

Small-world와 Scale-free 네트워크를 위한 조정 가능한 네트워크 생성 모델

이강원*, 엄희관*, 최혜진*

Tunable Network Generation Model for Small-World and Scale-Free Network

Kang-won Lee*, Hee-kwan Uhm*, Hye-jin Choe*

요약

본 연구에서는 파라미터들을 적당하게 조정함으로써 Small-world 성질과 Scale-free 성질을 나타내주는 일반적인 네트워크 생성 모델을 개발하였다. 네트워크 모델은 각각 파라미터를 갖는 PA(Preferential Attachment)와 TF(Triad Formation) 두 단계로 구성되어 네트워크 성장 과정이 현실세계에 존재하는 네트워크 성장과정과 유사하도록 하였다. 본 연구에서는 두 개의 파라미터 K 와 P_{TF} 를 변화 시켜나가면서 실제로 네트워크를 생성 한 후 ASPL, Clustering Coefficient, Power-law 성질 보유 여부 등을 조사함으로써 네트워크 생성 모델의 파라미터들이 Small-world 성질과 Scale-free 성질에 미치는 영향력을 분석하였다. 분석결과 PA 단계에서 사용한 파라미터 K 는 Power-law 성질 보유 여부를 결정 하는 것으로 나타나 제한한 모델에서 K 는 Scale-free 성질을 나타내주는 파라미터라는 것을 확인하였다. 또한 TF 단계에서 사용한 파라미터 P_{TF} 는 Clustering Coefficient에 직접적으로 영향을 미쳐서 P_{TF} 는 Small-world 성질을 어느 정도 나타내 줄 수 있다.

Key Words : Preferential Attachment, Triad Formation, Average Shortest Path Length, Clustering Coefficients, Power-Law

ABSTRACT

In this study we have developed a network generation model, which can show the properties of both small-world and(or) scale-free by tuning parameters. The network generation model consists of two steps: PA(Preferential Attachment) and TF(Triad Formation), which resembles the network growth process in real world. The effects of parameters on the small-world and scale-free properties are analyzed by investigating ASPL, clustering coefficient, and power-law property of generated networks. It is shown that parameter K for preferential attachment determines whether the network possesses the power-law property or not. Therefore, K is a parameter which represents the scale-free property. Also, it is shown that parameter P_{TF} for triad formation directly affects the clustering coefficient. Therefore it is a parameter which represents the small-world property.

* 이 연구는 서울과학기술대학교 교내 연구비의 지원으로 수행 되었습니다.

◆° First and Corresponding Author : Department of Industrial and Information Systems Engineering, Seoul National University of Science and Technology, kwlee@snut.ac.kr, 정희원

* 서울과학기술대학교 산업정보시스템공학과

논문번호 : KICS2017-04-098, Received April 4, 2017; Revised June 8, 2017; Accepted June 12, 2017

I. 서 론

많은 시스템들이 링크에 의해 연결되는 노드들로 구성되는 복잡계 네트워크(Complex Network)로 표현 될 수 있다. Internet은 개별 라우터들이 혹은 AS(Autonomous System)들이 노드가 되는 네트워크로 볼 수 있으며 WWW(World Wide Web)은 Web 사이트들의 노드로 구성되는 네트워크다. SNS(Social Network Service)에서는 개인이나 조직과 같은 행위자들이 노드들이며 행위자들 사이에 ‘이원적 유대’가 (예로 트위터나 페이스북과 같은 SNS는 팔로잉이나 친구맺기 등) 링크로 표현되는 네트워크다. 이외에도 뇌 신경이나 물질대사(Metabolic), 운송 시스템 등 사회과학이나 자연 과학, 공학 등 많은 분야의 시스템이 복잡계 네트워크로 표현 될 수 있다.

네트워크의 구조가 네트워크의 동적인 행태(Dynamic Behavior)에 어떻게 영향을 미치는 지에 대한 연구가 최근 들어 많은 관심을 받고 있다. 예로 SNS 네트워크에서 신제품이나 뉴스의 전파 속도 등의 분석이라든지 대규모 전력 공급 네트워크나 세계적인 금융 네트워크에서 한곳의 고장이나 부실이 어떻게 전파 되어 나가는지에 대한 분석을 위해서는 무엇보다도 네트워크의 구조적 특성에 대한 이해가 필수적이다.

네트워크의 구조에 대한 연구는 1990년대 후반까지 유클리디안 격자(Euclidean Lattice)와 같은 Regular 네트워크나 Erdos 와 Renyi^[1]의 Random 네트워크를 사용하여 진행 되어 왔다. 현실세계에 존재하는 실제적인 네트워크들은 Regular도 아니고 Random도 아닌 두 개의 네트워크 사이에 존재한다는 인식이 있어 왔다. 그러나 현실의 대규모 망에 대한 상세한 위상(Topological) 정보 부재라든지 등의 이유로 Erdos 와 Renyi의 Random 네트워크가 거의 반세기 동안 네트워크의 구조 분석을 위해 유일하고 가장 광범위 하게 사용되어온 모델이다.

1990년 후반에 들어서 두 개의 중요한 네트워크 성질, Small-world 성질과^[2] Scale-free 성질이^[3] 발표되었다. Small-world 성질은 임의 두 노드 사이의 낮은 평균 최단 거리(Average Shortest Path Length: ASPL)와 높은 Clustering Coefficient를 갖는데 SNS 네트워크나 생물학적 네트워크 등의 위상적 특성을 잘 나타낸다^[4,5]. Scale-free 성질은 노드의 차수(노드에 연결된 링크의 수)가 Power-law 분포를 따르는 특성을 갖고 있다. 이는 평균보다 매우 큰 차수를 갖는 허브 노드들이 다수 존재 한다는 사실을 나타내는데 WWW, 인터넷의 AS, Metabolic 네트워크 등의 위상

적 특성을 잘 나타낸다^[6,9].

