Proving logarithmic upper bounds for diameters of random graphs

Abbas Mehrabian

Pacific Institute for the Mathematical Sciences

CanaDAM 2015
University of Saskatchewan

Motivation

Many random graphs have diameters at most $O(\log n)$ ...

Generator	Diameter or Avg path len.	Community		Clustering coefficient	Remarks
		Bip. core vs size	$C(k)$ vs k		
Erdös–Rényi [1960]	$O(\log N)$		Indep.	Low, $CC \propto N^{-1}$	
PLRG [Aiello et al. 2000], PLOD [Palmer and Steffan 2000]	$O(\log N)$	Indep.		$CC \to 0$ for large N	
Exponential cutoff [Newman et al. 2001]	$O(\log N)$			$CC \to 0$ for large N	
BA [Barabási and Albert 1999]	$O(\log N)$ or $O(\frac{\log N}{\log \log N})$			$CC \propto N^{-0.75}$	
Initial attractiveness [Dorogovtsev and Mendes 2003]					
AB [Albert and Barabási 2000]					
Edge copying [Kleinberg et al. 1999], [Kumar et al. 1999]		Power-law			
GLP [Bu and Towsley 2002]				Higher than AB, BA, PLRG	Internet only
Accelerated growth [Dorogovtsev et al. 2001], [Barabási et al. 2002]				Non-monotonic with N	
Fitness model [Bianconi and Barabási 2001]					
Aiello et al. [2001]					
Pandurangan et al. [2002]					
Inet [Winick and Jamin 2002]					Specific to the AS graph
Forest Fire [Leskovec et al. 2005]	"shrinks" as N grows				
Pennock et al. [2002]					
Small-world [Watts and Strogatz 1998]	$O(N)$ for small N, $O(\ln N)$ for large N, depends on p			$CC(p) \propto (1-p)^3$, Indep of N	N = num nodes p = rewiring prob
Waxman [1988]					
BRITE [Medina et al. 2000]	Low (like in BA)			like in BA	BA + Waxman with additions
Yook et al. [2002]					
Fabrikant et al. [2002]					Tree, density 1
R-MAT [Chakrabarti et al. 2004]	Low (empirically)				

Chakrabarti & Faloutsos, ACM Computing Surveys 2006.

Motivation

Many random graphs have diameter $O(\log n)$ ...

1. Why is it important?
2. What is the reason?
3. How to prove such results?

Motivation

Many random graphs have diameter $O(\log n)$ ...

1. Why is it important?
2. What is the reason?
3. How to prove such results?

Expansion?

Motivation

Many random graphs have diameter $O(\log n)$ ...

1. Why is it important?
2. What is the reason?
3. How to prove such results?

Expansion? Works in many cases!

But, many random trees also have logarithmic diameter!

Our contribution

A technique for proving certain random graphs have diameter at most $O(\log n)$.

Note: no lower bounds today!

Generator	Diameter or Avg path len.	Community		Clustering coefficient	Remarks
		Bip. core vs size	$C(k)$ vs k		
Erdős–Rényi [1960]	$O(\log N)$		Indep.	Low, $CC \propto N^{-1}$	
PLRG [Aiello et al. 2000], PLOD [Palmer and Steffan 2000]	$O(\log N)$	Indep.		$CC \to 0$ for large N	
Exponential cutoff [Newman et al. 2001]	$O(\log N)$			$CC \to 0$ for large N	
BA [Barabási and Albert 1999]	$O(\log N)$ or $O(\frac{\log N}{\log \log N})$			$CC \propto N^{-0.75}$	
Initial attractiveness [Dorogovtsev and Mendes 2003]					
AB [Albert and Barabási 2000]					
Edge copying [Kleinberg et al. 1999], [Kumar et al. 1999]	$O(\log N)$	Power-law			
GLP [Bu and Towsley 2002]	$O(\log N)$			Higher than AB, BA, PLRG	Internet only
Accelerated growth [Dorogovtsev et al. 2001], [Barabási et al. 2002]				Non-monotonic with N	
Fitness model [Bianconi and Barabási 2001]					
Aiello et al. [2001]	$O(\log N)$				
Pandurangan et al. [2002]	$O(\log N)$				
Inet [Winick and Jamin 2002]					Specific to the AS graph
Forest Fire [Leskovec et al. 2005]	$O(\log N)$				
Pennock et al. [2002]					
Small-world [Watts and Strogatz 1998]	$O(N)$ for small N, $O(\ln N)$ for large N, depends on p			$CC(p) \propto (1-p)^3$, Indep of N	N = num nodes p = rewiring prob
Waxman [1988]					
BRITE [Medina et al. 2000]	Low (like in BA)			like in BA	BA + Waxman with additions
Yook et al. [2002]					
Fabrikant et al. [2002]					Tree, density 1
R-MAT [Chakrabarti et al. 2004]	Low (empirically)				

