A Comprehensive Approach to Detect Unknown Attacks via Intrusion Detection Alerts

ASIAN2007

December 9-11, 2007

Carnegie Mellon University Qatar Campus, Doha, Qatar

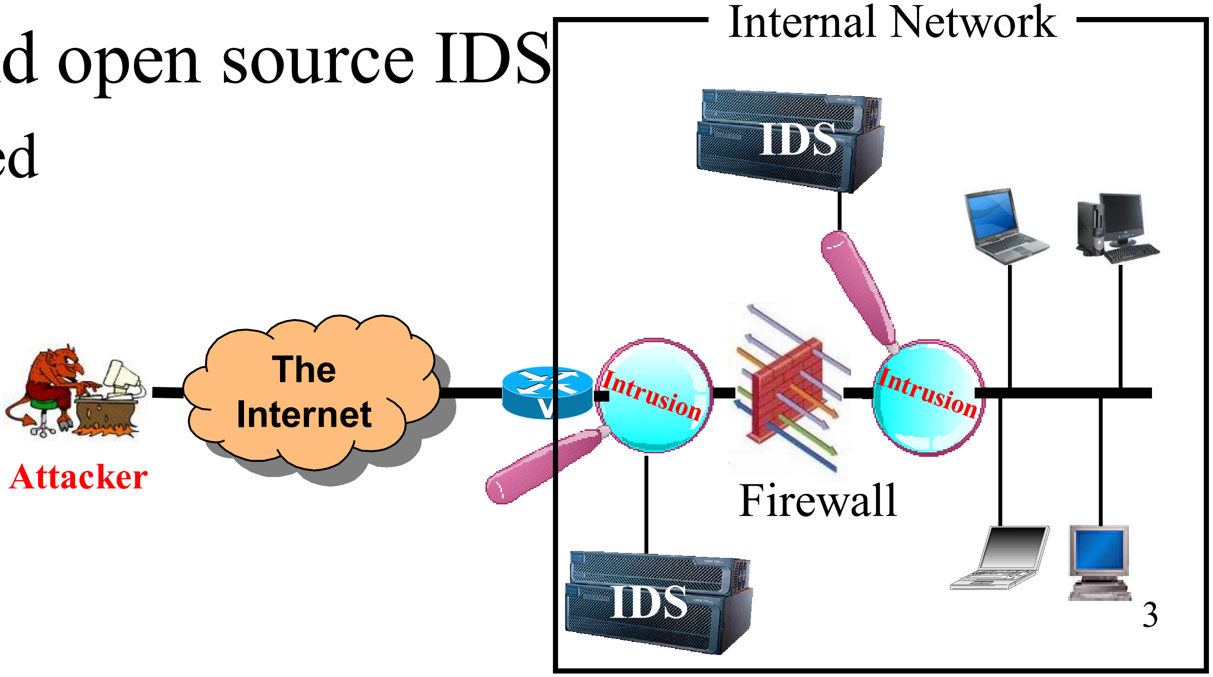
Jungsuk Song*, Hayato Ohba*, Hiroki Takakura**,
Kenji Ohira*, Yasuo Okabe** and Yongjin Kwon***
*Graduate School of Informatics, Kyoto Univ.
**Academic Center for Computing and Studies, Kyoto Univ.
***Information and Telecom. Eng., Korea Aerospace Univ.

Table of Contents

- Introduction
  - IDS(Intrusion Detection System)
  - Technical issues
- Our approach
  - Feature construction
  - Extracting representative points
  - Scoring
- Summary and future works