Small-world와 Scale-free 성질의 발견은 이러한 성질들을 갖는 Small-world와 Scale-free 네트워크 개발을 위한 네트워크 생성 모델(Network Generation Model)에 대한 연구를 촉발시켰다. Small-world 네트워크 모델은 Watts 와 Strogatz^[2]에 의해 처음으로 WS 모델이 제안 된 이후 여러 모델들이^[4,10,11] 후속으로 제안되었다. 그리고 Scale-free 네트워크 모델은 Barabasi 와 Albert^[3]에 의해 처음으로 BA 모델이 제안된 이후에 여러 다양한 모델들이^[12-16] 개발되어 왔다.

한편 Small-world 성질과 Scale-free 성질을 모두 갖는 네트워크 생성 모델도 개발되어 왔다. Small-world 네트워크 모델과 Scale-free 네트워크 모델 모두 ASPL이 네트워크 노드 크기에 Logarithm으로 증가한다는 공통점이 있지만 Small-world 네트워크 모델은 노드 차수가 power-law 분포를 따르지 않으며 Scale-free 네트워크 모델은 Clustering Coefficient가 높지 않다. Holme 등은^[17] power-law 분포를 따르며 동시에 Clustering Coefficient도 높은, 즉 Small-world 성질과 Scale-free 성질을 모두 갖는 네트워크 생성 모델을 개발 하였다. 이 모델을 토대로 Pasta 등은^[18]은 커뮤니티 구조를 갖는 네트워크 생성 모델을 개발 하였다. 이 모델들은 BA 모델을 기본으로 하고 Clustering Coefficient를 높이고자 Triad 형성 과정을 추가 하였다. Triad 형성 과정은 확률 P_t 로 수행되는데 P_t 를 적당히 변경함으로써 Clustering Coefficient를 조절해 나간다. 따라서 $P_t = 0$ 이면 이 모델은 바로 BA 모델이고 P_t 가 1에 가까울수록 Clustering Coefficient가 높아져 Small-world 성질을 나타나게 된다. 그런데 이 모델은 Power-law 성질을 나타내는 정도를 조정하는 방법이 없고 모델에 사용된 파라미터들의 변화에 따른 ASPL, Power-law 성질, Clustering Coefficient에 대한 영향력 분석등이 결여되어 있어서 모델의 유효성 및 검증이 충분하게 되어 있지 않다. 연구^[19]에서는 WS 모델을 기반으로 두 개의 입력 파라미터 K_0 와 α 를 적당하게 조절함으로써 Small-world 성질과 Scale-free 성질을 모두 갖는 네트워크 생성 모델을 개발 하였다. 그런데 이 생성 방법은 네트워크 성장 과정이 다소 인위적이어서 현실세계에 존재하는 네트워크 성장과정을 잘 표현하지 못한다는 문제점이 있다.

본 연구에서는 위에서 언급한 문제점, 즉 모델에 사용된 파라미터들이 Scale-free 성질과 Small-world 성질에 미치는 영향력 분석 결여와 비현실적인 네트워크 성장 과정 문제점을 보완하여 파라미터들을 적당

하게 조정함으로써 Small-world 성질과 Scale-free 성질을 나타내주는 새로운 네트워크 생성 모델을 개발하였다. 네트워크 모델은 각각 파라미터 K 와 P_{TF} 를 갖는 PA(Preferential Attachment)와 TF(Triad Formation) 두 단계로 구성되어 네트워크 성장 과정이 현실세계에 존재하는 네트워크 성장과정과 유사하도록 하였다. 아울러 본 연구에서는 두 개의 파라미터를 변화 시켜나가면서 실제적으로 네트워크를 생성한 후 ASPL, Clustering Coefficient, Power-law 성질 보유 여부 등을 조사함으로써 네트워크 생성 모델의 파라미터들이 Small-world 성질과 Scale-free 성질에 미치는 영향력을 분석하였다.

서론에 이어 2장에서는 본 연구에서 제안한 파라미터들을 적당하게 조정함으로써 Small-world 성질과 Scale-free 성질을 나타내주는 새로운 네트워크 생성 모델을 제안 하였다. 그리고 이 방법을 토대로 실제로 노드 1000개를 갖는 네트워크를 생성해 보았다. 3장에서는 2장의 방법을 따라 구축된 네트워크들의 분석을 위해 사용할 Small-world 성질과 Scale-free 성질을 나타내 주는 측도인 Clustering Coefficient, Average Shortest Path Length(ASPL)를 정의 하고 Power-law 성질을 설명하였다. 4장에서는 2장에서 제안한 네트워크 생성 모델의 유효성을 검증하기 위하여 모델의 파라미터들이 Small-world 성질과 Scale-free 성질에 미치는 영향력을 조사하였다. 이를 위하여 두 개의 파라미터를 변화시켜나가면서 네트워크들을 생성한 후 각 네트워크 별로 3장에서 설명한

분석 측도들을 유도 하였다. 마지막 결론은 5장에서 다루었다. 다음 (Fig. 1)에 본 논문의 구성을 나타냈다.

II. 새로운 네트워크 생성 모델

2.1 새로운 네트워크 생성 모델 제안

본 모델은 기본적으로 1) Preferential Attachment (PA)와 2) 연구^[17]에서 제안한 Triad Formation(TF) 두 개의 과정으로 이루어지는데 N 개의 노드를 갖는 네트워크를 생성한다. 본 연구에서는 Power-law 성질이 나타내는 정도를 조정하기 위하여 BA 모델의 PA(Preferential Attachment) 과정을 새롭게 제안하였으며 새롭게 추가되는 노드의 링크 개수도 상수가 아닌 지수분포를 가정하였다.

2.1.1 PA 과정

수정된 BA 모델은 초기, 성장, 그리고 수정된 BA 과정 세 개의 단계로 구성된다.