Chakrabarti & Faloutsos, ACM Computing Surveys 2006.

Our contribution

A technique for proving certain random graphs have diameter at most $O(\log n)$.

Our contribution

A technique for proving certain random graphs have diameter at most $O(\log n)$.
The idea is to couple with a random recursive tree such that graph's diameter $\leq$ tree's diameter.

Definition (random recursive tree)

Fix k. Initially we have a single node;
in every round a uniformly random node gives birth to k new children.

Pittel'94 proved a.a.s. the height is asymptotic to $e \log n$.
Height: maximum distance between a vertex and the root

Example 1: Preferential attachment graphs

Model definition

1. Fix m and start with a connected graph.
2. In each step a new vertex and m new edges are born.
3. One endpoint of each new edge is the new vertex,
   the other endpoint is sampled according to the degrees.

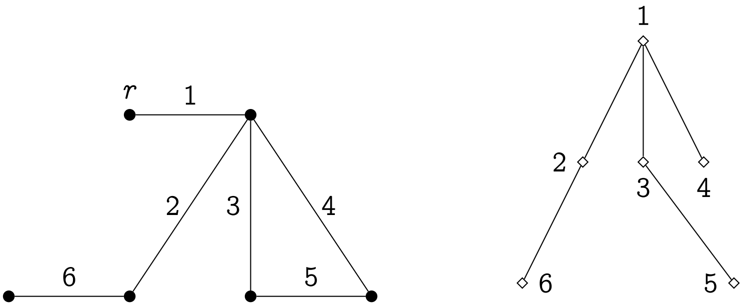
Model definition

1. Fix m and start with a connected graph.
2. In each step a new vertex and m new edges are born.
3. One endpoint of each new edge is the new vertex,
   the other endpoint is sampled according to the degrees.

The diameter is a.a.s. $O(\log n)$ if $m = 1$ [Pittel'94] and $O(\log n / \log\log n)$ if $m > 1$ [Bollobás and Riordan'04].
We prove a.a.s. the diameter $\leq 4e \log n$.

Proof sketch

1. Choose a vertex of the initial graph as the root.
2. Couple with a tree whose nodes correspond to the edges of the graph, such that the depth of each node $\geq$ the distance between the corresponding edge and the root.

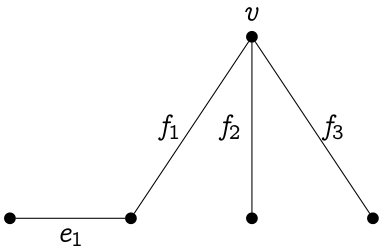


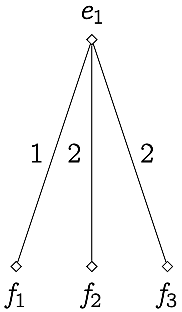
3. Creation of a new vertex (say $m = 3$)

Key observation: Sampling according to the degrees is equivalent to choosing a random endpoint of a uniformly random edge!

3. Creation of a new vertex (say $m = 3$)

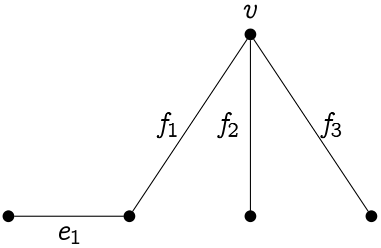
Key observation: Sampling according to the degrees is equivalent to choosing a random endpoint of a uniformly random edge!

