



특집 04

# 빅 데이터 기반의 소셜 네트워크 정보 확산 및 구조 분석 연구 동향

최병진·황용근·정교민 (서울대학교)

- 
- 목 차 »
1. 서 론
  2. 소셜 네트워크의 정보 확산 패턴 및 구조 분석 문제의 접근법
  3. 국내외 연구 동향
  4. 결 론
- 

## 1. 서 론

최근 정보 통신 기술의 발달로, 개인 또는 집단의 상호 의존적인 사회관계를 네트워크 구조로 이해하는 소셜 네트워크(Social Network)가 웹상에서 다양한 형태로 급성장하고 있다. 특히 페이스북(Facebook), 트위터(Twitter), 텀블러(Tumblr), 카카오토리(KakaoStory) 등 온라인 소셜 미디어(Online Social Media)의 부상은 새로운 정보 전파 네트워크로서 현대인의 정보 습득과 인간 관계의 새로운 방식으로 자리매김 하고 있다. 따라서 소셜 네트워크의 구조와 정보 흐름의 패턴을 이해하고 이를 제어하는 것은 이러한 미디어를 효과적으로 사용하기 위한 근본적인 문제이다.

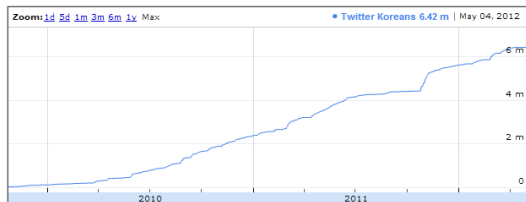
소셜 네트워크는 시간의 흐름에 따라 급격하게 진화하는 체계이기 때문에 각 구성 요소를 이해하는 것만으로는 완벽하게 설명되지 않으며, 망의 구조적 특성과 변화 과정을 이해해야 한다<sup>[26]</sup>. 소셜 네트워크들의 구조는 다음과 같은 다양한

공통된 성질들을 보인다. 임의의 두 구성원이 매우 짧은 경로로 연결되어 있는 좁은 세상(small world) 구조<sup>[24]</sup>, 연결이 많은 노드(node)인 허브(hub)들의 존재로 인한 불균등한 위상 구조, 서로 유사한 구성원들 간에 보다 긴밀히 연결되어 있는 공동체 구조 등이 그것이다<sup>[4]</sup>. 소셜 네트워크 내에서의 정보의 흐름은 이러한 구조적 특성에 기반하여 이해해야 하며, 이를 바탕으로 정보의 확산을 제어하는 것은 매우 중요한 문제이다. 소셜 네트워크 분석(SNA, Social Network Analysis)에 관한 연구는 높은 잠재적 가치를 지니고 있음에도, 아직까지 연구 초기 단계로서 전산, 사회학, 물리, 수학, 경제 등 여러 분야에서 다각도로 연구되고 있는 융합적인 연구 분야이다. 특히 대규모로 주어지는 데이터 분석(big data analysis)을 통한 정보 확산 패턴의 이해, 네트워크 구조에 대한 이해 및 네트워크 진화와 이에 대응하는 정보 흐름 제어 문제는 지난 수년간 세계적으로 급격히 연구되기 시작하였다.

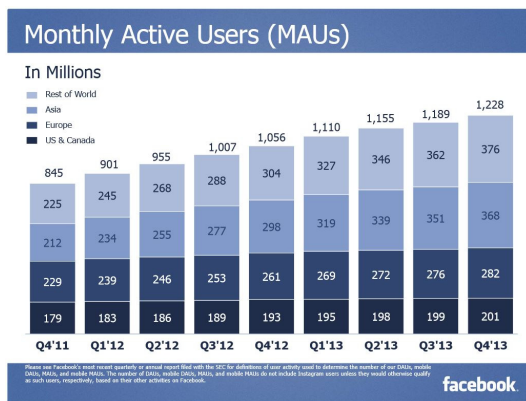
본 기고문에서는 소셜 네트워크의 정보 확산 패턴 및 구조 분석 문제의 접근법들에 대해 설명하고 각 분야에서의 국내외 연구 동향을 점검할 것이다.

