

Social Network Big Data 분석 기법과 응용

최병진, 황용근, 정교민
서울대학교

요약

최근 정보통신 기술의 발전과 더불어 급성장 중인 소셜 네트워크는 개인 혹은 집단간의 실제 사회적 관계를 네트워크 구조로 반영하고 있다. 소셜 네트워크들의 구조를 보다 정확하게 이해하고 소셜 네트워크 내에서 정보가 전파되는 패턴을 파악하기 위해 소셜 네트워크를 수학적으로 모델링하고, 이를 응용하여 소셜 네트워크 빅 데이터를 분석하는 다양한 연구가 이루어지고 있다. 본고에서는 소셜 네트워크의 구조 분석과 정보 확산 패턴 파악에 관한 주요 연구 사례들을 소개하고, 특히 소셜 빅 데이터 분석과 관련된 연구 주제 및 응용 사례들을 살펴보고자 한다.

I. 서론

페이스북 (Facebook), 트위터 (Twitter), 인스타그램 (Instagram), 밴드 (BAND) 등과 같은 온라인 소셜 미디어 (Online Social Media) 는 인간관계의 형성과 더불어 유용한 정보를 공유하는 새로운 채널로 자리잡았다. 정보 통신기술의 발전과 더불어 급부상한 소셜 네트워크 서비스 (Social Network Service) 는 특히 스마트폰과 같은 모바일 기기의 보급으로 인해 언제 어디서나 소셜 미디어를 이용할 수 있는 환경이 되면서 현대인의 일상생활에서 차지하는 비중이 높아졌다. 이러한 소셜 네트워크는 실제 사회의 복잡하고 상호 의존적인 개인 혹은 집단 간 관계를 네트워크 구조로 반영하는 성질을 가진다. 따라서 소셜 네트워크를 활용한 다양한 사회 문제를 해결 하기 위해서는 소셜 네트워크의 구조를 네트워크 관점에서 분석하고, 네트워크상에서 정보가 전파되는 패턴을 파악하는 것이 필요하다고 할 수 있겠다.

구조적 측면에서 소셜 네트워크들은 매우 다양한 형태를 보이나, 한편으로는 다음과 같은 공통적인 성질을 발견할 수 있다. 1) 서로 유사한 구성원들 간에 보다 긴밀하게 연결되어 있는 공동체 구조[4] / 2) 연결이 많은 노드(node)인 허브(hub)들의 존

재로 인한 불균등한 위상 구조 / 3) 임의의 두 구성원이 매우 짧은 경로로 연결되어 있는 작은 세상 (small world) 구조[21] 등이 있다. 이러한 구조적 특성과 함께 네트워크의 변화 과정을 동시에 파악하는 것이 소셜 네트워크 분석에 있어 매우 중요하다[22]. 시간의 흐름에 따라 급격하게 변화 및 진화하는 소셜 네트워크의 특성상 각 구성 요소만으로는 소셜 네트워크의 본질을 설명하기에 충분하지 않기 때문이다. 이와 같이 소셜 네트워크의 구조가 파악되고 나면 네트워크 안에서의 정보 흐름도 분석 및 예측이 가능해진다.

현재 초기 단계에 머물고 있는 소셜 네트워크 분석 (Social Network Analysis) 연구는 소셜 네트워크가 지니는 사회적 중요성을 고려했을 때 높은 잠재적 가치를 지닌 연구 분야이다. 또한 융합성을 띤 연구 분야이며 수학, 물리, 생물, 전산, 사회학, 경제 등의 분야에서 다각도로 연구가 이루어지고 있다. 여기에 최근 부각되고 있는 빅 데이터 분석 기법을 도입하여 규모가 큰 소셜 네트워크를 직접 분석하는 연구도 전 세계적으로 시도되고 있다.

본고에서는 우선 소셜 네트워크의 정보 확산 패턴과 구조를 분석하는 문제들의 여러 가지 접근 방법들을 설명할 것이다. 그리고 각 방법에 대한 국내·외 연구 동향을 점검하면서 빅 데이터 분석 기법과 관련된 연구들을 중점적으로 소개할 것이다.