- 초기: 네트워크: n_0 노드와 m_0 개의 링크로 이루어진 초기 네트워크를 임의로 구축한다.
- 성장: 매 단계마다 m 개의 링크를 갖는 노드 i 가 추가 되는데 노드의 개수가 N 개가 될 때 까지 반복된다. 이때 m 은 평균 M 을 갖는 지수분포를 따른다.
- 수정된 PA 과정: 새롭게 추가된 노드 i 는 네트워크에 존재하는 기존의 노드 j 에 연결되는데 노드 j 에 연결될 확률은 다음 식으로 주어진다.

$$P_j = \frac{d_j + \frac{K}{|S|}}{\sum_{i \in S} d_i + K} \quad (1)$$

본 연구에서 새롭게 제안한 식 (1)은 BA 모델^[3]의 연결 확률 P_j 를 변형 한 것으로 분모에 K , 그리고 분자에 $K/|S|$ 를 새롭게 추가하였다. d_j 는 노드 j 의 차수 즉 링크의 개수를 나타내며 S 는 현재 네트워크에 존재하는 노드들의 집합이고 $|S|$ 는 노드들의 개수를 나타낸다.

위식에서 $K = 0$ 이면 기존의 BA 모델을 나타낸다. 노드 i 가 노드 j 에 연결될 확률은 노드 j 의 차수에 비례하기 때문에 링크 수가 많은 노드에 연결될 확률이 높고 이는 ‘Rich Node Becomes Richer’ 라는 Preferential Attachment 과정을 따른다.

K 가 증가하게 되면 차수가 많은 노드에 연결될 확률이 점차 줄어들어 $K = \infty$ 가 되면 노드 i 가 네트워크

Proposal of new tunable network generation model (Chapter 2)
1) Preferential attachment (parameter K) 2) Triad formation(parameter P_{TF})
Definition of measures for network analysis (Chapter 3)
1) Clustering coefficient 2) Average shortest path length 3) Power-law
Verification of network generation model (Chapter 4)
1) Network generation according to K and P_{TF} (number of node = 1,000) 2) Calculation network measures for generated networks 3) Investigation of small-world and scale-free properties based on network measures

그림 1. 논문 구성도
Fig. 1. Paper Organization

크내의 노드 j 에 연결될 확률은 노드 차수와 무관하게 $1/|S|$ 로 일정하게 된다. 따라서 이를 통해 생성된 망은 Random Network에 가깝다.

K 는 0과 ∞ 사이의 값을 갖는데 0에 가까운 값을 갖으면 Preferential Attachment 성향이 강하게 나타나 생성된 망에서 Power-law 성질이 나타나게 된다. 반면에 K 값이 커질수록 Preferential Attachment 성향이 없어져 생성된 망에서 Power-law 성질을 기대할 수 없다. 이에 대한 구체적인 실증적 검증은 4장 생성된 네트워크의 분석에서 수행하였다.

2.1.2 TF 과정

TF 과정^[7]은 노드 i 가 PA 과정을 통해 노드 j 에 연결되면 노드 i 를 노드 j 의 이웃과 연결시키는 과정이다. 이는 한 노드가 다른 임의의 노드보다 ‘친구의 친구’ 노드에 연결될 확률이 높다는 사실에 기인하는데 네트워크 생성 과정이 현실세계에 존재하는 네트워크 성장 과정과 유사하게 된다. 한편 TF 과정을 통하면 네트워크의 Clustering Coefficient가 커질 것으로 기대된다.

이 과정은 확률 P_{TF} 로 수행되는데 이 과정이 수행되지 못하면 BA 모델의 수정된 PA 과정을 반복한다. TF 과정은 다음 단계를 따라 수행한다.

- 수정된 PA 과정에서 새롭게 생성된 노드 i 가 기존 노드 j 와 연결되면 노드 i 와 임의로 선택된 j 의 이웃을 연결한다.
- 만약 연결할 j 의 이웃 노드들이 존재하지 않으면, 즉 j 의 이웃 노드들이 이미 모두 i 와 연결 되어 있으면

대신 PA 과정을 수행한다,

다음 Fig. 2는 본 장에서 제안한 네트워크 생성 방법의 단계를 순서도를 나타냈다.

2.2 네트워크 생성 실례

$N = 1,000$ 개의 노드를 갖는 네트워크를 본 연구에서 제안한 방법인 (Fig.2)를 사용하여 생성하였다. 이를 위해 사용한 입력 자료는 다음과 같다.

- 초기 네트워크는 $n_0 = 3, m_0 = 3$ 을 가정하였으며 이는 다음 (Fig. 3)과 같다.
 - 새롭게 추가되는 노드들이 갖는 링크의 수는 평균 $M = 3$ 를 갖는 지수 분포를 따른다.
 - 수정된 BA 모델의 PA 과정에서 $K = 0$ 으로 가정 한다
 - TF 과정 수행 확률인 $P_{TF} = 0.6$ 로 가정한다.
- 위의 입력 자료를 토대로 생성한 네트워크를 다음 (Fig. 4)에 나타냈다.

