A robust approach of detecting anomalous hosts

Network Anomaly Diagnosis Workshop 2006

Tatsuya Mori, Ryoichi Kawahara, Noriaki Kamiyama
NTT Service Integration Laboratories, Tokyo, Japan

Keisuke Ishibashi
NTT Information Sharing Platform Laboratories, Tokyo
Japan

tatsuya@nttlabs.com

NTT

Motivation

- Anomalous host
  - A host whose behavior is *not* normal
    - worm-infected hosts, bots
      - vertical, horizontal scanning
    - attackers or victims of DDoS
    - misconfigured servers
    - etc…
- We need to find anomalous hosts
  - to protect users/customers
  - to learn their characteristics
    - to create a new ACL/signature
    - to learn the dynamics/mechamisms
    - to make workload models

Motivation cont.

- Problems:
  - anomalous activities could be **buried** under the normal activities
  - Super spreader $\neq$ worm-infected hosts
    - public stratum-1 NTP server, TLD DNS server
  - We need a robust approach of identifying anomalous hosts.

NTT

Idea

- Characterizing communication pattern of each host
  - no payload information
    - can be extracted from NetFlow record
  - anomalous hosts exhibits intrinsic communication pattern

- Naïve Bayes Classifier (NBC) for analyzing communication pattern
  - Simple approach
    - calculation cost is low while the accuracy is good
  - Robust identification
    - E.g., classifying spam messages

Communication pattern of each host

- We analyze source IPs in this work

- Definition of the communication pattern
  - For each source IP:
    - A1: # of dst IPs / # of flows (0～1)
    - A2: # of dst ports / # of flows (0～1)
    - A3: # of acked dst IPs / # of dst IPs (0～1)
    - A4: # of flows / # of packets (0～1)

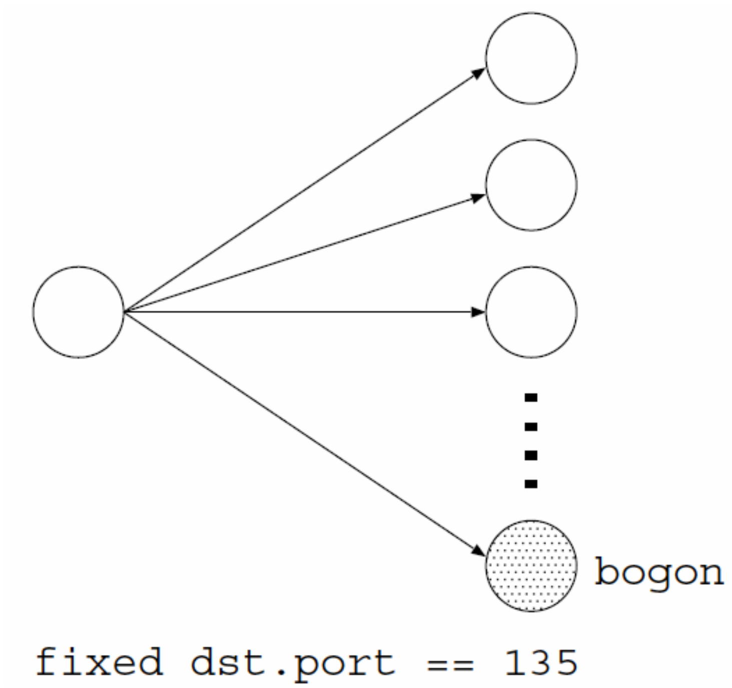
- Other metrics were also considered
  - bogon ratio, src ports statistics etc.
  - We chose the best 4 attributes in term of mutual information

Communication pattern of each host

- Merit
  - It can express concentration and dispersion easily
    - entropy-like statistics
  - more rich and flexible than entropy-based statistics
    - E.g., mean #pkts / flow, bogon ratio etc.
- Demerit
  - It needs to extract "cardinality" from massive flow data
    - Some counting techniques such as probabilistic counting or Bloom filter will be required.