II. 소셜 네트워크의 정보 확산 패턴 및 구조 분석 문제의 접근 방법

소셜 네트워크의 정보 확산 패턴 및 구조 분석 문제는 크게 두 가지의 방법으로 접근되어왔다. 한 가지 방법은 확산 패턴 및 구조의 근본적 원리에 대한 이해 및 수학적 모델링을 통한 변화 양상을 연구하는 것이고 다른 방법은 빅 데이터 분석 기법을 도입하여 소셜 네트워크 데이터를 직접 분석하여 얻은 결과를 통해 실제 소셜 네트워크에 대한 과학적 사실을 연구하는 것이다.

또한 소셜 네트워크에 대한 연구는 크게 1) 정보 확산 패턴

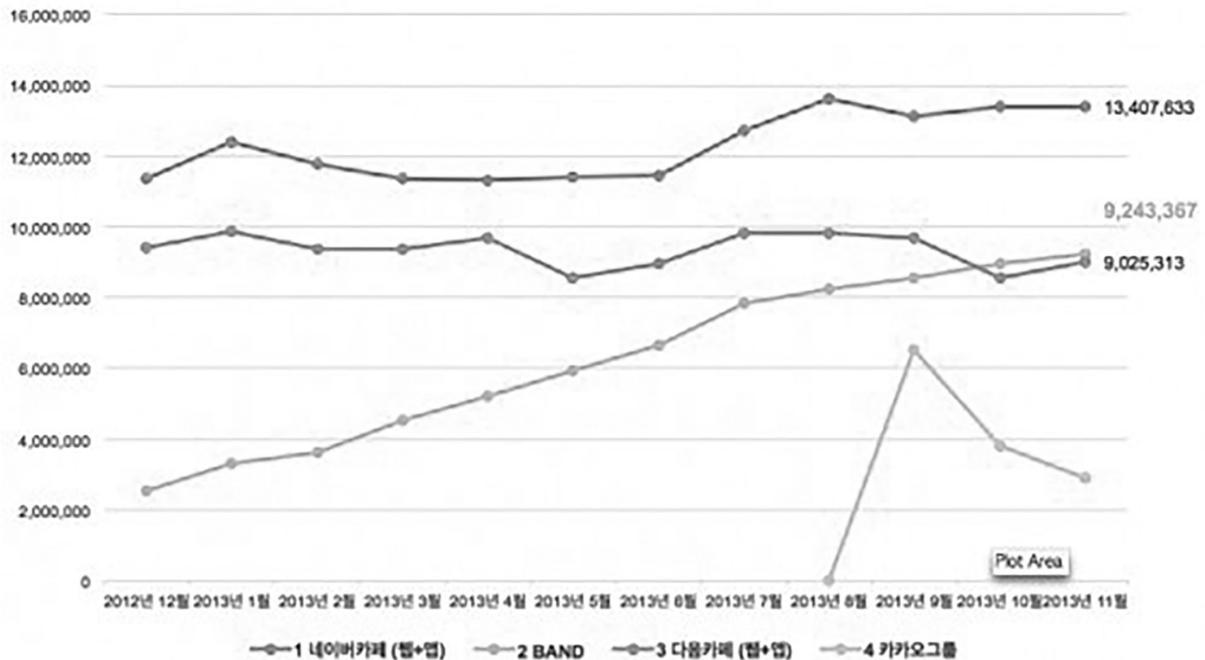


그림 1. 2012년~2013년 국내 주요 소셜 네트워크 커뮤니티 방문자 수 변화 [10]

분석, 2) 계층 및 집단구조 분석 그리고 3) 소셜 네트워크 진화 과정 예측의 세 가지 주제들로 나눌 수 있다[16]. 이들 연구 주제와 연구 사례들을 소개하자면 다음과 같다.

