보려는 연구가 활발하게 이루어졌다. 그 결과 과학자들이 이제까지 연구해 온 대부분의 대규모 복잡계 네트워크에서 허브가 발견되었다. 따라서 허브는 우리의 복잡하고 상호 연결된 세계의 보편적 구성요소인 것이다. 척도 없는 네트워크의 degree-노드가 가지는 링크 수-에 대한 노드의 수를 나타내는 막대그래프를 소위 로그-로그 플롯이라는 것에 대 봤을 때, 그림 2와 같은 모습을 보인다. 이는 링크의 분포가 멱함수 법칙(power law)이라고 불리는 수식을 정확하게 따르고 있음을 보여준다.

지난 십 수 년 동안의 연구 과정에서 과학자들은 자연이 어떤 경우에는 중형이 아니라 멱함수 법칙 분포에 따르는 양을 생성해 낸다는 사실을 인식하게 되었다. 멱함수 법칙 분포는 중형 커브 분포와 매우 다르다. 우선 멱함수 분포에는 정점(peak)이 없다. 멱함수를 따르는 그래프는 단조 감소 커브로서, 다수의 작은 사건들이 소수의 큰 사건들과 함께 발생한다는 것을 시사한다[5,26]. 각각의 멱함수 법칙은 그것에 특유한 지수(exponent)에 의해 규정되는데, 이것이 예를 들면 극히 인기 있는 웹 페이지가 인기가 적은 것들에 비해 어느 정도나 많이 있는가를 결정한다.

네트워크에 있어서 멱함수는 연결선 수의 분포를 서술하는 것이므로 이 지수는 흔히 연결선 지수(degree exponent)라고 불린다[27,28]. 웹에서의 이 연결선 지수

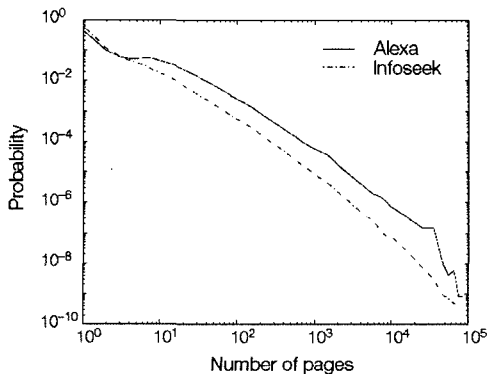


그림 2 Alexa와 Infoseek의 검색 로봇이 탐색한 사이트 들내에 분산된 페이지들의 로그-로그 플롯 그래프[5]

는 k 개의 들어오는 링크를 가진 웹페이지의 개수를 $N(k)$ 라고 할 때, $N(k)$ 는 $N(k) \sim k^{-\gamma} \times N$ 을 따른다. 여기서 γ 는 연결선 지수를 의미한다. 로그-로그 플롯 상에서 기울기(연결선 지수)는 $\gamma=2.1$ 에 가까운 값을 갖는다는 것을 보여주고, 나가는 링크에 대해서도 동일한 패턴을 보이는데 $\gamma=2.5$ 이다[5]. 일반적으로 연결선 지수 γ 의 값이 2~3사이의 값을 가지는 네트워크를 척도 없는 네트워크라 한다[28]. 앞에서 말했듯이 이런 네트워크를 분석하기위해서 레이아웃 알고리즘을 적용해서 시각적으로 나타내는데, 시각적으로 나타낼 때 척도 없는 네트워크에서 가장 중요한 구성요소인 허브를 잘 표현할 필요가 있다. 따라서 본 논문에서는 다음에 설명하는 기존의 여러 레이아웃과 다르게 허브를 잘 표현하는 그래프 레이아웃을 제안한다.

2.3 그래프 레이아웃 관련연구

그래프 레이아웃의 방법에는 크게 Spring embedded 방법과 Energy-Based Placement방법이 있다. 각각에 대해서 알아보면, 우선 Spring embedded방법은 노드와 노드사이에 링크가 있으면, 이것을 스프링으로 가정한다. 스프링은 노드들이 서로 너무 멀리 떨어지거나 너무 가까이 접근하지 않도록 적당한 길이와 강도를 갖도록 설정한다[7]. 에너지 기반의 방법들은 일반적으로 두 부분으로 나뉜다. 에너지 모델과 최소한의 전제에너지를 가지는 상태를 찾는 알고리즘으로 구분된다.

대표적인 Spring embedded방법들을 살펴보면, 가장 처음 Spring embedded 방법을 사용해서 그래프 레이아웃을 표현한 사람은 1984년 Eades에 의해서다. 그는 노드 간에 서로 밀치는 척력(斥力)을 다음과 같은 식으로 표현하였다.

repelling forces(척력) :

$$f_{rep}(p_u, p_v) = \frac{c_e}{\|p_v - p_u\|^2} \cdot \overrightarrow{p_u p_v}$$

위 식에서 c_e 는 반발 상수를 나타낸다. $u, v \in V$ 일 때, 비인접(非引接) 노드 u, v 의 모든 쌍들 사이에 존재하는 척력을 위의 식으로 나타냈다. 그리고 척력과 반대방향으로 작용하는 힘으로 스프링이 존재한다. 스프링은 인접한 노드들에 작용하는 힘으로 $u, v \in V$ 일 때, 인접 노드 u, v 의 모든 쌍들 사이에 존재한다. 이 모델에서는 보통의 스프링에 적용되는 Hooks spring law²⁾가 아니라, 좀 더 먼 거리의 노드들에는 더욱 약한 힘이 적용되는 대수적인(logarithmic) 스프링을 사용한다. 다음은 Eades의 스프링에 적용되는 수식이다.

