御界高级威胁检测系统

技术白皮书

目录

第一章　系统简介

1. 前言

随着移动设备的普及和云基础设施的建设，企事业单位积极拥抱数字化，加之万物互联 IoT 潮流的到来，网络安全在当前这个时间点需要重新定义。经典的攻击方式还未落幕，层出不穷的高级攻击方式已经上演，在安全形势不断恶化的今天，即使部署了各式样的安全产品，也无法做到预防所有的威胁。

在网络边界处阻止所有的威胁固然是我们最希望看到的情况，但是这种同步的拦截方式受限于较高的性能要求和较少的信息量，很难独立成为一个完整的企业安全解决方案。御界高级威胁检测系统（以下简称御界）在网络边界处采用镜像流量，旁路检测的方式，旨在发现企业受到的恶意攻击和潜在威胁。

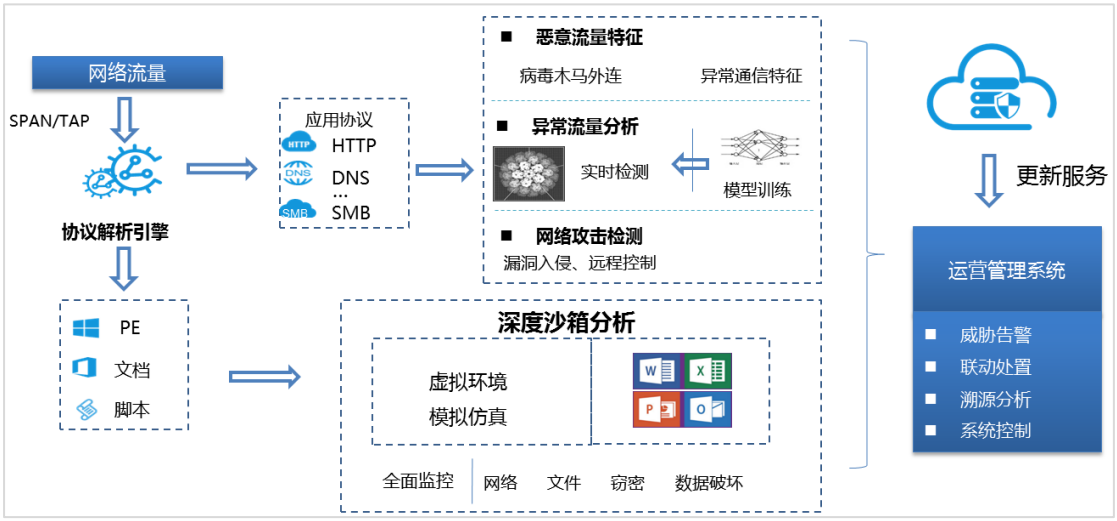
2. 核心能力

御界基于腾讯反病毒实验室的安全能力，依托腾讯在云和端的大数据，形成了强大且独特的威胁情报和恶意检测模型，凭借基于行为的防护和智能模型两大核心能力，高效检测未知威胁。也正因为丰富的数据和海量运算能力，在发现威胁之后的溯源上御界的数据平台也能提供一流的服务。

真正高质量的攻击，比如 APT 攻击，大多是以新面貌呈现，基于特征或者是黑样本库黑网址库的防御方式在对抗新的威胁之时远不如动态分析来的直接和有效。恶意代码层面可以通过变形和加密来做对抗，远控地址和远程 IP 也是全新的，攻击的入口也不尽相同，可以是邮件收发和消息处理等用户行为，也可以是通过系统或者软件漏洞进入，但是入侵，潜伏，远程连接，远程控制等过程从行为上看是有很大的相似和关联性的，基于行为的检测方案比基于特征的检测方案站在更高的维度上。智能模型也需要以行为传感器为基础，综合考量网络流量中的各种信息，协议，地址，端口，流量，连接频率等各种信息，来完成异常流量、异常行为的发现。

