Int'l Workshop on DDoS Attacks
and Defenses, KAIST, 2009

컴퓨터보안연구실
COMPUTER AND COMMUNICATION SECURITY LABORATORY
인터넷침해대응기술연구센터
Internet Intrusion Response Technology Research Center

Detecting DDoS Attacks and Worms using Randomness Check

Heejo Lee
(heejo@korea.ac.kr)
Computer and Communication Security Lab
Div. of Computer and Communication Engineering
Korea University
Sept. 29, 2009

(This is a joint work with Hyundo Park and Prof. Hyogon Kim at Korea University.)

LIBERTAS
JUSTITIA
VERITAS

Contents

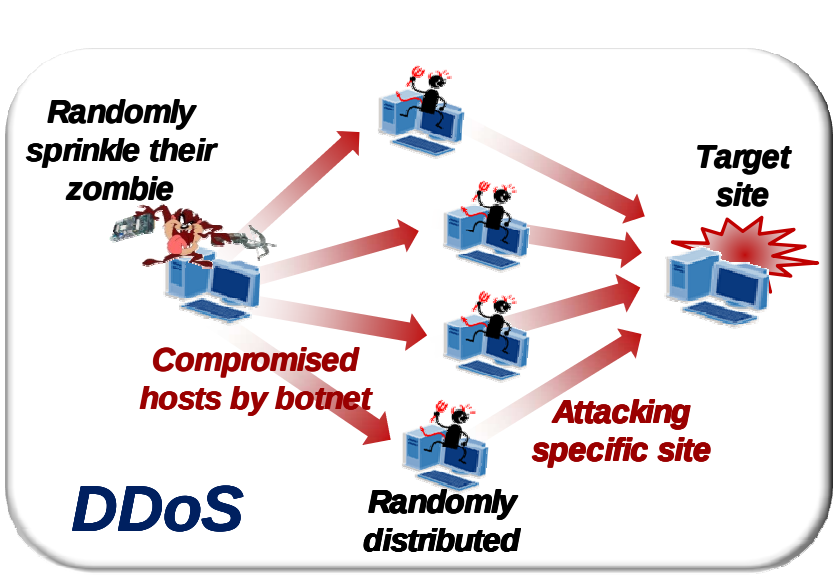
LIBERTAS
JUSTITIA
VERITAS

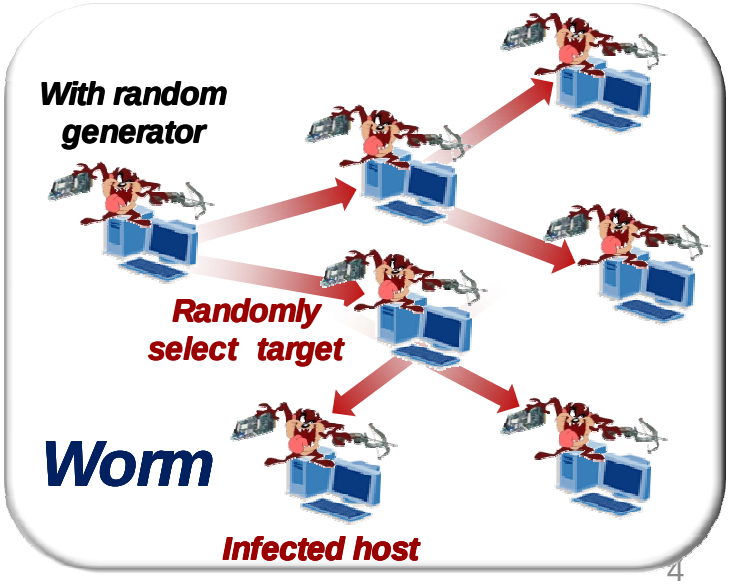
Research Motivation

- Faster detection
  - Earlier detection minimizes the damage from Internet attacks such as DDoS attacks and Internet worms
- Less computation
  - Smaller amount of computation allows to analyze more traffic in a high speed network
- Larger coverage
  - Developing a mechanism to detect more attacks is desirable

Randomness in the Internet Attack

- Internet attacks using a number of compromised hosts
  - Internet worm: selecting a next target with a random number generator
  - DDoS: sending attack traffic to a target using zombies distributed randomly in the Internet
- The feature of attack traffic
  - Internet worm: the destination addresses in network traffic are randomly distributed
  - DDoS: the clusters of zombies are randomly distributed