II-1. 정보 확산 패턴 분석

실제 사회에서는 웹 서비스뿐 만 아니라 여러 분야에서 다양한 서비스가 경쟁을 하는 모습을 볼 수 있는데, 그 가운데 어떤 것은 사용자를 확보하는데 성공하여 지속적인 성장을 이루는 반면, 상당수의 서비스들은 이에 실패하여 중국에는 서비스를 종료 것을 자주 볼 수 있다. 그런데 사업적으로 성공한 서비스와 그렇지 못한 서비스를 막상 살펴보면 제공하는 서비스 내용이 서로 비슷하고 기술적으로도 큰 차이가 없는 경우가 상당히 많다. 그렇다면 과연 무엇이 이들의 성패를 결정하는 데 중요하게 작용하는가에 대한 의문이 자연히 생길 수 있고, 이를 파악하는 것은 서비스 사업가에게 있어 매우 중요한 일이다.

소셜 네트워크의 경우를 살펴보면 많은 경우에 있어 서비스의 초기 사용자 수가 어떤 임계점 (tipping point) 에 도달했는지, 혹은 그렇지 못했는지의 여부가 향후 해당 서비스의 성패와 밀접한 연관을 가지는 것을 볼 수 있다. 이러한 현상은 정보의 흐름에 대해서도 비슷하게 관찰되는데, 정보 확산 초기에 놓여있는 상태에 따라 정보 확산이 전체 네트워크로 이루어지거나 반대로 그렇지 못하는 상전이 (phase transition) 현상을 볼 수

있다[1][4]. 이에 네트워크 정보 확산을 상전이 현상을 기반으로 수학적 모델링을 한 뒤, 해당 수학적 모델을 바탕으로 소셜 네트워크 데이터를 분석하여 정보 확산의 임계점 및 광범위 확산의 발생 확률 및 조건, 그리고 정보가 최종적으로 확산되는 규모를 추정하고자 하는 연구가 이루어지고 있다.

먼저 네트워크 정보 확산 과정에 대한 수학적 모델링 연구를 살펴보고자 한다. 기존에 널리 알려진 연구 결과 가운데 가장 대표적인 두 가지 모델을 살펴보자면, 먼저 Granovetter[13] 등에 의해 제안된 Linear Threshold(LT) 모델을 들 수 있다. LT 모델에서는 각 노드마다 정해진 정보 확산에 대한 어떤 문턱값 (threshold)에 의해 해당 노드가 다른 이웃 노드로부터 새로운 정보를 받아들일 것인지의 여부가 결정된다고 정의하고 있다. 다음으로 Goldenberg, Libai, and Muller[12]가 제시한 Independent Cascade(IC) 모델이 있는데 이 모델에서는 정보가 각 링크를 통해 독립적으로 확산되는 패턴을 가진다고 보고 이에 독립적인 확산 확률을 정의한다. 여기에 Kempe et al. [20]은 LT, IC 모델을 모두 포함할 수 있는 일반화된 threshold model을 제시하였고 이는 상전이 현상 연구가 발전하는 데 있어서 또 하나의 중요한 기점이 되었다. 하지만 한편으로 이들 모델을 기반으로 한 연구들은 주로 나무에 가까운 (locally tree-like) 그래프에서 수행[15]되어왔으며 대부분 정보 확산 확률이 고정된 경우[11]만을 고려했다는 한계를 가지고 있다.

