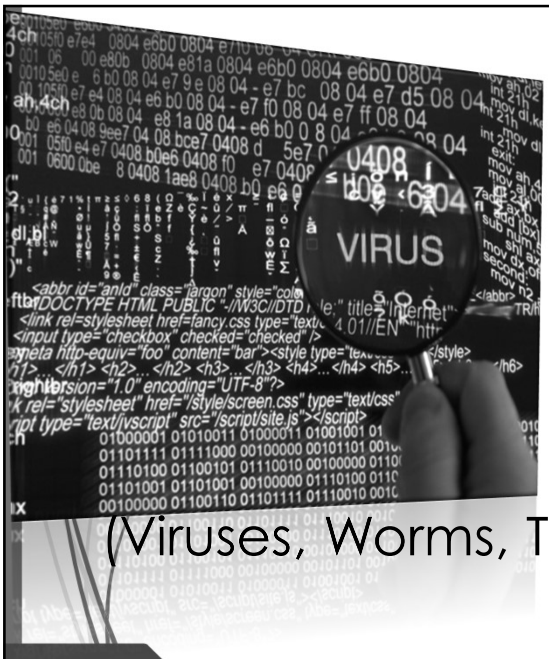


Research assignment 2

Malicious Software
(Viruses, Worms, Trojans, BotNets, Modelling)

Mohammad Sayad

University of Tehran

Introduction

➨ Social Engineering & its relation to infection propagation can be studied here. But they are not technical and we ignore them.

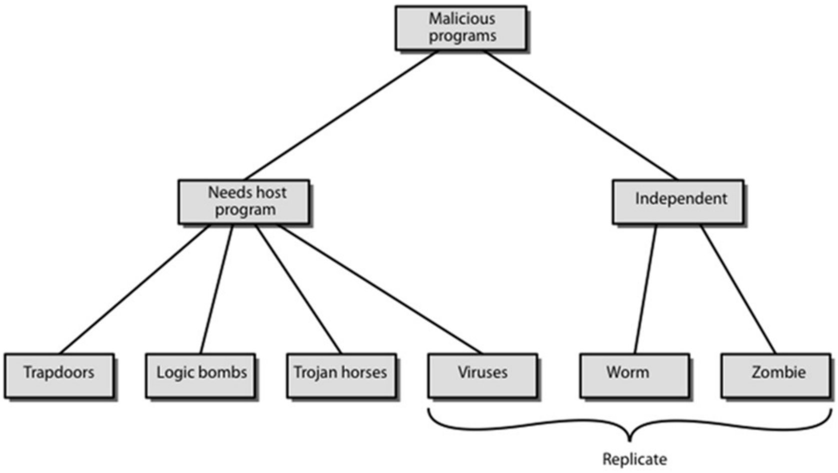
Viruses and other Malwares ویروسها و دیگر کدهای مخرب

- ➢ computer viruses have got a lot of publicity but they are only one of a family of **malicious software**
- ➢ effects usually obvious

➢have figured in news reports, fiction, movies (often exaggerated). They’re getting more attention than they deserve. They’re a concern though

Viruses and How they are detected in human body

Malicious Software Taxonomy - دسته بندی کد های مخرب (بد افزار)



Backdoor or Trapdoor

- It’s a secret entry point into a program
- allows those who know access bypassing usual security procedures
- It’s commonly used by developers
- It’s a threat when left open in production programs. It can be exploited by attackers
- very hard to block in OS
- requires good software development & update

Logic Bomb

- It’s one of oldest types of malicious software
- code embedded in legitimate program and activated when specified conditions met e.g.:
  - presence/absence of some file
  - particular date/time
  - particular user

Trojan Horse

- It's a program with hidden side-effects which usually looks attractive
  - e.g. game, software upgrade etc.
- When it is run, performs some additional tasks
  - allows attacker to indirectly gain access they do not have directly
- often used to propagate a virus/worm or to install a backdoor
- or simply to destroy data

Mobile Code

- A program/script/macro that runs on a collection of platforms, or sometimes, on large homogeneous collection (e.g. Windows)

- It's usually transmitted from a remote system to a local system & then executed on the local system
  - often to inject virus, worm, or Trojan horse
  - or uses exploits itself to gain unauthorized data access (e.g. root access)

Multiple-Threat Malware

- Malwares may operate in multiple ways
- A **multipartite virus** infects in multiple ways
  - e.g. multiple file types
- **Blended** attack uses multiple methods of infection or transmission (usually to maximize the speed of propagation).
  - e.g. Nimda was a worm, a virus, and a mobile code, altogether!
  - they sometimes use messengers and p2p file sharing to spread too.

Viruses ویروسها

- Piece of software that infects other programs
  - modifying them to include a copy of the virus
  - so it executes secretly when host program is run
- They’re developed for a specific operating system/hardware
  - taking advantage of their specific weaknesses
- A typical virus goes through the following phases in its life cycle:
  - Dormant
  - Triggering
  - Execution
  - Propagation

Virus Structure ساختار ویروس

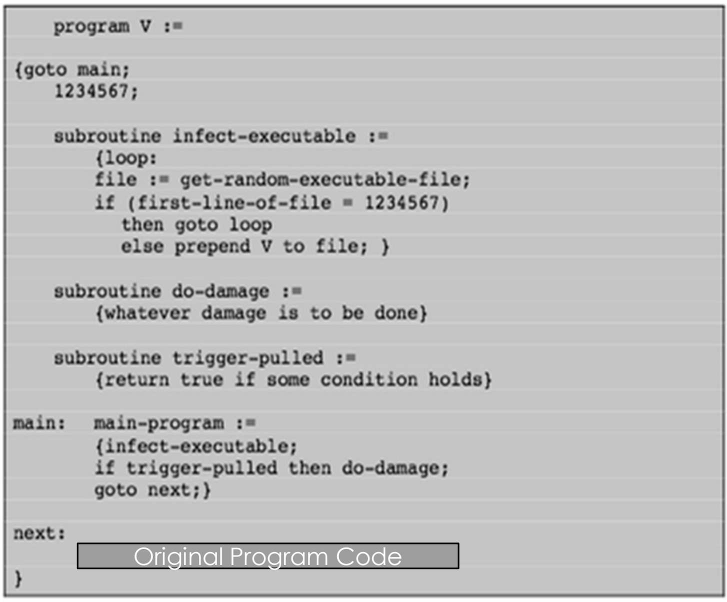
- Virus components:
  - infection mechanism : enables replication
  - trigger mechanism : activates the payload
  - payload : what it does
- Its place in file: prepended / postpended / embedded
- When infected program is run, virus code is executed first, and then the original program is run.

Viruses are packed in many ways inside the file. Upon running, it unpacks itself first with an unpacker.

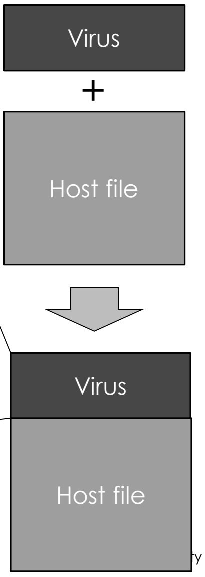
Virus Structure ساختار ویروس

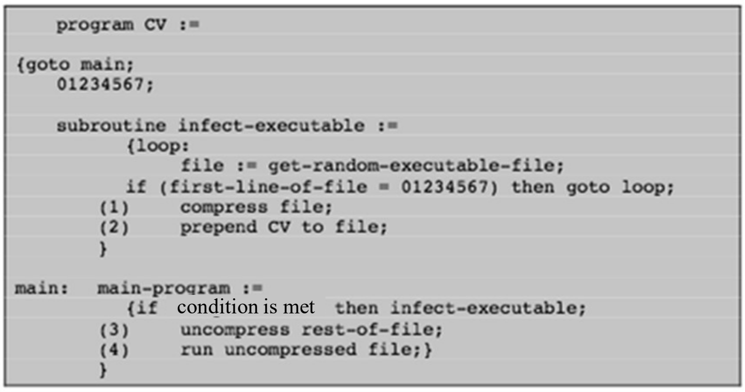
- There are 2 ways to prevent a virus spread:
  - We can block the initial infection → difficult, as there are many vulnerabilities
  - Or we can block its propagation → with strong or additional **access controls** in OS (e.g. Windows ILs or SELinux)

Pseudo-code of Virus that Attaches to the Beginning of a File

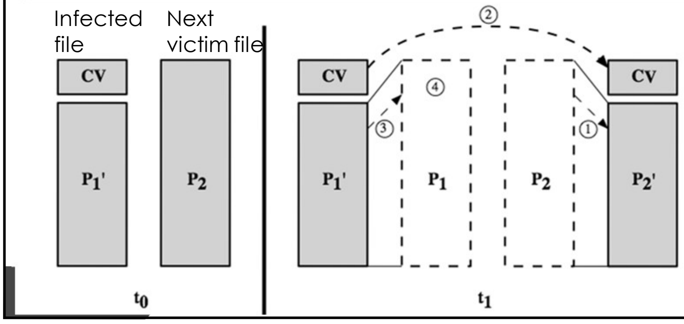


```
    program V :=

{goto main;
    1234567;

    subroutine infect-executable :=
        {loop:
        file := get-random-executable-file;
        if (first-line-of-file = 1234567)
           then goto loop
           else prepend V to file; }

    subroutine do-damage :=
        {whatever damage is to be done}

    subroutine trigger-pulled :=
        {return true if some condition holds}

main:   main-program :=
        {infect-executable;
        if trigger-pulled then do-damage;
        goto next;}

next:
      Original Program Code
}
```





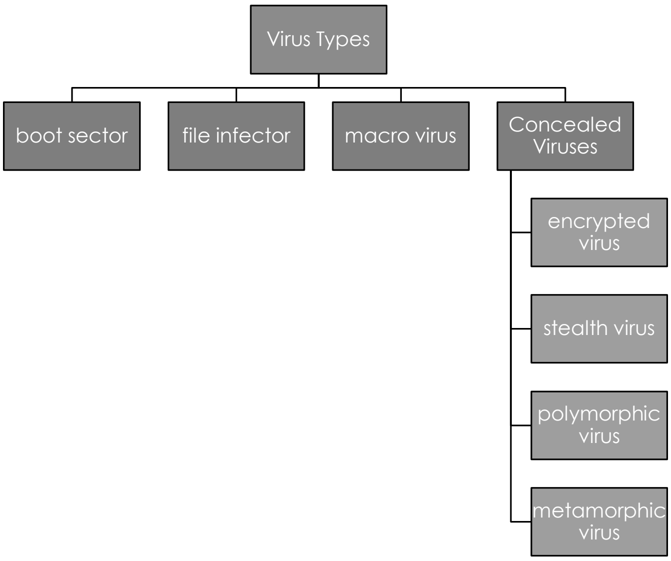
```
    program CV :=

{goto main;
    01234567;

    subroutine infect-executable :=
          {loop:
                file := get-random-executable-file;
          if (first-line-of-file = 01234567) then goto loop;
      (1)       compress file;
      (2)       prepend CV to file;
          }

main:   main-program :=
          {if  condition is met   then infect-executable;
      (3)       uncompress rest-of-file;
      (4)       run uncompressed file;}
          }
```