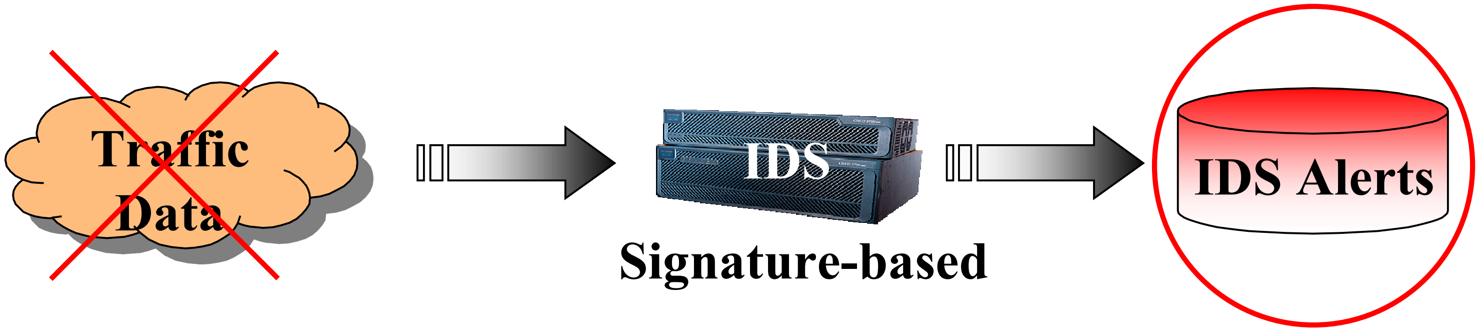
IDS(Intrusion Detection System)

- Combination of software and hardware that attempts to perform intrusion detection
- Raise the alerts when possible intrusion or suspicious patterns are observed
- Commercial and open source IDS
  - Signature-based



Technical Issues

- How to detect unknown attacks, i.e., 0-day attack
  - Signature-based IDS can detect only known attacks
- How to reduce false positives
  - 99% of the IDS alerts is false positive
  - Difficult to determine which alerts are unknown attacks or more dangerous
- Our Approach



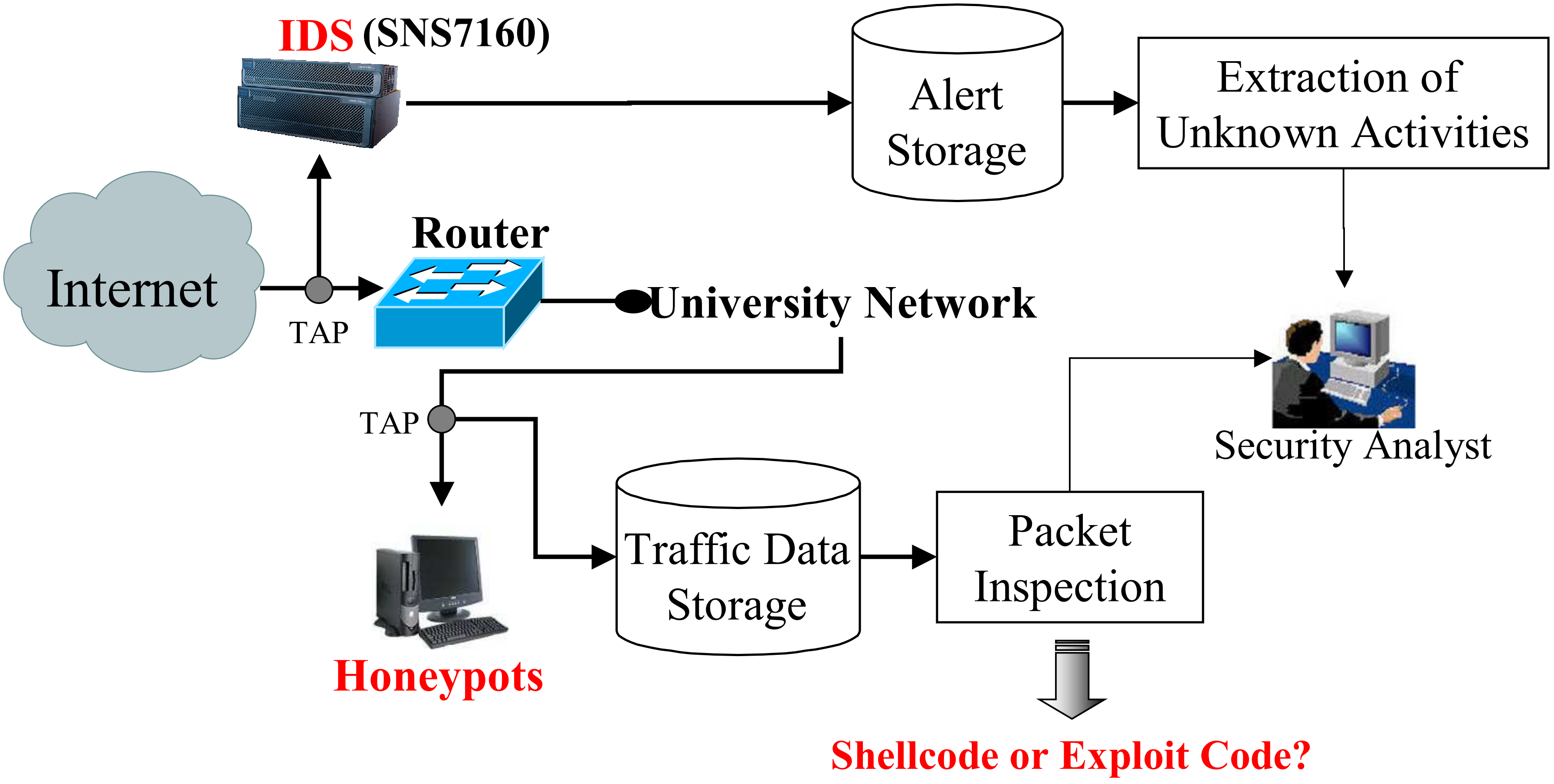
Why IDS Alerts?

