

Uganda Village Health Community Network Analysis

박지유, 윤수빈, 이윤지 in 네트워크 과학

INTRODUCTION

우간다에는 4500만 명의 인구가 있지만, 농촌과 저소득 지역의 인구는 양질의 의료로 접하기 어렵다.

우간다에는 여전히 많은 의료 문제가 남아있으며, 많은 가정이 기본적인 의료 서비스를 받지 못하는 경우가 많다.

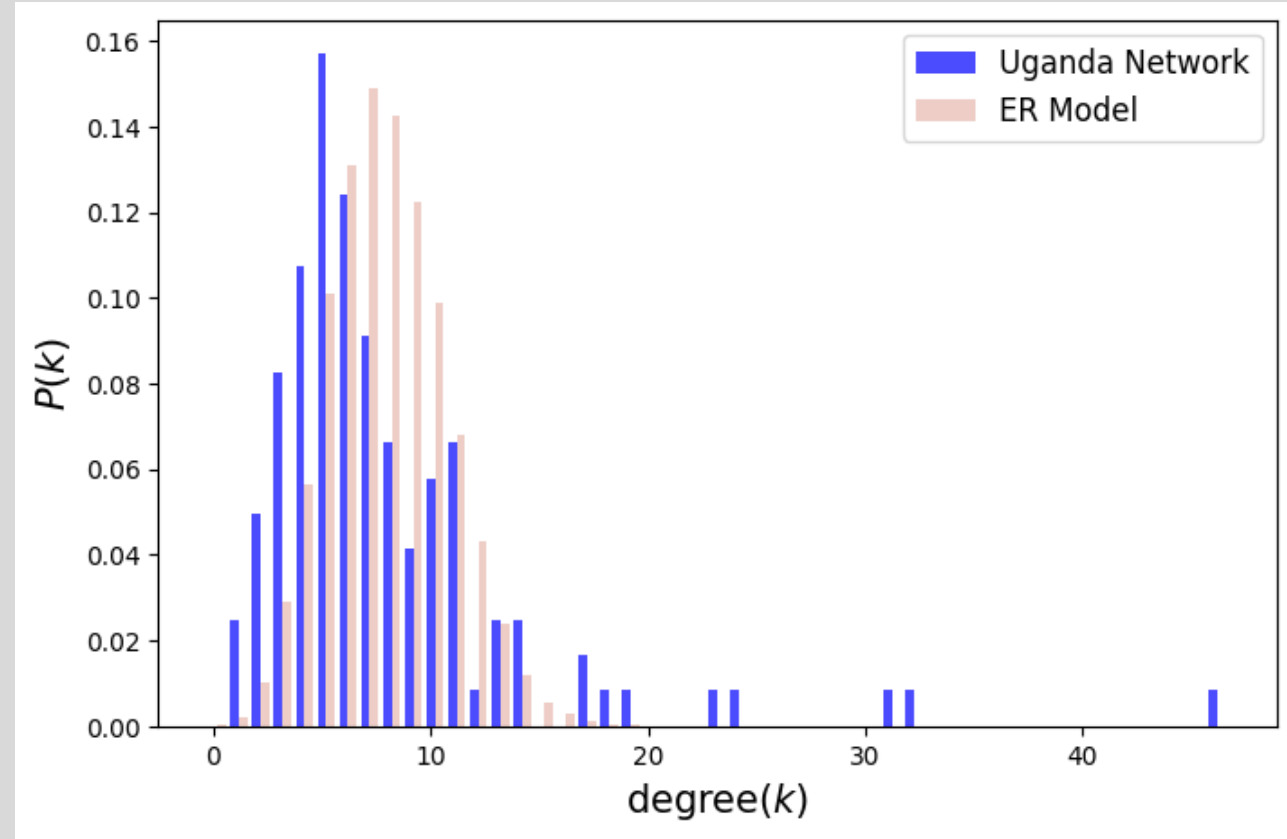
제한된 의료 접근에 우간다의 마을 보건팀은 가정과 공식 의료 서비스 사이의 연결 고리 역할을 한다.

