引擎深度预处理能力叠加机器学习的产品实践

安天移动政企产品部

执行体治理赋能与大模型辅助

01 安天移动AVL引擎能力介绍

安天移动AVL引擎——全球领先的终端安全防护能力

10亿

累计移动APP样本

30亿

全球智能终端覆盖量

60%

活跃智能设备覆盖率

覆盖全球移动智能设备30亿台，累计移动APP样本10亿级，国内活跃智能设备覆盖率 60%

安天移动AVL引擎——丰富的移动样本信息

作为一家移动安全厂商，涵盖超过 8000万 移动应用用安装文件

考虑不同版本和渠道发布，总APP文件标定数量超过 8亿

安天移动AVL引擎——移动应用安全分析运营体系

移动安全领域专家，业界领先的APP样本分析能力。依照国际、国内以及行业标准，帮助客户了解APP安全现状和风险，提供战略、技术层面的安全咨询服务，为客户建设全面、有效、的样本全面分析体系。

安全管理体系咨询

参照国际标准ISO27001、ISO27002、以及国内标准GB/T22080、GB/T22081，按照PDCA的完整管理过程，帮助客户建设、认证和维护完整、有效的信息安全管理体系（ISMS），为企业业务的运营提供充足的安全保障。

应用安全风险评估

包含但不限于本地接口安全评估、本地存储评安全估、应用通信安全评估、后端服务器安全评估、应用防护安全评估方案、可定制化业务安全评估，对所有评估发现的风险给出对应的安全处置和加固建议，帮助企业提升对信息系统的安全风险管理能力和安全保障能力。

情报能力服务

情报运营：持续对威胁情报线索收集、规纳整理记录，累积安全知识类型数据积累

情报分析：攻击的来源，传播途径、方式，攻击对象，攻击的技术手段，攻击的目的，影响的范围 。

情报持续追踪：对高级别APT组织威胁事件进行持续跟进分析

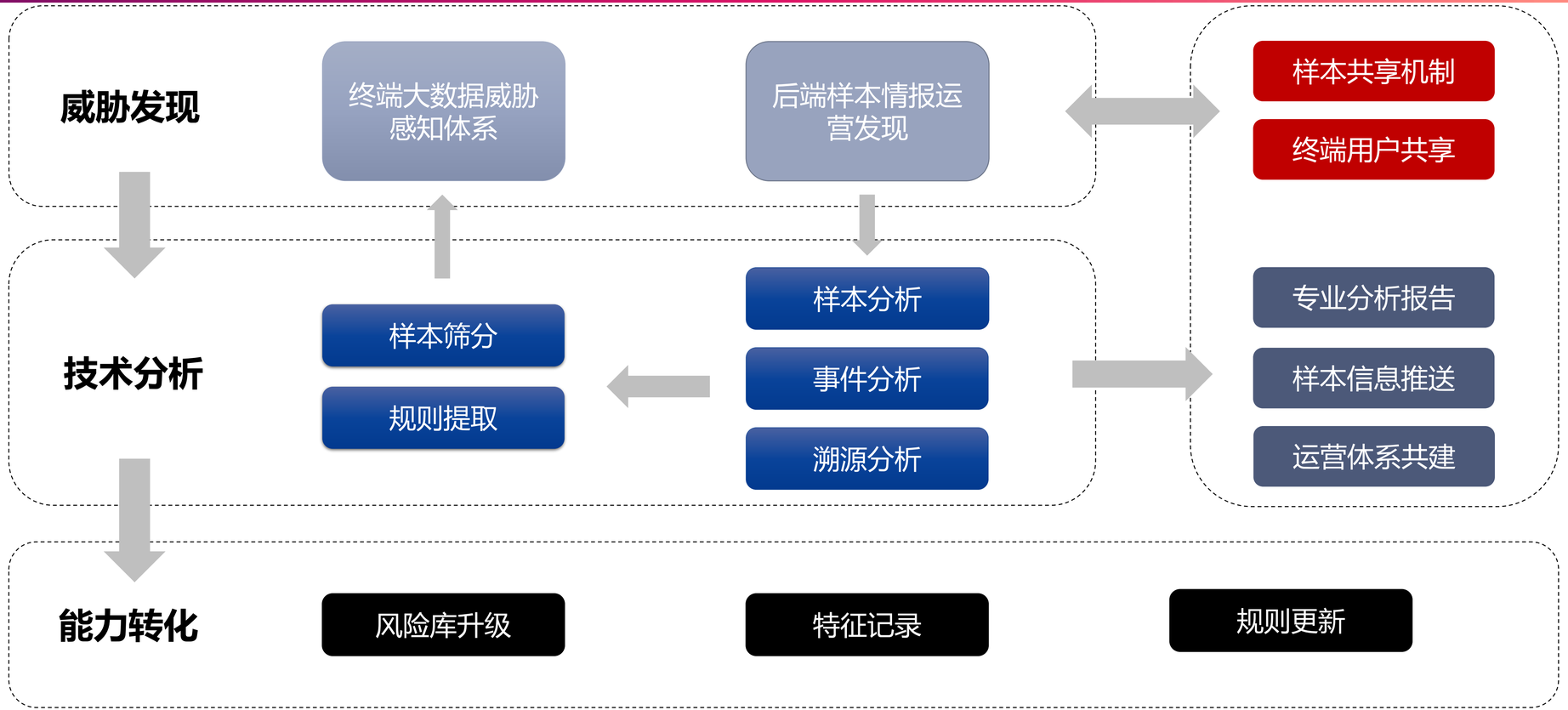
情报挖掘：收集公开情报，并针对有价值的情报进行二次分析分析样本最新捕获有价值的病毒家族，通过自建模型侦监测客户关注点挖掘情报。

