Misleading and Defeating Importance-Scanning Malware Propagation

Guofei Gu[1], Zesheng Chen[1], Phillip Porras[2], Wenke Lee[1]

[1]Georgia Institute of Technology

[2]SRI International

Outline

- Background
- White Hole: Design & Operation
- Misleading and Defeating Importance-Scanning Propagation
- Summary

Malware Propagation

- Email
- P2P media
- Drive-by download
- Scan-then-Exploit
  - fast
  - fully automatic, no need for human-interaction
  - remain one of the most successful, efficient and common propagation approaches

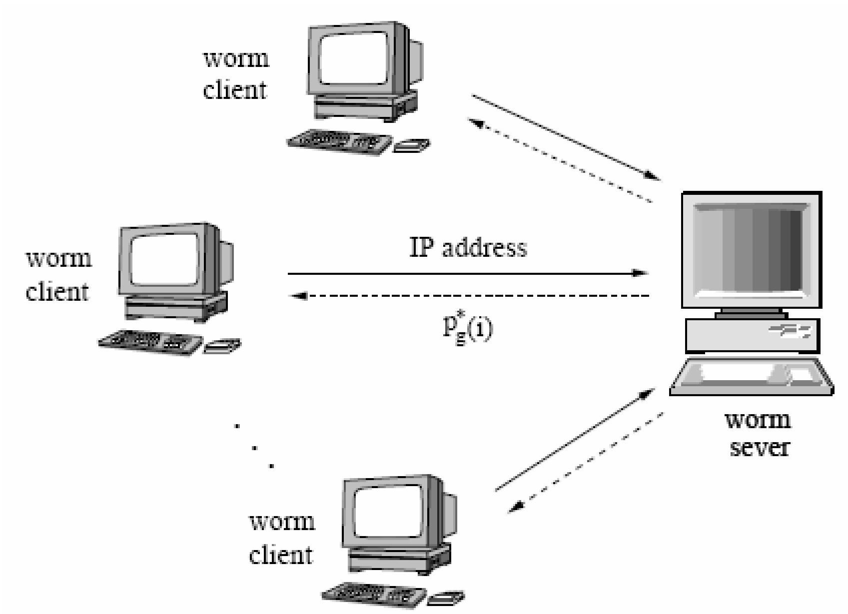
Malware Scanning Technique

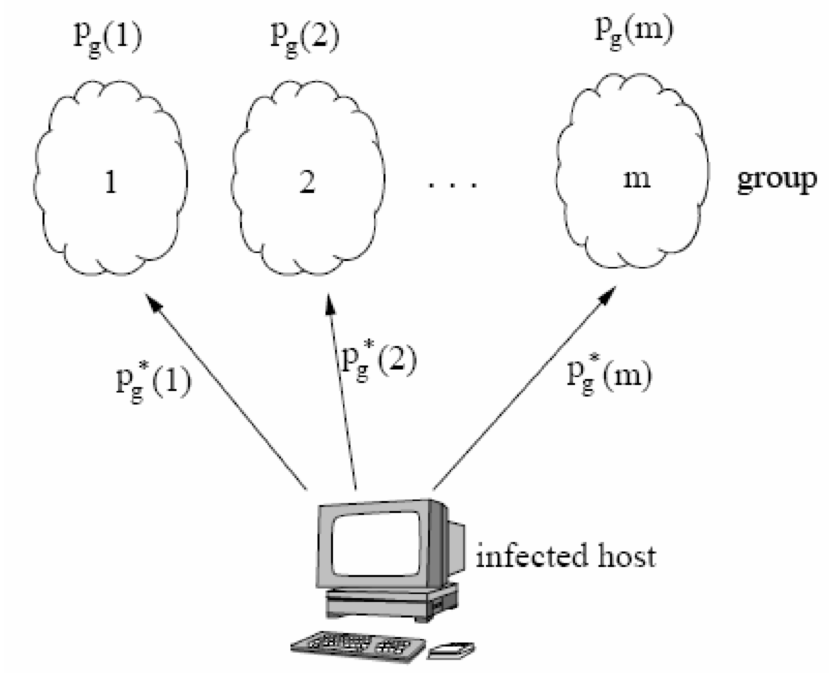
- Scanning strategies (from random scanning to more intelligent and targeted ways)
  - List based (e.g., flash worm)
    - carry on a detailed address list (IP or subnet)
    - obtain the list utilizing BGP information, or address sampling
    - fast, no waste of time on dark space
    - hard to carry a large list in practice
  - Probability based
    - carry on a probability distribution on different address space (subnets)
    - fast, and less information to carry
    - need to know the distribution