효율적인 접근을 위해서 마을 이웃 들간의 건강 커뮤니티 분석을 필수적이다.

RESEARCH RESULTS (1)

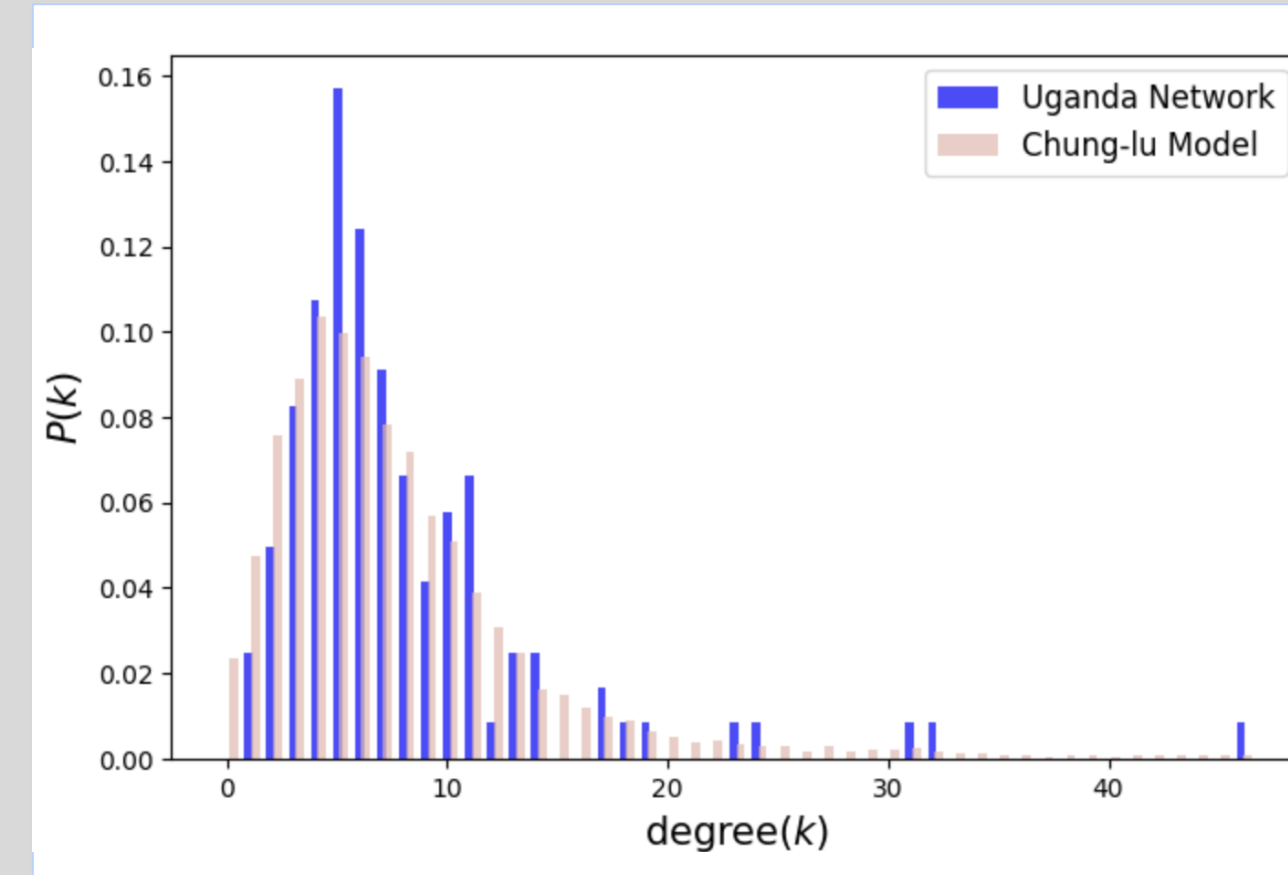
Github: <https://github.com/subinYoon/Uganda-health-trainers.git>

