威胁猎人
THREAT HUNTER

以情报构筑数字化安全基石

业务安全蓝军测评标准白皮书

（2024年版）

出品方：威胁猎人

（深圳永安在线科技有限公司）

THREATHUNTER

目录

前言

互联网黑灰产成长至今，已形成了团伙化、自动化、生态化、上下游严密配合的产业链网。据统计，互联网黑灰产对互联网企业及线下实体行业每年造成近千亿元的损失。网络黑灰产的从业群体伪装成正常的业务请求不断蚕食鲸吞着企业业务的利益，如掠夺新人红包的羊毛党、利用企业平台流量批量进行恶意营销诈骗的引流产业、破坏企业推荐计费规则的刷量作弊产业等。随着互联网黑产攻击模式的成熟，运作模式的可复制性越来越高，业务安全问题日渐突显。

当前，企业在处理业务安全问题时往往面临着以下挑战：

攻防信息不对等：企业对黑灰产攻击手段及变化缺乏深入了解，造成攻击发现处理滞后、周期长的局面。

策略效果评估难：策略下发后，难以全面评估策略效果、及时发现黑产新的绕过手段等。

评估体系缺失：不同于互联网基础安全，业务安全问题没有明确的边界。在基础安全中，多数攻击的危害程度已有像 OWASP Top10 这样的评估标准，而业务安全由于场景复杂，同时需要综合考虑用户体验及用户活跃，一直缺乏一套行之有效的评估体系。

与基础安全的严防死打不同，业务安全的目标通常不是杜绝攻击，而是将攻击限制在可控的范围内，从而确保业务的正常开展和业务规模的健康增长。

以营销活动场景对抗羊毛党举例，实体商品清库存的互联网促销活动、餐饮行业需要到店消费优惠折扣券等情况，对羊毛党的容忍度较高，业务安全的治理目标是将羊毛党控制在 50%

以下，而针对特定客群的拉新活动，则希望补贴尽可能落在真人用户上，对羊毛党容忍度较低——真人将有一定转化率成为长期用户，羊毛党的转化率几乎为 0，羊毛党占比过高将损失掉很多机会成本，那么业务安全的治理目标是将其控制在 10%以下。可以看到，即使都是营销活动场景，活动类型和规则不同，评估的目标也会不同。

因此，**企业的业务安全问题需要一套评估体系，能够数字化体系化地描述遭受攻击带来的危害程度及安全策略实施后的效果等。**

基于以上需求与目标，威胁猎人在 2020 年发布了国内首个《业务安全蓝军测评标准白皮书》，填补了业务安全行业长期以来缺失攻击危害及安全策略效果评估体系的空白。时隔四年，威胁猎人结合**第一版标准落地过程中遇到的挑战以及过去几年在各行业多家客户的业务安全蓝军实战经历，对第一版标准进行了修订和更新，发布了《业务安全蓝军测评标准白皮书（2024 年版）》。**

01

业务安全蓝军测评标准

一、业务安全蓝军测评标准

1.1 业务安全脆弱性评分(ISVS)

业务安全脆弱性评分（Interaction Security Vulnerability Score）是从攻击者视角出发，将企业的某个业务对象设定为攻击目标，使用黑灰产的攻击方式对该对象发起模拟攻击测试，还原攻击过程，并结合最终的攻击结果评估黑灰产攻击给该业务对象带来的危害。

ISVS 计算方式：

ISVS = Max（ 攻击效率 AE * 目标达成率 AR ）

- **攻击效率 AE(Attack Efficiency)**

测评对象或达成目标不同，评估攻击效率取值的方式往往也会不同，通常可以分别从物料获取效率和技术对抗效率这两个角度来考虑。物料包括攻击过程中需要用到的手机号资源、账号资源、IP 资源、设备资源、身份认证资源等，攻击者获取这些物料的成本越低，攻击效率越高，反之越低；技术则包括攻击过程中为了破解应用保护或绕过风控限制等，所用到的软件逆向技术、改机技术、改定位技术、人脸伪造技术等等，攻击者应用这些技术并攻击成功的门槛越低，攻击效率越高，反之越低。

