流量威胁狩猎圆桌论坛

时间：5.27

主持人：王超（360政企安全集团 技术总监）

演讲专家：

程光（东南大学网络安全学院院长）

姜政伟（中科院信工所威胁情报与威胁分析团队负责人）

黄玮尧（交通银行网络科技部专家）

曾小勇（宁静之盾成都科技有限公司CEO）

李薛（360政企安全集团高级安全专家）

参会人：王晨旭，陈江超

**议题一：俄乌战争对网络安全攻防，特别是流量分析技术的发展有哪些影响？**

程光：

俄乌战争从网络战方面可以分为两个阶段。第一个阶段发生在军事冲突前，为了赢得网络战的胜利，要摸清对方的关键信息基础设施的状态，探明政府军事信息系统的威胁和缺陷，深度看清网络空间的状态和实时变化。这一阶段的攻击主要以apt为主要实施攻击手段进行情报搜集、采集目标设施的信息（入真实网络地址、网络资源的情况等）以及高价值的系统的运作情况。第二阶段发生在俄乌军事战争之后，主要攻击手段包括，“匿名者”对俄罗斯的相关部门进行了DDoS攻击，对俄罗斯的IP的数据删除，对部分的系统进行了安全漏洞（比如“匿名者”入侵相关公司的信息，利用漏洞窃取大量的数据信息），还包括通过舆情进行舆论战攻击。

姜政伟：

在持续关注俄乌战争的过程中，发现了几个特点。1.美国以代理人的方式介入了网络战争 2. 双方舆情认知站占了主流。在网络攻防方面，以常规手段为主，出现了常规的网络攻击软件（如勒索软件、钓鱼邮件、僵尸网络等），漏洞也基本上是一些常规的命令执行，还有去年曝光的热门漏洞（Apache Log4j）。流量方面，主要表现是，规模化高强度的DDoS攻击，还有网站页面篡改（是使用频率很高的攻击手段）。在网络空间测绘层面，乌克兰在开展之后，遭到了持续的DDoS攻击，关基IP掉线率比例比较高，俄罗斯的多个关键基础设施也收到全球僵尸网络的攻击。对于流量分析方面的启示：1. 多种检测手段相结合。不同类型的流量应该用不同的检测方法。对于高速流量，我认为基于签名的检测是比较高校的检测方式之一，但是我们需要探索基于威胁情报、基于人工智能的通信行为建模识别的检测分析方法。2. 对于DDoS和Web应用攻击，如何才能做到准确的检测和有效的阻断。在俄乌战争中，出现了一种称之为众筹式的DDoS，攻击主导者之时在互联网上公开了俄罗斯相关的重点网站，让全球黑客组织自发攻击，甚至只需普通民众自发的访问这些网站，就形成了一种众筹式的DDoS。3. 平时常规流量的检测和战时大规模攻击流量的应急演练都得开展起来。

曾小勇：

我在俄乌战争过程中观察到了一些现象。比如数据擦除这种攻击手段，其前置条件是需要能够完全的控制系统（通过漏洞利用、apt攻击等手段），对数据进行擦除。

黄玮尧：

类似于俄乌战争中出现的DDoS攻击，我们应该如何防御DDos攻击呢？从银行的角度来说，DDoS攻击还是难以防范的。目前我们的抗击手段主要有两种。一种通过运营商大流量的清洗；第二种企业用购买的流量清洗设备进行清洗。

李薛：

从俄乌战争的情况来看，与2016年美国针对中东一个网站的攻击阶段和发展脉络有一点的相似性。在2016年公布的“发光交响乐”美国网络战军事行动中，第一个阶段也是采用DDoS或其他的一种分布式的攻击去摧毁网络的可用性。在第二个阶段，绘制了大量的恐怖组织的网络地图，对它充分做画像。第三阶段，针对这些画像，发起像鱼叉钓鱼、社会工程学的攻击。从三个阶段反看俄乌战争，对于抗击者来说，前期众筹DDoS的防御工作做的比较少。另一方面，我们发现对流量上产生了两个比较的需求，第一很多攻击都会隐藏在众筹的DDoS这些流量的下面，使得很多检测防御设备检测不到，流量检测设备怎么养能够更精确的去发现有价值的攻击行为。第二我们发现的这种有价值的行为，怎样快速的去响应，避免网络的资源被攻击下线。

