Aalto-yliopisto
Perustieteiden
korkeakoulu

Eksakti rekonstruktio stokastisessa lohkomallissa havaintomatriisin rivisummien avulla (aihe-esittely)

Mikael Niukkanen

06.11.2026

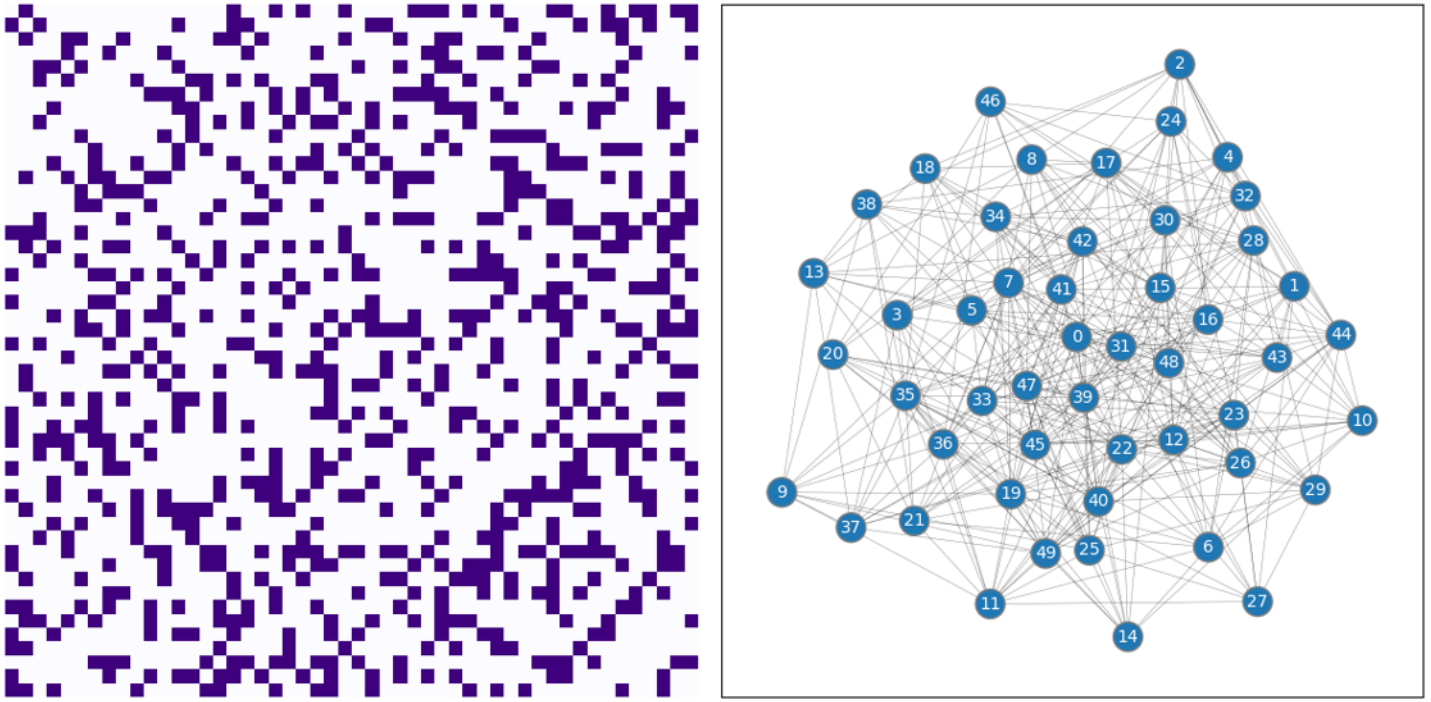
Ohjaaja: *prof. Lasse Leskelä*

Valvoja: *prof. Lasse Leskelä*

Työtä ei saa tallentaa tai julkistaa Aalto-yliopiston avoimilla verkkosivuilla. Kaikki oikeudet pidätetään.

