Network Theory

‘ 복잡계 네트워크 과학 ’ 이라는 책 내용을 정리. 최대한 수식은 제외한다.

1. 좁은 세상

1.1 밀그램의 좁은세상 실험

편지가 몇 사람을 거쳐 특정인에게 도착하는가? → 30%의 편지 도착, 평균 5.2단계 거침

1.2 한국에서의 좁은세상 실험

위 실험과 비슷하나 여러 통신방법 사용 → 목표 인물에게 도착된 연결은 15.7% (한국인들은 바빠서)
평균 4.6단계 거침

1.3 좁은세상이 왜 만들어질까

가장 간단한 원인은 한사람이 얼마나 많은 사람을 알고 지내는가와 관련.
솔라 풀의 연구에 따르면 3000~5000명

1.4 좁은세상 네트워크에 대한 스케치

Erodos와 Reny(ER)의 무작위 네트워크 소개

- N명의 사람(노드)이 있다고 가정하고 사람과 사람 사이에 p라는 확률로 연결선(인지관계, 링크) 있다고 생각하는 것
- 얼마나 잘 아는 사이인가 표현하기 위해 가중치를 줄 수도 있음
- 사람과 사람사이의 관계를 네트워크로 그려 집단이 어떻게 이루어지는지, 어떤 사람이 마당발인지 이해 가능

Watts와 Strogatz(WS)는 1998년 영화배우 네트워크, 선충의 신경만, 전력망을 연구

- 순서 없는 목록결집계수(이웃 노드들끼리 얼마나 잘 연결되어 있는지를 나타낸 양)라는 양 측정

이후 월드와이드 웹의 구조(웹페이지간의 연결)에 대한 관심을 가지게 됨

- 웹페이지 : 노드, 하이퍼링크 : 링크
- 한 노드에 연결된 링크의 수(도수)의 분포함수는 멱함수 법칙(P(k)~k^(-γ), γ는 임의의 실수)을 따른다.
  즉 도수가 큰 노드의 개수는 k^(-γ) 함수로 급격히 감소. 척도 없는 네트워크
- 도수가 매우 큰 노드를 허브라 부른다

네트워크는 사회학 뿐 아니라 생물학, 경제학, 컴퓨터공학 등에도 유용하게 쓰임

- 예 : 단백질 상호작용 네트워크, 신진대사 반응 네트워크

복잡계 네트워크 연구 분야는 구조적 특성 연구, 네트워크 위에서 일어나는 동역학 현상 연구, 네트워크 구조의 진화현상 연구로 구분 가능
모듈, 프랙탈네트워크, 보편성, 데이터 전송, 진화 현상등의 연구

2. 네트워크 분석

2.1 네트워크의 종류

- 시간에 따라 노드 수 증가하는 경우 **성장 네트워크**, 고정된 경우 **정적 네트워크**
- 노드 수 증가 속도보다 링크 수 증가 속도가 빠른 경우 **가속 성장 네트워크**
- 링크에 방향이 있는 경우 **방향성 네트워크**, 그렇지 않은 경우 **비방향성 네트워크**
  방향성 네트워크 중 하나인 신진대사 네트워크에서의 반응의 경우 역반응은 일어나지 않는 경우 존재
- 노드 사이의 링크가 있다 없다로 구분하는 경우 **이원성 네트워크**, 가중치가 있는 경우 **가중치 네트워크**
  두 사람 사이가 얼마나 잘 아는 사이인지 고려하는 경우 가중치 네트워크
- 노드 종류가 한가지이면 **단종 네트워크**, 두 가지 종류의 노드가 있고 링크는 다른 종류의 노드끼리만 연결가능하면 **이종분할 네트워크**

2.2 네트워크 구조 용어

- **도수** : 어떤 노드에 연결된 링크 수
- **평균거리(지름)** : 임의의 두 노드 간의 평균거리(한 노드에서 다른 노드로 가는데 지나가야 하는 최소 노드 수)
- **효율** : 수학적 정의만 나와 있으므로 생략
- **Betweenness Centrality(BC, 중앙성)와 부하(load)** : 네트워크 위에서 수송 현상을 연구하기 위해 소개된 양
- **두 이웃 도수 간의 상관관계** : 도수가 k인어느 한 노드의 이웃 노드들의 평균 도수.
  이 함수의 미분값의 부호에 따라 유사성 결합, 비유사성 결합으로 나눔
- **유사성 지표** : 이 값이 양수냐 음수냐에 따라 도수가 많은 노드와 적은 노드가 결합되는 경향이 많은지 적은지 알 수 있음
- **결집계수** : 이웃 노드들끼리 얼마나 잘 연결되어 있는지를 나타낸 양