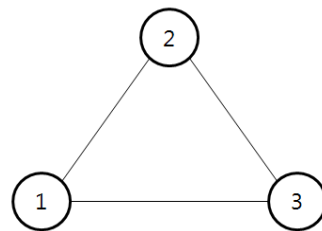


그림 3. 초기 네트워크
Fig. 3. Initial Network

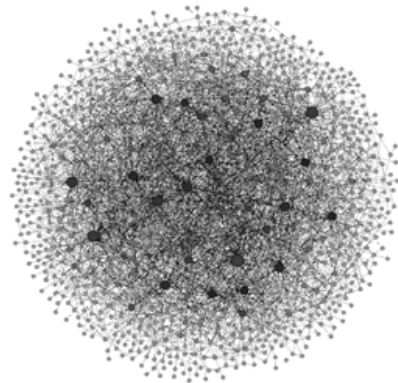


그림 4. 1,000 노드를 가진 네트워크 생성
Fig. 4. Generated Network with 1,000 Node

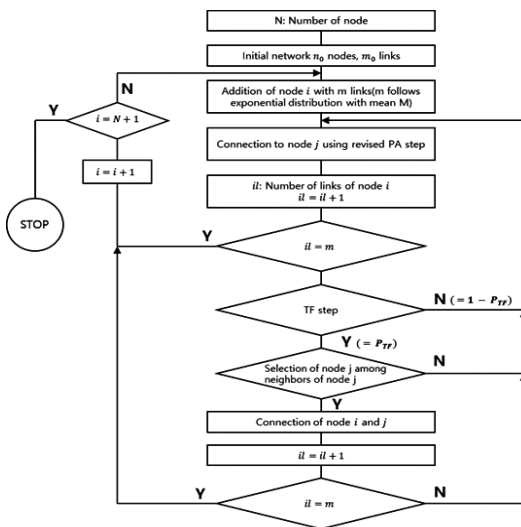


그림 2. 네트워크 발생 모델
Fig. 2. Network Generation Method

III. 네트워크 분석 척도

본 장에서는 네트워크의 Small-world 성질과 Scale-free 성질을 나타내주는 척도를 살펴보았다. 일

반적으로 Small-world 성질을 나타내주는 측도는 ASPL과 Clustering Coefficient이며 Scale-free 성질을 나타내 주는 측도로는 노드 차수가 Power-law 분포를 따르느냐 여부다.

3.1 Clustering Coefficient

SNS와 같은 네트워크에서는 노드들 사이의 연결이 랜덤하게 생성된 네트워크에 비해서 상대적으로 연결의 밀도가 높다고 알려져 있다. 노드들 사이에 연결의 밀도를 나타내는 척도로 Clustering Coefficient가 제안 되었다. Clustering Coefficient 계산은 노드의 세 개 쌍(Triplet)을 토대로 계산한다. 예로 노드 v 와 직접 연결된 노드의 개수를 k 개라고 할 때 모두 kC_2 개의 Triplet을 생성 할 수 있다. 이중 일부는 세 개가 모두 연결된 Closed Triplet이고 나머지는 연결이 안된 Open Triplet 이다. Clustering Coefficient는 네트워크에 존재하는 모든 Triplet의 개수중 Closed Triplet 개수의 비율로 정의 된다. 기본적으로 이 값은 0과 1 사이의 값을 갖는데 1에 가까울수록 특정 노드와 인접한 노드들끼리도 서로 연결될 확률이 높다는 것을 나타낸다. SNS와 같은 네트워크에서 보면 한 사람의 친구들은 그들 사이에 서로 친분이 있을 가능성이 매우 높다. 일반적으로 Small-world 성질을 나타내는 네트워크는 Clustering Coefficient가 높은 것으로 나타난다.

3.2 ASPL

평균 최단 거리를 나타내는 ASPL은 네트워크의 모든 노드들 간 최단 거리들의 평균으로 정의된다. L_{ij} 를 노드 i 와 j 사이의 최단거리를 나타낸다고 하면 ASPL은 다음과 같이 계산된다.

$$ASPL = \frac{\sum_{i > j} L_{ij}}{\binom{N}{2}}$$

ASPL은 위에서 논의한 Clustering Coefficient와 더불어 Small-world 성질을 나타내는데 사용된다. 일반적으로 Small-world Network는 동일한 노드 수를 갖는 다른 망에 비해 비교적 높은 Clustering Coefficient 값과 비교적 작은 ASPL 값을 갖는 것으로 나타났다.

3.3 Power-law

노드와 링크로 구성된 네트워크에서 노드의 차수,

즉 노드에 연결된 링크의 개수를 살펴봄으로서 네트워크의 구조적 특성을 파악 할 수 있다. 이때 노드의 차수 분포가 $y = x^{-\alpha}$ 형태의 Power-law 분포를 따르는 다수의 네트워크 들이 발견되었는데 이 분포의 특징은 지수 분포보다 훨씬 점진적으로 감소하기 때문에 관측치들이 매우 큰 값을 갖을 수 있고 이로 인해 평균값이나 대표값은 큰 의미를 갖지 못하는 매우 편향된(Skewed) 분포를 나타낸다는 것이다. 그래서 네트워크 노드의 차수가 평균 차수보다 매우 큰 허브(Hub)노드들이 존재하는데 그 차수가 특정 크기 (Scale)에 국한 되지 않기 때문에 이들 네트워크는 Scale-free 네트워크라고 불린다.

Faloutsos 등¹⁾은 1997년 말부터 1998년 말까지 BGP(Border Gateway Protocol) 라우팅 표로부터 획득한 데이터를 토대로 AS(Autonomous System) 수준의 Internet 망을 관측하였다. 그들은 노드(AS)의 차수가 Power-law 분포를 따르며 노드 차수 분포가 매우 편향되었다는 사실을 실험적으로 보였다. 실제로 노드의 85%가 평균 노드 차수보다 적은 노드 차수를 갖고 있으며 평균에 비해 매우 큰 노드 차수를 갖는 Hub 노드들이 다수 존재한다는 것도 확인 하였다.

노드 차수 분포가 Power-law를 따르는지 여부를 판단하는 방법은 노드 차수 분포를 Log-Log Scale로 나타낸 후 분포가 기울기 $-\alpha$ 를 갖는 직선식을 따르는지 보이는 것이다. 일반적으로 α 는 1과 3 사이의 값을 갖는다.

IV. 네트워크 분석

본 장에서는 2장에서 제안한 네트워크 생성 모델의 유효성을 검증하고자 한다. 이를 위해 본 연구에서 제안한 생성 모델의 두 파라미터인 K와 P_{TF} 가 Small-world와 Scale-free 성질에 미치는 영향력을 조사하였다.