Compression Virus
ویروسهای فشرده ساز

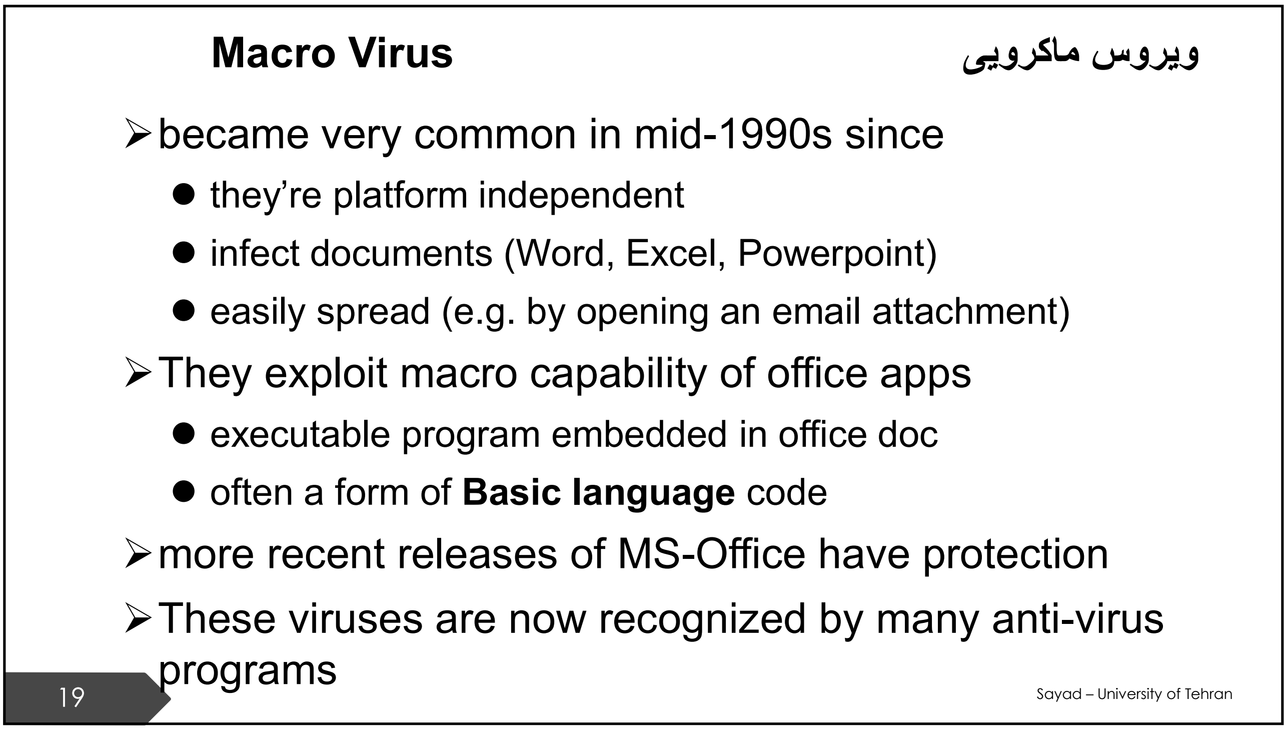
- ➢ Because appending viruses increase the file length, they can be detected. So some compress the file before infecting it.
- ➢ The virus decompresses the program before execution

Virus Classification دسته بندی انواع ویروسها



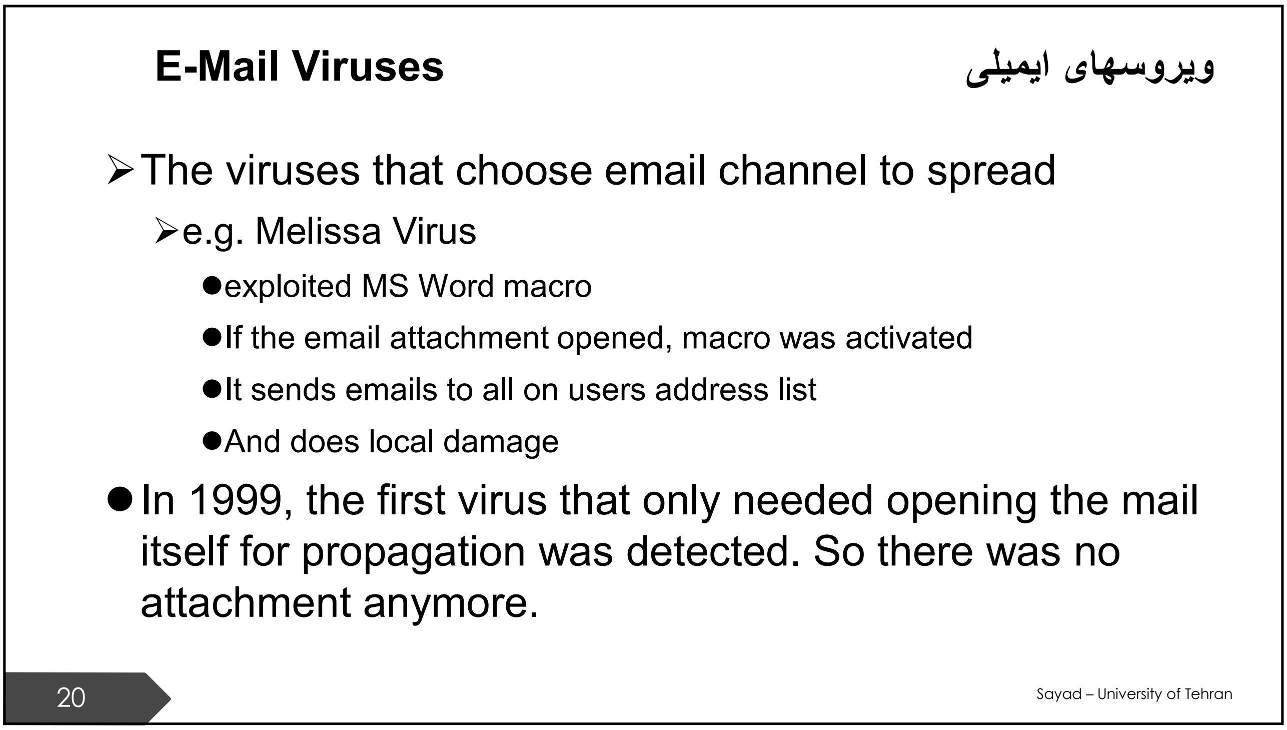
دسته بندی انواع ویروسها

- **Boot sector infector:** Infects a master boot record or boot record and spreads when a system is booted from the disk containing the virus.
- **File infector:** Infects files that operating system or shell consider to be executable.
- **Macro virus:** Infects files with macro code that is interpreted by an application.
- A virus classification **by concealment strategy** includes:
  - **Encrypted virus:** the virus creates a random encryption key, stored within the virus, and encrypts the remainder of the virus. When an infected program is invoked, the virus uses the stored random key to decrypt the virus. When the virus replicates, a different random key is selected.
  - **Stealth virus:** A form of virus explicitly designed to hide itself from detection by antivirus software. Thus,the entire virus, not just a payload is hidden.
  - **Polymorphic virus**: A virus that mutates with every infection, making detection by the “signature”of the virus impossible.
  - **Metamorphic virus:** As with a polymorphic virus ,a metamorphic virus mutates with every infection. The difference is that a metamorphic virus rewrites itself completely at each iteration, increasing the difficulty of detection. Metamorphic viruses may change their behavior as well as their appearance.



Macro Virus ویروس ماکرویی

- became very common in mid-1990s since
  - they’re platform independent
  - infect documents (Word, Excel, Powerpoint)
  - easily spread (e.g. by opening an email attachment)
- They exploit macro capability of office apps
  - executable program embedded in office doc
  - often a form of **Basic language** code
- more recent releases of MS-Office have protection
- These viruses are now recognized by many anti-virus programs



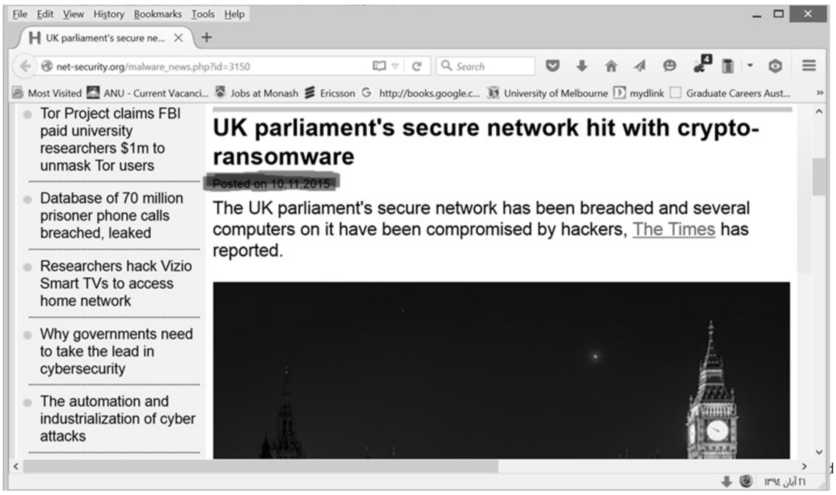
E-Mail Viruses ویروسهای ایمیلی

- The viruses that choose email channel to spread
  - e.g. Melissa Virus
    - exploited MS Word macro
    - If the email attachment opened, macro was activated
    - It sends emails to all on users address list
    - And does local damage
- In 1999, the first virus that only needed opening the mail itself for propagation was detected. So there was no attachment anymore.

Ransomware

بدافزار های باجگیر

Ransomware is a kind of virus that encrypts your files and asks for money in exchange for the key.



- Tor Project claims FBI paid university researchers $1m to unmask Tor users
- Database of 70 million prisoner phone calls breached, leaked
- Researchers hack Vizio Smart TVs to access home network
- Why governments need to take the lead in cybersecurity
- The automation and industrialization of cyber attacks

UK parliament's secure network hit with crypto-ransomware

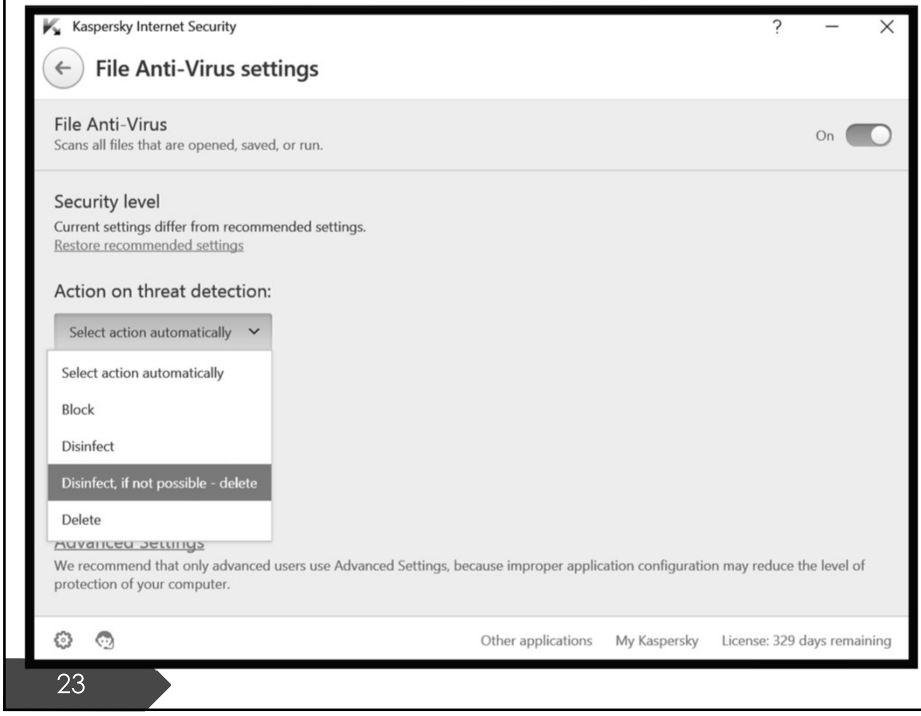
Posted on 10.11.2015

The UK parliament's secure network has been breached and several computers on it have been compromised by hackers, The Times has reported.

Virus Countermeasures

مقابله با ویروسها

- Prevention → ideal solution but difficult
- Realistically need:
  - detection
  - identification
  - removal
- If detected but can't identify or remove → Discard and replace infected program with the backup.