ER Histogram



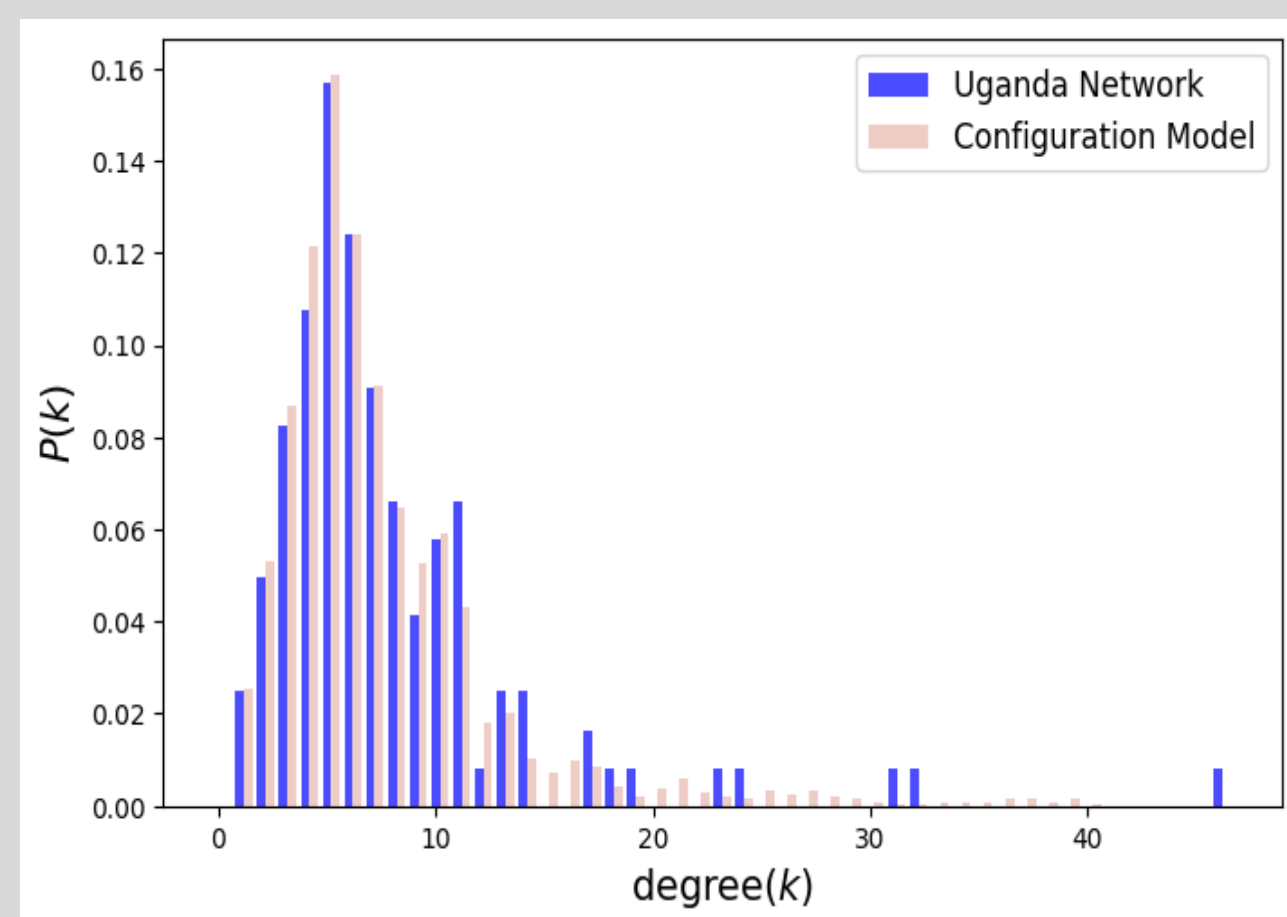
- 전체 차수가 평균 근처에 몰려 균일한 무작위 구조를 보임.
- 실제 네트워크의 허브 형태를 제대로 재현하지 못해 heavy-tailed한 구조가 약함.