Importance-Scanning Propagation

- Two stages
  - Learning stage: to uncover (vulnerable) address distribution by obtaining report from initial propagation or through network address sampling scanning
  - Importance-scanning stage: propagate using the (vulnerable) address distribution (**probability based scanning**)

Example Importance-Scanning Malware





Importance-Scanning Propagation (cont.)

- It is shown to be **faster** than using regular scanning *([Chen et al. WORM 2005])*

- It is shown to be **hard** to counteract using host-based defense (e.g., proactive protection and virus throttling) or IPv6 *([Chen et al. Infocom 2007])*

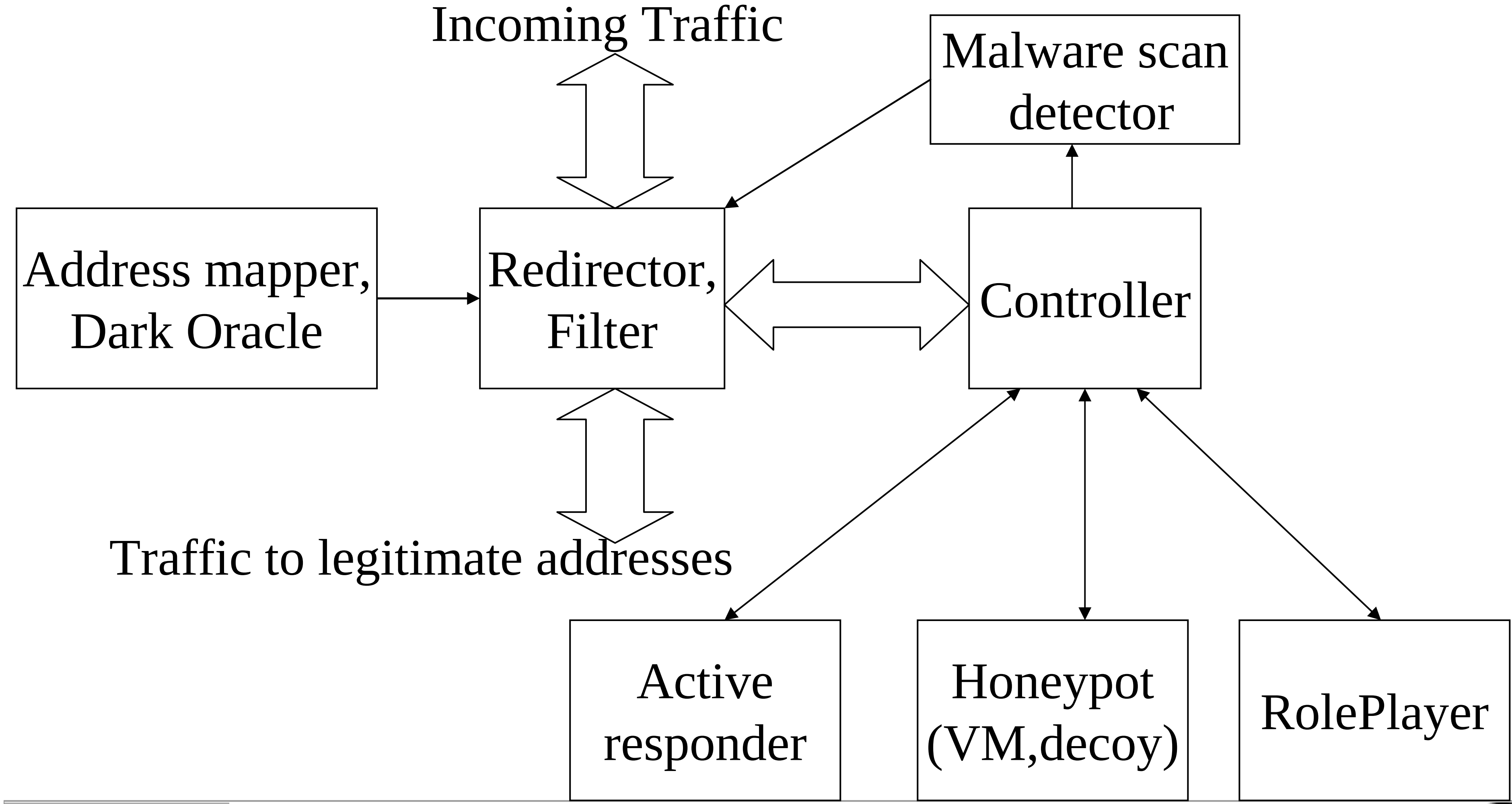
- New solution is needed ⟵ this work

Intuition of White Holes

- Hide a tree in a forest
  - Blend live targets in among phantom address (i.e., accept network connections to any addresses)

- Effect 1: reduce “regular” attacks on normal address space (as shown in OpenFire)
- Effect 2: mislead the learning of address distribution information
