It's a small world for random surfers

Abbas Mehrabian
amehrabi@uwaterloo.ca

University of Waterloo

6 September 2014
RANDOM, Barcelona

joint work with Nick Wormald

The random-surfer Webgraph model

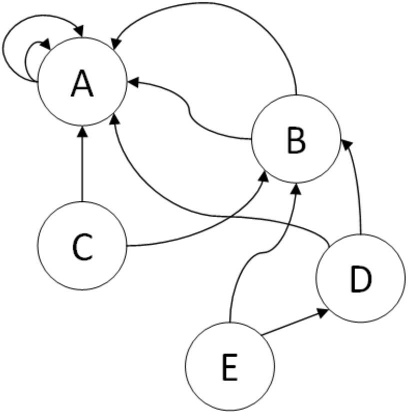
Model definition

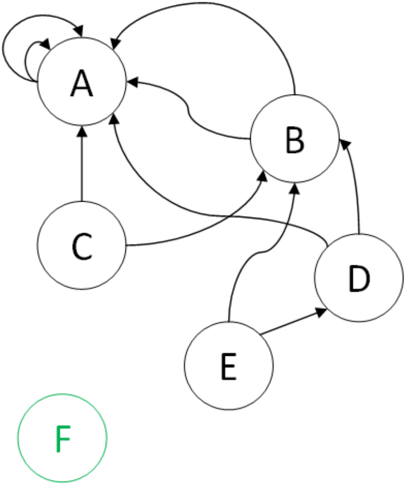
- ✓ Parameters: p and d
- ✓ Consider a pool of independent Geo(p) random variables.
- ✓ Build a random graph with out-degree d: start with one vertex with d loops, add a new vertex in each step.

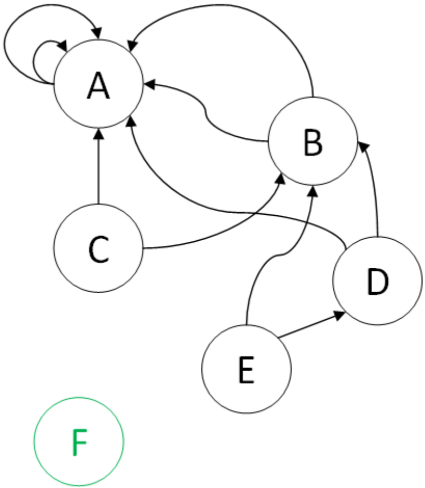
Model definition

- ✓ Parameters: p and d
- ✓ Consider a pool of independent Geo(p) random variables.
- ✓ Build a random graph with out-degree d: start with one vertex with d loops, add a new vertex in each step.