Chung-Lu Histogram



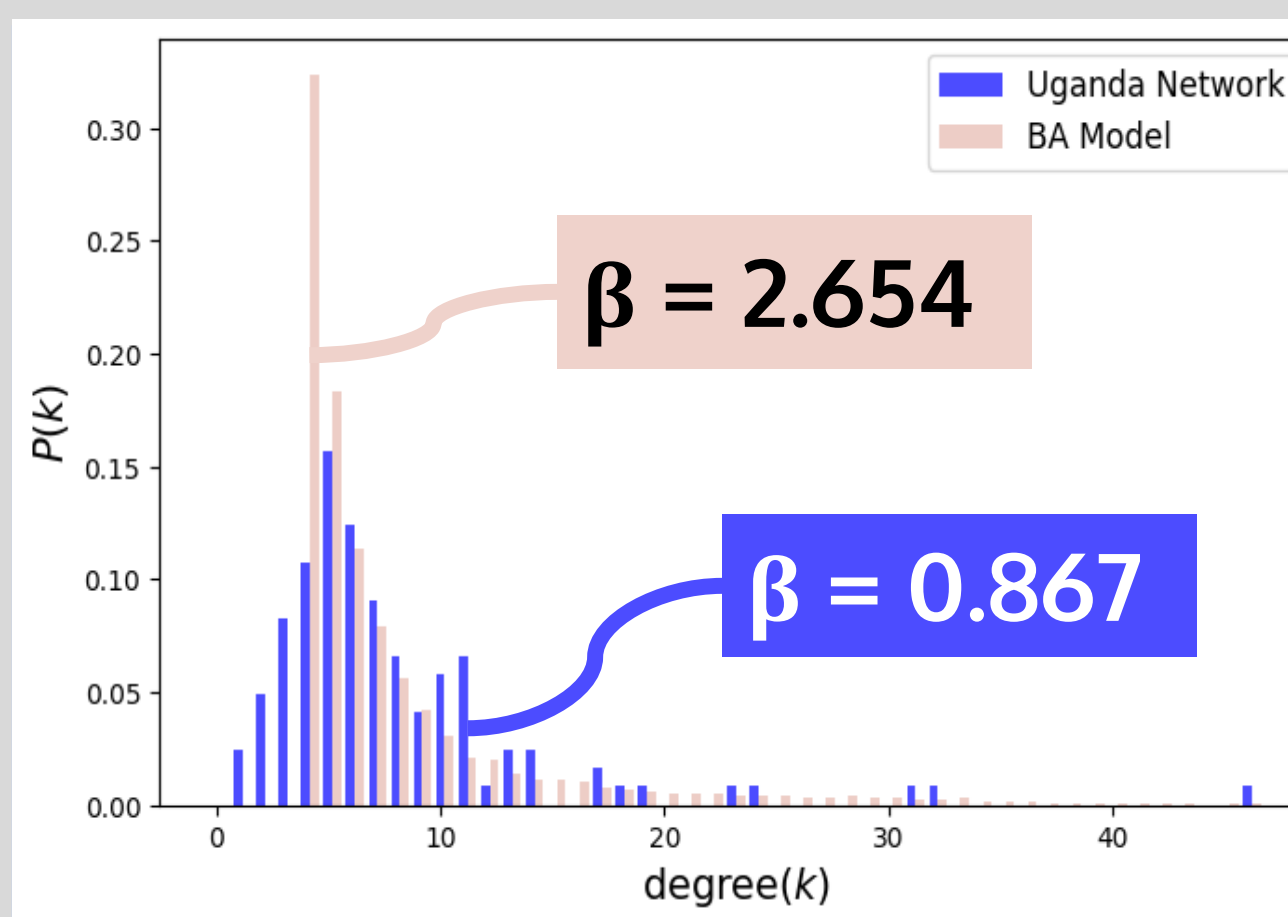
- 무작위 네트워크가 실제 네트워크보다 뭉개진 차수 분포를 가짐.
- 실제 네트워크의 특성에 따라 heavy-tailed한 구조를 보임.