**议题二：如何构建合理的安全运营体系，基于网络流量的威胁分析如何能够实现平战两用？**

姜政伟：

对于如何构建安全运营体系，在我们看来安全运营体系应该覆盖到资产的管理、防护的监测、响应处置、评估加固等能力，并且形成闭环，同时还要配备相应的人员队伍、软硬件设置和数据治理。有了这套东西，还得从基础层面、效果层面、价值层面做持续运营的一个安全运营的评价，再通过标准化（比如说制定SOP表中的操作流程），流程化（制定安全的运营流程），工程化（能力工具能够平台化），自动话（人机结合），这四方面完善和迭代运营体系。从流量分析的角度来看，运营体系中核心的部分应该包括，攻击的监测、威胁的处置、流量要素的存储能力、建立加密的流量分析、异常网络行为的挖掘和自动告警研判方面的能力。进一步来讲，安全运营体系还应该根据业务方关心的领域来发展有针对性的能力。对于平战两用的角度，我们认为要梳理出平时和战时的工作模式和内容的一些区别，然后将两者功能做深度的兼容（比如，平时做流量分析，可能只对攻击成功的情况需要去做处理，但战时不一样，只要发现有疑似的活着尝试攻击，就应该提前快速的预警和封堵）。此外，对于运营体系，如果真的做到了标准化、流程化、工程化和自动化，可以很自然的实现平时和战时的快速切换。只有在平时积累了大量的流量识别和通信行为的认知知识，才能在战争时期活着特别时期能够做到精准的防护和阻断。

黄玮尧：

平时要学习流量相关的知识，特别是流量的白名单的一些机制。目前分析流量的产品的主要问题，我们认为主要存在于分析流量的产品在分析完数据之后不会存储数据，而这些数据对于我们来说也是一种数据资产。只有在平时的运营中积累了一些基础，才能在战时有备无患。要在平时多演练。

程光：

对于网络流量做到平战两用，我们认为要考虑三个方面的问题：1. 要将流量进行存储。目前很多都没有原始数据，即使存储了，海量数据很难进行查询。希望能够建立一个数据存储和检索的系统。2. 在积累数据之后，我们要构建监管拥护的地址画像以及他的行为规律。3. 研究更多的分析数据的算法，能够对这些数据看得更明白。

李薛：

平战两个状态类似于“台上一分钟，台下十年功”。平时通过对存留的数据进行分析，积累数据、积累能力、积累技术，对一些情况能够形预案，一到战时，能够迅速的将积累的东西运用起来。360累计了全球的大量样本，300亿的样本，其中包括52亿的恶意样本，360本身也在对这些样本进行深入的挖掘，最近公布了40多个我们独立发现的境外APT组织，一旦到了战时，我们会将我们所掌握的有用信息以及云地一体化能力提供给有需要的部门。

曾小勇：

平时和战时是不一样的。在平时的时候，我们会使用黑名单的方式，在标记到识别出一些恶意流量的时候我们进行告警或者阻断。但是在战时，我认为在这种特殊时期，会更加偏向于白名单机制，让认识的流量或者明白其中行为的流量才会放行。这是我的一些见解。

**议题三：还量告警如何提高准确性，告警如何能做到让一般人快速研判？**

黄玮尧：

告警一般是基于特征进行判断，更深入的还有基于流量上下文、关联规则做判断。再深层次的技术基于大数据进行异常行为分析，这种情况下，大量的告警可以配合一些情报信息进行比对。目前为止，基于特征的告警误报率实际是比较高的，还需要利用正则方法基于上下文做一些关联，以提高告警的准确性。目前为止，当告警发生时安全厂商需要提触发告警的原因，如报文中的payload，来告诉方便一线安全人员检测。在实际分析中，主要关注APT情报在沙箱中的运行情况，提取正负样本进行训练，得到模型和规则。

姜政伟：

海量告警产生误报，可能是检测规则粒度太粗，或者是业务无关。从学术研究角度看，真实告警可以说是攻击事件在检测系统中的表达，它不是孤立的。完整的攻击可以映射为有多条日志和告警所构成的攻击场景，所以攻击场景构建是提高准确率和加速人工研判的方法。具体而言，可以从告警和业务数据的一些时空、行为的特征对告警做关联。然后把有共同攻击意图的告警，整合在一起。把一些重复的行为数据做优化。另外也结合ATTCK框架对攻击行为做标注，再加机器学习和专家知识做成完整的攻击链。在业务实战，对于特定的安全场景，大量告警本身就是无效的，不能都推送出来让用户去研判。首先要保证检测质量，如果好似基于规则检测的话，规则来源可靠才能配置上线，并且需要对上线规则做反馈修正。其次应该减少不属于该业务关注领域内的检测规则。第三是结合威胁情报做过滤，加上一些融合规则，筛选出用户关注的告警。

