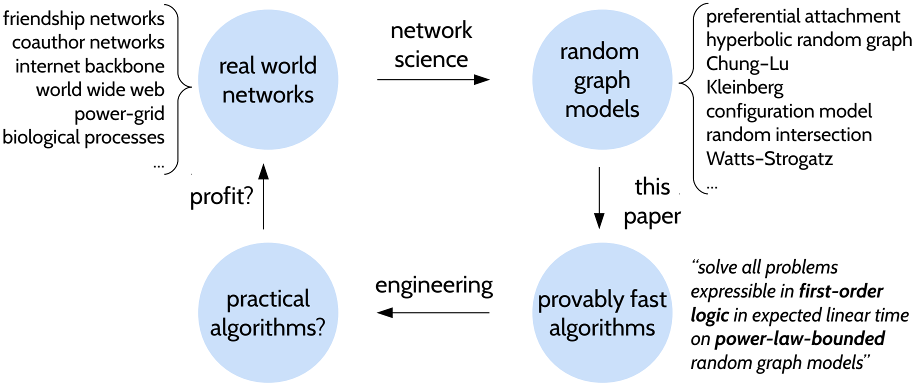
First-Order Model-Checking in Random Graphs and Complex Networks

Jan Dreier, Philipp Kuinke, Peter Rossmanith

Theoretical Computer Science
RWTH Aachen University, Germany

ESA 2020

Motivation


friendship networks
coauthor networks
internet backbone
world wide web
power-grid
biological processes
...
real world networks
network science
random graph models
preferential attachment
hyperbolic random graph
Chung-Lu
Kleinberg
configuration model
random intersection
Watts-Strogatz
...
this paper
provably fast algorithms
"solve all problems expressible in **first-order logic** in expected linear time on **power-law-bounded** random graph models"
engineering
practical algorithms?
profit?

Logic

Many problems can be expressed in first-order (FO) logic, e.g.,

- independent set of size k:

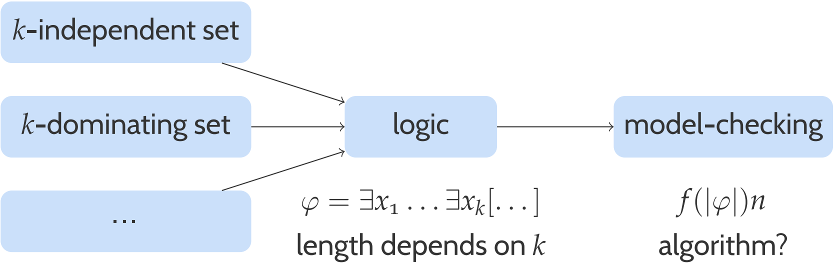
$$\exists x_1 \ldots \exists x_k \bigwedge_{i,j} x_i \neq x_j \wedge x_i \not\sim x_j$$

- dominating set of size k:

$$\exists x_1 \ldots \exists x_k \forall y \bigvee_{i} y \sim x_i \vee y = x_i$$

Can be solved in $n^{O(k)}$ (ETH: essentially optimal)

Model-Checking

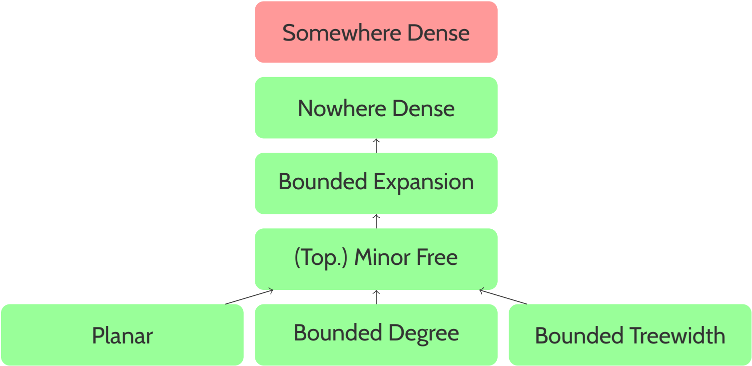


model-checking problem

Input:	graph G and a sentence φ
Parameter:	$\lvert\varphi\rvert$
Problem:	is φ true in G?