注：关于攻击效率指标取值高低参考，详见附件。

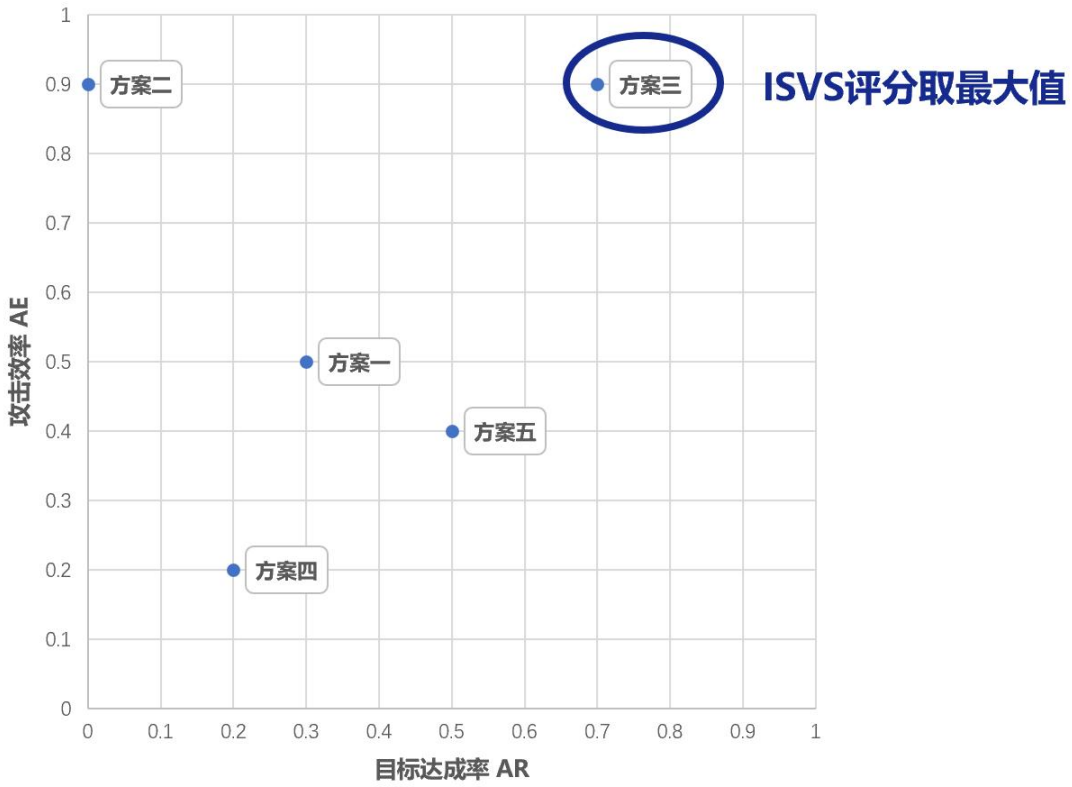
- **目标达成率 AR(Achievement Rate)**

目标达成率 AR 是指该攻击方案的最终测试结果达到预期攻击目标的比率。达到或超过预期攻击目标，取值为 1。

如果是定量的攻击目标，比如预期在指定时间段内攻击成功 10000 次，实际攻击成功 8000 次，则目标达成率为（8000 / 10000）= 0.8；如果是非定量的攻击目标，比如预期是绕过测评对象的人脸识别，实际结果也是绕过成功，则目标达成率为 1，否则目标达成率为 0；如果需要在限定的条件下才能达成目标，比如只有低版本安卓系统才可以攻击成功，则可以结合实际情况进行取值 0 到 1 之间。

以上两个指标各自评估出一个[0, 1]区间的得分，相乘便得到 ISVS 评分，取值区间也是[0, 1]。由于在测评过程中可能会采用多个攻击方案进行测评，需要综合多个攻击方案的 ISVS 评分，取最大值（基于木桶短板原则）。ISVS 分数越高，则说明测评对象面临该攻击方案时越脆弱，若黑灰产使用该方案对业务发起攻击，造成的危害越大。