Aalto-yliopisto
Perustieteiden
korkeakoulu

Esimerkki: erään satunnaisverkon toteuma



Kuva: Eric Bridgeford, Alex Loftus, and Joshua Vogelstein
© Copyright 2022. https://docs.neurodata.io/graph-stats-book/representations/ch5/single-network-models_ER.html

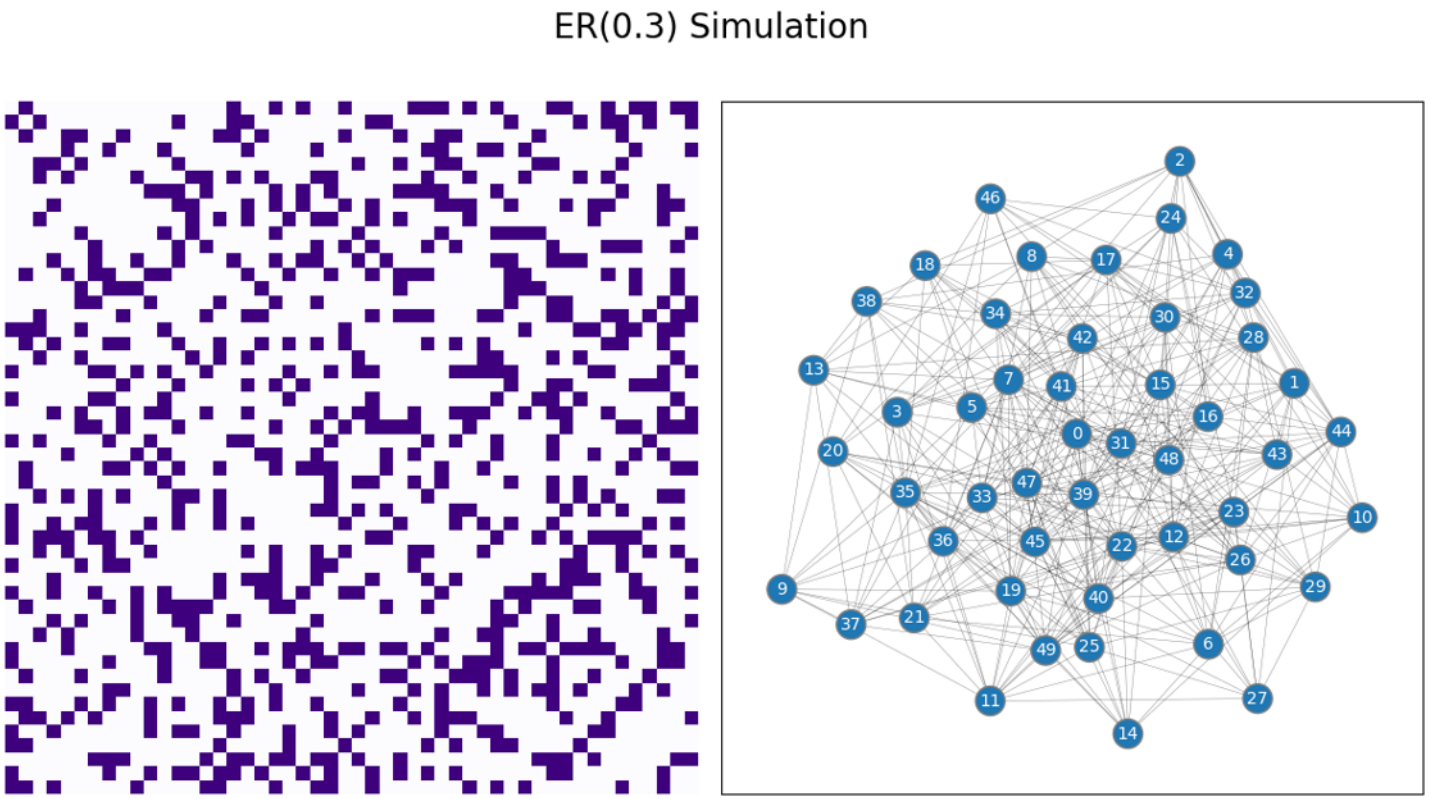
Satunnaisverkot ja satunnaisverkkomallit

- Satunnaisverkot
  - Verkkoarvoisia satunnaismuuttujia (vrt. reaaliarvoinen tai vektoriarvoinen). Tässä esittelyssä vain suuntaamattomia verkkoja
- Satunnaisverkkomallit
  - Tiettyjä satunnaisverkkojen jakaumia
  - esimerkkejä:
    - Erdős–Rényi-malli, merkitään ER(n, p)
      - n solmujen lukumäärä, p linkkitodennäköisyys
    - Stokastinen lohkomalli, merkitään SBM(n, π, Q)
      - Kullakin solmulla oma *yhteisö*
      - n solmujen lukumäärä, π yhteisötodennäköisyydet, Q linkkitodennäköisyydet

