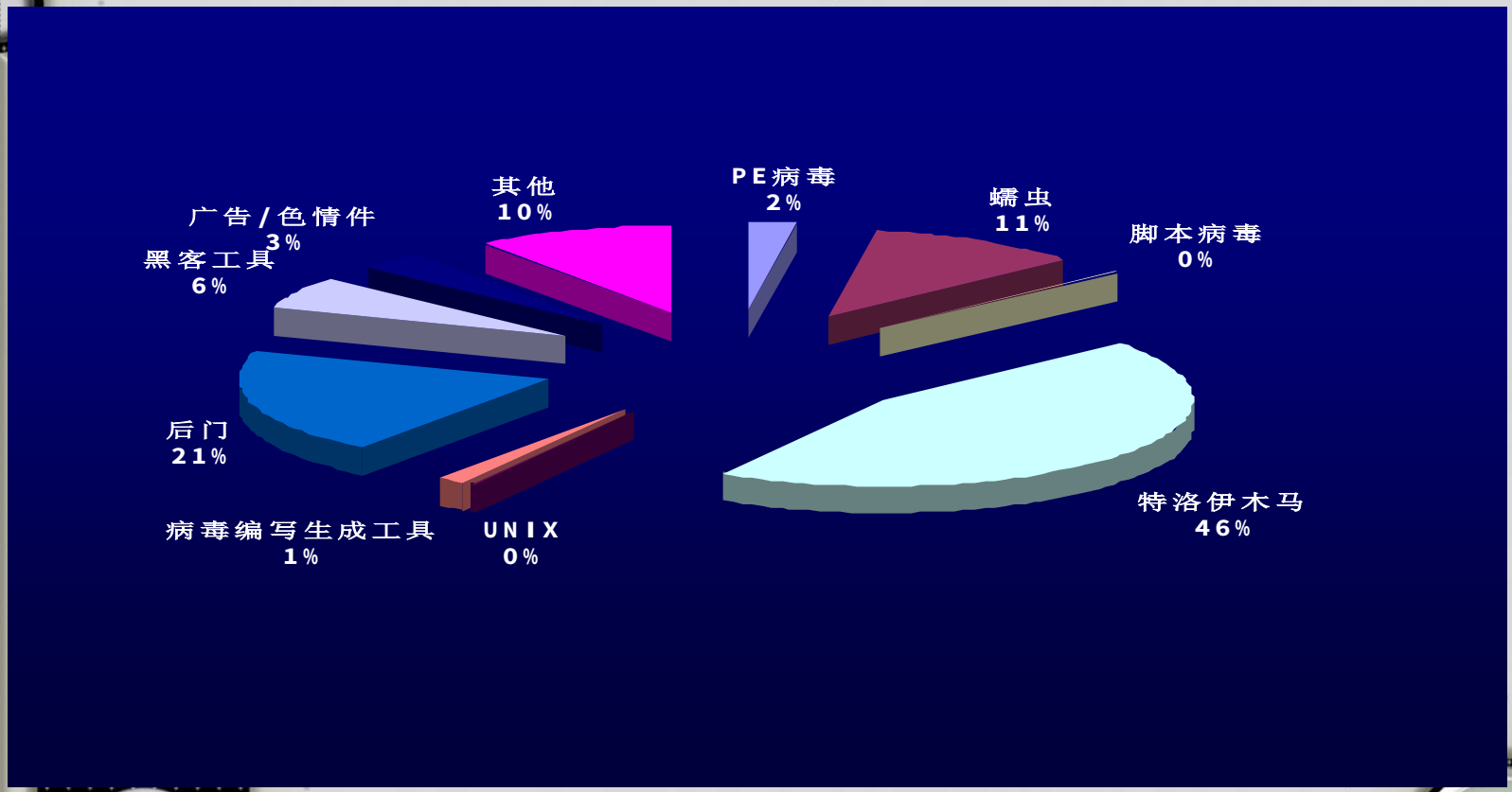
Virus Detection System
VDS

seak@antiy.net

Outline

- **The pop trend of virus in 2004**
- **Quality of the IDS**
- **Mechanism of the VDS**
- **Data processing**