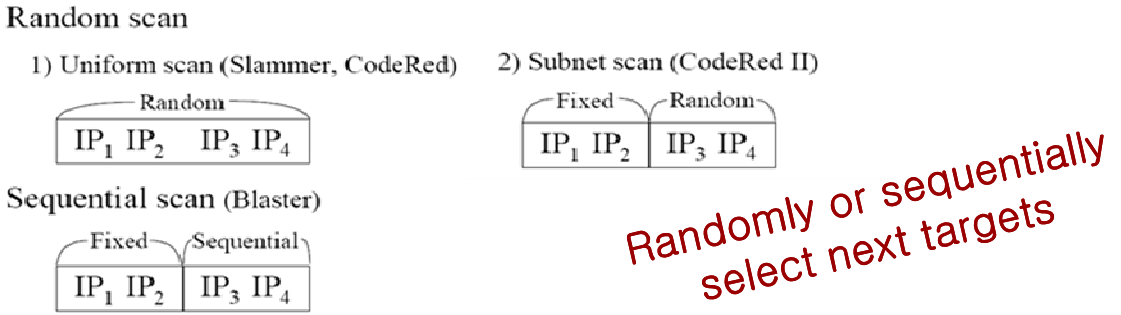




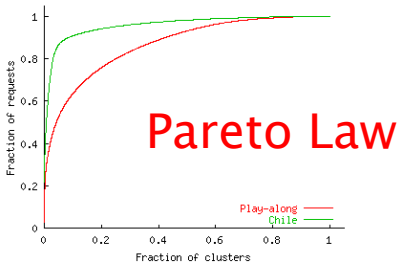
Internet Attacks and Traffic Randomness

LIBERTAS
JUSTITIA
VERITAS

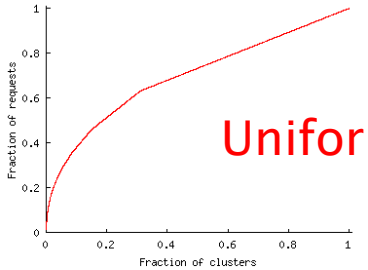
- Internet worm
  - An Internet worm uses a random generator to select target hosts.



- DDoS attacks and FE (Flash Event)
  - The distribution of clusters of zombies follows the Pareto law under FE, and is uniformly distributed under DDoS attacks



The distribution of clusters under FE



The distribution of clusters under DDoS

J. Jung, B. Krishnamurthy, and M. Rabinovich, "Flash crowds and denial of service attacks: Characterization and implications for CDNs and web sites", World Wide Web, May 2002.

LIBERTAS
JUSTITIA
VERITAS

Randomness Checks using Matrix

- Randomness checks using the rank value of a matrix [1][2]
  - Using Gaussian elimination to calculate the rank value
  - The rank value: The number of non-zero rows after applying Gaussian elimination
- The probability of the rank value of a random binary *m*x*n* matrix

$$2^{r(n+m-r)-nm}\prod_{i=0}^{r-1}\frac{\left(1-2^{i-n}\right)\left(1-2^{i-m}\right)}{\left(1-2^{i-r}\right)} \quad \text{where } r=1,2,\ldots,\min(m,n)$$

- Finding the threshold of the rank value of a random matrix

$$2^{r(n+m-r)-nm}\prod_{i=0}^{r-1}\frac{\left(1-2^{i-n}\right)\left(1-2^{i-m}\right)}{\left(1-2^{i-r}\right)}=P$$

Applying log function $\log_2\left(2^{r(n+m-r)-nm}\prod_{i=0}^{r-1}\frac{\left(1-2^{i-n}\right)\left(1-2^{i-m}\right)}{\left(1-2^{i-r}\right)}\right)=\log_2 P$

simplified as $(m-r)^2>\log_2\frac{1}{P}$

The threshold of a 256X256 matrix is fixed 252 with 99.99%

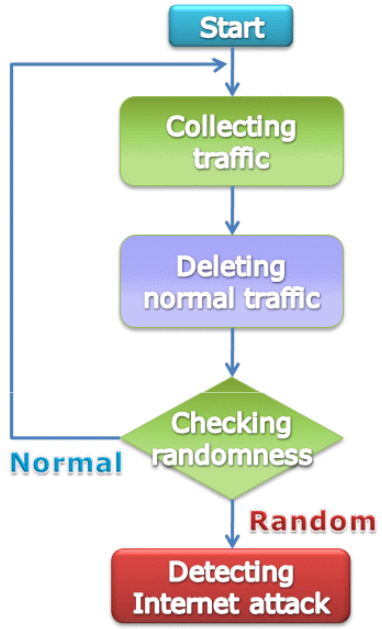
[1] G. Marsaglia and L.H. Tsay, "Matrices and the structure of random number sequences," Linear Algebra Appl., vol.67, pp.147–156, 1985.