- Effect 3: convert the advantage of importance-scanning (the predictable affinity) to a potential vulnerability against it *(explained later)*

White Hole Architecture



White Hole Operation: General Idea

- A set of responders, honeypots, roleplayers to handle suspicious connections
  - Provide *more faked* live address information
- Malware scan detection (in the learning stage) to locate scanner and filter scans to legitimate space
  - Provide *less true* live address information
- Tarpit technique (e.g., LaBrea) to stick tcp-based malware
  - Slow down or even stop propagation (more biased information, more stuck connections)
  - Extremely effective for importance-scanning propagation

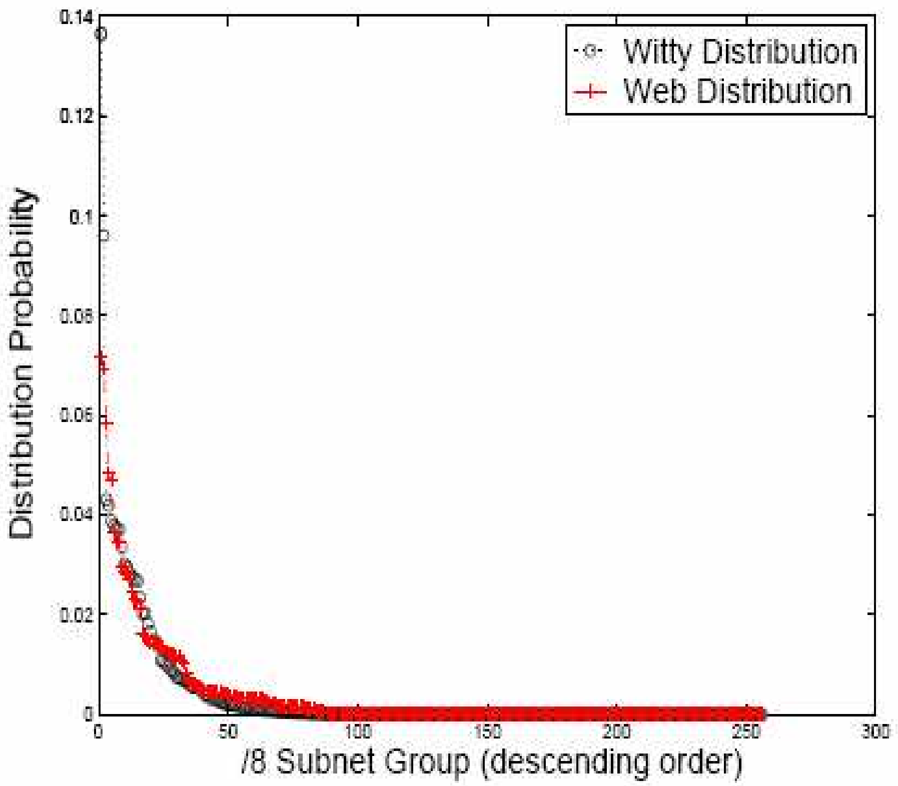
Misleading Importance-Scanning

- <u>*Infection rate*</u>: the average number of infected vulnerable hosts per unit time by a single malware at *early* propagation
  - A BGP worm speeds up 3.5 times than a regular IPv4 worm
  - An importance-scanning propagation has even higher infection rate