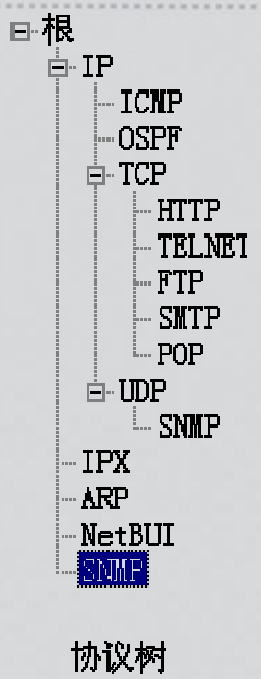
20047 kinds of new virus in 2004



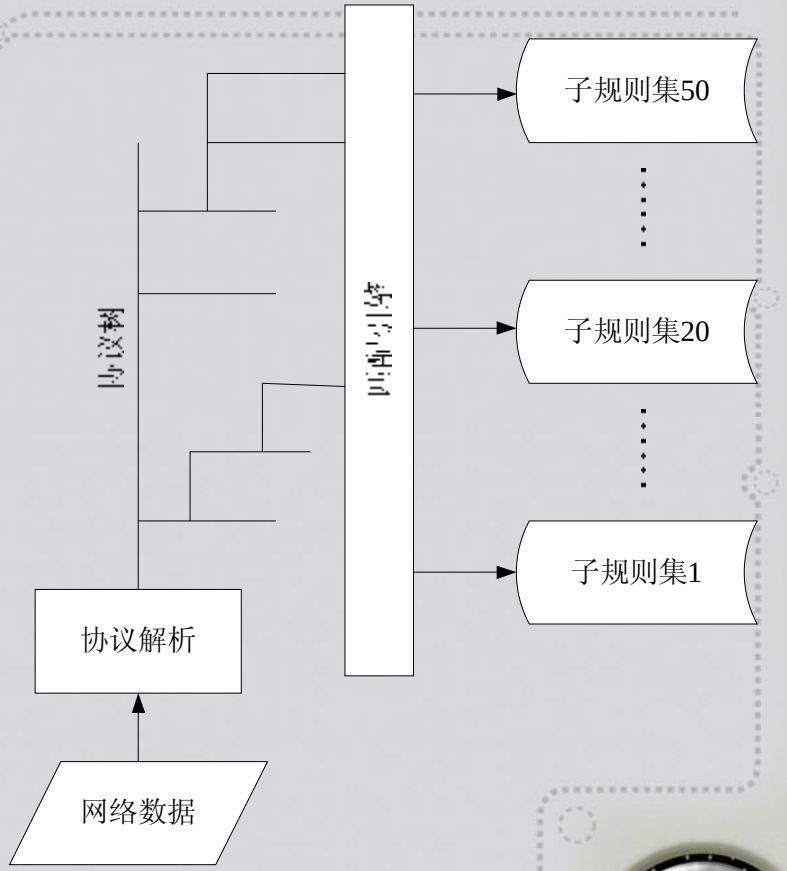
Outline

- The pop trend of virus in 2004
- Quality of the IDS
- Mechanism of the VDS
- Data processing

How traditional IDS works



协议树



- Accurate protocol resolution
- Small rule sets，short characteristics matching
- No more than 500 rules in one rule set

Confronted with virus, IDS retreating?

- Year 2000，in sum 10350 viruses came out，backdoor 1029。
- Snort x.x.x
- 05/21/2001
- Backdoor.rules 127 rules
- Virus.rules 87 rules，targeting at mail worms，detecting mail attachment name、extended name、topic
- Year 2004，in sum 20047 virues，backdoor 4010。
- Snort 2.3.3
- 03/01/2005
- Backdoor.rules，82 rules
- Virus.rules 1 rule，attachment extended name detecting

Unified software designing

- **Unified design: In case of dealing with the extensively complicated events, we should classify the events and unify one or more of the processing modules by using expandable data structure and data set.**
- **AV Ware: Target objects' diffluence.**
- **IDS: Protocol's diffluence.**

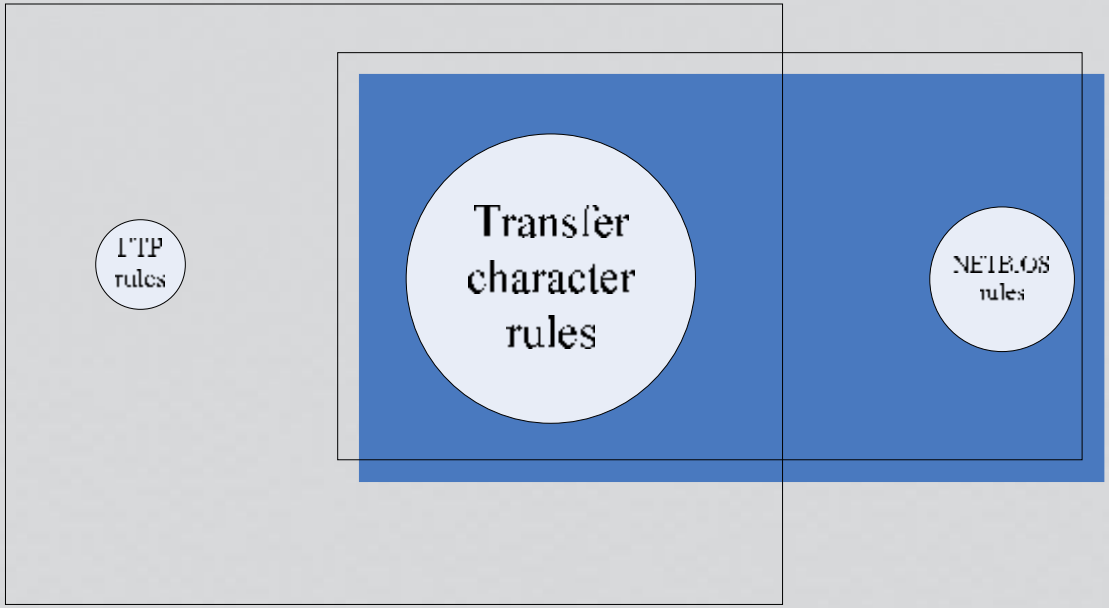
AVML and Snort

- Echo virus(id=″B00801″;type=″Backdoor″;os=″Win32″;format=″pe″;name=″bo″;version=″a″;size=″124928″;Port_listen=on[31337];content=|81EC0805000083BC240C05000000535657557D148B842424050000008BAC242005000050E9950500000F85800500008B|;delmark=1)

- alert tcp $EXTERNAL_NET any -> $HOME_NET 21 (msg:"Backdoor.bo.a Upload"; content: |81EC0805000083BC240C05000000535657557D148B842424050000008BAC242005000050E9950500000F85800500008B |;)

- alert tcp $EXTERNAL_NET any -> $HOME_NET 139 (msg:"Backdoor.bo.a Copy"; content: |81EC0805000083BC240C05000000535657557D148B842424050000008BAC242005000050E9950500000F85800500008B |;)。

