

Complex Dynamical Networks:

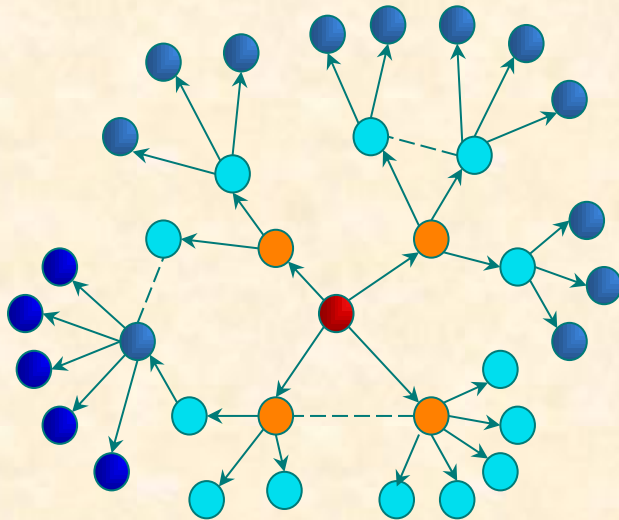
Lecture 5: Spreading Dynamics

EE 6605

Instructor: G Ron Chen



Epidemic Spreading

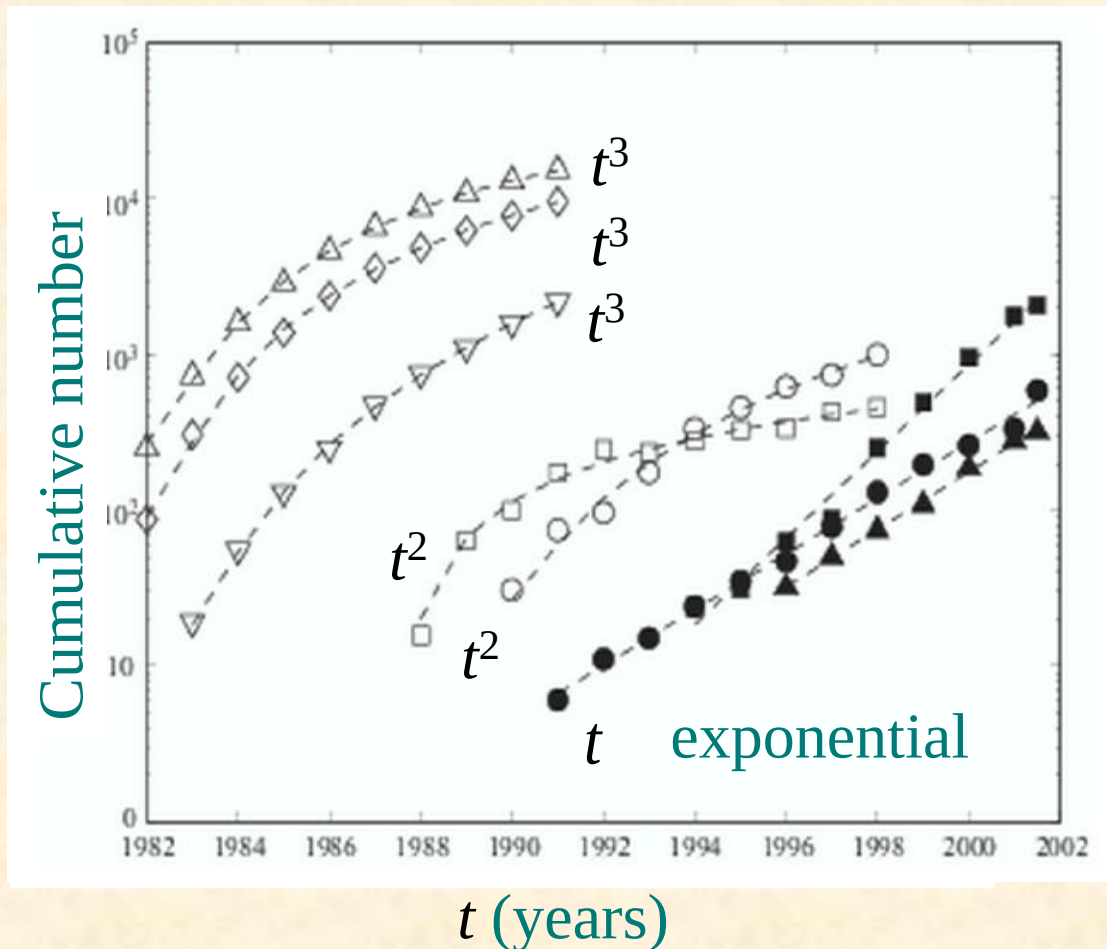


Typically through contact ...

Human Travel Map:



Example: AIDS Epidemics



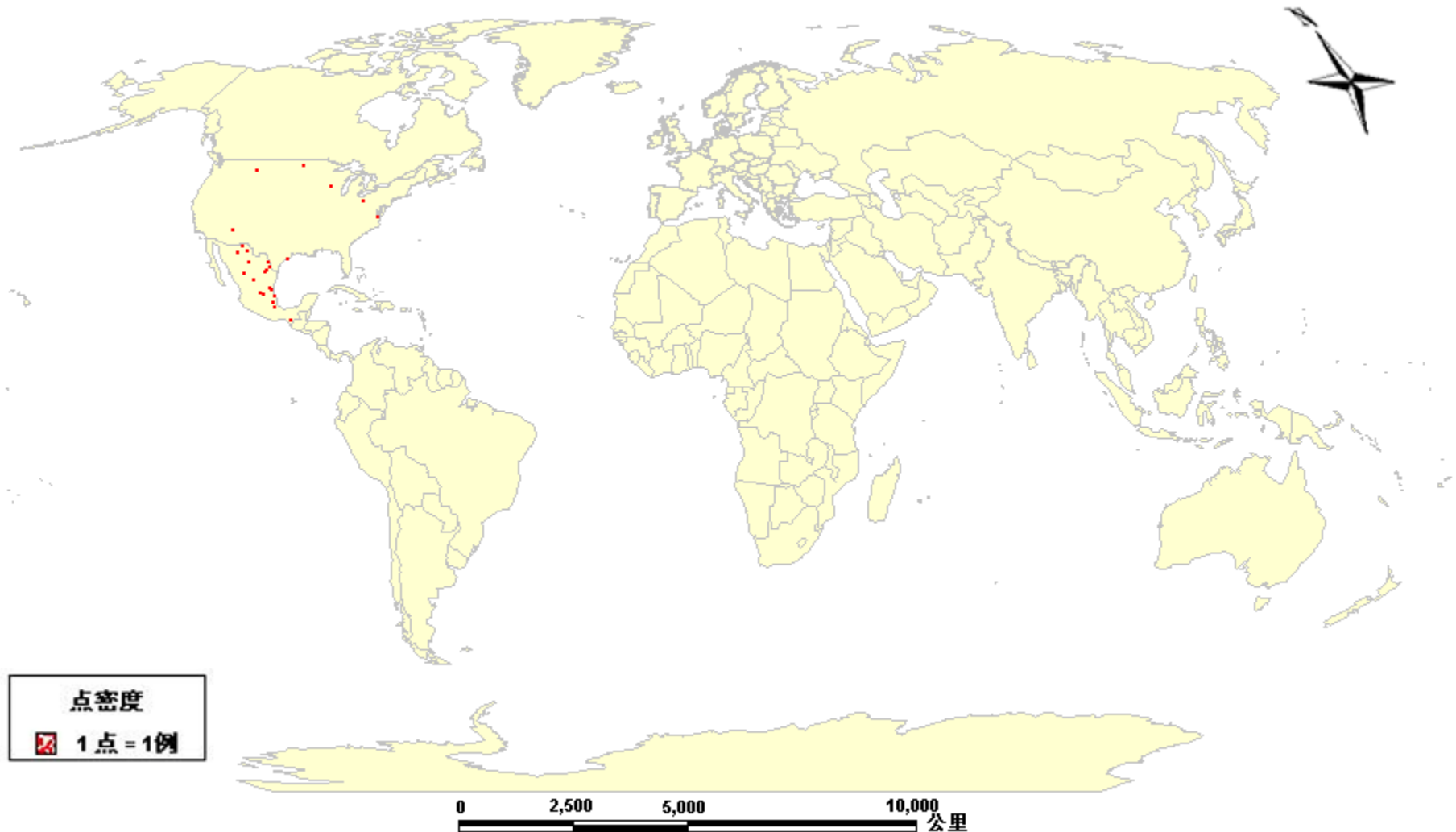
- $\triangle$ New York - Hom
- ∇ New York - Het
- $\diamond$ San Francisco - Hom
- $\circ$ South Africa
- $\square$ Kenya
- $\blacktriangle$ Georgia
- $\blacksquare$ Latvia
- $\bullet$ Lithuania

Szendroi & Czanyi, Proc. Royal Society of London, B 2004

Example: H1N1

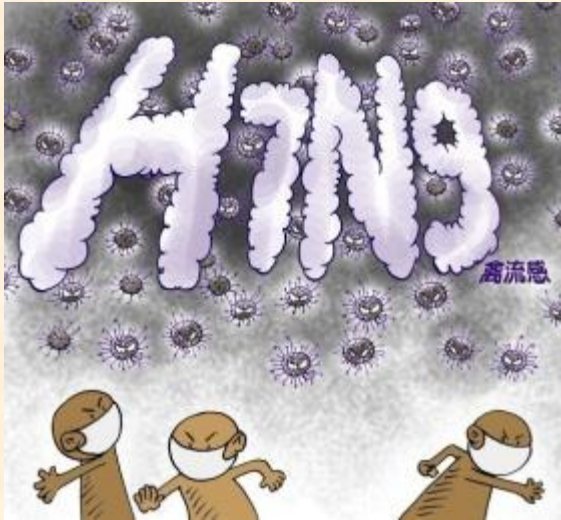
全球甲型H1N1流感疫情

2009年4月24日

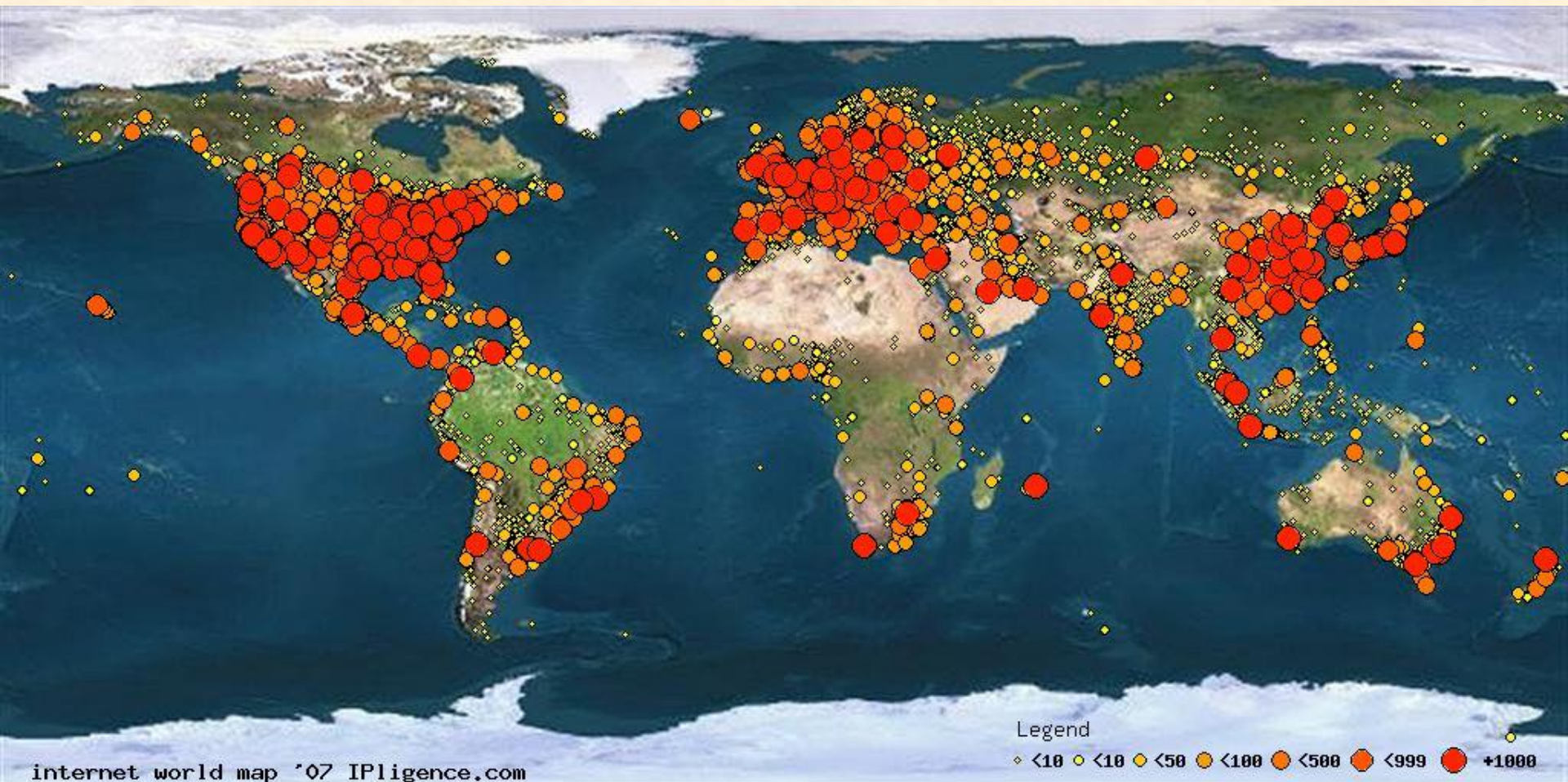




Coming Again?



Internet Computer Networks



Computer Virus Spreading

- The first virus capable of infecting PCs was perhaps the **Brain** virus developed in **Pakistan** in 1986.
- **I-Love-You** bug began in **Philippines** on May 4, 2000, and spread across the world in one day (traveling through **Hong-Kong** to Europe then to USA), infecting more than 78 million computers worldwide in just 4 days, i.e., about 10% of all computers connected to the Internet, and causing about US\$5.5 billions in damage.
- It was estimated that there were more than 48,000 identified viruses on the Internet worldwide in year 2000 alone.
- In 2000-2001, well-known viruses: **Lover Letter**, **Nimda**, and **Sircam** ...
- It was estimated that more than 80% computers were attacked by viruses in **China** in 2004.
- Also in 2004, the infamous **Worm Sasser** attacked several hundred thousands of computers over the world within just a couple of weeks.
- In 2009, **Conficker** (worm) remotely installs software on infected machines.
- Many other examples ... still today! [\[VirusList\]](#)

More Examples ...

In 2008, for example, security software maker “Symantec in Mountain View” in California detected 1.6 million new threats from computer viruses and other malicious software (In 2007, only 600,000 or so)

Such attacks will become more common and more sophisticated in the future

Statistically, unsolicited junk e-mail (‘spam’) accounts for about 90–95% of all e-mails sent

Such trash will become more common and more annoying in the future

Computer Viruses and Worms

- ❖ Computer **viruses** are usually referred to as some small **computer programs** that can reproduce themselves by infecting other programs and computers, which continuously grow and spread out.
- ❖ When a virus is active inside a computer, it is able to copy itself in many different ways into the codes of some programs of the computer.
- ❖ When the infected computer program is run into another computer, typically the code of the virus is executed first thus continues to infect other programs in the new computer.
- ❖ This process repeats endlessly, leading to the collapse of a local or even global network of computers eventually, causing tremendous technological and economical disasters.

Computer Viruses and Worms

Major computer viruses may be roughly classified into:

- **Boot-sector viruses**, which infect the boot sectors of floppies and hardware devices
- **File viruses**, which infect application programs
- **Macro viruses**, which infect data files directly
- **Hybrid types of viruses**, in a certain combination of the above basic ones, which infect some special technologies or applications such as Java, ActiveX, and HTML, etc.
- **Trojan horses** are malicious file or program, which can make copies of themselves, steal information, or harm their host computer systems

For Example: Last week

Name	Type	
<u>Android.iBanking</u>	Trojan	03/07/2014
<u>Trojan.Ransomcrypt.H</u>	Trojan	03/05/2014
<u>Android.Dendoroid</u>	Trojan	03/04/2014
<u>W32.Nakcos</u>	Virus	03/04/2014
<u>Trojan.Ransomcrypt.G</u>	Trojan	03/03/2014
<u>Backdoor.Comdinter</u>	Trojan	03/04/2014
<u>W32.IRCBot!gen3</u>	Worm	03/03/2014
<u>W32.Meteit!gen1</u>	Trojan, Worm	03/03/2014
<u>Trojan.Gampass!gen4</u>	Trojan	02/28/2014
<u>Linux.SSHKit!gen1</u>	Trojan	03/01/2014
<u>Trojan.Mebroot!gen6</u>	Trojan	03/01/2014

[Data Source](#) (Updated Daily)

Trojan

可利用DOS命令进行自我删除的恶意木马变种被发现

国家计算机病毒应急处理中心通过对互联网的监测发现，近期出现一种恶意木马程序变种 T r o j a n __ S w i s y n . C F D 。该变种在对操作系统进行入侵感染破坏后，利用 D O S 命令进行自我删除，阻止防病毒软件对其进行查杀。

该变种运行后，会自身复制到受感染操作系统指定文件夹下，重命名为可执行文件，设置其属性为“只读”、“隐藏”。与此同时，该变种会创建进程快照，查找受感染操作系统中指定文件名进程，如发现该文件进程正在运行，则弹出对话框，显示防病毒软件正在升级的虚假提示信息。

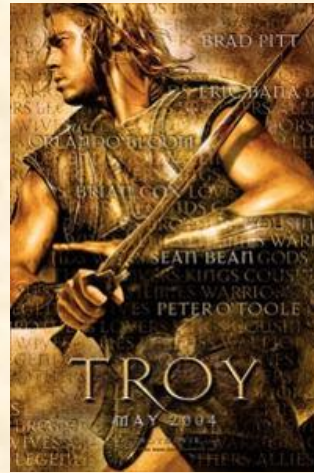
另外，该变种会迫使受感染操作系统主动访问指定的恶意 W e b 网址，最终变种可以获取受感染操作系统的本机信息（诸如：计算机名、操作系统版本、处理器类型、内存大小等），随即发送到恶意攻击者指定的 W e b 服务器上，致使受感染操作系统接受远程恶意代码指令。

针对已经感染该恶意木马程序变种的计算机用户，专家建议立即升级系统中的防病毒软件，进行全面杀毒。对未感染的用户建议打开系统中防病毒软件的“系统监控”功能，从注册表、系统进程、内存、网络等多方面对各种操作进行主动防御。

“Achilles' Heel”



Corfu, Greece

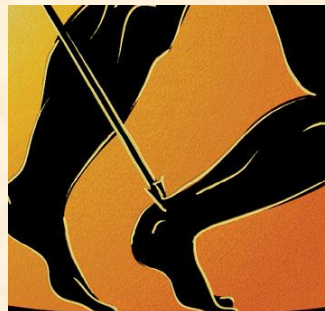


2004 film: TROY

特洛伊



Trojan Hours



Story

Robustness and Fragility of Scale-Free Networks

“Achilles’ heel”

R. Albert, H. Jeong, A. L. Barabasi, *Nature*, 406, 387-482 (2000)



In Greek mythology, when Achilles was a baby, his mother Thetis took him to the River Styx, which was supposed to offer powers of invulnerability, and dipped his body into the water. But as mother held Achilles by the heel, his heel was not washed over by the water of the magical river. Achilles grew up to be a man of war who survived many great battles. On one day, a poisonous arrow shot at him was lodged in his heel, killing him shortly after.