정보 확산 현상을 이해하는 것과 더불어 정보 흐름 패턴 분석에 있어서 중요한 또 하나의 문제는 영향력 최적화 문제이다. 이는 기본적으로 정보 흐름의 영향력을 최대화 (influence maximization), 즉 소셜 네트워크 상에서 정보가 어떤 조건일 때 가장 잘 전파될 수 있는가를 알아보는 문제로 정의할 수 있는데, 일반적으로 영향력 최적화는 풀기 매우 어려운 문제 (NP-hard) 로 표현된다. 이에 대해 실제로 Kempe et al.[20]은 정보 확산의 영향력을 나타내는 영향력 함수가 앞서 설명한 LT, IC 모델에서 모두 서브모듈러 (submodular) 함수로 표현 가능하며, 이를 최대화 하는 문제는 NP-hard임을 증명하였다. 따라서 영향력 최대화 알고리즘 개발에 있어서 규모가 큰 네트워크를 표현한 소셜 네트워크 빅 데이터를 빠른 시간에 처리 가능하게 (scalable) 알고리즘을 설계하는 것이 주요한 연구 과제로 대두되고 있다. 초기에 Kempe et al.[20], Leskovec et al.[23] 등은 서브모듈러 함수의 특성을 활용, 영향력 최적화 문제의 최적해에 대한 근사치를 보장하는 탐욕적 알고리즘 (greedy algorithm) 을 제안했는데, 이 알고리즘을 통해 얻은 최적해의 정확도는 비교적 우수하나 알고리즘 자체가 몬테-카를로 시뮬레이션 (Monte-Carlo Simulation) 에 의존하기 때문에 최적값을 얻기 위해서는 그만큼 알고리즘 수행 시간이 과도하게 길어지게 되어 실세계 네트워크에 대한 빅 데이터에 대해서는 이들 알고리즘의 실용성이 저해되는 한계가 있다. 그래서 이후 실행 속도를 높일 수 있는 휴리스틱 (heuristic) 알고리즘들이 제안되고 있으며, 대표적인 알고리즘으로는 Chen et al.[7]이 제안한 PMIA(Prefix-excluding Maximum Influence Arborescence)를 들 수 있다. 비록 이 알고리즘은 각 노드끼리 긴밀히 연결된 네트워크에서는 알고리즘 수행 시간이 상당히 길어지는 등 네트워크 구조에 다소 민감하며 메모리 사용량이 매우 높다는 단점을 가지고 있지만 현재 알려진 연구 결과 중 비교적 효율적인 휴리스틱 알고리즘에 속한다. 이외에도 소셜 네트워크 빅 데이터를 처리할 수 있도록 영향력 최대화 문제 알고리즘의 성능을 높이는 연구가 지금도 지속적으로 진행되고 있다.

영향력 최대화 문제를 실세계 소셜 네트워크 빅 데이터에 대해 보다 효과적으로 수행할 수 있게 된다면 상품의 마케팅 전략 수립과 같은 정보 확산이 관여되는 다양한 분야에 적용되어 관련 산업의 활성화와 같은 긍정적인 효과를 기대할 수 있을 것이다.

II-2. 계층 및 집단 구조 분석

소셜 네트워크를 네트워크 연결 구조의 관점에서 분석하는 것도 소셜 네트워크 분석 연구의 핵심 중 하나라고 할 수 있다. 네

트워크의 구조는 크게 계층 구조 (hierarchical structure) 와 집단 구조 (community structure) 를 보는 두 가지의 관점으로 접근이 가능하다.

그 가운데 계층 구조의 분석 연구는 그 중요성에 비하여 다른 소셜 네트워크 연구 분야 대비 연구 성과가 다소 부족한 편이다. 기존 계층 구조에 대한 연구는 불평등 측도에 대한 연구, 분할 알고리즘에 대한 연구가 주로 이루어져 왔다.

불평등 측도 측면에서는 물질적인 자산이나 자본의 불평등에 집중한 계층 연구가 주로 이루어지고 있는데 대표적인 것이 지니 계수 (Gini index) 관련 분석이다. 지니 계수는 불평등 정도를 직관적으로 파악하는 것이 용이하기 때문에 많이 활용되고 있다. 하지만 지니 계수를 산출하는 데 있어서 전체적인 사회 네트워크 구조에 대해서는 고려하지 않기 때문에 소셜 네트워크 구조 분석을 반영하는 개선의 여지가 남아있다.