Redundant scan for diffluence

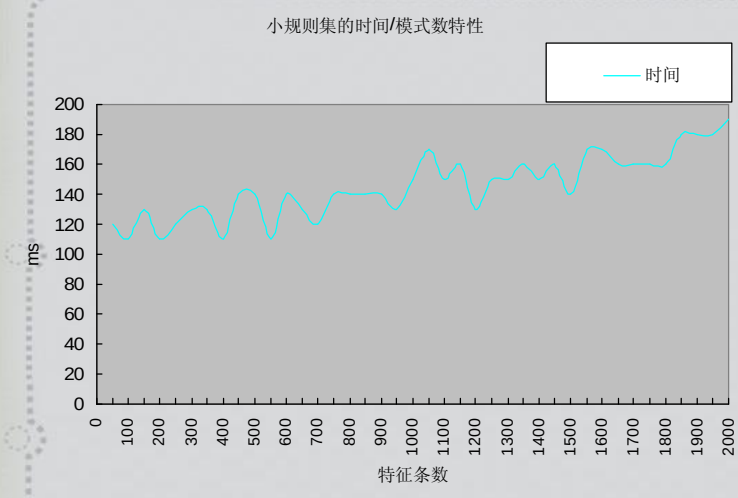


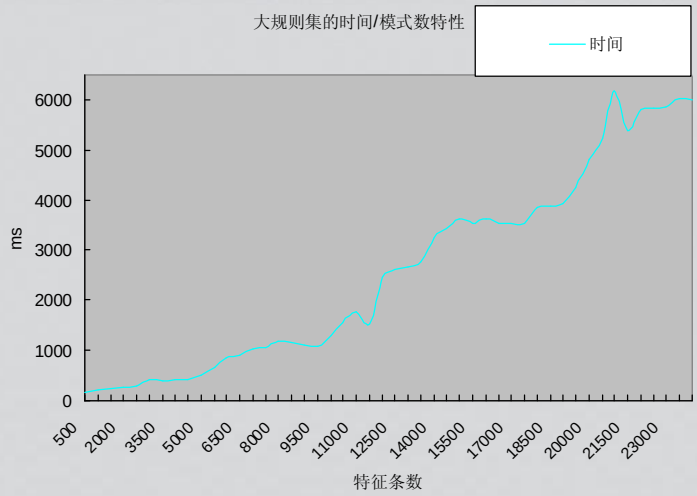
Pressure of the rule set scale

type	quantity
Email worm	2807
IM-worm	172
P2P-worm	1007
IRC-worm	715
Other worm	675
total	5376

- **Besides worm, there are over 20,000 kinds as the Trojan, Backdoor, etc… related to the network.**
- **The corresponding rule set may exceed 30,000 records.**

The pressure of efficiency





Test by snort of the latest version, good efficiency by small set VS dramatically downfall by large set

Efficiency pressure brought by the increasing rule sets, is the fundamental pressure for IDS when adopting the anti-virus mechanism

The network prospective, detecting level, no small granular virus locating of IDS, co[illegible]t the reason why it fails to assume such pressure

Outline

- The pop trend of virus in 2004
- Quality of the IDS
- Mechanism of the VDS
- Data processing

What is the crux?

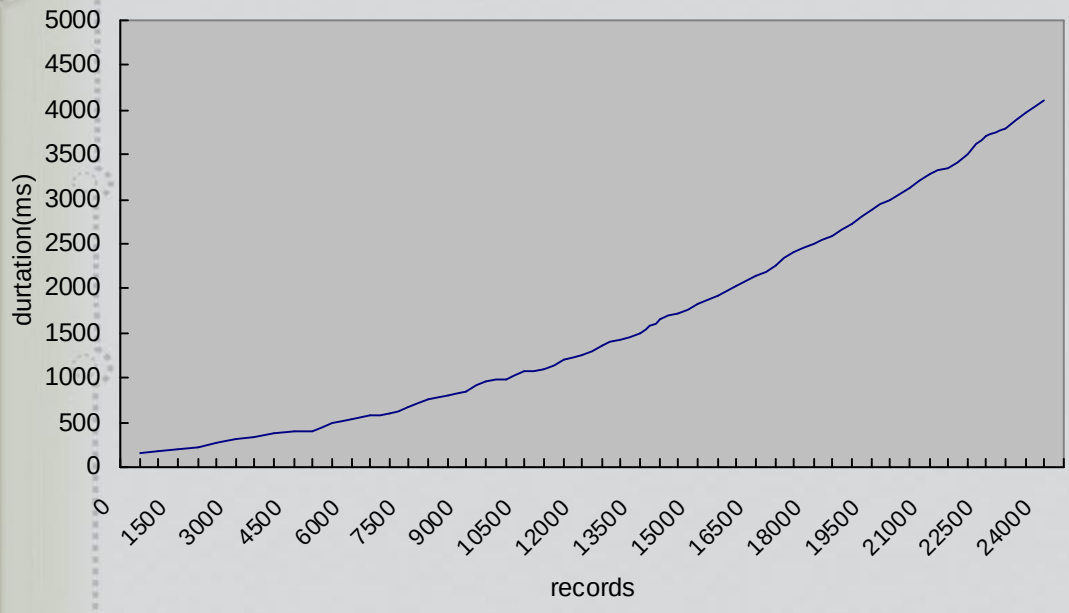
- **The efficiency pressure is the major pressure in network virus detecting**
- **The new unification model focuses on matching speed and granularity, its construction is algorithm optimization oriented.**

How to unifiy

- **The network flow falls into three categories according to its content**

category	example：
Direct matching	Nomal scanning ,attacking, transfering
Preprocessing Required	URL (case insensitive) MAIL (coding)
Specific Algorithm Required	script