Kaspersky Internet Security

File Anti-Virus settings

File Anti-Virus
Scans all files that are opened, saved, or run.

On

Security level
Current settings differ from recommended settings.
Restore recommended settings

Action on threat detection:

Select action automatically

- Select action automatically
- Block
- Disinfect
- Disinfect, if not possible - delete
- Delete

Advanced Settings
We recommend that only advanced users use Advanced Settings, because improper application configuration may reduce the level of protection of your computer.

Anti-Virus Evolution

تکامل آنتی ویروسها

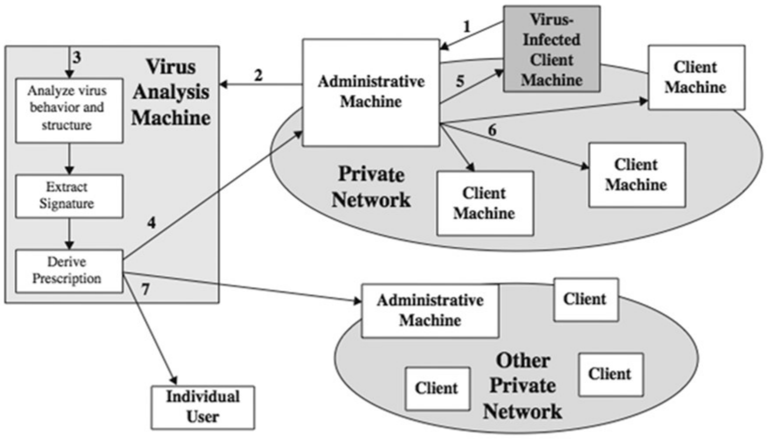
➢Antivirus Generations:

- Generation 1: signatures
  - Use a piece of virus code that doesn't change to detect it (similar to RNAi in nature)
- Generation 2: Heuristics
  - Use heuristic methods to detect if a file is infected, like finding the general pattern viruses use inside the file.
- Generation 3: Identity Actions
  - Antiviruses stay in memory and watch programs behavior. They know specific behaviors which belong to viruses.
- Generation 4: Combination Packages
  - Combination of the above plus new technologies, e.g. access control to prevent changes of important executables.

Generic Decryption رمزگشایی عمومی

- GD is a method to detect encrypted viruses or those who change their look.
- GD scanner virtually runs the programs
  - it has an emulator to run the program
  - It has a signature-based detection engine
  - It has an emulation control module that stops the run frequently to check for the virus signatures
- Using GD, polymorphic, stealth and encrypted viruses are detected.
- Sometimes GD is simply called the "**Emulator**",

Digital Immune System



An architecture to detect a virus and react from IBM and Symantech

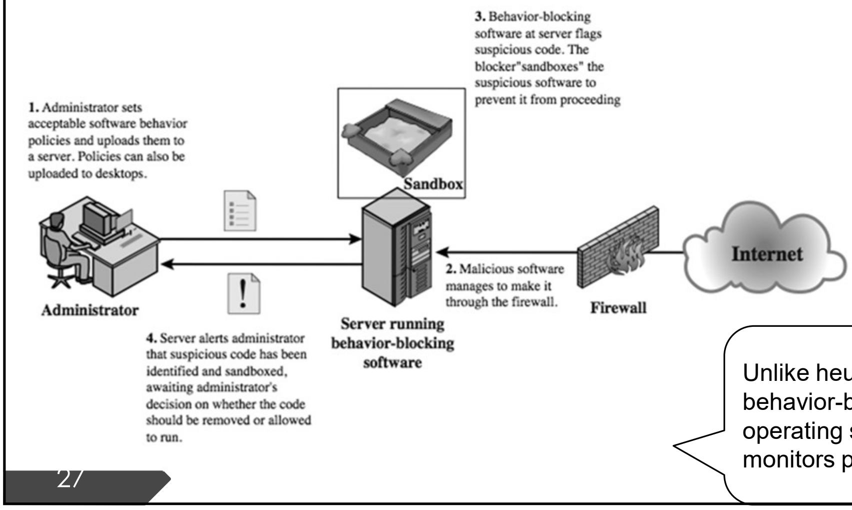
Once a virus is initially detected, it's sent to the "Virtual Analysis Machine" by the administrative machine. After the removal method is found, it's given to the administrative machines to teach clients to how to remove it.

This is now being used in home antiviruses too.

Suspicious files are sent to the company.

Behavior-Blocking Software نرم افزار کنترل مبتنی بر رفتار

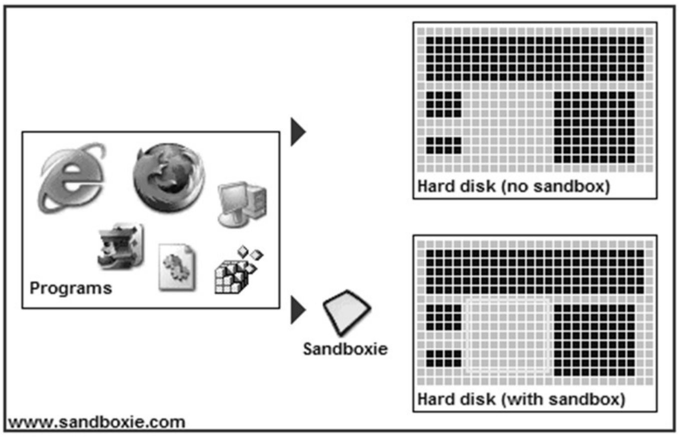
This is to block the execution of programs with suspicious behaviour. Normal behavior is defined by the administrator.


3. Behavior-blocking software at server flags suspicious code. The blocker"sandboxes" the suspicious software to prevent it from proceeding
1. Administrator sets acceptable software behavior policies and uploads them to a server. Policies can also be uploaded to desktops.
Sandbox
Internet
2. Malicious software manages to make it through the firewall.
Firewall
Administrator
Server running behavior-blocking software
4. Server alerts administrator that suspicious code has been identified and sandboxed, awaiting administrator's decision on whether the code should be removed or allowed to run.

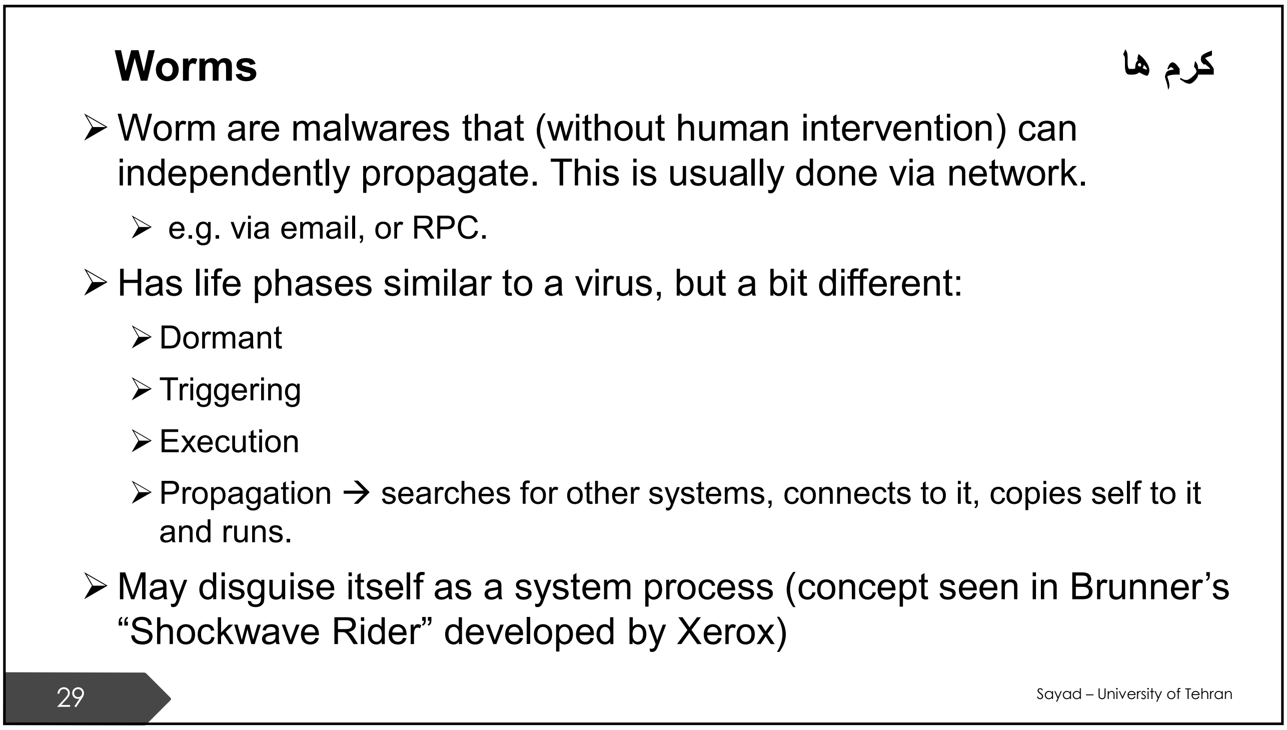
Example: if a program wants to change the system configuration, it's blocked and its behavior is reported to the administrator as well as the antivirus program.

Unlike heuristics or fingerprint-based scanners, behavior-blocking software integrates with the operating system of a host computer and monitors program behavior in real-time.

جعبه شن Sandboxing



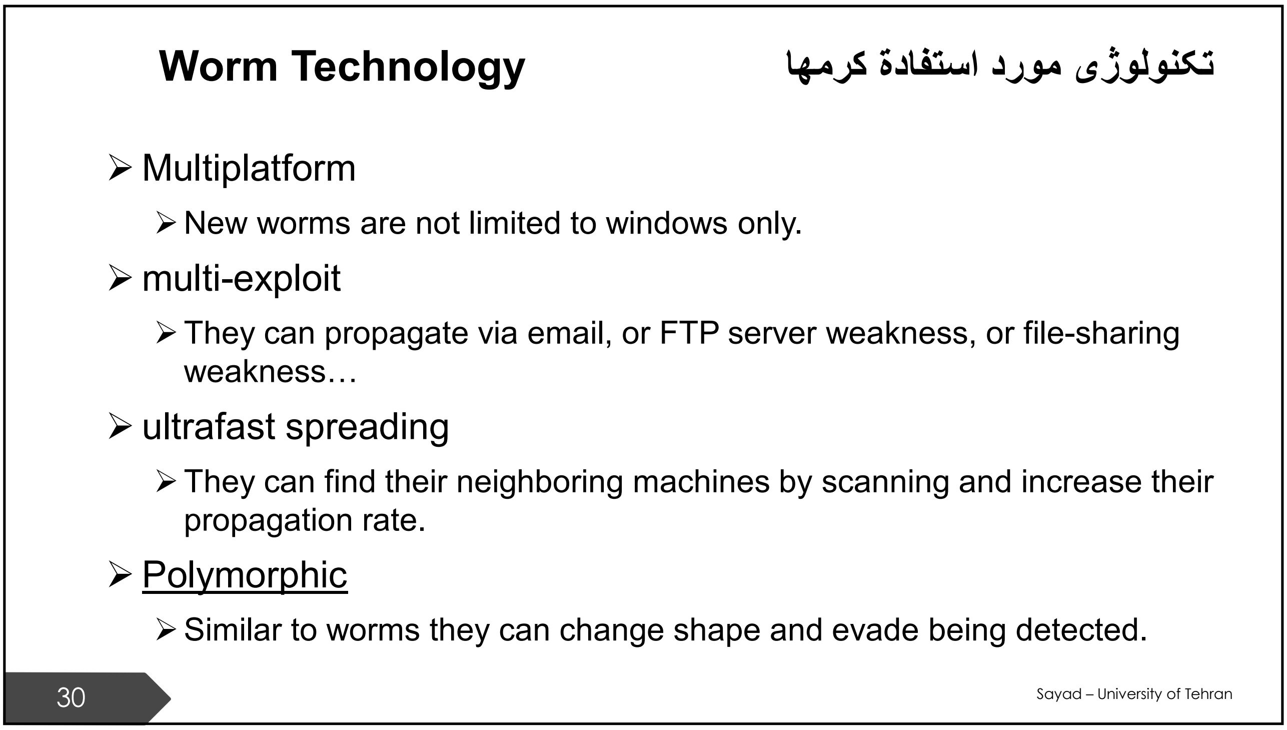
- A way to run the programs in a virtual and controlled environment. You can undo anything.
- Usually virtual machines are used for this purpose.