## 2. 소셜 네트워크의 정보 확산 패턴 및 구조 분석 문제의 접근법

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 양상 연구이다. 다른 한 가지 방법은 머신러닝 등의 기법을 활용한 직접적인 소셜 네트워크 데이터의 분석을 통해 과학적 사실을 연구하는 것이다. 이 두 가지 방식은 상호보완적이어서 통합적이고 심



(그림 1) 2009년-2012년 한국 트위터 유저 수 변화<sup>[10]</sup>



(그림 2) 2011년-2013년 페이스북 월별 액티브 유저 수 변화<sup>[11]</sup>

도있는 이해와 연구가 필요하다. 이러한 소셜 네트워크에 대한 연구는 크게 세 가지 문제들로 구분된다<sup>[19]</sup>.

### 2.1 소셜 네트워크 정보 흐름 패턴 분석

우리는 종종 기술적으로 큰 차이가 없어 보이는 여러 서비스가 어떤 것은 시장에서 실패하고 어떤 것은 성공하는 것을 보곤 한다. 이러한 현상은 소셜 네트워크 상에서도 관찰할 수 있다. 서비스의 초기 유저 수가 특정한 임계점(tipping point)에 도달했는지의 여부가 이를 결정하게 된다. 서비스뿐만 아니라 일반적인 정보에 대해서도 초기 상태에 따라 특정한 정보가 전체 네트워크로 확산되거나 거의 확산되지 못하는 상전이(phase transition) 현상이 관찰된다<sup>[4]</sup>. 이 현상의 발생 조건 또는 원인을 파악한다면 특정 정보의 확산 규모의 예측이 가능해지고 상품의 마케팅 전략 수립이나 전염병 확산 방지와 같은 분야에 활용 가능할 것이다. 이를 위해 네트워크 정보 확산 모델 및 실제 데이터 분석을 통해 임계점 및 광범위 확산의 발생 확률 및 조건을 규명하고, 최종 확산 규모 분석을 시행하는 연구가 이루어지고 있다.

정보 흐름 패턴 분석의 또 하나의 중요한 문제는 한정된 자원을 활용하여 정보 흐름에서의 영향력을 최적화시키는 문제이다. 이 문제는 일반적으로 매우 어려운 문제(NP-hard)로, 따라서 이 문제의 풀이를 위해 대규모 네트워크에서 빠른 시간에 수행되는-scalable) 알고리즘의 개발이 요구된다.

### 2.2 계층 및 집단 구조 분석

네트워크의 연결 구조 분석은 소셜 네트워크 분석의 중요한 축이다. 네트워크의 구조 분석은

두 가지 측면, 계층 구조의 분석과 집단 구조의 분석으로 접근할 수 있다. 계층 구조 연구는 소셜 네트워크 분석에서의 중요성에 비해서 연구가 거의 되지 않은 편이다. 계층 구조는 사회 불평등 척도 등과 연관되어 있어 계층 구조 분석 및 연구를 통해 소셜 네트워크의 정보 전달 불평등 척도 측정 및 완화 알고리즘의 개발이 가능하다.

집단 구조 연구는 계층 구조 연구가 거의 이루어지지 않은 것과 반대로 상당히 활발하게 이루어져왔다. 네트워크 집단 구조는 연구에 따라 다양한 방식으로 이해되어왔다<sup>28,21</sup> 네트워크에서 구조뿐만 아니라 노드의 내용을 반영하는 기법에 관한 연구도 이루어지고 있다<sup>17,29</sup>. 더욱 확장성 있고 정확도가 높은 집단 검출 기법이 계속 연구되고 있으며, 이러한 것은 더욱 실제와 맞는 네트워크 구조 모델링 및 정보 흐름 분석 등으로 다양하게 활용이 가능하다.