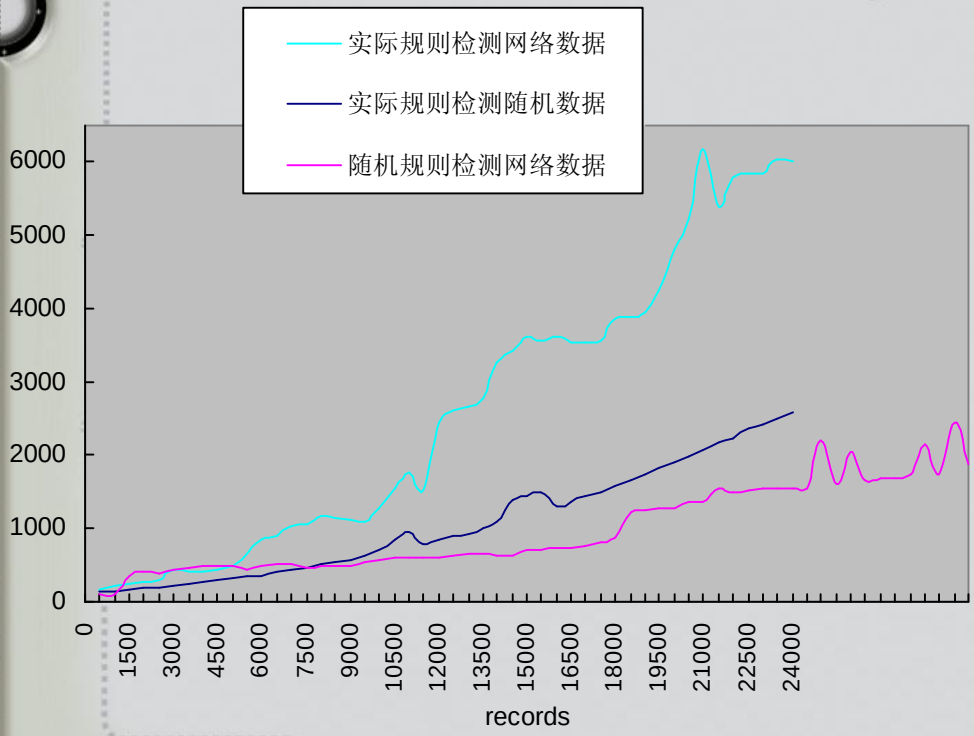
Algorithm optimization（1）



In a situation that the quality of rules is smaller than 6,000, it is not obvious that a linearity counted of time and record increases. But about 10,000 records , it begin to present reverse rising , cause the sudden drop of performance , until it is not available。

- **The influence of time matching by changing the quantity of records**

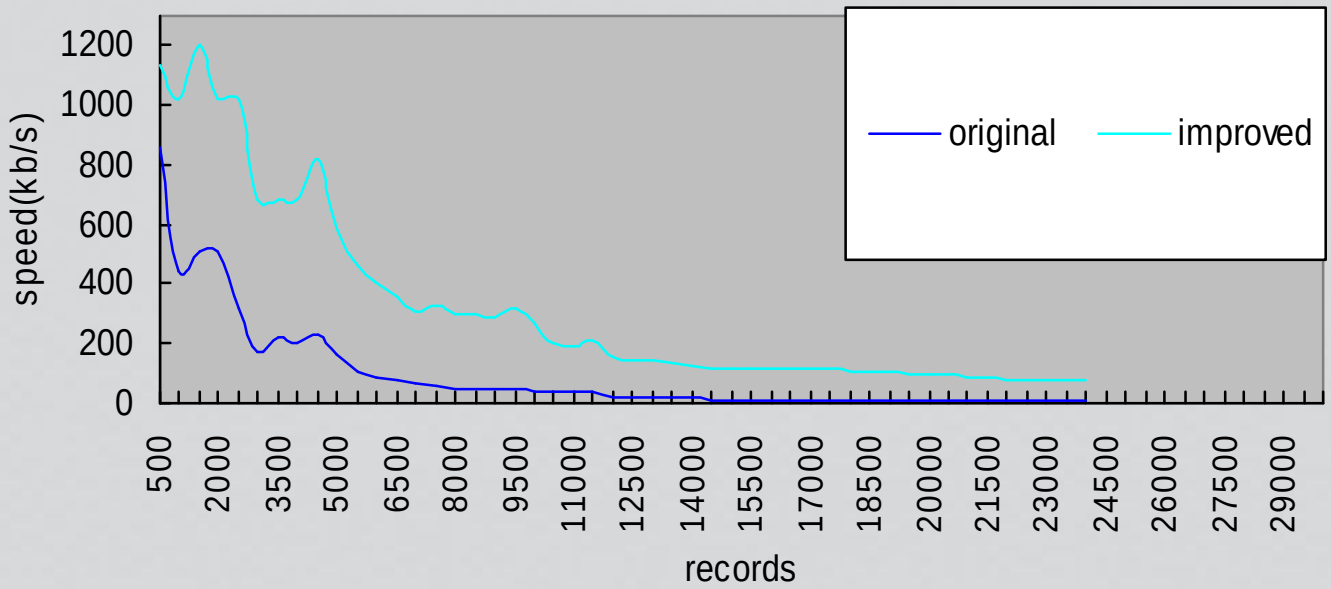
Algorithm optimization （2）



- **Scan speed is also related to the matching data and quality of patterns. Because of the approximation between the virus characteristic, can not present the characteristic of the random distribution. And so does the network data. So they all make effect to the matching situation.**