네트워크 구조상의 분할 알고리즘에 있어서는 권력 관계를 반영하는 네트워크가 주어질 때, 여기에서 실제 계층 구조를 검출해내는 문제를 생각할 수 있는데, 이는 소셜 네트워크를 가중치 방향 그래프로 표현한 뒤, 여기서 계층간 연결의 최소화와 계층 내 연결의 최대화를 달성하는 분할 알고리즘을 찾는 문제로 풀어낼 수 있다. 즉, 가중치 방향 그래프에서의 위상정렬 문제로 변환 가능하다. 다만 이 문제는 NP-hard 문제로, 아직 초기 단계의 연구에 머물러있다.

계층 구조 연구와는 달리 집단 구조 연구는 비교적 활발하게 이루어져 왔다. 네트워크 집단 구조 분석에 관한 연구는 여러 방법들을 통하여 이루어져 왔는데 [24] [2], 집단 구조를 검출해내는 방법에 대하여 모듈화 기반 알고리즘, 고유벡터 분석, 무작위 행보 알고리즘 등이 개발되어 있다. 또한 집단 검출에 있어서 네트워크 구조뿐만 아니라 네트워크 각 노드들의 내용을 반영하는 집단 검출도EM (Expectation-Maximization) 기반 알고리즘 [14], 행렬 분해 기반 알고리즘 [25] 등도 개발되어왔다. 최근에는 overlapping community detection 방식의 구조 검출 알고리즘 연구도 활발하게 진행되고 있다 [27] [28].

집단 구조 검출에 있어서 최근 주목받는 방법으로 Blondel et al.[9] 등이 제시하는Modularity Maximization(MM)에 기반한 집단 구조 검출이 있다. 여기서 modularity는 같은 그룹 내에서는 많은 연결을 가지고 다른 그룹과는 적은 연결을 가질수록 높게 나타내는 값이다. MM 기법에서는 이 modularity를 최대화시키는 방법을 통하여 집단 구조를 검출해낸다.

이러한 방법 이외에도 커널 방식의 SPML (Structure-

Preserving Metric Learning) [26] 과 같은 알고리즘이 개발되었다. 커널 기반 군집화에서는 네트워크 연결 구조와 내용상의 유사도를 반영한다. 이는 특정 차원의 유클리드 공간으로 임베딩하는 것과 같은 개념인데, 이산적 구조인 네트워크에 대해 연속성을 요구하는 작업을 가능하게 하는 장점이 있다. 대표적으로 유클리드 공간상에서 적용 가능한 k-means 군집화 알고리즘을 네트워크 구조에 대해서도 바로 사용할 수 있도록 하고, 높은 차원으로의 임베딩을 통해 비선형적 분포에 대한 처리도 선형적으로 근사화시킨 이후에 해결할 수 있다. 다만 커널 기법의 특성상 계산복잡도가 높은 한계가 존재하기 때문에 최근 많이 다루는 실세계 소셜 네트워크 빅 데이터에 적용하기 위해서는 확장성 측면에서의 개선이 요구되고 있다.

집단 검출 기법은 집단 검출의 정확성과, 네트워크 크기에 대한 확장성을 지금보다 향상시키기 위한 방향으로 계속 연구되고 있다. 이를 통해 규모가 큰 실세계 소셜 네트워크에서도 비교적 정확하게 집단 구조를 파악하여 이를 다양한 관련 분야에서 활용이 가능하게 하는 것이 이들 연구의 궁극적인 목표이다.

II-3. 소셜 네트워크 진화 과정 예측

우리가 생활하는 실제 세계는 고정되어있지 않고 시간의 경과에 따라 계속 변화하는 특성을 가진다. 소셜 네트워크도 이와 마찬가지로 시간에 따른 지속적이고 빠른 변화를 겪는다. 그렇기 때문에 네트워크에서 아직 발견되지 않은 정보 또는 시간의 진행에 따른 네트워크의 진화를 예측하는 것이 네트워크 분석에 있어서 현재 네트워크의 특성을 분석하는 것 이상으로 중요하다고 생각할 수 있다. 네트워크의 진화는 다음과 같은 양상으로 설명이 가능한데, 구조적 측면에는 각 노드 (node) 나 노드 사이의 연결(link)이 생성 혹은 소멸되는 현상을 들 수 있고, 내용 (contents) 적 측면에서는 네트워크상의 정보들이 확산되거나 사라지는 등의 변화로 볼 수 있다. 즉 네트워크 상에서의 연결 예측, 네트워크 복원과 트렌드 예측 등이 네트워크 진화에 대한 대표적인 연구 주제라고 할 수 있겠다.