[2] G. Marsaglia, "DIEHARD: A battery of tests of randomness," 1996. http://stat.fsu.edu/~geo/diehard.html

Early Detection Mechanism using Randomness Checks

LIBERTAS
JUSTITIA
VERITAS

- Early detection mechanism for large scale Internet attacks
  - **Collecting traffic**
    - Constructing the traffic matrix
  - **Deleting the normal traffic**
    - Deleting the normal traffic in the constructed traffic matrix
    - Using the matrix operation, as XOR and AND
  - **Checking randomness**
    - Checking randomness of the matrix, deleted the normal traffic
    - Checking randomness with the rank value



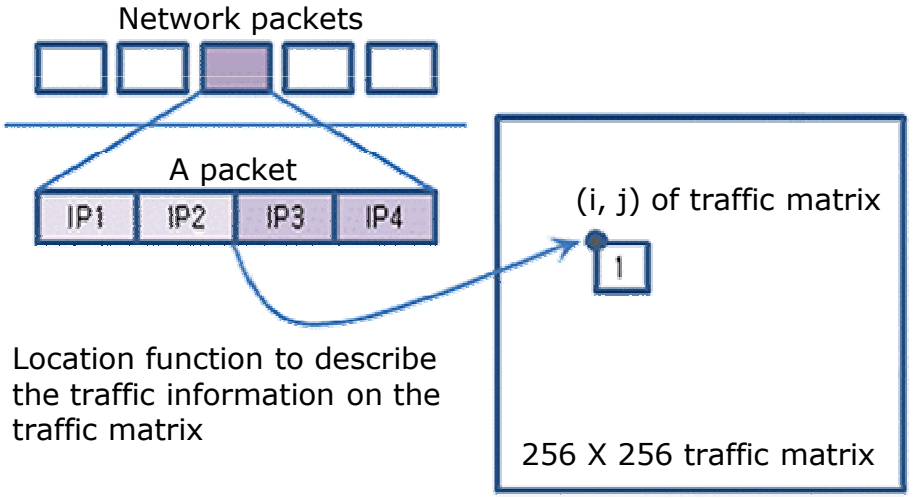
- **ADUR**(Anomaly Detection Using Randomness check)
  Detecting unknown worms using randomness check of the distribution of destination IP addresses in the network traffic

- **FDD**(FE and DDoS Distinguisher)
  Distinguishing between FE and DDoS using randomness check of the distribution of clusters among zombie machines

LIBERTAS
JUSTITIA
VERITAS

Traffic Matrix Construction

- Constructing a traffic matrix on the monitoring network
  - Describing the distribution of the traffic on the matrix
- Description method (using location function)
  - ADUR: Describing the distribution of destinations on the matrix
  - FDD: Describing the distribution of clusters of zombies on the matrix



The location function of ADUR

$$i = IP1 \oplus IP3$$
$$j = IP2 \oplus IP4$$

The location function of FDD

$$i = IP3$$
$$j = IP2$$

LIBERTAS
JUSTITIA
VERITAS

Eliminating Normal Traffic

- Features of traffic under attack
  - Normal connections between hosts remain their connection during over a time unit (e.g., 1sec) to communication
  - Infected machines send attack traffic during a short period of time toward the target hosts

    (ex: scanning, various Flooding attack traffic, and etc.)

