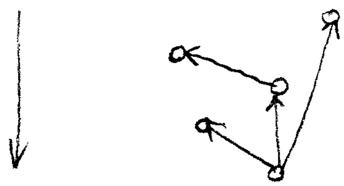
Lecture 4 (1/18/2018)
Scale-Free Networks – Why?

Price model (1965) – Network of scientific papers

Two ingredients (i) growing network
(ii) "Rich-get-richer" effect
(Herbert Simon, 1955, Economy)



t

\# of nodes : $n(t)$
– an increasing function of time

Acyclic network

$p_q(n)$ — fraction of nodes in the network of size n with in-degree q

Now add a single new node

Probability that an existing node i gets a link (citation)

$$\sim (q_i + a), \quad a > 0 \text{ constant}$$

$$\text{Probability} = \frac{q_i + a}{\sum_i (q_i + a)} = \frac{q_i + a}{n\langle q\rangle + na} = \frac{q_i + a}{n(c+a)}$$

$\langle q\rangle$ ← average in-degree $\equiv c$

Expected # of new citations to all nodes with degree q: $n\,p_q(n)\cdot c \cdot \frac{q+a}{n(c+a)}$

But $\underbrace{q-1}_{} \to q$

A node of in-degree $(q-1)$, if it receives a citation

Expected # of such nodes: $n\,p_{q-1}(n)\cdot c \cdot \frac{q-1+a}{n(c+a)}$

– Represents $\oplus$ to $p_q(n)$

Similarly $q \to q+1$: – loss or $\ominus$ to $p_q(n)$

$$n\,p_q(n)\cdot c\,\frac{q+a}{n(c+a)}$$

$$\Rightarrow \boxed{(n+1)\,p_q(n+1) = n\,p_q(n) + \frac{c(q-1+a)}{c+a}\,p_{q-1}(n) - \frac{c(q+a)}{c+a}\,p_q(n)}$$

– Master equation for the dyn. evo. of $p_q(n)$

Solutions: $n \to \infty$ $\quad p_q = p_q(\infty)$

$$(n+1)\,p_q(\infty) - n\,p_q(\infty) = p_q(\infty)$$

$$\Rightarrow \quad p_q = \frac{c}{c+a}\left[(q-1+a)\,p_{q-1} - (q+a)\,p_q\right]$$

$p_0 = ?$ no such thing as $p_{-1}(n)$ – 2nd term in Master eq. doesn't appear

when a new node is added to the network, it has in-degree zero

$$\Rightarrow \quad (n+1)\,p_0(n+1) = n\,p_0(n) + 1 - \frac{ca}{c+a}\,p_0(n)$$

$$\Rightarrow \quad p_0(\infty) = 1 - \frac{ca}{c+a}\,p_0(\infty)$$

$$\Rightarrow \quad p_0 = \frac{1+a/c}{a+1+a/c}$$

Iteration: $p_q = \dfrac{(q+a-1)(q+a-2)\cdots a}{(q+1+a+a/c)\cdot(q+a+a/c)\cdots(2+a+a/c)}\cdot\dfrac{1+a/c}{1+a+a/c}$

Gamma function: $\begin{cases} \Gamma(x) \equiv \int_0^\infty t^{x-1}e^{-t}\,dt \\ \Gamma(x+1) = x\,\Gamma(x) \quad \text{for all } x>0 \end{cases}$

$$\Rightarrow \quad \frac{\Gamma(x+n)}{\Gamma(x)} = (x+n-1)(x+n-2)\cdots x$$

Thus $\quad (q+a-1)(q+a-2)\cdots a = \dfrac{\Gamma(q+a)}{\Gamma(a)}$

$$(q+1+a+\tfrac{a}{c})(q+a+\tfrac{a}{c})\cdots(2+a+\tfrac{a}{c})(1+a+\tfrac{a}{c})$$

$x \equiv 1+a+a/c$

$n \equiv 1+q$ $\quad \longrightarrow \quad \dfrac{\Gamma(q+1+1+a+a/c)}{\Gamma(1+a+a/c)}$

Solution in terms of Gamma function:

$$p_q = \left(1+\frac{a}{c}\right)\frac{\Gamma(q+a)\,\Gamma(a+1+a/c)}{\Gamma(a)\,\Gamma(q+a+2+a/c)}$$

Euler's Beta function $\quad B(x,y) = \dfrac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$

$$\Rightarrow \quad p_q = \frac{B(q+a,\,2+a/c)}{B(a,\,1+a/c)}$$

Asymptotic behavior:

Large x : $\quad \Gamma(x) \approx \sqrt{2\pi}\,e^{-x}\,x^{x-\frac{1}{2}}$

$$B(x,y) \approx \frac{e^{-x}\,x^{x-1/2}}{e^{-(x+y)}(x+y)^{x+y-\frac{1}{2}}}\,\Gamma(y)$$

$(x+y)^{x+y-\frac{1}{2}}$

$= x^{x+y-\frac{1}{2}}\left(1+\frac{y}{x}\right)^{x+y-\frac{1}{2}}$

large x $\approx x^{x+y-\frac{1}{2}}\,e^{y}$ $\quad \Rightarrow \quad B(x,y) \approx x^{-y}\,\Gamma(y)$