下图是 XX 产品面对五种攻击方案时 ISVS 取值的散点图分布：



XX 产品-业务安全蓝军测评结果

1.2 ISVS 评分的参考意义

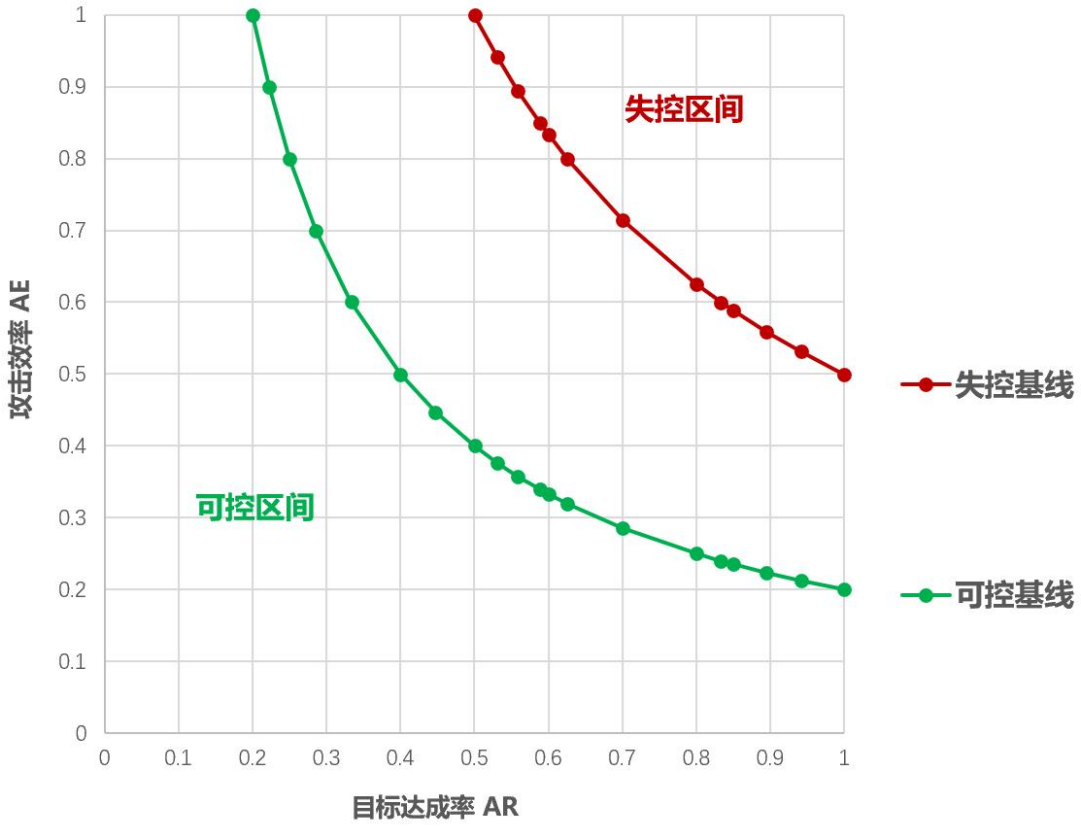
企业在进行业务安全脆弱性评分(ISVS)时不是单纯追求低分，而是从攻击者视角，客观反映出当前业务安全的水位，及时暴露短板，并以此推动整改和治理。

1.2.1 基线对比

前文提到业务安全的目标不是杜绝攻击，而是将攻击限制在可控的范围内。如果有明确的定义，比如使用黑产广泛掌握的攻击方式（攻击效率 AE 取值 1）发起 1000 次攻击，攻击成功次数低于 200 属于可控，超出 500 属于失控，则可以建立两个 ISVS 评分基准值：可控