먼저 네트워크 생성 모델의 순서도인 Fig. 2를 토대로 Java 프로그래밍을 이용하여 직접 네트워크를 구축한 후 3장에서 제안한 네트워크 분석 측도를 유도 하였다. 네트워크 크기는 $N = 1,000$ 으로 설정하였으며 새롭게 추가되는 노드가 갖는 링크수는 평균 10의 지수분포를 따른다고 가정하였다.

본 연구에서 제안한 네트워크 생성 모델은 두 개의 파라미터 K와 P_{TF} 로 구성된다. K는 Power-law 성질을 나타내는 파라미터로 생성된 네트워크의 Scale-free 성질에 영향을 미친다. P_{TF} 는 Clustering Coefficient 크기를 결정하는 파라미터로 구축된 네트

워크의 Small-world 성질에 영향을 미친다.

먼저 4.1 장에서 K 가 구축한 네트워크에 미치는 영향력을 분석하기 위하여 $K = 0, 10, 100, 1,000, 10,000$ 로 변화 시켜 나가면서 네트워크를 구축한 후 ASPL과 Clustering Coefficient 그리고 Power-law 성질을 조사하였다 또한 P_{TF} 와의 상호작용(Interaction)을 조사하기 위하여 위의 과정을 $P_{TF} = 0.1, 0.3, 1$ 세 개의 확률값 하에서 각각 수행하였다.

그리고 4.2 장에서 P_{TF} 가 구축한 네트워크에 미치는 영향력을 분석하기 위하여 $P_{TF} = 0, 0.1, 0.3, 0.5, 1$ 로 변화 시켜 나가면서 네트워크를 구축한 후 ASPL과 Clustering Coefficient 그리고 Power-law 성질을 조사하였다 또한 K 와의 상호작용을 조사하기 위하여 위의 과정을 $K = 0.1, 1,000, 10,000$ 세 개의 값 하에서 각각 수행하였다.

본 연구에서 유도한 모든 네트워크 분석 측도는 모든 조합(K 의 영향력 분석을 위한 15개 조합과 P_{TF} 가 미치는 영향력을 분석하기 위한 15개 조합)에 대해 각각 10개의 네트워크를 생성한 후 이들 10개 네트워크들의 측도를 평균하여 사용하였다.

4.1 K 가 구축한 네트워크에 미치는 영향력 분석

K 가 네트워크에 미치는 영향력을 ASPL, Clustering Coefficient, 그리고 Power-law 성질 보유 여부 측면에서 살펴보았다.

4.1.1 ASPL

$P_{TF} = 0.1, 0.3, 1$ 세 개의 확률값 하에서 $K = 0, 10, 100, 1,000, 10,000$ 로 변화 시켜 나가면서 네트워크를 구축한 후 ASPL를 구하여 이를 <Table 1>에 정리하였다.

위 <Table 1>에서 볼 수 있듯이 P_{TF} 에 무관하게 K 가 증가함에 따라 ASPL은 아주 미세하게 증가한다, K 가 작은 값일 경우는 PA로 인해 Hub 노드가 생겨난다. Hub 노드의 존재는 임의의 두 노드사이의 거리를 단축시켜 ASPL 값을 감소한다. 반면에 K 값이 커지면

표 1. K 에 따른 ASPL
Table 1. ASPL according to K

	$K= 0$	$K= 10$	$K= 100$	$K=1,000$	$K=10,000$
$P_{TF} = 0.1$	2.64	2.67	2.69	2.70	2.70
$P_{TF} = 0.3$	2.64	2.66	2.70	2.71	2.70
$P_{TF} = 0.5$	2.63	2.68	2.70	2.71	2.71

면 Hub 노드는 존재 하지 않는다. 그런데 새롭게 추가되는 노드는 임의로 선정된 기존의 노드와 연결되는데 이것도 일정 부분 임의의 두 노드들 사이의 거리를 단축시키는 효과가 있기 때문에 Hub 노드가 존재하지 않는다고 해서 ASPL이 급격하게 증가하지 않게 된다.

K 값에 따른 ASPL 변화는 P_{TF} 값에 따라 변화가 없기 때문에 K 와 P_{TF} 사이의 상호 작용은 존재하지 않는다.

4.1.2 Clustering Coefficient

$P_{TF} = 0.1, 0.3, 1$ 세 개의 확률값 하에서 $K = 0, 10, 100, 1,000, 10,000$ 로 변화 시켜 나가면서 네트워크를 구축한 후 Clustering Coefficient를 구하여 이를 <Table 2>에 정리하였다.

어떤 P_{TF} 값 하에서도 $K = 0$ 일 때 Clustering Coefficient가 가장 크고 K 가 증가하면 Clustering Coefficient 값은 감소해 나간다. K 가 작은 값일 경우는 PA로 인해 Hub 노드가 생겨난다. 이때 임의의 두 노드가 Hub 노드와 연결될 가능성이 높고 두 노드끼리 연결만 되면 Closed Triplet이 생성되어 Clustering Coefficient 값은 증가한다. 반면에 K 값이 커지면 Hub 노드는 존재 하지 않는다. 그런데 새롭게 추가되는 노드는 임의로 선정된 기존의 노드와 연결되는데 이 두 노드가 다른 노드와 함께 Closed Triplet을 형성할 가능성은 낮아 Clustering Coefficient 값은 낮아지게 된다.

표 2. K 에 따른 군집화 계수
Table 2. Clustering Coefficient according to K

	$K= 0$	$K= 10$	$K= 100$	$K=1,000$	$K=10,000$
$P_{TF} = 0.1$	0.17	0.12	0.09	0.09	0.08
$P_{TF} = 0.3$	0.19	0.15	0.16	0.11	0.11
$P_{TF} = 0.5$	0.21	0.16	0.14	0.14	0.14

4.1.3 Power-law

$P_{TF} = 0.1, 0.3, 1$ 세 개의 확률값 하에서 $K = 0, 10, 100, 1,000, 10,000$ 로 변화 시켜 나가면서 네트워크를 구축한 후 노드 차수분포가 Power-law 분포를 따르는지 여부를 검증하여 결과를 <Table 3>에 정리하였다.

