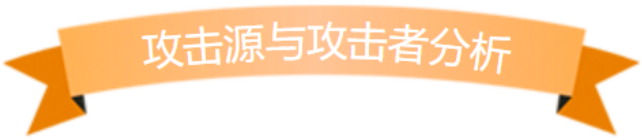
python网络攻击总参三部_GitHub - BLKStone/datacon: datacon比赛方向三-攻击源与攻击者分析writeup...

weixin_39621870 于 2020-11-23 12:42:29 发布 79 收藏 1
文章标签： python网络攻击总参三部
DataCon-方向三-攻击源与攻击者分析-writeup

BlueThanos战队



排名	团队名称	单位名称	得分
4	安全数据3组	平安科技（深圳）...	87分
5	Satori	浙江大学	84分
6	大王有条小毛驴	武汉大学/成都信息...	83分
7	ART-ML	ART	79分
8	dataFLY000	信息工程研究所	77分
9	独角兽	无	74分
10	p@433	中国平安财产保险...	70分

0x01比赛要求

本题设置了多个维度的网络行为数据，涉及到不同类不同维度的数据源，包含web告警信息，ip基础信息，域名信息，whois信息，日常访问行为信息，终端行为信息等。考察选手如何通过多维度的数据源体系化的描绘一个攻击者，设计并建立一套分析方法，综合各维度数据对攻击者进行分析，描绘出可能对大会威胁最大的攻击者。