먼저 연결 예측 분야부터 살펴보면, 네트워크에서 연결이 동적으로 변화하는 양상을 예측하는 방식으로 유사도 부여 방법이 주로 연구되어 왔다. 여기서는 각 노드들 사이에 유사도를 부여하는데, 이 유사도가 높을수록 두 노드 사이를 이어주는 연결이 존재할 가능성이 높다고 가정하는 것이다. 따라서 여기서는 어떻게 유사도를 부여하느냐가 매우

중요해지는데, 유사도를 부여 방식에는 크게 두 가지 방식이 있다. 첫 번째로 이웃 노드의 숫자로 유사도를 부여 방식이다. 대표적으로 Common Neighbors, Jaccard, Adamic/Adar[6] 유사도 방법을 들 수 있다. 두 번째 방법으로는 노드끼리 서로 연결된 모든 경로 (Path) 를 고려하여 유사도를 부여할 수 있는데, Katz[19], Simrank[15], Hitting Time 기법과 Google의 Page Rank에서 변형된 Rooted Page[5] 유사도 부여 방식이 이에 해당된다. 최근에는 Supervised Random Walk를 통해 네트워크 구조 및 노드와 에지의 특성을 모두 고려한 방법[3]도 제시되었다.

한편 트렌드 예측은 네트워크의 연결 관계보다 네트워크의 구성 내용적인 측면에 초점을 맞추어 네트워크가 진화하는 과정을 예측하는 것이라 할 수 있다. 여기에는 먼저 데이터를 기반 (data-driven) 으로 하여 내용을 분석하는 접근법이 있는데, 자연언어 처리(Natural Language Processing) 를 통하여 페이스북 등의 소셜 네트워크나 블로그 등에 게시된 글의 내용을 자동으로 분석하여 처리하게끔 구성할 수 있다. 이러한 분석 가운데서 필자의 주관적 의견, 이를테면 특정 주제에 관하여 긍정적이나 부정적으로 느끼는 감정을 검출할 수 있는데 이를 감성분석 (sentiment analysis) 라고 한다. 감성분석에서는 소비자가 작성한 제품의 이용후기 와 같은 텍스트를 대상으로 진행할 수 있는데, 텍스트에서 감성을 추출하는 방법으로는 토픽 모델의 일종인 LSA(Latent Semeatic Analysis)나 분류 기법의 하나인 SVM(Support Vector Machine) 등을 사용한다. 다만 일반적인 감성분석 기법들은 네트워크 구조를 고려하지 않기 때문에 이후 MRF(Markov Random Field)를 이용해 네트워크 구조와 내용을 모두 반영하는 토픽 모델[8]도 제시되었지만 확장성의 측면에 있어서 한계를 지니고 있다는 단점을 가지고 있다

한편 모델 기반 (model-driven) 접근 방법은 정보 흐름의 패턴을 모델링하여 트렌드를 예측하는 것을 의미한다. 모델 기반 접근법에 대한 연구는 정보 확산 모델에서 정보 전파에 대한 임계점을 분석하여 특정 정보의 확산을 촉진 혹은 억제하기 위해 필요한 조건은 무엇인지, 또는 긍정적인 정보와 부정적 정보가 같이 확산될 때 어떤 것이 보다 우세하게 확산되는지[18] 등을 파악하는 것 등을 들 수 있다. 이에 트렌드 예측을 위한 여러 모델이 제시되었지만 바이럴 마케팅과 같은 실제 분야에서처럼 경쟁적 정보가 공존하는 경우나 정보의 수용 단계가 구분되는 경우와 같은 현상을 표현하는 데 있어서는 분명한 한계를 보이고 있다. 그리고 기존 연구에서는 확산되는 정보 자체의 특징이 고려되지 않아 현실에서 루

거나 유행 등이 확산되는 과정을 효과적으로 표현하기 어려운 문제가 있다.