样本全面评估 线索追溯

样本分析能力：使用反编译、Hook、调试、沙箱等工具对样本进行动静态分析。可准确分析样本功能，行为，恶意性，数据传输特征，影响范围。溯源样本开发者和同源样本挖掘。

数据分析能力：大数据平台进行数学模型建立，对数据进行统计分析。可完成1.设备人物画像 2.设备排查3.态势感知。

安天移动AVL引擎——基于海量发现与协同实现威胁精准治理



安天移动AVL引擎——全球领先的终端安全防护能力

1 **测评样本库：**
全球随机2500个恶意样本

2 **安天AVL测评环境：**
本地扫描（无需网络环境）

3 **测评标准：**
恶意样本检出率，实时和常规

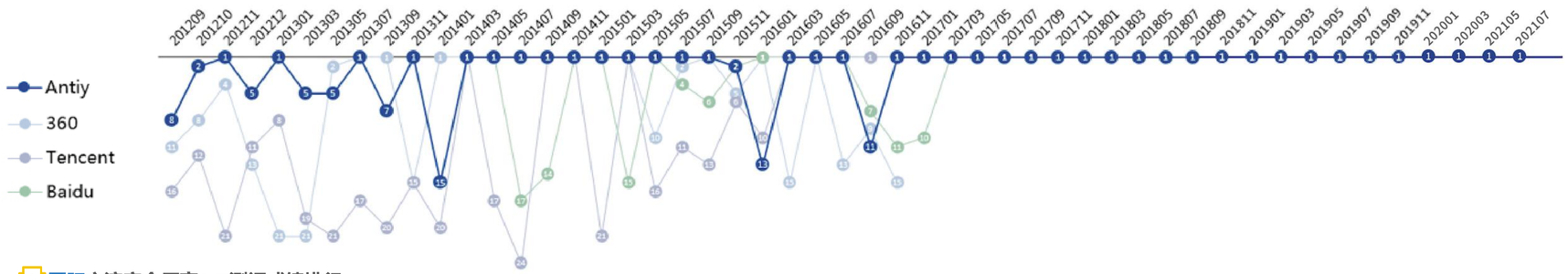
4 **参测产品：**