Robustness versus Fragility

Illustration:

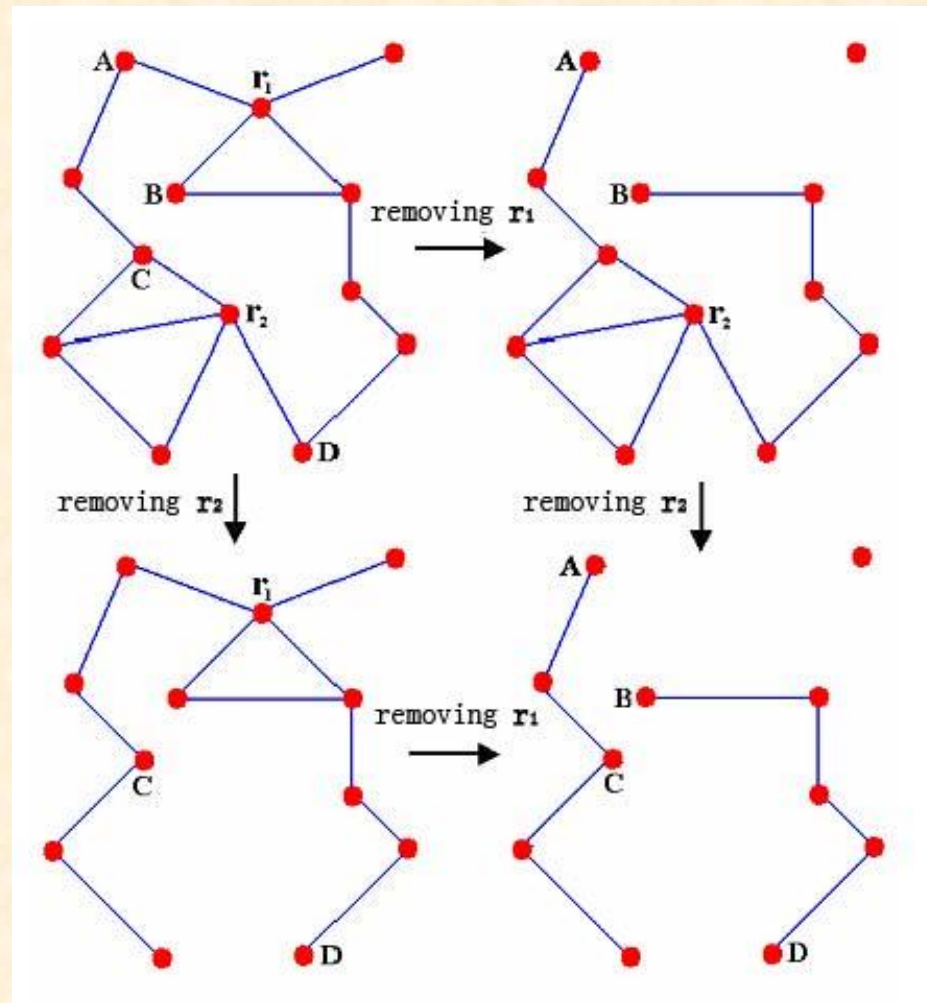
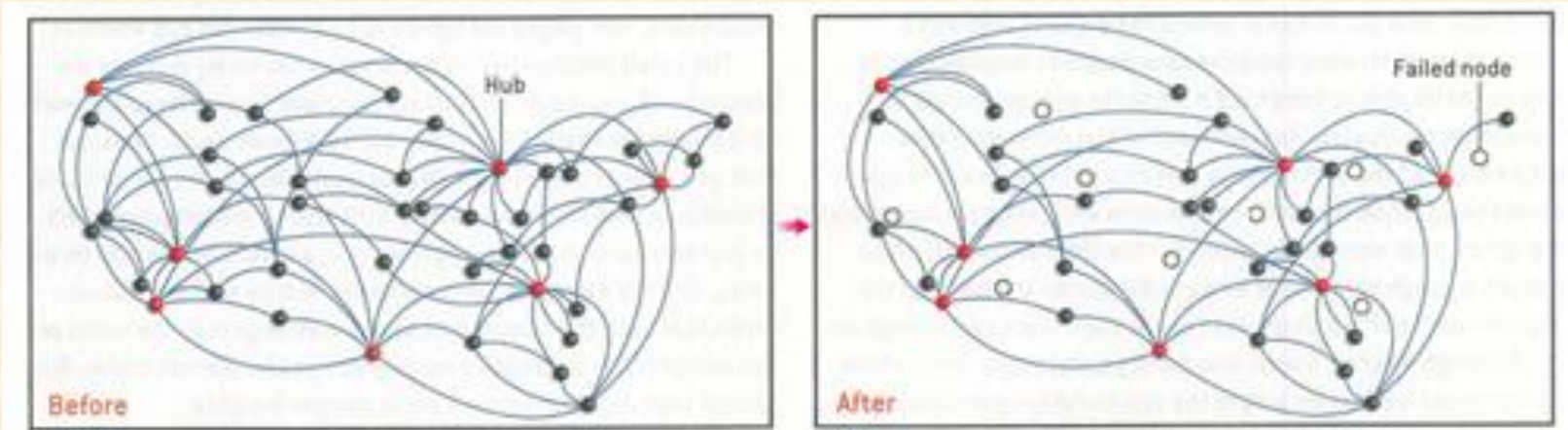
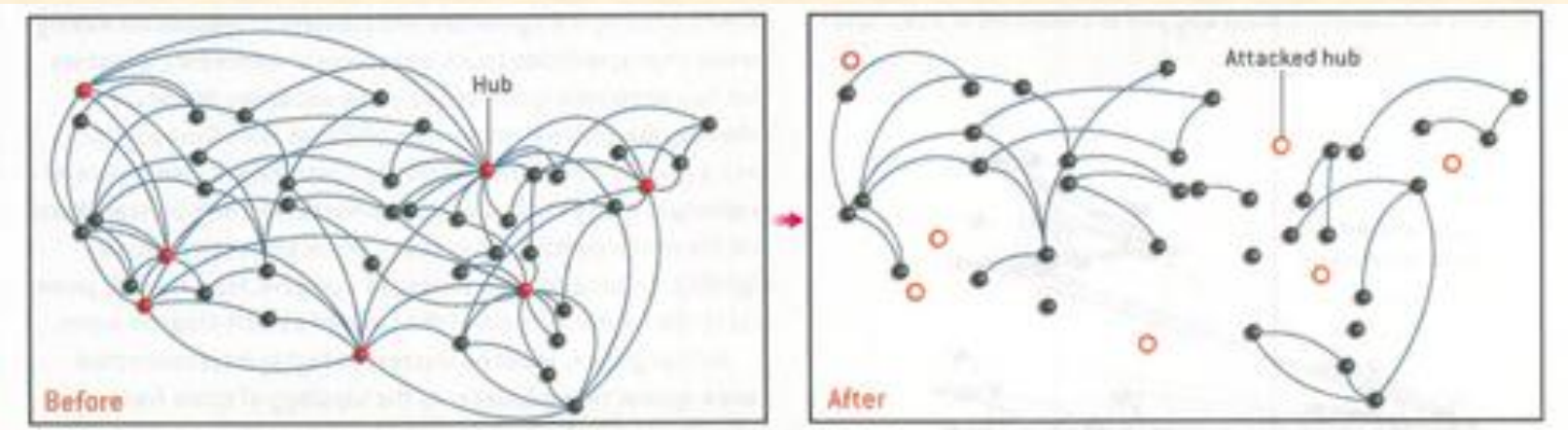


Illustration Example:

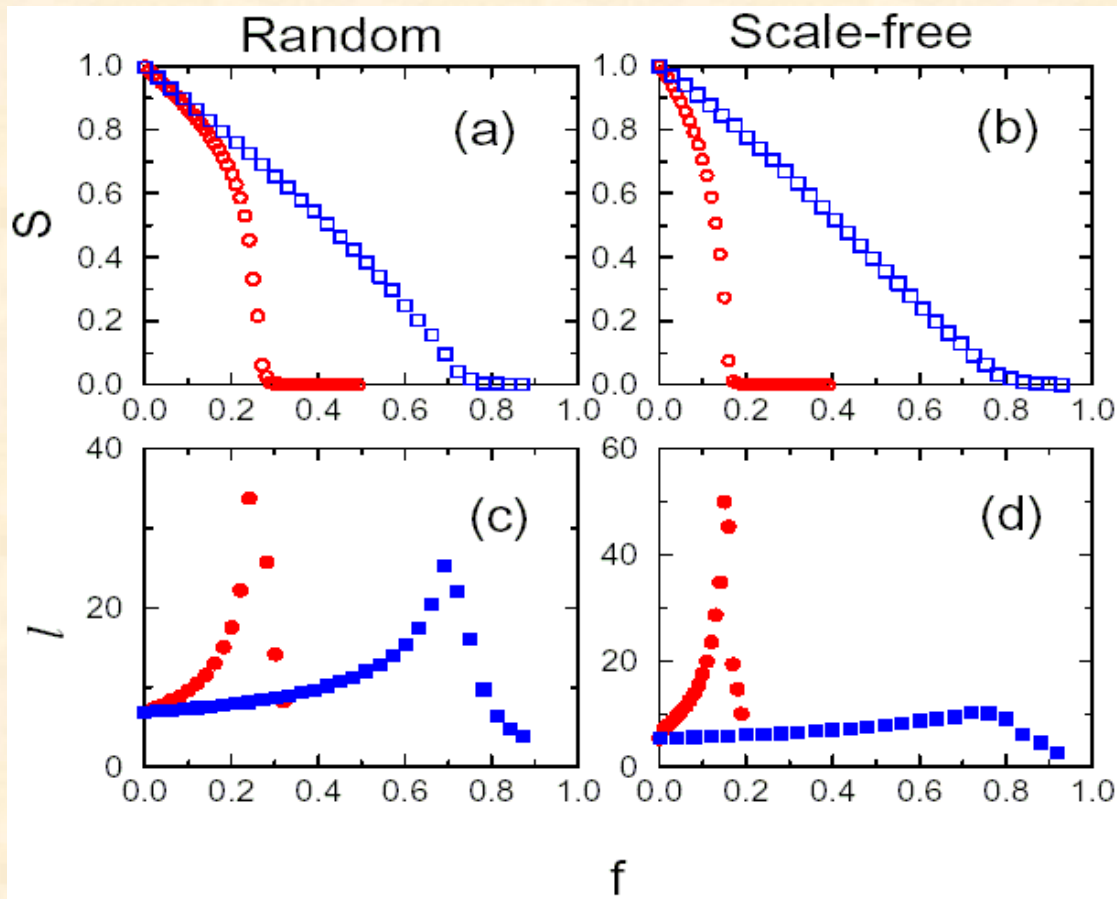


(a) Random failures



(b) Intentional attacks

Simulation Example:



S – size of largest subgraph

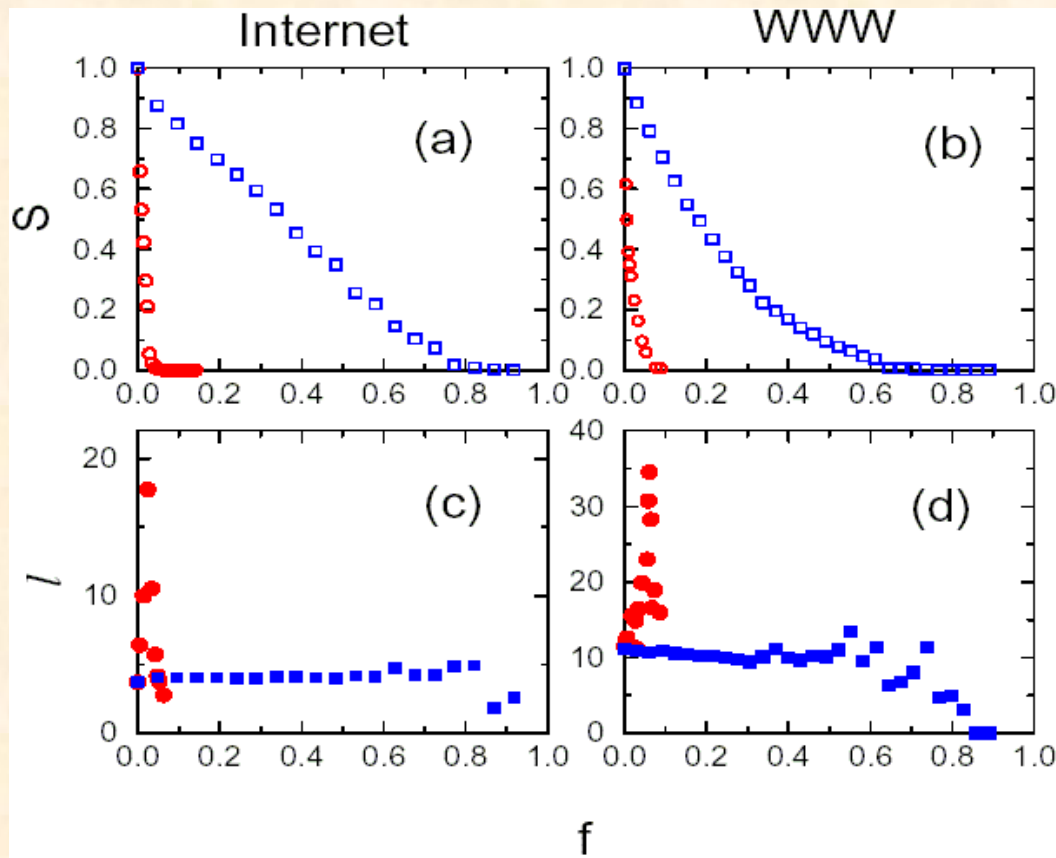
l – average path length

f – fraction of attacks

R. Albert, H. Jeong, A. L. Barabasi,
Nature, 406, 387-482 (2000)

Robustness and fragility of ER random-graph and BA scale-free networks:
 (a) and (c): ER random-graph networks; (b) and (d): BA scale-free networks;
 squares—random removal of nodes; circles—intentional removal of nodes

Internet and WWW: Simulation Results



S – size of largest subgraph

l – average path length

f – fraction of attacks

R. Albert, H. Jeong, A. L. Barabasi,
Nature, 406, 387-482 (2000)

Robustness and fragility of the Internet and WWW against intentional attacks:

(a) and (c): Internet; (b) and (d): WWW

squares—random failures, circles—intentional attacks

Computer Viruses and Worms

- ❖ The first computer **worm** was created in 1988 by the now-infamous Robert Tappan Morris
- ❖ On November 2, 1988, Robert Morris, Jr., a graduate student in Computer Science at Cornell, wrote an experimental, self-replicating, self-propagating program called a *worm* and injected it into the Internet.
- ❖ Computer worms are most aggressive cyber-organisms (larger and more sophisticated programs) with much more powerful abilities to attack computers.
- ❖ They shut down the Internet e-mail systems that got clogged with infected e-mails propagating from the worm.
- ❖ A worm is capable of sending itself to all e-mail addresses in the e-mail address book of the computer, which received an infected e-mail. This makes worms very effective in spreading over the Internet.

Two Incidents on facebook

<< See all Yahoo! Tech News

Koobface, Other Worms Target Facebook Friends (NewsFactor)

Posted on Thu Mar 5, 2009 11:31AM EST

Add articles about technology to your My Yahoo! 

- As Facebook works to make itself more relevant and timely for its growing member base with a profile page makeover, attackers seem to be working overtime to steal the identities of the friends, fans and brands that connect through the social-networking site.

Indeed, Facebook has seen five different security threats in the past week. According to Trend Micro, four new hoax applications are attempting to trick members into divulging their usernames and passwords. And a new variant of the Koobface worm is running wild on the site, installing malware on the computers of victims who click on a link to a fake YouTube video.

The Koobface worm is dangerous. It can be dropped by other malware and downloaded unknowingly by a user when visiting malicious Web sites, Trend Micro reports. When attackers execute the malware, it searches for cookies created by online social networks. The latest variant is targeting Facebook, but earlier variants have also plagued MySpace.

KOOFACE WORM TARGETS FACEBOOK

Malware Worm Spreading on Facebook – 45,000 Passwords Stolen So Far

January 5, 2012

in Internet Safety & Privacy

Like

7,597 people like this.



Seculert issued a warning today that the Ramnit worm, which has traditionally targeted financial login credentials, is now targeting Facebook users. At the time

摘要

一. 中国网民信息安全总体现状

- ✧ 整体上，我国信息安全环境不容乐观，有 74.1% 的网民在过去半年内遇到过信息安全问题，总人数达 4.38 亿。
- ✧ 手机垃圾短信/骚扰短信、骚扰电话发生比例仍然较高，在整体网民中发生比例分别达 59.2%、49.0%；其次为手机欺诈/诱骗信息，发生比例为 36.3%；再次为手机恶意软件，发生比例达 23.9%；其它方面，假冒网站/诈骗网站发生比例为 21.6%，中病毒或木马发生比例为 17.6%，个人信息泄露发生比例为 13.4%，账号或密码被盗发生比例为 8.9%。

中国顶级域名根服务器故障 大部分网站受影响

据多家DNS服务商透露，今日下午3点，全国所有通用顶级域的根服务器出现异常，导致国内大部分用户无法正确解析域名，对全国互联网链接造成系统性影响。

根服务器主要用来管理互联网的主目录。全世界只有13台，这13台根域名服务器中名字分别为“A”至“M”，其中10台设置在美国，另外各有一台设置于英国、瑞典和日本。

“简单的说，如果我们要访问baidu.com这个网站，先要指向根服务器，根服务再将用户指向.com服务器，.com的解析服务器再把用户指向baidu.com。”一位DNS技术专家解释说，这次的问题仅出现在中国，说明全球根服务器并未出现问题，问题很可能是国内网络运营商。

DNS (Domain Name System) 域名系统

中国手机安全报告：移动安全问题已全面爆发

360互联网安全中心11日发布《2013年中国手机安全报告》显示，2013年是移动安全问题进入全面爆发的新阶段：2013年安卓（Android）平台新增恶意程序样本67.1万个，较2012年增长4.4倍；用户感染恶意程序9747万人次，较2012年增长了88.3%，其中资费消耗、恶意扣费、隐私窃取和诱骗欺诈类恶意程序的感染量最高。

《报告》显示，骚扰电话、垃圾短信依然活跃。360互联网安全中心2013年共收到用户举报垃圾短信总量约为4.46亿条，360手机卫士全年共为用户拦截各类手机垃圾短信971亿条。2013年被用户标记的各类骚扰电话号码约4.22亿个，被标记的骚扰电话号码日均活跃数量为230多万个。其中，“响一声”电话以51%的比例位居骚扰电话数量的首位，广告推销电话占比为21%，而诈骗电话的占比约为18%。

《报告》显示，2013年用户手机感染恶意程序后面临直接经济损失的可能性达到了近七成。2013年还出现了“欺诈信使”“支付鬼手”“尸潮”“扣费黑帮”和“不死木马”等一系列新型高危木马，严重威胁用户手机安全。（原标题：《2013年中国手机安全报告：移动安全问题已全面爆发》）

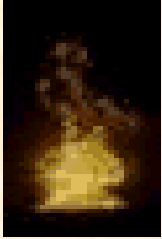


Posting Detail



Title :	Zero-day malware新形電腦病毒
Body :	Zero-day malware Dear colleagues and students, On 22 January, CityU found that a new computer virus (zero-day malware) invaded the University's email system and hacked the login names and passwords of some faculty and staff. The malware was successfully removed before it could spread further. There is no evidence that any sensitive information was compromised. Upon learning about the incident, the University immediately contacted the police, informed the Office of the Privacy Commissioner for Personal Data, and asked all affected faculty and staff to change their passwords. To further bolster the security of our computer systems, experts from our vendor have been engaged to review our email infrastructure and management processes. In addition, all faculty, staff and students are required to change their passwords regularly, and are reminded to encrypt their files when sending sensitive information through emails. Office of the Chief Information Officer 23 January 2014 新形電腦病毒 各位同事、同學： 城大在1月22日發現電郵系統被一種新形的電腦病毒（zero-day malware）入侵，盜取了一些教職員的登入名稱及密碼。大學在病毒進一步入侵系統或其他敏感資料外洩之前，已成功將之移除。 校方得悉事件後已即時向警方報案及知會個人資料私隱專員公署，並隨即通知所有受影響的用戶更改密碼。城大非常重視大學電腦系統的安全，已要求服務供應商檢討電郵系統的架構及管理程序，進一步加強保安的措施，全校教職員及學生必須定時更改密碼，並在使用電郵傳送敏感資料時須為檔案加密。 資訊總監辦公室 2014年1月28日