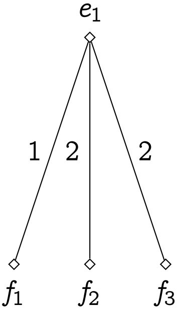




3. Creation of a new vertex (say $m = 3$)

Key observation: Sampling according to the degrees is equivalent to choosing a random endpoint of a uniformly random edge!





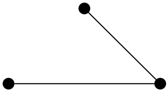
Thus, the diameter is a.a.s. at most $4e \log n$.

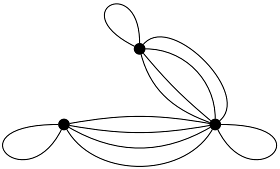
Slight generalization

What if the attachment probabilities are linear functions of the degrees?

Slight generalization

What if the attachment probabilities are linear functions of the degrees?
Say prob. of attaching to $v = 2\deg(v) + 1$





Example 2: The Cooper-Frieze model

The Cooper-Frieze model

- ✓ In an step, either a new vertex is born and edges are added from it to the existing graph,
  or edges are added between the existing vertices.
- ✓ The number of added edges is a bounded random variable.
- ✓ One endpoint of each added edge is either the new vertex,
  or a uniformly random vertex,
  or sampled according to degrees.
- ✓ The other endpoint is either a uniformly random vertex or sampled according to degrees.

[Cooper and Frieze'01]

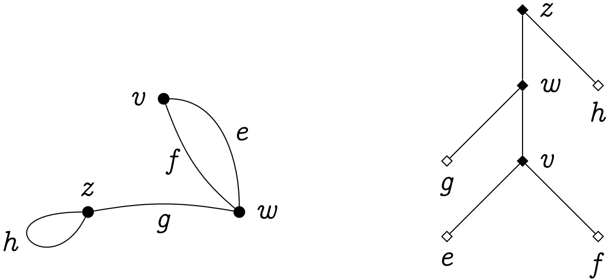
The Cooper-Frieze model

- ✓ In an step, either a new vertex is born and edges are added from it to the existing graph,
  or edges are added between the existing vertices.
- ✓ The number of added edges is a bounded random variable.
- ✓ One endpoint of each added edge is either the new vertex,
  or a uniformly random vertex,
  or sampled according to degrees.
- ✓ The other endpoint is either a uniformly random vertex or sampled according to degrees.

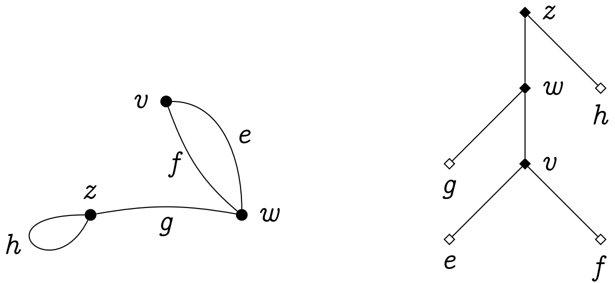
[Cooper and Frieze'01]

We prove a.a.s. the diameter is $O(\log n)$. (NEW RESULT!)

Proof sketch



Proof sketch



In every step, either a uniformly random white node gives birth, or a uniformly random black node gives birth.

The height is still logarithmic.

New results

Theorem (M'14+)

The following random graph models have diameter $O(\log n)$ a.a.s.

- ✓ *The (edge) copying model [Kumar,Raghavan,Rajagopalan,Sivakumar,Tomkins,Upfal'00]*
- ✓ *Aiello-Chung-Lu models* *[Aiello,Chung,Lu'01]*
- ✓ *The Cooper-Frieze model* *[Cooper,Frieze'01]*
- ✓ *The generalized linear preference model* *[Bu,Towsley'02]*
- ✓ *The PageRank-based selection model [Pandurangan,Raghavan,Upfal'02]*
- ✓ *Directed scale-free graphs* *[Bollobás,Borgs,Chayes,Riordan'03]*
- ✓ *The forest fire model* *[Leskovec,Kleinberg,Faloutsos'05]*

The PARID model of Deijfen, van den Esker, van der Hofstad and Hooghiemstra has diameter $O(\log^2 n)$ a.a.s. if the initial degrees' distribution has an exponential decay.