Leadership Statistics in Random Structures

Eli Ben-Naim† and Paul Krapivsky‡

†*Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545*

‡*Department of Physics, Boston University, Boston, MA 02215*

The largest component ("the leader") in evolving random structures often exhibits universal statistical properties. This phenomenon is demonstrated analytically for two ubiquitous structures: random trees and random graphs. In both cases, lead changes are rare as the average number of lead changes increases quadratically with logarithm of the system size. As a function of time, the number of lead changes is self-similar. Additionally, the probability that no lead change ever occurs decays exponentially with the average number of lead changes.

E. Ben-Naim and P. L. Krapivsky, *Europhys. Lett.* **65**, 151 (2004).

Problem

An ensemble of random growing trees (RT) or random growing graphs (RG), starting from N nodes.

Questions

- What is the size of the leader, i.e., the largest component $l(t, N)$?
- What is the number of lead changes $L(t, N)$ as a function of time t and system size N?
- What is the number of total lead changes $L(N)$ as a function of system size?

Motivation

- Data storage algorithms in computer science (RT) [1,2].
- Collision processes in gases (RT) [3].
- Random networks (RG) [4,5].
- Polymerization and Gelation (RG) [6].

Random Trees

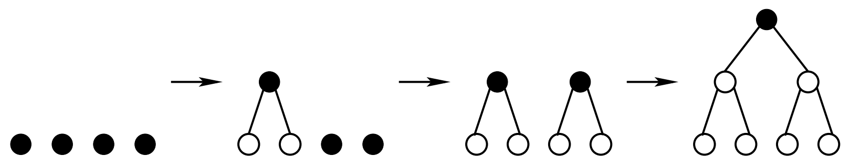

FIG. 1. Illustration of the tree merger process with system size $N = 4$.

- Start with N single-leaf trees.
- Pick two **trees** at random and merge them.

Size Distribution

- Distribution of component of size k at time t is $c_k(t)$
- Rate Equation

$$\frac{d}{dt}c_k = \sum_{i+j=k} c_i c_j - 2cc_k$$

- Solution subject to initial condition $c_k(0) = \delta_{k,0}$.

$$c_k(t) = \frac{t^{k-1}}{(1+t)^{k+1}}$$

- Exponential scaling distribution (asymptotically)

$$c_k(t) \simeq k_*^{-2}\Phi(k/k_*), \qquad \Phi(x) = e^{-x}$$

- Typical size: $k_* \sim t$.

Leadership Statistics

- **Leader Size $l(t,N)$:** Obtain leader size from cumulative distribution $u_k = N\,\Sigma_{j=k}^{\infty}\,c_k$ using $u_l = 1$. Asymptotically,

$$l \simeq t \ln \frac{N}{t}.$$

- **Number of lead changes $L(t,N)$:** Obtain the rate by which the leader is surpassed from the rate of change in the cumulative distribution $\frac{d}{dt}L = \frac{d}{dt}u_k|_{k=l}$.

$$L(t,N) \simeq \ln t \ln N - \frac{1}{2}(\ln t)^2.$$

- **Unusual scaling form:** involves scaling function $F(x) = x - \frac{1}{2}x^2$.

$$L(t,N) = (\ln N)^2\,F(x), \qquad x = \frac{\ln t}{\ln N} = \frac{\ln k_*}{\ln N}$$

- **Total number of lead changes $L(N)$:** is obtained by noting that the condensation time, the time to form a single tree of size N is simply $t_f \simeq N$. Using $x = 1$,

$$L(N) \simeq A(\ln N) \qquad A = \frac{1}{2}.$$

Numerical Simulations

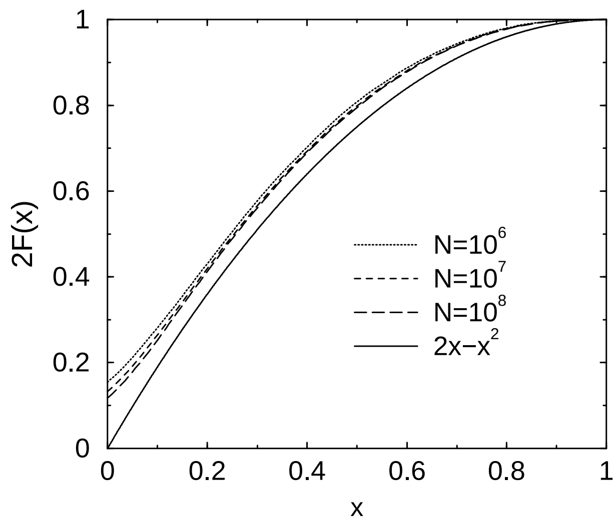

FIG. 2. The normalized time dependence of the number of lead changes for random trees, $L(t,N)/L(N)$, versus the scaling variable $x = \ln t/\ln N$. The simulation data, representing an average over 10^3 Monte Carlo runs, is compared with the theoretical prediction $2F(x) = 2x - x^2$.

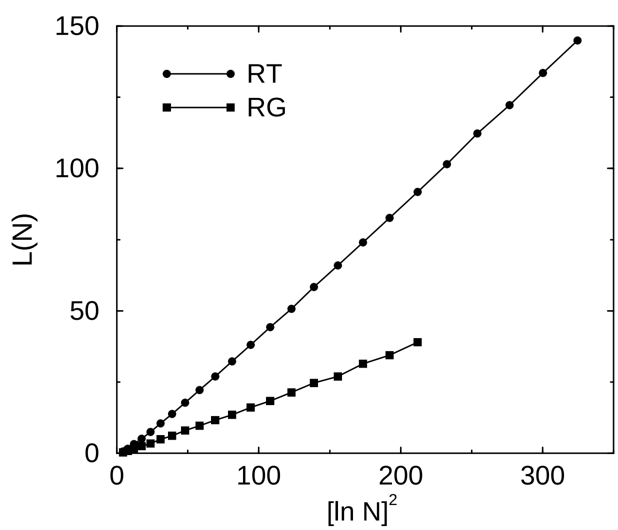

FIG. 3. The total number of lead changes $L(N)$ versus the system size N. Shown are simulation results for Random Trees (RT) and Random Graphs (RG) representing an average over 10^4 realizations.