基线 0.2 和失控基线 0.5。评分在区间[0, 0.2)说明将攻击限制在可控的范围内，评分在区间(0.5, 1]说明已经失控，急需加强安全建设。

以散点图形式对可控区间和失控区间进行直观展示：

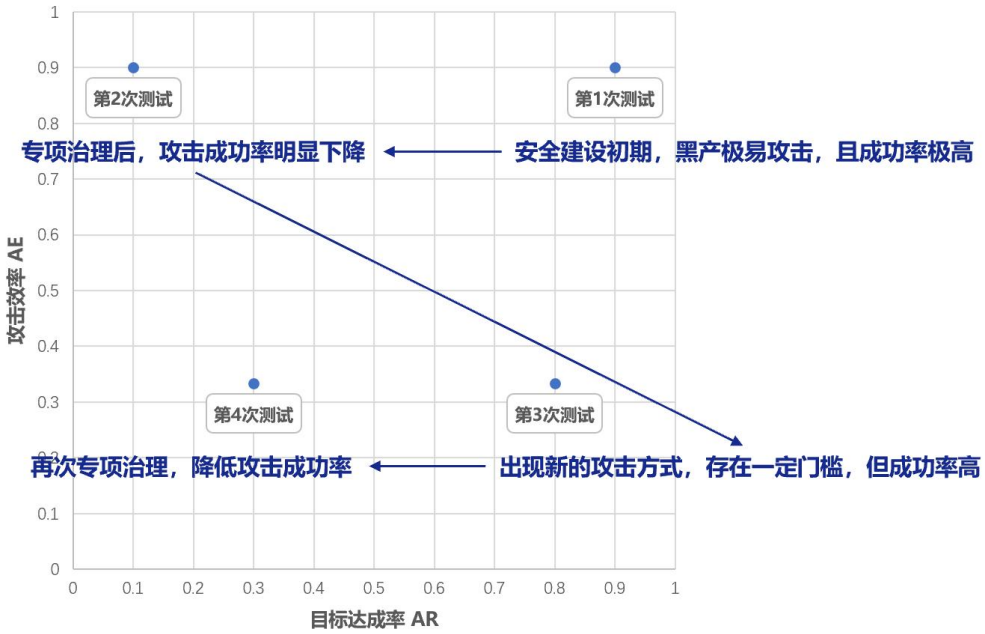


XX 产品-业务安全蓝军测评基线

1.2.2 纵向对比

测评对象和预期目标不变的情况下，ISVS 评分也会随着业务安全水位的变化和黑灰产攻击方式的升级迭代而变化，因此业务蓝军测评需要周期性地进行，通过 ISVS 评分的纵向对比来反映安全建设工作的实际成效，以及是否出现了新的短板。

以散点图形式不同阶段 ISVS 评分进行直观展示：

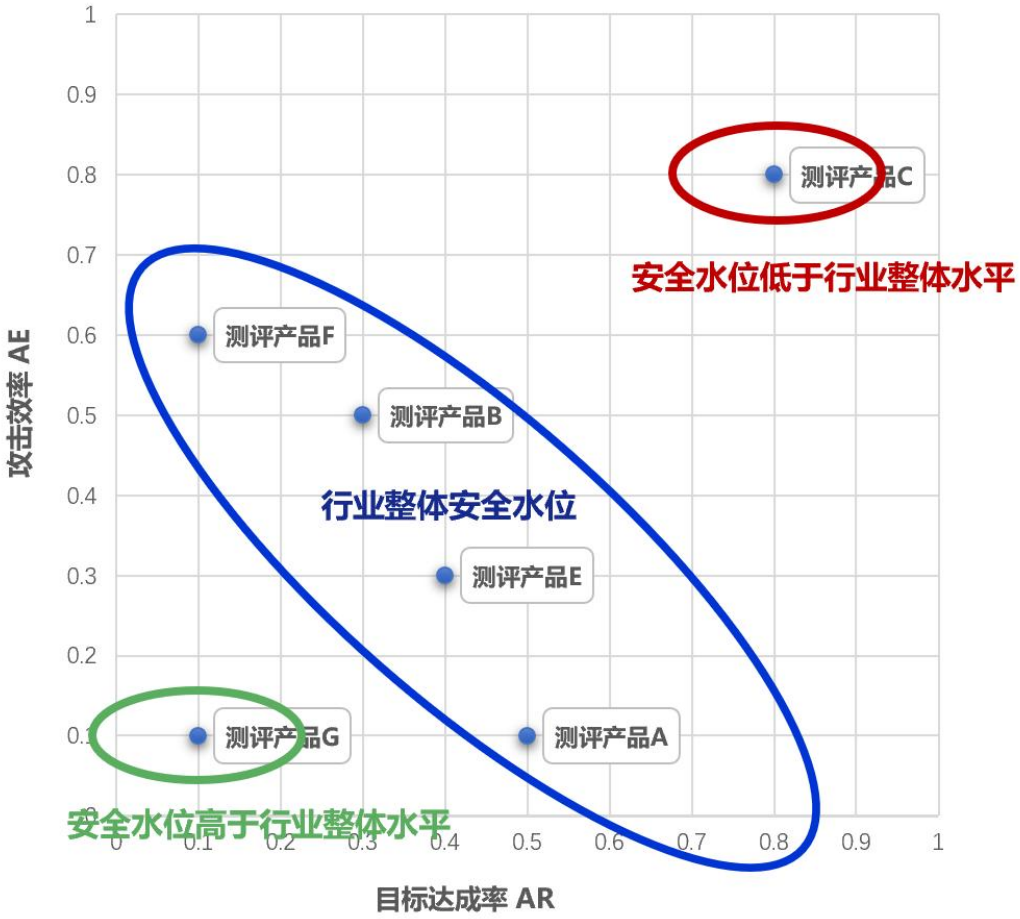


XX 产品-业务安全蓝军纵向测评

1.2.3 横向对比

跟行业内的同类业务进行对比测评，并通过 ISVS 评分来反映测评对象在行业内的安全水位。如果相比行业内的同类业务，ISVS 评分偏高，则需要尽快加强安全建设，因为相对薄弱的业务必然会成为黑灰产的重点攻击对象。