识别出攻击IP

建立一种分析方法与系统大致确定IP与人的关系。

提交分析方法设计文档(pdf)，需包含完整的处理流程图和描述。并提交实现的系统源码，要求可复现。以复现计算生成的结果为准。

建立一套分析方法与系统，从攻击目的和攻击能力层面对攻击者进行分析。

提交分析方法设计文档(pdf)，算法设计原理，模型构建。并提交实现的系统源码，要求可复现。以复现计算生成的结果为准

0x02整体思路

对日志和拓展数据预处理

由于数据量操作时间较多，所以数据聚合与处理用pandas进行，大大加快了效率。

通过对日志数据的预处理，得到IP侧相关数据和domain侧相关数据，方便协同调取使用分析。

通过正则（主要）和机器学习处理。

数据处理类型

数据处理中间结果及逻辑

攻击矩阵

识别的标签

识别的操作类型

识别的攻击类型

机器学习方法使用

制定IP关联规则

主要8个规则

建立模型进行IP聚类

设计的算法依照超市选货的相关方式，对IP相关性聚类成攻击者

在日志数据预处理结果和进行完IP攻击聚类的基础上，进行攻击者人物画像

常规信息

目的分析

能力分析

主要分为1)攻击的广度与深度、2)攻击复杂性分析、3)漏洞利用能力、4)攻破防护能力、5)反溯源能力。 前两个是定性分析，后三个可定量分析。

构建模型，量化能力等级

0x03系统设计

1.日志数据预处理

1)数据处理类型

定义了IP行为块的概念，识别操作类型、脚本类型、agent类型、标签、攻击类型。

单目标IP行为块：在持续时间段内，IP对单个目标进行的一系列的攻击操作。

IP行为块：在持续时间段内，IP进行的一系列的攻击操作。主要分类4个类型：单网站的漏洞扫描、同漏洞批量扫描、web渗透、高级web渗透。

操作源语识别：针对webshell控制中进行的统一化。PS：未实现（时间不够）

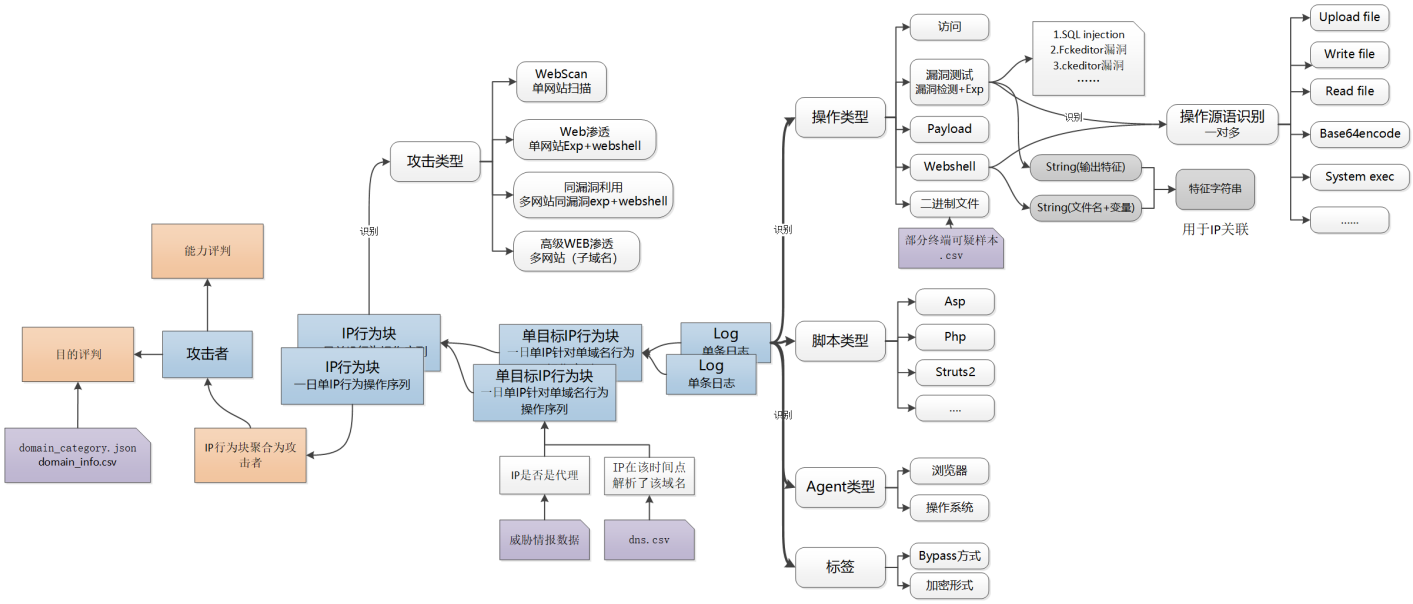
标注：

蓝色块：数据模型

橙色块：达到的目标

紫色块：外部数据源

灰色块：拓展数据



2）数据处理中间结果及逻辑

分攻击侧和目标服务器侧两类中间数据集，支撑后期分析使用。

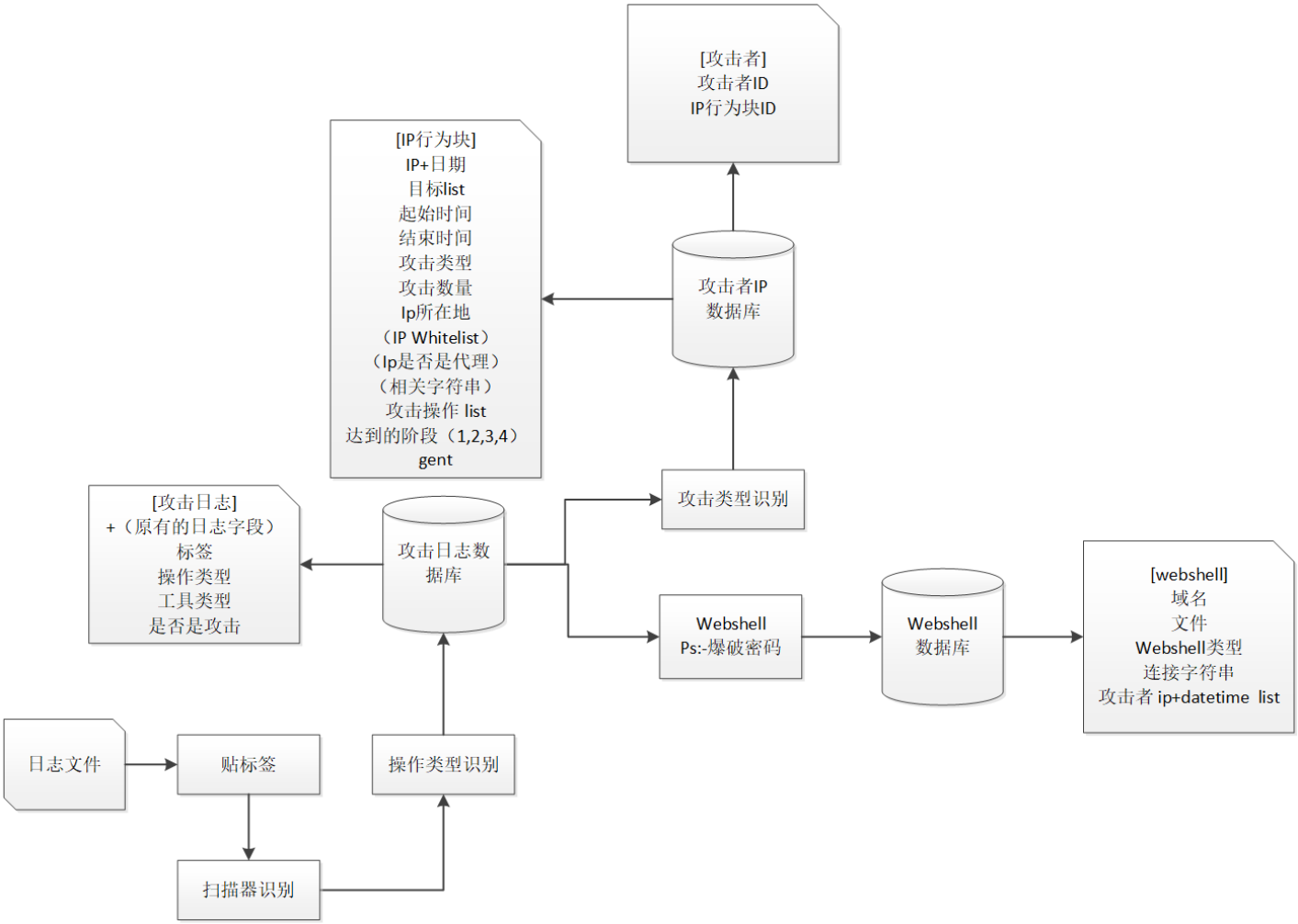
主要数据处理为生成聚类和分析能力的逻辑。进行贴标签，扫描器识别，操作类型识别（详情见3攻击矩阵），webshell识别归类，攻击类型识别，以及相关数据统计。