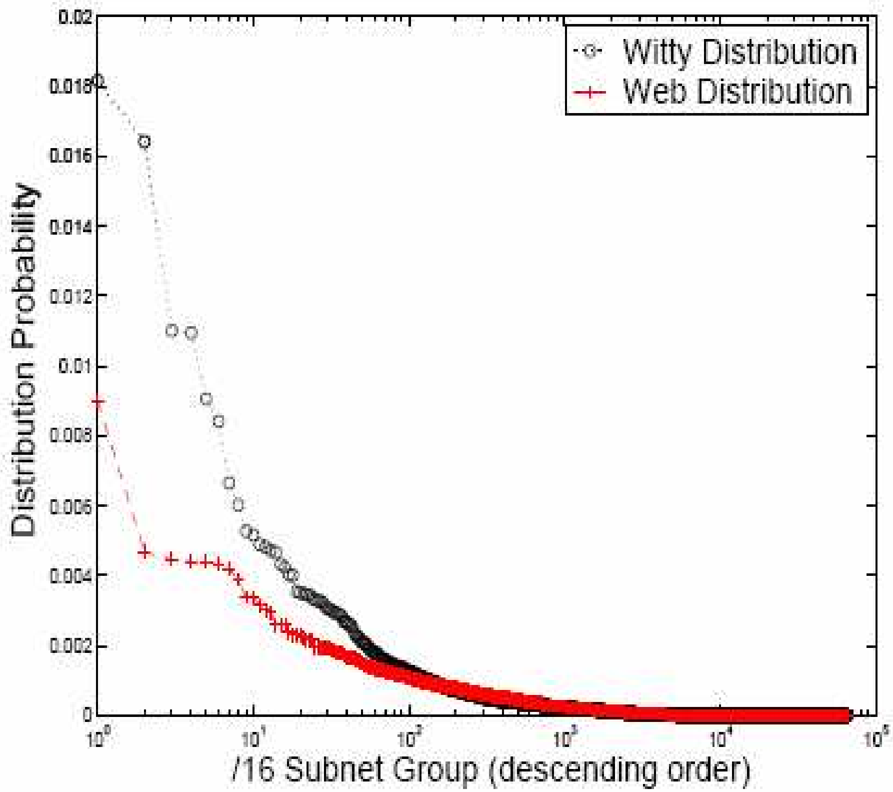
- White holes decrease the infection rate of importance-scanning propagation with a factor of **$(N\beta+U)/(N\beta)$**
  - N: # vulnerable hosts on Internet
  - U: # addresses used by white holes
  - β: correct estimation probability of true vulnerable hosts (due to wide deployment of address blacklisting)

- Misleading U: due to faked live addresses
- Misleading N: due to scan detection & filtering