Worms

کرم ها

- Worm are malwares that (without human intervention) can independently propagate. This is usually done via network.
  - e.g. via email, or RPC.
- Has life phases similar to a virus, but a bit different:
  - Dormant
  - Triggering
  - Execution
  - Propagation → searches for other systems, connects to it, copies self to it and runs.
- May disguise itself as a system process (concept seen in Brunner's "Shockwave Rider" developed by Xerox)



Worm Technology

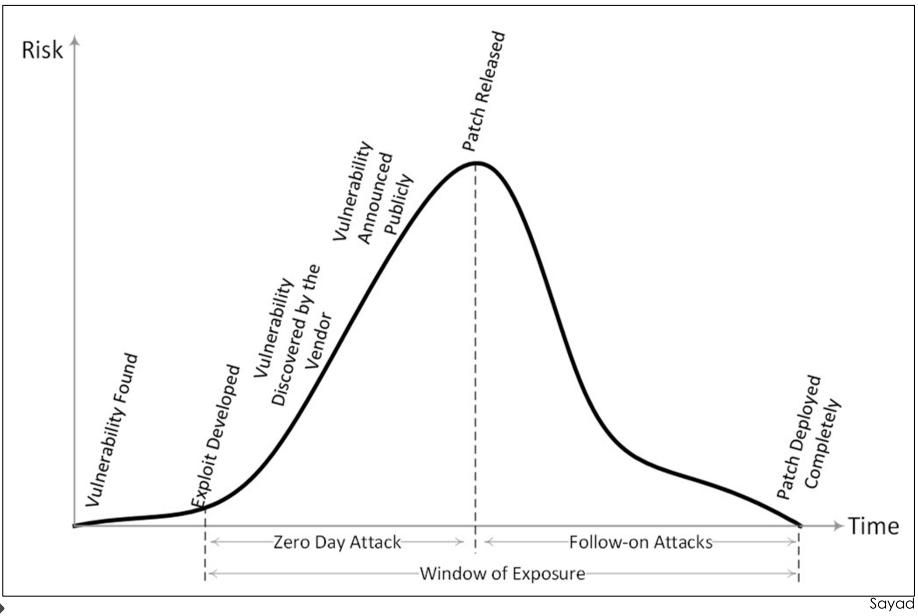
تکنولوژی مورد استفادة کرمها

- Multiplatform
  - New worms are not limited to windows only.
- multi-exploit
  - They can propagate via email, or FTP server weakness, or file-sharing weakness…
- ultrafast spreading
  - They can find their neighboring machines by scanning and increase their propagation rate.
- Polymorphic
  - Similar to worms they can change shape and evade being detected.

Worm Technology تکنولوژی مورد استفادهٔ کرمها

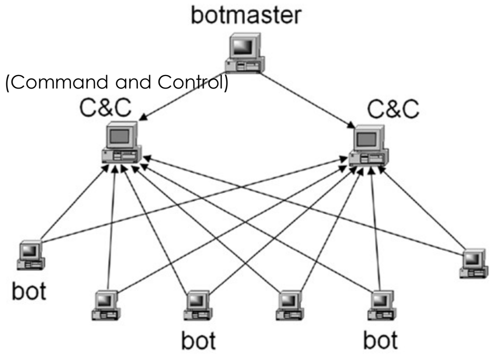
- Metamorphic
  - They can change their behavioral patterns too.
- Transport vehicles
  - They can be a means of delivering something, e.g. a code to quickly launch DDoS attack.
- Zero-day exploit
  - Those which unexpectedly spread, use vulnerabilities that are not announced yet.

Window of Vulnerability, Zero Day Threats & Patching



Botnets شبکه های بات

- A combination of robot and netwok. Botnet is a set of connected programs (via internet) that cooperate to do a single task (e.g. DDoS attack or spamming). They usually take orders from a central station called “Bot Master”



- To make such a network, malwares are used. A malware is developed to infect computers. Later they can use the network for any purpose.
- “Botnet” is also called “zombie army”

Worm Countermeasures روشهای مقابله با کرمها

- Worm detection method are similar to virus ones.
- Once a worm gets into a system, antiviruses can detect it.
- Unlike viruses, worms increase network traffic.

Worm Countermeasures روشهای مقابله با کرمها

Countermeasures:

- signature-based worm scan filtering

پس از شناسایی رفتار و امضای منحصر به فرد کرم، جلوی اسکن وی و آلودگی بیشتر شبکه را می گیرند.

- filter-based worm containment

پیامهای رد و بدل شده را وارسی میکنند تا ببینند کد خود کرم در آنها (در حال انتقال) هست یا نه.

- payload-classification-based worm containment

بجای گشتن به دنبال کد کرمهای مشخص، بسته های رد و بدل شده را برای شکل غیر عادی از داده وارسی میکنند.

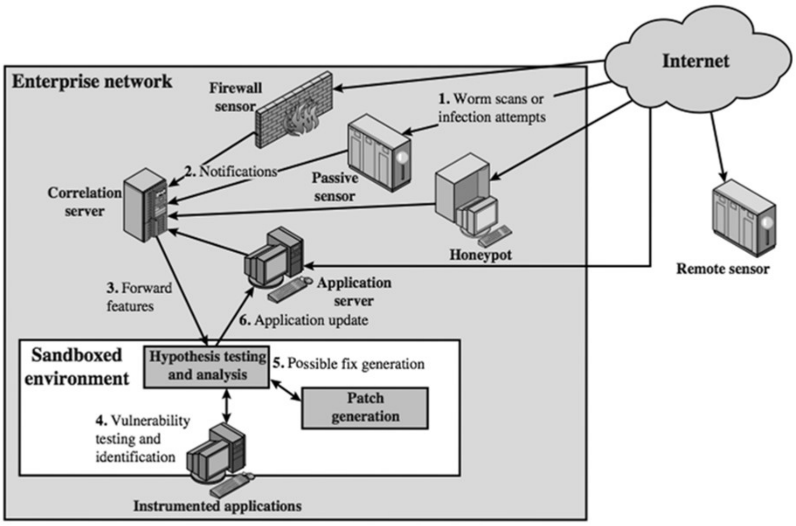
- threshold random walk scan detection

با کنترل اتصالاتی که یک کامپیوتر برقرار میکند و از روی تصادفی بودن و تعدد آن به فعالیت Scanner یک کرم در آن کامپیوتر پی میبرند.

- rate limiting and rate halting

برای ترافیک خروجی و یا تعداد اتصالات برقرار شده یک کامپیوتر محدودیت میگذارند. درصورت مشاهده امضای کرم با روش اول، میتوانند کامپیوتر مورد نظر را مسدود کنند.

Network-based Worm Defense دفاع شبکه ای در برابر کرمها



It's a set of external and internal sensors (incl. Honeypots, Firewalls, …) to detect suspicious information and send them to a correlation server to make comparisons. Correlation Server has a big picture of the network. If the initial suspicion is correct, it's sent to a sand-box like analyzer to find cure/patch.

This is similar to an SOC (Security Operation Center).

Honeypot ظرف عسل

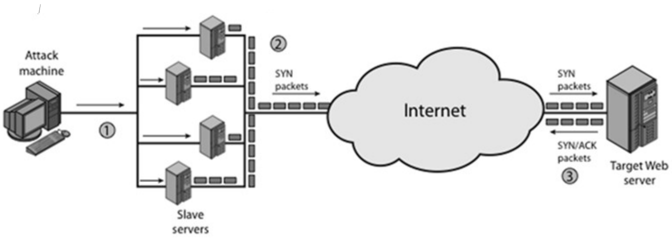
- It’s a way to collect infected samples as well as spams and also detect the origins. These computers play the role of baits for viruses, worms and spams. They are exposed to infection deliberately. But these machines are monitored constantly. Every now and then, information is collected from these machines. They are called honeypots.
- Honeypot samples:
  - Bubblegum Proxypot. An open proxy honeypot for deceiving and detecting spammers.
  - BackOfficer Friendly: BOF is a free Windows based honeypot can listen on only 7 ports
  - Deception Toolkit: DTK was the first OpenSource honeypot, that emulates a variety of listening services. Its primary purpose is to deceive human attackers.

Distributed Denial of Service Attacks (DDoS)

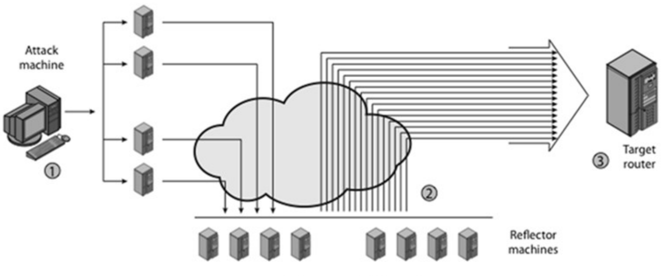
- DDoS is the distributed form of DoS attack.
- Like DoS, the goal is to stop the service by sending too many requests or too much traffic, but from multiple machines.
- Botnet owner can turn bot computers into zombies.
- It’s hard to prevent and since it’s distributed, finding the real source is even harder.

Distributed Denial of Service Attacks (DDoS)

DoS توزیع یافته، Availability یک سیستم را با تهاجم همگانی تعداد زیادی کامپیوتر دیگر هدف قرار میدهد.