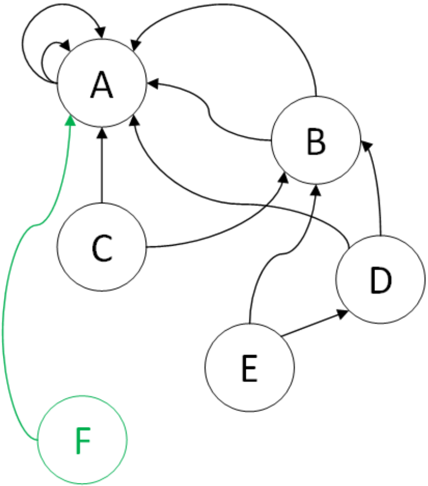
Say $p = 1/2$ and $d = 2$


A
B
C
D
E

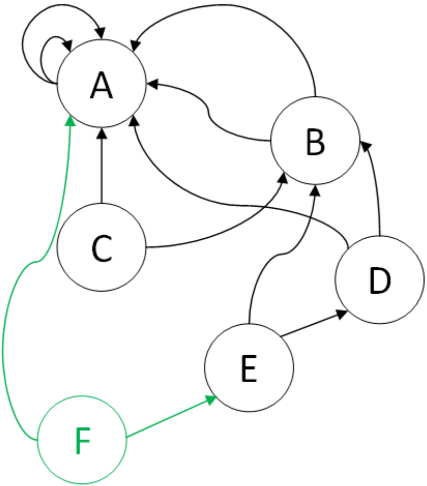

A
B
C
D
E
F



C, 2

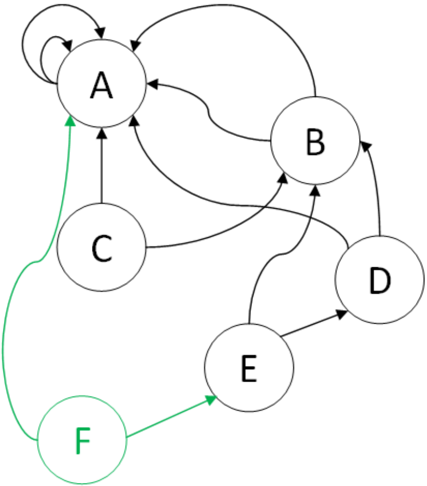


C, 2



C, 2

E, 0



C, 2

E, 0

Equivalently, PageRank-based selection

Previous work

- ✓ Pandurangan, Raghavan, Upfal'02: PageRank-based selection model definition, experimental results
- ✓ Blum, Chan, Rwebangira'06: random-surfer model definition, experimental results
- ✓ Chebolu and Melsted'08: observed the models are equivalent, partial analysis of degree sequence

Previous work

- ✓ Pandurangan, Raghavan, Upfal'02: PageRank-based selection model definition, experimental results
- ✓ Blum, Chan, Rwebangira'06: random-surfer model definition, experimental results
- ✓ Chebolu and Melsted'08: observed the models are equivalent, partial analysis of degree sequence

We studied the diameter of the underlying undirected graph...

Our result

Theorem

A.a.s. the diameter of the underlying graph $\leq (8e^p/p)\log n$

Our result

Theorem

A.a.s. the diameter of the underlying graph $\leq (8e^p/p)\log n$

✓ The small-world phenomenon holds for this model.

Our result

Theorem

A.a.s. the diameter of the underlying graph $\leq (8e^p/p)\log n$

- ✓ The small-world phenomenon holds for this model.
- ✓ No dependence on the out-degrees: due to the proof technique.

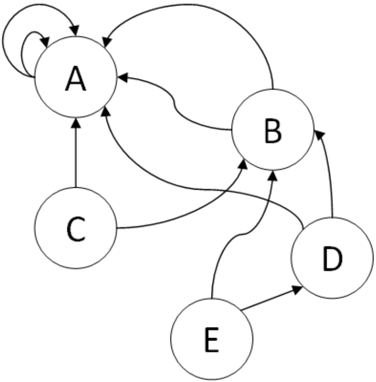
Our result

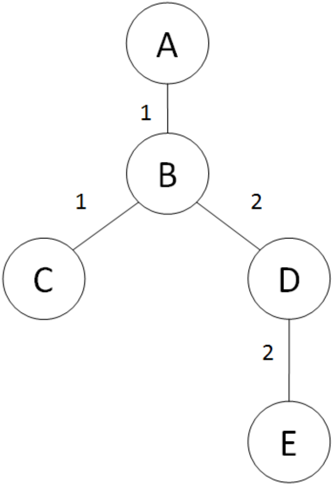
Theorem

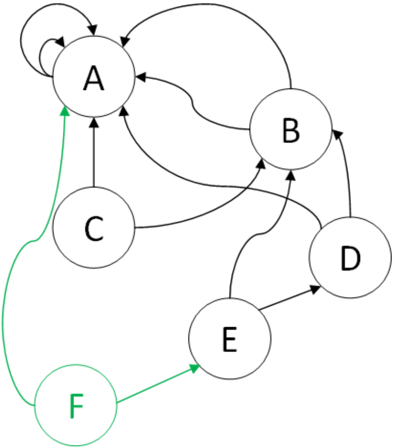
A.a.s. the diameter of the underlying graph $\leq (8e^p/p)\log n$

- ✓ The small-world phenomenon holds for this model.
- ✓ No dependence on the out-degrees: due to the proof technique.