Esimerkki: Erdős–Rényi-malli

- Verkon solmut linkittyvät satunnaisesti ja riippumattomasti
- ER(n, p): n solmua, linkkitodennäköisyys p
- Mahdollinen kysymys: milloin G ~ ER(n, p) on yhtenäinen (*connected*)?
- Vastaus: G asymptoottisesti melkein varmasti yhtenäinen, kun $p > \frac{(1+\varepsilon)\ln n}{n}$ eli kun keskimääräinen aste kasvaa O(log(n)). Toisaalta melkein varmasti epäyhtenäinen, kun $p < \frac{(1-\varepsilon)\ln n}{n}$.

Toteuma mallista ER(n = 50, p = 0.3)

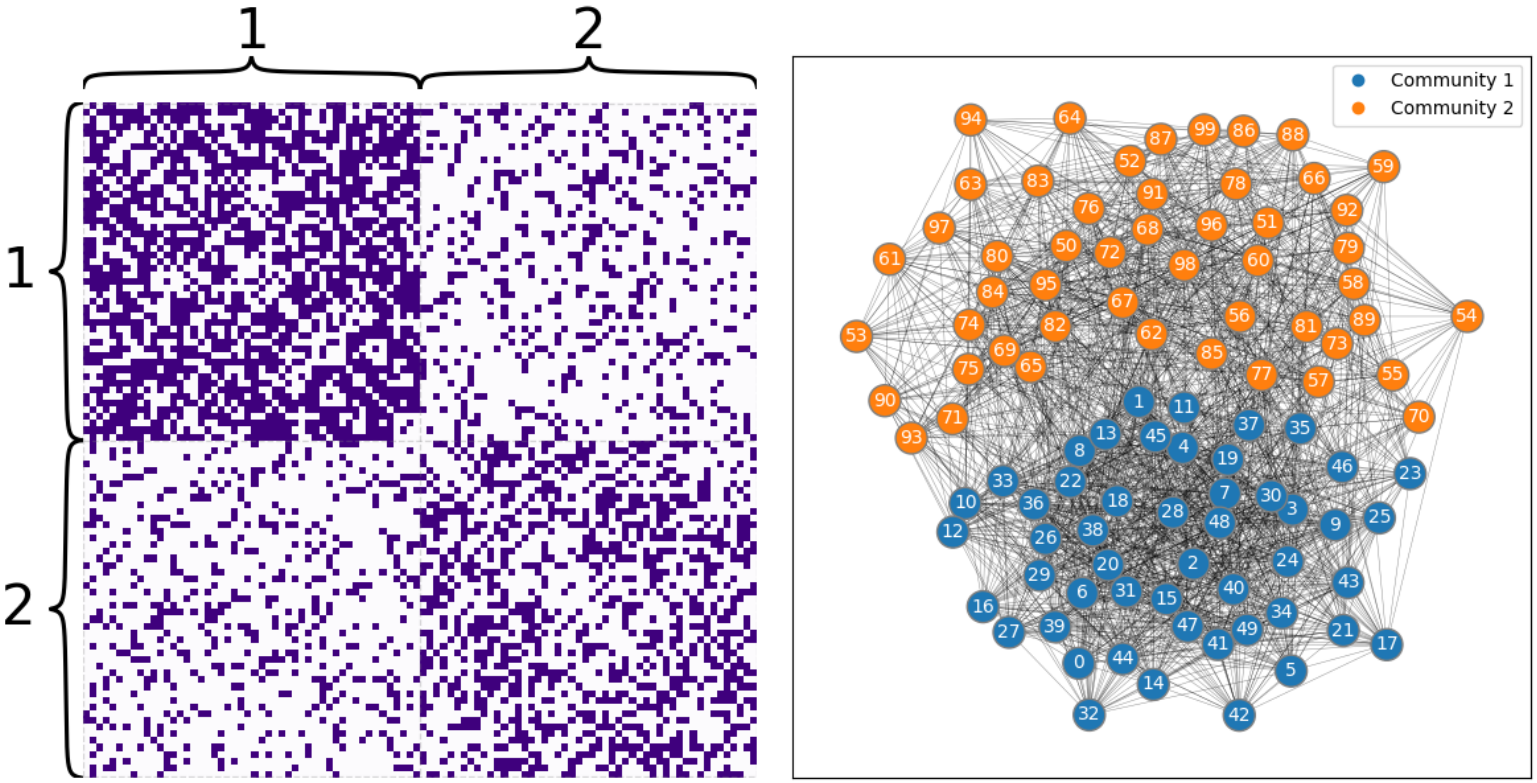


Kuva: Eric Bridgeford, Alex Loftus, and Joshua Vogelstein