- Attackers try to hide their activities
- Many security devices, including IDS, are deployed
- Anyone can easily get many IDS products and free software

⇓

- They cannot hide their activities completely because there are wide variety of security devices, including IDS
- They sometimes try to raise alerts intentionally by sending well-crafted packets so that they induce IDS operator’s misjudgment
- After that, unknown attacks are started to the targeted vulnerability
- Their combination and frequency are quite different from already-known attack activities

Experimental Environment



Example of IDS alerts(SNS7160)

Features	Data	Time	Incident ID	Number	Severity	Reliability	Signature ID	Src address
	Src port	Dst address	Dst port	Protocol	Interface	Start	End	Event Name
Example 1	2006-08-01	00:01:02	44ce1a3a9b0bc73d	395	6	10	220016	10.133.226.96
	22	61.144.21.34	58363	TCP	re1000g1	5	5	Missing SSH2 Key Exchange
Example 2	2006-08-01	00:01:02	44ce1b369732683a	1	4	8	501141	210.12.21.105
	61689	10.36.116.6	80	TCP	re1000g0			TCP FIN-ACK Portsweep

- Original Features

  ⇒ not enough to extract attacker’s ingenious conduct
- Statistical Features
  - to extract hidden and unusual patterns

Incident ID : a group of the alerts that are considered as correlated attacks by IDS.