Configuration Histogram



- 실제 네트워크와 무작위 네트워크 모두에서 차수 분포가 거의 동일하게 나타남.
- 모두에서 heavy-tailed한 차수 분포가 관찰됨.

BA Histogram

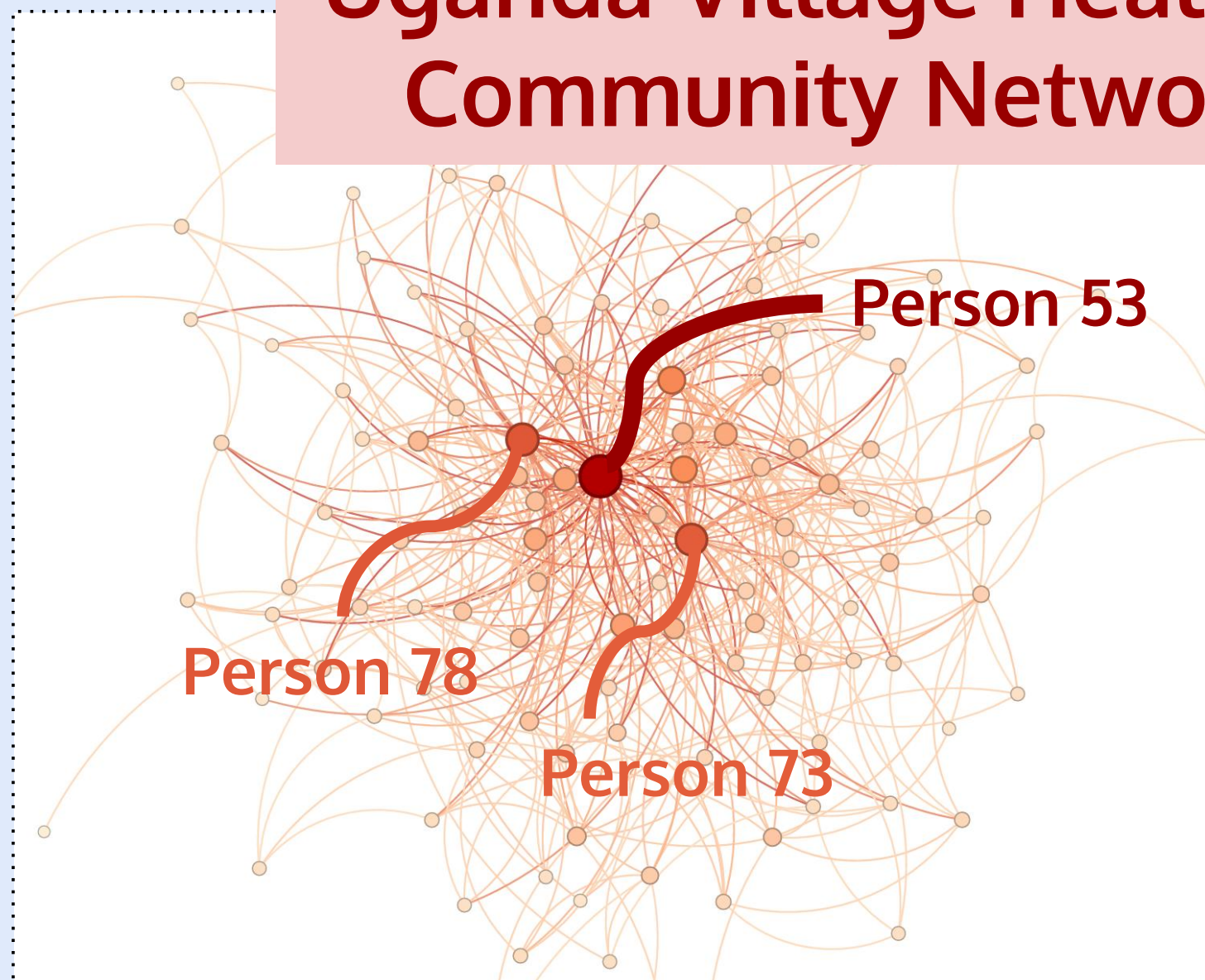


- 선호적 연결로 인해 허브가 과도하게 형성됨.
- BA β 값 ≈ 2.654 → 허브 과도 형성
- Original β 값 ≈ 0.87 → 실제 네트워크는 훨씬 완만한 power-law

Data Description

- 우간다 7 마을의 이웃들 간 건강 상담에 대한 데이터로 노드(node)의 개수는 121개, 엣지(edge)의 개수는 473개이다.
- 네트워크의 밀도(density)는 0.065 정도로 sparse한 편이다. 하지만 현실 사회 네트워크로서 매우 정상적인 값이다.
- 네트워크의 지름(dimaster)은 5로 이는 네트워크가 비교적 잘 연결되어 있음을 뜻하며, 우간다 이웃들 간의 건강 상담이 매우 연결성있음을 의미한다. (small world 특성)
- 차수 상관계수는 -0.12로, 고차수와 저차수 노드가 연결되는 경향성이 나타난다. 이는 일반적인 사회 네트워크에서 드문 패턴이다.
- 또한 삼각형(triangle)의 개수는 255개로, 겹치는 상담 관계가 많이 존재함을 보여준다.

Uganda Village Health Community Network



출처 : Chami, G. (2017). *Uganda Village Networks 2013*. Apollo - University of Cambridge Repository. <https://doi.org/10.17863/CAM.15616>

Analytical Measures

ER 모델

- 모든 노드 쌍이 동일한 확률 p 로 연결되는 가장 기본적인 무작위 네트워크 모델이다.
- 연결 패턴의 규칙성을 확인해 볼 수 있다.

Chung-Lu 모델

- 실제 네트워크의 차수 분포를 기댓값 형태로 유지하는 랜덤 모델이다.
- 원본 그래프의 속성이 차수 분포 때문에 나타나는지 확인 가능하다.

Degree Centrality

- 노드가 가진 연결의 수를 측정한다.
- 노드가 네트워크 내에서 얼마나 활발하게 연결되어 있는지를 나타낸다.

Configuration 모델

- 실제 네트워크의 정확한 차수 분포를 유지하면서, 연결 정보는 무작위 재배치 생성하는 랜덤 모델이다.
- 실제 네트워크의 구조가 단순한 차수 분포로 설명 가능한지 볼 수 있다.