Another Example



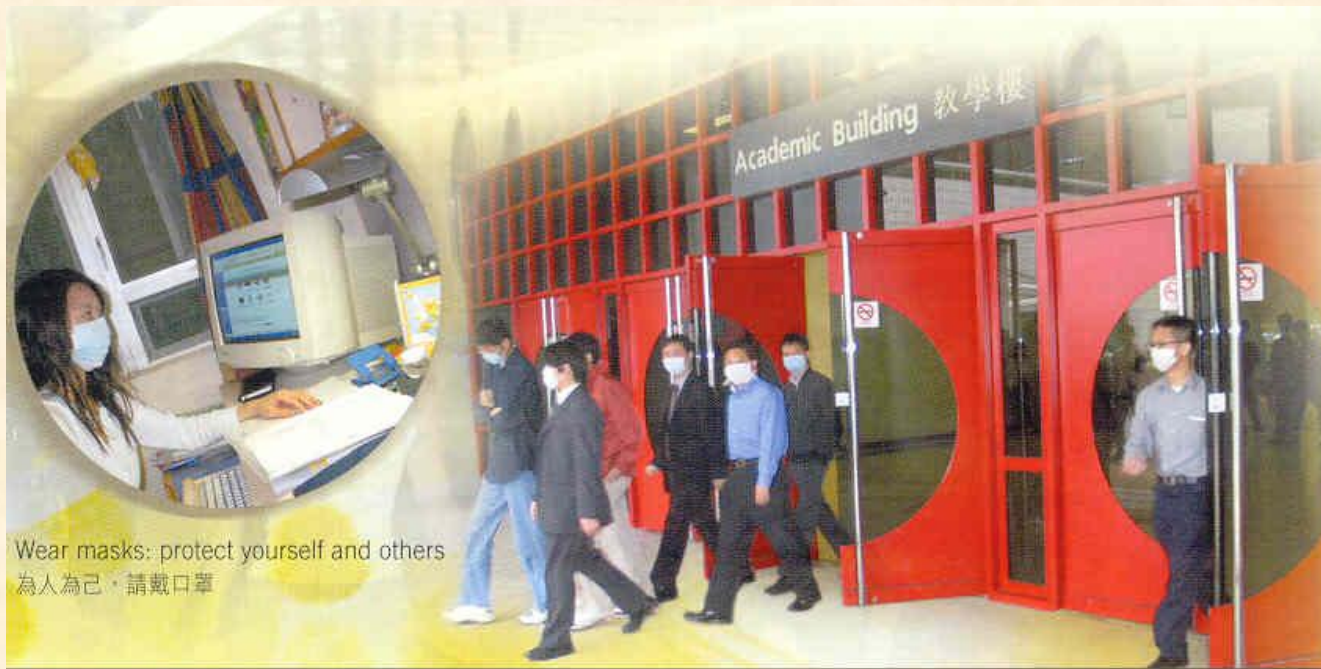
Flame (malware)

Flame, also known as Flamer, sKyWIper, or Skywiper, is modular computer malware discovered in 2012 that attacks computers running the Microsoft Windows operating system.

Reportedly, the program is being used for targeted cyber espionage (spy) in some middle eastern countries.

June 19, 2012, Washington Post published an article that Flame was jointly developed by U.S. National Security Agency, CIA and Israel's military at least five years ago as part of a classified effort code-named Olympic Games to collect intelligence in preparation for cyber-sabotage aimed at slowing Iranian nuclear efforts.

Mathematical Modeling of Epidemic Spreading



CityU of HK, 2003

SARS — First Major New Infectious Disease of the 21st Century

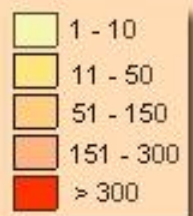
- ❖ Severe Acute Respiratory Syndrome (SARS) is a highly infectious and potentially lethal atypical form of pneumonia that begins with deceiving common flu-like symptoms
- ❖ On 16 November 2002, the first known case of SARS was discovered in Guangdong province in southern China
- ❖ The number of reported SARS cases has increased exponentially, prompting the World Health Organization (WHO) to issue a global alert on 12 March 2003

SARS

Cases Distribution in

Affected Areas

Legend



Rec.	Country	Cumulative Cases	New Cases	No. of Death	No. of Patients Discharged
1	China	2914	161	131	1299
2	China, Hong Kong SAR	1557	30	138	710
3	Singapore	199	1	23	137
4	Canada	142	0	18	77
5	China, Taiwan	66	17	0	25
6	Vietnam	63	0	5	53
7	United States	41	2	0	0
8	Italy	8	4	0	3
9	Thailand	7	0	2	5
10	Germany	7	0	0	6



Map created with ArcIMS.

0 7562km



Zoom in --Click on the map and hold the left mouse button down, draw a extend you want to zoom in, and the release the mouse.



Epidemic Models



Epidemic Models

Consider a population of N connected individuals, described by a network of N nodes.

Let the number of healthy nodes be $S(t)$, infected nodes be $I(t)$, and recovered nodes be $R(t)$ at time $t \geq 0$, initially $S(0) = S_0$, $I(0) = I_0$, and $R(0) = R_0 = 0$.

For convenience, one may normalize these variables by dividing them by N , so they also have the meaning of “density” (percentage), with $S(t) + I(t) + R(t) = 1$ for all $t \geq 0$.

Assume that the lifetime of a virus or disease is much shorter than the lifetime of the nodes in the population; therefore, birth and death of nodes are not taken into consideration in the following study.

Suppose that the probability of a node becoming “infected” from “susceptible” is ν , and the probability of a node being cured and becomes “susceptible” again is δ . Then, define the virus *effective spreading rate* as

$$\lambda = \frac{\nu}{\delta} \tag{5-1}$$

One extreme case is that all cured nodes immediately become susceptible again, which corresponds to $\delta = 1$, thus $\lambda = \nu$ = probability of a node becomes infected.

SI Model

Suppose $R(t) = 0$ for all $t \geq 0$. Consider the rate of increase in the density of infected nodes, from time t to time $t + \Delta t$, which is given by

$$\frac{I(t + \Delta t) - I(t)}{\Delta t} \approx [\lambda S(t)]I(t)$$

In the limit $\Delta t \rightarrow 0$, one has $\frac{dI(t)}{dt} = \lambda S(t)I(t)$. Since $S(t) = 1 - I(t)$, this becomes the so-called Logistic model:

$$\frac{dI(t)}{dt} = \lambda[1 - I(t)]I(t), \quad I(0) = I_0 \quad (5-2)$$

This equation has a solution

$$I(t) = \frac{1}{1 + (I_0^{-1} - 1)e^{-\lambda t}} \quad (5-3)$$

Clearly, $I(t) \rightarrow 1$ as $t \rightarrow \infty$, meaning that all nodes are infected eventually. This is true if no infected nodes are cured and then recovered in the process.

SIS Model

Again, consider the case of $R(t) = 0$. But, now, suppose that the rate of recovery from “infected” back to “susceptive” is δ , namely, $S(t) = \delta I(t)$. It then follows from (5-2) that

$$\frac{dI(t)}{dt} = \lambda[1 - I(t)]I(t) - \delta I(t), \quad I(0) = I_0 \quad (5-4)$$

This equation has a solution

$$I(t) = \begin{cases} \left\{ \frac{1}{1 - \frac{\delta}{\lambda}} + \left(I_0^{-1} - \frac{1}{1 - \frac{\delta}{\lambda}} \right) e^{-\left(1 - \frac{\delta}{\lambda}\right)t} \right\}^{-1}, & \delta \neq \lambda \\ \frac{1}{\lambda t + I_0^{-1}}, & \delta = \lambda \end{cases} \quad (5-5)$$

There are three cases:

- (i) $\delta > \lambda$: This means that recovering is faster than infecting. It follows from (5-5) that $I(t) \rightarrow 0$ as $t \rightarrow \infty$, implying that the virus will die out eventually.
- (ii) $\delta < \lambda$: This means that recovering is slower than infecting. It follows from (5-5) that $I(t) \rightarrow \left(1 - \frac{\delta}{\lambda}\right)$ as $t \rightarrow \infty$, implying that a certain fraction of nodes will remain being infected. In particular, if $\delta \ll \lambda$ then $I(t) \sim 1$ as $t \rightarrow \infty$, implying that almost all nodes will be infected eventually.
- (iii) $\delta = \lambda$: This is a threshold, determining whether or not the virus will spread.

SIR Model

Now, suppose $R(t) > 0$, namely, some nodes will be cured therefore recovered

In this case, one can similarly establish the following model:

$$\begin{cases} \frac{dI(t)}{dt} = \lambda S(t)I(t) - \delta I(t) \\ \frac{dS(t)}{dt} = -\lambda S(t)I(t) \\ \frac{dR(t)}{dt} = \delta I(t) \\ I(0) = I_0, \quad S(0) = S_0, \quad R(0) = 0 \end{cases} \quad (5-6)$$

By qualitative analysis or numerical solutions, one can find the following:

- (i) $S_0 > \frac{\delta}{\lambda}$: Virus will spread out for some time, but eventually die out.
- (ii) $S_0 \leq \frac{\delta}{\lambda}$: Virus will die out quickly.

Threshold Theory in Epidemic Spreading

The SIR epidemic spreading model on a network

Consider a population of N individuals, called nodes hereafter, each is either in the Susceptible (S), or Infected (I), or Recovered (R) state.

Let $\rho_{S,k}(t)$, $\rho_{I,k}(t)$ and $\rho_{R,k}(t)$ be the densities of the S, I and R states of nodes of degree k at time t , respectively.

Then, these densities are related by means of the normalization condition

$$\rho_{S,k}(t) + \rho_{I,k}(t) + \rho_{R,k}(t) = 1 \text{ for all } t \geq 0$$

Define the total number of recovered nodes at time t by

$$R(t) = \sum_k P(k) \rho_{R,k}(t)$$

where $P(k)$ is the node-degree distribution. Denote its prevalence (steady state) by

$$R_\infty = \lim_{t \rightarrow \infty} R(t)$$

Threshold Theory in Epidemic Spreading

The SIR epidemic spreading model on a network

At the mean-field level, for undirected random-graph type of uncorrelated sparse networks, the above densities satisfy the following system of differential equations:

$$\begin{aligned}\frac{d\rho_{I,k}(t)}{dt} &= -\rho_{I,k}(t) + \lambda k \rho_{S,k}(t) \Theta(t) \\ \frac{d\rho_{S,k}(t)}{dt} &= -\lambda k \rho_{S,k}(t) \Theta(t) \\ \frac{d\rho_{R,k}(t)}{dt} &= \rho_{I,k}(t)\end{aligned}\tag{1}$$

where λ is the effective spreading rate and $\Theta(t)$ represents the average density of infected nodes pointed at by any given edge.

Threshold Theory in Epidemic Spreading

The SIR epidemic spreading model on a network

To find $\Theta(t)$ --

In general, the probability that an edge points to an infected node with degree k is proportional to $kP(k)\rho_{I,k}(t)$

Notice that the infected node pointed by the edge has previously received virus through an edge that cannot be used for virus transmission anymore (since its originating node had been recovered), the above probability should be modified to be $(k-1)P(k)$. Consequently, one has

$$\Theta(t) = \frac{1}{\langle k \rangle} \sum_k (k-1)P(k)\rho_{I,k}(t) \quad (2)$$

A combination of system (1) and equation (2), along with the natural initial conditions

$$\rho_{R,k}(0) = 0, \quad \rho_{I,k}(0) = \rho_k^0 \quad \text{and} \quad \rho_{S,k}(0) = 1 - \rho_k^0$$

completely define the SIR model of virus spreading on any random-graph type of uncorrected network with the node-degree distribution $P(k)$.

Threshold Theory in Epidemic Spreading

The SIR epidemic spreading model on a network

Assume that initially infected nodes are uniformly distributed, namely, $\rho_k^0 = \rho^0$ for all k at time $t = 0$. In this case, in the limit $\rho^0 \rightarrow 0$, namely, starting from very few infected nodes, one has $\rho_{I,k}^0(0) \approx 0$ and $\rho_{S,k}(0) \approx 1$. Under these assumptions, the second and third equations in system (1) can be directly integrated, yielding

$$\rho_{S,k}(t) = e^{-\lambda k \phi(t)} \quad \text{and} \quad \rho_{R,k}(t) = \int_0^t \rho_{I,k}(\tau) d\tau \quad (3)$$

where

$$\phi(t) = \int_0^t \Theta(\tau) d\tau = \frac{1}{\langle k \rangle} \sum_k (k-1) P(k) \rho_{R,k}(t)$$

which, based on (3), gives

$$\begin{aligned} \frac{d\phi(t)}{dt} &= \frac{1}{\langle k \rangle} \sum_k (k-1) P(k) \rho_{I,k}(t) \\ &= \frac{1}{\langle k \rangle} \sum_k (k-1) P(k) [1 - \rho_{R,k}(t) - \rho_{S,k}(t)] \\ &= 1 - \frac{1}{\langle k \rangle} - \phi(t) - \frac{1}{\langle k \rangle} \sum_k (k-1) P(k) e^{-\lambda k \phi(t)} \end{aligned} \quad (4)$$

Threshold Theory in Epidemic Spreading

The SIR epidemic spreading model on a network

If equation (4) can be solved, then one can obtain the total epidemic prevalence (steady-state) $\rho_{S,k}(\infty)$ as a function of $\phi_\infty = \lim_{t \rightarrow \infty} \phi(t)$, so that by $\rho_{R,k}(\infty) = 1 - \rho_{S,k}(\infty)$ one obtains

$$R_\infty = \sum_k P(k) [1 - e^{-\lambda k \phi_\infty}]$$

However, for a general node-degree distribution $P(k)$, equation (4) actually cannot be solved. Nevertheless, one can still obtain some useful information about the asymptotic behavior of the virus spreading. Since $\rho_{I,k}(\infty) = 0$ and consequently $\lim_{t \rightarrow \infty} d\phi(t)/dt = 0$, from equation (4) one obtains the following steady-state equation:

$$\phi_\infty = 1 - \frac{1}{\langle k \rangle} - \frac{1}{\langle k \rangle} \sum_k (k-1) P(k) e^{-\lambda k \phi_\infty}$$

which has a fixed-point solution $\phi_\infty = 0$.

Threshold Theory in Epidemic Spreading

The SIR epidemic spreading model on a network

In order to have a nonzero solution for ϕ_∞ , which would mean a prevalence $R_\infty > 0$, it is necessary that

$$\left. \frac{d}{d\phi_\infty} \left[1 - \frac{1}{\langle k \rangle} - \frac{1}{\langle k \rangle} \sum_k (k-1) P(k) e^{-\lambda k \phi_\infty} \right] \right|_{\phi_\infty=0} \geq 1$$

This condition implies that

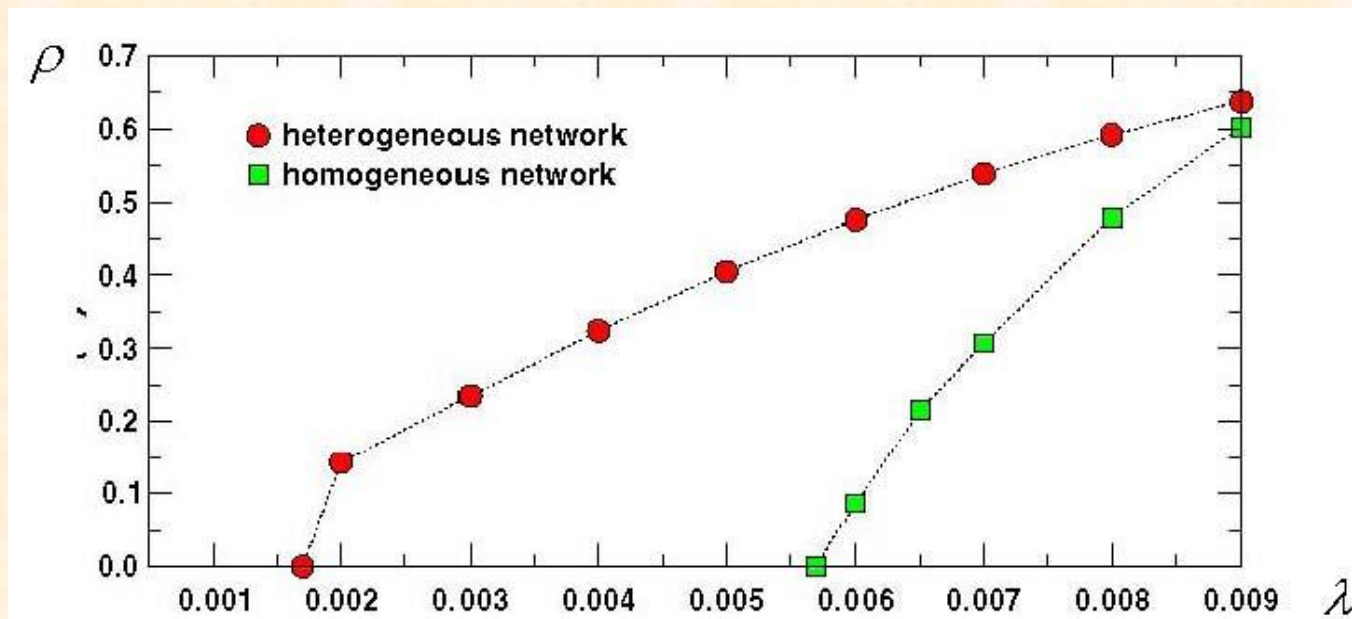
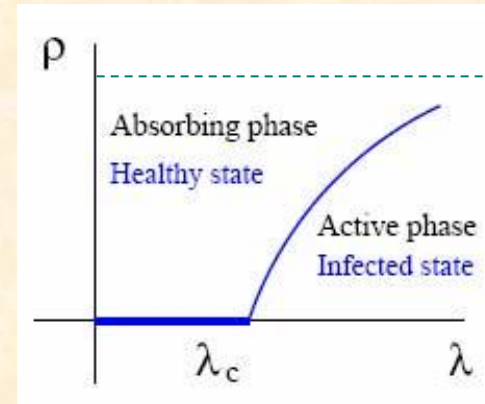
$$\frac{\lambda}{\langle k \rangle} \sum_k k(k-1) P(k) \geq 1$$

yielding the epidemic threshold

$$\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle} \quad (5)$$

below which $R_\infty = 0$ and above which $R_\infty > 0$.