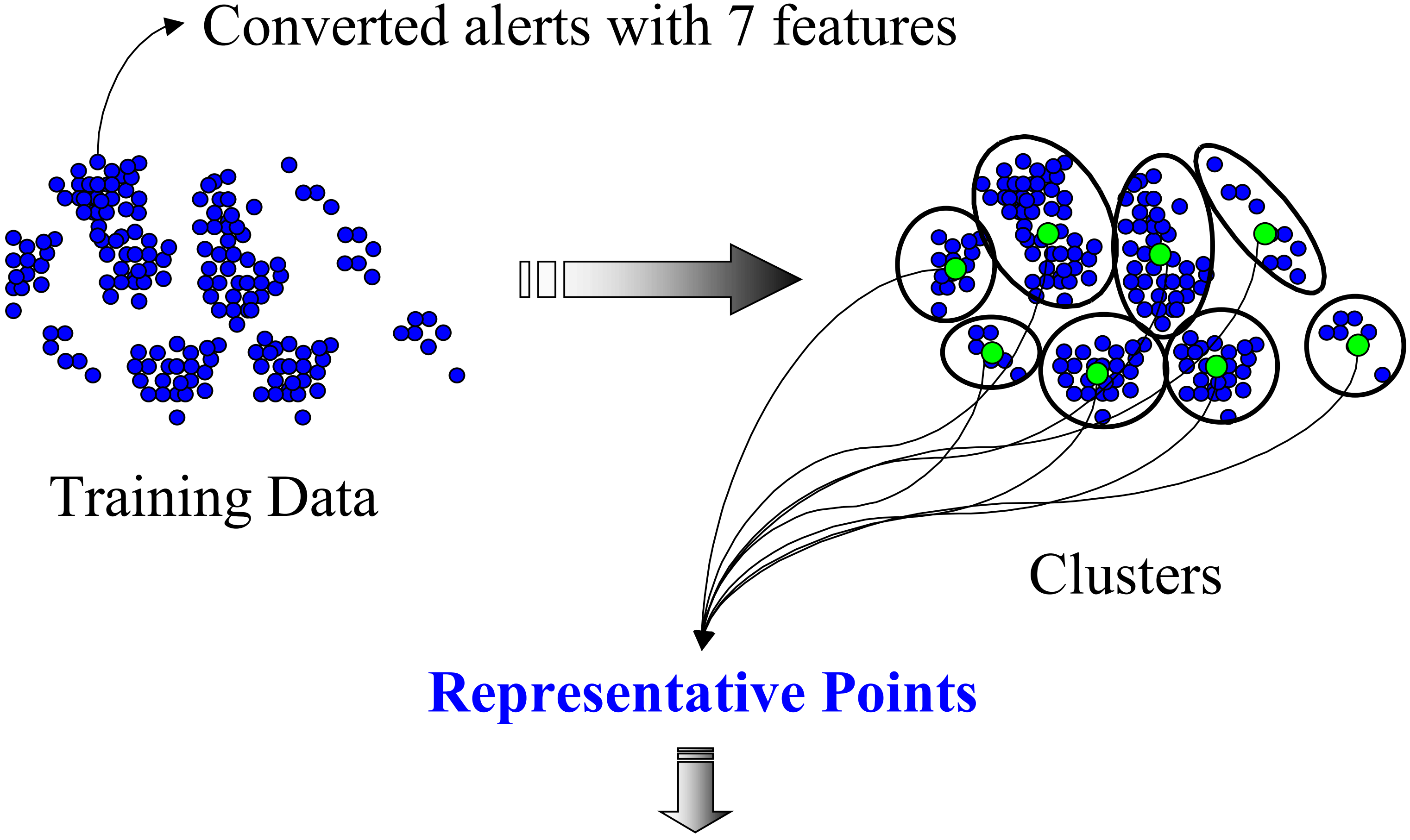
Feature Construction

- Basic features
  - Source address and port
  - Destination address and port
- Additional features(Incident ID-based)
  1. num_same_incident
     - Number of alerts with the same incident ID as the current alert
     - Detection of the attacks that consist of a large number of simultaneous connections such as DoS attack and Probing attack
  2. num_diff_alert
     - Number of different kinds of alerts within an incident group
  3. kind_sequent_alert
     - Kinds of alerts that appear after the current alert
     - Detection of new or excessive combinations of IDS alerts

Extracting Representative Points

- Initialization: Regard all the points in the training data as members of an initial cluster
- Repeat the following steps *l* times
  - Selection: select two points from each cluster randomly, and regard them as new representative points
  - Assignment: assign each alert in the training data to the closest representative point (generation of cluster)
  - Updating: update every cluster's representative point with the average of its members
- 2^l representative points are obtained