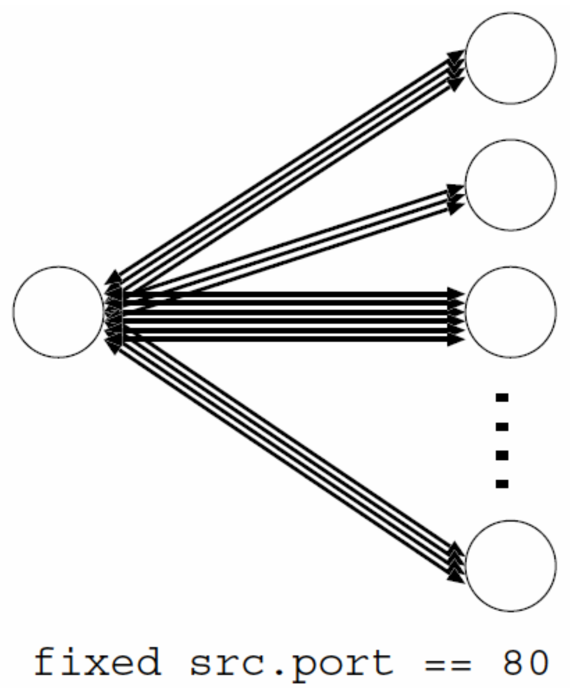
Typical examples (mental model)

Worm-infected host



$A^* = \{1,0,0,1\}$

Large-scale web server



$A^{**} = \{0,1,1,0\}$

Measured data

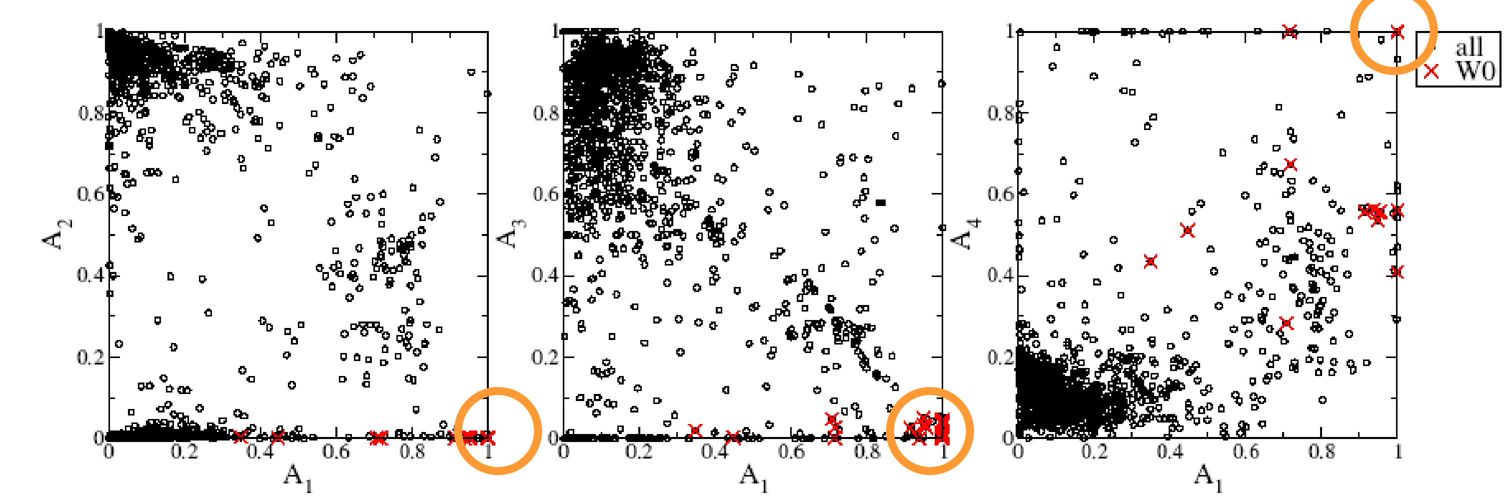
- 1Gbps Internet backbone link
- 5 minutes of measurement
- Number of source host = 116,889 (H)
- Number of host that generates more than 300 flows = 1,340 (H')
  - Mean flow generation rate >= 1 fps

Communication pattern

W0:
A set of heuristically extracted worm-infected hosts

protocol	destination port	example of worms
TCP	135	Blaster
TCP	139	Welchia
TCP	1433	Slammer
UDP	1434	Slammer

1340 hosts were examined; # of W0 hosts was 26



NBC (1)

- Machine learning method based on Bayesian inference
- Supervised classification technique
- Used in many network applications
  - Spam filtering, passive OS fingerprinting, intrusion detection etc.
- Calculation cost is low
  - linear to data size
- Robust classification

NBC (2)

- **learning: likelihood probability**
  - the probability that the attribute vector of a host is A, given that the class of the host is Ci,

$$P\left(\mathbf{A}|C_i\right) = P\left(A_1, A_2, \ldots|C_i\right)$$

- **classification : a priori probability**
  - the probability that the class of a host is Ci, given that the attribute of the host is A (measured attribute)
  - Bayse theorem + assumption of independence