- The matrix operation effect
  - XOR operation deletes the entries with the same values
  - AND operation maintains only the same values

$$\begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \oplus \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

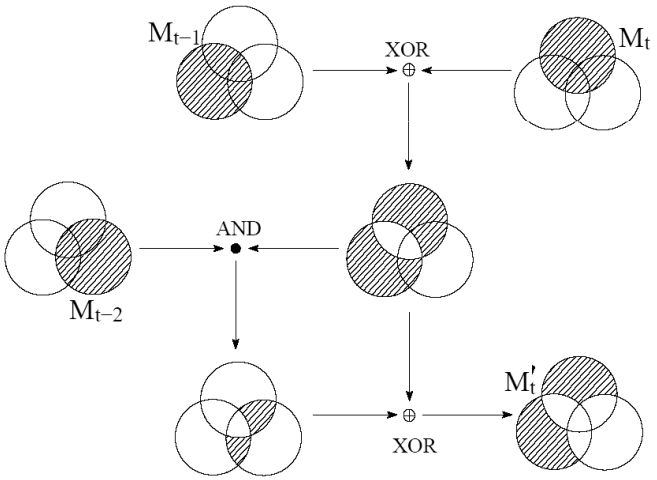
XOR effect

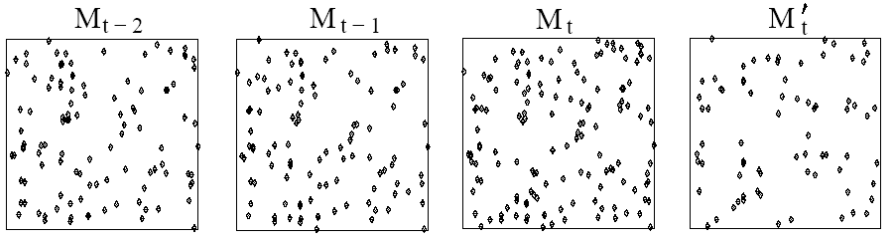
$$\begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \bullet \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

AND effect

Eliminating Normal Traffic

- M_t, M_{t-1}, and M_{t-2} are the traffic matrices constructed at time t, t-1, and t-2, respectively
- Using matrix operations such as AND and XOR, the normal traffic is reduced on the traffic matrix, thus M_t' can be constructed