Goal: linear fpt run time $f(|\varphi|)n$

Sparse Graph Classes



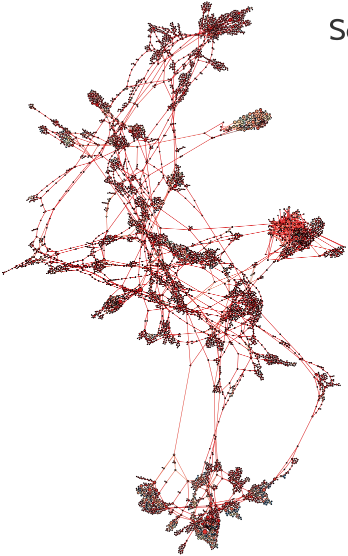
MSO model-checking on bounded treewidth in $f(|\varphi|)n$.

[Courcelle 1990]

FO model-checking on nowhere dense graph classes in $f(|\varphi|)n^{1+\varepsilon}$.

[Grohe, Kreutzer, Sieberz 2011]

The Real World



Some central properties:

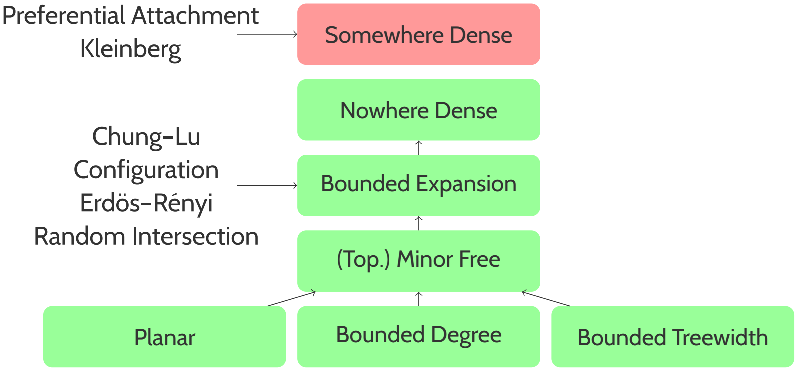
- **Skewed degree distribution**
  Fraction of vertices with degree k proportional to $k^{-\alpha}$ with $2 \leq \alpha \leq 3$
- **Clustered**
  If we have a common friend we are likely friends as well
- **Small-world property**
  Everyone is close to everyone

Random Graph Models

Random graph model: probability distribution over graphs

- Preferential attachment model
- Hyperbolic random graph model
- Chung–Lu model
- Erdös–Rényi model
- Configuration model
- Random intersection graph model
- Watts–Strogatz model
- …

Previous Results



[Grohe 2001], [Farrell et. al. 2015], [Demaine et. al. 2019], [Dreier et. al. 2020]

Our Result

A random graph model is *3-power-law-bounded* if (roughly speaking):

- fraction of vertices with degree k is $O(k^{-3})$
  real networks: typically $k^{-\alpha}$ with $2 \leq \alpha \leq 3$

optimal

- unclustered
  real networks: typically clustered

???

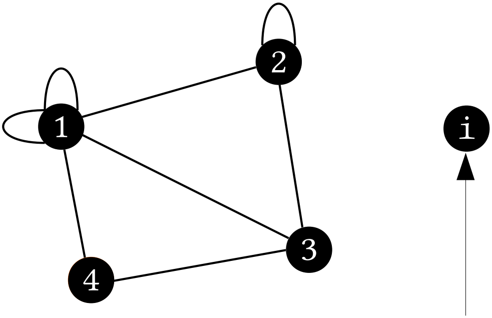
Theorem

Given a first-order sentence φ and a graph G sampled from a 3-power-law-bounded model, one can decide whether φ is true on G in expected time $f(|\varphi|)n^{1+\varepsilon}$ for every $\varepsilon > 0$.

Big Question: model-checking on clustered models?

Example: Preferential Attachment Model

Introduced by Barabási and Albert in 1999 to explain the structure of the world wide web.