第二章　系统架构

1. 架构设计



御界高级威胁检测系统架构图

御界检测系统，以原始网络流量作为输入源，原始网络流量可以通过 SPAN 或者 TAP 方式输入。

经过 **TFA 协议解析引擎**分析(简称 **TFA 引擎**)，可以解析原始网络流量中的应用层协议，同时还原应用协议中包含的文件内容。

TFA 引擎支持多种协议的解析，包括 HTTP、SMTP、DNS、FTP、SMB，NFS，SSH，TLS 等， TFA 引擎文件还原能力，覆盖了可以产生威胁的任何文件类型，比如 PE 文件、脚本文件、压缩包、Office 文档等。

对于加密协议后者其他私有协议，TFA 引擎会解析流量的跨维特征，维度包括：时间维度（流量发生时间、流量持续时长等）、空间维度（流量节奏、数据包在流中分布等）这些跨维信息将会被推送到**算法模型模块**进行综合分析，智能鉴定。

对于可解析的协议，TFA 引擎会深度解析协议内容，解析结果将会根据协议类型分发到不同的检测模块进行鉴定，目前已有的检测模块包括 3 个：

恶意模型匹配模块通用的鉴定模块，针对全协议进行检测。恶意模型匹配模块内置多个检测模型，包括 APT 检测模型、病毒木马检测模型、数据泄漏检测模型，可以完成对 APT 攻击，病毒木马攻击，数据泄漏等多种异常流量的快速检测。

异常域名检测模块，针对恶意域名的特征进行离线学习，产生高启发的识别模型，识别模型用于对各种协议中的域名进行预测式的检测，在恶意域名生效前域名智能检测模块就已经可以进行预判。

网络攻击检测模块，是针对入侵攻击的检测模块，专业度高，识别稳定，内置 27 类攻击的检测规则。

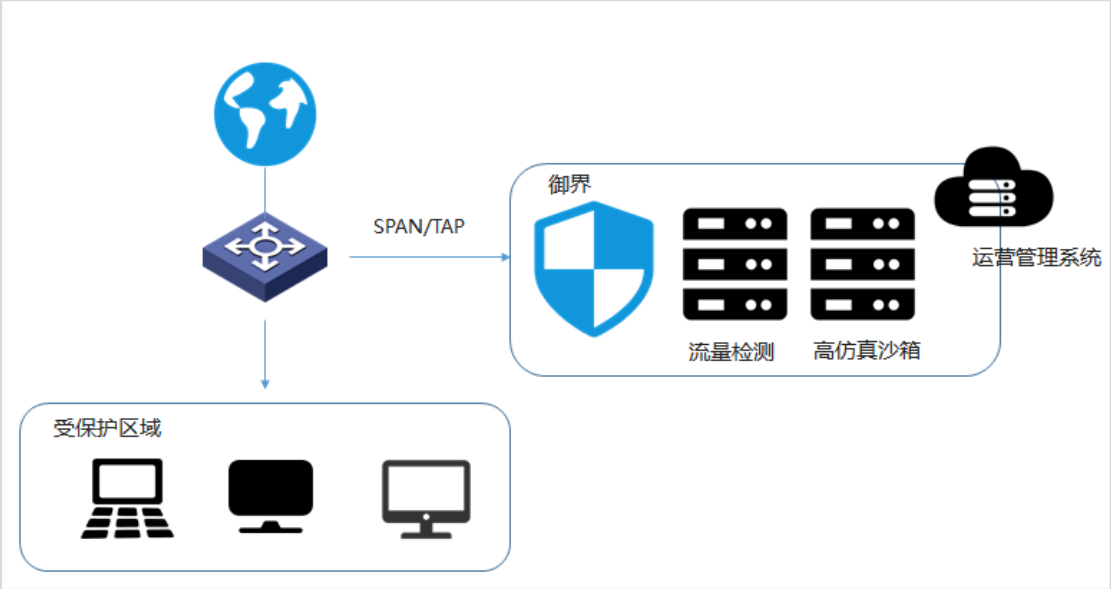
TFA 引擎在深度解析协议内容的同时，也会还原协议中包含的文件内容，还原的文件类型覆盖了可以产生威胁的任何文件类型，比如 PE 文件、脚本文件、压缩包、Office 文档等。这些对网络有潜在的风险文件，将会被推送至**高仿真沙箱**，高交互行为沙箱具有丰富的软件环境，高强度的检测对抗手段，同时还会针对不同的恶意样本做针对性的行为触发，基于这些能力高交互沙箱能获取完整的行为报告。