2) Hooks spring law의 완전한 수식은 다음과 같다.

$$f = \left[-k_s (\|x_a - x_b\| - r) + k_d (v_a - v_b) \right] \frac{x_a - x_b}{\|x_a - x_b\|}$$

logarithmic spring :

$$f_{spring}(p_u, p_v) = c_s \cdot \log \frac{\|p_u - p_v\|}{l} \cdot \frac{1}{\|p_u - p_v\|}$$

위의 수식에서 스프링의 실제 길이가 본래의 길이보다 크거나 작은지가 스프링 힘의 방향을 결정한다. 상수 c_s 는 스프링의 강도를 조정하는 파라미터이다. Eades 이후 1991년 Fruchterman과 Reingold는 레이아웃의 계산에 있어서, 여러 면에서 속도향상을 보이도록 Spring embedded 방법을 개선했다. 우선 힘(force)을 좀 더 빠르게 평가하도록 하고, 척력은 모든 노드들의 쌍(pair) 사이에 적용시키고, 스프링 대신 인력(引力)을 추가하였다. 인력은 인접한 노드들 사이에 적용되며 인접한 노드에서는 인력과 척력의 조합이 스프링과 같은 힘으로 산출된다. 1995년에 Frick가 기본적인 스프링 모델에 커다란 변화를 보였는데, 이 모델 역시 알고리즘의 속도를 향상시키고 레이아웃의 질을 높이도록 개선되었다. 척력과 인력에서 제곱근이 발생하지 않도록 했으며, 높은 degree를 가진 노드들의 당기는 힘을 효과적으로 감소시키는 장점이 있다. 그리고 새로운 중력(gravitational force)이 식에 도입되었다. 중력은 더 많은 링크를 가진 노드들이 더 강하게 인접한 노드들을 끌어당기도록 되어 있다.

다음은 Energy-Based Placement 방법에 대해서 살펴 보면, force-directed 방법들은 모델이 force 시스템이고, 각 노드에 미치는 전체 힘의 합이 제로 값인 균형 상태를 찾는 알고리즘으로 되어 있다. 힘은 에너지의 감소 변화이기 때문에, force directed 모델은 지역적인 최소 에너지를 찾는 에너지 모델로 쉽게 변경될 수 있다. 대부분의 힘과 에너지 모델들은 노드들이 많은 수의 노드들과 상호 관계성을 가진다면, 이론적인 거리로 평균하는 그래프 레이아웃에 적합하다. 에너지 모델의 목적은 인간의 시각으로, 그려진 그래프의 속성을 추론하도록 만드는 것이다. 추론하기에 좀 더 효과적이고 유용한 결과를 만들기 위해서 작고 일정한 에지(edges)의 길이, 잘 분산된 노드들, 잘 분리된 클러스터와 같은 분명한 특성을 레이아웃에서 요구한다. 본 연구에서는 레이아웃 알고리즘에 요구하는 분명한 특성은 허브이다. 다음은 일반적인 에너지 모델의 수식이다.

$$\sum_{u,v \in E} (f \|p_u - p_v\|) - \sum_{u,v \in V^{(2)}} g \|p_u - p_v\|$$

이 모델에서 $f(x)$ 는 거리 x 를 가지는 인접한 두 노드 사이에 존재하는 인력으로 해석될 수 있다. 그리고 $g(x)$ 는 거리 x 를 가지는 인접한 두 노드 사이에 존재하는 척력으로 볼 수 있다. 여기에 적용되는 이론은 매우 간단한데, 서브그래프 G_1 과 G_2 가 매우 높게 결집된 쌍인지 아니면 약하게 결집된 쌍인지를 그래프로 그려주

다. 두 개의 서브 그래프는 n_1 과 n_2 노드들을 가진다 ($n_1, n_2 \in \mathbb{N}_+$). 그리고 두 서브그래프 사이에는 n_e 개의 링크(edges)들이 있다($n_e \in \mathbb{N}_+$, $n_e \ll n_1 n_2$). 이 두 서브그래프가 매우 강하게 결집되어 있다면 매우 가까운 거리에 위치될 것이다. 가장 최근의 에너지 모델은 2003년 Noack에 의해서 r -PolyLog 에너지 모델이 제안되었다. 이 r -PolyLog 에너지 모델은 가중치 링크를 가진 그래프로 쉽게 일반화 될 수 있다[22]. Noack의 에너지 모델은 다음과 같은 수식으로 표현된다.

$$U_{r-PolyLog}(p) = \sum_{u,v \in E} (w_{u,v} \|p_u - p_v\|)^r - \sum_{u,v \in V^{(2)}} \ln \|p_u - p_v\|$$

여기에서 $w_{u,v}$ 은 인접한 노드 u, v 의 링크에 존재하는 가중치(加重值)로 해석될 수 있다. 그리고 가장 클러스터를 잘 표현하는 Noack의 에너지 모델은 차수가 ($r \approx 1$)일 때, 즉 $Lin-Log$ 에너지 모델이다. 차수가 2이상이면 노드 u, v 의 링크에 가중치가 없는 비가중치 모델이 되어서 네트워크상의 링크 길이가 일정한 에너지 모델로, 클러스터를 거의 표현하지 못하는 이전의 에너지 모델과 같아진다. 그래서 차수가 ($r=3$)일 때, 3-PolyLog Energy Model은 Fruchterman-Reingold로 해석할 수 있다[22,29]. Noack의 에너지 모델($Lin-Log$)은 클러스터를 잘 표현하기는 하지만, 모든 네트워크의 특성을 잘 나타내는 것은 아니다. 특히 척도 없는 네트워크의 허브를 표현하는 데는 적합하지 않다. 따라서 본 논문에서는 척도 없는 네트워크를 위한 그래프 레이아웃 알고리즘을 제안한다.

3. 척도 없는 네트워크를 위한 그래프 레이아웃 알고리즘

그래프 레이아웃에 사용되는 모델 중에서 force directed와 energy based model은 일반적인 무방향 그래프를 직선으로 그리는 가장 인기 있는 방법들이다. 이 방법들은 상대적으로 구현하기가 쉽고, 많은 그래프에서 만족할 만한 결과들을 보인다[30].

앞에서 말한 바와 같이 기존의 좁은 세상 네트워크의 레이아웃 방법들은 주로 클러스터 중심의 레이아웃 알고리즘을 사용한다. 하지만 우리 주변에 존재하는 복잡한 네트워크 중에서 흔히 발견되는 척도 없는 네트워크를 표현하기에, 기존의 에너지 모델로는 표현하기가 매우 어렵다. 왜냐하면 척도 없는 네트워크에는 클러스터가 중요한 구성요소가 아니라 허브가 중요한 구성요소이기 때문이다. 따라서 본 논문에서는 일반적인 에너지 모델에서 척력을 나타내는 수식부분을 허브의 특성을 반영하도록 고치는 것이다. 표 2에서는 제안한 그래프 레이아웃의 전체 프로그램 수행 순서의 의사코드를 나타낸다.