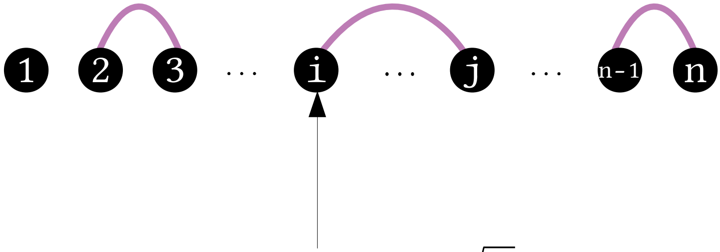


$$\text{expected degree} \approx \sqrt{\frac{n}{i}}$$

Example: Chung–Lu Model

A more direct way to get a fixed degree distribution.

$$\frac{1}{\sqrt{2\cdot 3}} \qquad \frac{1}{\sqrt{i\cdot j}} \qquad \frac{1}{\sqrt{(n-1)\cdot n}}$$



$$\text{expected degree} \approx \sqrt{\frac{n}{i}}$$

α-power-law-boundedness

A random graph model with vertices $1, \ldots, n$ is 3-*power-law-bounded* if the probability that some subset of edges $E \subseteq \binom{1,\ldots,n}{2}$ is present is at most

$$\log(n)^{O(|E|^2)} \prod_{ij \in E} \frac{1}{\sqrt{i \cdot j}}.$$

α-power-law-boundedness

A random graph model with vertices $1, \ldots, n$ is 3-*power-law-bounded* if the probability that some subset of edges $E \subseteq \binom{1,\ldots,n}{2}$ is present is at most

$$\log(n)^{O(|E|^2)} \prod_{ij \in E} \frac{1}{\sqrt{i \cdot j}}.$$

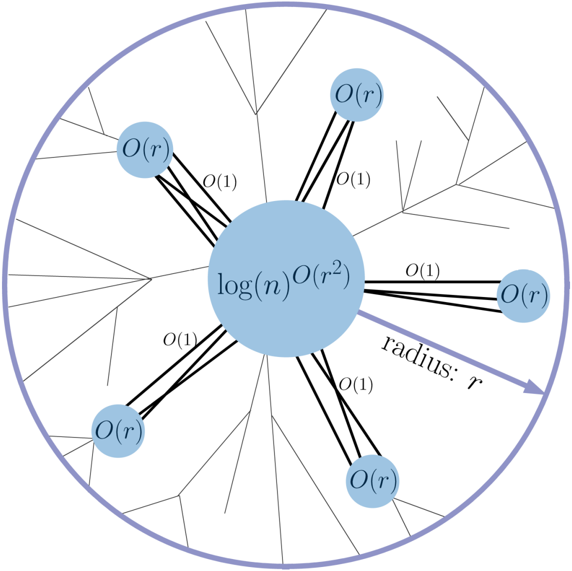
✓

- Preferential attachment model
- Chung–Lu model
- Erdös–Rényi model
- Configuration model
- …

✗

- Hyperbolic random graph model
- random intersection model
- Watts–Strogatz model
- Kleinberg model
- …

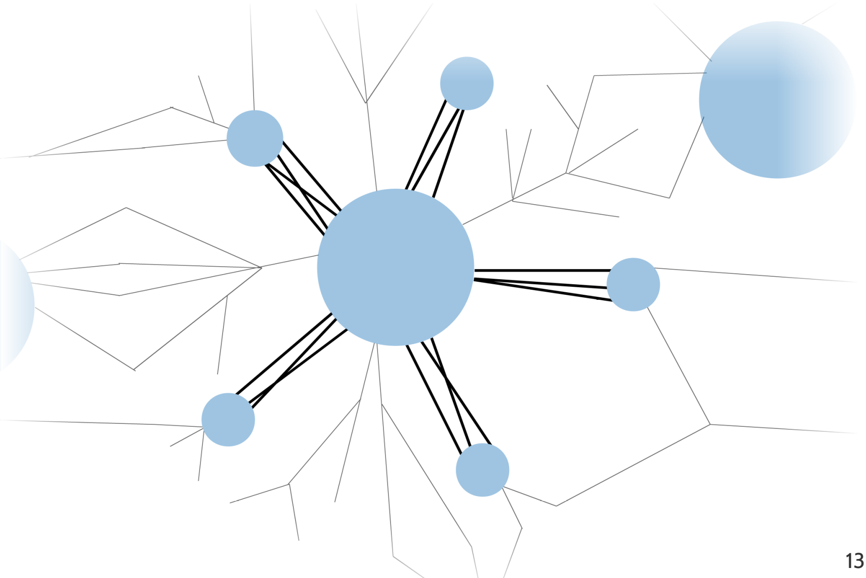
Asymptotic Structure of 3-power-law-bounded models