流量引擎所有的产出结果（日志解析内容、协议检测结果、文件行为日志）将会被上传到运营管理系统，进行统一管理。

算法模型模块基于收集的内容会采用智能学习的方式对汇总数据进行综合判断，智能检测。

客户在管理系统中使用**可视化模块**和**告警模块**对安全问题进行跟踪。

2. 产品部署模式



御界高级威胁检测系统的部署模式

高扩展性

御界检测系统支持单机部署，所有的功能模块(TFA、检测模块、沙箱、SOC)部署在单机服务器上。

如果要扩展御界的吞吐能力，您可以选择集群部署，例如将高交互沙箱模块独立部署，这将大幅度提高未知威胁的检测能力。

多模块自由组合

御界检测系统支持模块独立部署，例如您可以去除高交互沙箱模块，这将大幅度提高系统吞吐，同时也能满足基本的安全需求。

低成本

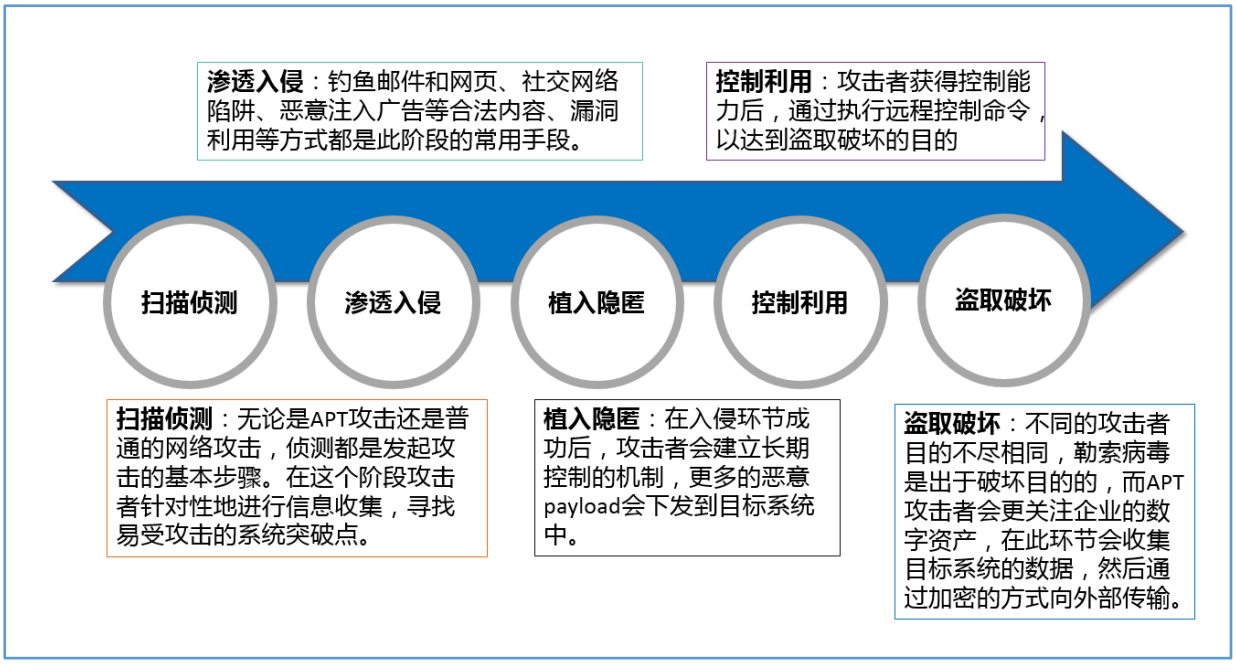
我们提供免硬件的部署方式，客户无需支付额外的硬件成本。

3. 产品规格

	在建议的硬件配置环境下
流量处理性能	2000mbps – 3000mbps
沙箱数量动态调整	支持
文件分析性能	3 万-5 万/天
日志存储周期	6 个月
网络协议支持	HTTP、SMTP、SMB、FTP、DNS、SSH 等常用网络协议
文件类型支持	PE 文件、脚本文件、Office 文档、PDF 文档、图片、压缩包等
操作系统支持	WindowsXP、Windows 7、Linux 等
管理系统	支持控制台，B/S 管理等方式

第三章　系统特点

1. 攻击链条检出



对企业内网的一般攻击过程

传统的企业安全产品可能更着眼于上述流程中的一个点或者两个点进行防护，比如端上的安全产品重点关注 payload，防火墙重点关注远程连接的地址，端和防护墙的防护是完全割裂的。腾讯御界系统从网络边界处的流量入手，通过对流量中的文件信息和连接信息进行综合分析，提供一种面向攻击链的检测体系，在每个攻击环节，部署检测探针，收集特征信息，交由算法模型汇总判定。

每次网络攻击的发生都是存在时序性和关联性的，腾讯御界可以从时间维度和空间维度将整条攻击链还原出来，可以让网络管理人员了解整个攻击流程，从而评估资产风险。并且可以根据攻击链条发现内网安全的薄弱环节进行加固。