以散点图来展示测评对象在行业内的安全水位情况：



XX 行业-业务安全蓝军横向测评

02

业务安全蓝军测评案例

二、业务安全蓝军测评案例

2.1 虚假安装蓝军测评案例

虚假安装主要出现在业务营销推广当中，某些推广渠道会跟黑灰产勾结，伪造虚假的应用安装量，从而骗取企业的推广费用。

2.1.1 定义测评对象和攻击目标

测评对象：某社交应用

预期攻击目标：攻击 1 周，每个攻击方案平均每天成功完成 2000 次虚假安装

2.1.2 制定攻击方案

基于黑灰产研究所掌握的情报信息，从攻击者视角来看，虚假安装目前主要有两类攻击路径：

1）在真实手机或模拟器上通过脚本模拟点击的方式进行应用的安装和启动，并通过群控批量控制多台设备，通过改机技术篡改设备的 IMEI、安卓 ID、Mac 地址等参数来伪造新设备。

2）破解应用的接口协议，编写自动化程序伪造应用安装和启动的接口请求，在请求参数中填入虚假的设备信息，在没有实际安装应用的情况下直接欺骗业务后台。

根据以上两种攻击路径，制定出四种攻击方案，如下：

方案一：通过模拟器来伪造多台手机设备；

方案二：使用真实手机，并通过软件改机的方式来伪造新设备；

方案三：使用真实手机，并通过定制 ROM 改机的方式来伪造新设备；

方案四：破解应用接口协议，直接伪造安装和启动的接口请求。

2.1.3 执行攻击并评估 ISVS

攻击方案	方案说明	AE	AR	ISVS
方案一	该应用具备较强的模拟器检测能力，测试多款模拟器都无法正常启动该应用，即使根据黑灰产所掌握的方法刻意抹去模拟器特征也不能成功，因此攻击效率 AE 取值为 0，ISVS 评分也为 0。	0	/	**0**
方案二	该应用具备较强的前端对抗能力，测试多款改机软件，均被应用检测出注入/Hook 等痕迹导致无法正常启动该应用，因此攻击效率 AE 取值为 0，ISVS 评分也为 0。	0	/	**0**
方案三	定制 ROM 改机通过篡改系统底层源码来篡改设备信息，导致应用无法对其进行有效的对抗，测试可以安装并启动应用，同时每次改机后都会认为是一台新设备的安装。但定制 ROM 改机在测评期间并没有被黑产广泛使用，且需要购买专门的设备，因此攻击效率 AE 取值为 0.5，最终平均	0.5	0.6	**0.3**

	每天完成的虚假安装次数为 1200 次，目标达成率 AR 为(1200/2000) = 0.6，ISVS 评分为 0.5*0.6 = 0.3。			
方案四	虽然协议攻击是黑产常见的攻击方式，但是该应用对接口协议做了加密，网络抓包只能看到加密后的内容无法直接进行协议伪造；同时该应用做了代码保护，无法直接反编译加密算法也无法直接调用加密函数。最终在高级逆向分析工程师投入一个月左右的时间才完成了协议算法的破解，这对于黑灰产来说技术对抗效率是比较低的，攻击效率 AE 取值为 0.1。破解之后平均每天可以完成超过 2000 次的虚假安装，目标达成率 AR 为 1，ISVS 评分为 0.1*1 = 0.1。	0.1	1	**0.1**

结合方案一到方案四的 ISVS 评分，测评对象的综合 ISVS 评分为 **0.3**，同时可以给出评语：

测评对象在防御虚假安装上具备较强的安全能力，黑灰产无法通过模拟器或软件改机伪造新设备来制造虚假安装，而伪造协议的攻击方式对于黑灰产有着很高的技术门槛，可以采用定期更换算法的方式来进一步限制这类攻击。对于定制 ROM 改机不具备检测能力，考虑到该方式会被黑产越来越多地使用，需要尽早进行针对性防范。