네트워크 예측 문제에 관한 연구 성과들도 폭 넓은 분야에 응용이 되고 있는데 비단 소셜 네트워크 서비스에 포함된 친구, 상품 등에 대한 추천 시스템뿐만 아니라 복잡계 네트워크로 해석할 수 있는 생물학에서의 단백질 상호작용[17] 추측에까지 적용되고 있다. 뿐만 아니라 트렌드에 민감한 여러 기업들의 마케팅 전략이나 경영 지원 업무, 그리고 선거 혹은 공공 캠페인의 전략을 수립하는 데에도 이용이 가능하며 나아가 새로운 인간관계의 형성 패턴에 대한 이해 및 예측을 가능하게 한다.

III. 결론

지금까지 소셜 네트워크의 정보 확산 패턴 및 구조 분석 연구에 대한 접근법을 알아보았고 소셜 네트워크 빅 데이터를 적용하기 위해 어떤 알고리즘들이 연구되고 있는지 살펴보았다. 사회적 비중과 중요성이 높아지고 있는 온라인 소셜 미디어에 걸맞게 소셜 네트워크를 깊이 이해하기 위한 여러 연구가 폭넓게 진행 중인 가운데, 한편으로는 알고리즘의 확장성과 계산 복잡도 측면에서 기존 연구 결과들은 아직 부족한 점이 많다.

실제 사회에서의 개인 및 집단 간 관계가 투영된 소셜 네트워크를 올바르게 이해하고, 이런 소셜 네트워크의 성질을 다양한 분야에 활용, 이로부터 파생 가능한 여러 사회적 이득을 얻는 것이 소셜 네트워크 분석 연구의 목표이다. 이러한 목표를 달성하기 위해서는 결국 규모가 큰 실세계 네트워크에 대해 구조 분석 및 정보 확산 제어 등이 가능해야 한다. 그러므로 실세계 네트워크를 정확하게 반영할 수 있는 모델 못지 않게 이를 기반으로 실세계의 소셜 네트워크를 효과적으로 분석할 수 있는 빅 데이터 처리 기반 문제 해결 알고리즘의 중요성이 더욱 커지고 있으며, 따라서 향후 이들 분야에서 지속적으로 활발한 연구가 이루어 질 것으로 기대된다.

참고 문헌

[1] N. Alon and J. H. Spencer, The Probabilistic Method, Wiley, 1992, xiii+254 pp.
 [2] A. Arenas, A. Díaz-Guilera, and C. J. Pérez-

Vicente, Synchronization Reveals Topological Scales in Complex Networks, Phys. Rev. Lett. 96 114102, 2006.

- [3] L. Backstrom and J. Leskovec, Supervised Random Walks: Predicting and Recommending Links in Social Networks, In proc. of the 4th ACM WSDM conference, 2011.
 [4] A. Barabási, Linked: The New Science of Networks, Perseus Publishing, April 2002.
 [5] S. Brin and L. Page, The Anatomy of a Large-scale Hypertextual Web Search Engine, Computer Networks and ISDN Systems, 30(1-7):107-117, 1998.
 [6] G. Salton and M. J. McGill. Introduction to Modern Information Retrieval, McGraw-Hill, 1983.
 [7] W. Chen, C. Wang, and Y. Wang, Scalable Influence Maximization for prevalent viral marketing in large scale social networks, In proc. of the 16th ACM SIGKDD, 2010.
 [8] Y. Sun, J. Han, J. Gao, and Y. Yu, iTopicModel: Information Network-Integrated Topic Modeling, In proc. of the 10th IEEE ICDM, 2009.
 [9] V. D. Blondel, J. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks, J. Stat. Mech 2008(10), P10008, 2008
 [10] Bloter.net, [그래프] 주요 커뮤니티의 모바일 체류시간, <http://www.bloter.net/archives/173396>
 [11] J. C. Miller, Percolation and Epidemics in Random Clustered Networks, Phys. Rev. E 80, 020901, 2009.
 [12] J. Goldenberg, B. Libai, and E. Muller, Talk of the network: A complex systems look at the underlying process of word-of-mouth, Marketing Letters. Vol. 12, No. 3, pp.211-223, 2001.
 [13] M. Granovetter, Threshold Models of Collective Behavior, American journal of sociology. Vol. 83, No. 6, pp.1420-1443, 1978.
 [14] T. Yang, R. Jin, Y. Chi, and S. Zhu, Combining Link and Content for Community Detection: a Discriminative Approach, In proc. of the 15th SIGKDD, 2009.
 [15] D. E. Whitney, Dynamic Theory of Cascades