Extracting Representative Points



Converted alerts with 7 features

Training Data

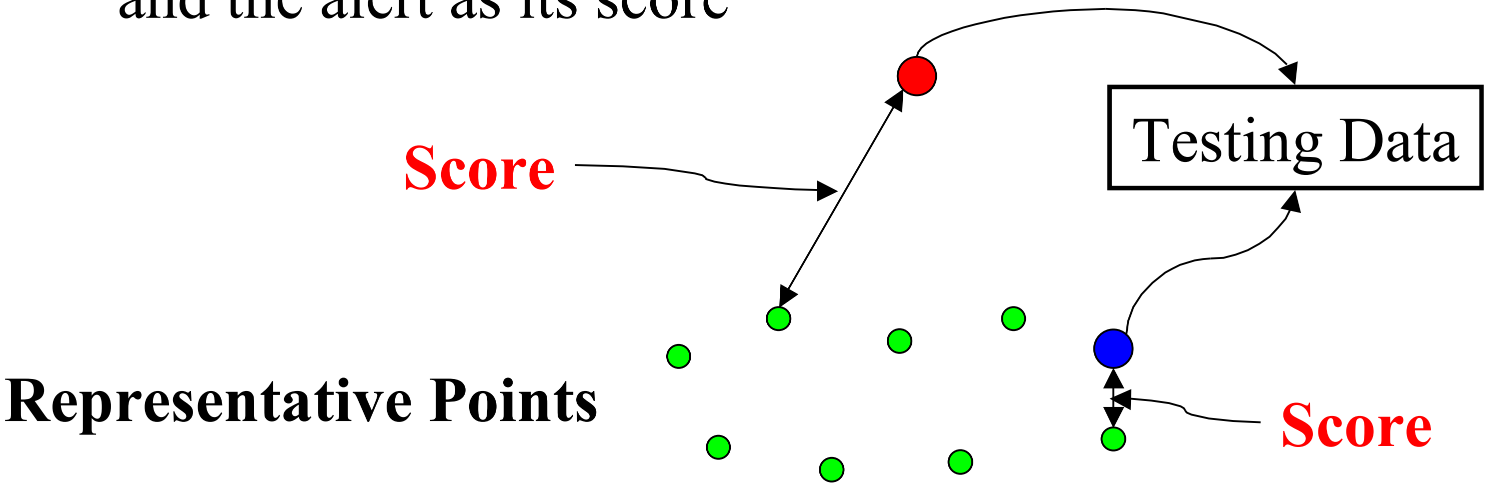
Clusters

Representative Points

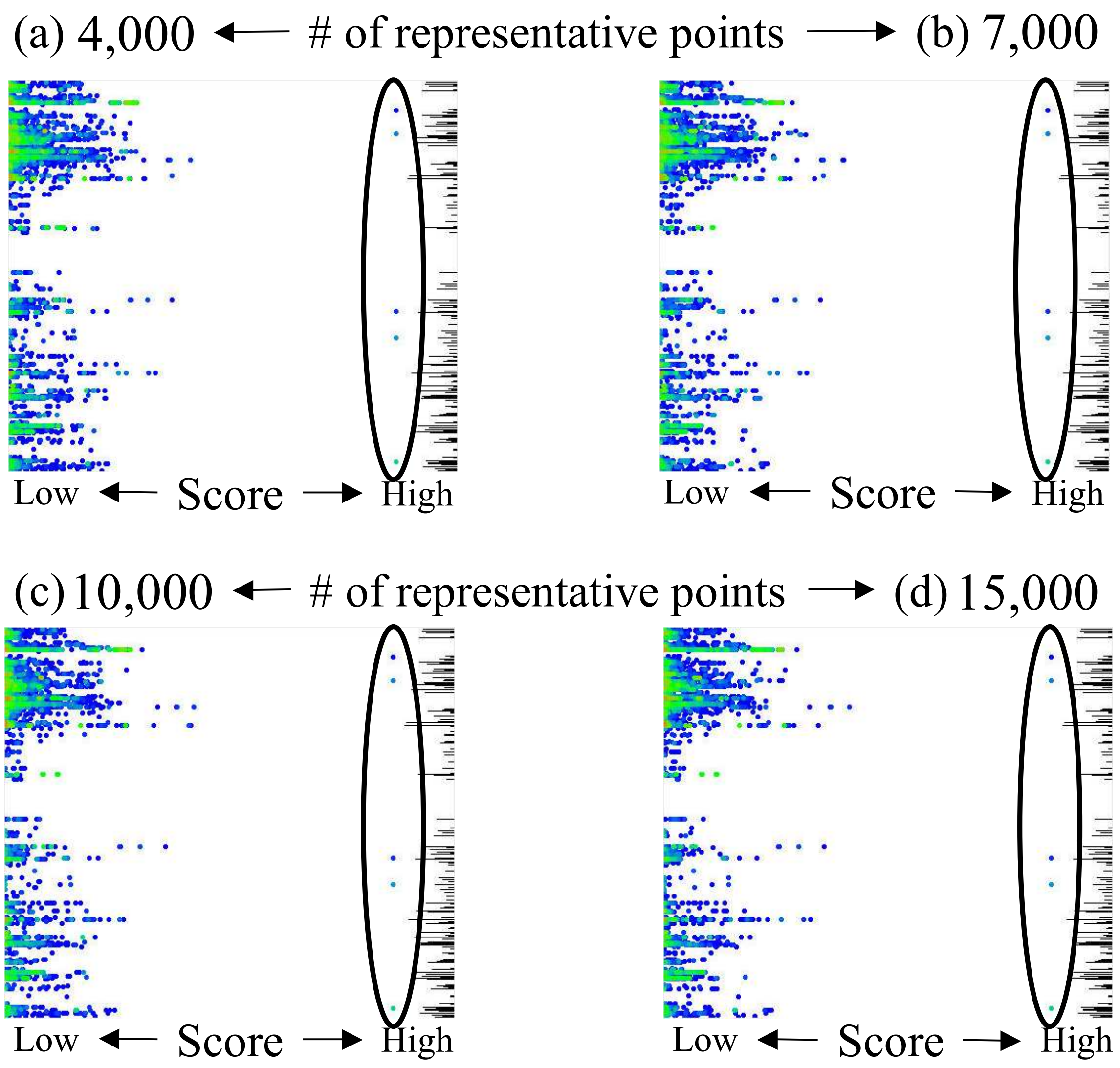
Usual false positives or known attacks

Scoring

- Assign a score to each alert of the testing data to reflect how anomalous it is
- Process
  1. measure the distance between all the representative points and the alerts of the testing data
  2. find out the closest representative point for each alert