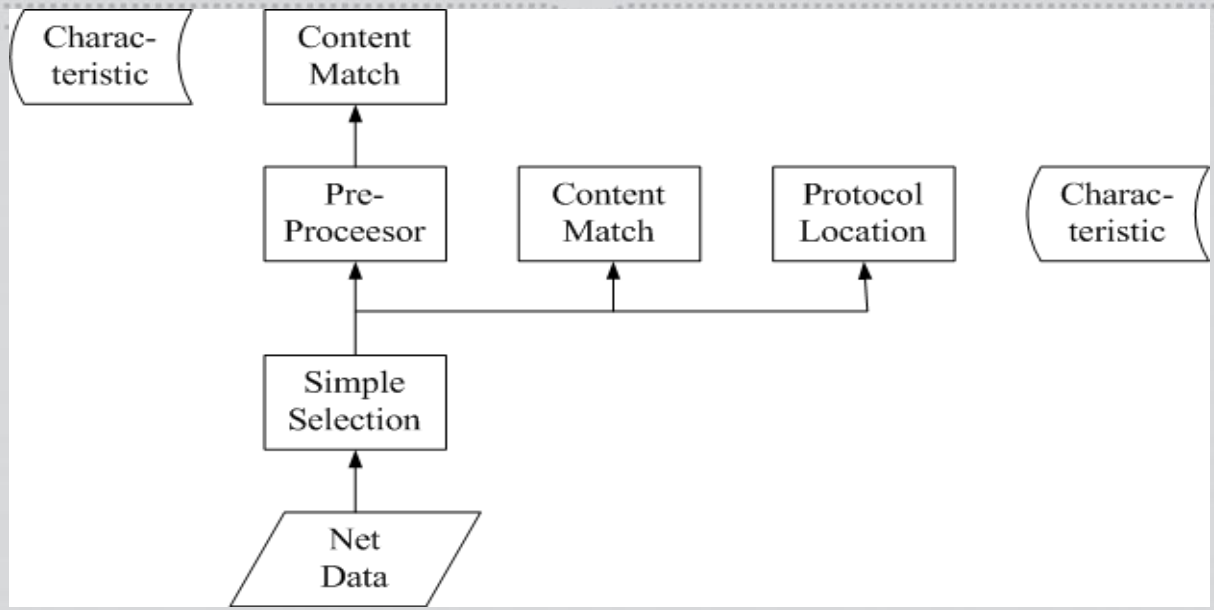
- Scan methods and object's influence on the data

Algorithm optimization （3）



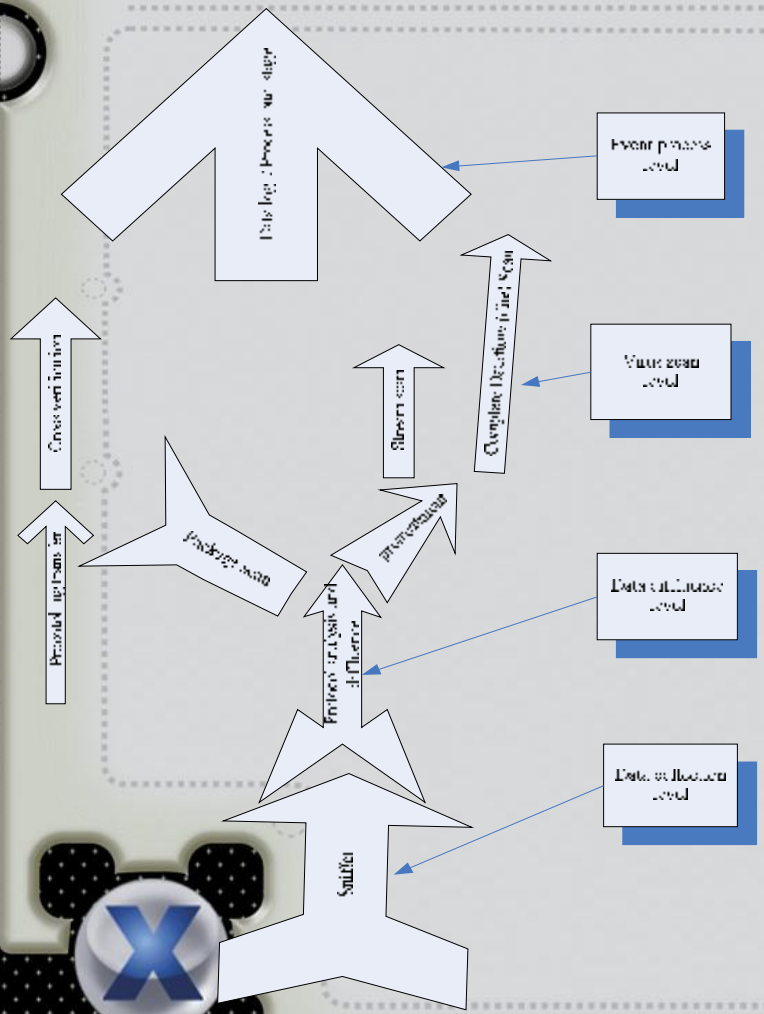
- The influence of efficiency by limit the approximation of the virus's characterist