— power-law in x with exponent y

For large value of q — power-law tail

$$p_q \sim (q+a)^{-\gamma} \sim q^{-\gamma} \quad (q \gg a)$$

Exponent $\quad \gamma = 2 + \dfrac{a}{c}$

Scale-Free Networks: Absence of Epidemic Threshold

VOLUME 86, NUMBER 14 PHYSICAL REVIEW LETTERS 2 APRIL 2001

Epidemic Spreading in Scale-Free Networks

Romualdo Pastor-Satorras[1] and Alessandro Vespignani[2]
[1]*Departament de Física i Enginyeria Nuclear, Universitat Politècnica de Catalunya, Campus Nord, Mòdul B4, 08034 Barcelona, Spain*
[2]*The Abdus Salam International Centre for Theoretical Physics (ICTP), P.O. Box 586, 34100 Trieste, Italy*
(Received 20 October 2000)

The Internet has a very complex connectivity recently modeled by the class of scale-free networks. This feature, which appears to be very efficient for a communications network, favors at the same time the spreading of computer viruses. We analyze real data from computer virus infections and find the average lifetime and persistence of viral strains on the Internet. We define a dynamical model for the spreading of infections on scale-free networks, finding the absence of an epidemic threshold and its associated critical behavior. This new epidemiological framework rationalizes data of computer viruses and could help in the understanding of other spreading phenomena on communication and social networks.

- SIS model for computer virus
- Susceptible-Infected-Susceptible
- λ - probability for a susceptible node to be infected
- Outbreak: $\lambda > \lambda_c$
- λ_c - epidemic threshold

Epidemic Threshold ~ Inverse of Second Moment

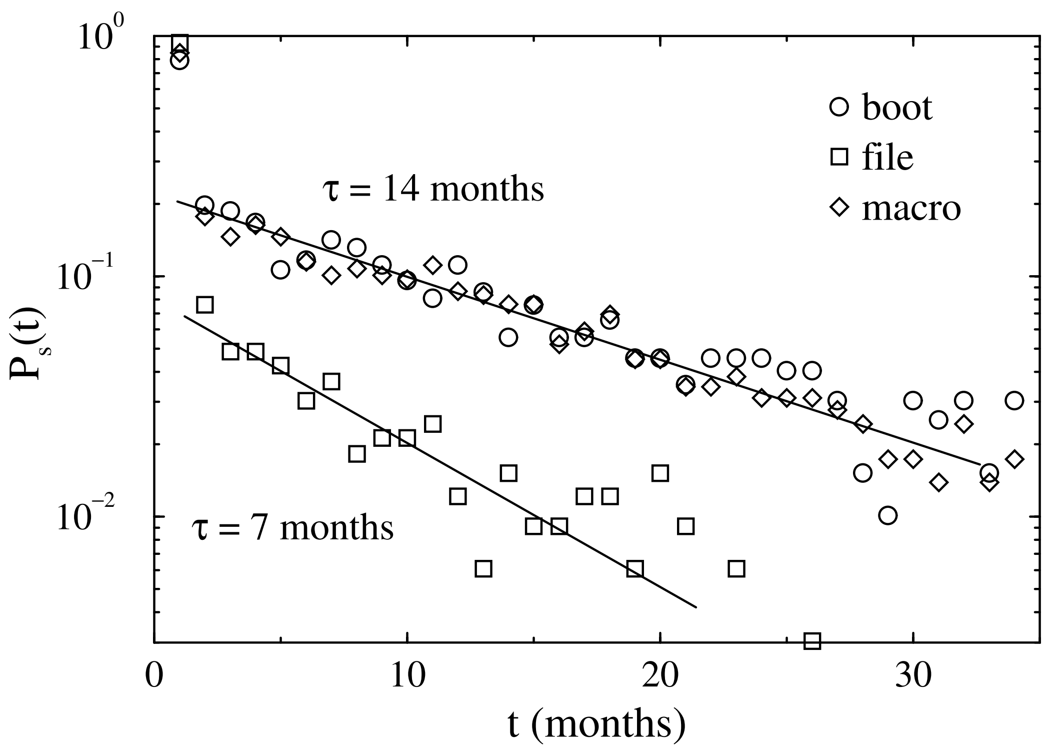


FIG. 1. Surviving probability for viruses in the wild. The 814 different viruses analyzed have been grouped in three main strains [9]: file viruses infect a computer when running an infected application; boot viruses also spread via infected applications, but copy themselves into the boot sector of the hard drive and are thus immune to a computer reboot; macroviruses infect data files and are thus platform independent. The presence of an exponential decay is evident in the plot, with characteristic time τ.

- Prevalent existence of virus – very long lifetime
- This shouldn't be the case due to the widespread use of antivirus software, which reduces the spreading rate to near zero values
- Resolution: epidemic threshold = 0

Mean-field theory predicts:

$$\lambda_c \sim 1/\langle k^2 \rangle$$