### 2.3 소셜 네트워크 진화 과정 예측

빠르게 변화하는 실제 세계와 같이 소셜 네트워크 또한 시간에 따라 변화하는 특성을 가진다. 따라서 네트워크 분석에서는 현재의 특성을 분석하는 것뿐만 아니라 아직 발견되지 않은 정보 또는 시간이 진행됨에 따른 네트워크의 진화를 예측하는 것은 매우 중요한 문제라고 할 수 있다. 네트워크의 진화라는 것은 구조적인 측면인 노드(node)와 연결(link)의 생성 및 제거와 내용(contents)적 측면인 네트워크상의 정보들이 확산되고 사라지는 변화를 같이 말한다. 네트워크 진화에 관한 연구로는 대표적으로 트렌드 예측(trend prediction)과 연결 예측(link prediction)<sup>25</sup>이 있다. 이러한 예측 연구들은 트렌드에 민감한 여러 기업들의 경영 전략 수립에 이용할 수 있고, 선거 혹은 공공 캠페인의 전략에도 이용 가능하

며, 또한 새로운 인간관계의 형성 패턴에 대해 이해 및 예측을 가능하게 한다.

## 3. 국내외 연구 동향

최근 연구 결과들을 살펴 볼 때, 소셜 네트워크 관련 분야는 컴퓨터 과학 분야 중에서도 가장 활발하게 연구되는 분야 중 하나이며 국외뿐만 아니라 국내에서도 많은 연구진들이 연구에 기여하고자 하고 있다. 여기에서는 위에서 제시한 3가지 주요 연구 주제에 해당하는 기존 연구 결과를 살펴보고, 향후 해결해야 하는 과제들이 무엇인가에 대해 이야기하고자 한다.

### 3.1 소셜 네트워크 정보 흐름 패턴 분석

그동안 네트워크상의 정보 확산 현상을 설명하는 많은 모형들이 제시되어 왔으며, 대표적으로 두 가지 모델이 논의되고 있다. 첫 번째로 Linear Threshold(LT) 모델이 있는데 이는 Granovetter<sup>16</sup> 등에 의해 제안되었으며 각 노드가 새로운 정보를 받아들일 것인지 여부는 해당 노드마다 정해진 문턱값(threshold)에 의해 결정된다. 다음으로 Goldenberg, Libai, and Muller<sup>14</sup>가 제시한 Independent Cascade(IC) 모델이 있는데 이는 각 링크를 통해 독립적으로 정보가 확산될 확률을 정의하여 확산 패턴을 표현한다. Kempe et al.<sup>123</sup>은 LT, IC 모델에 대해 모두 적용 가능한 일반화된 threshold model을 제시하였고 이를 통해 정보 확산의 상전이 현상 연구가 많이 진행될 수 있었다. 하지만 이러한 연구들은 주로 나무에 가까운(locally tree-like) 특수한 구조의 그래프에서 수행되어왔으며<sup>18</sup> 구조를 고려한 분석에서는 대부분 정보 확산 확률이 고정된 경우만을 고려<sup>12</sup>하였다. 이들 연구는 따라서 실세계의 네트워크에서

발생하는 정보 확산 현상을 설명하는 데에는 분명한 한계가 있다.