Architecture of VDS



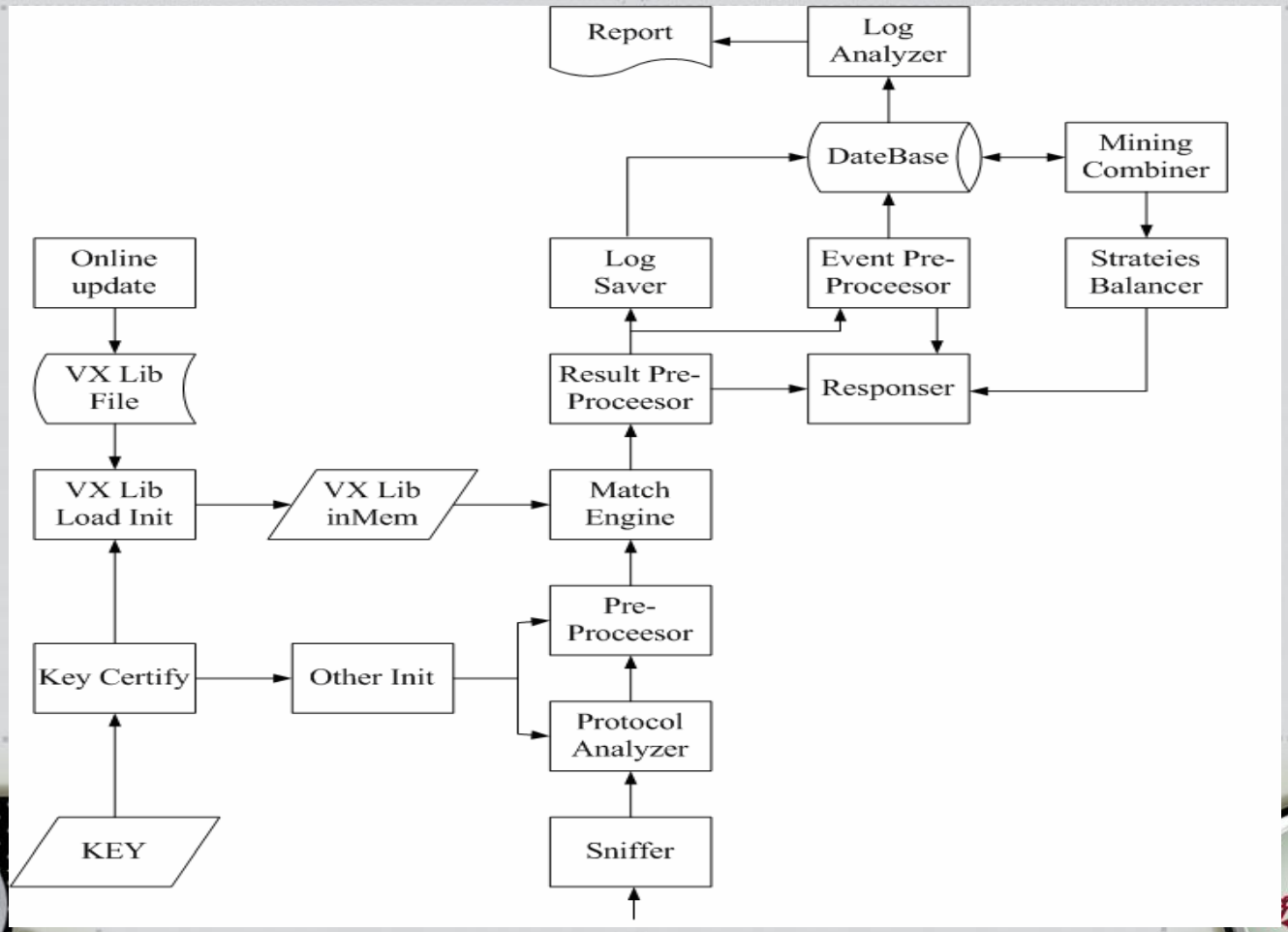
- **Unitary model is regarding the match speed and the granularity of matching — matching is the foremost.**
- **Classifying network traffic data into three types：data matched on binary level，data needing pre-teat and data needing algorithm specified.**

Dataflow direction and the Level of virus detection

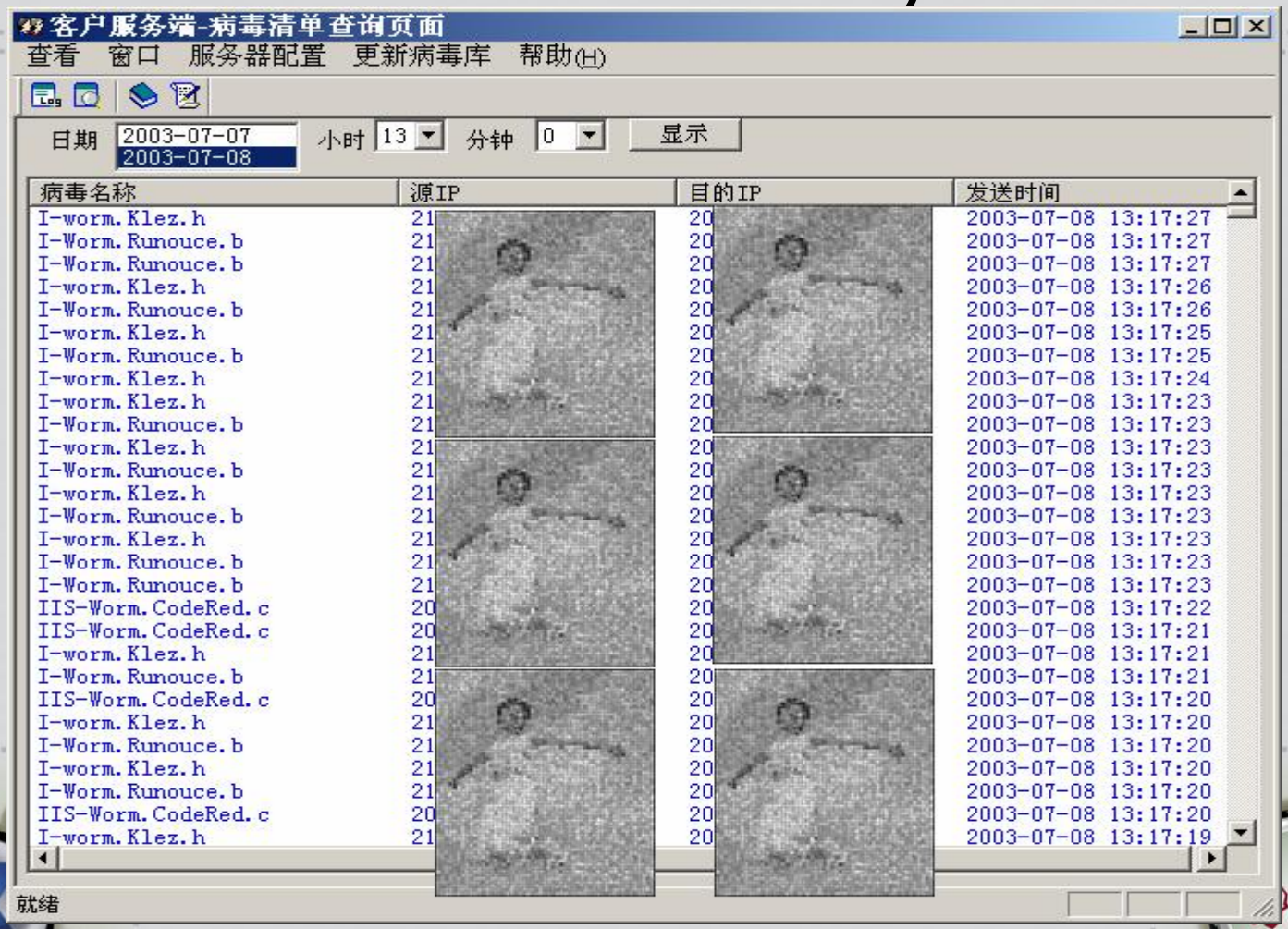


- **Divided into 4 levels: collection、diffluence、detection and process**
- **Provide package scan、incomplete data scan And complete data scan.**