2. 发现 APT

御界的沙箱技术以腾讯哈勃分析系统为核心，加之场景化的策略。高仿真、深度分析是哈勃的主要特点，模拟真实用户环境，模拟用户交互，模拟网络环境，从而激发样本行为。其中自研的监控模块，经过了海量样本行为分析的检验，监控广度与深度覆盖的全面性在和国内外各种动态分析系统的对比中均处于领先地位。经过多年的积累，目前哈勃沙箱系统已经成为 google virustotal 指定合作伙伴，是 google virustoal 上唯一的动态行为分析能力的提供者。

行为安全评级的核心

传统的轻量沙箱或者动态分析系统，仅仅把动态分析作为辅助判定的一种手段。而哈勃分析系统通过多种不同的方式，对经过动静态分析的日志进行自动化的解析和处理，根据日志的内容对样本的安全等级进行判定。在保证极低的误报率前提下，识别率处于国际领先水平。

3. 异常流量感知

御界异常流量分析系统利用机器学习和行为检测模型，可以精准识别网络会话异常、上行流量异常、下行流量异常，并可以识别蠕虫 DDoS 攻击、IP 扫描、端口扫描、勒索病毒传播、web 服务探测、邮件服务器探测等若干种恶意行为。基于大数据的能力，可以有效识别异常敏感的网络通信，自动关联网络内安全状况，智能识别异常流量。

4. 勒索病毒检测

御界依托哈勃沙箱动态行为分析技术与动态环境深度模拟技术，可以对最新的敲诈勒索类病毒以及变种进行识别，并且无需像传统 IDS/IPS 一样依靠定期升级规则库。动态环境模拟技术可以使敲诈勒索病毒在沙箱系统内运行时，智能分析并且满足勒索行为触发的所需条件，从而依靠动态行为分析技术识别。

5. 威胁情报

御界依托腾讯安全大数据中心，采集海量样本和哈勃分析集群产生的分析数据，从而生成威胁情报，通过在线或离线升级服务的方式，输送给御界各个子系统。

6. 漏洞攻击检测

基于特征的漏洞检测技术，只能应对历史漏洞的扫描和攻击。面向攻击链的检测方式和深入全面的动态分析技术，使得御界在面对 0day 漏洞的攻击时，有更多的应对办法。在浏览器漏洞，操作系统漏洞，Office 漏洞等方面，腾讯反病毒实验室积累了丰富的经验，检测方法和模型在御界中都得以施展。