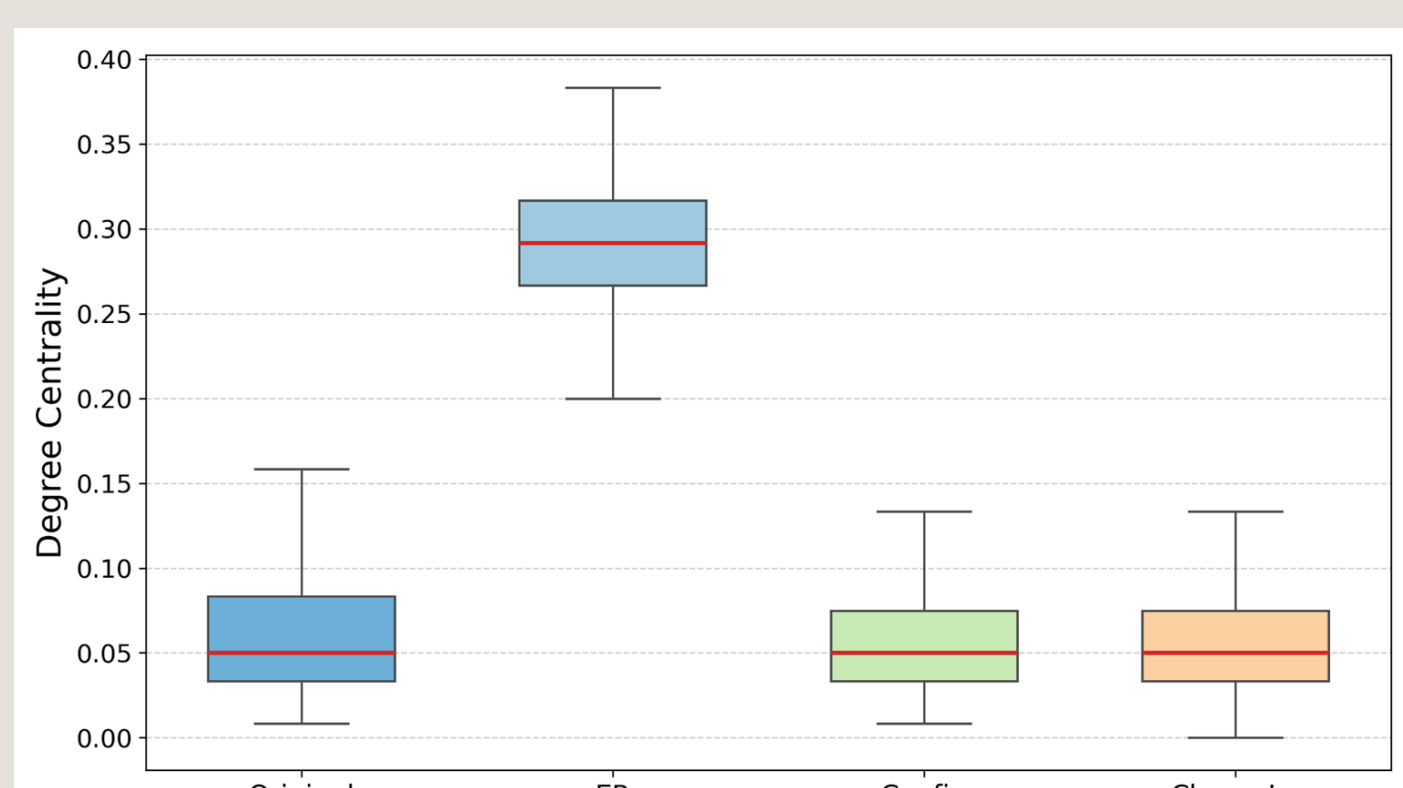
BA 모델

- 성장 + 선호적 연결을 따르는 모델이다.
- 실제 네트워크에서 보이는 허브 구조가 이런 원리로 생긴 것인지 확인할 수 있다.