한편 정보 흐름 최적화 문제는 영향력 최대화 (influence maximization) 문제로 정의하는데, 이는 정보 확산 패턴을 이해하는데 있어 상전이 현상 분석과 함께 반드시 필요하다. 이 문제에 대해 Kempe et al.<sup>[23]</sup>은 LT, IC 모델에서의 영향력은 서브모듈러(submodular) 함수로 표현할 수 있으며 이를 최대화 하는 문제는 매우 어려운 문제 (NP-hard)임을 증명하였다. 이러한 영향력 최대화 문제의 해를 구하는 데 있어 Kempe et al.<sup>[23]</sup>, Leskovec et al.<sup>[27]</sup> 등은 최적해에 대한 근사치를 보장하는 탐욕적 알고리즘(greedy algorithm)을 제안했는데, 해의 정확도 측면에서는 비교적 우수한 결과를 얻을 수 있지만 몬테-카를로 시뮬레이션(Monte-Carlo Simulation)을 이용하는 근본적인 한계 때문에 실제 데이터에 적용하기에는 계산 속도가 느리다는 문제점이 있다. 따라서 실행 시간 향상을 위한 휴리스틱(heuristic) 알고리즘도 많이 연구되고 있는데 알려진 연구 성과 가운데 가장 효율적인 휴리스틱 알고리즘으로는 Chen et al.<sup>[7]</sup>이 제안한 PMIA(Prefix-excluding Maximum Influence Arborescence) 알고리즘을 들 수 있다. 다만 이 알고리즘의 경우 네트워크 구조에 매우 민감하여 긴밀히 연결되어 있는 네트워크에서는 실행 속도가 현저히 저하되는 등 일정한 성능을 얻어내기 힘들며, 메모리 자원을 많이 사용한다는 단점을 가지고 있다.

### 3.2 계층 및 집단 구조 분석

사회적 위상 구조의 전역적인 구조(global structure)를 파악, 분석하는 것은 연결 구조 예측, 트렌드 예측 등의 관점에서 중요한 문제이다. 네트워크의 노드들이 형성하는 주된 전역 구조는

집단 구조(community structure)와 계층 구조(hierarchical structure)를 들 수 있다.

한편 이런 사회적 연결 구조를 고려한 계층 구조에 대한 연구는 사회적으로 매우 중요한 의미를 가진다. 예를 들면 실제 사회의 계층 구조를 완화하기 위해 네트워크 구조를 개선하는 데 드는 비용의 최적화 문제의 해결책도 제시 가능하다. 그럼에도 불구하고 이와 관련된 기존 연구들은 국내뿐만 아니라 해외에서도 그 사례가 적은 편이며, 다음과 같이 교환 이론(exchange theory)과 불평등 측도(measurement of inequality)에 관한 연구, 그래프의 분할(partitioning) 알고리즘에 대한 연구로 국한되어 왔다.

#### 3.2.1 불평등 측도 연구

사회적 불평등, 특히 물질적인 자산이나 자본의 불평등에 초점을 둔 연구도 많이 행해져 왔다. 자본분배의 불평등 측면에 대한 연구 가운데 대표적인 것으로 지니 계수(Gini index)를 들 수 있는데, 지니 계수(Gini index)란 로렌츠(Lorenz) 곡선으로 불리는 자본의 누적(cumulative) 분포와  $y=x$  선 사이의 면적을 구하는 것으로 면적이 넓을수록 빈부격차가 큼을 나타낸다. 지니 계수는 매우 직관적인 계량법이라는 특징 덕분에 실제로 여러 국가의 불평등 측도를 계산하는 등에 많이 활용되고 있지만 사회의 네트워크 구조나 전역적인 관계 구조는 전혀 고려하지 않는 맹점을 가지고 있다.

#### 3.2.2 네트워크 구조상의 분할 알고리즘에 대한 연구

권력관계를 반영하는 어떤 네트워크에서 실제 계층 구조를 찾아내는 문제를 생각해 볼 수 있다. 이는 알고리즘 관점에서 가중치 방향 그래프(weighted directed graph)가 주어져 있을 때 계층

간 연결을 최소화 하고 계층 내 연결을 최대화하는 분할 알고리즘을 찾아내는 것으로 이해가 가능하다. 결국 이 문제는 가중치 방향 그래프에서 위상정렬(topological sorting)을 하는 문제로 변환될 수 있으며 이는 매우 어려운 문제(NP-hard)임이 알려져 있다. 이 문제를 푸는 알고리즘에 대한 연구는 초기 단계에 머물러 있다.