Proof idea: build a coupling with a weighted random recursive tree

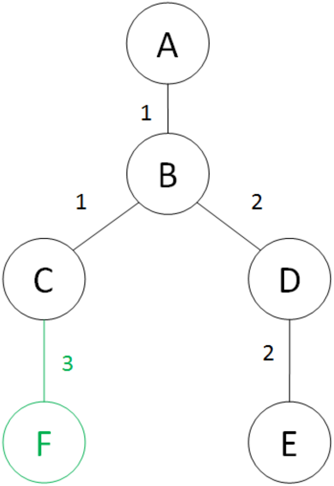

A
B
C
D
E

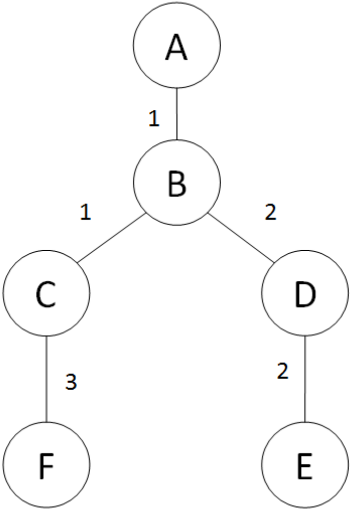

A
1
B
1
2
C
D
2
E



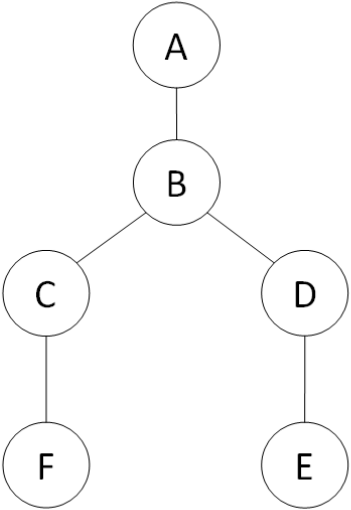
C, 2

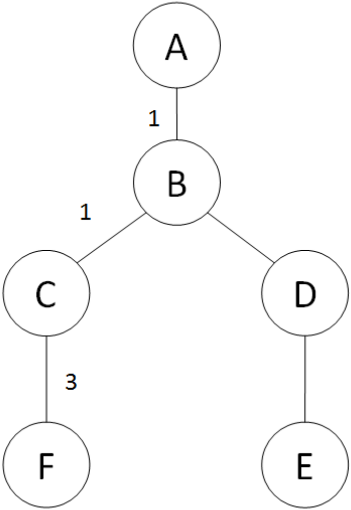
E, 0

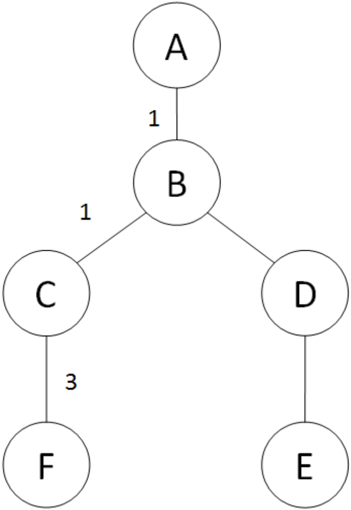



A
1
B
1
2
C
D
3
2
F
E

The height of the unweighted tree $\sim e \log n$ [Pittel'94]



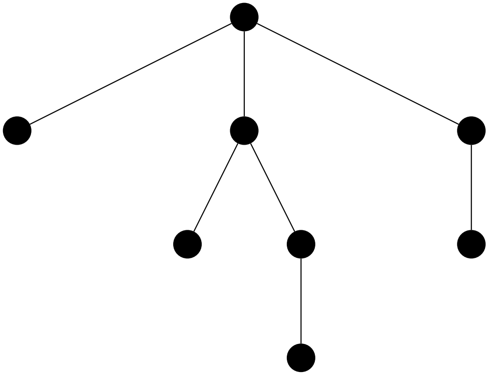

A
1
B
1
C
D
3
F
E

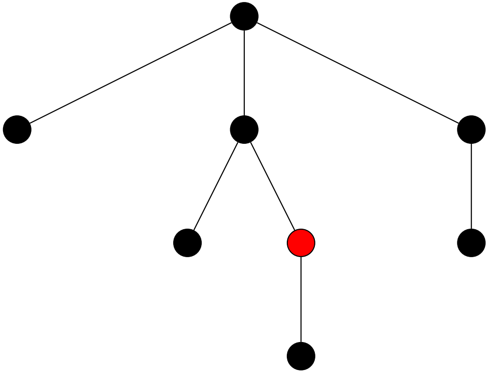


Theorem

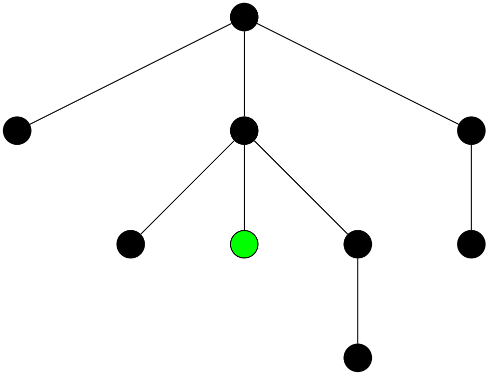
A.a.s. the height of the weighted tree $\leq (4e^p/p)\log n$