표 2 제안한 그래프 레이아웃 전체 알고리즘 의사코드

```

/* 레이아웃을 적용할 Network에서 N은 전체 노드 수를 나타내고,
* k는 각 노드들의 평균 링크 수를 나타낸다.
*/
Network = Create_Network(N, k);
//네트워크의 Power law의 연결선 지수를 계산한다.
gamma = Power_law(Network);
if(gamma >=2 && gamma <=3)
{ Hub_Layout 수행 } //제한한 에너지 모델 알고리즘 수행
else { Cluster_Layout 수행 }
// Noack의 Lin_Log 에너지 모델 알고리즘 수행

Methode Create_Network(nodes, mean_edges)
{ // 척도 없는 네트워크 생성
  for(i=0; i <= nodes; i++){
    node를 생성;
    for(i=0; i <= mean_edges*nodes; i++){
      // 두 노드를 선택, 각 노드들은  $\Pi(k_i) = \frac{k_i}{\sum_j k_j}$ 의 확률로 선택;
      // 선택된 노드 사이에 링크 추가;
    }
  }
}

Methode Power_law(Network)
{
  log_D[math.log(degree)]; // degree의 log 값
  log_N[math.log(Node_num)]; // 노드 수의 log 값

  // 최소자승 선형근사식으로  $P(k) = \frac{n_k}{N} = k^{-\gamma}$ 의 연결선 지수
  //  $\gamma$ 을 계산;
}

```

척도 없는 네트워크 레이아웃 절차

• 네트워크 식별 단계 : 척도 없는 네트워크인지를 알기 위해서는 분석할 네트워크에서 노드와 Degree의 관계가 멱함수 법칙($P(k) = k^{-\gamma}$)을 따르는지를 조사한다. 좀 더 상세하게 설명하면, 네트워크에서 k개의 링크를 가진 노드가 존재할 확률 $P(k)$ 는, $P(k) = \frac{n_k}{N} = k^{-\gamma}$ 로 표현할 수 있다. 여기에서 n_k 는 k개의 링크를 갖는 노드의 수이고, N은 전체 노드의 개수이다. 위 식의 의미는 연결선 지수(γ)의 값이 2인 네트워크에서, 링크의 개수(degree)가 5인 노드가 존재할 확률은 $5^{-2} = 1/25 = 4\%$ 이다. 즉 확률적으로 100개의 노드를 가진 네트워크가 있으면 노드 4개 정도가 5개의 링크를 가질 것임을 나타낸다. 마찬가지로 10개의 링크를 갖는 노드들이 존재할 확률은 1%이므로 네트워크에서 1개 정도의 노드가 존재할 수 있다. 그럼 100개의 링크를 갖는 노드가 1개 정도 존재 하려면 몇 개의 노드들로 네트워크가 구성되어야 하는가? 이 식에 대

입하면, 10,000개의 노드를 가진 네트워크라면, 노드 1개 정도는 100개의 링크를 가질 수 있음을 알 수 있다. 멱함수 법칙의 연결선지수를 구하기 위해서 양변에 상용로그(\log_{10})를 취해서 로그-로그 플롯에 표시한다. 이 플롯에 표시되는 점들의 기울기를 구하기 위해서 이산 최소자승근사(최적의 선형근사)법을 이용하면 되는데, 이 기울기로 바로 연결선지수를 구할 수 있다. 다음은 최소자승근사에서 기울기(a)를 구하는 공식이다.

$$\begin{aligned}
 & \text{최소자승 선형근사식 : } \sum_{i=1}^m [y_i - (ax_i + b)]^2 \\
 & a = \frac{m \left(\sum_{i=1}^m x_i y_i \right) - \left(\sum_{i=1}^m x_i \right) \left(\sum_{i=1}^m y_i \right)}{m \left(\sum_{i=1}^m x_i^2 \right) - \left(\sum_{i=1}^m x_i \right)^2}
 \end{aligned}$$

이 플롯에서 y축의 값은 ($\log_{10} rank$)이거나 ($\log_{10} degree$)로 만들 수 있다. 그래서 m은 최대 $\log_{10} rank$ or $\log_{10} degree$ 값을 나타낸다. x축은 $\log_{10} nodes$ 이다. 하지만 본 논문에서는 주로 y축의 값을 $\log_{10} degree$ 로 실험하였다.

• 레이아웃 적용 단계 : 네트워크 식별 단계에서 구한 멱함수 법칙의 연결선 지수(γ)의 값이 2와 3사이에 존재하면, 본 논문에서 제안한 허브 중심의 레이아웃(Hub_Layout)을 적용하고, 그렇지 않으면 클러스터 중심의 레이아웃(Cluster_Layout)을 적용한다. 멱함수의 계산 결과 연결선 지수의 값이 2미만이면, Cluster_Layout을 수행하는데, Cluster_Layout은 Noack의 식에서 차수(r)의 값을 1로 하는 Lin-Log(Clustering Energy Model)알고리즘을 그대로 이용해서 그래프 레이아웃을 그린다. 다음은 Lin-Log알고리즘 식이다.

$$U_{1-RelyLog}(p) = \sum_{u,v \in E} (w_{u,v} \| p_u - p_v \|^2)^{1-\gamma} - \sum_{u,v \in V^{(2)}} \ln \| p_u - p_v \|$$

네트워크에 클러스터가 존재하고, 클러스터가 네트워크의 중요한 특성을 의미한다면, Noack의 Lin-Log 에너지 모델은 매우 우수한 그래프 레이아웃을 표현할 것이다[34]. 하지만, 허브가 존재하고 허브가 네트워크에서 중요한 구성 요소인 네트워크라면, Noack의 에너지 모델은 허브를 클러스터 내부로 응집한 모습으로 그려져 허브를 파악하기 힘들게 할 것이다. 따라서 이럴 경우에는 제안한 알고리즘으로 허브를 표현하여야 한다.