주어진 네트워크에서 집단 구조를 검출하는 연구는 모듈화 기반(modularity-based) 방법과 고유벡터 분석(spectral analysis), 무작위 행보(random walk) 알고리즘 등이 개발되어 있다. 그리고 네트워크 구조만을 고려하는 집단 검출 알고리즘 이외에도 구조 및 노드의 내용(contents)을 모두 반영하는 집단 검출도에 대한 EM (Expectation-Maximization)-기반 알고리즘<sup>[17]</sup>, 행렬 분해(matrix factorization)<sup>[29]</sup> 기반 알고리즘 등이 연구되어 왔다. 최근에는 노드 간의 구조적·내용적 유사도를 거리로 나타내어 학습하는 커널 방식의 SPML (Structure-Preserving Metric Learning) 알고리즘<sup>[9]</sup>이 주목받고 있다. 그 중 커널 기반의 군집화는 노드들을 연결 구조와 내용상 유사도를 반영한다. 이는 특정 차원의 유클리드 공간으로 임베딩하는 것과 같은 개념인데, 이산적 구조인 네트워크에 대해 연속성을 요구하는 작업을 해 줄 수 있는 장점이 있다. 예를 들면, 커널을 통해 임베딩을 한 후 유클리드 공간에서 k-평균 군집화와 같이 흔히 사용되는 알고리즘을 바로 적용할 수도 있다. 여기에 높은 차원으로 임베딩함으로써 비선형적 분포에 대해서도 선형적으로 근사화하여 효과적으로 처리할 수 있는 특징도 있다.

이처럼 커널 기반의 군집화는 비선형적 이산 데이터를 다룰 수 있기에 소셜 네트워크의 노드를 군집화하기에 적합하지만, 커널 기법의 특성상 임베딩되는 실제 좌표를 알 수 없기 때문에 높은 계산 복잡도를 요구한다. 이는 군집화의 실용

성을 크게 제한하는데 최근 많이 다루는 소셜 네트워크들은 그 크기가 매우 크기(big data) 때문이다. 따라서 커널 기반의 네트워크 군집화 알고리즘에 있어서는 계산 복잡도를 낮추는 것이 중요한 과제로 떠오르고 있다.

### 3.3 동적 변화와 트렌드 예측

소셜 네트워크에서의 진화 과정 예측 문제는 네트워크의 변화 히스토리 및 특성에 대한 정보를 통해 앞으로 생겨날 연결이나 관찰되지 않은 연결을 예측하는 문제로 볼 수 있다. 이는 전체적인 네트워크 구조 변화뿐만 아니라, 특정 노드에 대한 연결성도 예측할 수 있어서 개인적 연결 예측(link prediction)과 네트워크 복원 (reconstruction) 등에 두루 쓰일 수 있다. 이러한 기법은 소셜 네트워크 서비스의 추천 시스템뿐만 아니라 복잡계 네트워크인 생물학에서의 단백질 상호작용<sup>[20]</sup> 추측에 이르기까지 다양한 분야에서 응용되고 있다.

기존의 동적 변화 예측 방식은 각 노드 사이에 유사도(similarity)를 부여한 다음 그 값이 높은 순으로 연결이 있을 가능성이 높다고 가정한다. 여기서 유사도를 부여하는 방식은 크게 두 가지 방법이 있다. 첫 번째로 공통된 이웃(neighbor)들의 숫자로 유사도를 부여하는 방식이 있는데 대표적으로 Common Neighbors, Jaccard, Adamic/Adar<sup>[6]</sup> 유사도 방법이 있다. 두 번째로 노드사이의 모든 경로(path)를 고려하는 방식이 있는데 대표적으로 Katz<sup>[22]</sup>, Hitting Time과 Google의 Page Rank에서 변형된 Rooted Page Rank<sup>[5]</sup> 유사도 방법이 있다. 특히 경로를 고려한 방법 중에서는 지도된 무작위 행보(Supervised Random Walks)를 이용한 네트워크 구조와 노드와 에지(edge)의 특성 모두를 고려하여 정확성을 향상시킨 방법이 최근에 제시되기도 하였다<sup>[3]</sup>.