→Ci is obtained by the maximum a priori probability (MAP)

$$P\left(C_i|\mathbf{A}\right) = \frac{P\left(C_i\right)}{P\left(\mathbf{A}\right)} \prod_j P\left(A_j = a_{jk}|C_i\right)$$

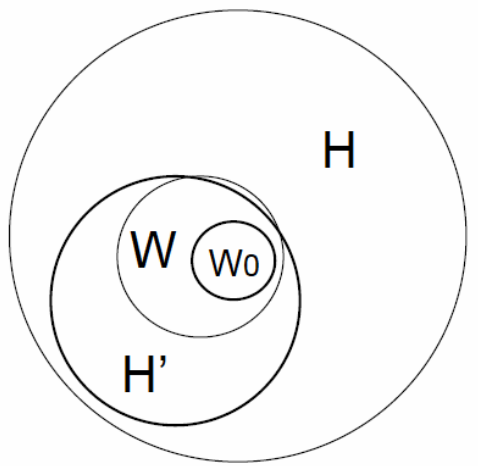
Identification procedure

- D_1: training data
- D_2: test data
- C_1: class of worm-infected hosts
- C_2: class of other hosts

- Learning: Calculate the likelihood probability $P(A|Ci)$ for D_1
- Classification: calculate the a priori probability $P(Ci|A)$ for D_2

How to train the classifier in reality?

- Problem : we don't have complete labeled data
  - ideal labeling : W → C1, H'\W → C2
- Solution :
  - heuristic labeling : W0 → C1, H'\W0 → C2



H: total hosts
H': target hosts
W: worm-infected hosts
W_0: heuristically extracted worm-infected hosts

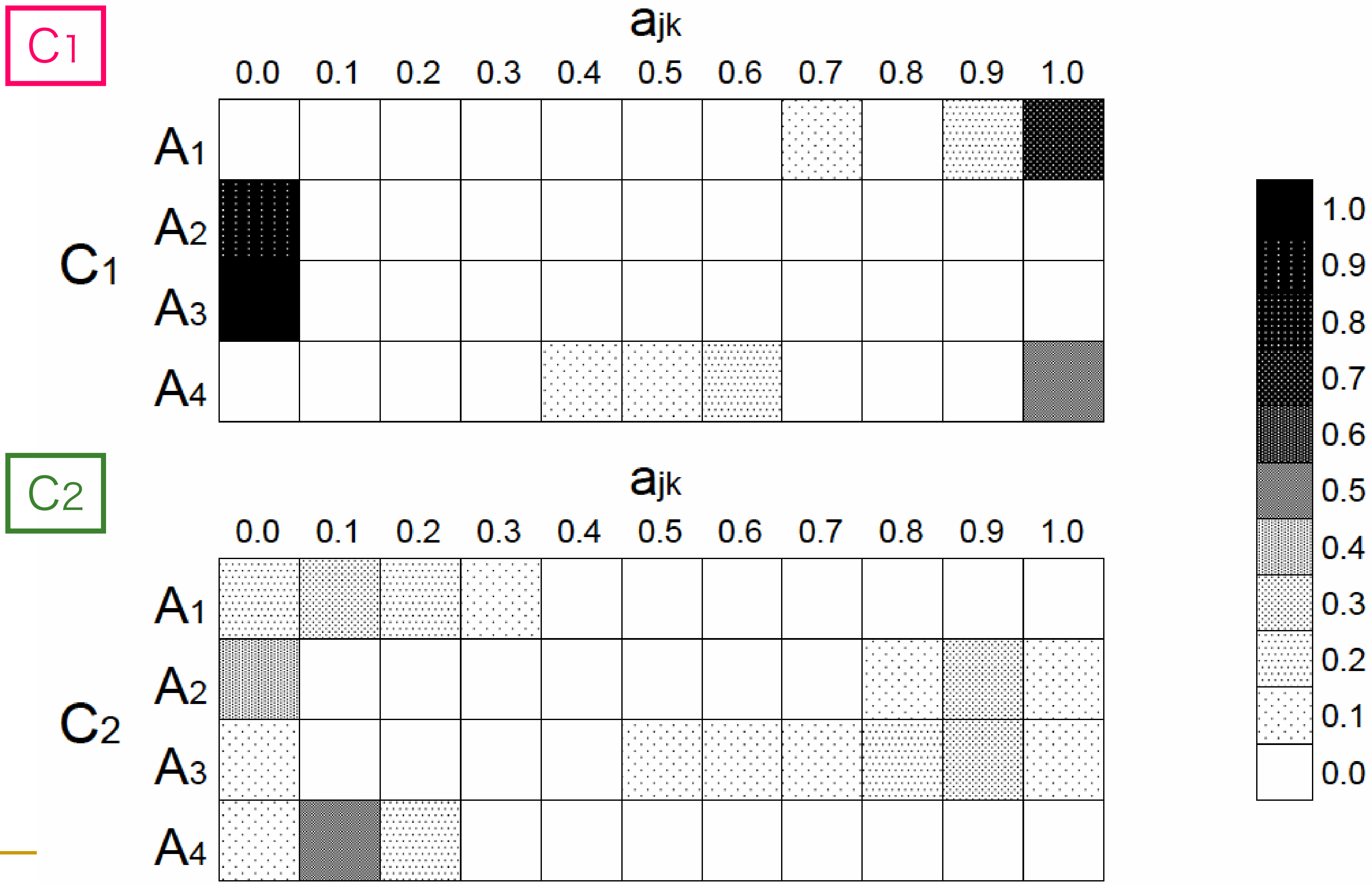
If the #H' >> #W_0, this assumption holds