(a) Distributed SYN flood attack



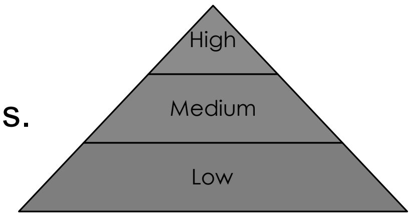
(a) Distributed ICMP attack

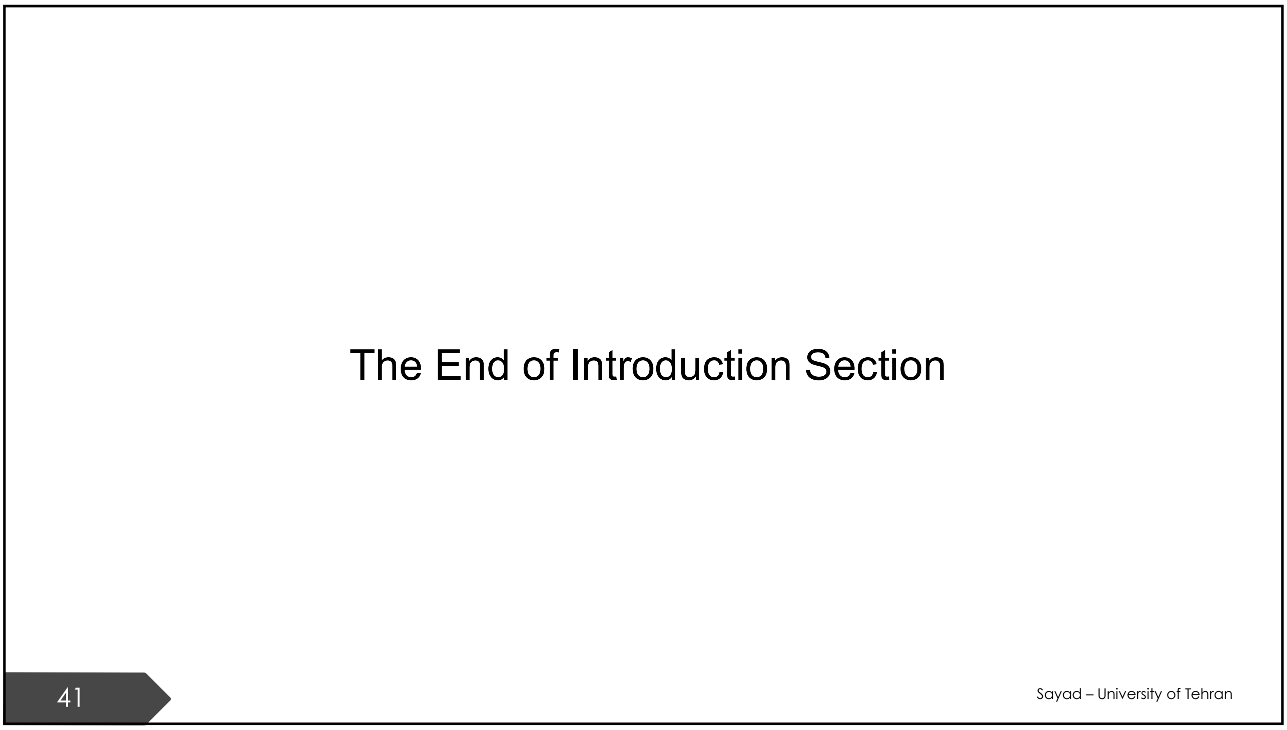
روش اول (با استفاده از مثال SYN Flooding):
۱- دشمن کنترل تعداد زیادی کامپیوتر را بدست میگیرد.
۲- هر یک از آنها تعداد زیادی بسته SYN را با آدرس IP مبداء جعلی به مقصد Web Service قربانی ارسال میکند
۳- قربانی باید برای همه این آدرسها SYN+ACK را ارسال کند و منتظر ACK آنها بماند اما تعداد Session های باز او محدود است و آدرسها نیز چون جعلی بودند هرگز پاسخ نمیدهند. بنابراین سیستم از سرویس دهی خارجی میشود.
به این گونه روشها Direct DDoS Attack گفته میشود.

روش دوم (با مثال Distributed ICMP Attack):
۱- دشمن کنترل تعداد زیادی کامپیوتر را بدست میگیرد.
۲- هر یک از آنها تعداد زیادی بسته ECHO به مقصد کامپیوتر های تصادفی و سالم در اینترنت میفرستد اما آدرس مبدا بسته های ECHO را آدرس IP قربانی قرار میدهد. بدین شکل قربانی با پاسخهای ICMP بمباران شده و از دسترس خارج میشود.
به این گونه روشها Reflection DDoS Attack میگویند.

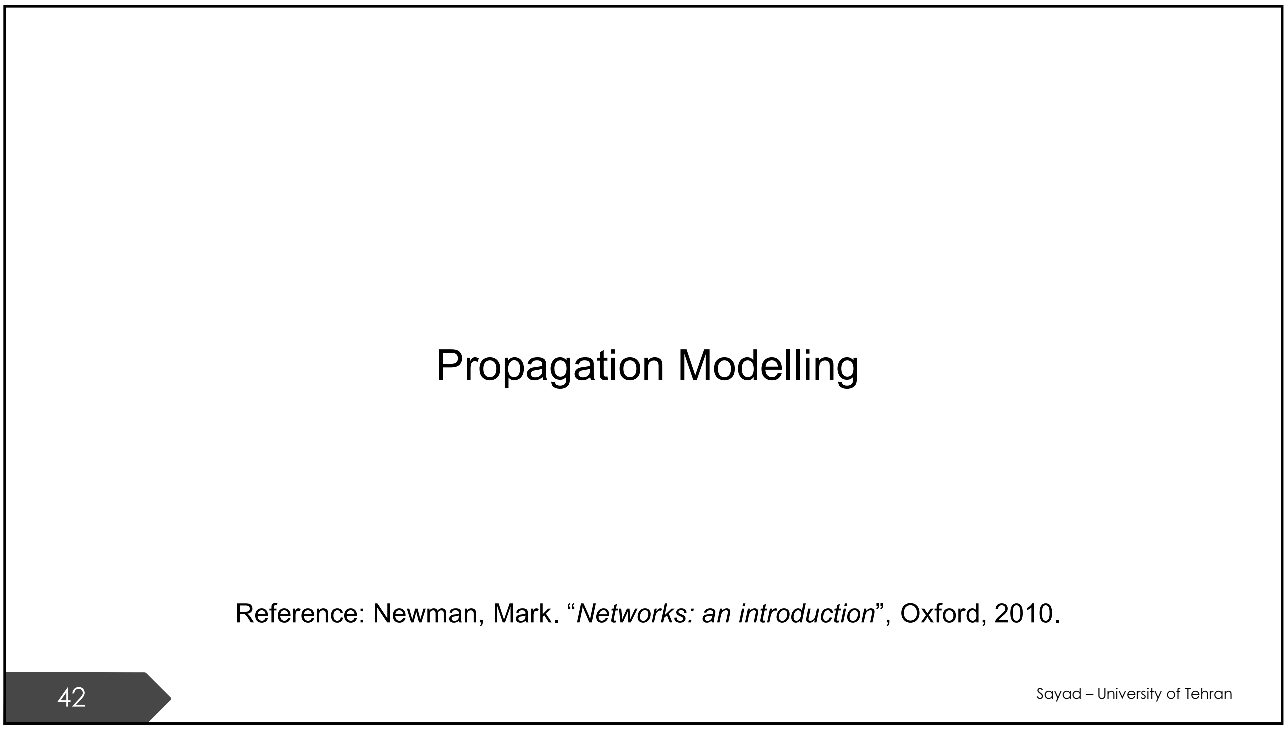
DoS and DDoS Attacks

- Three major defense layers:
  - Prevention:
    - Massive infection prevention and capacity planning in a way that service provisioning is not stopped during an attack (FIPS 199)
  - Detection and Filtering the attack
  - Detection of source, for future preparations.





The End of Introduction Section



Propagation Modelling

Reference: Newman, Mark. “*Networks: an introduction*”, Oxford, 2010.

Preliminaries

- Epidemic models attempt to capture the dynamics in spreading of a disease (or computer virus, idea, product adoption, etc.).
  - Some Key Questions:
    - How do viruses spread in computer networks?
    - Will a disease become epidemic?
    - Who are the best people to vaccinate?
    - Will a given YouTube video go viral?
    - Whom should we focus on during marketing in order to maximize product sales?

Classic Epidemic Models

- Classic Epidemic Models (fully mixed)
  - SI
  - SIR
  - SIS
  - SIRS
- Epidemic Models over Networks
  - Scale-free Network Model for SIS
  - A General Network Model for SIS

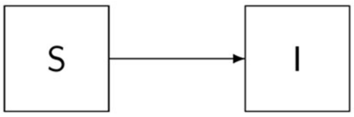
Many Papers on These

Some Definitions

- In a **fully mixed** network, every individual (node) has an equal chance of coming into contact with every other individual (node) in the population.
  - Something like a full-mesh network
- Fully mixed networks are not real. We use this assumption to simplify the equations. But recent papers do not.

SI Model (Susceptible – Infective)

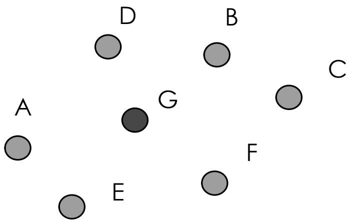
- Assumptions:
  - Let $S(t)$ be the number of individuals who are *susceptible* to sickness at time *t.*
  - Let $X(t)$ be the number of individuals who are *infected* at time *t*.
  - Total population size is *n.*
  - Contact with infected individuals causes a susceptible person to become infected.
  - An infected, never recovers, and stays infected, and infectious to others.
- There are 2 states: Susceptible & Infected



- Parameters:
  - β (infection rate): The probability of getting infection after contact per unit of time (or the number of contacts one makes per unit of time)

(Arias & Cancho, 2015)

The SI Model



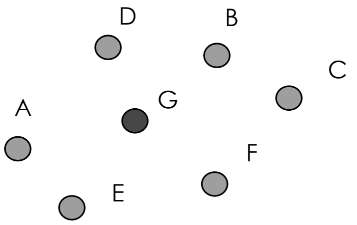
$$\frac{dX(t)}{dt} = \beta \frac{S(t)X(t)}{n} \quad \& \quad \frac{dS}{dt} = -\beta \frac{SX}{n}$$

- ➨An infected node contacts β other nodes per unit of time.
- ➨$S(t)/n$ is the probability that the one we contact is susceptible.
- ➨So from an infected node's point of view, $\beta S/n$ of contacted nodes will be susceptible and will get infected.
- ➨This was for a single infected node → all infected nodes infect $\mathrm{X} \times \beta S/n$ susceptibles per unit of time altogether.

Notice the fully-mixed network assumption

(Newman, 2010)

The SI Model – a different view



$$\frac{dX}{dt} = \beta \frac{SX}{n} \quad \& \quad \frac{dS}{dt} = -\beta \frac{SX}{n}$$

- ➨If we are infected, S/n is the probability that the one we meet is susceptible.
- ➨XS/n is the average number of susceptible nodes the infected nodes meet per unit of time.
- ➨$\beta SX/n$ is the average number of susceptible people that get infected from all infected nodes per unit of time.
  - ➨ Note that we've assumed each infected node only contacts 1 person per unit of time and β is a probability in this view.

SI Solution

- For simplicity, let’s define:

  $s = S/n$ and $\mathrm{x} = X/n$

- At any time, $S(t) + X(t) = n$ → $s + x = 1$
- Thus the equation is turned into:

$$\frac{dx}{dt} = \beta s x \quad \rightarrow \quad \frac{dx}{dt} = \beta(1-x)x$$

The solution to this differential equation (known as the “*logistic growth equation*”) is a *logistic growth curve:*

$$x(t) = \frac{x_0 e^{\beta t}}{1 + x_0(e^{\beta t} - 1)} \qquad \text{Subject to} \quad x(0) = x_0$$

Solving the logistic growth equation

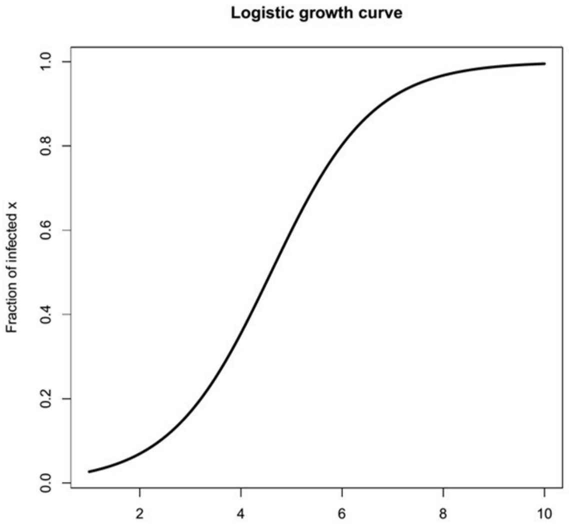
$$\begin{aligned}
\frac{dx}{dt} &= \beta(1-x)x \\
\Longleftrightarrow \int_{x_0}^{x} \frac{1}{(1-x)x}\,dx &= \int_0^t \beta\, dt \\
\Longleftrightarrow \int_{x_0}^{x} \frac{1}{(1-x)}\,dx + \int_{x_0}^{x} \frac{1}{x}\,dx &= \beta t - \beta 0 \\
\Longleftrightarrow \int_{x_0}^{x} \frac{1}{(1-x)}\,dx + \int_{x_0}^{x} \frac{1}{x}\,dx &= \beta t \\
\Longleftrightarrow \ln\frac{1-x_0}{1-x} + \ln\frac{x}{x_0} &= \beta t \\
\Longleftrightarrow \ln\frac{(1-x_0)x}{(1-x)x_0} &= \beta t
\end{aligned}$$