3.1 제안한 알고리즘

이 절에서는 제안한 허브 중심의 레이아웃 알고리즘의 구현에 대해서 설명한다. 기존의 Noack의 클러스터 중심의 레이아웃 알고리즘으로는 허브를 가지는 척도 없는 네트워크를 그렸을 때, 허브가 네트워크에 존재하더라도 그것과 관계없이 이 네트워크도 높은 절집계수

를 가진 네트워크이기 때문에 매우 강한 결집력을 가진 클러스터들로 네트워크가 표현된다. 따라서 허브가 이 클러스터 안쪽에 놓여지기 때문에 허브들을 잘 구별하기가 어렵다. 제안한 Hub_Layout은 일반적인 에너지 모델에서 척력을 나타내는 수식부분을 허브의 특성을 반영하도록 고치는 것이다. 다음은 제안하는 에너지 모델의 수식이다.

$$f_{Energy}(p_u, p_v) = \sum_{u,v \in E} (f \| p_u - p_v \|) - \sum_{u,v \in V^{(H)}} (g \| p_u - p_v \|)$$

$$g(p_u, p_v) = (1/length) * D_{uv}$$

앞의 수식에서 u 와 v 노드 사이에는 적용되는 두 가지 힘, 즉 인력과 척력(허브성 척력)이 존재하는데, 허브들끼리는 서로 밀치고 그러지 않은(degree가 낮은) 노드들은 가깝게 그러지도록 D_{uv} 항을 도입해서 허브를 잘 표현하는 알고리즘으로 수정하였다. 아래 그림 3에서처럼 거리가 다른 두 노드이지만, 각각의 degree는 같다. 이럴 경우 노드 서로 간에 미치는 힘의 크기는 거리에 반비례해서 달라져야한다. 예를 들면, 물리적인 힘 중에서 전자력을 보면, 거리에 따라 서로 당기는 힘과 미는 힘의 세기가 달라지기 때문이다. 따라서 두 노드에 미치는 힘의 영향은 $(1/거리) * D_{uv}$ 라는 거리에 반비례 관계로 적용되어야 한다.

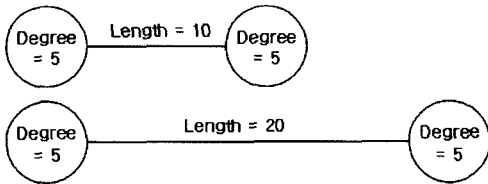


그림 3 D_{uv} 항은 거리와 반비례 관계

D_{uv} 항의 구체적인 작용은 두 노드의 degree에 따라 다르게 작용해야 한다. 즉 degree가 모두 다 높은 노드 간에는 허브성 척력이 강하게 작용해야 하고, 또 노드 모두 degree가 낮다면 노드 간에는 허브성 척력이 약하게 작용해야 한다. 그리고 두 노드 중 한 노드는 degree가 높지만 다른 노드는 degree가 낮을 경우에도 척력이 약하게 작용해야만 한다. 이와 같은 작용을 하도록 degree와 척력에 관한 식을 다음과 같이 표현할 수 있다.

$$D_{uv} = ((k_u + k_v) - (|k_u - k_v|))^{\gamma}$$

위의 식에서 $(k_u + k_v)$ 부분은 허브성에 비례해서 척력이 작용하도록 하기 위한 것이고, $(|k_u - k_v|)$ 부분은 서로 허브성이 차이가 나는 노드(특히 허브성이 매우 낮은 노드)와 허브 노드들 간의 척력을 감소시키기 위해서 도입한 부분이다. 다음은 제안한 식이 노드들 간의 거리에

반비례해서 척력이 적용되며, 거리가 일정 비율로 증가할 때 척력의 크기도 일정한 비율로 감소함을 증명하고, 네트워크의 규모에 관계없이 허브성에 비례해서 노드들 간에 척력이 작용함을 증명한다.

정리 1. 두 노드사이의 허브성 척력 $g(p_u, p_v)$ 이 거리에 반비례한다.

증명 1. $g(p_u, p_v)$ 의 정의에 의해서 당연히 성립한다. QED.

정리 2. 동일한 거리를 갖는 두 노드 u, v 에 대하여 각 노드의 degree가 각각 a 배 증가하면 D_{uv} 는 a^{γ} 배가 된다.

증명 2. 노드 u, v 의 degree가 각각 k_u, k_v 이면, 두 노드 사이에 존재하는 허브성 척력은 $D_{uv} = ((k_u + k_v) - (|k_u - k_v|))^{\gamma}$ 이다. 이제 u', v' 의 degree를 각각 $a * k_u, a * k_v$ 라 하면 두 노드 사이에는 $D_{u'v'} = (a * (k_u + k_v) - a * (|k_u - k_v|))^{\gamma}$ 인 값이 허브성 척력이다. 그러면 $D_{u'v'} = (a * ((k_u + k_v) - (|k_u - k_v|)))^{\gamma}$ 이고, 지수의 분배법칙에 의해서 $D_{u'v'} = a^{\gamma} * ((k_u + k_v) - (|k_u - k_v|))^{\gamma}$ 이고. 따라서 $D_{u'v'} = a^{\gamma} * D_{uv}$ 이다. 따라서 동일한 거리를 갖는 두 노드 u, v 에 대하여 각 노드의 degree가 각각 a 배 증가하면 D_{uv} 는 a^{γ} 배가 된다. QED.