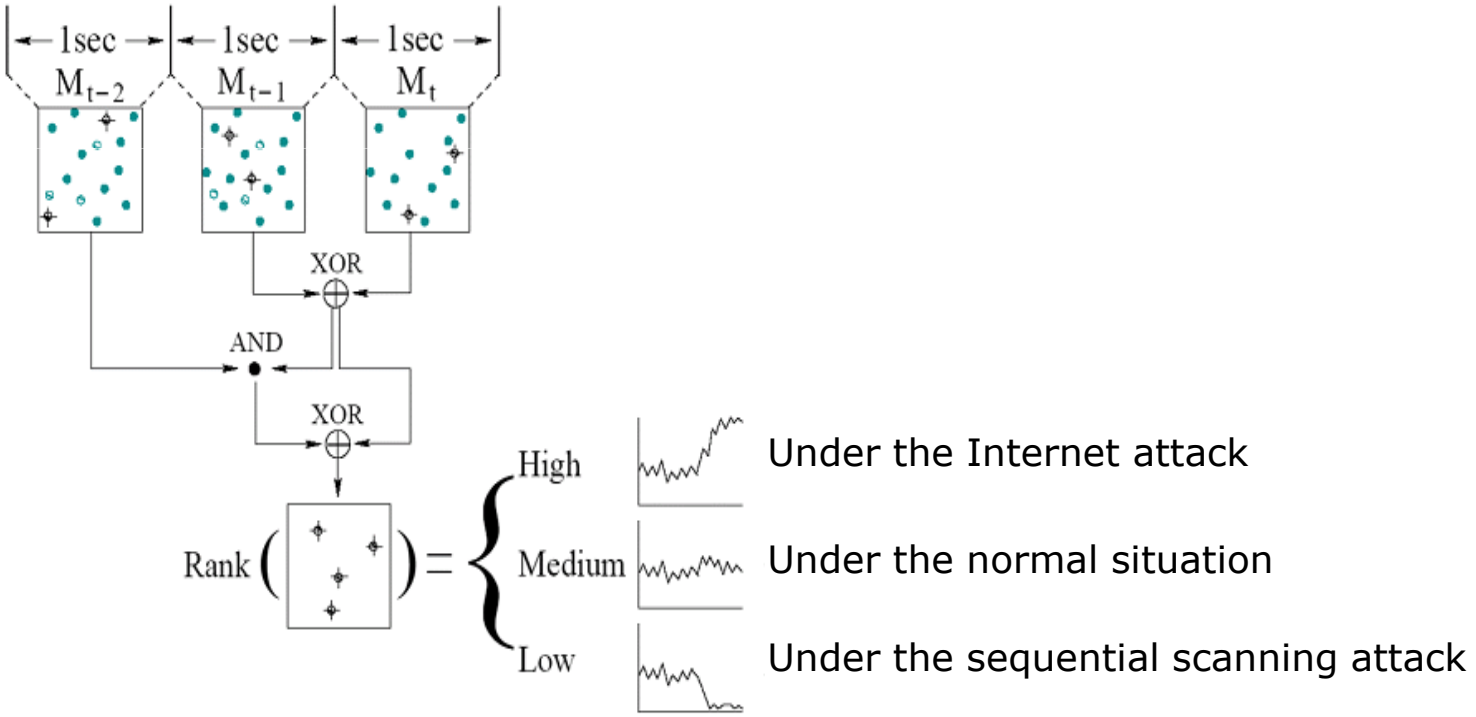


The effect of matrix operations in a /16 campus network traffic

$$M_t' = M_{XOR}(t) \oplus (M_{XOR}(t) \bullet M_{t-2})$$
$$M_{XOR}(t) = M_t \oplus M_{t-1}$$

Randomness Check using Rank Values

- The rank value of the traffic matrix as a key indicator
  - If the rank is medium, the network situation is normal
  - If the rank is high, the network situation is under attack
  - If the rank is low, the network has a worm with sequential scanning



LIBERTAS
JUSTITIA
VERITAS

Sequential Scanning Attack

- A sequential scanning attack is also detectable by the side effects of Gaussian elimination
- Sequential scanning makes a large portion of elements of the matrix become one
- If all elements in a row are 1, the row in the matrix become 0 after Gaussian elimination

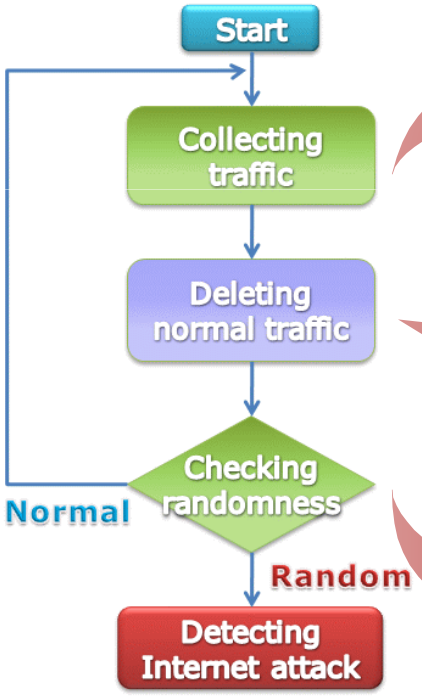
$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

The side effect of Gaussian elimination

LIBERTAS
JUSTITIA
VERITAS

Constructing Traffic Matrix for ADUR

- Detecting unknown worms using randomness check of the distribution of destination IP addresses



The location function of ADUR

$$i = IP1 \ominus IP3$$
$$j = IP2 \oplus IP4$$

Deleting the normal traffic with matrix operations, XOR and AND

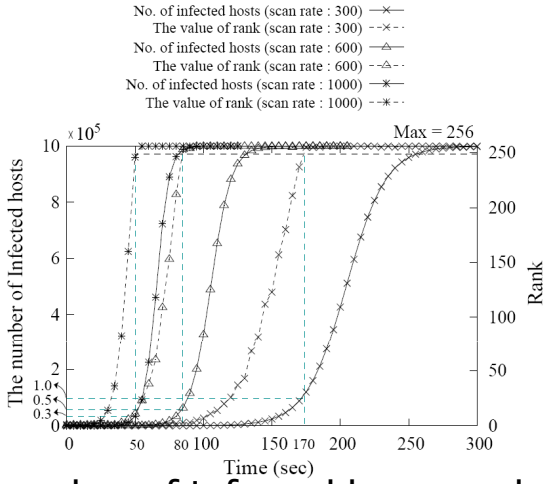
$$M'_t = M_{XOR}(t) \oplus (M_{XOR}(t) \bullet M_{t-2})$$
$$M_{XOR}(t) = M_t \oplus M_{t-1}.$$