Solving the logistic growth equation

$$
\begin{aligned}
\ln \frac{(1-x_0)x}{(1-x)x_0} &= \beta t \\
\iff \frac{(1-x_0)x}{(1-x)x_0} &= e^{\beta t} \\
\iff \frac{x}{1-x} &= \frac{x_0 e^{\beta t}}{1-x_0} \\
\iff \frac{1-x}{x} &= \frac{1-x_0}{x_0 e^{\beta t}} \\
\iff \frac{1}{x} = \frac{1-x_0}{x_0 e^{\beta t}} + 1 &= \frac{1-x_0+x_0 e^{\beta t}}{x_0 e^{\beta t}} \\
\iff x &= \frac{x_0 e^{\beta t}}{1-x_0+x_0 e^{\beta t}}
\end{aligned}
$$

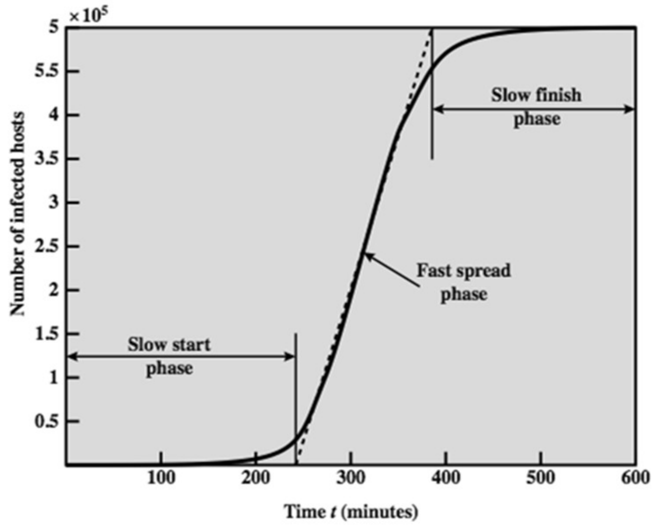
Logistic Equation Growth (SI Model)

$$x(t) = \frac{x_0 e^{\beta t}}{1-x_0+x_0 e^{\beta t}}$$



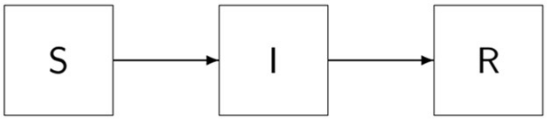
(Newman, 2010)

Example of Solution to SI (Infected nodes growth)



SIR (Susceptible – Infective – Removed) Model

- In this model, nodes are susceptible, then get infected, but some die (or recover and get immunized).
- Again we assume a fully-mixed network.
- There are three states: Susceptible, Infected & Removed/Recovered
- Parameters are:
  - β (infection rate): The probability of getting infection after contact, per unit of time (or you can still think of it as the No. of contacts made per unit of time).
  - γ (recovery rate): The probability of recovery from infection, per unit of time.



S

I

R

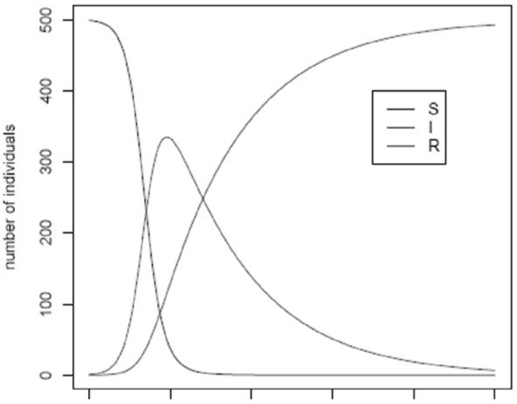
The SIR Model – Differential Equations

Can you tell why?

$$\frac{ds}{dt} = -\beta s x \ , \qquad \frac{dx}{dt} = \beta s x - \gamma x, \qquad \frac{dr}{dt} = \gamma x$$

$$s + x + r = 1$$

- The solution to these equations is not analytically tractable (does not have a closed-form formula). But numerical solutions exist and look like this:



(see Newman's book for further details on the equations)

SIR Solution

- To solve SIR equations, we eliminate x from the first and last equations:

$$\frac{ds}{dt} = -\beta s x \ , \quad \frac{dr}{dt} = \gamma x \quad \Rightarrow \quad \frac{1}{s}\frac{ds}{dt} = -\frac{\beta}{\gamma}\frac{dr}{dt}$$

Integrate from both sides with respect to t:

$$s = s_0 e^{(-\frac{\beta r}{\gamma})}$$

Where s_0 is the value of s at t=0. We assumed there is no one at the recovered state at t=0 (i.e. $r_0 = 0$).

SIR Solution

➨Now we combine the previous result and $x = 1 - r - s$ and put it in $\frac{dr}{dt} = \gamma x$:

$$\frac{dr}{dt} = \gamma(1 - r - s_0 e^{-\frac{\beta r}{\gamma}})$$

If we can solve this equation for r, then we can find s from the previous slide equation, and then x from $x = 1 - s - r$. The solution for the above equation is :

$$t = \frac{1}{\gamma}\int_0^r \frac{du}{1 - u - s_0 e^{-\frac{\beta u}{\gamma}}}$$

But there’s no closed-form for this integration. That’s why we will plot the results numerically.

SIR Solution

➨However, this equation can be solved when $t \rightarrow \infty$. Then $\frac{dr}{dt} = 0$.

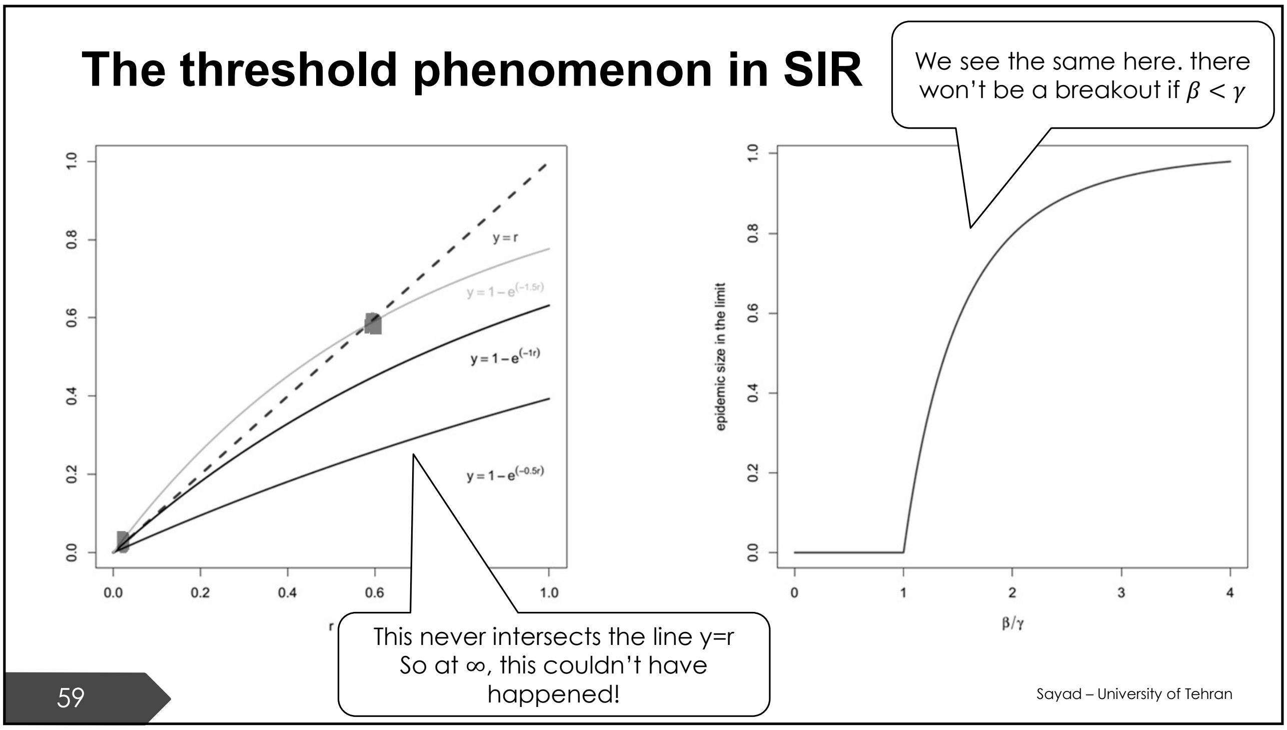
$$\frac{dr}{dt} = \gamma(1 - r - s_0 e^{-\frac{\beta r}{\gamma}})$$

Thus:

$$r_\infty = 1 - s_0 e^{-\frac{\beta}{\gamma} r_\infty}$$

and if we assume everybody is susceptible at t=0, and there’s only one infected node, then $s_0 \cong 1$:

$$r_\infty = 1 - e^{-\frac{\beta}{\gamma} r_\infty}$$



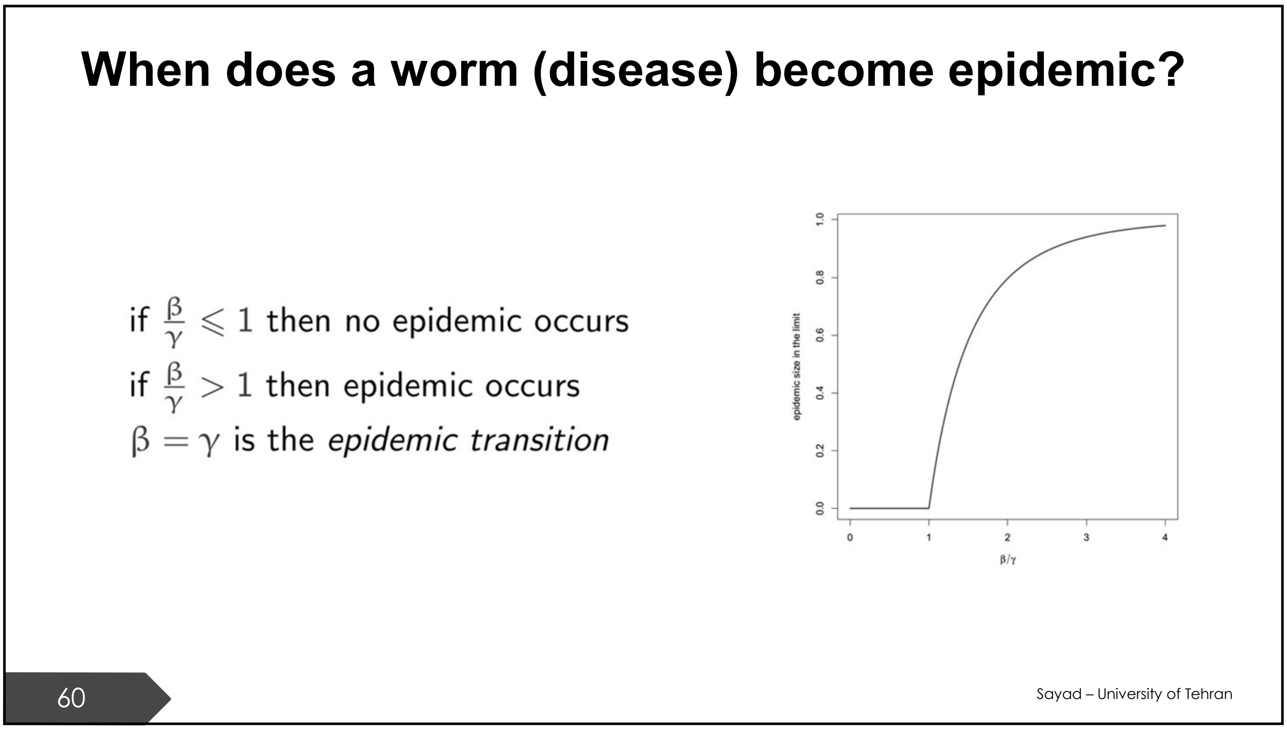
The threshold phenomenon in SIR

We see the same here. there won't be a breakout if $\beta < \gamma$

This never intersects the line y=r
So at ∞, this couldn't have happened!