P_{TF} 값에 무관하게 $K = 0$ 과 10에서는 Power-law 분포를 따르는 것으로 나타났으며 $K = 100$ 부터는

표 3. K 에 따른 Power-law 성질과 지수
Table 3. Power-law Property according to K and Exponent

	K=0	K=10	K=100	K=1,000	K=0,000
$P_{TF} = 0.1$	0, -1.2	0, -1.2	1, -1.0	1, -1.0	1, -1.0
$P_{TF} = 0.3$	0, -1.2	0, -1.2	1, -1.1	1, -1.0	1, -1.0
$P_{TF} = 0.5$	0, -1.2	0, -1.2	1, -1.1	1, -1.1	1, -1.0

(0: Power-law 분포를 따름, 1: Power-law 분포를 따르지 않음)

Power-law 분포를 따르지 않는 것으로 나타났다. 이는 K가 작을 때는 PA로 인해 평균 노드 차수에 비해 노드 차수가 매우 큰 노드들이 존재하고 이로 인해 Power-law 성질이 나타난다. K가 커지면 새롭게 추가 되는 노드와 연결될 노드가 임의로 선정되기 때문에 노드 차수가 매우 큰 노드들이 존재하지 않게 되어 Power-law 성질이 나타나지 않는다.

$P_{TF} = 0.1$ 하에서 Power-law 분포를 따르는 $K = 0$ 과 따르지 않는 $K = 1,000$ 과 $10,000$ 의 노드 차수 분포를 Log-Log scale로 (Fig. 4)에 나타냈다. (Fig. 5)에서 볼 수 있듯이 Power-law 분포를 따르는 $K = 0$ 경우는 차수가 작은 일부 노드를 제외하면 기울기가 -1.185인 직선 형태를 따르는 것을 볼 수 있다. 반면에 Power-law 분포를 따르지 않는 $K = 1,000$ 과 $10,000$ 의 경우는 노드 차수 분포가 직선 형태와는 상당한 거리가 있다.

4.2 P_{TF} 가 구축한 네트워크에 미치는 영향력 분석
 P_{TF} 가 네트워크에 미치는 영향력을 ASPL,

Clustering Coefficient, 그리고 Power-law 성질 보유 여부 측면에서 살펴보았다.

4.2.1 ASPL

$K = 0, 1,000, 10,000$ 세 개의 값 하에서 $P_{TF} = 0, 0.1, 0.3, 0.5, 1$ 로 변화 시켜 나가면서 네트워크를 구축한 후 ASPL을 구하여 이를 <Table 4>에 정리하였다.

<Table 4>에서 보듯이 K와 무관하게 P_{TF} 에 따른 ASPL 변화는 미미해 보인다. $P_{TF} = 0$ 이면 TF 과정 대신 PA 과정을 수행하고 $P_{TF} = 1$ 이면 PA 과정 대신 TF 과정을 수행하는데 두 개의 과정이 ASPL에 미치는 영향력은 차이가 매우 미미하다.

표 4. PTF 에 따른 ASPL
Table 4. ASPL according to PTF

	$P_{TF} = 0$	$P_{TF}=0.1$	$P_{TF}=0.3$	$P_{TF}=0.5$	$P_{TF}=1$
K=0	2.64	2.64	2.64	2.63	2.63
K=1,000	2.69	2.70	2.71	2.71	2.71
K=10,000	2.69	2.70	2.70	2.70	2.71

4.2.2 군집화 계수

$K = 0, 1,000, 10,000$ 세 개의 값 하에서 $P_{TF} = 0, 0.1, 0.3, 0.5, 1$ 로 변화 시켜 나가면서 네트워크를 구축한 후 Clustering Coefficient를 구하여 이를 <Table 5>에 정리하였다.

<Table 5>에서 보듯이 K와 무관하게 P_{TF} 증가에 따라 Clustering Coefficient는 커지는 것으로 나타났다. P_{TF} 가 증가하면 Closed Triplet을 만들어 나가는 TF 과정을 수행 할 확률이 높아지므로 Clustering Coefficient가 커지게 된다.

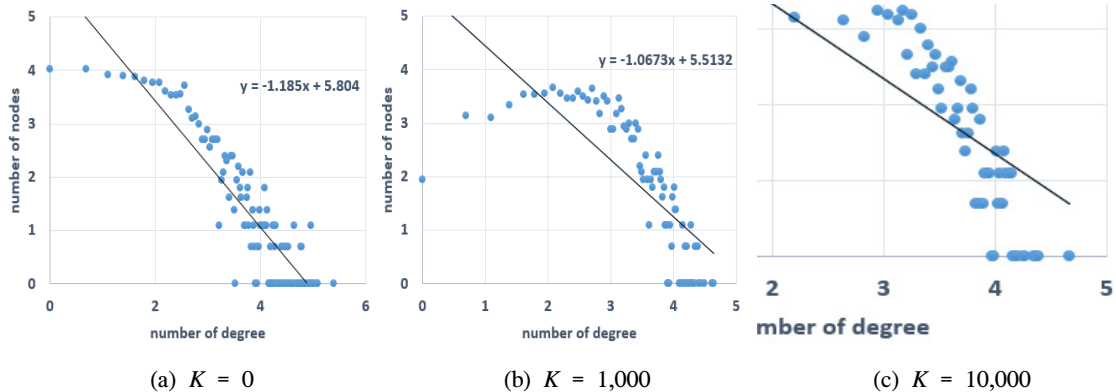


그림 5. $P_{TF} = 0.1$ 하에서 노드 차수 분포
Fig. 5. Node Degree Distribution(Log-Log scale) under $P_{TF} = 0.1$

표 5. P_{TF} 에 따른 군집화 계수
Table 5. Clustering Coefficient according to P_{TF}

	$P_{TF}=0$	$P_{TF}=0.1$	$P_{TF}=0.3$	$P_{TF}=0.5$	$P_{TF}=1$
$K=0$	0.12	0.17	0.19	0.20	0.21
$K=1,000$	0.04	0.09	0.11	0.12	0.14
$K=10,000$	0.04	0.08	0.11	0.12	0.14

4.2.3 Power-law

$K = 0, 1,000, 10,000$ 세 개의 값 하에서 $P_{TF} = 0, 0.1, 0.3, 0.5, 1$ 로 변화 시켜 나가면서 네트워크를 구축한 후 노드 차수가 Power-law 분포를 따르는지 여부를 검증하여 결과를 <Table 6>에 정리하였다.