2.2 人脸识别绕过蓝军测评案例

人脸识别绕过主要出现在一些关键的身份认证场景当中，比如网络犯罪分子在掌握了受害者的账号密码之后，还需要绕过人脸识别，才能成功进行异地登录、转账等敏感操作。

2.2.1 定义测评对象和攻击目标

测评对象：某金融应用

预计攻击目标：成功绕过应用的人脸识别，完成非本人的账号登录

2.2.2 制定攻击方案

基于黑灰产研究所掌握的情报信息，从攻击者视角来看，绕过人脸识别有两个关键步骤，分别是生成虚假的人脸视频和控制手机摄像头播放人脸视频。

生成人脸视频有两种方式，分别是：

1）人脸活化：借助 CrazyTalk 等软件，基于受害者的照片生成眨眼、点头、摇头等动作的视频；

2）人脸替换：借助 DeepFaceLab 等软件，将提前录制好的视频中的人脸，替换成受害者的人脸。

控制手机摄像头播放人脸视频有三种方式，分别是：

1）拍摄电脑屏幕：在电脑上播放人脸视频，手机摄像头直接拍摄电脑屏幕；

2）手机摄像头劫持：购买黑灰产定制的过人脸手机，劫持摄像头的视频流；

3）云手机虚拟摄像头：利用云手机的虚拟摄像头功能，可以播放指定的视频。

因此一共组合出 2*3 = 6 种攻击方案，如下：

方案一：人脸活化+拍摄电脑屏幕；

方案二：人脸活化+手机摄像头劫持；

方案三：人脸活化+云手机虚拟摄像头；

方案四：人脸替换+拍摄电脑屏幕；

方案五：人脸替换+手机摄像头劫持；

方案六：人脸替换+云手机虚拟摄像头。

2.2.3 执行攻击并评估 ISVS

攻击方案	方案说明	AE	AR	ISVS
方案一	使用人脸活化方式制作出来的视频，无法通过检测，因此方案一的目标达成率 AR 取值为 0，ISVS 评分也为 0。	/	0	**0**
方案二	同方案一	/	0	**0**
方案三	同方案一	/	0	**0**
方案四	摄像头对着电脑屏幕拍摄，无法通过检测，因此方案四的目标达成率 AR 取值为 0，ISVS 评分也为 0。	/	0	**0**
方案五	劫持摄像头并播放人脸替换后的视频，可以通过检测并登录成功，因此方案五的目标达成率 AR 取值为 1。	0.5	1	**0.5**

	但这种攻击方案需要购买专门的过人脸手机，且不保证稳定性，攻击效率 AE 取值为 0.5，ISVS 评分为（0.5 * 1）= 0.5。			
方案六	云手机虚拟摄像头播放人脸替换后的视频，可以通过检测并登录成功，因此方案六的目标达成率 AR 取值为 1。虽然云手机使用门槛不高，但该攻击方式在测评期间并没有被黑灰产掌握，综合攻击效率 AR 取值为 0.2，ISVS 评分为（0.2 * 1）= 0.2。	0.2	1	**0.2**

结合方案一到方案六的 ISVS 评分，测评对象的综合 ISVS 评分为 **0.5**，同时可以给出评语：

测评对象对于防御人脸识别绕过存在一定的安全盲区，虽然通过人脸活化技术制作的视频无法绕过检测，但通过人脸替换技术制作的视频，配合特定手机劫持摄像头以及云手机虚拟摄像头，均可以成功绕过检测。一旦被黑灰产恶意利用，可能会给用户带来巨大的资产损失，因此需要尽早进行针对性防范。

03

业务安全蓝军测评类型

三、业务安全蓝军测评类型

前文提到互联网黑灰产成长至今已形成了团伙化、自动化、生态化、上下游严密配合的产业链网，随着网络黑产攻击模式的成熟，运作模式的可复制性越来越高，业务安全问题日渐突显。当下，企业在处理业务安全问题时面临着攻防信息不对等、策略效果难评估等挑战，业务安全蓝军测评服务应运而生。