Virus Spreading Threshold



Epidemic Threshold on Homogenous Networks

Consider the SIS model: $S \rightarrow I \rightarrow S \rightarrow I \rightarrow S \rightarrow \dots$

- ❖ Suppose that the probability of a node becoming “infected” from being “susceptible” is ν , and the probability of a node being cured and becomes “susceptible” again is δ . Define the virus effective spreading rate as

$$\lambda = \frac{\nu}{\delta}$$

- ❖ Let the density of infected population at time t be $\rho(t)$, with $\rho(t) \rightarrow \rho \quad (t \rightarrow \infty)$

Epidemic Threshold on Homogenous Networks

Assume:

- ❖ **Homogeneity:** The network is homogeneous (such as an random-graph network or a small-world network), For simplicity, assume that each node has degree $k_i \approx \langle k \rangle$
- ❖ **Homogenous mixing:** The infection strength is proportional to the population density; or equivalently, both ν and δ are constant
- ❖ **The lifetime of the disease** is much shorter than the **lifetime of the nodes** in the population, so birth and death of nodes are not taken into consideration

Epidemic Threshold on Homogenous Networks

Then, the epidemic spreading equation is:

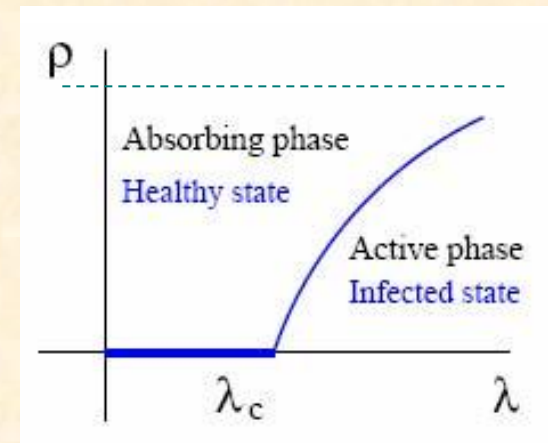
$$\frac{d\rho(t)}{dt} = -\rho(t) + \lambda\rho(t)[1 - \rho(t)] \cdot \langle k \rangle$$

It yields, as $t \rightarrow \infty$

$$\rho = \begin{cases} 0 & \lambda < \lambda_c \\ \frac{\lambda - \lambda_c}{\lambda} & \lambda \geq \lambda_c \end{cases}$$

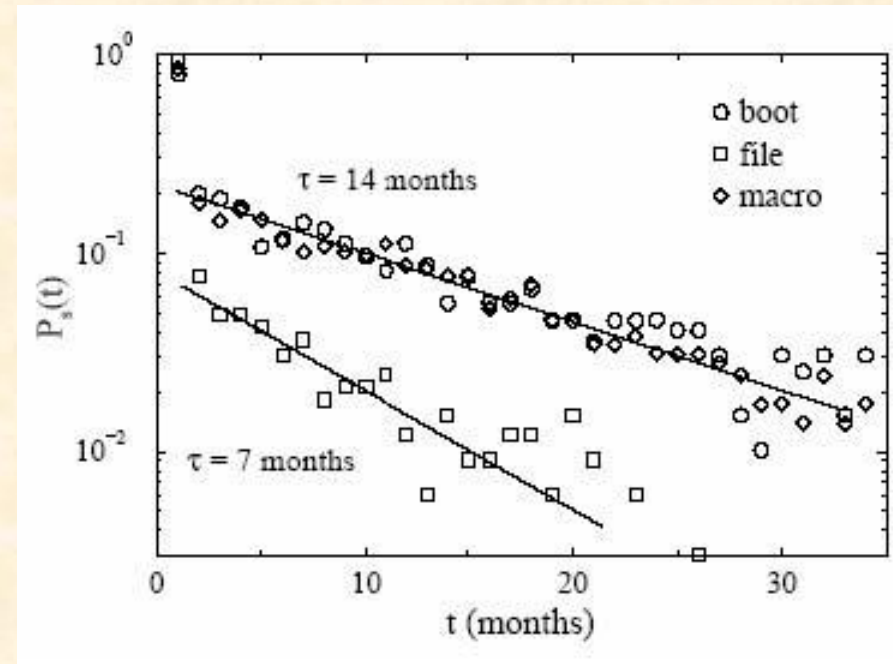
with the spreading threshold

$$\lambda_c = \frac{1}{\langle k \rangle}$$



Epidemic Threshold on Homogenous Networks

- ❖ Define the virus surviving probability $P_s(t)$ to be the percentage of nodes that are alive since birth till the current time t
- ❖ Figure shows the computer virus surviving probabilities in the 50 months from February 1996 to March 2000, attacked by boot-sector, file and macro viruses, respectively.



(Pastor-Satorras and Vespignani, 2001)

Epidemic Threshold on Scale-Free Networks

The virus spreading equation is modified to:

$$\frac{\partial \rho_k(t)}{\partial t} = -\rho_k(t) + \lambda k [1 - \rho_k(t)] \Theta(\rho_k(t))$$

Here, $\Theta(\rho_k(t))$ is the probability of an edge connecting to an infected degree- k node.

Solution:

$$\rho_k = \frac{k \lambda \Theta(\lambda)}{1 + k \lambda \Theta(\lambda)}$$

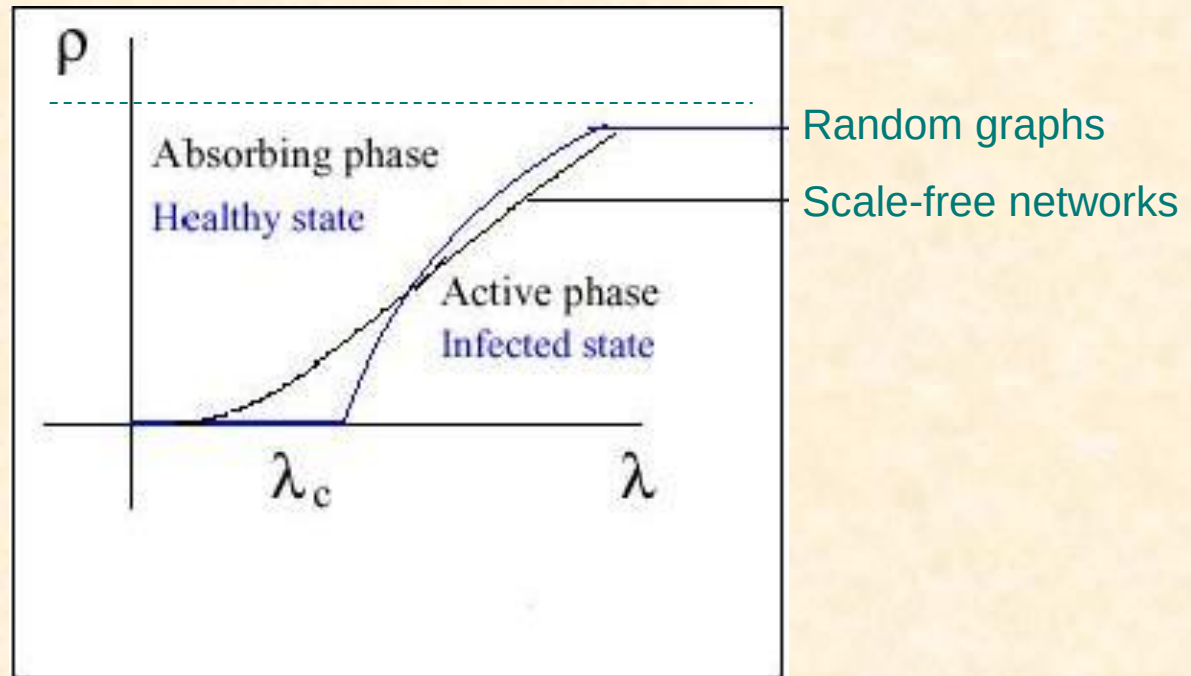
Threshold:

$$\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle}$$

→ Higher-degree nodes have higher probabilities to be infected.

→ as $N \rightarrow \infty$ → $k \rightarrow \infty$
→ $\lambda_c \rightarrow 0$

Epidemic Threshold on Scale-Free Networks



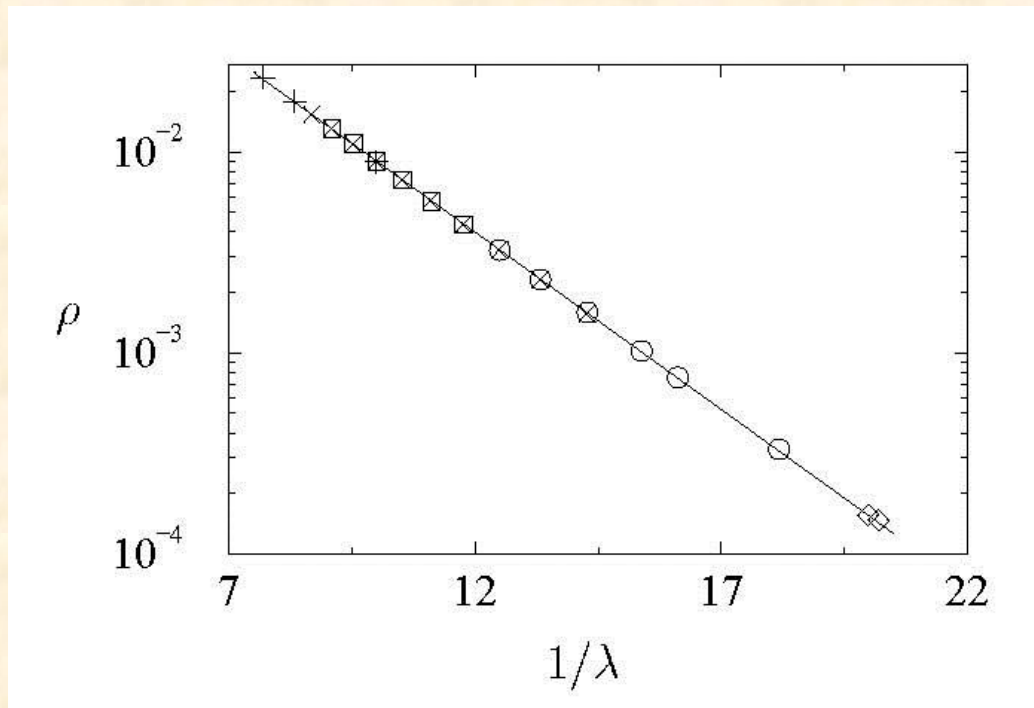
λ is the spreading rate
 λ_c is the threshold

- Classical theory: $\lambda_c > 0$
- Scale-free network: $\lambda_c = 0$

(Pastor-Satorras, 2001)

Epidemic Threshold on Scale-Free Networks

In a BA network, ρ and $1/\lambda$ have a power-law relation



+	$N = 10^5$
square	$N = 5 \times 10^5$
x	$N = 10^6$
circle	$N = 5 \times 10^6$

A BA network with
different sizes

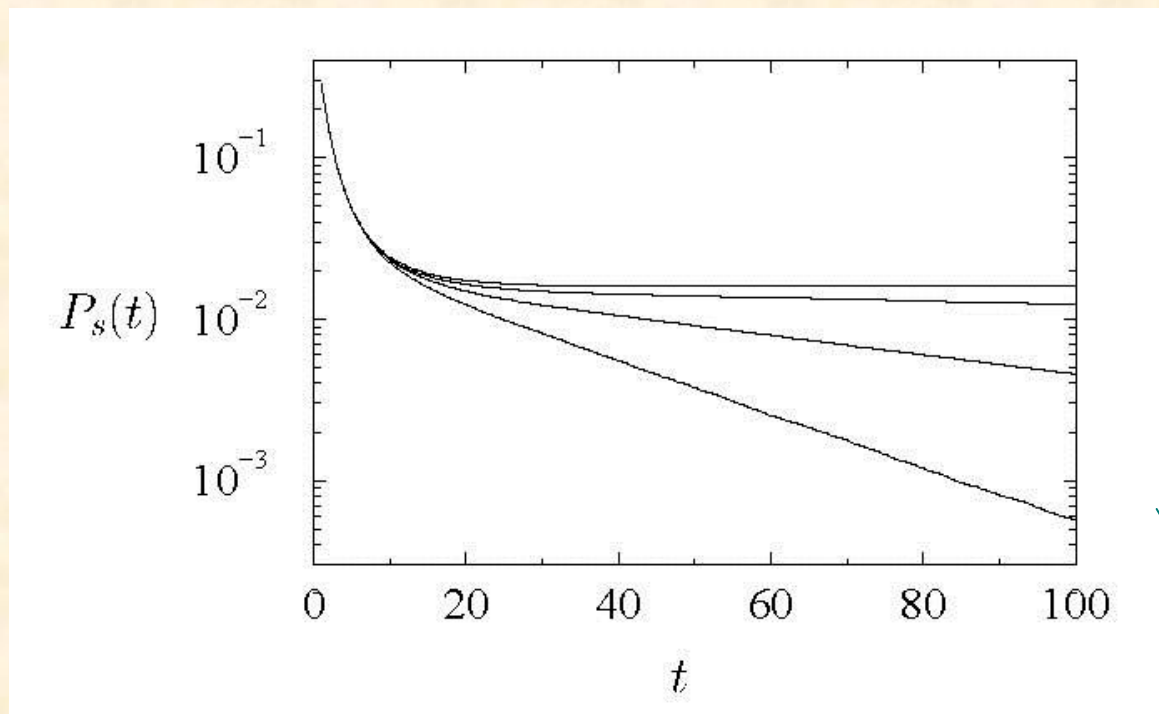
Different-sized networks follow the same power law

(Pastor-Satorras and Vespignani, 2003)

Epidemic Threshold on Scale-Free Networks

In the same BA network:

Evolution of virus surviving probability



Network sizes
increase from top
to bottom

BA network with
different sizes

(Pastor-Satorras and
Vespignani, 2003)

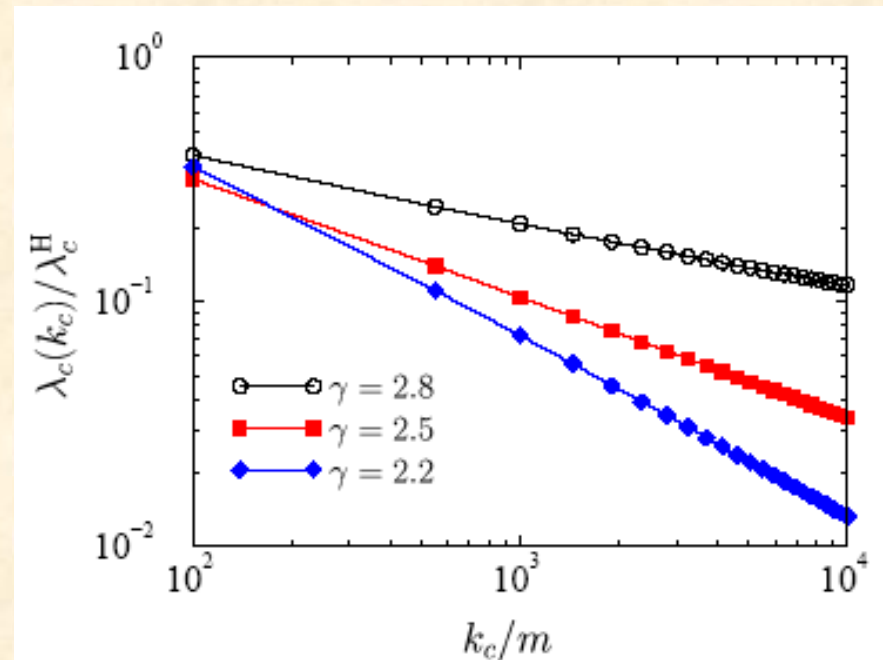
→ The bigger the network, the less virus to survive

Epidemic Threshold on Scale-Free Networks

Finite-size effect (for a not-too-large N (so, not-too-large k):

It shows a comparison of the epidemic threshold values on a finite-size scale-free network and on a corresponding homogeneous network, λ_c versus λ_c^H

It can be seen that the threshold of the former is only about 1/10 of the latter, and that as k_c (i.e., N) increases, the threshold can still tend to zero, implying the fragility of finite-sized scale-free networks against virus spreading.



(Pastor-Satorras and Vespignani, 2003)

Immunization on Scale-Free Networks

- ❖ Since scale-free networks are fragile to virus attacks, causing wide and serious outbreak, immunization becomes especially important for this type of networks
- ❖ There are three effective immunization strategies:
 - Random Immunization (or, Uniform Immunization)
 - Targeted Immunization (or, Selective Immunization)
 - Acquaintance Immunization

Random Immunization

Let the **immunity** (density of immunized nodes) be denoted by g .

Random immunization strategy: the immunization threshold is

$$g_c = 1 - \frac{\lambda_c}{\lambda}$$

where $\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle}$ is the effective epidemic threshold, giving

$$g_c = 1 - \frac{1}{\lambda} \frac{\langle k \rangle}{\langle k^2 \rangle}$$

Clearly, $N \rightarrow \infty \rightarrow \langle k^2 \rangle \rightarrow \infty \rightarrow \lambda_c \rightarrow 0 \rightarrow g_c \rightarrow 1$

→ in order to immune a scale-free network, almost every node has to be immunized

→ random immunization for scale-free networks is inefficient and expensive

Targeted Immunization

Targeted immunization strategy: to immune the most highly connected nodes, thereby reducing or completely blocking virus spreading

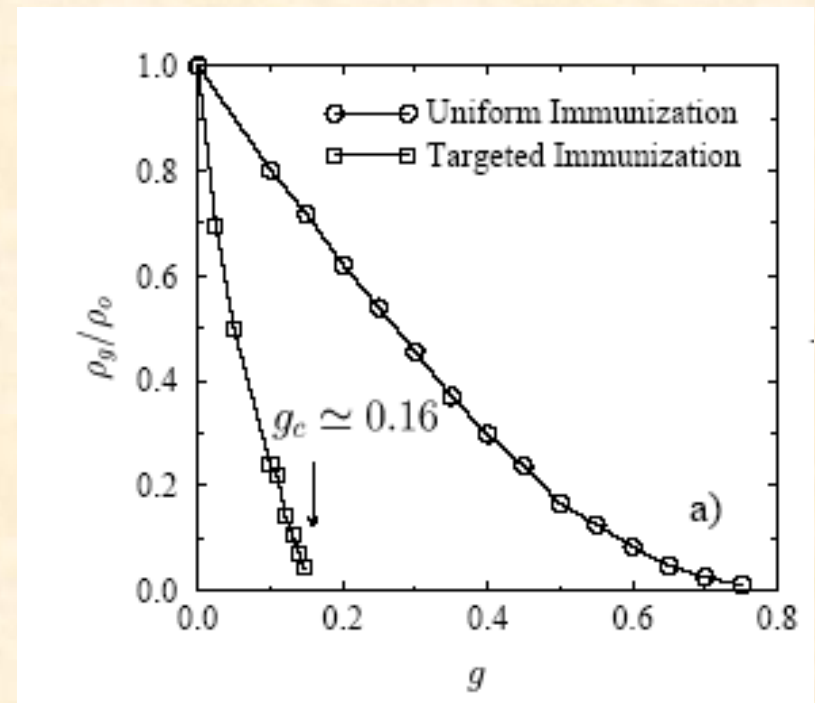
For BA scale-free networks, the immunization threshold is

$$g_c \sim e^{-\frac{2}{m\lambda}}$$

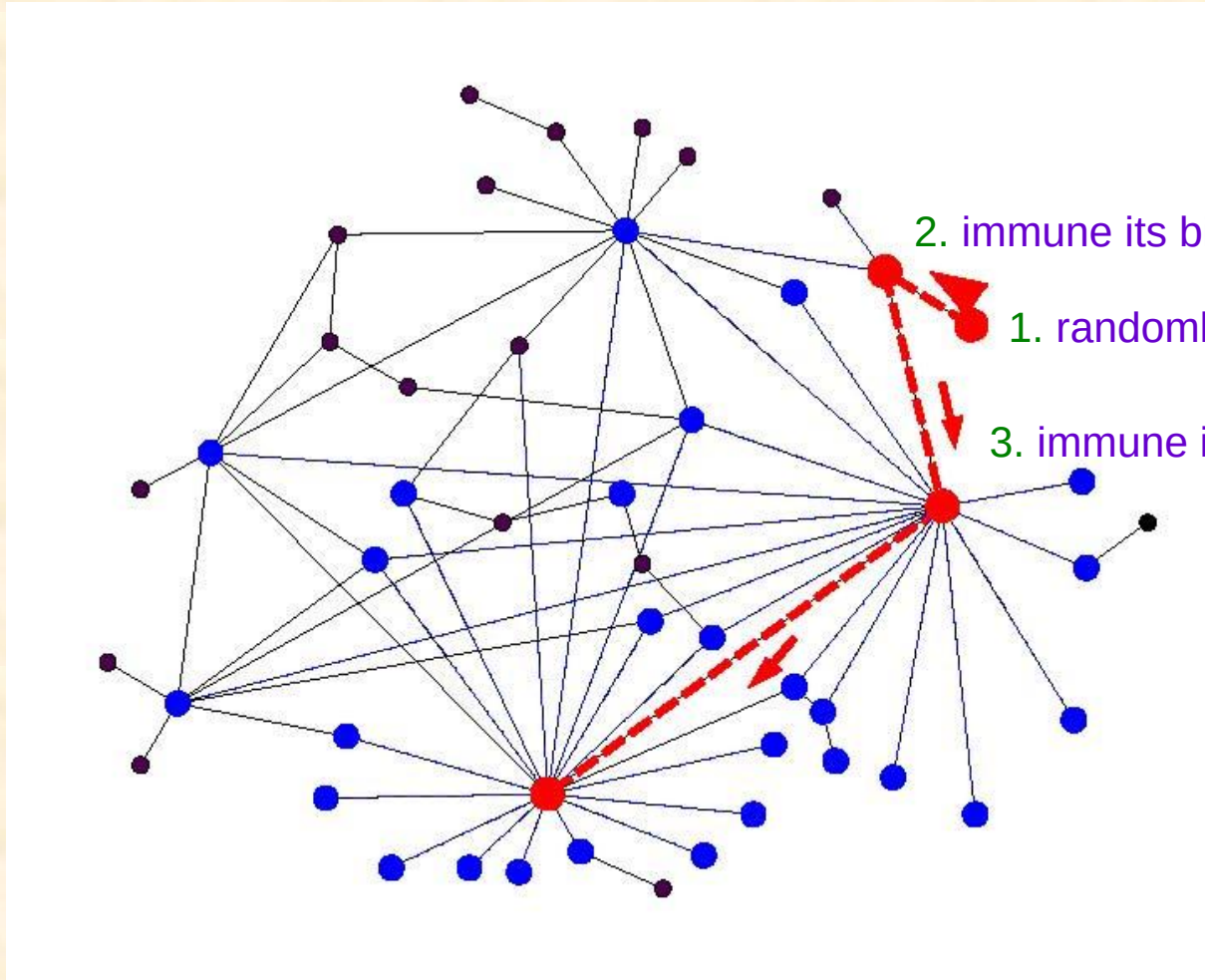
→ a very small immunization threshold can be obtained for a very large range of effective spreading rate λ