Non-Uniformly Distributed (Vulnerable) Hosts on Internet

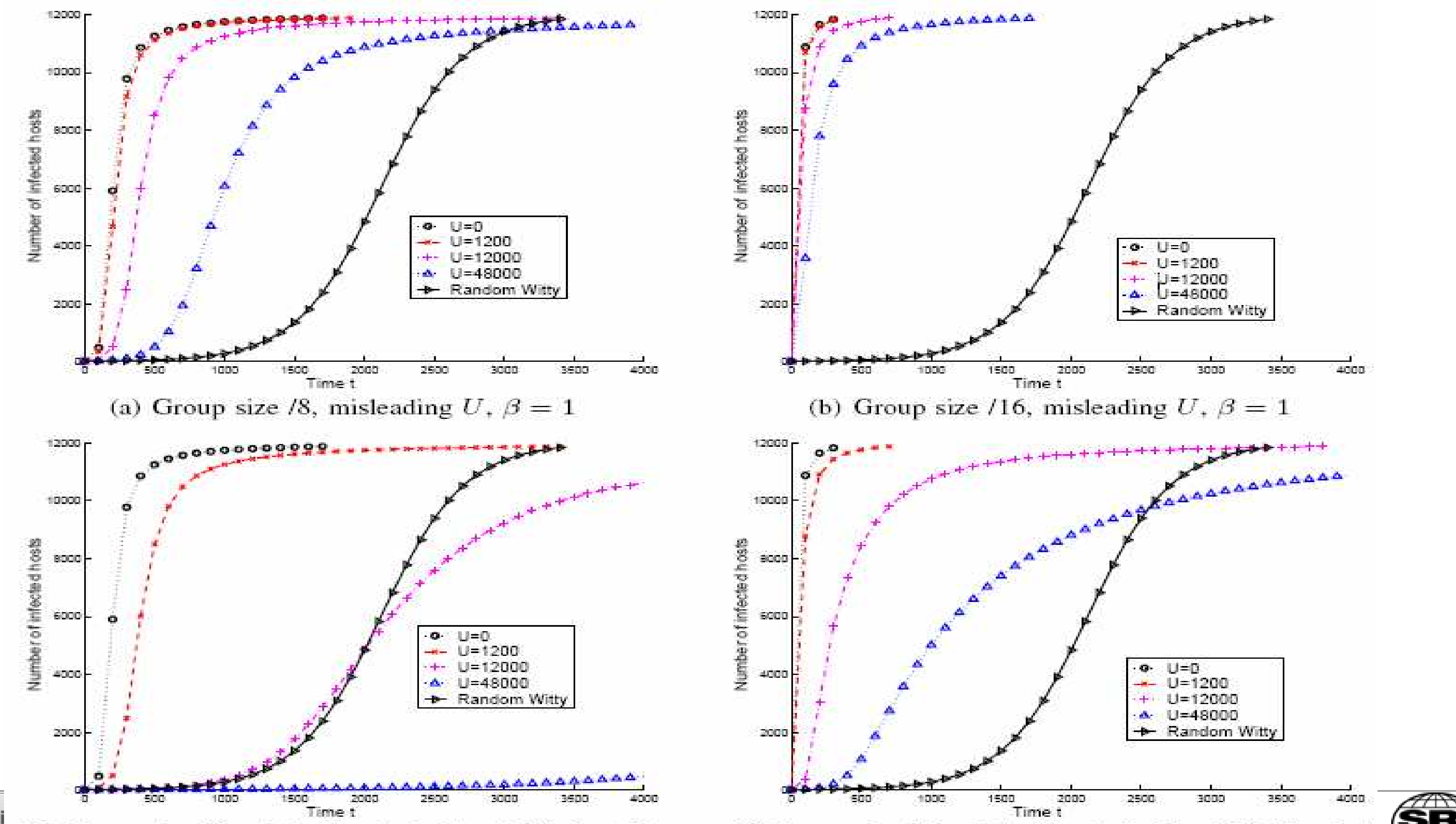


(a) In /8 subnet group



(b) In /16 subnet group (X-axis in log scale)