Algorithm

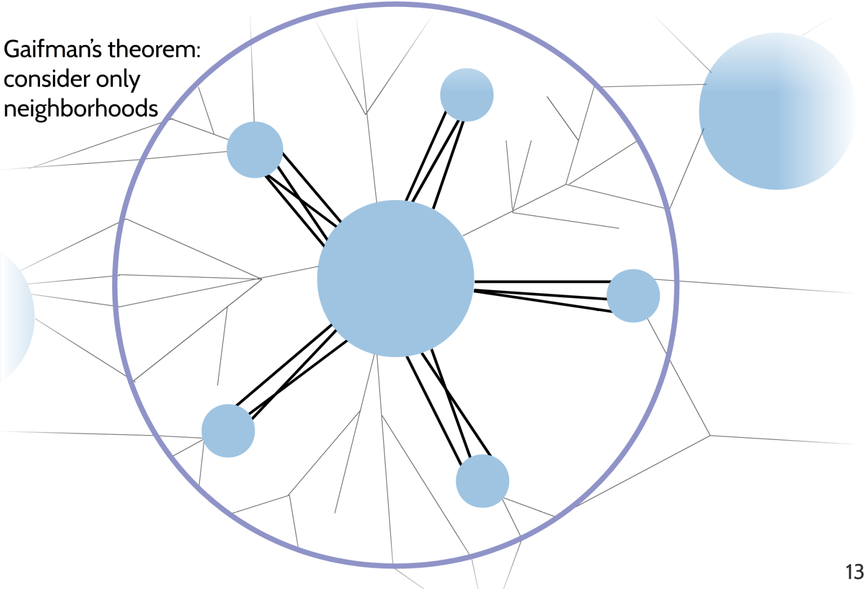
Input: graph sampled from 3-power-law-bounded model