中间数据集说明：

攻击日志中间结果为基础数据

webshell中间结果主要用于关联攻击者，描述攻击目标情况

IP行为块中间结果主要用于攻击者目的和行为分析



3）攻击矩阵

攻击矩阵中的技术内容就是本文中的操作类型。

借鉴att&ck模型，依照本数据情况进行定义生命周期如下图。

由于漏洞检测和漏洞利用测试仅通过当前数据难以区分归为一类。

阶段3的定义目的主要在于寻求漏洞测试和持久化的桥梁，以此构建成攻击链。

web攻击矩阵				
非攻击行为（non-attack）1	漏洞测试（expliot）2	漏洞利用（payload）3	持久化（webshell）4	后渗透（bin）
访问	SQL注入漏洞	上传webshell asp	chopper	
爬虫	iis6.0解析漏洞	上传webshell php	antsword	
	struts2漏洞远程执行	上传webshell jsp	weevely	
	XXE漏洞	上传黑页	other	
	XSS漏洞	提示漏洞字符串（安全监测）		
	文件包含漏洞			
	代码托管配置信息导致源码泄露			
	备份文件导致源码泄露			
	异常文件访问			
	thnkphp5路由漏洞			
	thinkphp3.2漏洞			
	thinkphp漏洞远程执行			
	一句话webshell爆破			
	fckeditor绕过限制上传漏洞			
	ewebeditor绕过限制上传漏洞			
	其他类型绕过限制上传漏洞			
	查询工具代码执行漏洞			
	CGI MOD命令执行漏洞			
	dedecms-download文件漏洞			

操作类型：

sql_injection,SQL注入漏洞利用

vul_asp_resolve,iis6.0解析漏洞

vul_struts2_rce,struts2漏洞远程执行

vul_xss,XSS漏洞利用

vul_include_fie,文件包含漏洞

vul_code_leak,代码托管配置信息导致源码泄露

vul_dede_plus_download,dedecms-download文件漏洞远程执行

vul_thinkphp_5_route_rce,thnkphp5路由漏洞远程执行

vul_thinkphp_3.2_rce,thinkphp3.2漏洞远程执行

vul_thinkphp_rce,thinkphp漏洞远程执行

vul_backup_rar,备份文件导致源码泄露

vul_search_tool_rce,查询工具代码执行漏洞

vul_fck_upload,fckeditor绕过限制上传

vul_eweb_upload,ewebeditor绕过限制上传

...

4）识别的标签

标签的识别主要是通过正则匹配实现的，webshell的识别，可以运用机器学习的方法，方法是通过脚本采集chopper、蚁剑等webshell的http request，然后抽取特征作为训练集合，训练构建模型进行识别。

主要分为4大类标签：

attack类标签

共计80余条，属于攻击行为特征指纹，包括攻击类型、bypass方法等。

部分规则：

\'*(\s|\+|/*\.**/)*(or|and)(\s|\+|/*\.**/)+ vul_sql_injection NaN

(=|\s+|\++|\(|\)|')select(\s+|\++|\(|\)) sql_select NaN

...

info类标签

共计10余条拓展信息，但不具有攻击特征。

部分规则：

\<\?xml\s* xml

\

$$\exists a \in A\ b \in B\ c \in C, (k \text{ in } a) + (k \text{ in } b) + (k \text{ in } c) > 1$$

<

<<

<

<