- on Finite Clustered Random Networks with a Threshold Rule, Phys. Rev. E 82, 066110, 2010.
- [16] 정하웅, 강병남, 복잡계 네트워크에 대한 최근 연구 동향, 물리학과 첨단기술, 2007년 10월호.
- [17] F. Kuhn, K. Panagiotou, J. Spencer, and A. Steger, Synchrony and Asynchrony in Neural Networks, In proc. of the 21st SODA, 2010.
- [18] B. Karrer and M. E. J. Newman, Random Graphs Containing Arbitrary Distributions of Subgraphs, Phys. Rev. E 82, 066118, 2010.
- [19] L. Katz. A New Status Index Derived from Sociometric Analysis, Psychometrika, 18(1):39-43, 1953.
- [20] D. Kempe, J. Kleinberg, and É. Tardos, Maximizing the Spread of Influence Through a Social Network. In proc. of the 9th ACM SIGKDD, 2003.
- [21] J. Kleinberg, Navigation in a Small World, Nature 406:845, 2000.
- [22] J. Leskovec and J. Kleinberg and C. Faloutsos, Graphs over Time: Densification Laws, Shrinking Diameters and Possible Explanations, In proc. of the 11th ACM SIGKDD, 2005.
- [23] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. Van-Briesen, and N. S. Glance, Cost-effective Outbreak Detection in Networks, In proc. of the 13th ACM SIGKDD, 2007.
- [24] M. E. J. Newman and M. Girvan, Finding and Evaluating Community Structure in Networks, Phys. Rev. E 69, 026113, 2004.
- [25] S. Zhu, K. Yu, Y. Chi, and Y. Gong, Combining Content and Link for Classification Using Matrix Factorization, In proc. of the 30th SIGIR, 2007.
- [26] B. Shaw, B. Huang and T. Jebara, Learning a Distance Metric from a Network, In proc. of the 25th NIPS, 2011.
- [27] S. Lim, S. Ryu, S. Kwon, K. Jung and J. Lee, LinkSCAN*: Overlapping Community Detection Using the Link-Space Transformation, ICDE 2014, pp292-303, 2014
- [28] G. Palla, I. Derenyi, I. Farkas, and T. Vicsek, Uncovering the overlapping community structure of complex networks in nature and society, Nature, vol. 435, pp. 814 - 818, 2005.

약 력



최 병 진

2014년 KAIST 전기 및 전자공학과 학사
2014년~현재 서울대학교 전기정보공학부
석·박사 통합과정
관심분야: Social Network 모델링 및 분석, 추천
시스템



황 용 군

2014년 한양대학교 전자정보시스템공학전공 학사
2014년~현재 서울대학교 전기정보공학부
석·박사 통합과정
관심분야: 추천 시스템, Social Network 분석,
Deep Learning



정 교 민

2003년 서울대학교 수학과 학사
2009년 MIT 수학과 박사
2009년~2013년 KAIST 전산학과 교수
2013년~현재 서울대학교 전기정보공학부 교수
관심분야: 기계학습, 추천 시스템, Social Network
정확화산 분석 및 예측