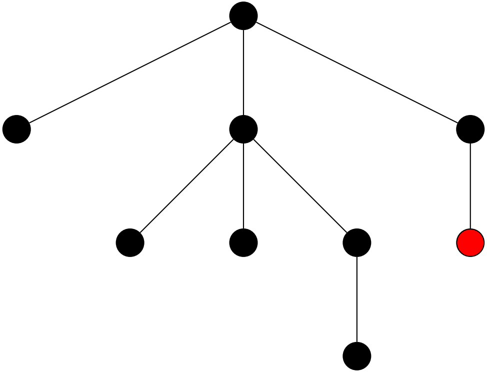
Random-surfer trees ($d = 1$)



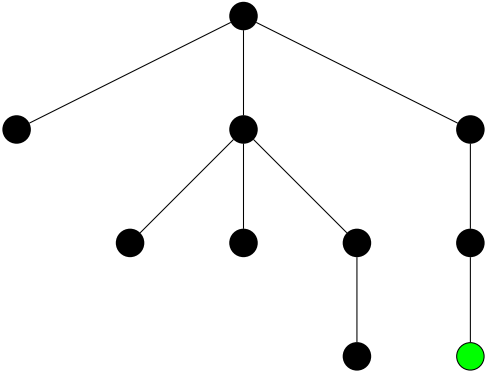


next geom.r.v. $= 1$





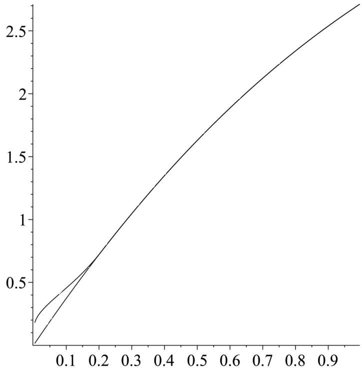
next geom.r.v. $= 0$



Our result

Theorem

A.a.s. the height is between $(L(p) - o(1)) \log n$ *and* $(U(p) + o(1)) \log n$*, and the diameter is between twice these values.*

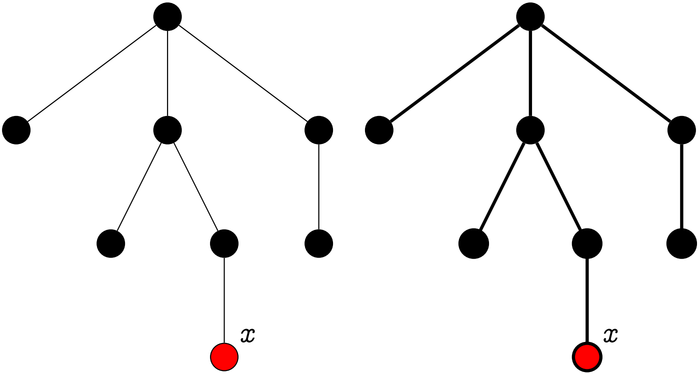


Our result

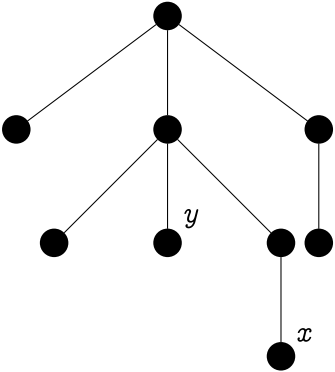
Theorem

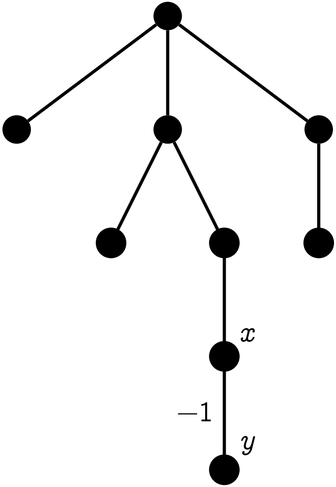
A.a.s. the height is between $(L(p) - o(1)) \log n$ *and* $(U(p) + o(1)) \log n$*, and the diameter is between twice these values.*