59

Sayad – University of Tehran



When does a worm (disease) become epidemic?

if $\frac{\beta}{\gamma} \leqslant 1$ then no epidemic occurs

if $\frac{\beta}{\gamma} > 1$ then epidemic occurs

$\beta = \gamma$ is the *epidemic transition*

60

Sayad – University of Tehran

Some notable stuffs about SIR

- Once a node is infected, it might recover with the probability of γ per unit of time (i.e. per timeslot).
- So just like flipping a coin till getting Head, we can calculate the length of time τ that an infected individual is likely to remain infected before recovering by Geometric distribution.
- If we break τ into small pieces of $\delta\tau$, the probability of not recovering during τ is :

$$\lim_{\delta\tau \to 0} (1 - \gamma\delta\tau)^{\tau/\delta\tau} = e^{-\gamma\tau}$$

and the probability that the individual remains infected this long and then recovers in the interval between τ and $\tau + d\tau$ is:

$$p = e^{-\gamma\tau}\gamma d\tau$$

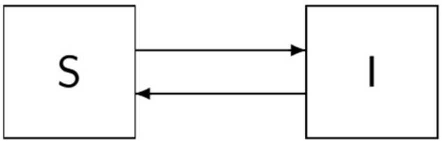
SIR – Basic Reproduction Rate R_0

- R_0 is the average number of additional people that a newly infected person passes the disease onto, before they recover☟.
  - $R_0 > 1$ means each infected person infects more than 1 person and hence the epidemic grows exponentially (at the early stages).
  - $R_0 < 1$ makes the epidemic shrink (die gradually).
  - $R_0 = 1$ marks the epidemic threshold between the growing and shrinking regime.
  - In the SIR model: $R_0 = \beta/\gamma$ (Eq. 17.16 of your textbook)

☟ This is defined for the early stages of the epidemic and so one can assume that most people are in the susceptible state.

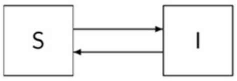
SIS (Susceptible – Infective – Susceptible) Model

- There are two states in SIS model. People who get infected, will recover but are still susceptible. Just like Flu.



- The SIS parameters are:
  - β (infection rate): probability of contagion after contact, per unit of time
  - γ (recovery rate): The probability recovery from an infection per unit of time

The SIS Model



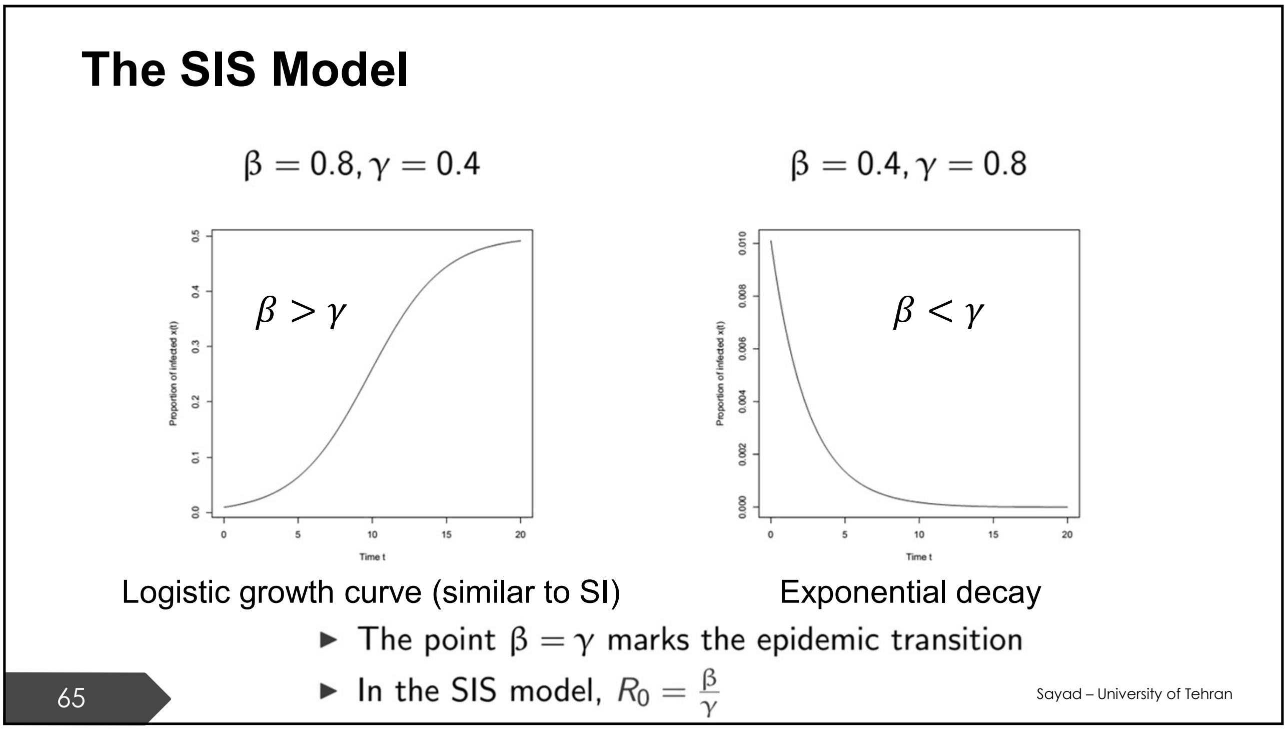
$$\frac{ds}{dt} = \gamma x - \beta s x \qquad \frac{dx}{dt} = \beta s x - \gamma x$$

$$s + x = 1 \quad (S + X = n)$$

Eq. 17.17a to 17.22 of Newman's book:

$$x(t) = \frac{x_0(\beta - \gamma)e^{(\beta-\gamma)t}}{\beta - \gamma + \beta x_0 e^{(\beta-\gamma)t}}$$

For which we have assumed x_0 is a small number initially (in a big network).



The SIS Model

$\beta = 0.8, \gamma = 0.4$

$\beta > \gamma$

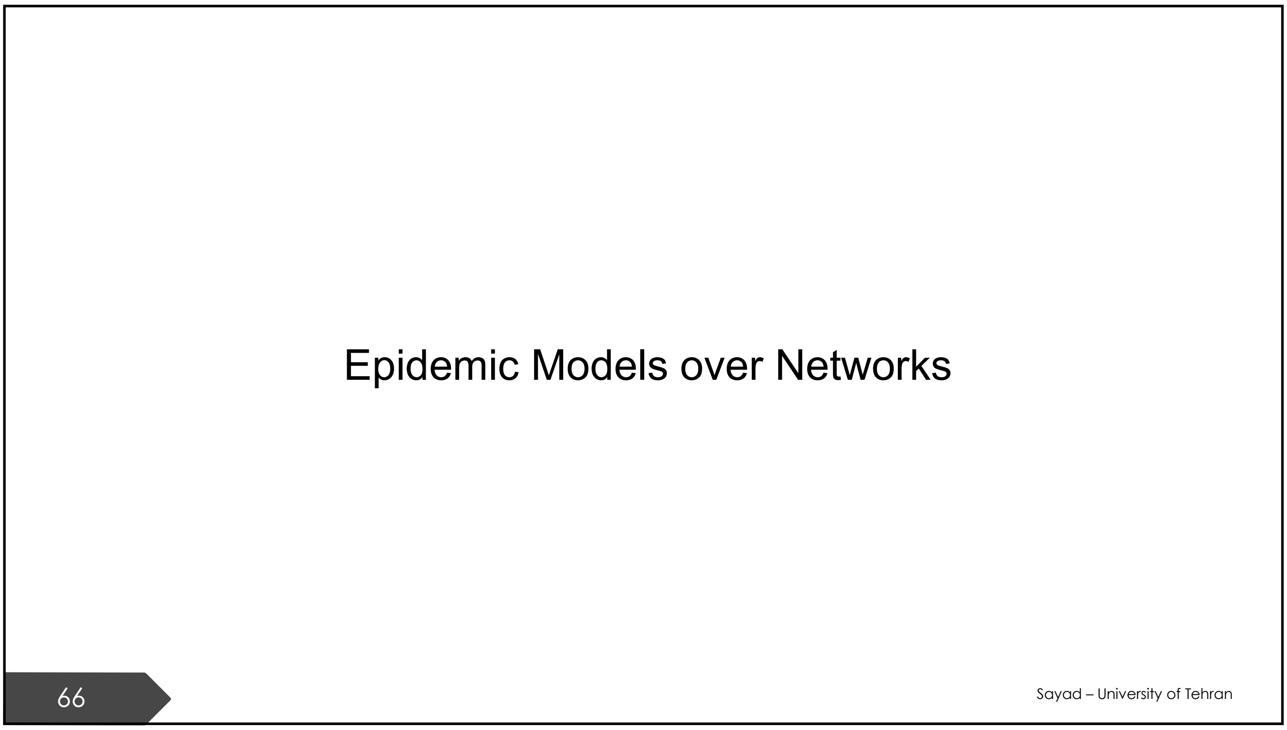
$\beta = 0.4, \gamma = 0.8$

$\beta < \gamma$

Logistic growth curve (similar to SI)

Exponential decay

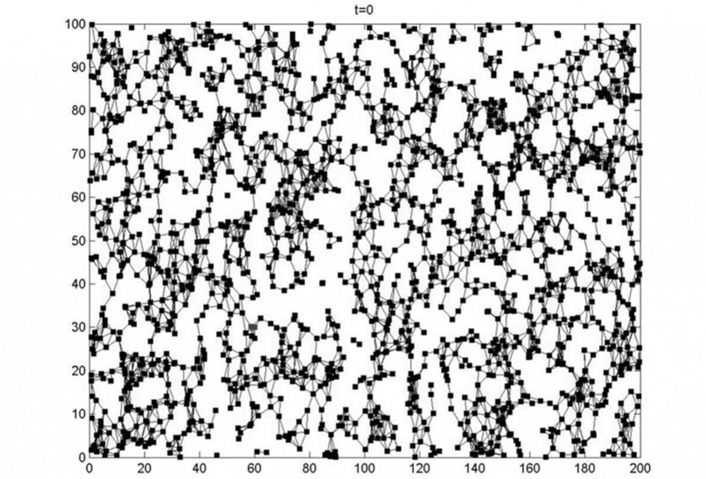
- The point $\beta = \gamma$ marks the epidemic transition
- In the SIS model, $R_0 = \frac{\beta}{\gamma}$



Epidemic Models over Networks

Propagation Modelling in Distributed Networks

- Applications:
  - Source Identification
  - Designing Fault-tolerant, robust and stable networks
  - Risk evaluation of future complex cyber physical systems



“One the race of worms and patches”, Sayad et al., 2016

Epidemic Models over Networks

➨Now imagine a more realistic network, which is not fully mixed!

➨**Homogeneous Network:**

➨Assume all nodes have degrees very close to $\langle k \rangle = E\{k\}$

➨We can re-write the equation of the epidemic models taking into account that individuals have approximately $\langle k \rangle$ possibilities of infection from neighbors.