Checking randomness with the rank value

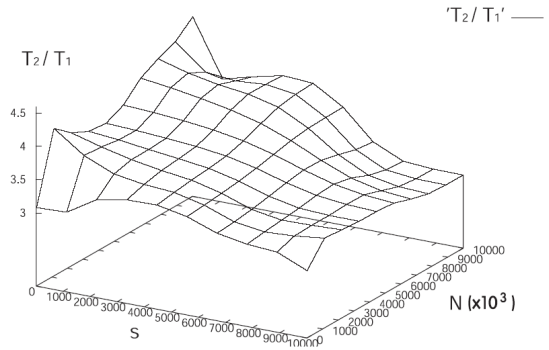
Effectiveness of ADUR

LIBERTAS
JUSTITIA
VERITAS

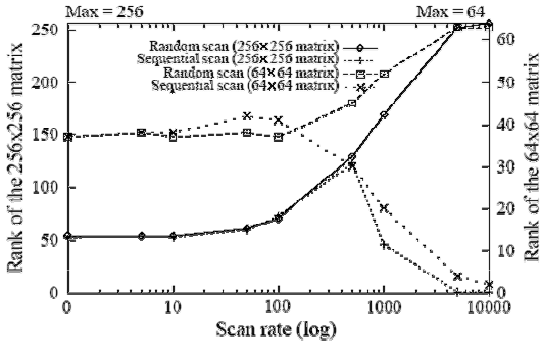
- Rank values under random and sequential scanning worm propagation situation by worm propagation model (AAWP)



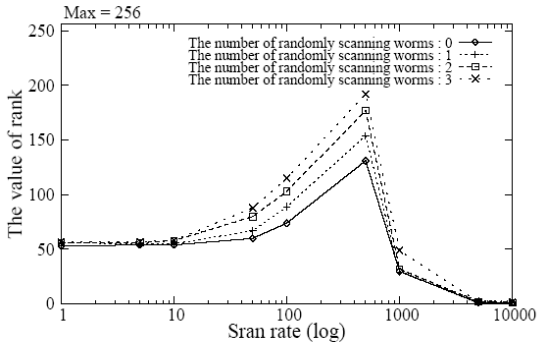
The number of infected hosts and the corresponding rank as a function of time



Three times earlier than 90% infection of vulnerable hosts



(a) Rank values as the function of a scan rate



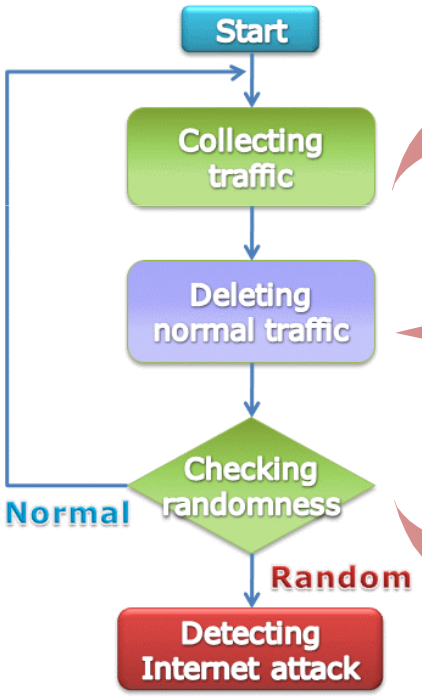
(b) Rank values as the function of a scan rate

Comparing the ranks of sequentially or randomly scanning schemes

LIBERTAS
JUSTITIA
VERITAS

Constructing Traffic Matrix for FDD

- Distinguishing between FE and DDoS using randomness check of the distribution of clusters of zombies



The location function of FDD

$i = IP3$
$j = IP2$

Deleting the normal traffic with matrix operations, XOR and AND

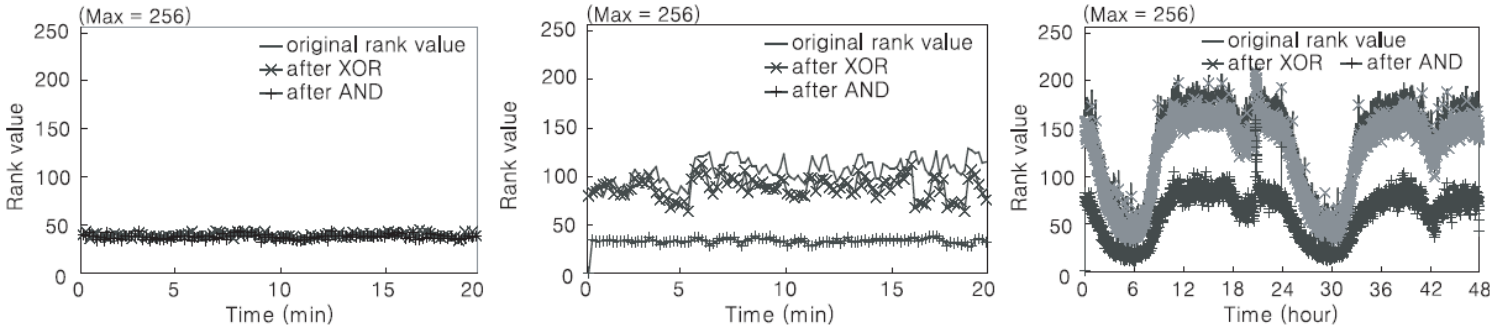
$M_t' = M_t \text{ XOR } M_{t-1}$
$M_t' = M_t \text{ AND } M_{t-1}$