3. 무작위 네트워크

- 1959년 Erdo와 Reny(ER)가 소개
- 정적네트워크
- 앙상블 개념을 이용한 확률적인 방법 이용
- 노드 간에 링크가 존재할 확률 p를 조절하면서 네트워크가 어떻게 변하는가 연구
- p가 증가함에 따라 많은 노드들을 연결하는 큰 송이가 언제 만들어질 것인가에 대해 큰 관심
    - 스미기 전이(percolation transition)이 언제 일어날 것인가
- 스미기 전이가 일어나기 시작하는 경우는 평균 도수가 1에 가까운 경우
- 각 노드당 연결된 링크 수의 분포는 푸아송 함수를 따른다
- 지름, 결집계수 계산, 노드수가 증가하는 ER 네트워크의 경우 어떻게 되는가?

4. 좁은세상 네트워크

- Watts와 Strogatz(WS)의 연구

- 영화배우 네트워크, 전력망, 선충의 신경망에서의 평균 거리, 결집계수 측정
- 임의의 두 노드를 링크로 연결하는게 아니라 인접한 이웃의 노드들과 연결한 뒤 한 링크를 선택하고 한 노드를 고정한 채 다른 하나는 다른 노드로 재배열. 이 때 재배열되는 노드는 제한없이 선택되고 이미 연결된 노드는 배제.
- 실재 네트워크에서 얻은 결집계수 결과를 만족

5. 척도 없는 네트워크

도수의 분포가 멱함수를 따른다

5.1 월드 와이드 웹

- 나가는 링크 수의 분포함수의 지수는 계에 따라 달라지나 들어오는 링크 수의 분포함수의 지수는 γ ≒ 2.1로 거의 일정
- 임의의 두 웹 페이지의 평균 거리는 19(19번 클릭만하면 지구상의 모든 웹 페이지에 도달 가능)
    - 사실 월드 와이드 웹은 방향성이 있으므로 평균 거리는 엄밀하게 잘 정의할 수 없다
- 시간이 갈수록 나무구조 형태에서 고리가 있는 구조로 바뀌는 것이 변화하는 일반적인 패턴

5.2 인터넷

- Autonomous System(AS) : 대학교 외부와 통신하는 대표 라우터 레벨의 한 노드가 그 예
- AS를 노드로 갖는 네트워크에서의 링크 수의 분포함수의 지수는 γ ≒ 2.15 ~ 2.2
- 좀 더 낮은 단계의 라우터가 노드가 되는 네트워크에서는 γ ≒ 2.48
- Goh, et al. 의 연구에 따르면 인터넷 AS계의 노드 수 N은 N(t) = N(0)exp(a_1 * t)로 증가
- 링크 수도 마찬가지로 증가 (N(t) = N(0)exp(a_2 * t), a_1 ≒ 0.029(1), a_2 ≒ 0.034(2))
    - 가속 성장 네트워크
- 결집계수가 C ≒ 0.18 ~ 0.3 정도로 같은 노드 수, 링크 수로 만든 무작위 네트워크의 결집계수 C ≒ 0.001정도보다 매우 크다.
    - 집단구조를 이루고 있기 때문

5.3 영화배우 네트워크

- 영화배우와 그들이 출연한 영화가 함께 조합되어 이루는 네트워크
    - 두 종류의 노드로 구성(A형 : 영화배우, B형 : 영화)
- 다른 종류의 노드들끼리만 연결 → 이종분할 네트워크
- 노드 수가 많고 데이터들이 정확하기 때문에 사회연결망을 연구하는데 좋은 예
- 링크 수가 많은 노드가 존재하는 뚱뚱한 고리(fat-tailed)현상 존재 (P(k) ~ (k + k_0)^(-γ))

5.4 논문 인용 네트워크

- A라는 논문이 B라는 논문을 인용하게 되면 A → B 연결망이 구축된다
    - 방향이 있는 네트워크
- 노드 수가 증가하는 성장하는 네트워크
- 링크 수가 많은 노드는 오래된 논문일 것임을 간단히 예상할 수 있다

5.5 신진대사 반응 네트워크

- 생명체가 살아가는 데 필요한 에너지를 공급하기 위해 세포 속에서 작용하는 일련의 화학반응인 신진대사 반응을 네트워크로 표시한 것
- 이종분할 네트워크(A : 화학물, B : 화학반응의 촉매)