Closeness Centrality

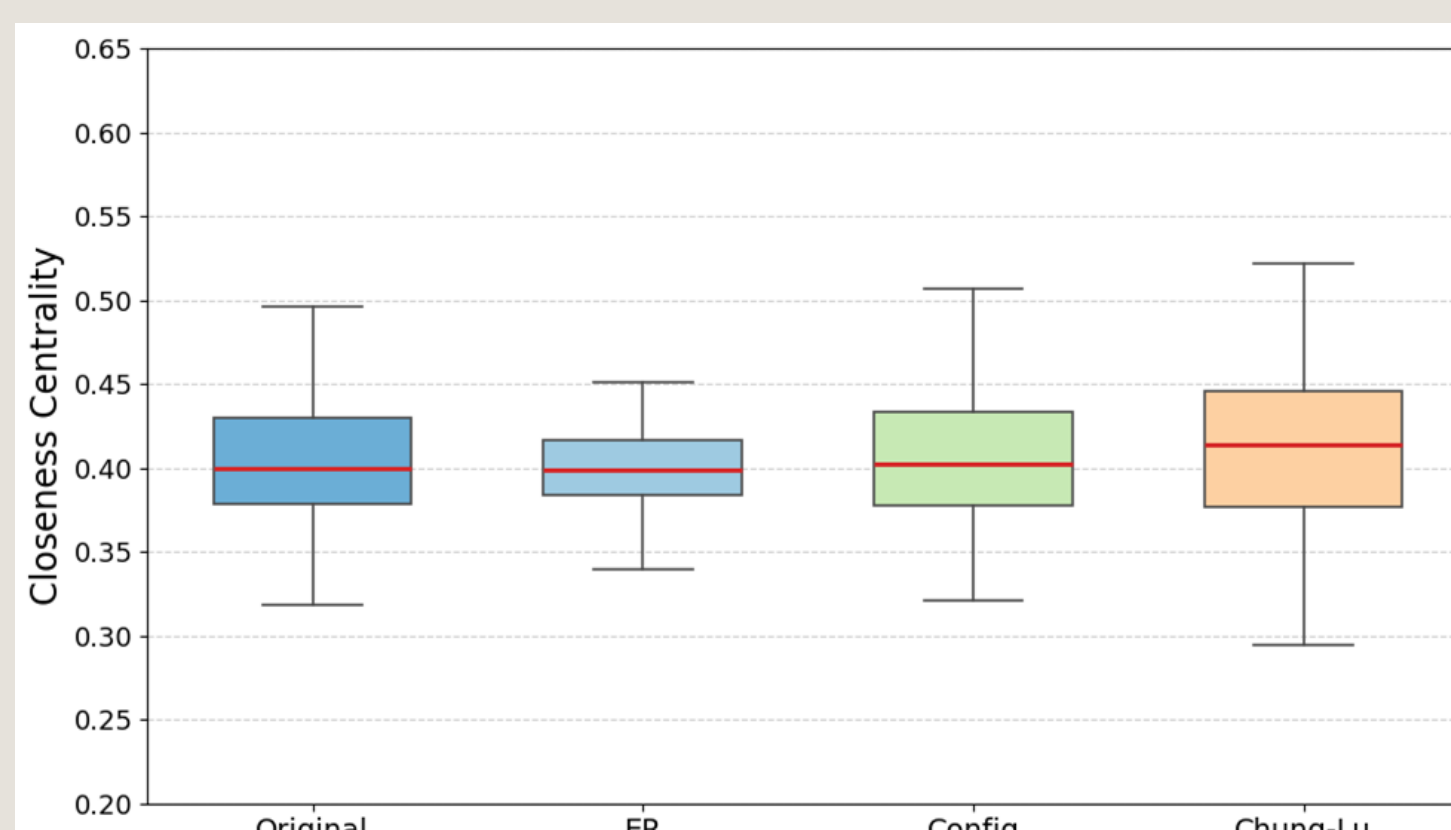
- 특정 노드가 다른 노드들에게 얼마나 빠르게 접근할 수 있는지 나타내는 지표이다.
- 정보를 빠르게 확산시킬 수 있는 중심 노드를 찾을 수 있다.

RESEARCH RESULTS (2)



Degree Centrality

- Chung-Lu model과 Configuration model이 낮은 중앙값과 넓은 범위로 원본 그래프와 유사한 분포를 보임.
- ER model은 무작위 연결로 인해 모든 노드가 균등한 연결성을 가지므로 다른 모델보다 높은 중앙값을 가짐.



Closeness Centrality

- Configuration 모델은 차수 분포와 연결 구조를 비슷하게 재현해 원본 네트워크와 가장 유사한 closeness 분포를 보인다.
- ER과 Chung-Lu 모델은 허브 구조나 연결 패턴을 재현하지 못해 실제 거리 구조를 나타내지 못함.

CONCLUSION

- Configuration 모델이 세 모델 중 실제 네트워크의 구조를 잘 나타내는 것으로 보아, 차수 분포가 핵심 요인임을 확인했다.
- 또한 연결 방식은 선호적이기 보다 무작위성에 의해 형성되는 경향을 보였다.
- 이와 함께, 네트워크에는 뚜렷한 허브의 존재를 확인하였다.