知彼：攻击者情报

蚂蚁集团网商银行信息安全部
分享人 ：柳星（404notfound）

目录

一、自我介绍 Who

二、背景和目的 Why

三、攻击者情报维度 What

四、体系化解决方案 How

五、结果及解读 Do

六、展望 Outlook

Part1 Who am I ?

我是谁 打哪来 往哪去

我是谁

ID：404notfound
名字：柳星，2020年硕士毕业于西安电子科技大学
博客：https://4o4notfound.org
Github：https://github.com/404notf0und

打哪来

工作单位：蚂蚁集团网商银行信息安全部

网商银行

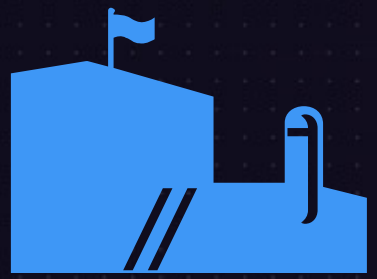
往哪去

研究方向：专注于「安全智能化」

Ai

Part2 为什么做？

现状



纵深防御

- 陈兵边界
- 梯度阻击

攻击检测

- 依赖先验知识
- 对抗已知攻击

伪静态策略

- 安全专家完成检测及响应策略

问题及思考🤔

陈兵边界

被动防御；
视界有限；

→ 更主动？

攻击检测

易被绕过；
对抗未知攻击能力有限；

→ 更高维？

伪静态策略

人力成本高昂；
策略周期滞后于攻击活动；

→ 更智能？

攻击者情报：一个中心、两个基本点

基本点1：数据

- 边界->全网
- 被动->主动

一个中心：人

- 攻击检测->攻击者识别
- 已知攻击->未知攻击

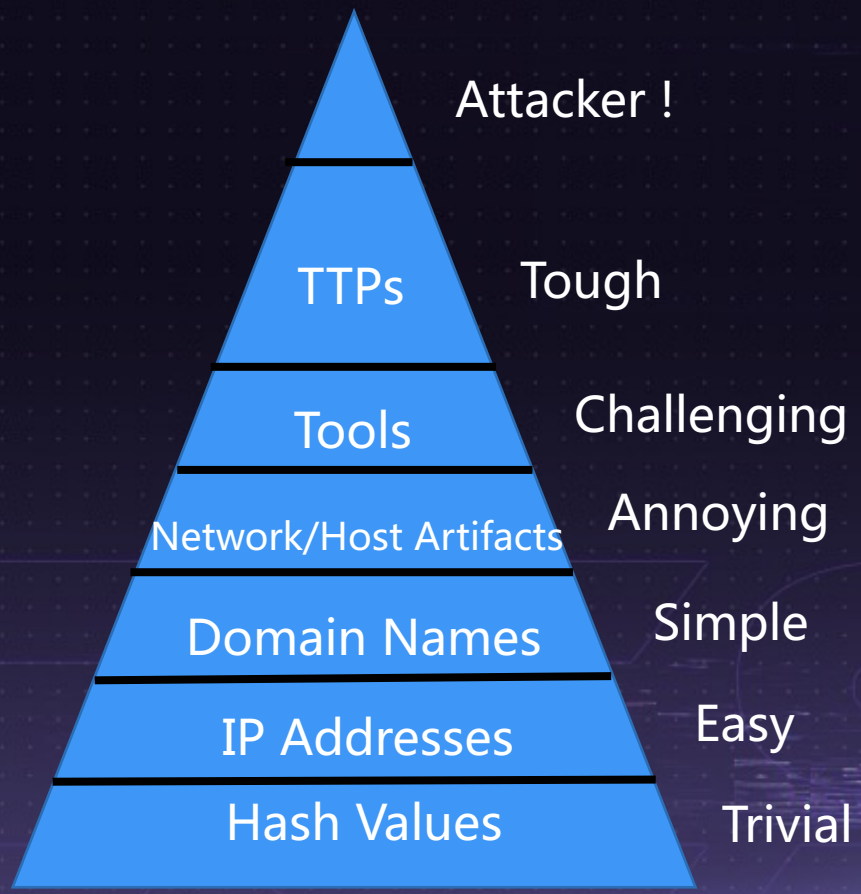