© Copyright 2022. https://docs.neurodata.io/graph-stats-book/representations/ch5/single-network-models_ER.html

Stokastinen lohkomalli

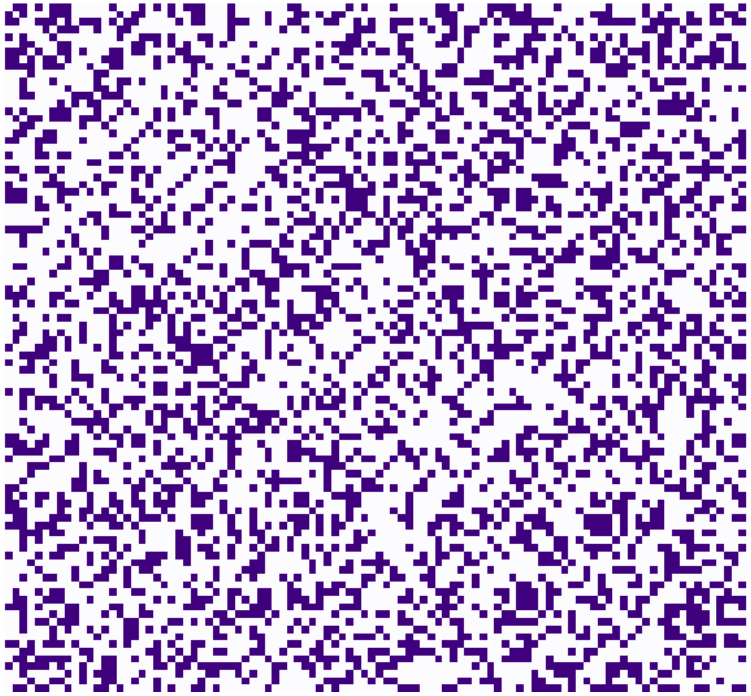
- Suuntaamattoman verkon solmut linkittyvät satunnaisesti ja riippumattomasti
- Verkossa *k* yhteisöä, yhteisöt osittavat solmujoukon
  - π_i todennäköisyydellä solmu u yhteisössä C_i
  - Yhteisövektorilla σ merkitään solmujen yhteisöindeksit
    - $\sigma_u = i,$ kun solmu $u \in C_i$
- Solmujen *i,j* linkkitodennäköisyys riippuu niiden yhteisöistä
  - **Pr**[{*u,v*} ∈ E] = $\mathbf{Q}_{\sigma(u)\sigma(v)}$
  - $\mathbf{Q}$ on $k \times k$ matriisi, joka sisältää linkkitodennäköisyydet yhteisöstä toiseen