→ for scale-free networks, targeted immunization has a much smaller immunization threshold than random immunization, hence is more effective

Satorras and Vespignani (2003)



Acquaintance Immunization



Idea:

1. → 2. → 3. →

2. immune its biggest neighbor

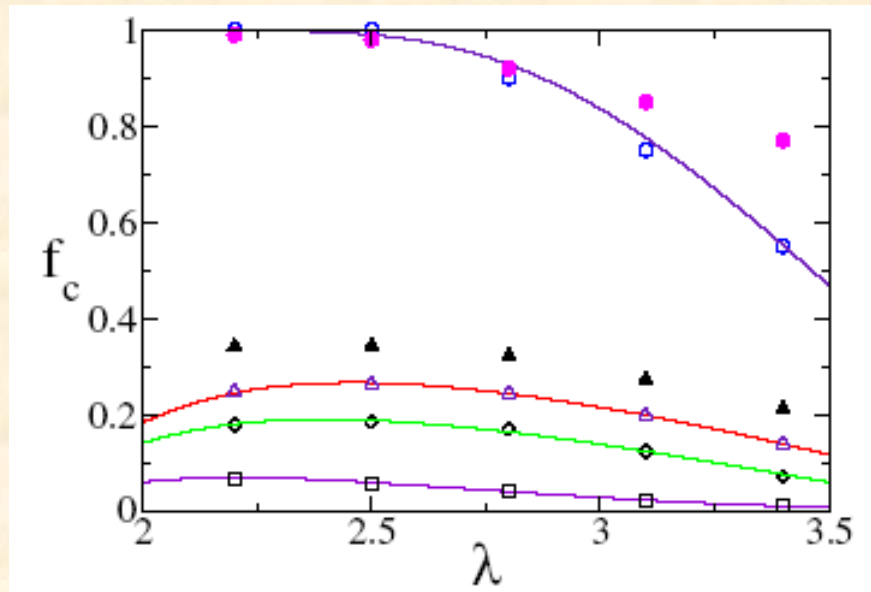
1. randomly select one node

3. immune its biggest neighbor

Immunization on Scale-Free Networks

Acquaintance immunization strategy: to randomly select a fraction of nodes from the population, and for each node in the fraction select a biggest one of its neighbors for immunization

This scheme is a localized strategy -- only requires information about the randomly selected node and its neighbors, but not the global network



empty circles – random
empty triangles – acquaintance
empty diamonds – two-acquaintance
empty squares – targeted

solid circle and solid triangles
— assortative networks
(random) (acquaintance)
 f_c — immunization threshold

Madar et al. (2004)

Computer Virus Spreading over the Internet

Now, let N be the number of susceptible nodes in the network, $a = a(t)$ be the fraction of infected nodes at time t , and K be the rate of infection

Since the total number of infected nodes is Na , and an infected node infects a total of $K(1 - a)$ nodes each time, the rate of total nodes to be infected in the next short time duration, dt , will be $(Na)K(1 - a)$, giving

$$\frac{d(Na)}{dt} = (Na)K(1 - a) \quad \rightarrow \quad \frac{da}{dt} = Ka(1 - a)$$

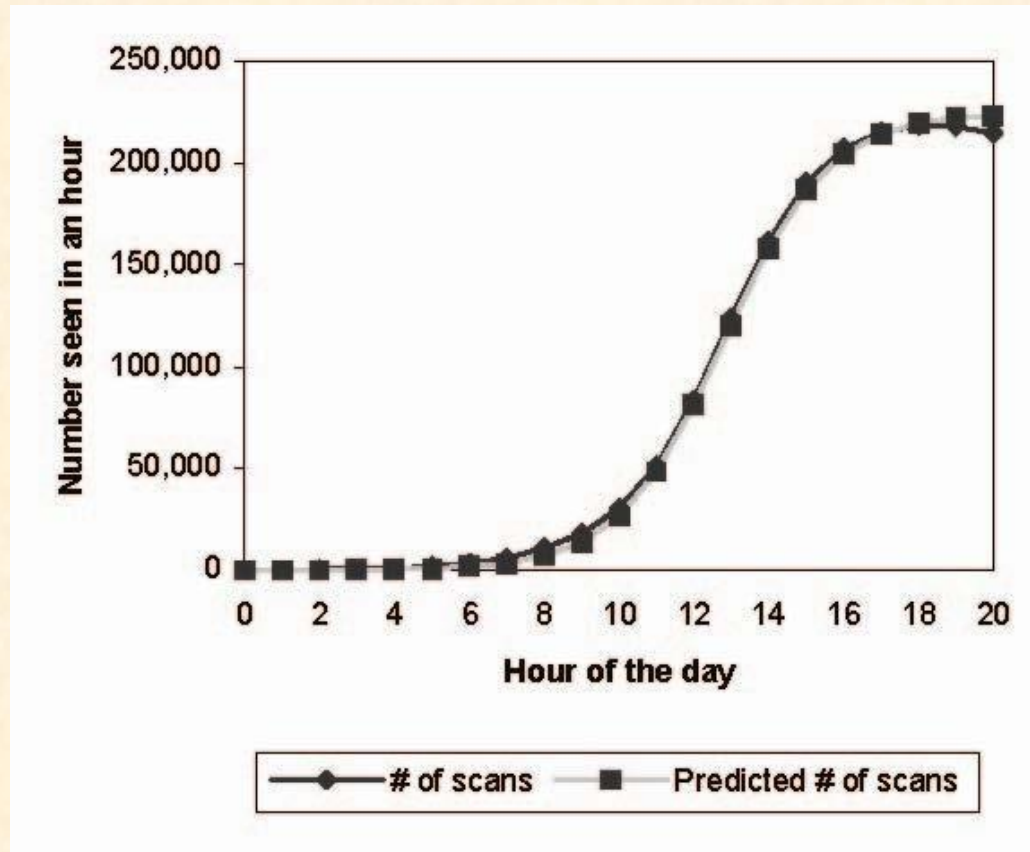
which is the **logistic equation**, with solution

$$a = \frac{e^{K(t-t_0)}}{1 + e^{K(t-t_0)}}$$

Clearly, $t \rightarrow \infty \rightarrow a \rightarrow 1 \rightarrow$ all nodes will be infected eventually

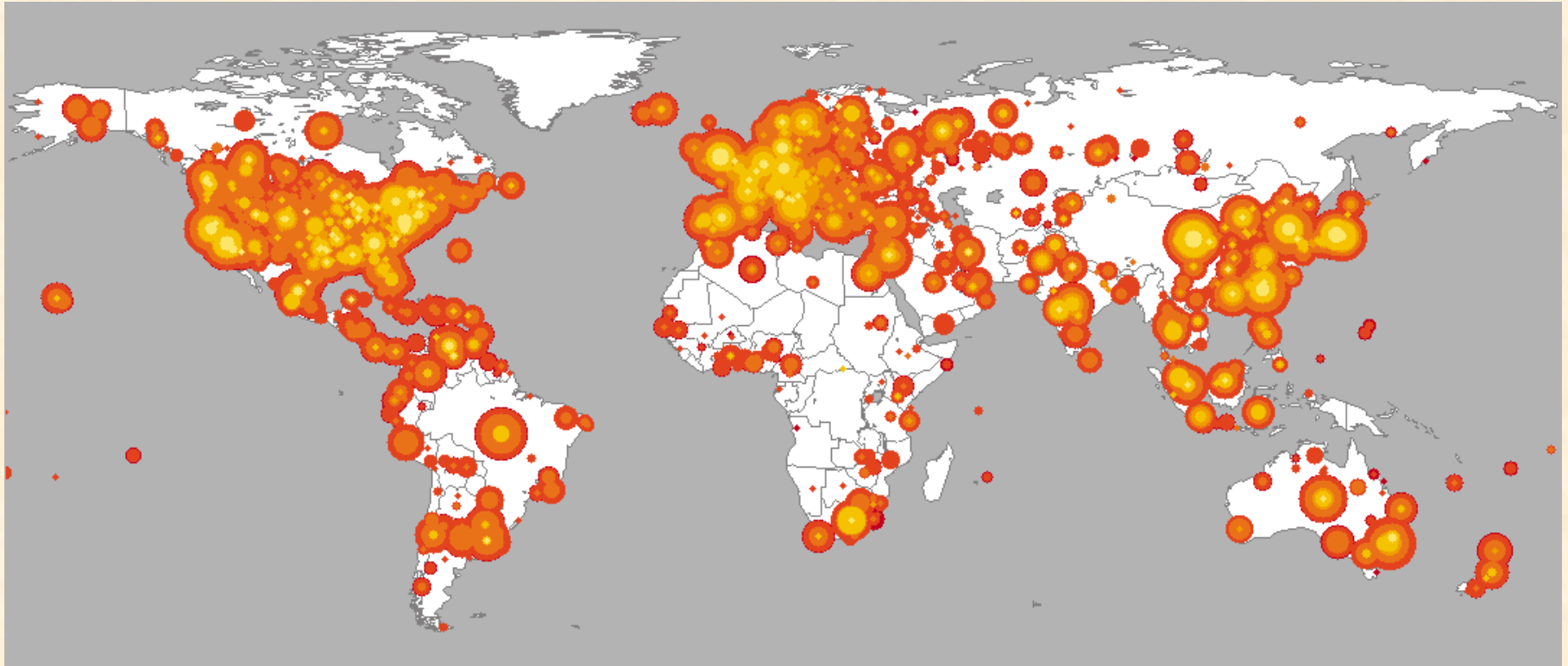
This formula of a can be used to predict a computer virus spreading

Computer Virus Spreading over the Internet



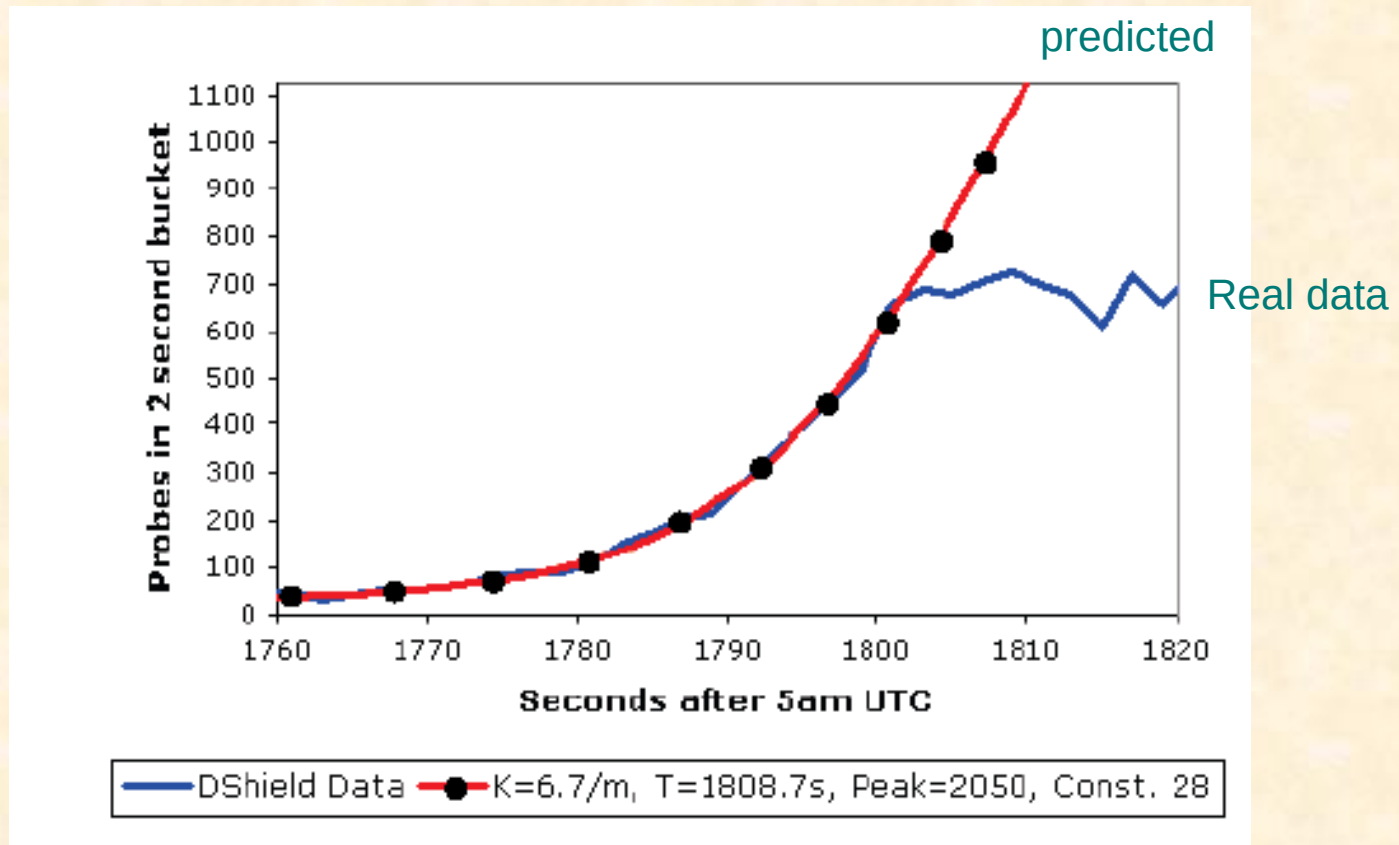
Code-Red Worm: Hourly probe rate data during the reemergence
on 1 August 2002 (staniford et al., 2002)

Code-Red Worm Spreading



Code-Red Data

Modeling the Computer Worms



Slammer Worm: real data versus model prediction

19 July 2001 (staniford et al., 2002)

Modeling the Computer Worms

A Compartment-Based Model of Computer Worms

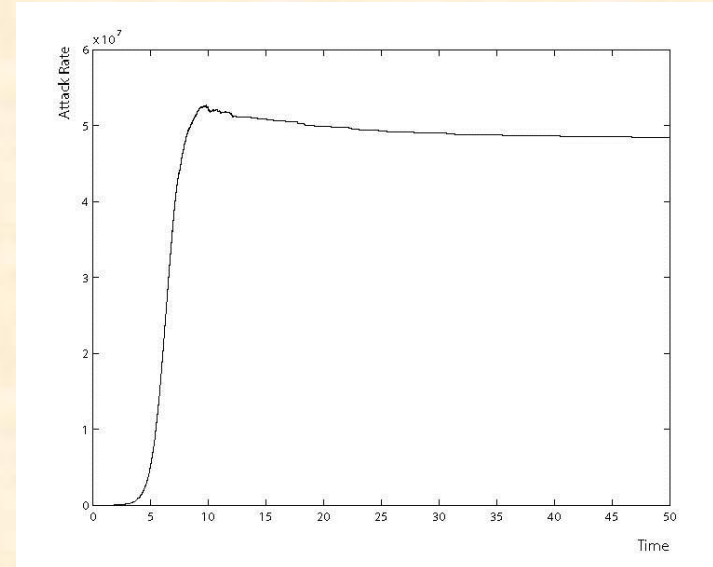
Let N_i be the number of susceptible hosts in the i th AS
and a_i be the fraction of the infected hosts among them

Let K be the (constant) average rate of spreading spread

Then, the model is

$$da_i / dt = \left[a_i K \frac{N_i}{N} + \sum_{\substack{j=1 \\ j \neq i}} \frac{N_j}{N} a_j K \right] (1 - a_i)$$

$$i = 1, \dots, N$$



Serazzi and Zanero (2004)

Spreading Models of Email Viruses

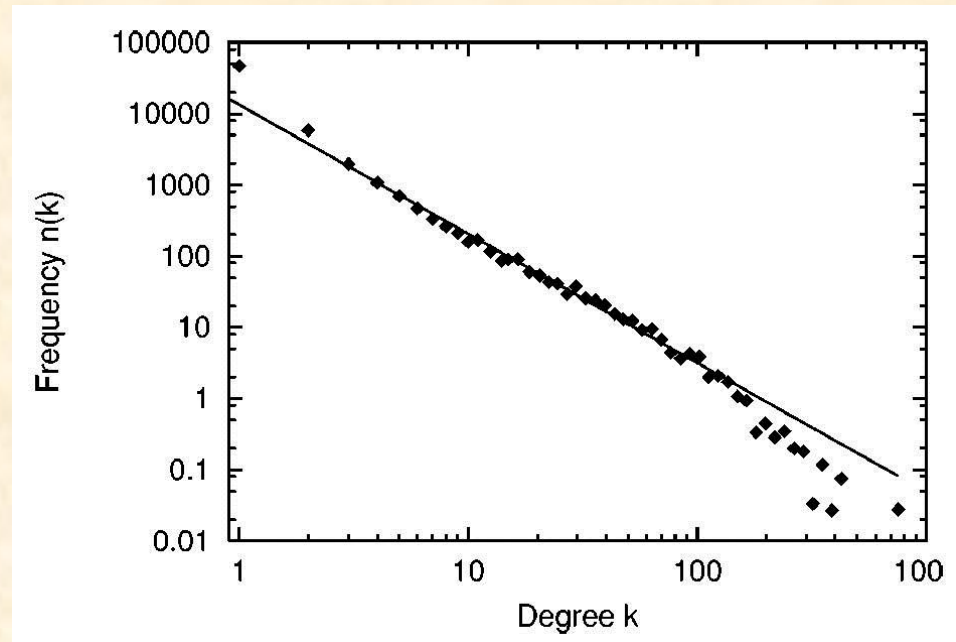
❖ Typical examples:

- Milessa (1999)
- I-Love-You (2000)
- Nimda (2001)
- Win32/Sircam (2001)

❖ Chinese versions:

- Worm.Klez.cn.b (2002)
- Worm.SoBig.c (2003)
- Worm.Mimail.C (2003)
- SCO bomb (2004)

➤



Node-degree distribution of the email network
in Kiel University, Germany, 2002

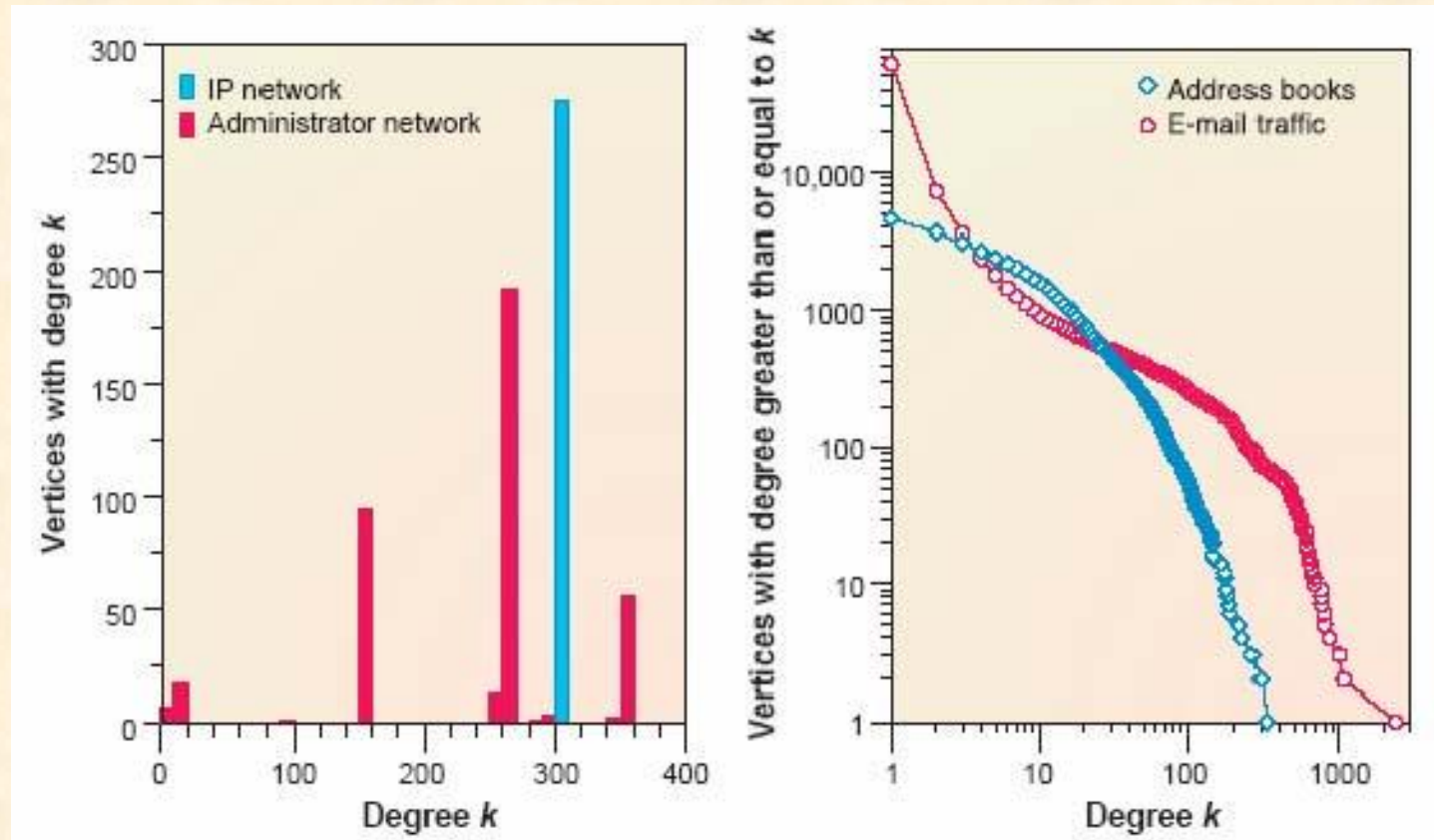
$N = 59812$ $\langle k \rangle = 2.88$ $\gamma = -1.81$

Ebel et al (2002)

On the contrary -

Virus Spreading Can also Change the Network Topology

Different virus replication and spreading patterns may lead to their different traveling-path topologies on the underlying networks



Balthrop et al. (2004)

It is desirable to have a control strategy that is immune to the topological changes

BREAK

10 minutes

Cascading Reactions

- ❖ The phenomenon of *cascading reactions (failures)*, is quite similar to *virus spreading* over various complex networks
- ❖ **Similarity:** They typically lead to collapse of a large portion of the network or even the entire network, called *avalanche*
- ❖ **Differences:**
 - As long as the network falls apart into pieces, the network is considered failed, while in virus spreading infected individuals are the ultimate concerns
 - In cascading failures, loads (weights) on nodes are taken into main consideration, but usually not so in virus spreading

Real Examples of Cascading Failures

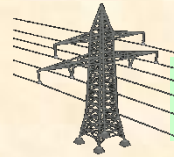


Internet

October 1986:
the first documented
Internet collapse
due to congestion



Drop in speed
of a factor 100



Power grid

August 1996:
sag of just one
electric line in
Oregon USA



Blackout affected 4
million people in 9
different states

August 2003:
local failure
in Ohio USA



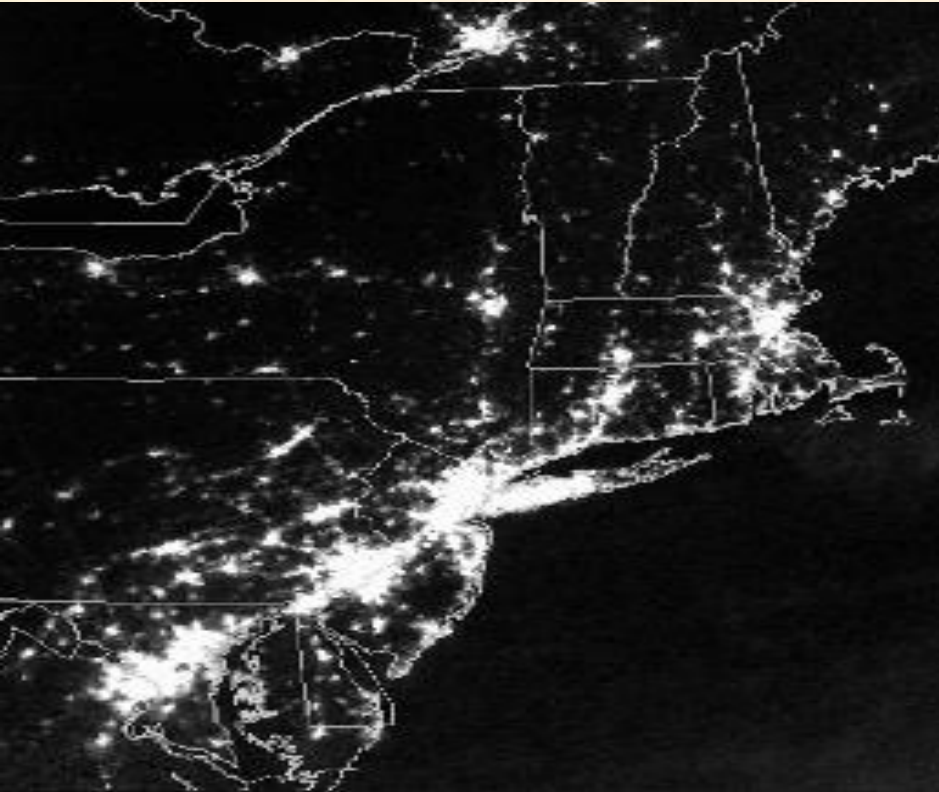
Largest blackout
in the US history

Some Typical Incidents of Power Blackouts

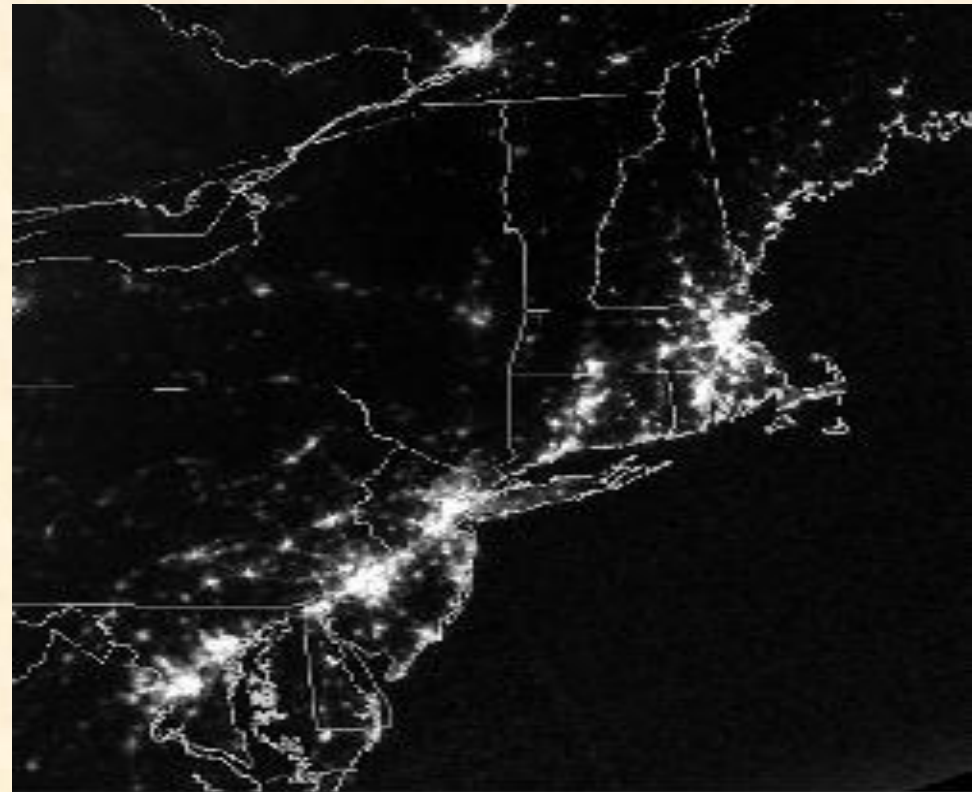
Data Source

People Affected (Millions)	Location	Date
670	India	2012-7-30
3.8	USA (IN, OH, WV, PA ... DC)	2012-6-29
...	...	...
100	Indonesia, Java-Bali	2005-08-18
55	Italy, Southern	2003-09-28
50	North America, Northeastern	2003-08-14
60	Brazil, Southern and Southeastern	1999-03-11

Snapshot: 15 August 2003 Power Blackout



Before blackout

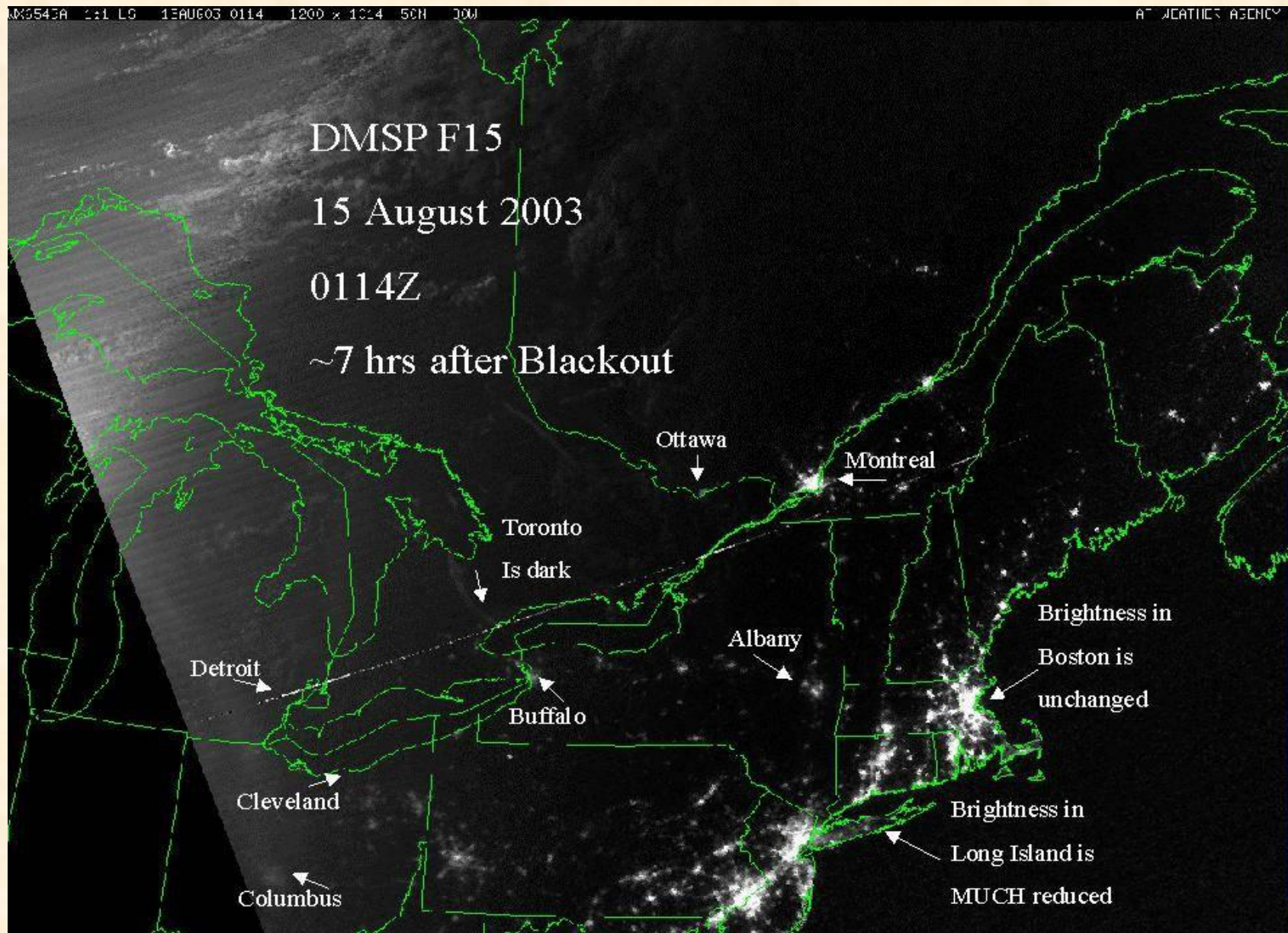


During blackout

Satellite images

Reference

A detailed map



Will It Happen Again ?



... 1959, 1961, 1965, 1977, 2003, ... 2012, 2013?

Major Power Blackout Records

Unfortunately, the answer is YES

印度再断电半数地区瘫痪 国民指因腐败政客贪婪所致

2012年08月01日 07:44

来源：财经网 作者：张阳

1人参与 0条评论

打印

转发

字号TT



7月31日，印度乘客在焦急地等火车。因为断电，印度的很多铁路陷入瘫痪。

7月31日，印度的断电危机更加严重。印度电力部门一名高级官员表示，周二又有多多个地区的电网出现故障，导致全国一半地区没有电力供应，其中包括首都新德里和加尔各答等众多大城市，受影响人口达6亿。

据印度《经济时报》31日报道，当天下午，印度全国有一半的地区没有电力供应，这被称为印度史上最大的断电危机。目前，所有连接北部、东部和东北部地区的发电站均已关闭。在首都及多个邦，地铁服务均暂停，铁路系统一名发言人说，全国有多达500列火车因为断电陷入困境。

August 01, 2012 RECORDER REPORT 0 Comments

BUSINESS RECORDER

Hundreds of millions of people across India were left without power on Tuesday in one of the world's worst blackouts, trapping miners, stranding train travellers and plunging hospitals into darkness when grids collapsed for the second time in two days.

Stretching from Assam, near China, to the Himalayas and the north-western deserts of Rajasthan, the outage covered states where half of India's 1.2 billion people live and embarrassed the government, which has failed to build up enough power capacity to meet soaring demand. "Even before we could figure out the reason for yesterday's failure, we had more grid failures today," said R. N. Nayak, chairman of the state-run Power Grid Corporation.

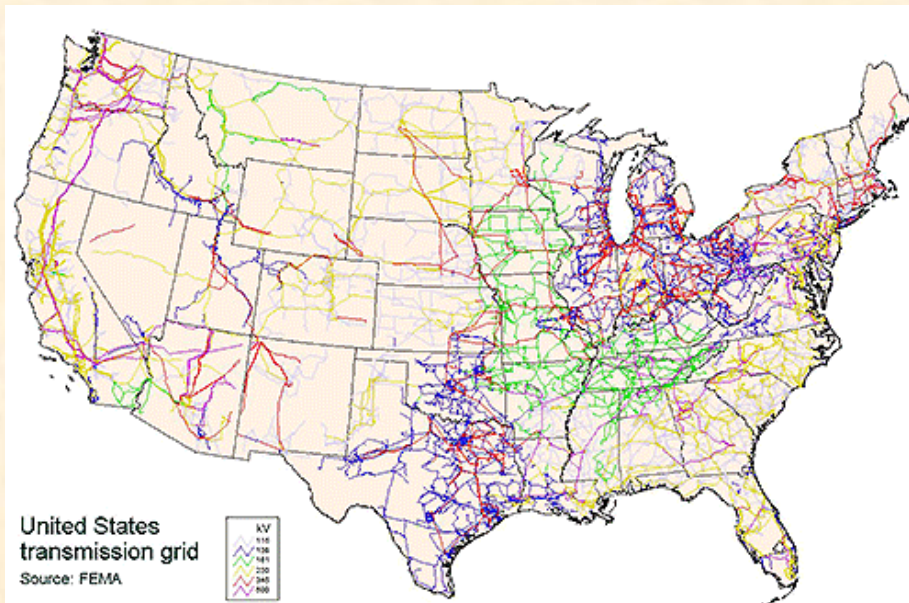
Prime Minister Manmohan Singh had vowed to fast-track stalled power and infrastructure projects as well as introduce free market reforms aimed at reviving India's flagging economy. But he has drawn fire for dragging his feet. By nightfall, power was back up in the humid capital, New Delhi and much of the north, but a senior official said only a third was restored in the rural state of Uttar Pradesh, itself home to more people than Brazil.



Power blackout caused traffic jam

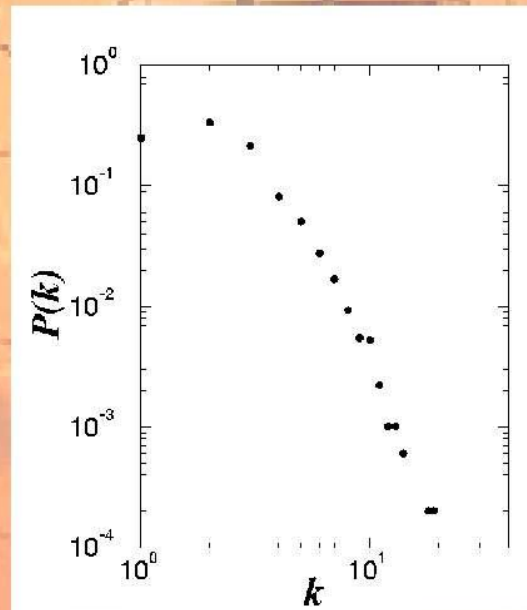
Why?