Effect of Misleading: Witty-Vulnerable-Distribution



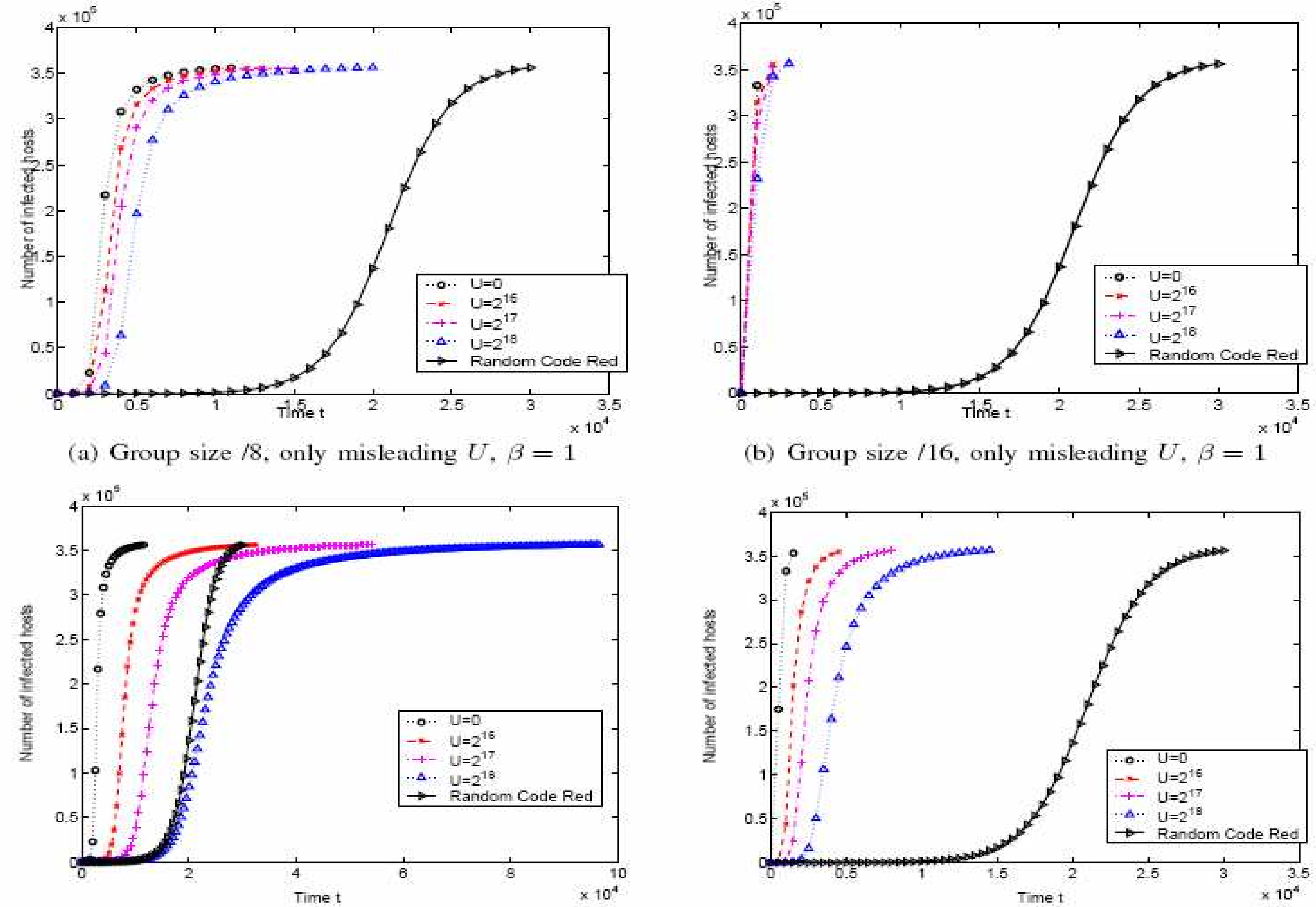
(a) Group size /8, misleading U, $\beta = 1$

(b) Group size /16, misleading U, $\beta = 1$

(c) Group size /8, misleading both N and U, $\beta = 0.1$

(d) Group size /16, misleading both N and U, $\beta = 0.1$

Effect of Misleading: Web-Distribution



(a) Group size /8, only misleading U, $\beta = 1$

(b) Group size /16, only misleading U, $\beta = 1$

(c) Group size /8, misleading both N and U, $\beta = 0.1$

(d) Group size /16, misleading both N and U, $\beta = 0.1$

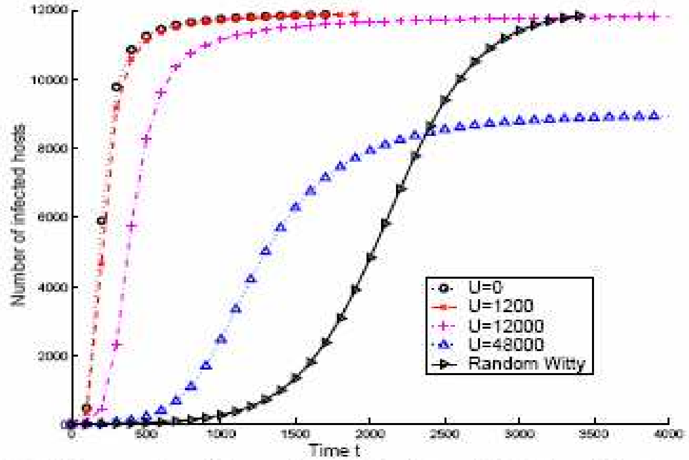
Defeating Importance-Scanning

- Further use tarpit technique in white holes
  - Stick tcp-based malware for a long time
  - Underlying reason to slow down propagation
    - there is a limitation on the number of ***concurrent*** connections a host can keep