Checking randomness with the rank value

15

Effectiveness of FDD

- Rank values of the traffic matrix under FE



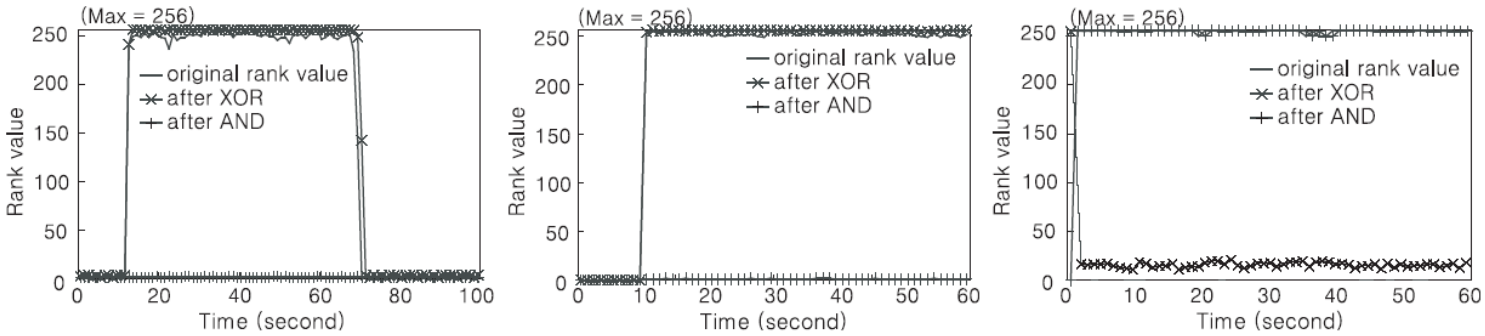
Rank values under FE

  - FE01 and FE02 are gathered on the Internet gateway between the United States and Korea
    - FE01 : numerous clients sent a large number requests to a server to download newly issued versions of java scripts for their personal websites and blogs (http://files.cometsystems.com)
    - FE02 : the traffic of Microsoft Windows update website that attracted a large number of requests when an accumulated patch to Windows Internet Explorer was released
  - MBC : the traffic of the biggest private broadcast company in Korea

Effectiveness of FDD

LIBERTAS
JUSTITIA
VERITAS

- Rank values of the traffic matrix under DDoS attacks



Rank values under DDoS attacks

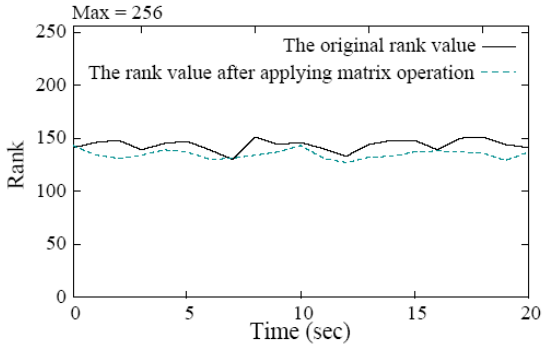
  - DDoS01 and DDoS02 are gathered on the Internet between the United States and the Korea
    - Extracted source spoof DDoS attacks

      H. Kim, S. Bahk, and I. Kang, "Real-time visualization of network attacks on high-speed links," IEEE Network Magazine, vol. 18, pp. 30-39, 2004.
  - Non source spoofed DDoS : generating network simulator (NS-2) with general website traffic by CAPBELL[1]