아래의 그림 4에 나타난 다양한 degree의 크기를 가지는 노드들을 통해서 제안한 식이 허브성 척력을 잘 나타내고 있는지를 예를 들어 설명하면, 연결선 지수 γ 가 2.5일 경우 그림 4의 (a)와 (b)의 관계를 살펴보면, D_{uv} 값이 각각 5.66, 316의 값을 가진다. 이 경우 1:55.9의 비율을 보이고, 그림 4의 (c)와 (b)의 관계도 각각 5.66, 316의 값을 가진다. 이 경우 1:55.9의 비율을 보인다. 또한 그림 4의 (d)와 (e)의 관계는 각각 32, 1778.8의 값을 가진다. 이 경우 1:55.9의 비율을 보인다. 이런 결과는 허브들 간의 척력이 그렇지 않은 경우에 비해 특별히 강하게 작용함을 증명한다. 그리고 위의 정리 2의 예를 들면, 그림 4의 (b)와 (e)와 (f)의 관계를 보면 모두 크기가 다른 허브로 볼 수가 있는데, 각각의 노드들은 두 배씩 degree가 증가한다. 이들 간의 D_{uv} 값을 구하면, 각각 316, 1788, 10119의 값을 가진다. 이 경우 1:5.66:32의 비율을 보인다. 5.66:32는 다시 1:5.66이 나온다. 따라서 degree의 값이 2배씩 변화를 가질 때 일정한 비율(1:5.66)로 바뀌게 됨을 증명한다. 즉 degree = 1 : 2 일 때, $D_{uv} = 1 : 2^{2.5} \approx 5.6568$ 의 값을 가진다.

본 논문에서 제안한 알고리즘은 네트워크의 규모에 관계없이 허브성 척력이 적용되도록 네트워크의 규모(전체 노드수와 링크수)에 따른 이상적인 거리를 계산한다. 이에 관한 내용은 논문[31]을 참조하여 구현하였다. 이

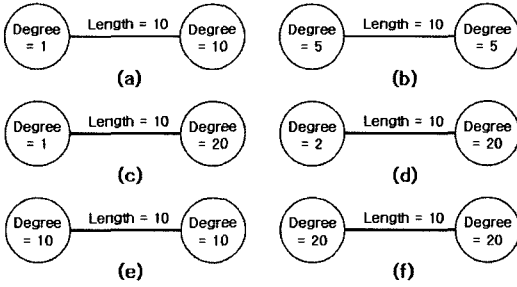


그림 4 다양한 Degree를 가지는 노드들

상과 같이 정리 1, 2에 의하여 제안한 알고리즘은 허브성 척력이 거리에 반비례하고, 각각의 degree가 증가함에 따라서 일정한 크기로 증가하며, 네트워크의 규모에 관계없이 적용되는 것이 특징이다.

4. 실험

본 논문에서는 네트워크에 허브가 존재할 때 허브를 잘 표현하는 것이 목적이므로 허브가 존재하지 않는 네트워크에 대해서는 실험을 하지 않는다. 따라서 실험에 사용한 네트워크는 다음과 같은 과정을 통해서 만들어진다[32].

- 단계 1 : 링크를 가지지 않는 N 개의 노드가 네트워크를 구성한다.
- 단계 2 : 링크를 추가하기 위해서 두 노드를 선택한다. 이때 각각의 노드(i)는 링크 수(k)에 따라 선택될 확률($\Pi(k_i)$)이 결정된다. 선택될 확률 계산식은 다음과 같다.

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}$$

- 단계 3 : 네트워크에 원하는 링크의 수만큼 단계 2를 반복한다.

위 과정은 Barabasi의 네트워크의 성장이론에 근거해서 작성되었다. 본 논문의 실험에서 사용한 네트워크는 Barabasi의 모델을 이용해서 만든 가상의 네트워크이다.

4.1 실험에 사용한 네트워크의 특성

이 절에서는 실험에 사용한 네트워크의 특성에 대한 설명이다. 실험에 사용된 네트워크는 척도 없는 네트워크로서 Barabasi의 모델을 이용해서 만들어 졌다. Random Network는 실험에 사용된 네트워크와 같은 수의 노드와 링크를 갖는 완전한 무작위 네트워크이다. 이 내용들이 표 3에 정리되어 있다.

결집계수를 보면 척도 없는 네트워크는 0.214로 무작위 네트워크의 0.05에 비해 상당히 높은 값을 나타내며, 두 네트워크의 평균 최단거리는 거의 차이가 나지 않는다. 그러나 최대 degree는 척도 없는 네트워크가 31로

표 3 실험에 사용된 척도 없는 네트워크와 같은 수의 노드와 링크를 가지는 무작위 네트워크의 비교

	척도 없는 네트워크	무작위 네트워크
전체 노드 수	100	100
전체 링크 수	250	250
최단평균거리(L)	3.02	2.998
결집계수(C)	0.214	0.05
최대 링크 수	31	10
연결선지수(γ)	2.82	3.63

표 4 실험에 사용된 네트워크의 특성

순위(Rank)	링크 수(Degree)	노드들의 개수
1	1	49
2	2	10
3	3	7
4	4	4
5	5	4
6	6	2
7	7	1
8	8	2
9	9	2
10	10	1
11	11	1
12	12	4
13	13	1
14	14	3
15	15	1
16	16	1
17	18	1
18	19	1
19	21	1
20	23	1
21	24	1
22	30	1
23	31	1

매우 많은 링크를 가지는 노드가 존재하는 반면, 무작위의 네트워크에서는 그런 노드가 없다. 표 4에서는 실험에서 사용된 네트워크의 특성을 상세히 나타내고 있다. 이 표 4의 값을 이용해서 그림 5~8을 그렸을 때 척도 없는 네트워크의 특성인 멱함수 그래프 분포를 보이고 있다. 하지만, 무작위 네트워크를 생성해서 같은 실험을 한 결과 링크 수에 따른 노드들의 그래프가 종모양의 커브를 그리는 분포를 보였다.