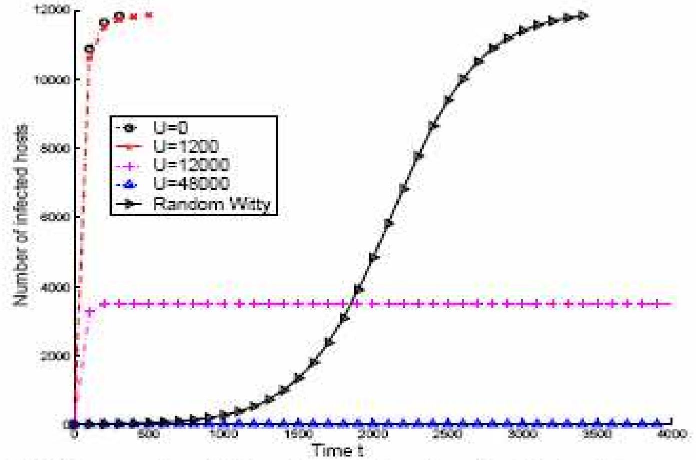
- Importance-scanning tends to scan more on dense space (the advantage of spreading faster)

- More scans to white holes → more will be trapped → less capability to spread → slow down → stop

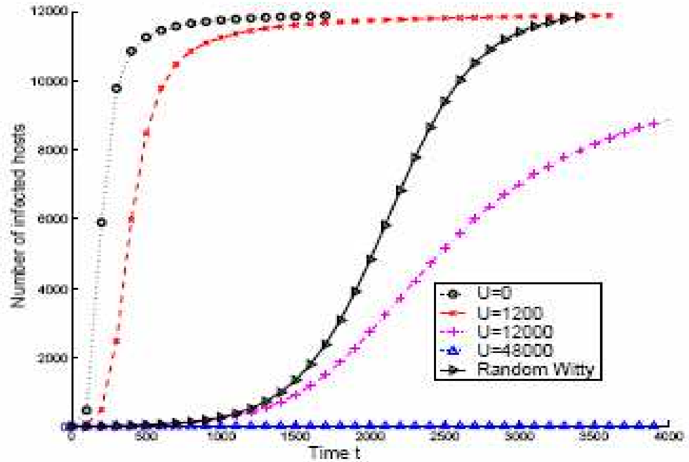
Effect of Defeating: Witty-Vulnerable-Distribution



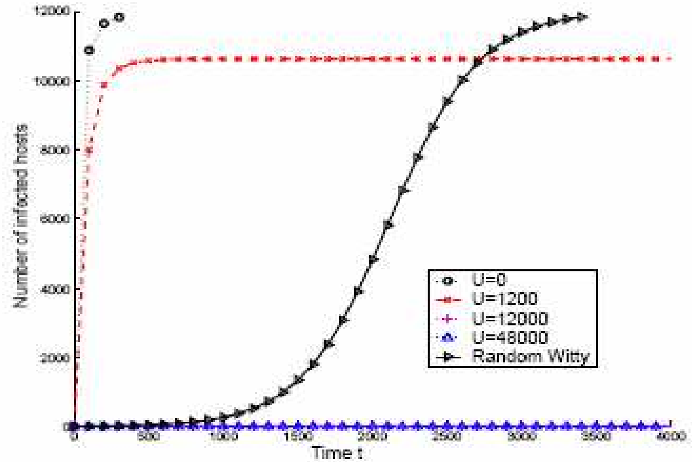
(a) Group size /8, only misleading U ($\beta = 1$), plus tarpit