第四章　核心功能模块

1　TFA 检测引擎

御界流量系统具有强大的网络协议识别和检测能力，基于 TFA 协议解析引擎，支持丰富的协议类型解析，包括 HTTP，FTP，SMTP，SMB、DNS、SSH、TLS 等，能够在协议解析层面对网络流量进行识别。引擎对文件类型可以进行精准识别，可以甄别图片、压缩包、可执行文件、网页文件、脚本文件等多种文件类型并进行针对性处理，对于压缩包会尝试解压并提取其中的子文件进一步分析，甚至对于一些加密的压缩包也具备一定破解解密的能力具体而言。

入侵特征模型检测模块

依托腾讯多年在网络入侵方面的技术沉淀，实时跟踪全世界互联网络中发生的网络入侵事件，分析入侵攻击过程及原理，及时发布漏洞攻击的行为特征及规则，加强入侵检测的能力，有效预警网络入侵事件。

异常流量模型检测模块

异常流量分析模块将网络数据流作为数据输入源，利用机器学习与大数据的技术，鉴定数据流背后的行为企图，及时发现潜在的网络攻击行为。该模块将资产关系自动识别纳入到学习范畴，并结合管理员指定的资产信息，利用资产属性及相互关系，智能识别非法访问流量，对不符合资产属性及访问关系的网络

流量，进行敏感协议检查，有效提升风险警示能力。采用多种聚类分析，结合时间特性、资产特性和行为特性，有效甄别蠕虫类网络攻击的攻击源和被攻击方，并通过协议图谱及报文图谱识别非规则类的潜在问题。

异常域名检测模块

依托多年网络信息数据的采集积累，通过聚类算法，大数据分析，对网络流量中的域名进行精确识别，阻断恶意流量访问。不拘泥于已有存量库，可以动态感知网络态势，将新的挂马网站和恶意域名归纳入库，具有实时性强，更新速度快等特点。

网络攻击检测模块

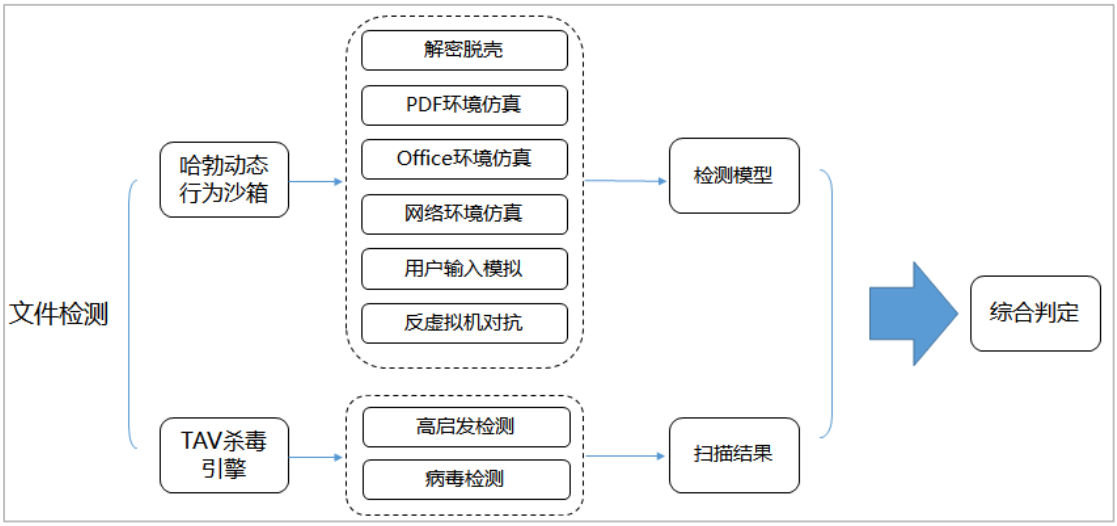
通过日志智能关联分析技术，基于统计学习模型，针对网络流量参数值长度，字符分布，参数顺序，访问频率，访问时间间隔等等参数，通过异常打分模型将多个特征维度异常值融合，得到一个打分结果。同时结合基于文本分析的机器学习模型，使用基于隐马尔科夫模型(HMM)的参数值异常检测对攻击行为进行最终判定，可以有效检测 XSS 攻击，SQL 注入，暴力破解，缓冲器溢出攻击，文件上传漏洞，命令执行漏洞等。