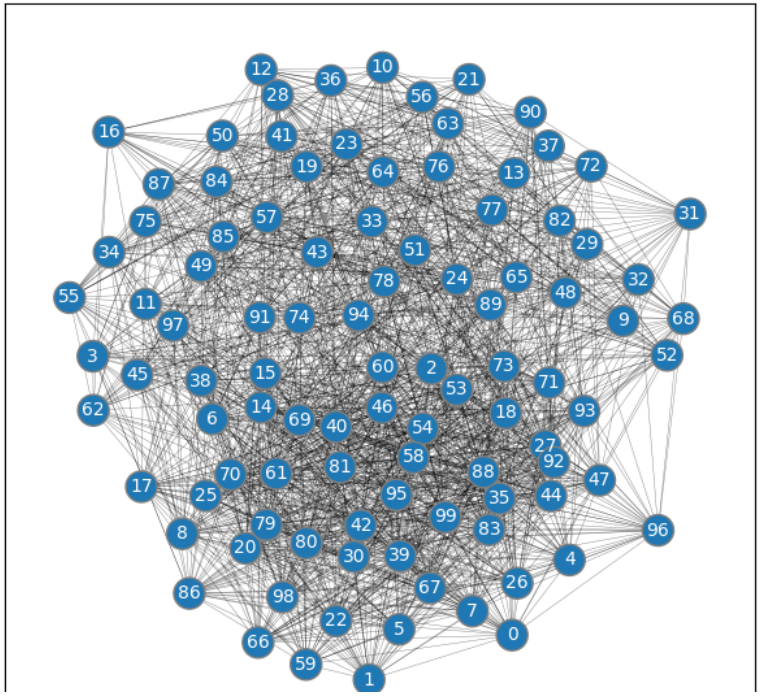
Toteuma mallista SBM(n=100, π = [0.5, 0.5], Q = [0.6, 0.2; 0.2, 0.4]) (1/2)



Kuva: Eric Bridgeford, Alex Loftus, and Joshua Vogelstein
© Copyright 2022. https://docs.neurodata.io/graph-stats-book/representations/ch5/single-network-models_SBM.html

Toteuma mallista SBM(n=100, π = [0.5, 0.5], Q = [0.6, 0.2; 0.2, 0.4]) (2/2)





Kuva: Eric Bridgeford, Alex Loftus, and Joshua Vogelstein
© Copyright 2022. https://docs.neurodata.io/graph-stats-book/representations/ch5/single-network-models_SBM.html

Yhteisöntunnistus ja eksakti rekonstruktio

- Yhteisöntunnistus
  - Yritetään rekonstruoida verkon yhteisörakenne σ
- Eksakti rekonstruktio
  - Yhteisörakenteen rekonstruktio $\hat{\sigma}$ on virheetön
  - $\Pr[\hat{\sigma} = \sigma] \rightarrow 1$ kun $n \rightarrow \infty$
  - tunnettuja tuloksia

Tunnettu tulos

- Eksaktin rekonstruktion ehto tunnetaan SBM:n yleisessä tapauksessa (Abbe, Sandon 2015)
- Ehto on informaatioteoreettinen: kertoo, milloin optimaalinen algoritmi kykenee eksaktiin rekonstruktioon ja milloin ei
- Ehdon oletus: algoritmilla on käytössään verkon vierusmatriisi (tai vastaava informaatio)

Tutkimusaukko

- Entä jos verkosta havaitaan vain solmujen asteluvut (vierusmatriisin rivisummat)?
- Kirjallisuudessa tilannetta tarkasteltu vain vähän
- Huomio: jos yhteisöjen odotetut asteet ovat samat, asteet eivät erottele niitä lainkaan

Tavoitteet ja rajaukset

- Tavoite
  - Johtaa ehto eksaktille rekonstruktiolle, kun verkosta havaitaan vain asteluvut
  - Kirjallisuuskatsaus eksaktiin rekonstruktioon yleisessä tapauksessa
  - Mahdollinen simulaatiotutkimus
- Rajaukset
  - Yhteisöjä 2-3 kappaletta
  - Asteluvut skaalautuvat O(log(n)) – verkko on yhtenäinen

Tietolähteet ja menetelmät

- Tietolähteet
  - Emmanuel Abbe ja Colin Sandon. Community detection in general stochastic block models: fundamental limits and efficient algorithms for recovery. IEEE, 2015.
  - Muu aiheeseen liittyvä kirjallisuus
- Menetelmät
  - Approksimaatiot
    - Asteluvut likimain riippumattomia
    - Astelukujen approksimointi Poisson-jakaumalla
    - Poisson-jakaumien häntätodennäköisyyksien approksimointi
  - Aiemmat tulokset soveltuvin osin

Aikataulu

- Kirjallisuuskatsaus 1-2 kk
- Ehdon johtaminen eksaktille rekonstruktiolle 1 kk
- Työn ensimmäisen version kirjoittaminen 3 kk
- Työn unohtaminen arjen keskellä N kk
- Työn viimeistely 1-2 vk
- Päätulosten esittäminen 27.11.2026