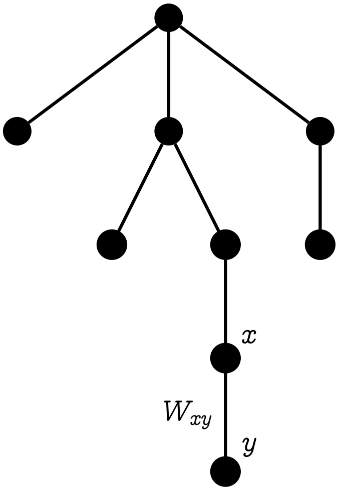
Proof idea: eliminate the complicated attachment rule by introducing weights,
then adapt a powerful technique of Broutin and Devroye'06 for analysing heights of weighted random trees



next geom.r.v. $= 2$ next geom.r.v. $= 2$

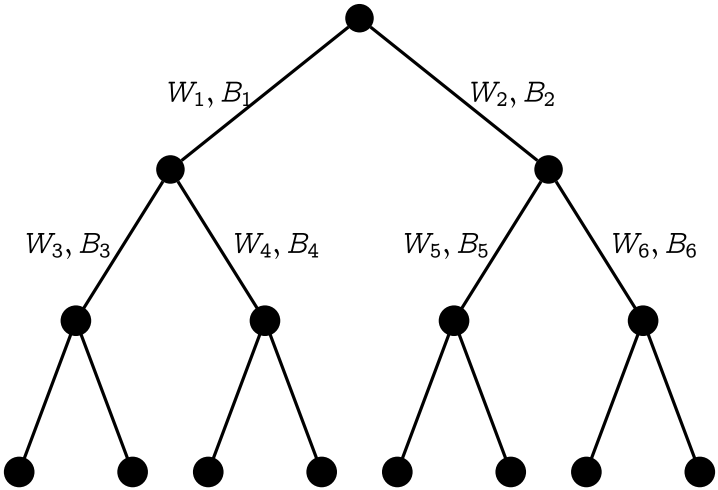

y
x


x
−1
y



$$W_{xy} = 1 - \text{Geo}(p)$$

The technique of Broutin and Devroye



Poissonization, and use the theory of large deviations

Full statement of our result

Given p and $\varepsilon > 0$, a.a.s. the height is between $(L(p) - \varepsilon)\log n$ and $(U(p) + \varepsilon)\log n$, and the diameter is between twice these values.

Let $p_0 \approx 0.206$ be the unique solution in $(0, 1/2)$ to

$$\log\left(\frac{1-p}{p}\right) = \frac{1-p}{1-2p} .$$

Let s be the solution in $(0, 1)$ to

$$s \log\left(\frac{(1-p)(2-s)}{1-s}\right) = 1 .$$

Then,

$$L(p) = \exp(1/s)s(2-s)p ,$$

and

$$U(p) = \begin{cases} L(p) & \text{if } p_0 \le p < 1 \\ \left(\log\left(\frac{1-p}{p}\right)\right)^{-1} & \text{if } 0 < p < p_0 . \end{cases}$$

Open problems

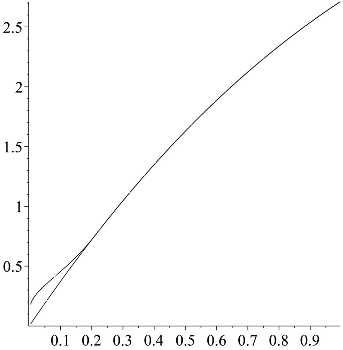
✓ Diameter of random-surfer Webgraph $\leq (8e^p/p) \log n$.

Open problems

- ✓ Diameter of random-surfer Webgraph $\leq (8e^p/p) \log n$.
  Logarithmic lower bound?
  What about the maximum finite directed distance?!

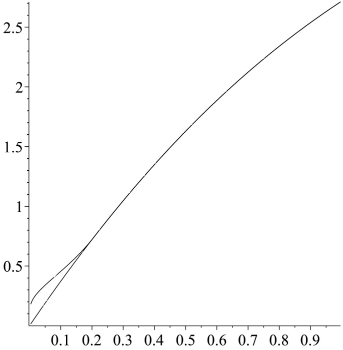
Open problems

- ✓ Diameter of random-surfer Webgraph $\leq (8e^p/p)\log n$.
  Logarithmic lower bound?
  What about the maximum finite directed distance?!
- ✓ Diameter of random-surfer tree is between $(2L(p) - o(1))\log n$ and $(2U(p) + o(1))\log n$



Open problems

- ✓ Diameter of random-surfer Webgraph $\leq (8e^p/p)\log n$.
  Logarithmic lower bound?
  What about the maximum finite directed distance?!
- ✓ Diameter of random-surfer tree is between $(2L(p) - o(1))\log n$ and $(2U(p) + o(1))\log n$



$$W_{xy} = \max\{1 - \text{Geo}(p), 1 - W_x\}$$