그림 6은 순위(Rank)에 따른 노드수를 멱함수 법칙 그래프로 나타냈을 때, 연결선 지수 $\gamma=2.82$ 를 나타내고 있다. 따라서 이 네트워크는 척도 없는 네트워크라 불릴 수 있다. 왜냐하면, 그림 5의 각 축에 대한 로그 값을 취했을 때 그래프에 나타나는 모양이 선형으로 표현된다. 이럴 경우 제안한 허브 중심의 그래프 레이아웃 알

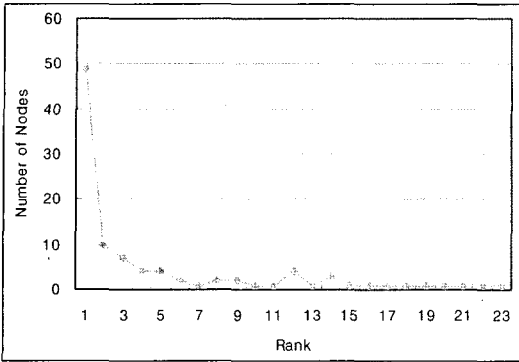


그림 5 순위(Rank)에 따른 노드 수를 나타내는 그래프

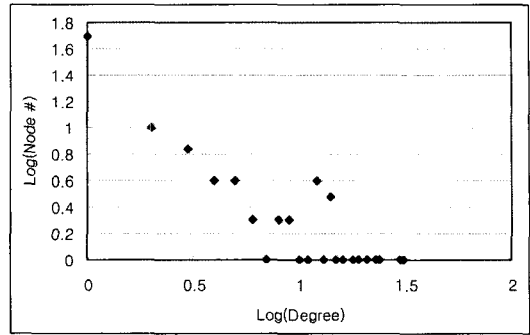


그림 8 링크 수(Degree)에 따른 멱함수 법칙 그래프

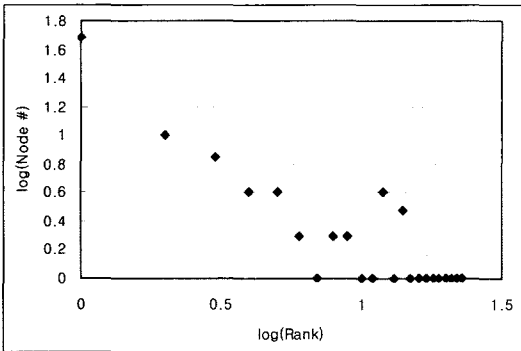


그림 6 순위(Rank)에 따른 멱함수 법칙 그래프

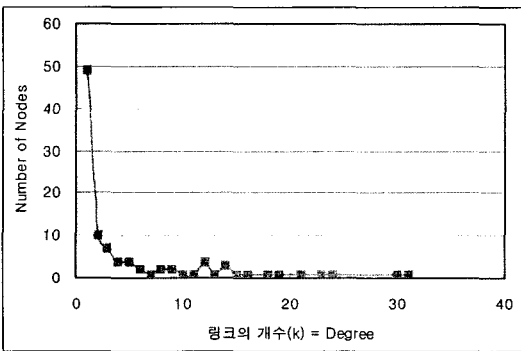


그림 7 링크 수(Degree)에 따른 노드들의 그래프

고리즘으로 이 네트워크를 나타내는데 사용해야 한다. 이럴 경우 레이아웃에서 허브가 다른 노드에 비해 확연히 잘 구별되어 네트워크의 특성을 분석하는데 도움이 된다.

그림 8은 링크 수(Degree)에 따른 노드수를 멱함수 법칙 그래프로 나타냈을 때, 연결선 지수 $\gamma=2.76$ 을 나타내고 있다. 따라서 이 네트워크도 척도 없는 네트워크라 볼 수 있다. 이와 같이 노드가 degree와 rank 두 경우 모두에서 멱함수 법칙을 따르는 척도 없는 네트워크

특성을 보였다. 하지만 본 연구에서 제안한 알고리즘에서는 링크 수(Degree)를 주로 사용한다.

4.2 실험 결과

그림 9는 실험에 사용된 네트워크의 연결선 지수 γ 값이 2이상이므로 허브 중심의 레이아웃으로 그렸다. 이 레이아웃에서 허브들이 외곽으로 적절히 분산되어 그려지고 있다. 따라서 허브들의 정보를 파악하는데 제안한 알고리즘이 매우 유용하다. 만약 이 그래프의 레이아웃을 그리는데, 기존의 클러스터 중심으로 표현하는 레이아웃 알고리즘을 적용해 본다면, 그림 10에서 보여주는 것처럼 중심에 노드들이 뭉쳐져 척도 없는 그래프에서 알고자 하는 허브들을 분석하기가 매우 힘들다. 그림 10에서 중심에 노드들이 뭉친 이유는 실험에 사용한 네트워크의 결집계수가 허브를 중심으로 상당히 높기 때문에 당연히 클러스터 중심의 그래프 레이아웃 알고리즘(Noack의 *Lin-Log*)에서는 허브들을 강한 결집을 보이는 클러스터로 나타낸다. 그림 11은 Noack의 식을 이용하지만, 링크에 가중치를 주지 않은 경우의 레이아웃 모

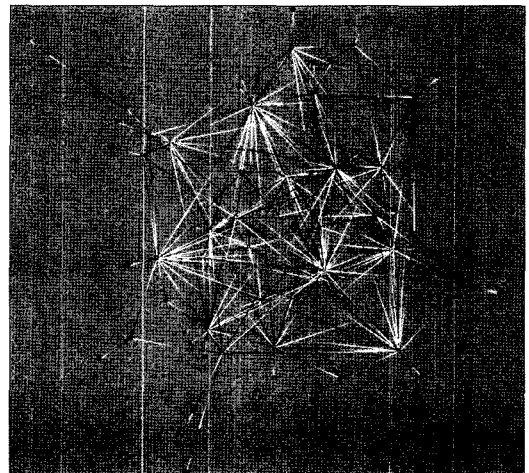


그림 9 제안한 알고리즘으로 그린 그래프 레이아웃

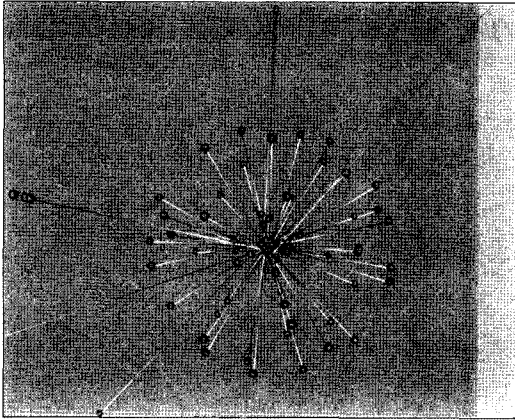


그림 10 Noack *Lin-Log*로 그린 그래프 레이아웃

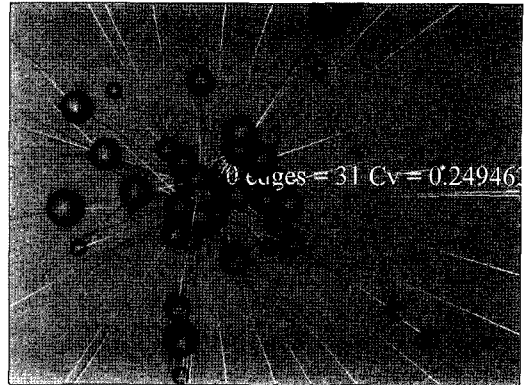


그림 13 *Lin-Log* - 허브를 구별하기 어렵다.