Homogeneous SI Model

Naive Modelling:

Here, think of β as the probability of passing infection to a neighbor per contact per unit of time

$$\frac{dX}{dt} = \beta\langle k\rangle X\frac{S}{n} \quad \rightarrow \quad \frac{dx}{dt} = \beta\langle k\rangle x(1-x) \quad , \qquad \frac{ds}{dt} = -\beta s(1-s)$$

Solution:

$$x(t) = \frac{x_0 e^{\beta\langle k\rangle t}}{1 - x_0 + x_0 e^{\beta\langle k\rangle t}}$$

- ➨ Similar behaviour to the non-networked model (because practically $\beta\langle k\rangle = \beta_{old}$.
- ➨ Growth of infection depends on both β and $\langle k\rangle$.

Homogeneous SIR Model

$$\frac{ds}{dt} = -\beta\langle k\rangle sx \qquad \frac{dx}{dt} = \beta\langle k\rangle sx - \gamma x \qquad \frac{dr}{dt} = \gamma x$$

Epidemic thresholds:

- ► if $\frac{\beta}{\gamma} \leqslant \frac{1}{\langle k\rangle}$ then no epidemic occurs
- ► if $\frac{\beta}{\gamma} > \frac{1}{\langle k\rangle}$ then epidemic occurs

Similar behaviour to the fully-mixed model (because practically $\beta\langle k\rangle = \beta_{old}$.

Homogeneous SIS Model

$$\frac{ds}{dt} = \gamma x - \beta\langle k\rangle sx \qquad \frac{dx}{dt} = \beta\langle k\rangle sx - \gamma x$$

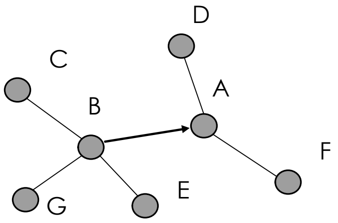
Solution: $$x(t) = x_0 \frac{(\beta\langle k\rangle - \gamma)e^{(\beta\langle k\rangle - \gamma)t}}{\beta\langle k\rangle - \gamma + \beta\langle k\rangle x_0 e^{(\beta\langle k\rangle - \gamma)t}}$$

➨Similar behaviour to the fully-mixed model

➨Epidemic threshold of $\frac{\beta}{\gamma} > \frac{1}{\langle k\rangle}$ (same as SIR)

A More Realistic Model

(Degree-based Approximation from Pastor-Satorras, 2001 and Newman,2010)



➨ Let $s_k(t)$ and $x_k(t)$ be the probabilities that a node with degree k is susceptible and infected, respectively.

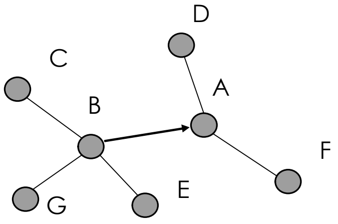
➨A is susceptible. To get infected, it should get it from one of its neighbors (e.g. B).

➨If B is infected, it must have got it from one of its neighbors, But we must exclude A as it's susceptible and couldn't have infected B!

- ➨B will have the same probability of being infected at the current time as the average vertex with degree one less.

A More Realistic Model

(Degree-based Approximation from Pastor-Satorras, 2001 and Newman, 2010)

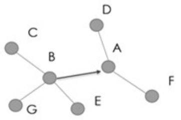


- So B's probability of infection depends upon its excess degree, the number of edges it has other than the edge we followed from A to reach it.
  - B's probability of infection is thus x_k, but k indicates the excess degree, not the normal degree. If the normal degree pmf is shown by p_k, then the excess one is q_k.

Newman, Eq. 13.46 → $q_k = \frac{(k+1)p_{k+1}}{\langle k \rangle}$

A More Realistic Model (SI)

(Degree-based Approximation from Pastor-Satorras, 2001 and Newman, 2010)



- Now, consider the probability that node A becomes infected between t and $\mathrm{d}t$. The probability that a neighbour of A is infected (at t) is:

$$v(t) = \sum_{k=0}^{\infty} q_{k'} x_{k'}(t)$$

If a neighbor is infected, then the probability that the disease is transmitted to A in the given time interval is $\beta\ dt$. Therefore, the probability of transmission from a single neighbor is $\beta v(t)\ dt$ and the probability of transmission from any neighbor is $\beta k v(t)\ dt$, where k is now the number of A's neighbors. We also require that A itself be susceptible, which happens with the probability of $s_k(t)$. So our final probability that A becomes infected is $\beta k v s_k\ dt$. Thus :

$$\frac{ds_k}{dt} = -\beta k v s_k$$

A More Realistic Model (SI)

(Degree-based Approximation from Pastor-Satorras, 2001 and Newman, 2010)

- The equation can be solved exactly as:

$$s_k(t) = s_0 e^{-\beta k \int_0^t v(t')dt'} = s_0 \times \left(e^{-\beta \int_0^t v(t')dt'}\right)^k$$

Newman 17.66 : $t \to \infty$ → $\left(e^{-\beta \int_0^t v(t')dt'}\right) \approx e^{\beta\left(1-\frac{p_1}{\langle k\rangle}\right)t}$

Now that we have $s_k(t)$, we can find the total number of infected individuals at t:

$$x(t) = \sum_{k=1}^{\infty} p_k x_k(t) = \sum_{k=1}^{\infty} p_k(1 - s_k(t))$$ ✓

A More Realistic Model (SIS)

(Degree-based Approximation from Pastor-Satorras, 2001 and Newman, 2010)

- Now imagine an SIS model:

$$\frac{dx_k}{dt} = \beta k v(\beta)(\overbrace{1 - x_k}^{s_k}) - \gamma x_k$$

in which we have emphasized the dependency of v on β. But it's the same v.

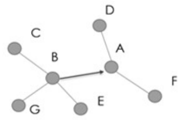
- In the stationary state (i.e. $\forall k\ \frac{dx_k}{dt} = 0$), after solving we obtain:

$$x_k = \frac{k\beta v(\beta)}{\gamma + k\beta v(\beta)}$$

Which means higher degree nodes are more susceptible to get infected.

A More Realistic Model (SIS)

(Degree-based Approximation from Pastor-Satorras, 2001 and Newman, 2010)



- And similar to before:

$$x(t) = \sum_{k=1}^{\infty} p_k x_k(t)$$

- We have an interesting work here. In the scale-free model of [Barabasi and Albert, 1999], we have $p_k = 2m^2/k^3$ and so we obtain in this case (without loss of generality at $\gamma = 1$):

$$v(\beta) = \frac{e^{-\frac{1}{m\beta}}}{\beta m} \qquad and \qquad x \approx 2e^{-\frac{1}{m\beta}}$$

which shows an interesting result → we can observe that there is no epidemic threshold for (infinite) scale-free networks.

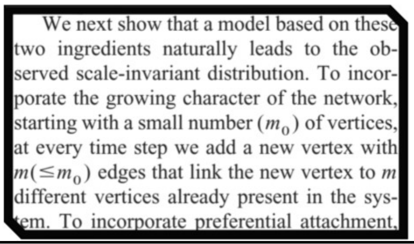
What's m? What does scale-free network mean?

- A **scale-free network** is a network whose degree distribution follows a power law, at least asymptotically. That is, the fraction *P*(*k*) of nodes in the network having *k* connections to other nodes goes for large values of *k* as :

$$P(k) \propto k^{-\alpha}$$

- where α is a parameter whose value is typically in the range 2 to 3, although occasionally it may lie outside this bound.

Barabasi proposes a way of making such networks. m is a parameter in his paper.



We next show that a model based on these two ingredients naturally leads to the observed scale-invariant distribution. To incorporate the growing character of the network, starting with a small number (m_0) of vertices, at every time step we add a new vertex with $m(\leq m_0)$ edges that link the new vertex to m different vertices already present in the system. To incorporate preferential attachment,

One More Note – A General Network Model for SIS

- ➨Assume **A** is the adjacency matrix of the underlying network and **A_{ij}** is the entry corresponding to the edge between nodes i and j
- ➨Assume A is symmetric (contagion goes in both ways) and has dimension n x n (n is the population size)
- ➨Chakrabarti et al. have proven a theorem for SIS on general networks in 2008.

One More Note – A General Network Model for SIS

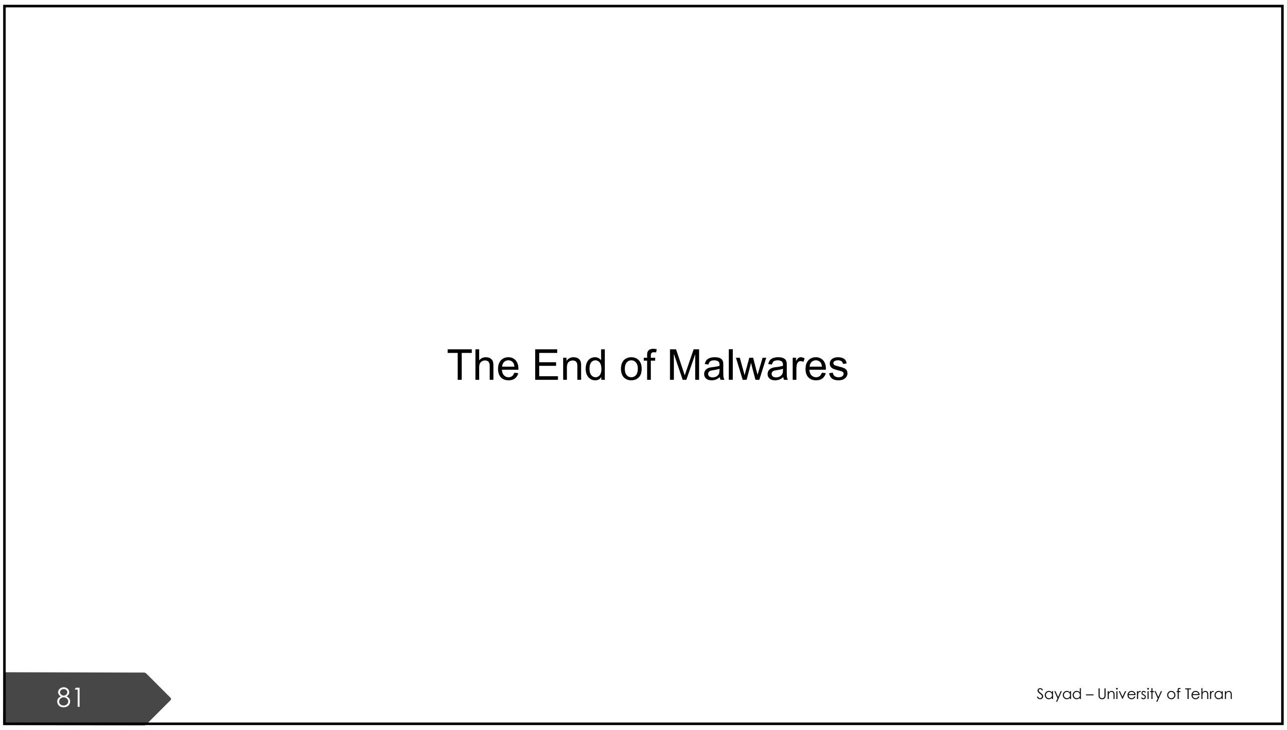
Theorem (Chakrabarti et al. 2008)

The epidemic threshold of the SIS model over arbitrary networks is:

If $\frac{\beta}{\gamma} > \frac{1}{\lambda_1}$ → epidemic occurs

If $\frac{\beta}{\gamma} < \frac{1}{\lambda_1}$ → no epidemic occurs

where λ_1 is the largest eigenvalue of the underlying contact network (**A**).



The End of Malwares

81

Sayad – University of Tehran