진화 과정 예측 문제의 주요 과제는 크게 두 가지를 들 수 있는데, 첫 번째로 알고리즘의 정확성과 확장성(scalability)을 향상시키는 것이 있겠고 두 번째는 진화 과정 예측을 다방면으로 응용하는 것이다. 연결 예측은 위에 제시된 바와 같이 여러 가지 방법들이 존재하지만 정확도 향상이 필요하다. 정확성 측면에서는 최근에 나온 지도된 무작위 행보 알고리즘이 가장 우수한 편이지만 링크가 있을 상위 20개의 노드를 맞추는 정확도가 40%에 불과하므로 개선의 여지가 많다고 볼 수 있다<sup>3)</sup>. 그리고 온라인 소셜 네트워크처럼 빅 데이터(big data) 분석을 효과적으로 수행하기 위해서는 네트워크가 집단 구조를 갖는 특성을 이용한 확장성(scalable) 있는 알고리즘 개발이 필요하다. 이와 더불어 실제 세계에 연결 예측을 적용할 수 있는 새로운 주제를 발굴하는 것도 중요한 과제이다.

구조적인 측면에 대한 부분이 연결 예측이라면 여기에 구성과 내용(contents)의 측면까지 고려하여 동적인 변화를 예측하는 것이 트렌드 예측이다. 트렌드 예측의 예시로서 최근에 각광 받는 쿨헌팅(coolhunting)을 들 수 있는데, 이는 데이터 기반(data-driven)의 접근 방법으로서 현존하는 문화적 유행에 대해 관찰 및 예측을 하는 일련의 방법론을 가리킨다<sup>13)</sup>. 또한 이러한 방식을 블로그 공간 등, 소셜 네트워크를 대상으로 자동화하기 위해 자연언어 처리(NLP: Natural Language Processing) 기법을 주로 사용한다. 여기에서 특히 특정 주제에 대한 필자의 주관적인 의견을 검출해 내는 작업이 감성 분석(sentiment analysis)이다. 감성 분석은 소비자가 남긴 제품의 구매 후기와 같은 텍스트를 대상으로 주제 모델(topic model)의 하나인 잠재 의미 분석(LSA: Latent Semantic Analysis)과 분류(classification) 기법의 하나인 지지 벡터 머신 (SVM: Support Vector

Machine) 등을 사용한다. 그러나 대부분의 감성 분석 기법들은 네트워크 구조를 고려하지 않는다는 맹점이 있다. 이에 Sun et al.<sup>8)</sup>는 마르코프 임의 장(MRF: Markov Random Field)을 이용해서 네트워크 구조와 콘텐츠를 모두 반영하는 주제 모델을 제시했지만 네트워크의 크기가 클 경우 알고리즘 수행에 있어서 매우 오랜 시간을 요하기 때문에 확장성에 한계를 가지고 있다.

모델 기반(model-based) 측면에서는 정보 흐름 패턴을 모델링하여 앞으로의 트렌드를 예측하는 분석 방법으로 접근되어 왔다. 정보 확산 모델에서의 임계점 분석을 통해 어떤 조건 하에서 특정 정보의 확산을 촉진 혹은 억제할 수 있을지, 또는 긍정적/부정적 정보가 공존하여 확산되는 경우 어떠한 정보가 우세할 것인지<sup>21)</sup> 등이 연구되어 왔다. 하지만 실제 소셜 네트워크를 통한 바이럴 마케팅이나 루머의 확산 등에 있어서 둘 이상의 경쟁적인 정보가 공존하는 경우나 정보에 대한 수용 단계가 구분되는 경우 등에 대한 분석은 기존의 모델들로 설명하기에는 한계가 있다. 또한, 데이터 기반의 방식과 달리 확산되는 정보 자체의 특징을 고려하지 못하기 때문에 현실 세계의 루머나 유행 등의 확산을 잘 설명하지 못하는 한계가 있다. 따라서 이 두 접근 방식의 한계들을 해결할 수 있는 상호 보완적인 접근 방법의 개발이 향후 과제로 떠오르고 있다.