如果要让一般人达到快速研判，一是要提升准确率，二是降低告警数量，将专家经验固化成操作界面。三是要融合多个数据源作为依据，像我们现在建立的国家威胁情报平台，就可以聚合多个厂商的数据，用于全场的告警运营。四是做一些APT样本分析，也能产出比较可靠的情报。此外，告警展示应该比较直观，能过给出系统预判的结果和结论，比如这次攻击是成果了还是失败了。给出攻击的重要程度、危害程度。

程光：这个问题实际上是怎么在海量数据中检测出小样本，我们实验室前期从网络流量中识别Tor流量，与这个问题相似。我们发现采用二阶段的识别效果可能好一些，第一阶段专注于正常信息识别，过滤掉大量正常信息。第二阶段专注于告警信息识别。这种方法可以提升告警准确率。针对告警信息复杂的问题，可以采用告警分离，精细识别告警信息类型。根据下文信息，判断攻击所处攻击链的位置，预测下一步攻击行为，从而提高快速研判的能力。

李薛：从实践角度看，更多情况下是为了看资产有没有被攻击者攻陷，检测应该从发现攻击行为到攻击者攻击到这些资产。对不同的攻击行为进行分类，比如扫描IP一般情况下不理会，战时可能直接封禁。通过告警的自动化研判，整理出被攻陷的资产。这些是应该首先被处理的。至于怎么让一般人快速研判，首先可以对特征进行高亮，放在ATTCK框架里面。其次，基于上下文情况，把多个阶段的攻击聚合在一起，告诉安全人员目前处在killchain的哪个阶段。最后可以用一些自动话扫描工具经常性进行内网扫描，加深对资产的了解。

**议题四：加密流量对流量分析有哪些挑战和先进技术成果？**

曾小勇：

针对标准是SSL、TLS加密，可安装特定的根证书对加密流量进行解密再分析。对通信节点的IP、域名结合威胁情报进行判断。彻底的解决这个问题是比较困难的，可以根据网络协议的会话规律特征做一些工作。比如HTTP加密进行Get请求发起的请求包和响应包的大小和频率是不一样的，可以结合一些人工智能模型进行粗粒度的研判。

李薛：加密化成为了不可逆的趋势，对于生产网的流量，可以把加密流量在内网通过证书解密的方法进行SSL卸载，然后做检测。 在不可解密的情况下，大多还是看流量的特征，比如思科从流量的频率进行检测，即上下行流量的占比。我们360认为主要有3个难点，1是现在的带宽很大，在这种在保证性能的情况下进行分析。我们现在是通过白名单的方法过滤掉正常的证书、IP，对剩下的流量利用机器学习发方法进行过滤。2几乎攻击链的每一环都会产生加密流量，在这种情况下检测多阶段的攻击，实际是需要复合型模型。3是加密流量的检测结果解释性不足，需要通过威胁情报进行一些解释。

程光：

加密流量的加密化主要体现在协议作用域的加密和安全协议的复杂交互特征这两方面。目前的研究主要集中在加密流量分类和加密协议的逆向，分类主要包括加密流量的业务分类、应用分类、视频内容分类、加密网页识别、VPN工具识别，匿名网络分类、未知协议识别。主要是技术：1是对流量特征进行分析，这种需要对海量流量进行分析。2是利用深度学习模型，这种可以在某段时间内有效。

姜政伟：

加密流量识别主要的难点：1是特征信息不足，可用特征减少；2是检测模型的概念飘移，是机器学习学习中的场景问题，这个问题在检测加密流量的时候更为严重，因为网络攻防是不断变化的。概念漂移会使训练好的模型准确率降低。3是标注样本严重缺乏；4是实时检测不太现实，消耗资源太大。现有的研究方法：1是基于签名的检测，在TLS1.2之前比较好用；2基于指纹的检测，抽取TLS1.3的指纹；3是基于机器学习的检测方法；4是利用深度学习，学习流量中的时间和空间特征。目前真实应用中存在的问题：1是对通讯数据的质量要求比较高，如果会话包丢失超过一定比例判断就有问题；2是可解释性和可验证性不足；3是误报率比较高。

黄玮尧：

作为银行部门我们最主要关注高可靠性，我们会配套一些终端EPP的方法进行综合解密。