业务安全蓝军测试人员将模拟真实黑产，使用真实环境中的黑产资源、攻击工具和方法，在可控范围内进行非破坏性的还原攻击测试，复现攻击细节，帮助企业了解以下内容：

- 产业链现状：得到当前产业链攻击的投入成本、产出利益、产业链的攻击规模及攻击企业自身的黑产的攻击规模等。辅助企业从成本等多维度优化对攻击行为的处理方式，提高防守效率，加高攻击者门槛等。
- 攻击手段细节：得到当前黑产的一手数据及攻击手段细节，辅助企业合理及时调整风控策略及规则。

当前，业务安全蓝军测评主要有以下几种类型：

3.1 基础测评

基础测评适用于业务安全建设早中期，这个阶段企业往往对黑灰产缺乏了解，不清楚业务是否有遭受黑灰产攻击以及是否攻击成功，也不清楚黑灰产攻击造成的危害，因此需要通过专业的业务安全蓝军测评来建立对黑灰产认知，并针对性进行安全建设。

为了帮助客户进一步了解黑灰产，除了输出业务安全蓝军测评报告之外，还可以根据客户需求，输出相关的黑灰产产业链研究报告。

3.2 周期测评

周期测评适用于长期开展的业务功能或营销活动，同时业务安全建设到中期之后，由于业务发展的变化和黑灰产的变化，需要通过周期性的测评来反映当前的业务安全水位，是否有提升或下降，评估各项安全策略的具体效果的同时，及时发现新出现的业务安全盲区。

3.3 行业横评

指定相同场景下的相同测试项或相似测试项，选定同行业不同公司，对相同测试项同时进行测评，并对其进行横向对比。此类测评会存在一定限制，对于需要取得授权才能进行的测试项，在未获得授权时不能进行测评。

04

业务安全蓝军测评场景

四、业务安全蓝军测评场景

4.1 营销活动作弊场景

补贴、新人红包、助力提现等活动是企业常见的营销手段，也是黑灰产普遍会关注的利益点。不同的活动和目标面对的羊毛党攻击风险及容忍度不同，测评的预期目标也会有所不同，以下是一些典型的测评项目举例：

1、特定优惠券批量获取类活动

测试项目：某平台 5/10/15 元新人券获取变现；

预期攻击目标：1 小时内完成 100 次账号注册，账号要求可成功获取并使用优惠券。

2、拼团类活动

测试项目：某平台的老带新拼团活动；

预期攻击目标：1 小时内完成 30 组拼团，进入发货流程视为发起羊毛攻击成功，考虑实体商品变现渠道前提下。

3、助力邀请拉新提现类活动

测试项目：某平台邀请好友得现金活动；

预期攻击目标：老号拉新模式拿到最高奖励，且可以批量执行，提现成功记为攻击成功。

此外还有几点需要注意下：

1、营销活动可能出现在应用中，也可能出现在小程序中或者 H5 页面中，对于不同的端，黑灰产的攻击效率也会不同；

2、黑灰产攻击营销活动最终要通过某种渠道进行变现，因此测评时可以根据客户需求，将最终成功变现的数量作为攻击目标。

4.2 业务刷量场景

业务刷量主要出现在社交、内容、电商等平台，通过刷指定对象的访问量、点击量、播放量、点赞量、评论量等，一方面可以骗取平台的激励，另一方面可以提升排名获得更多曝光量或资源倾斜，对平台生态造成破坏。以下是一些典型的测评项目举例：

1、刷访问/点击/播放量

测试项目：某视频网站指定视频；

预期攻击目标：1 天内完成 10000 次视频播放，且不能被后端风控识别为虚假刷量。

2、刷点赞/回复/评论量

测试项目：某社交平台指定帖子；

预期攻击目标：1 天内完成 500 次回帖，且不能被后端风控识别为虚假刷量。

4.3 广告刷量场景

广告刷量是广告行业一直就有的潜规则，按照不同的结算方式，通过刷广告的曝光量、点击量、下载量、安装量等，骗取广告主的广告费用。而对于广告主而言，往往也并不要求做到100%的真量，因此测评项目主要围绕以下两个需求来设计：

1、限制大规模刷量

通过测评来评估黑灰产是否可以通过群控、伪造协议等方式大规模制造虚假的广告曝光量、点击量等。

2、对比渠道质量

通过测评获取的数据，计算广告主投放的各个渠道中假量的占比，作为渠道质量对比的重要依据。

4.4 人脸识别绕过场景

人脸识别作为最为重要的身份认证手段，确保执行登录、转账、支付等敏感业务动作时是本人在操作。一旦人脸识别被绕过，意味着最可靠的防御机制被突破，难以再对后续危害进行有效的阻断。因此对于金融、支付等跟用户资产强关联的应用，针对人脸识别绕过场景的蓝军测评，是非常必要的。