NTT

Statistics of measured data

	$\mathbf{D_1}$	$\mathbf{D_2}$
Total # of packets	53,811,483	52,207,374
$n(\mathbf{H})$	116,889	115,690
$n(\mathbf{H}')$	1,340	1,358
$n(\mathbf{W_0}) = n(\mathbf{C_1})$	26	25
$n(\mathbf{C_2})$	1,314	1,333

Likelihood probability P(A|C)



Confusion matrix

		estimated class	
	$\rightarrow$	$\widehat{C_1}$	$\widehat{C_2}$
labeled class	C_1	24	1
	C_2	22	1,311

- Classified 46(=24+22) hosts as C_1
  - Newly found 22 <u>suspicious</u> hosts (== W*)
- Misidentified 1 host as C_2 (false negative)
  - The host seems to communicate with a honeypot

Analysis of communication of W*

combination	# of src hosts	# of dst hosts	# of packets	name of worms etc.
(TCP,445)	9	19,821	37,830	Sasser
(TCP,6129)	2	8,916	10,199	DameWare scan
(UDP,1026)	2	8,838	10,661	MS Messenger spam
(UDP,1027)	2	8,810	10,659	MS Messenger spam
(TCP,3306)	2	4,963	8,855	MySQL UDF Worm (Bot)
(ICMP,—)	1	1,939	1,939	Welchia
(UDP,137)	4	1,730	1,735	Qaz, OpaSoft
(TCP,15118)	1	1,192	2,520	Dipnet/Oddbob
(TCP,135)	1	298	529	Blaster, Lovsan
(TCP,9898)	1	180	180	Dabber, Doomran

Extracted hosts that sends the packets with some of these combinations to more than 300 destination addresses (1 address / sec)
→20 of 22 hosts matched the condition
the rest 2 hosts are also likely to be worm-infected

Results for other data

- Training data = D1
- Test data = D3
  - measured at different network, 100Mbps international backbone link, WIDE MAWI dataset

Estimated class

Labeled class

$\rightarrow$	$\widehat{C_1}$	$\widehat{C_2}$
C_1	33	0
C_2	15	74

NTT

Results for other data cont.

- Number of newly found suspicious hosts = 15
- Found *unknown* combination
  - TCP 445, ICMP, UDP 1028, TCP 80, UDP 137 etc.
    - udp.1028 : Kilo, SubSARI or messenger spam (?)

- 14 of 15 hosts sent the packets with some of these combination to more than 900 dst hosts (1 address / seconds)

- It is very likely that those newly found hosts are worm-infected.

summary

- Method of identifying anomalous hosts was presented
- idea
  - communication pattern of each host
  - leveraged the NBC
- Validation through the measured data (for worm-infected hosts)
  - newly found (unknown) worm-infected hosts
  - correctly classify the hosts whose attribute vectors deviated from the typical pattern.
    → thus, it is robust
  - there exists estimation error
    - Needs for better data set

NTT

Future/ongoing work

- Labeling
  - in this work, we trained the classifier with heuristic labeling
  - use the labeled + unlabeled data set for NBC
    - Class of unlabeled item can be estimated with the EM algorithm + NBC [NIGAM 99]

- Improvement of the training data
  - Coping with the honeypot etc.

- Counting distinct elements efficiently
  - data streaming approach
    - K. Ishibashi et al., "Finding top N hosts in cardinality", IEEE NetDB 2006
    - T. Mori et al., "NetDelta", submitted

Acknowledgement

This work was supported by Ministry of Internal Affairs and Communications, Japan.