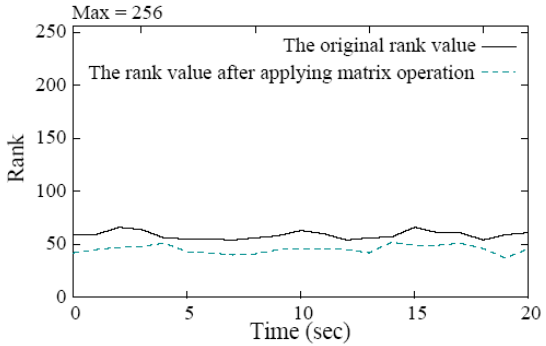
[1] A. Feldman, A.D. Gilbert, P. Huang, and W. Willinger, "Dynamics of IP traffic: A study of the role variability and the impact of control", ACM SIGCOMM, 1999.

LIBERTAS
JUSTITIA
VERITAS

Rank Values under FE



(c) CAIDA01



(d) CAIDA02

– CAIDA01 and CAIDA02 : traffic from CAIDA that record the traffic from an ISP located at the Equinix data center in Chicago, connected to Seattle, through a OC-192 pipe

  - CAIDA01 : The average numbers of new connections and packets per second are 19 and 988, and total number of packets during 20second is 19754
  - CAIDA02 : The average number of new connections and packets per second are 2 and 729, and total number of packets during 20 seconds 14588

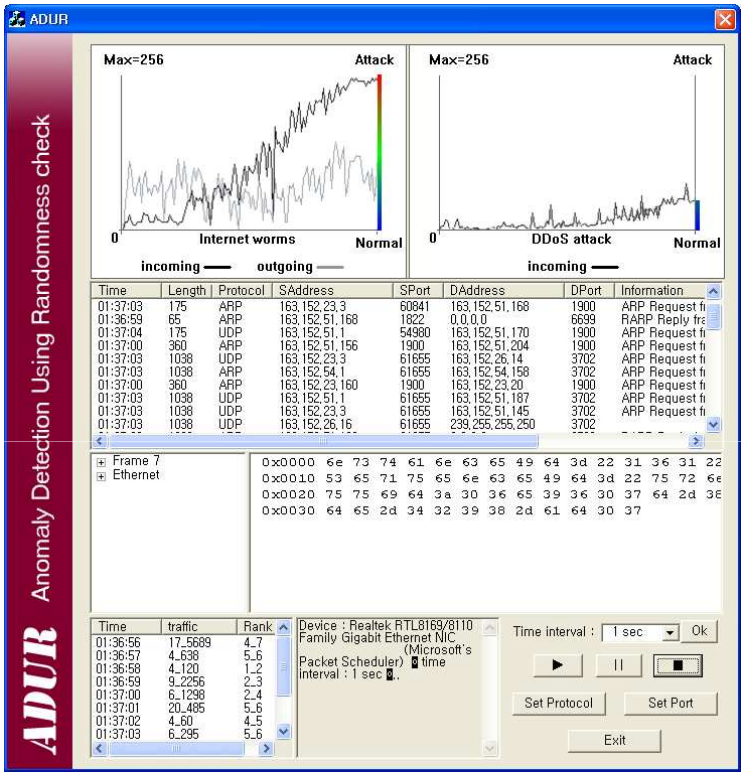
LIBERTAS
JUSTITIA
VERITAS

Conclusion

- Internet attacks detection mechanism using randomness checks
  - Detecting unknown Internet attacks using widespread compromised hosts in the Internet
- Contributions
  - Memories effects (8Kbyte in the case of a 256X256 matrix)
  - Easily applying to the system
  - Represented a value (rank)
  - The flexibility of the size of matrix
  - The fixed threshold
- Future works
  - Developing the mechanism that can forecast the future Internet attacks.
  - Developing the mechanism that can extract to only DDoS or FE traffic to block attack traffic when FE and DDoS are mixed at same time.

LIBERTAS
JUSTITIA
VERITAS

http://ccs.korea.ac.kr/ADUR



- ADUR
  - H. Park, H. Kim, and H. Lee, "Is Early Warning to Imminent Worm Epidemic Possible?", IEEE Network Magazine, Sep. 2009.
- FDD
  - H. Park, P. Li, D. Gao, H. Lee and R. H. Deng, "Distinguishing between FE and DDoS Using Randomness check", Information Security Conference(ISC), LNCS, Vol. 5222, pp. 131-145, Sep. 2008.

LIBERTAS
JUSTITIA
VERITAS

Q & A

- Thank you for your attention!!