全球主流的移动反病毒产品

连续平均检出率
全球第一

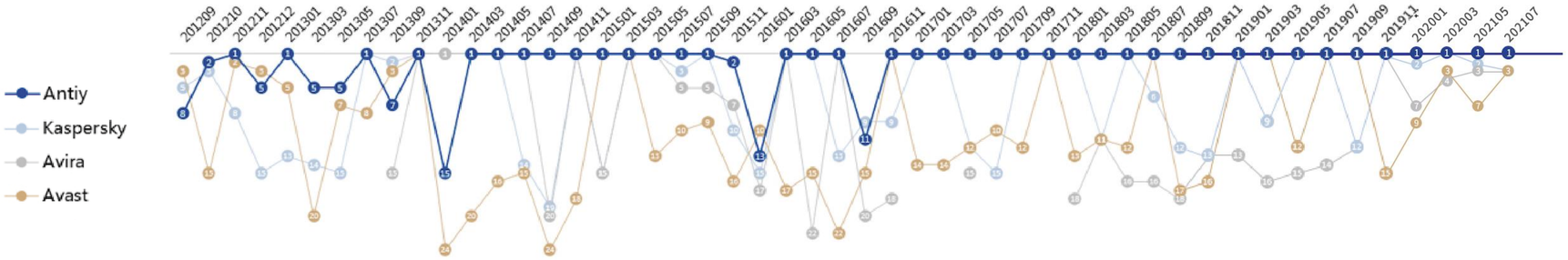
安天（Antiy）
99.95%

Avira	Kaspersky	Avast
99.80%	99.76%	99.58%
百度	腾讯	360
99.52%	99.34%	98.91%

国内主流安全厂商AVT测评成绩排行



国际主流安全厂商AVT测评成绩排行



02

AI对移动反病毒引擎的增强

采用AI增强移动反病毒引擎的检出

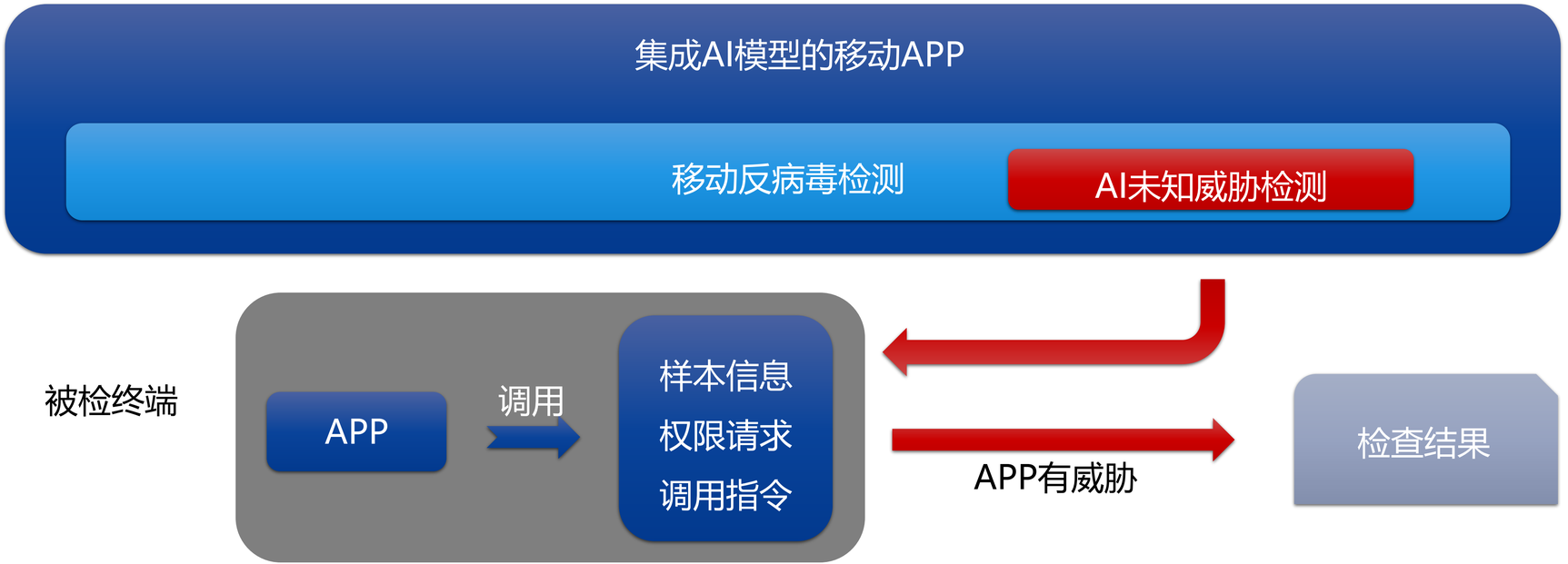
现有移动反病毒引擎追求尽量少的资源开销，尽可能覆盖全部病毒特征，又快又准地判定移动终端中存在的病毒，避免病毒造成的危害，保障99%以上检出率，0.001%的误报率