System structure



Data efficiency



Virus data output from Harbin Institute of Technology on July 8 , 2003.

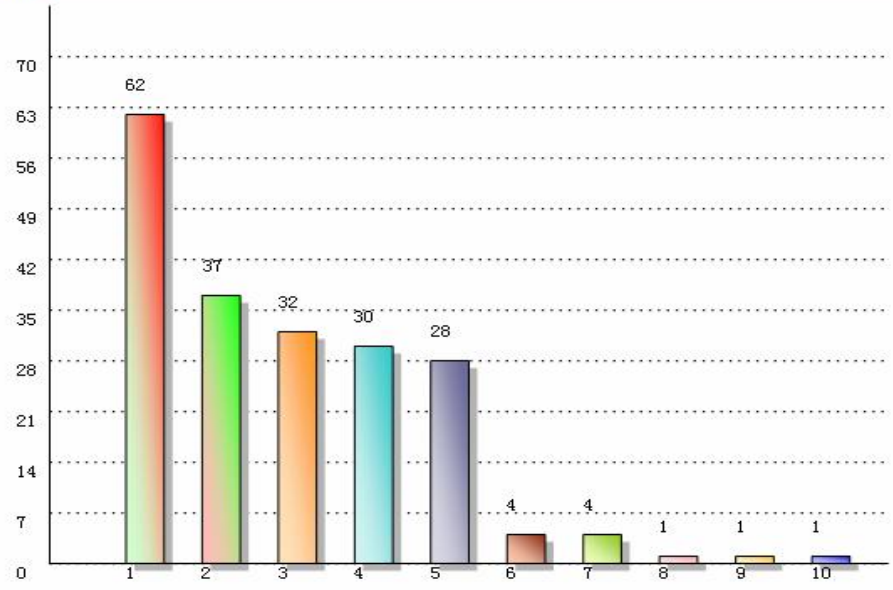
Statistics of the 26th week in 2005

2005年26周邮件蠕虫监测结果统计报告

检出次数排行榜

名次	病毒名称	进入内网比例	检出次数	病毒流量(byte)	感染列表
1	Email-Worm.Win32.Bagle.af	100%	62	0	受感染主机 受攻击主机
2	Email-Worm.Win32.LovGate.ad	100%	37	0	受感染主机 受攻击主机
3	Email-Worm.Win32.LovGate.ae	0%	32	0	受感染主机 受攻击主机
4	Email-Worm.Win32.LovGate.w	100%	30	0	受感染主机 受攻击主机
5	Email-Worm.Win32.LovGate.w	0%	28	0	受感染主机 受攻击主机
6	Email-Worm.Win32.NetSky.c	100%	4	0	受感染主机 受攻击主机
7	Email-Worm.Win32.LovGate.q	100%	4	0	受感染主机 受攻击主机
8	Email-Worm.Win32.Zafi.d	100%	1	0	受感染主机 受攻击主机
9	Email-Worm.Win32.Bagle.af	0%	1	0	受感染主机 受攻击主机
10	Email-Worm.Win32.NetSky.z	100%	1	0	受感染主机 受攻击主机
		总计	200	0	

检出次数统计图



Unknown virus forewarning system

发现病毒体传输次数排行榜:

名次	病毒名	发现次数
1	I-worm.Klez.h	42217
2	I-Worm.UNKnow	2548
3	TrojanDropper.Win32.Small.j	4
4	I-Worm.Nimda	2
5	Backdoor.Netbus.160.a	1
6	Trojan.Win32.HDBreaker	1

- **Detect a unknown worm （I-Worm.Unknow ）increasing notablely on June 5, 2003. and on June 6 it was proved to be the virus:I-worm.sobig.f.**

Outline

- **The pop trend of virus in 2004**
- **Quality of the IDS**
- **Mechanism of the VDS**
- **Data processing**

VDS related Researching status

- **Network worm detecting based on GrIDS**
- **Detecting methods based on PLD hardware**
- **Detecting based on HoneyPot**
- **Worm VS Worm**
- **Without detecting the known virus, all worms are unknown**
- **VDS brings in engineering methods, enables precise location of network virus event**

Event Processing (1)

- DEDL，Detection Events Description Language.
- By using the symbol description mode, defined the network event into a format criterion, and support the ability of common condition- deriving.
- Defined elements: event type、event ID、source IP、target IP、event time、 such more than 20 key elements.