(b) Group size /16, only misleading U ($\beta = 1$), plus tarpit

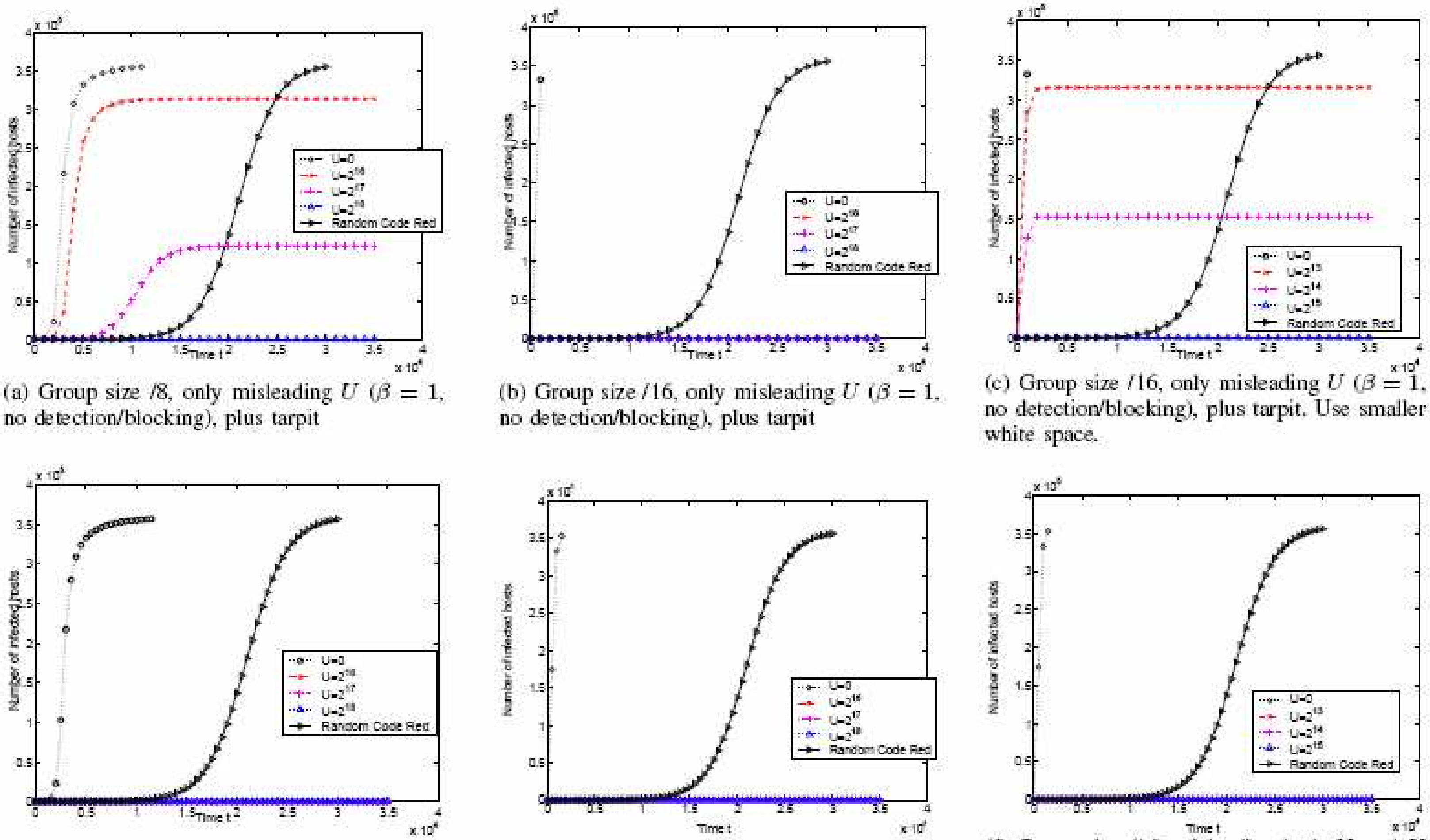


(c) Group size /8, misleading both N and U ($\beta = 0.1$), plus tarpit



(d) Group size /16, misleading both N and U ($\beta = 0.1$), plus tarpit

Effect of Defeating: Web-Distribution



(a) Group size /8, only misleading U ($\beta = 1$, no detection/blocking), plus tarpit

(b) Group size /16, only misleading U ($\beta = 1$, no detection/blocking), plus tarpit

(c) Group size /16, only misleading U ($\beta = 1$, no detection/blocking), plus tarpit. Use smaller white space.

(d) Group size /8, misleading both N and U ($\beta = 0.1$), plus tarpit

(e) Group size /16, misleading both N and U ($\beta = 0.1$), plus tarpit

(f) Group size /16, misleading both N and U ($\beta = 0.1$), plus tarpit. Use smaller white space.

Related Work

- Internet monitoring: Telescope, iSink …
- Malware/worm detectionn: Kalman filter based, DSC, …
- Honeypot/honynet: honeyfarm, GQ …
  - Besides special functionality, white hole can also serve general-purpose honeynet functionalities
- Openfire: reduce regular attacks on normal address space
  - White holes use several different response/detection techniques, and address importance-scanning malware propagation

Summary and Future Work

- White hole
  - address a new generation of malware propagation strategies – importance-scanning
  - Exploit the advantage of importance-scanning to against it
  - Use a relatively small space with satisfactory effect
- Need to further study:
  - White hole dissuasion vs. attraction (game-theoretic analysis in plan)
  - Distributed deploy strategy

Q &A

Thank you!