5.6 단백질-단백질 상호작용 네트워크

- 단백질이 노드, 두 단백질이 서로 상호작용을 하는 경우 링크 존재
- 뚱뚱한 고리(fat-tailed)현상 존재

6. 척도 없는 네트워크 모형

6.1 바라바시-알버트(BA) 모형

- 성장하는 네트워크
- 새로이 생겨난 노드는 m개의 링크를 기존 노드와 연결

6.2 적합성 모형

- BA모형 변형
- 링크 수가 큰 노드는 링크 수가 작은 노드와 연결되는 성질이 있어서 좀 더 실제 네트워크에 가깝다

6.3 후버만-아다미치 모형

- 노드들의 개수 N(t)는 시간에 대하여 지수적으로 증가
- 최근 이 모형을 기본으로 하여 인터넷의 진화현상을 설명하려는 연구가 수행됨

6.4 스태틱 모형

- ER의 무작위 네트워크 모형을 좀 더 일반화, 성장모형이 아님
- 가중치 네트워크
- 도수 k가 큰 영역에서 큰 영역에서도 도수 분포함수는 멱함수 꼴에 잘 맞는 장점을 가지고 있다
- 연결선수 간의 상관관계가 없는 경우이기 때문에 실재하는 네트워크의 성질 조사에 적합하지 않음

6.5 Configuration 모형

- 역시 ER의 무작위 네트워크 일반화
- 네트워크의 도수 분포에 따라 각 노드들에게 도수 배정

6.6 적응 모형

- 링크가 끊어지거나 새로 생김 → 재배열 현상

- 적응모형의 경우 재배열현상시 새로 연결된 노드의 도수가 이전의 노드의 도수보다 큼
    - 시간이 지남에 따라 더 좋은 환경으로 적응

6.7 단백질 상호작용 네트워크 모형

- 노드, 연결선 복제, 다변화, 돌연변이의 3가지 기작

6.8 사회네트워크에 대한 모형

- 총 노드 수가 고정되어 있다고 생각하는 것이 좋다
- 연결선수의 분포가 멱함수에 따르지 않고 최빈값이 존재
- 사람들 사이에서는 유유상종성, 결집성이 매우 높다
- 모든 사회연결망에서 나타나는 특징은 아님
    - 예 : 영화배우 네트워크
- 각 개인은 여러 그룹에 속해 있을 수 있다

7. 척도 없는 네트워크의 Ultrasmall 성질

- $2 < \gamma < 3$ 일 때 평균 거리 <D> 는 근사적으로 log log N인데 이는 log N보다 작게 축척되므로 이 때의 네트워크를 ultrasmall 네트워크라 부른다
- 도수가 가장 많은 허브를 중심에 두고 허브의 나뭇가지에 도수가 많은 순서로 각 노드를 붙여 나감
- 허브에서 최외각에 있는 노드까지의 거리가 가장 짧은 네트워크

8. 복잡계 네트워크에서 스핀 모형의 상전이 현상

8.1 평형 앙상블

8.2 Potts 모형

8.3 분배함수

8.4 Static 모형을 이용한 Potts 모형 상전이 현상

8.5 Helmholtz 자유에너지

9. 스미기 전이

10. 복잡계 네트워크의 견고성 및 연쇄적 파멸 현상

10.1 복잡계 네트워크의 견고성

- 의도적으로 외부에서 허브를 공격하는 경우 척도 없는 네트워크가 무작위 네트워크보다 취약하다
  - 신진대사 반응 네트워크나 단백질 상호작용 네트워크 등에 적용 가능

10.2 연쇄적 파멸 현상

- 척도 없는 네트워크에서는 두 노드 간의 거리가 매우 좁기 때문에 한 노드의 피해는 다른 노드에 곧바로 영향을 끼친다
  - 도미노 현상에 의한 피해
  - 예 : 2003년 8월 미국 동북부에서 일어난 전력망의 연쇄 정전 사태

10.3 모래사태 모형

- Bat-Tang-Wiesenfeld(BTW) 모형을 척도 없는 네트워크에 적용
- BTW 모형 : 모래가 쌓이는 경우를 예로 들면
  - 노드 i 에 쌓인 모래의 높이를 h_i라 하고 높이가 사태 문턱 z_i에 도달하면 모래가 무너져 이웃의 노드로 옮겨진다
  - 이 작업을 반복하면서 모든 노드들의 높이가 문턱 높이 이상이 되는 경우가 생길 수 있으므로 둘레가 되는 도수가 1인 노드에서 일정한 비율로 모래 알갱이를 제거시켜 준다
  - 외부로부터 모래가 주입되었을 때 불안정한 노드가 몇 개가 생기는가에 관심

11. 복잡계 네트워크에서의 수송 현상

인터넷에서의 패킷 전달, 세포 내의 신호 전달, 사람들 사이의 정보 전달 등

11.1 최단 경로 수송과 부하

의도된 수송 → 일반적으로 최단 경로 수송

From:
http://obg.co.kr/doku/ - **OBG WiKi**

Permanent link:
http://obg.co.kr/doku/doku.php?id=physics:network_theory

Last update: **2020/11/29 14:09**