Cascading Failures on Scale-Free networks



ELECTRICAL POWERGRID

nodes: generators, transformers
edges: transmission lines



The electrical power grid of the western United States

$$\gamma \cong 4$$

Cascading Failures: Models and Analysis

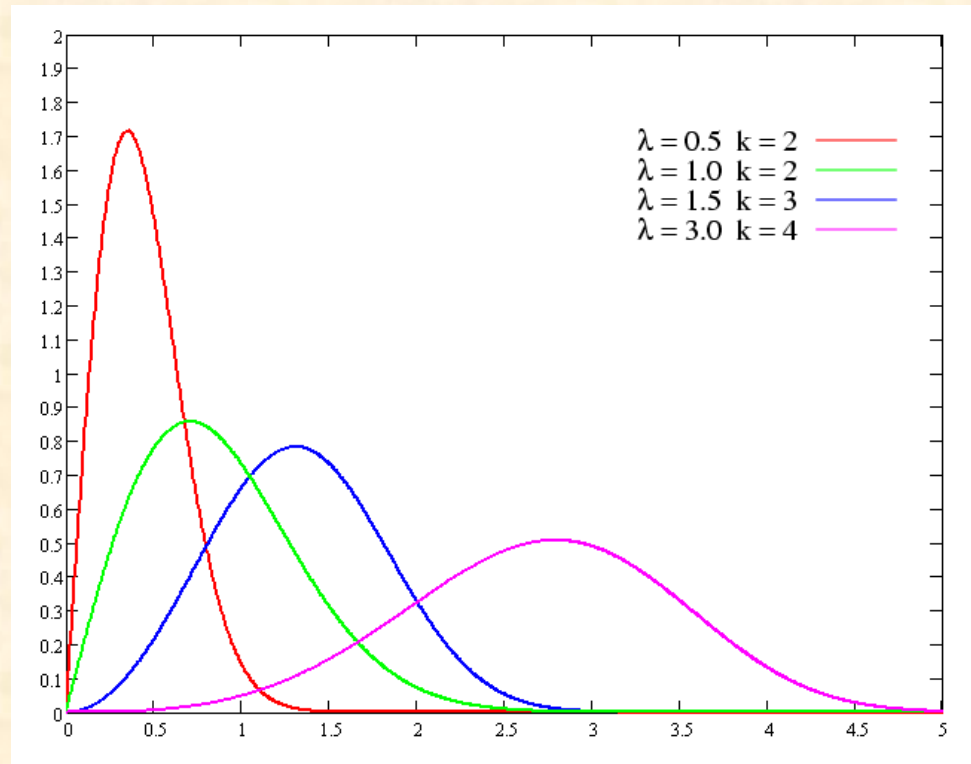
Models Based on Node Dynamics

- ❖ The Fiber-Bundle Model (FBM) is a conceptual graph framework, typically a BA model, for cascading failure analysis
- ❖ In the FBM, a set of large “fibers” (clusters of nodes) are located on the sites of a lattice, and each site is randomly assigned a security threshold sampled from a given probability $P(\sigma_i) = 1 - e^{-(\sigma_i)^\rho}$ distribution — typically the Weibull (cumulative) distribution
- ❖ Then, this set of nodes are loaded against their loading security thresholds $\sigma_i > 0$
- ❖ For a network load σ at a node, if $\sigma_i < \sigma$ then this node is **failed**
The individual load carried by each failed node is then equally redistributed through connections (edges) to their neighbors
- ❖ This redistribution may induce secondary failures, which in turn may trigger tertiary failures, and so on, leading to collapse

Recall: Weibull distribution

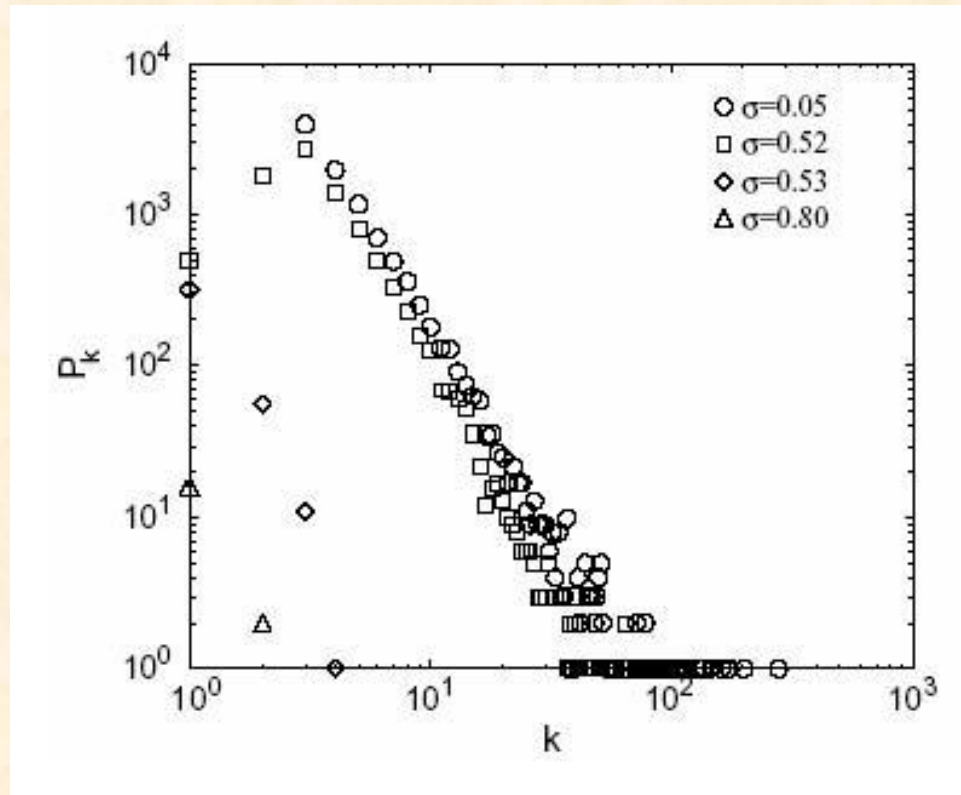
$$f_k(x) = \frac{k}{\lambda} \left(\frac{x}{\lambda} \right)^{k-1} e^{-(x/\lambda)^k}$$

k and λ are constant parameters



Cumulative distribution function: $1 - e^{-(x/\lambda)^k}$

Fiber-Bundle Model Built on a BA Network



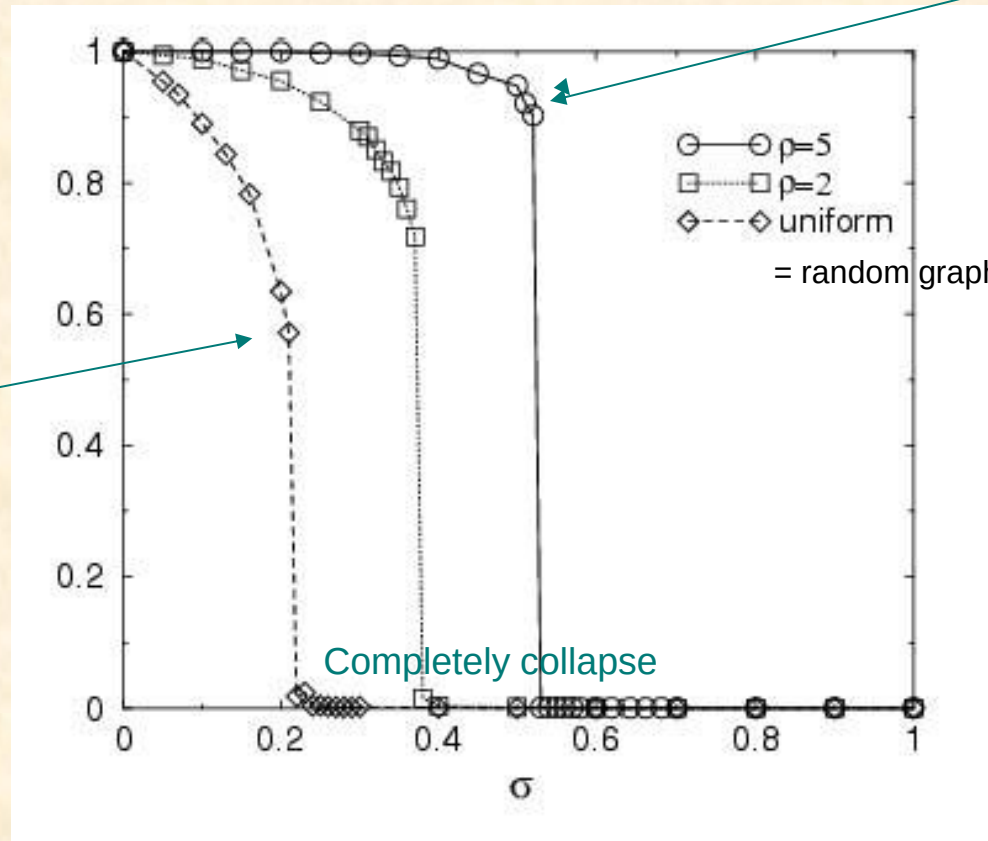
Power-law distributions of degree- k nodes for small σ

Moreno et al. (2002)

Fiber-Bundle Model

Fraction of connected sub-nets

less severe
(network not collapses unless 40% nodes are disconnected)



more severe

(whole network collapses even when only 10% nodes are disconnected)

= random graph

BA scale-free network: $N = 10,000$ $P(\sigma_i) = 1 - e^{-(\sigma_i)^p}$

Moreno et al. (2002)

Node-Betweenness Model

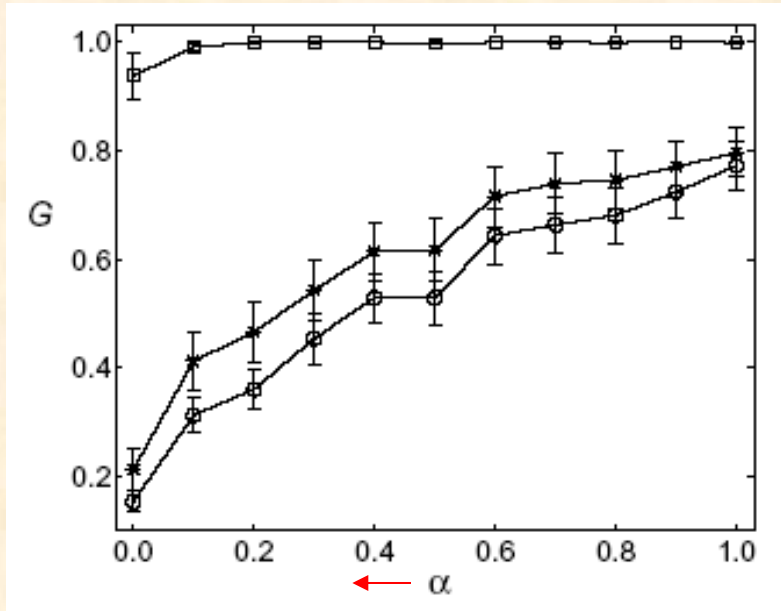
- ❖ Let the maximum-load capacity of node i be proportional to its initial loading L_i :

$$C_i = (1 + \alpha)L_i \quad i = 1, 2, \dots, N$$

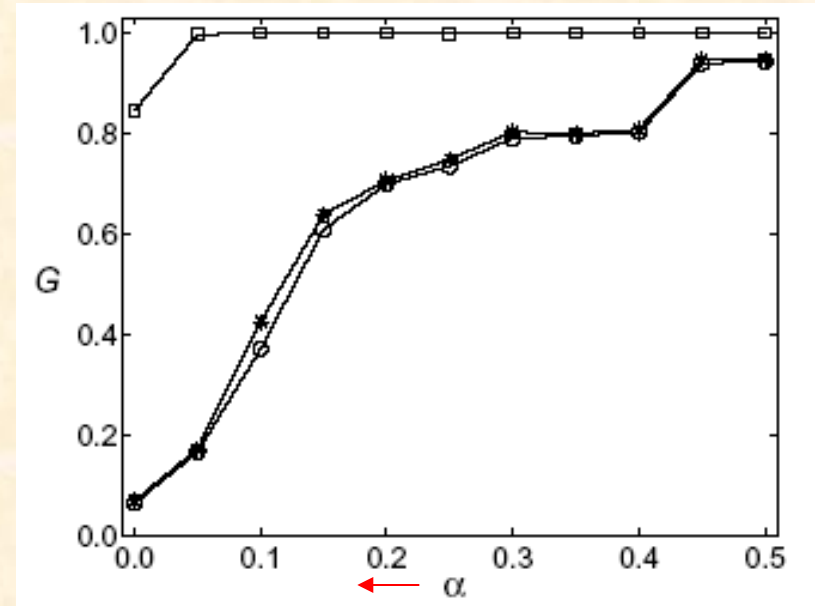
where constant $\alpha \geq 0$ is the tolerance parameter

- ❖ Normally, there are free flows traversing on the network
- ❖ When a node fails to work, traffic cannot go through it, so the node will be removed
- ❖ Then, the total traffic flows are uniformly re-distributed over the rest of the network
- ❖ This generally changes the loads of the remaining nodes and the distribution of the shortest paths
- ❖ The task is to study when and how the network collapses

Node-Betweenness Model



Scale-free network model



AS-level real Internet

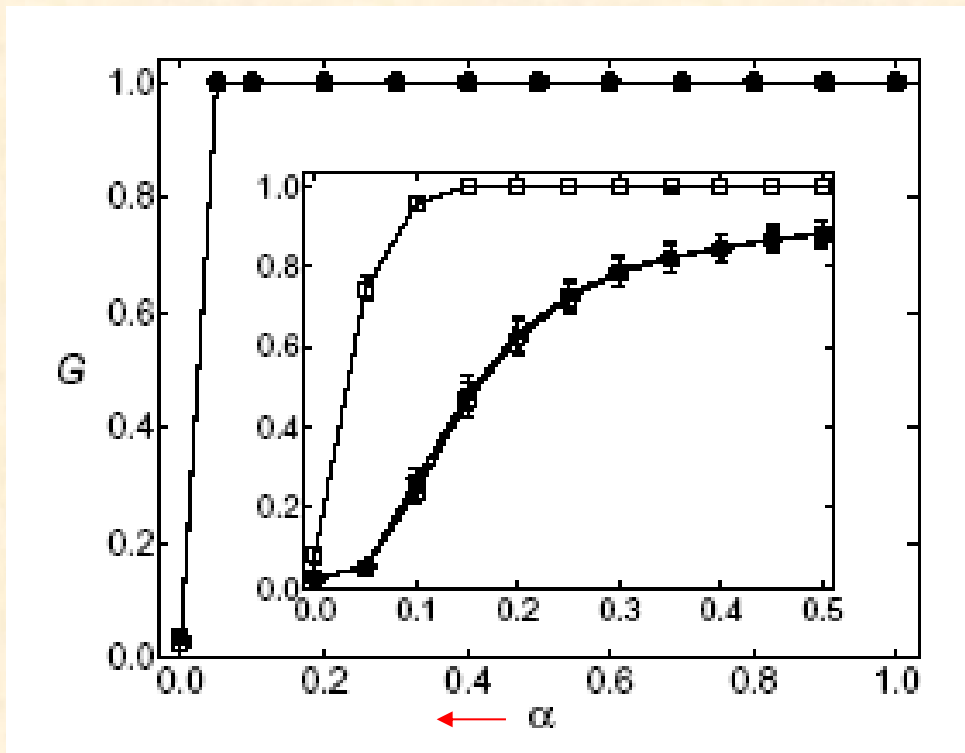
G is the fraction of remaining connected sub-nets over the whole network
[square curves correspond to random node failures, star curves to largest-degree node failures, and circle curves to biggest-betweenness node failures]

Observation 1: smaller tolerance $\rightarrow$ more severe failure

Observation 2: largest-degree/betweenness failures is more severe than random failures

Motter et al. (2002)

Node-Betweenness Model



Large picture:
Random network

Small picture
Scale-free network

Motter et al. (2002)

Comparison of cascading failures on random and scale-free networks

square – random attack, circles – intentional attack. $N = 5,000$ $\langle k \rangle \approx 3$

Observation 1: For random graphs, both random and intentional attacks perform similarly

Observation 2: For scale-free networks, intentional attacks are more severe than random attacks

Cascading Failures: Models and Analysis

Models Based on Edge Dynamics

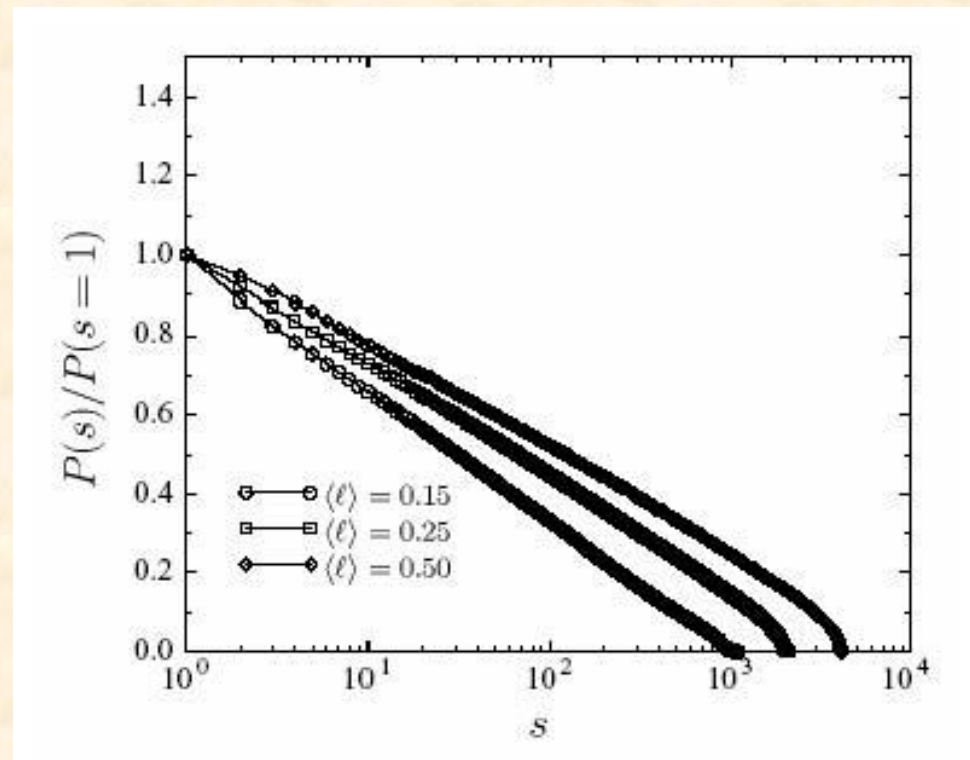
Consider loads on edges

Let s be the number of overloaded edges

Let $P(s)$ be the cumulated failures (normalized)

Let $\langle \ell \rangle$ be the average edge loading over the whole network

Observation: it is basically independent of average edge loading (because the three curves are similar and yet $0.15 < 0.25 < 0.5$)



Moreno et al. (2003)