#### 4. 결론

지금까지 소셜 네트워크의 정보 확산 패턴 및 구조 분석 연구에 대한 접근법 및 연구동향을 살펴보았다. 이처럼 소셜 네트워크에 있어서 정보 확산 현상 및 소셜 네트워크의 구조를 규명하고자 많은 연구가 이루어지고 있으며, 네트워크에 대한 모델링적 측면에서의 수학적 분석 기술에

실세계의 빅 데이터 연산 및 처리 측면에서의 머신러닝 기반의 분석 기술을 함께 적용하는 것이 향후 주요한 연구 주제로 떠오를 것으로 기대된다.

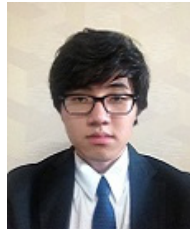
### 참 고 문 헌

- [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] B. Shaw, B. Huang and T. Jebara, Learning a Distance Metric from a Network, In proc. of the 25th NIPS, 2011.
- [10] OikoLab, 트위터 한국인 인덱스, <http://tki.oiko.cc/service/count>, 2012.
- [11] David Cohen, 4Q EARNINGS CALL: Facebook's Mobile Ad Revenue Of \$1.25B Accounts For 53% Of Total Revenue; Monthly Active Users Now Total 1.23B, 296M Mobile-Only, Allfacebook.com, 2014
- [12] J. C. Miller, Percolation and Epidemics in Random Clustered Networks, Phys. Rev. E 80, 020901, 2009.
- [13] P. Gloor and S. Cooper, Coolhunting: Chasing Down the Next Big Thing. American Management Association, ISBN 0814473865, 2007
- [14] 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.
- [15] B. Markovsky, D. Willer and T. Patton, Power Relations in Exchange Networks, American Sociological Review, 2011
- [16] M. Granovetter, Threshold Models of Collective Behavior, American journal of sociology. Vol. 83, No. 6, pp.1420-1443, 1978.
- [17] 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.
- [18] D. E. Whitney, Dynamic Theory of Cascades on Finite Clustered Random Networks with a Threshold Rule, Phys. Rev. E 82, 066110, 2010.
- [19] 정하웅, 강병남, 복잡계 네트워크에 대한 최근 연구 동향, 물리학과 첨단기술, 2007년 10월호.
- [20] F. Kuhn, K. Panagiotou, J. Spencer, and A. Steger, Synchrony and Asynchrony in

Neural Networks, In proc. of the 21st SODA, 2010.

- [21] B. Karrer and M. E. J. Newman, Random Graphs Containing Arbitrary Distributions of Subgraphs, Phys. Rev. E 82, 066118, 2010.
- [22] L. Katz, A New Status Index Derived from Sociometric Analysis, Psychometrika, 18(1):39-43, 1953.
- [23] D. Kempe, J. Kleinberg, and É. Tardos, Maximizing the Spread of Influence Through a Social Network, In proc. of the 9th ACM SIGKDD, 2003.
- [24] J. Kleinberg, Navigation in a Small World, Nature 406:845, 2000.
- [25] J. Kleinberg and D. Liben-Nowell, The Link Prediction Problem for Social Networks, In Journal of the American Society for Information Science and Technology, 58(7):1019-1031, May 2007.
- [26] 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.
- [27] 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.
- [28] M. E. J. Newman and M. Girvan, Finding and Evaluating Community Structure in Networks, Phys. Rev. E 69, 026113, 2004.
- [29] 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.

저 자 약 력



**최 병 진**

이메일 : bjchoe@snu.ac.kr

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



**황 용 군**

이메일 : wangcho2k@snu.ac.kr

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





**정 교 민**

이메일 : [kjung@snu.ac.kr](mailto:kjung@snu.ac.kr)

- 2003년 서울대학교 수학과(학사)
- 2009년 MIT 수학과(박사)
- 2009년~2013년 KAIST 전산학과 교수
- 2013년~현재 서울대학교 전기정보공학부 교수
- 1995년 국제수학올림피아드(IMO) 금메달 수상
- 관심분야: Big data Trend 예측, 추천 시스템, Social Network 정보확산 분석 및 예측