移动未知威胁样本，基于免杀、漏洞利用、触发式启动、加壳、重打包等方式进行对抗，意图减弱反病毒引擎的检出判定

为了提高对该类移动未知威胁样本的检出，保证覆盖率和时效性，我们引入AI机器学习技术进行检测识别的增强。

AI对移动反病毒引擎增强产品实践

应用程序在设备上运行时，会触发一系列的行为，如API的调用、权限请求、硬件访问等。通过收集大量应用程序样本信息，分析它们的行为特点，利用机器学习训练各种类型的行为模型，是AI未知威胁实时防护动态检测应用行为的基础。通过检测应用程序运行时的行为序列，由AI模型输出判定结果，从而向用户反馈应用程序的安全风险等级。



AI反病毒产品对移动未知威胁的挖掘发现实践

某客户利用AI增强的反病毒产品形态，检查手机终端，试图检查“飞马病毒”的影响，
其中我们分析出 1个疑似“飞马病毒”变种，这在情报发现的及时性有一定的抢先。

飞马病毒命名

Trojan/Android.pegasus.a[prv,rmt,spy,opc]

Trojan/Android.pegasus.a[prv,rmt,spy,gen]

AI对移动反病毒引擎增强的效果

测试方案	反病毒引擎	反病毒引擎+AI	实际分析
1万个黑样本	检出率 100%	检出率 100%	
10万个全新随机样本	5032个样本判黑	5137个样本判黑	5034 个黑样本
5137个判黑样本	判黑率 98%	判黑率 100%	103 个非黑

我们选取移动样本进行对比测试，结果评估：

- ✓ AI赋能不影响原反病毒引擎的检出能力
- ✓ 结合AI能发现更多可疑样本，测试出现 2个 新的风险行为样本
- ✓ 利用AI不影响检出能力，但会产生误判，将正常样本判黑

03 存在的问题和风险

AI在移动反病毒中存在的问题

AI模型确实有助于快速筛选和识别移动未知威胁样本，但经过多次测试发现仍有很严重的问题：

极易发生误报

对于小众特定类型的应用，往往由于行为特征的相似性容易识别成威胁应用。

AI在移动反病毒中存在的问题

模型测试

	precision	recall	f1-score	support
B	0.82	0.86	0.84	48265
W	0.86	0.83	0.85	52989
accuracy			0.84	101254
macro avg	0.84	0.84	0.84	101254
weighted avg	0.84	0.84	0.84	101254

模型	检出率	易错误报率
对比基准：逻辑回归	84%	100%
决策树 xgboost	91%	75%
贝叶斯	76%	78%
FNN	90%	75%
Wide&Deep	93%	77%
DeepFM	**98%**	**80%**
xDeepFM	94%	78%