- Processing methods
- Tech-based Internal combine
- Parallel-type combine
- Analysis-based Parallel combine
- Radiant-type combine
- Convergence-type combine
- Chain-type combine

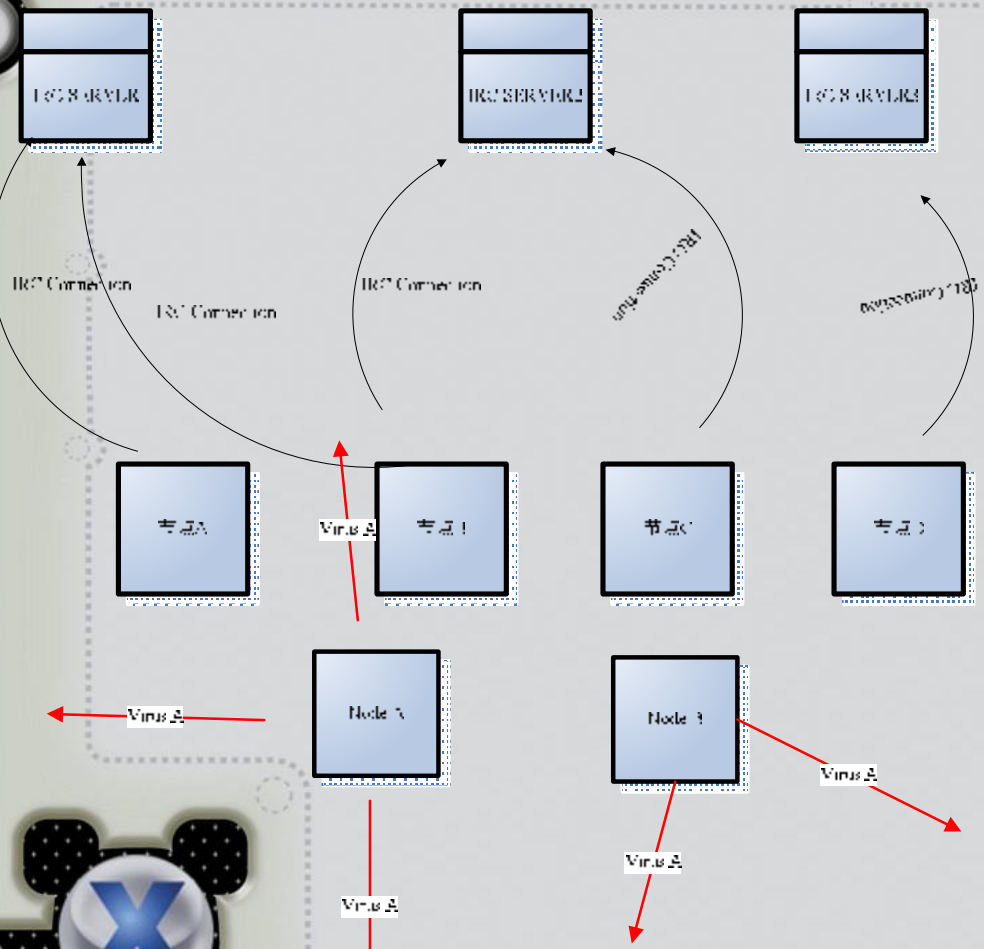
Event Processing (2)

If existNet_Action(RPC_Exploit)[IP(1)->IP(2);time(1)]
Net_Action(RPC_Exploit) [IP(2)->IP(3) ;time(2)]
and
time(2)>time(1)
than
Net_Action(RPC_Exploit) [IP(1)-> IP(2) -> IP(3)]

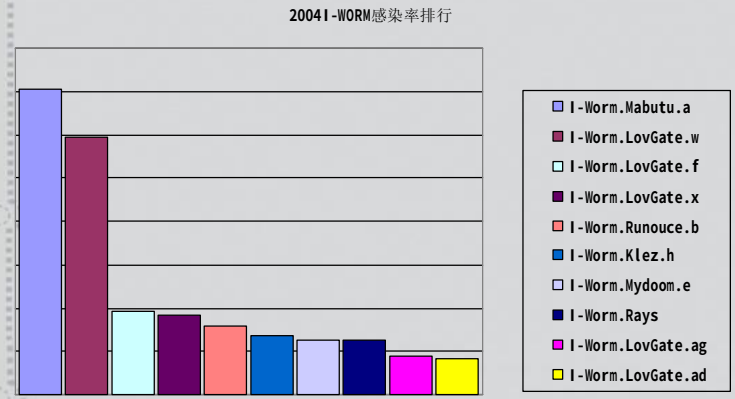
Behavior Classify

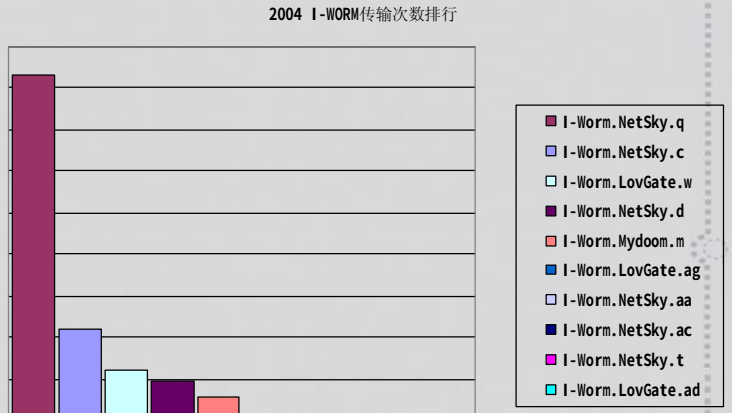
DEDL events	AVML regulations about diagnostic behavior
Net_Action(act)[IP(1),IP(2):445; ;time(1)] Net_Action(act)[IP(1),IP(3):445; ;time(1)] .... Net_Action(act)[IP(1),IP(12):445; ;time(1)] Net_Action(Trans,Worm.Win32.Dvldr)[IP(1)->IP(12);time(1)]	Virus_act_lib Virus seek(id=″W02872″;dport=139,445;trans=netbios)

Data processing mining



The meaning of data analysis





Vague relationship between infection times and number of infected nodes

Infection is most efficient via trusted chain

Counterstrike by data: virus to virus

Reflections

- **Monitor the network virus has been explored academically and productively, it has extended to be a new technology in its direction.**
- **The way made up of attack and defense is from the world of certain to the world of freedom. — we are on the road.**

Thank You !