Algorithm



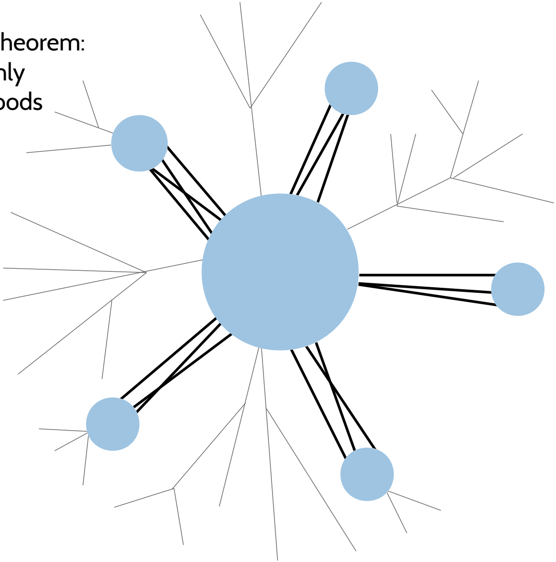
Algorithm

Gaifman's theorem:
consider only
neighborhoods



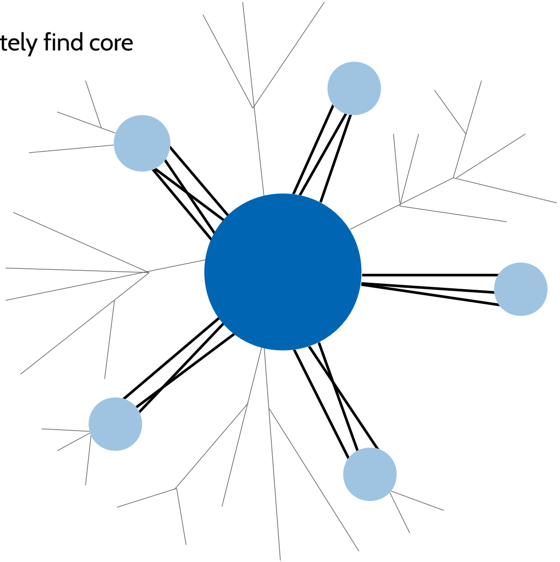
Algorithm

Gaifman's theorem:
consider only
neighborhoods



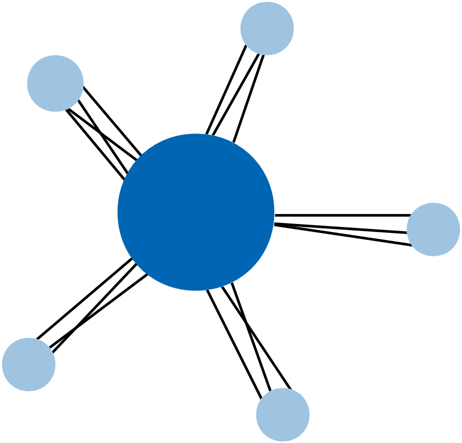
Algorithm

approximately find core



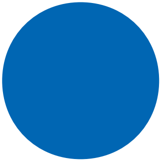
Algorithm

prune trees



Algorithm

prune protrusions

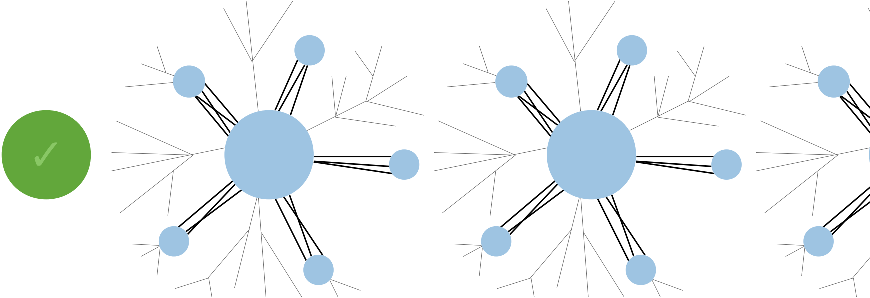


Algorithm

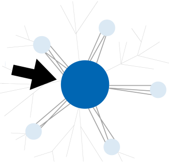
use brute force on core

Algorithm

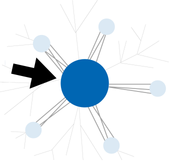
repeat for every neighborhood



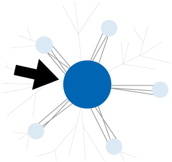
Proof Sketch

Runtime: $n^{O(1)} \cdot \sum_{x=1}^{n} \Pr\big[$  $= x\big] \cdot x^{|\varphi|}$

To get a run time of $f(|\varphi|)n^{O(1)}$ we bound

$\Pr\big[$  $\geq x\big]$ for every x.

Proof Sketch

- Assume core  is large in some neighborhood.
- 
- Then there is an edge set $E \subseteq \binom{1,\ldots,n}{2}$ from a collection Π of "blocking sets" such that $E \subseteq E(G)$.
- By definition of 3-power-law-boundedness:
  $\Pr[E \subseteq E(G)] \leq \log(n)^{O(|E|^2)} \prod_{ij \in E} \frac{1}{\sqrt{i \cdot j}}$.
- Union bound: Any "blocking set" form Π present with probability is at most
  $\sum_{E \in \Pi} \Pr[E \subseteq E(G)] \leq \sum_{E \in \Pi} \log(n)^{O(|E|^2)} \prod_{ij \in E} \frac{1}{\sqrt{i \cdot j}}$.

Summary

A random graph model is *3-power-law-bounded* if (roughly speaking):

- fraction of vertices with degree k is $O(k^{-3})$
  real networks: typically $k^{-\alpha}$ with $2 \leq \alpha \leq 3$

optimal

- unclustered
  real networks: typically clustered

???

Theorem

Given a first-order sentence φ and a graph G sampled from a 3-power-law-bounded model, one can decide whether φ is true on G in expected time $f(|\varphi|)n^{1+\varepsilon}$ for every $\varepsilon > 0$.

Big Question: model-checking on clustered models?