我们进行了多轮AI模型优化迭代，检出率可以达到98%以上，基本不会遗漏；

但我们针对已知非病毒的易错样本也进行了一轮测试，误报发生率都能达到80%。

AI在移动反病毒中存在的问题

对于病毒的直接判定的用户场景而言，这种不准确的问题是致命的。

滥杀无辜

自相残害

04 问题的补救

利用AI模型对移动未知威胁进行快速筛选

针对未知威胁样本善于伪装、传播范围小的特点检出渗透等级低的样本，筛选出高价值样本

- ✓ 基于庞大的底层数据全面比对出可疑样本
- ✓ 预设规则简洁有力，快速圈定疑似未知威胁样本
- ✓ 操作简单

通过移动反病毒引擎分析能力体系做进一步补充

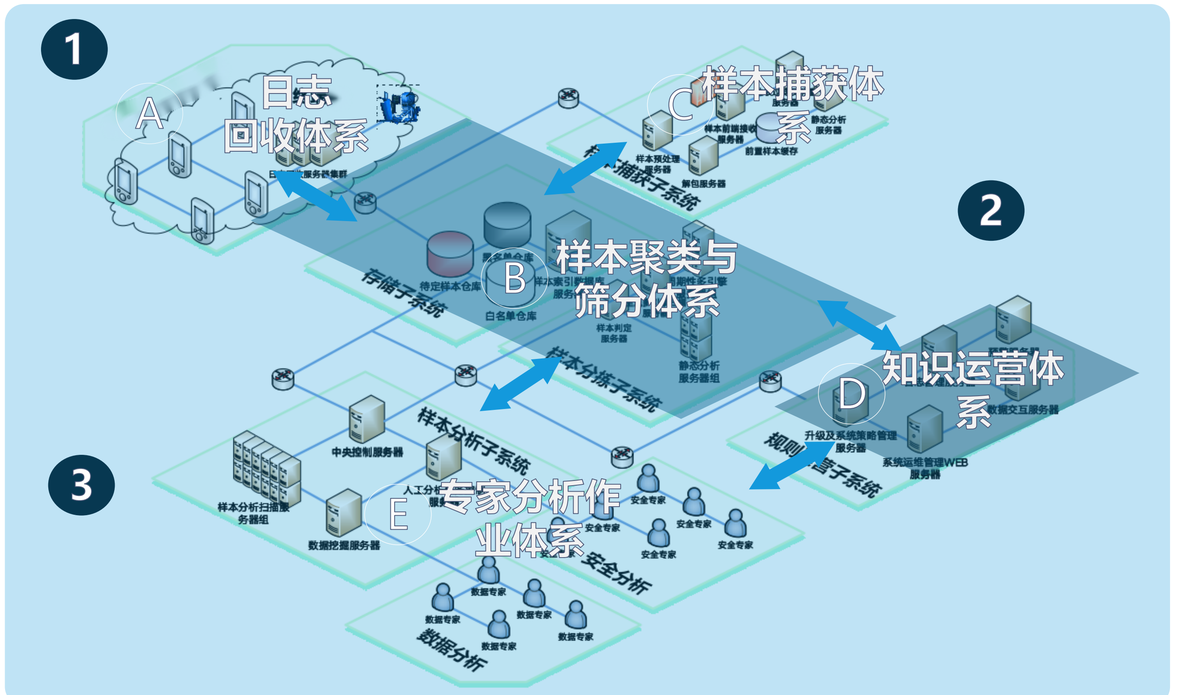
对可疑样本进行进一步分析运营，务实样本的恶意性。

平台人机自动化知识运营体系

- 千万级样本存储容量
- 日处理样本十万级

安全和算法专家作业体系

- 近百人的专家团队（安全分析运维、安全研究和数据科学）
- 日产出聚类规则上百个，辐射上万样本检出
- 对样本多维信息做精细标定



AI是移动AVL引擎中增强型帮手

AI作为移动AVL引擎的增强型帮手，在提升反病毒能力方面发挥着作用。在安天的掌控下，它通过深度学习和机器学习技术，实现了对亿级移动应用样本集合进行高效处理和分析。借助引擎深度预处理能力，AI可以从这些样本中提取出数千种移动恶意样本的维度特征，并为模型训练提供有价值的数据。安天长期积累的移动应用分析运营能力，修正和补救因AI产生的误报。这使得我们能够更及时地发现和应对APT变种的移动未知威胁。总而言之，AI在移动AVL引擎中充当着一个强大的协助者，通过其计算能力和智能化特点，将辅助引擎提高反病毒防护水平，进一步保障移动安全。

网络空间威胁对抗与防御技术研讨会
暨 第十一届安天网络安全冬训营

移动反病毒引擎“玩转”AI

安天冬训营 wtc.antiy.cn

执行体治理赋能与大模型辅助

北向守望