随着 ChatGPT 等 AI 技术的发展，制作的人脸视频会越来越逼真，这将是一个长期攻防对抗的过程，因此测评需要周期性地进行。

05

业务安全蓝军测评流程

五、业务安全蓝军测评流程

1、沟通测试需求：双方就测评对象和预期攻击目标等进行沟通说明；

2、蓝军测评小组输出推荐测试项文档：测评小组根据客户需求输出推荐的测试项及测试评估标准等；

3、讨论确定测试对象和预期攻击目标；

4、签订合同及授权协议；

5、实施测试；

6、验收测试结果。

附：攻击效率指标取值高低参考

1．物料获取效率

1）手机号

获取方式	攻击效率
短信验证码存在暴破或回显等漏洞，导致可以随便填写手机号	高
通过黑灰产接码平台获取的手机号可以进行注册、登录和业务操作	高
通过黑灰产私密对接获取的手机号可以进行注册、登录和业务操作	中
只能使用真人在用的手机号进行注册、登录和业务操作	低

使用查杀分离的方式，只有在黑手机号注册的账号进行恶意业务操作的时候才进行限制	低

2）账号

获取方式	攻击效率
新注册的账号没有任何限制，可执行敏感业务操作如批量发帖、回帖等	高
新注册的账号无法执行敏感业务操作，需要积累到一定的活跃度	中
可以使用机器养号的方式积累账号的活跃度	高
无法使用机器养号，只能人工养号或购买真人账号	低
账号更换设备或地域进行登录时，无需进行任何验证	高
账号更换设备或地域进行登录时，需要进行短信等验证	中
可以通过 PS 伪造营业执照并成功注册企业账号	高

3）IP

获取方式	攻击效率
无 IP 访问频率、常见登录地域 IP 等风控限制策略	高
使用代理 IP、秒拨 IP 等方式可以绕过风控	中
使用代理 IP、秒拨 IP 等方式可以绕过风控，但部分攻击被识别或拦截	中

使用代理 IP、秒拨 IP 等方式难以绕过风控	低

4）设备

获取方式	攻击效率
可以使用常见的模拟器来伪造出多台设备	高
可以使用多开分身等工具来伪造出多台设备	高
可以使用软件改机的方式来伪造出多台设备	高
可以使用定制 ROM 改机的方式来伪造出台设备	中
可以使用硬件改机的方式来伪造出多台设备	低

2. 技术对抗效率

1）协议破解和伪造

破解和伪造方式	攻击效率
通过抓包工具截获的可以直接进行重放攻击	高
通过常见的反编译工具可以直接还原出协议请求的相关代码	高
协议请求相关的代码做了加固，破解难度低	高
协议请求相关的代码做了加固，破解难度高；但可以直接调用相关函数	中
协议请求相关的代码做了加固，破解难度高；且无法直接调用相关函数	低

2） 前端对抗

对抗方式	攻击效率
可使用 XPosed/LSPosed 等常见框架注入应用进程进行攻击	高
可使用黑灰产魔改后的框架注入应用进程进行攻击	中
可使用注入系统进程或服务的方式进行攻击	中
以上所有对抗方式都无法正常进行攻击	低

3）人机识别对抗

对抗方式	攻击效率
未使用人机识别验证码或使用简单的图形验证码	高
使用了滑块/点选等较为复杂的验证码，需要借助黑灰产打码平台	中
使用群控工具、按键精灵/AutoJs 等脚本工具或 WebDriver/Selenium 等 Web 自动化工具编写模拟人工操作的脚本，可以攻击成功	高
使用群控工具、按键精灵/AutoJs 等脚本工具或 WebDriver/Selenium 等 Web 自动化工具编写模拟人工操作的脚本，部分攻击被识别或拦截	中
使用群控工具、按键精灵/AutoJs 等脚本工具或 WebDriver/Selenium 等 Web 自动化工具编写模拟人工操作的脚本，很难攻击成功	低

威胁猎人 THREAT HUNTER

以 情 报 构 筑 数 字 化 安 全 基 石

关注“威胁猎人”

深圳永安在线科技有限公司（品牌名：威胁猎人）

官方服务热线：400-809-3699

官方合作邮箱：marketing@threathuter.cn

官网地址：www.threathunter.cn