2 文件还原模块

流量文件还原模块可以通过分光、镜像等方式部署于企业网关旁路，通过 TFA 协议解析引擎实时分析和解析网络协议数据包，还原数据文件，之后通过文件分析鉴定模块进一步识别。

流量文件还原模块对还原的文件类型可以进行精准识别，可以甄别图片、压缩包、可执行文件、网页文件、脚本文件等多种文件类型并进行针对性处理，对于压缩包会尝试解压并提取其中的子文件进一步分析。

3 文件分析检测模块



文件检测流程

文件分析鉴定模块用于对流量还原提取的文件进行分析鉴定，及时感知安全威胁。文件分析鉴定模块分为两大组成部分，一个是哈勃动态行为沙箱，一个是 TAV 自研杀毒引擎。

哈勃动态行为沙箱

哈勃动态行为沙箱是一套高仿真环境下的文件行为检测系统，依托于虚拟机技术，在充分模拟文件运行环境的情况下，通过行为序列和机器学习算法对样本进行全面分析，可以有效鉴定经过加壳，代码混淆等高级变形技巧的病毒木马及其新变种，不仅可以识别传统恶意文件，对于 APT 攻击样本这种变化繁复，隐蔽性强，威胁大的新一代攻击也具有极高的识别鉴定能力。

哈勃动态行为沙箱具有如下技术特点：

- 支持丰富的分析环境包括 WindowsXP/Windows7/Windows10/Android/Linux。
- 支持全面的文件类型，包括可执行文件 EXE，Office 文档（Word，Excel，PPT 等），PDF 文件，HTML 网页文件，脚本文件（bat，js，vbs，swf 等），APK 包，以及 RAR、ZIP、7z 等各种压缩包文件。
- 对于加壳加花反调试的样本，具有脱壳解密能力，以及对抗反虚拟机技术的能力。
- 具有网络模拟能力，能够监控样本文件的网络通信，可以模拟断网环境下的网络交互，充分暴露样本的网络行为。
- 具有模拟用户模拟点击和输入的能力，对于存在交互页面的样本，例如安装包，钓鱼页面等，可以自动模拟点击按钮和用户输入，从而触发样本更多的动态行为。
- 通过用户层和驱动层的 Hook 技术，对样本的动态行为序列进行监控，通过陷阱文件的模拟可以有效识别近几年呈爆发趋势的勒索类病毒。

TAV 杀毒引擎

杀毒引擎是杀毒软件等信息安全产品的核心技术之一。腾讯反病毒实验室凭借自身雄厚的实力，在核心技术等方面不断完善，最终通过多年技术积累，孵化出了 TAV 自研杀毒引擎。

TAV 自研杀毒引擎依靠领先的特征码匹配、复合类文档拆解、脱壳、模拟执行、机器学习等技术，目前已成为国内唯一通过自主技术获得 Check Mark、VB100、AV-C 等国际著名测评机构认证的大满贯引擎。

TAV 杀毒引擎具有如下技术特点：

- 在传统特征码查杀的基础上，增加高启发分析能力，可以在汇编指令层面模拟执行程序代码，可以有效对抗加壳类样本。
- 拥有海量样本基因库，精准查杀各类病毒，通过遗传学算法等机器学习能力，对病毒最新变种也可以快速识别，对感染形样本的识别检出效果显著。
- 查杀速度快，秒级响应，识别能力强，误报率低。