  3. regard the distance between the closest representative point and the alert as its score



Visualization of Unknown Activities



(a) 4,000 ← # of representative points → (b) 7,000

(c) 10,000 ← # of representative points → (d) 15,000

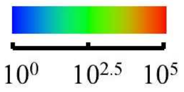
-.Training Data: August, 2006

-.Testing Data: November, 2006

-.Vertical axis: Signature IDs

-.Horizontal axis: Score

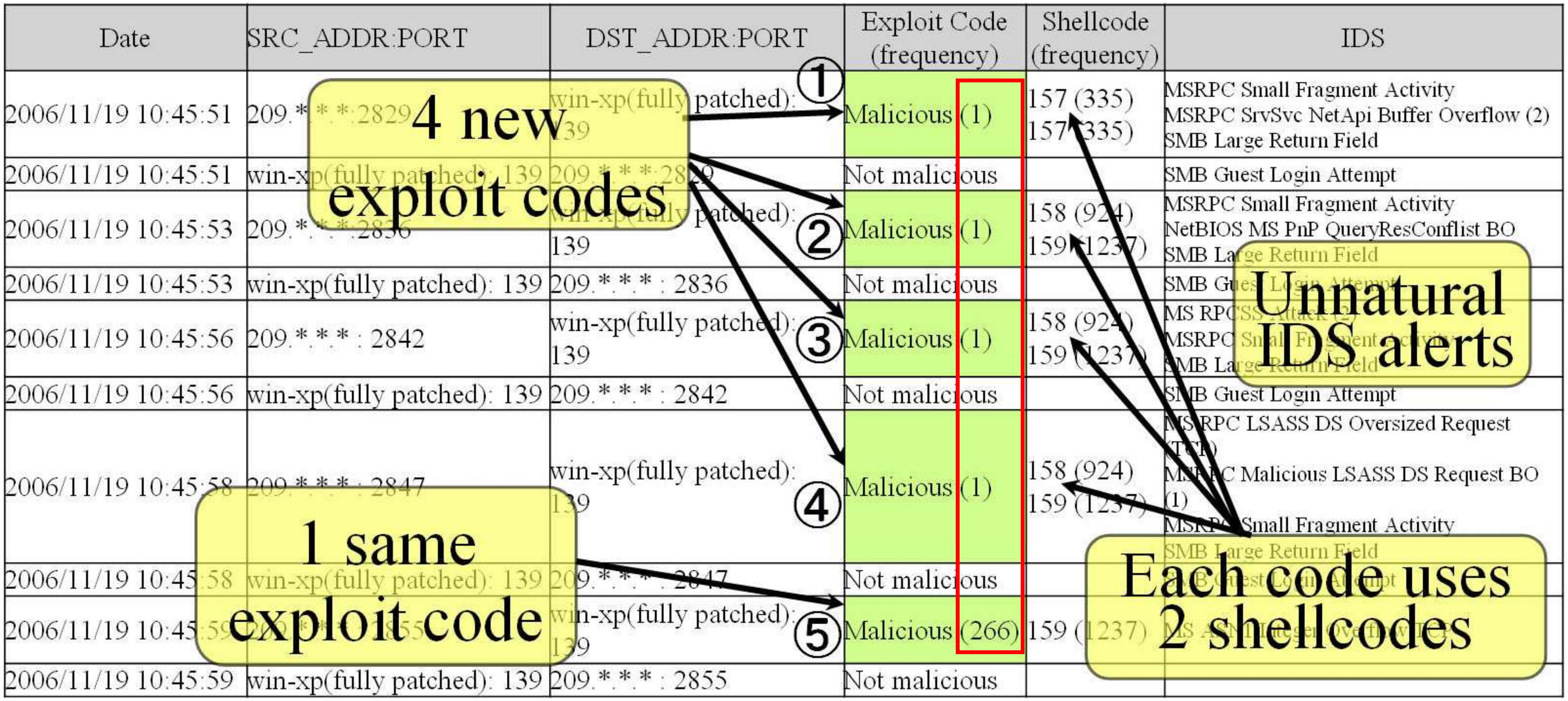
-.# of IDS alerts:



-.Scores of them are insensitive to the number of the representative points

-. Most right side points correspond to unknown activities with a high possibility

Example of Unknown Activities



Date	SRC_ADDR:PORT	DST_ADDR:PORT	Exploit Code (frequency)	Shellcode (frequency)	IDS
2006/11/19 10:45:51	209.*.*.*:2829	win-xp(fully patched): 139	Malicious (1)	157 (335) 157 (335)	MSRPC Small Fragment Activity MSRPC SrvSvc NetApi Buffer Overflow (2) SMB Large Return Field
2006/11/19 10:45:51	win-xp(fully patched): 139	209.*.*.*:2829	Not malicious		SMB Guest Login Attempt
2006/11/19 10:45:53	209.*.*.*:2836	win-xp(fully patched): 139	Malicious (1)	158 (924) 159 (1237)	MSRPC Small Fragment Activity NetBIOS MS PnP QueryResConflist BO SMB Large Return Field
2006/11/19 10:45:53	win-xp(fully patched): 139	209.*.*.* : 2836	Not malicious		SMB Guest Login Attempt
2006/11/19 10:45:56	209.*.*.* : 2842	win-xp(fully patched): 139	Malicious (1)	158 (924) 159 (1237)	MS RPCSS Attack (2) MSRPC Small Fragment Activity SMB Large Return Field
2006/11/19 10:45:56	win-xp(fully patched): 139	209.*.*.* : 2842	Not malicious		SMB Guest Login Attempt
2006/11/19 10:45:58	209.*.*.* : 2847	win-xp(fully patched): 139	Malicious (1)	158 (924) 159 (1237)	MSRPC LSASS DS Oversized Request (TCP) MSRPC Malicious LSASS DS Request BO (1) MSRPC Small Fragment Activity SMB Large Return Field
2006/11/19 10:45:58	win-xp(fully patched): 139	209.*.*.* : 2847	Not malicious		SMB Guest Login Attempt
2006/11/19 10:45:59	209.*.*.* : 2855	win-xp(fully patched): 139	Malicious (266)	159 (1237)	MS ASN1 Integer Overflow (TCP)
2006/11/19 10:45:59	win-xp(fully patched): 139	209.*.*.* : 2855	Not malicious		

4 new exploit codes

1 same exploit code

Unnatural IDS alerts

Each code uses 2 shellcodes

-. Attacker used 4 new exploit codes and 1 same exploit code

-. First 4 exploit codes used 2 shellcodes, and IDS triggered 3 or 4 different alerts

-. These combinations of IDS alerts are unnatural

-. Attacker is developing his shellcodes that are combined by the existing shellcodes

-. These activities were caused by Allaple worms

Summary and Future Works

- Method to extract unknown activities from IDS alerts
  - Example and Visualization of extracted unknown activity

- Future works
  - Universal Feature Construction Method
    - Not all vendors provide Incident ID
    - Building mechanism is different from each other
    - Additional features using only “Basic features”
    - Basic features
      - source address and port, destination address and port, and detection time..
  - Detection of unknown activities which do not raise any alert

Thank you for your attention!