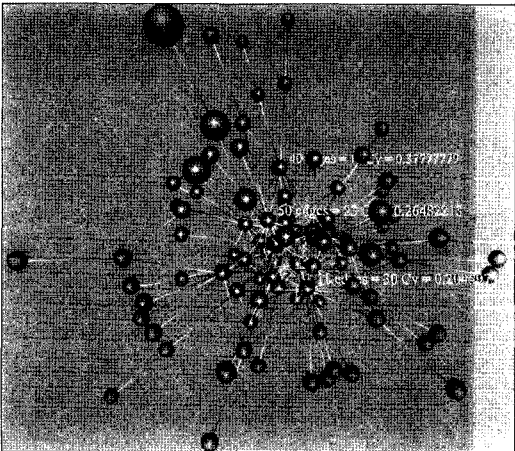


그림 11 차수(r)가 3인 3-*Poly-Log*

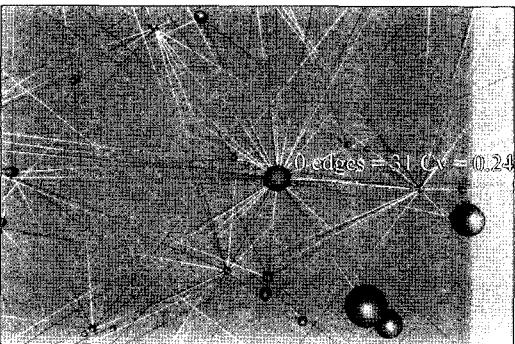


그림 12 제안한 알고리즘-허브를 잘 구별할 수 있다.

습을 보여주고 있다. 이 그림에서는 클러스터와 허브를 잘 표현하지 못했다. 그림 12와 13에서 $\max \text{degree}=31$ 을 가지는 허브를 두 그래프 레이아웃에서 찾을 때, 제안한 알고리즘에서는 허브들이 척력에 의해 멀리 떨어

져 있어서 분명하게 구별할 수 있는데 비하여, *Lin-Log*에서는 클러스터를 중심으로 내부에 뭉쳐져 있어서 구별하기가 매우 힘들다.

5. 결론

네트워크는 공학이나 자연과학은 물론이고 기업 이사회 인맥이나, 논문의 공동저자, 심지어 영화배우망 등과 같은 사회과학의 여러 분야를 연구하는데 중요하게 사용되는 모델이다. 그리고 그것을 시각적으로 나타내는 그래프 레이아웃은 네트워크 분석에서 매우 중요하며 다양한 알고리즘도 많이 연구되었다. 그런데 자연에는 척도 없는 네트워크가 아주 많다[26]. 이런 척도 없는 네트워크에는 허브가 존재하는데 기존의 알고리즘은 클러스터를 잘 표현하는 정도이지 허브의 표현은 매우 어렵다. 그래서 본 논문에서는 척도 없는 네트워크를 잘 표현하는 그래프 레이아웃 알고리즘을 제안했다. 본 논문에서 제안한 알고리즘에서 허브들 간에 작용하는 허브성 척력이 거리에 반비례하고, 허브들의 degree가 a 배 증가하면, 허브사이에 작용하는 척력의 크기는 a^γ (γ 는 연결선 지수)배가 된다. 또한, 전체 노드수와 전체 링크수에 따라 적용되는 힘의 크기를 조정하는 계수를 두어서 네트워크의 규모에 관계없이 허브성 척력이 적용되는 특성이 있다. 제안한 알고리즘은 먼저 허브의 존재를 식별한 후 그 정보를 토대로 허브끼리 밀도록 하는 것을 원리로 한다. 허브의 존재는 네트워크가 척도 없는 네트워크인지를 구분하는 것과 같다. 따라서 네트워크에 존재하는 노드들의 degree값이 멱함수 법칙의 분포를 따르고 연결선 지수의 값이 2와 3사이에 존재하면 허브가 존재한다고 판단한다. 이렇게 허브의 존재가 확인되면 Noack의 에너지 모델에서 차수 r 의 값을 3으로 하고, 척력에 다음의 항 ($D_{uv} = ((k_u + k_v) - (|k_u - k_v|))^{\gamma}$)을 추가하여 허브끼리 밀치게 했다. 여기서 k 는 노드의

링크 수이고, γ 은 연결선 지수이다. 제안한 알고리즘이 기존의 클러스터링 중심의 레이아웃 알고리즘(Noack의 *Lin-Log*)에 비해서 허브를 잘 표현하는지에 대한 실험을 해보았다. 실험에 사용한 네트워크가 연결선 지수 $\gamma=2.76$ 을 나타내므로 허브를 가지고 있는 척도 없는 네트워크로 볼 수 있다. 따라서 이 네트워크에 제안한 그래프 레이아웃을 적용했을 때, 그림 9와 그림 12에서 보이는 바와 같이, 그림 10, 11, 13과 비교해서 허브를 매우 쉽게 구별하고, 네트워크 분석에 도움이 됨을 알 수 있다. 그리고 향후 연구방향은 좀 더 네트워크의 분류를 세분화해서, 각각의 네트워크에 적합한 그래프 레이아웃을 자동적으로 적용시키는 방법에 대해서 연구한다.