Distribution of number of lead changes

- The probability $P_n(t,N)$ that there are n lead changes at time t is Poissonian (assuming no correlations build up). Thus, it is characterized by the average number of lead changes $L(t,n)$.

$$P_n(t,N) = \frac{[L(t,N)]^n}{n!}\, e^{-L(t,N)}.$$

- The survival probability of the first leader $S(t)$, the probability that the initial leader is never overtaken, equals $P_0(t,N)$

$$S(N) = \exp[-L] \simeq \exp\left[-A(\ln N)^2\right].$$

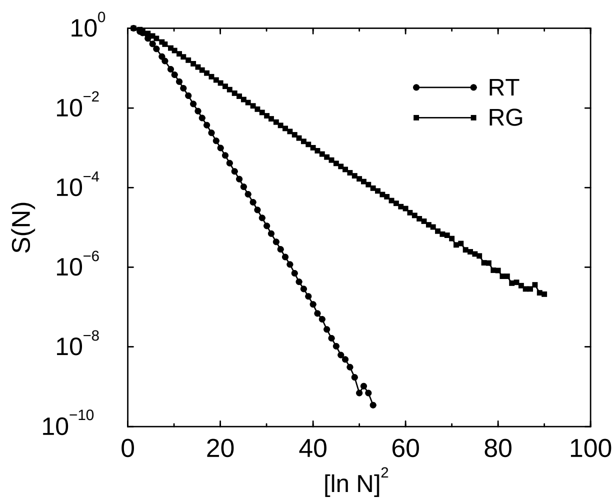

FIG. 4. The survival probability of the initial leader $S(N)$ versus the system size N. The number of realizations was 10^{10} and 10^8 for random trees and random graphs, respectively.

Random Graphs

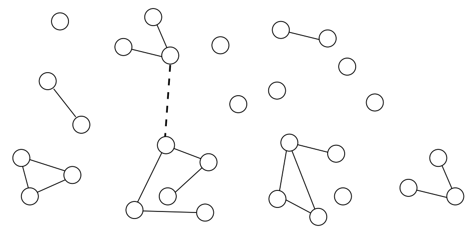

FIG. 5. Illustration of the random graph growth process.

- Start with N single-node graphs.
- Pick two **nodes** at random and merge their respective **graphs**.

Size Distribution

- Distribution of components of size k at time t, is $c_k(t)$, satisfies the rate equation $\frac{d}{dt}c_k = \frac{1}{2}\sum_{i+j=k} ijc_ic_j - kc_k$
- Solution subject to initial condition $c_k(0) = \delta_{k,0}$

$$c_k(t) = \frac{(kt)^{k-1}}{k \cdot k!}e^{-kt}$$

- Scaling distribution (asymptotic)

$$c_k(t) \simeq k_*^{-5/2}\Phi(k/k_*), \qquad \Phi(x) \propto x^{-5/2}e^{-x/2}.$$

- Gelation time: $1 - t_g \sim N^{-1/3}$
- Typical size: $k_* \sim (1-t)^{-2}$ as $t \to 1$; Giant comp. size $\sim N^{2/3}$

Leadership Statistics

- Obtained from size distribution as in the random tree case.

- **Leader size $l(t,N)$:**

$$l \simeq \frac{2}{(1-t)^2}\, \ln[N(1-t)^3].$$

- **Number of lead changes $L(t,N)$:**

$$L(t,N) \simeq 2\, \ln N\, \ln\frac{1}{1-t} - 3\left[\ln\frac{1}{1-t}\right]^2 .$$

- **Unusual scaling form:** involves $F(x) = 2x - 3x^2$.

$$L(t,N) \simeq (\ln N)^2\, F(x), \qquad x = \frac{\ln\frac{1}{1-t}}{\ln N} = \frac{1}{2}\frac{\ln k_*}{\ln N}$$

- **Total number of lead changes $L(N)$:**

  Obtained from $1 - t_g \sim N^{-1/3}$ or $x = 1/2$

$$L(N) \simeq A(\ln N) \qquad A = \frac{1}{3}.$$

- **Probability no lead change occurs:**

$$S(N) \simeq \exp[-\frac{1}{3}(\ln N)^2]$$

Conclusions

- Similar laws characterize random trees and random graphs.
- Lead changes are rare, their total number grows as $(\ln N)^2$.
- Unusual scaling behavior, with scaling variable $\ln k_*/\ln N$.
- Probability no lead change occur decays as $\exp[-A(\ln N)^2]$.

References

1. H. M. Mahmoud, *Evolution of Random Search Trees* (John Wiley & Sons, New York, 1992).
2. D. E. Knuth, *The Art of Computer Programming, vol. 3, Sorting and Searching* (Addison-Wesley, Reading, 1998).
3. R. van Zon, H. van Beijeren, and Ch. Dellago, Phys. Rev. Lett. **80**, 2035 (1998).
4. B. Bollobás, *Random Graphs* (Academic Press, London, 1985).
5. S. Janson, T. Łuczak, and A. Rucinski, *Random Graphs* (John Wiley & Sons, New York, 2000).
6. P. J. Flory, *Principles of Polymer Chemistry* (Cornell, Ithaca, 1953).