표 6. P_{TF} 에 따른 Power-law 성질과 지수
Table 6. Power-law Property according to P_{TF} and Exponent

	$P_{TF}=0$	$P_{TF}=0.1$	$P_{TF}=0.3$	$P_{TF}=0.5$	$P_{TF}=1$
$K=0$	0, -1.2	0, -1.2	0, -1.2	0, -1.2	0, -1.2
$K=1,000$	1, -0.9	1, -1.0	1, -1.0	1, -1.0	1, -1.0
$K=10,000$	1, -0.9	1, -1.0	1, -1.0	1, -1.0	1, -1.0

(0: Power-law 분포를 따름, 1: Power-law 분포를 따르지 않음)

$K = 0$ 일 경우 P_{TF} 와는 무관하게 모두 Power-law 분포를 따르는 것으로 나타났다. $K = 0$ 하에서 P_{TF} 가 0.1, 0.3, 그리고 1인 경우 노드 차수 분포를 Log-Log scale로 (Fig. 5)에 나타냈다. (Fig. 6)에서 볼 수 있듯이 세 경우 모두 차수가 작은 일부 노드들을 제외하면

기울기가 대략 -1.18인 직선 형태를 따르는 것을 볼 수 있다. 이는 K 가 작을 때는 PA로 인해 평균 노드 차수에 비해 노드 차수가 매우 큰 노드들이 존재하고 이로 인해 Power-law 성질이 나타나기 때문이다.

K 가 1,000이나 10,000에서는 P_{TF} 와는 무관하게 모두 Power-law 분포를 따르지 않는다. K 가 커지면 새롭게 추가되는 노드와 연결될 노드가 임의로 선정되기 때문에 노드 차수가 매우 큰 노드들이 존재하지 않게 되어 Power-law 성질이 나타나지 않는다.

4.3 분석 결과

위의 4.1과 4.2의 분석결과를 바탕으로 본 연구에서 제안한 네트워크 생성 모델의 두 파라미터인 K 와 P_{TF} 가 Small-world와 Scale-free 성질에 미치는 영향력을 다음에 정리하였다.

4.3.1 파라미터 K

2.1장의 1) PA 과정에서 예측한대로 K 가 작은 값을 갖을 경우에는 생성된 네트워크가 Scale-free 성질을 나타내고 커지면 Scale-free 성질이 나타나지 않는다. 이는 본 연구에서 제안한 네트워크 생성 모델이 파라미터 K 를 적절하게 조정함으로써 Scale-free 네트워크를 만들 수 있다는 것을 나타낸다. 아울러 K 는 Hub 노드 생성과 밀접하게 연관되어 있어 Clustering Coefficient에는 영향을 미치지 않지만 ASPL과는 무관한 것으로 나타났다.

4.3.2 파라미터 P_{TF}

2.1장의 2) Triad Formation 과정에서 예측한대로 P_{TF} 가 커지면 Clustering Coefficient가 증가하는 것으로 나타나 P_{TF} 는 Small-world 성질을 여는 정도 표현

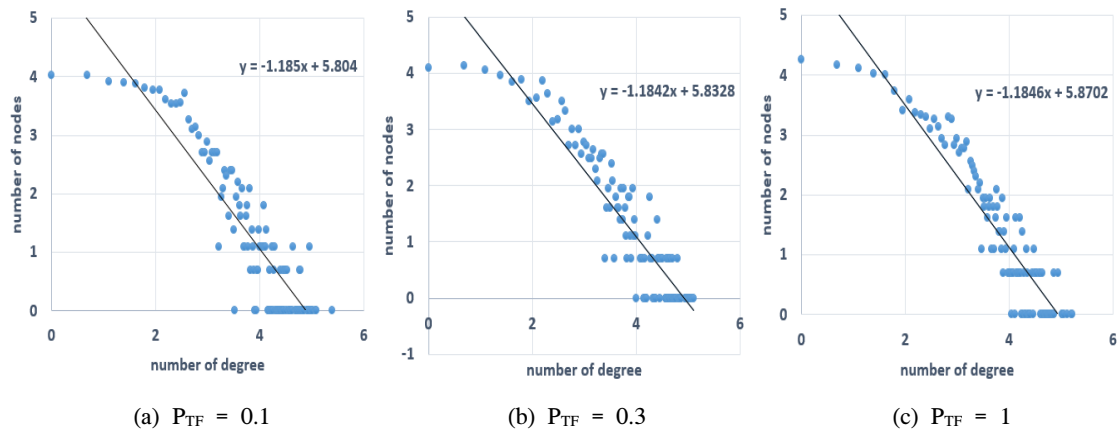


그림 6. $K = 0$ 하에서 노드 차수 분포
Fig. 6. Node Degree Distribution(Log-Log scale) under $K = 0$

할 수 있다는 것을 나타냈다. 그러나 P_{TF} 는 Power-law 성질과는 무관해 Scale-free와는 연관이 없는 것으로 나타났다.

V. 결론

본 연구에서는 파라미터들을 조정함으로써 Small-world 성질과 Scale-free 성질을 나타내 줄 수 있는 네트워크 생성 모델을 개발하였다, 본 연구에서 제안한 모델은 K 와 P_{TF} 두 개의 파라미터를 갖으며 네트워크가 점진적으로 성장해 나가는 PA 과정과 노드들의 밀집도를 증가시키는 TF 과정 두 개의 단계로 구성된다.