참 고 문 헌

- [1] Gray William Flake, *The Computational Beauty of Nature*, A Bradford Book, the MIT Press, 1998.
- [2] S. Wasserman and K. Faust, *Social networks analysis*, Cambridge University Press, Cambridge (1994).
- [3] D. Watts and S. H. Strogatz, "Collective dynamics of small world networks," *Nature(London)* 393, 440, 1998.
- [4] B.A. Huberman and L.A. Adamic, "Growth dynamics of the world-wide web," *Nature* 401, 131, 1999.
- [5] R. Albert, H. Jeong, and A.-L. Barabasi, "Diameter of the world-wide web," *Nature* 401, 130-131, 1999.
- [6] Charles J. Alpert and Andrew B. Kahng, "Recent directions in netlist partitioning: A survey," Technical report, Computer Science Department, University of California at Los Angeles, 1995.
- [7] M. Kaufmann, D. Wagner (Eds.), *Drawing Graphs: Methods and Models*, Lecture Notes in Computer Science. Springer Berlin / Heidelberg. Volume 2025, 2001.
- [8] Ivan Herman, Guy Melançon, and M. Scott Marshall, *Graph visualization and navigation in information visualization*. volume 6(1), pages 24 - 43, 2000.
- [9] Josep diaz, Jordi Petit and Maria Serna, "A Survey of Graph Layout Problems," *ACM Computins Surveys*, Vol.34, No.3, pp.313-356 September 2002.
- [10] I. Herman, G. Melançon, MS Marshall. "Graph Visualization and Navigation in Information Visualization : a Survey," In: *IEEE Transactions on Visualization and Computer Graphics*, 6(1), pp. 24-43, 2000.
- [11] I. F. Cruz and J. P. Twarog, "3D Graph Drawing with Simulated Annealing", *Proceedings of the Symposium on Graph Drawing GD '95*, Springer-Verlag, pp.162-165, 1995.
- [12] R. Davidson and D. Harel, "Drawing Graphs Nicely Using Simulated Annealing," *ACM Transaction on Graphics*, Vol.15, No.4, pp.301-331, 1996.
- [13] P. Eades, "A Heuristic for Graph Drawing," *Congressus Numerantium*, Vol.42, pp.149-160, 1984.
- [14] C.-S Jeong and A. Pang, "Reconfigurable Disc Trees for Visualizing Large Hierarchical Information Space," *Proceedings of the IEEE Symposium on Information Visualization (InfoViz'98)*, IEEE CS Press, 1998.
- [15] G. di Battista, P. Eades, R. Tamassia, and I.G. Tollis, "Algorithms for drawing graphs: an annotated bibliography," *Computational Geometry: Theory and Applications*, Vol.4, No.5, pp.235-282, 1994.
- [16] Andre M.S. Barreto, Helio J.C. Barbosa, "Graph Layout Using a Genetic Algorithm," sbm, p. 179, VI Brazilian Symposium on Neural Networks (SBRN'00), 2000.
- [17] J. Kleinberg, "The small-world phenomenon: An algorithmic perspective," In *Proc. 32nd ACM Symposium on Theory of Computing*, 2000.
- [18] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Mod. Phys.* 74, 4797, 2002.
- [19] M.E.J. Newman, "The structure and function of complex networks," *SIAM Review* 45, 167256, 2003.
- [20] A.-L. Barabasi and R. Albert. Emergence of scaling in random networks. *Science*, 286: 509~512, 1999.
- [21] P. Eades and Q.-W. Feng, "Multilevel Visualization of Clustered Graphs," *Proceedings of the Symposium on Graph Drawing GD '96*, Springer-Verlag, pp.101- 112, 1997.
- [22] A. Noack. "An energy model for visual graph clustering," In *Proc. 11th Int. Symp. on Graph Drawing*, pages 425 - 436. Springer-Verlag, 2003.
- [23] M. Faloutsos, P. Faloutsos, and C. Faloutsos, On power-law relationships of the internet topology, *Computer Communications Review* 29, 251 (1999).
- [24] Mark, Buchanan, "Ubiquity: the science of history... or why the world is simpler than we think," 2001.
- [25] D. Watts. "Small world: the dynamics of networks between order and randomness," Princeton, 1999.
- [26] M. E. J. Newman, "Power laws, Pareto distributions and Zipf's law," *Contemporary Physics* 46, 323-351, 2005.
- [27] Mark E. J. Newman, Stevens H. Strogatz, and Duncan J. Watts. "Random graphs with arbitrary degree distributions and their applications," *Physics Reviews E*, 64, 2001.
- [28] Bollobas B, Riordan O, Spencer J, Tusnady G, "The Degree Sequence of a Scale-Free Random Graph Process," *Random Structures and Algori-*

- thms 18, 279-290, May 2001.
- [29] A. Noack. "Energy models for drawing clustered small-world graphs," Technical Report 07/03, Institute of Computer Science, Brandenburg University of Technology at Cottbus, 2003.
 - [30] Ulrik Brandes. Drawing on physical analogies. In M. Kaufmann and D. Wagner, editors, Drawing Graphs, LNCS 2025, pages 71-86. Springer-Verlag, Berlin, 2001.
 - [31] Thomas M. J. Fruchterman, Edward M. Reingold, "Graph Drawing by Force-directed Placement," Software-Practice and Experience archive Volume 21(11), November 1991.
 - [32] Bing Wang, Zhongzhi Zhang, Huanwen Tang, Zhilong Xiu. "Evolving Scale-Free Network Model with Tunable Clustering," International Journal of Modern Physics B November 16, 2005.



조 용 만

1996년 강릉대학교 전자계산학과 졸업 (이학사). 1996년~2001년 한전KDN(강릉지사) 근무. 2002년 강릉대학교 컴퓨터공학과 졸업(공학석사). 2002년~2004년 방송통신대학교 조교. 2002년~현재 강릉대학교 컴퓨터공학과 박사과정. 2005년~

현재 강릉대학교 해양센서네트워크시스템기술연구센터 연구원. 관심분야는 인공생명, 인공지능, 복잡계, 해양 센서네트워크



강 대 원

1985년 연세대학교 수학과 졸업(이학사)
1988년 고려대학교 전산학과 졸업(이학사). 1991년 고려대학교 수학과(이학석사). 1996년 고려대학교 컴퓨터학과(이학박사). 1996년~현재 강릉대학교 컴퓨터공학과 교수. 관심분야는 인공생명, 인공

지능, 소프트 컴퓨팅, 복잡계, 로보틱스