Cascading Failures: Models and Analysis

A model based on both node and edge dynamics

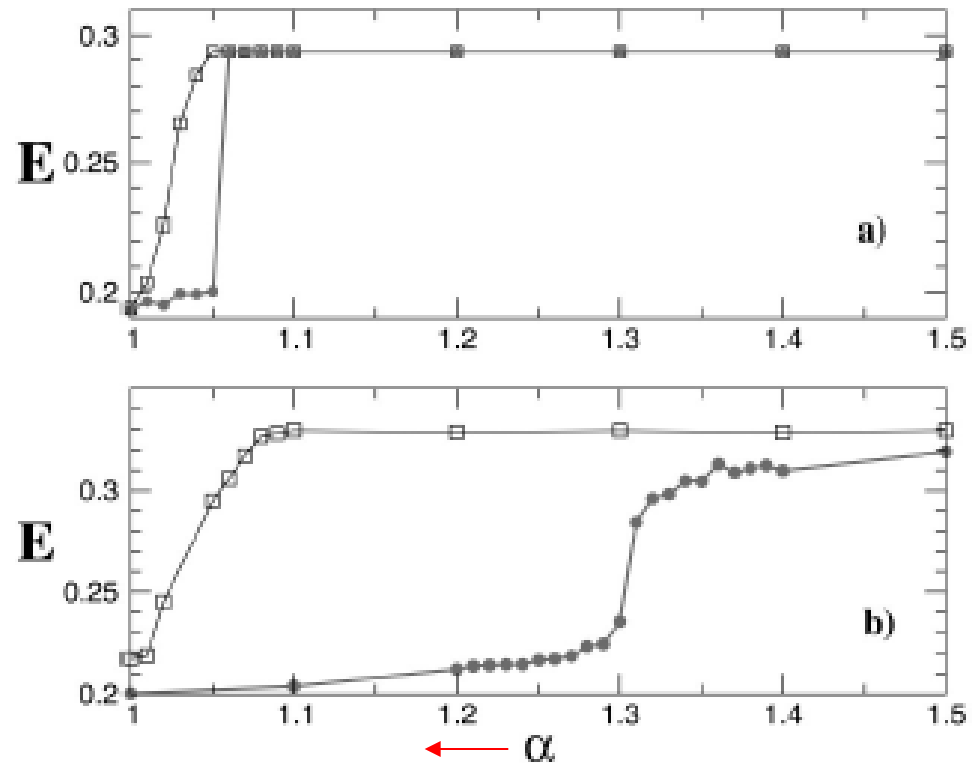
$e_{ij} \in [0,1]$ – efficiency of edge (i,j) , e.g., bandwidth

E – average efficiency:

$$E(G) = \frac{1}{N(N-1)} \sum_{i \neq j} e_{ij}$$

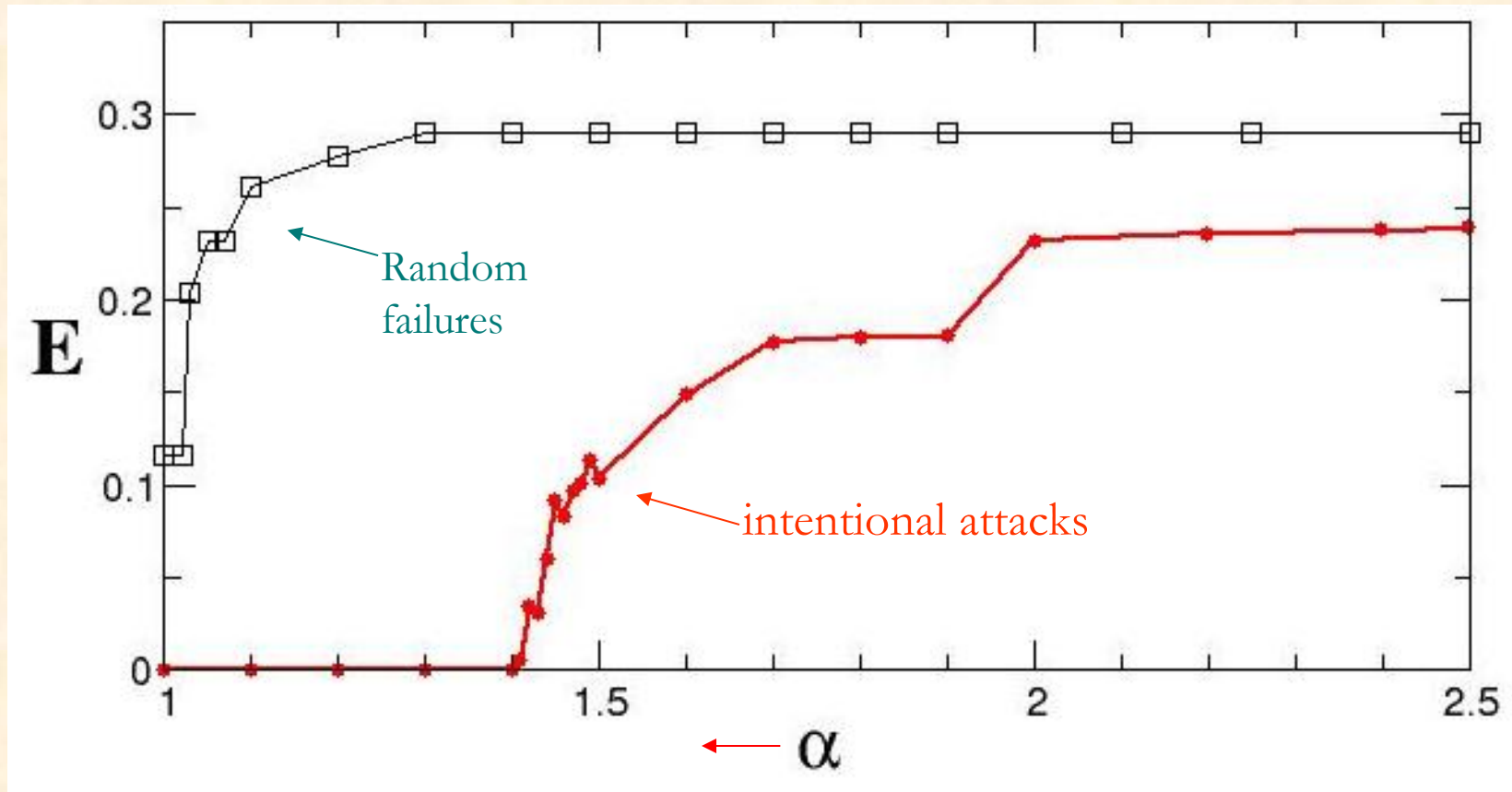
$L_i(t)$ – loading
 $C_i = \alpha \cdot L_i(0)$ – capacity
 $\alpha \geq 1$ – tolerance

Crucitti et al. (2004)



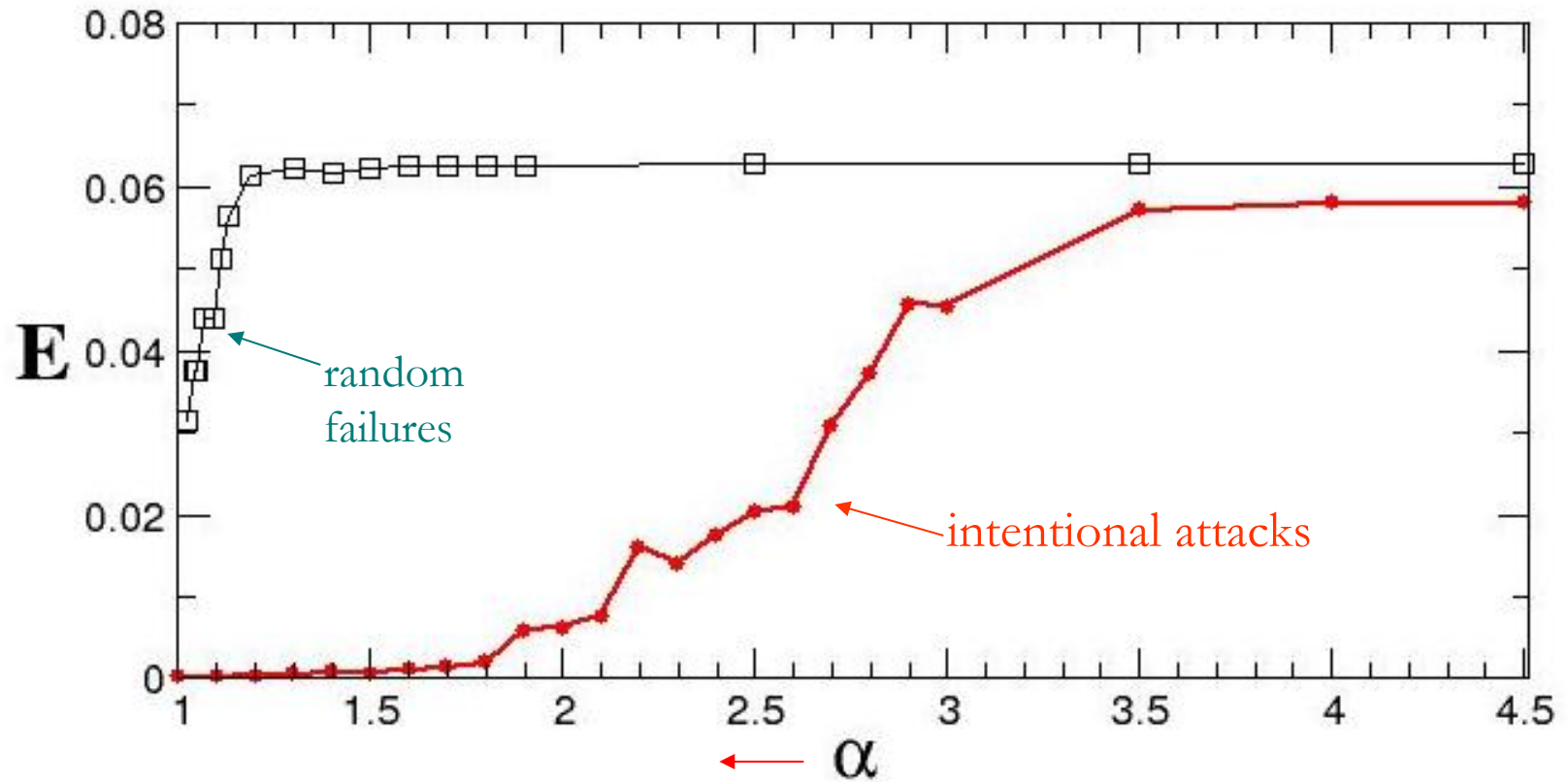
a) ER network, b) BA network. [squares – random node failures, circles – biggest node failures]

AS-level Internet (Real Data)



Crucitti et al. (2004)

US Power Grid (Real Data)



Crucitti et al. (2004)

Models based on Load Re-distributions

Consider a BA network with N nodes

Initialize node i with a load $L_i = ak_i^\alpha$, where a and α are tunable constant parameters.

When node i fails during a dynamical process, then its initial load will be redistributed to every of its neighboring nodes.

The probability that a load from node i is redistributed to node j is given by

$$\Pi_j = \frac{k_j^\alpha}{\sum_{l \in \Omega_i} k_l^\alpha}, \text{ where } \Omega_i \text{ is the neighborhood of node } i.$$

Thus, the additional load received by node j is

$$\Delta L_{ji} = L_i \Pi_j = L_i \frac{k_j^\alpha}{\sum_{l \in \Omega_i} k_l^\alpha}$$

Models based on Load Redistributions

Assume: Each node has a capacity $C_i = TL_i$ of holding loads, with a constant $T \geq 1$.

Thus, if the accumulated loads of a node j is over its capacity, namely, $L_j + \sum_{i \in \Omega_j} \Delta L_{ji} > C_j$, then this node j will fail.

As a result, its total load will be redistributed to all its neighbors, which may cause cascading failures.

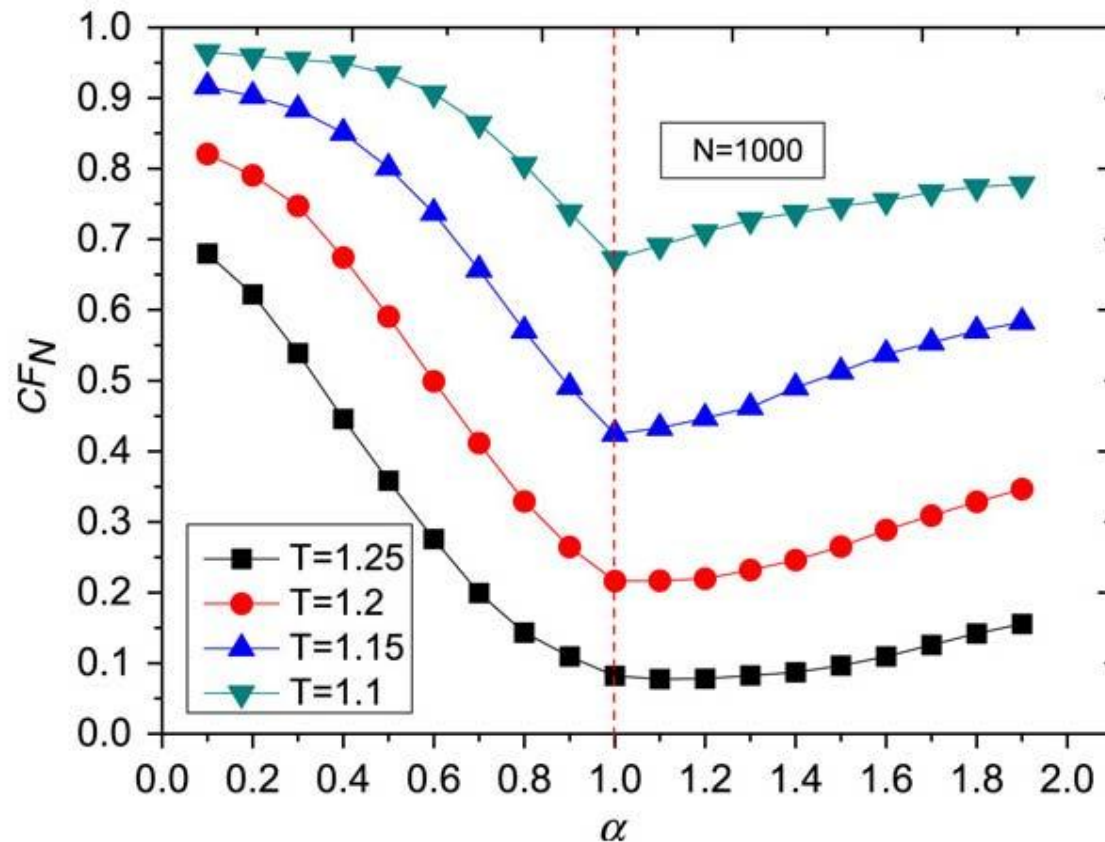
To measure the ability of the network against cascading failures, define an index CF_i for each node i , to be the total number of failed nodes after the failing process is ended.

The average index over the whole network:

$$CF_N = \frac{1}{N(N-1)} \sum_i CF_i$$

-- the smaller the better

Wang *et al.*, Physica A (2008)



Simulation Results: when $\alpha = 1$, in the case of the same T , all CF_N reach the minimum, implying the strongest robustness of the network against cascading failures.

Wang et al., Physica A (2008)

Interdependent Networks

SCIENTIFIC AMERICAN™

Winner of the 2011 National Magazine Award for General Excellence

[News & Features](#) [Blogs](#) [Multimedia](#) [Education](#) [Citizen Science](#) [Topics](#)

[Home](#) » [Scientific American Magazine](#) » [July 2011](#)

[Feature Articles](#) | [Technology](#) 12 [Like](#)



Hacking the Lights Out: The Computer Virus Threat to the Electrical Grid

Computer viruses have taken out hardened industrial control systems. The electrical power grid may be next

By David M. Nicol | June 20, 2011

Interdependent Networks

A large-scale complex network in the real world typically is a network of networks: an integrated network of many different kinds of interdependent networks.

Example: human mobility networks consist of short-range interconnected flows by cars, trains and other means of transportation networks, interconnected with long-range commuting flows like airline flights, on top of which are communication networks (mobile phone calls and Internet message exchanges).



Interdependent Networks of human mobility in the North America

blue – short-range

brown – long-range

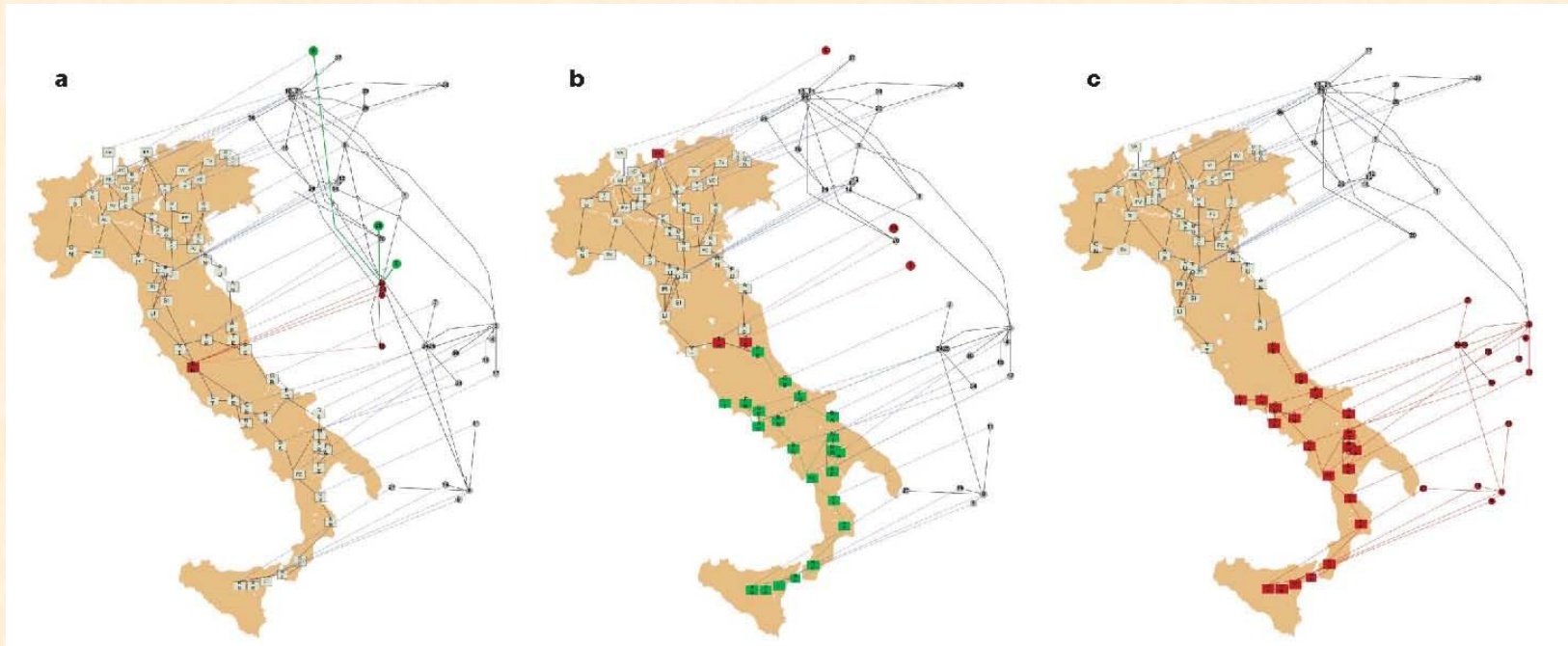
Vespignani, *Nature* (2010) 464: 984-985

Cascading Failures on Interdependent Networks

- ❑ In a tightly interconnected network infrastructure, the failure of nodes in one network not only can lead to consequent failures of some other nodes in the same network but also can lead to failures of nodes in another network, which in turn can cause further failures in the first network, ultimately leading to collapse of the entire networked system.
- ❑ A broader degree distribution increases the vulnerability of interdependent networks to random failures of nodes, which is opposite to the scenario in a single heterogeneous network.
- ❑ Need to consider interdependence among different networks in an integrated network in designing its strong and robust ability against random attack and failures.

Buldyrev et al. *Nature* (2010) 464: 1025-1028

A Real-World Example: Italy Blackout in 2003



a One power station failed (red) → a few associated Internet servers failed (red), and some Internet servers are isolated (green)

b The above failed Internet servers (red) and associated edges are removed → related power stations failed (red), and some power stations became isolated (green)

c The above failed power stations (red) and Internet servers (red) are removed → both networks failed (both red) → system collapsed

Buldyrev et al. *Nature* (2010) 464: 1025-1028

Cascading Failures: Models and Analysis

❖ Other Models:

- Binary influence model
- Sand-pile model
- OPA model -- ORNL-PSerc-Alaska model used to study the blackout dynamics in power transmission grids, where ORNL stands for Oak Ridge National Laboratory; PSerc stands for Power Systems Engineering Research Center; both in USA
- CASCADE model
-

❖ Cascading Failures in Coupled Map Lattices

End