K 와 P_{TF} 의 다양한 조합에 따라 네트워크들을 생성한 후 그들의 ASPL, Clustering Coefficient, Power-law 성질 보유 여부 등을 조사함으로써 두 파라미터들이 Small-world 성질과 Scale-free 성질에 미치는 영향을 분석하였다.

PA 과정에서 사용한 파라미터 K 는 Power-law 성질 보유 여부를 결정 하는 것으로 나타났다. 아울러 K 는 Hub 노드의 존재 유무를 결정하게 되는데 이로 인해 K 값에 따라 Clustering Coefficient 값이 변하는 것을 확인하였다. 하지만 K 값의 변화에 따른 ASPL의 변화는 매우 미미하였다. TF 과정에서 사용한 파라미터 P_{TF} 는 Clustering Coefficient에는 직접적으로 영향을 미치지만 P_{TF} 에 따른 Power-law 성질 보유 여부나 ASPL에는 변화가 없었다.

이제 이 두개의 파라미터를 적당하게 조정함으로써 Small-world 네트워크 또는 Scale-free 네트워크 혹은 두 개의 성질을 모두 갖는 네트워크를 생성 할 수 있다. 예로 Scale-free 네트워크를 생성하기 위해서는 K 를 작게 설정하여 Hub 노드 생성을 촉진시키면 되고 Small-world 네트워크를 생성하기 위해서는 P_{TF} 값과 K 값을 작게 유지하여 Clustering Coefficient를 증가시키고 ASPL을 작게 만들면 된다.

보다 정교한, 특정한 값의 ASPL과 Clustering Coefficient를 갖는 Customized 된 네트워크 생성을 위해서는 파라미터 값들과 네트워크 측도들 사이의 수리적인 관계를 명확히 밝힐 필요가 있는데 이에 대한 연구는 추후 연구 과제로 남긴다.

References

[1] P. Erdos and A. Renyl, "On random graph," *Mathematica*, vol. 6, pp. 290-297, 1959.
 [2] D. J. Watts and S. H. Strogatz, "Collective

dynamics of small-world," *Nature*, vol. 393, pp. 440-442, 1998.

[3] A. Barabasi and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, pp. 509-512, 1999.
 [4] X. F. Wang and G. Chen, "Complex network; small-world, scale-free and beyond," *IEEE Cir. and Syst. Mag.*, pp. 6-20, First Quarter 2003.
 [5] S. Wang, et al., "Dynamic analysis of biochemical network using complex network method," *Thermal Sci.*, vol. 19, no. 4, pp. 1249-1253, 2015.
 [6] M. Faloutsos, et al., "On power-law relationship of the Internet topology," *ACMSIGCOMM*, pp. 251-262, 1999.
 [7] D. I. Oh and K. W. Lee, "Study on the characteristics of the Korea Internet AS-level topology using node degree and node connectivity metrics," *J. KICS*, vol. 38, no. 6, pp. 417-426, Jun. 2013.
 [8] H. Lee, "The Power of power law: a view from VOD service," *J. KICS*, vol. 38, no. 8, pp. 607-614, Aug. 2013.
 [9] D. I. Oh and K. W. Lee, "Analysis of korea internet AS-level topology," *J. KICS*, vol. 37, no. 10, pp. 901-911, Oct. 2012.
 [10] R. D. Caux, et al., "Dynamic, small-world, social network generation through local agent interactions," *Complexity*, vol. 19, no. 6, pp. 44-53, 2014.
 [11] D. Barmpoutis and R. M. Murray, "Networks with the smallest average distance and the largest average clustering," arXiv:1007.4031 [q-bio.MN], 2010.
 [12] C. Dangalchev, "Generation models for scale-free networks," *Physica A*, vol. 338, p. 659, 2004.
 [13] P. L. Krapivsky, S. Redner, and F. Leyvraz, "Connectivity of growing random networks," *Phys. Rev. Lett.*, vol. 85, no. 4629, 2000.
 [14] E. Ravasz and Barabasi, "Hierarchical organization in complex networks," *Phys. Rev. E*, vol. 67, no. 026112, 2003.
 [15] G. Caldarelli, et al., "Scale-free networks from

- varying vertex intrinsic fitness,” *Phys. Rev. Lett.*, vol. 89, no. 258702, 2002.
- [16] D. Krioukov, “Hyperbolic geometry of complex networks,” *Phy. Rev. E*, vol. 82, no. 036106, 2010.
- [17] P. Holme and B. J. Kim, “Growing scale-free networks with tunable clustering,” *Phy. Rev. E*, vol. 65, no. 026107, 2002.
- [18] M. Q. Pasta and Z. Jan, “Tunable and growing network generation model with community structure,” arXiv:1310.8396v2[cs.SI]Oct. 2014.
- [19] K. W. Lee, J. H. Lee, and H. Z. Choe, “Generalized network generation for small-world network and scale-free network,” *J. KICS*, vol. 41, no. 7, pp. 754-764, 2016.

이 강 원 (Kang-Won Lee)



1980년 : 서울대학교 공과대학
산업공학(공학사)
1982년 : 서울대학교 대학원(공
학석사)
1985년 : Kansas State Univ.
(공학박사)
2017년~현재 : 서울과학기술대

학교 산업정보시스템공학과 교수
<관심 분야> 정보통신망 분석, 이동망 성능 분석

엄 희 관 (Hee-Kwan Uhm)



2017년 : 서울과학기술대학교
산업정보시스템 공학과(학사)
<관심 분야> 데이터 마이닝,
컴퓨터 시뮬레이션

최 혜 진 (Hye-Jin Choe)



2017년 : 서울과학기술대학교
산업정보시스템 공학과(학사)
<관심 분야> 데이터 마이닝,
컴퓨터 시뮬레이션