基本点2：算法

- 人工->人工+智能
- 伪静态->可动态

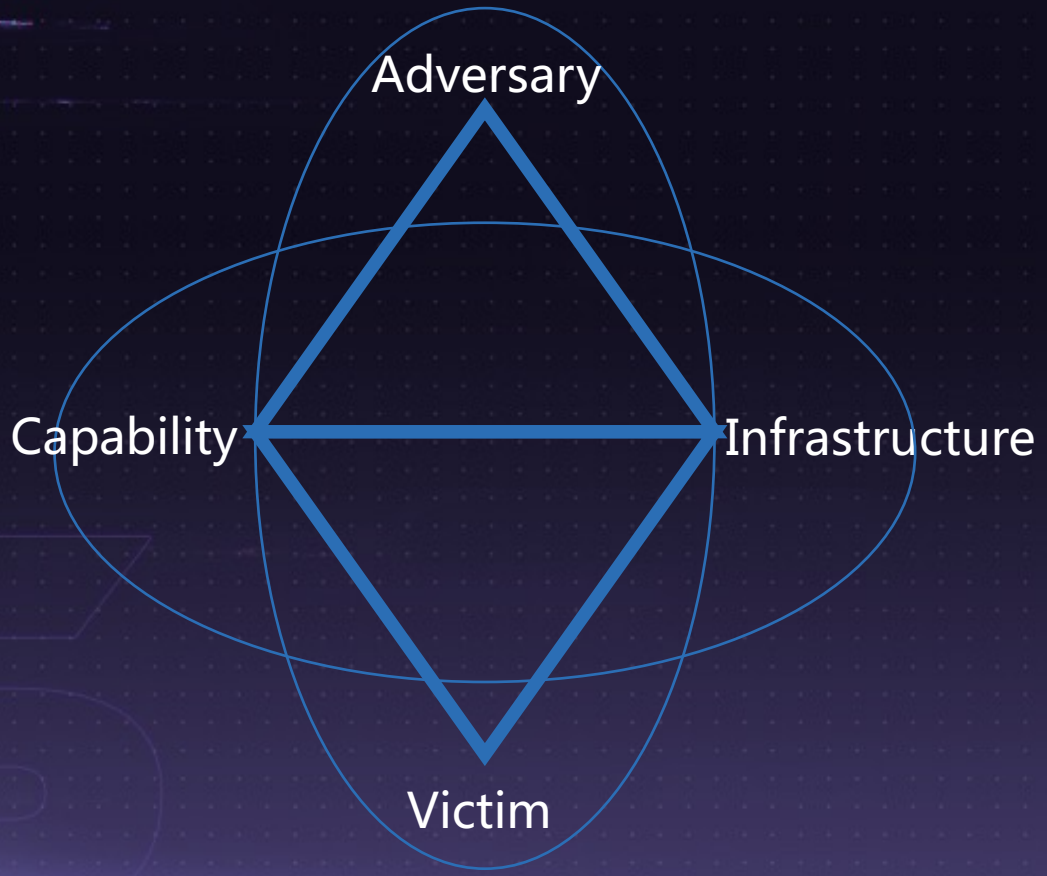
Part3 做什么？

业界相关理论和实践

Pyramid of Pain



The Diamond Model



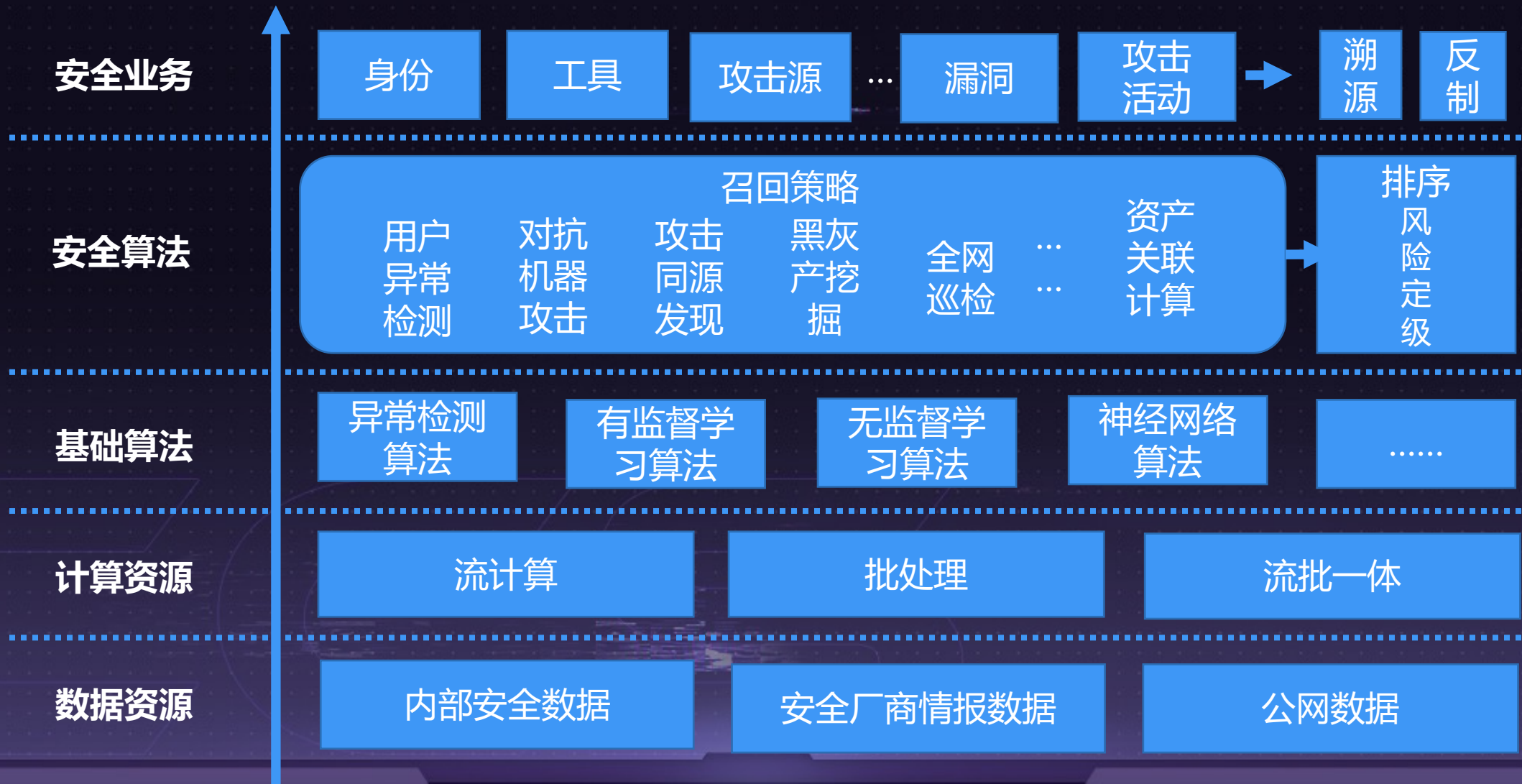
目标拆解

攻击者情报维度：不重、不漏原则

Part4 怎么做？

解决方案

基于数据和算法驱动的攻击者情报体系



1.内部感知

数据

- 全流量数据
- 内部安全产品数据
- 策略数据

特征

- 行为序列词嵌入向量特征：word2vec、bert
- 时间+空间维度：（历史、当前+低阶、高阶）统计特征

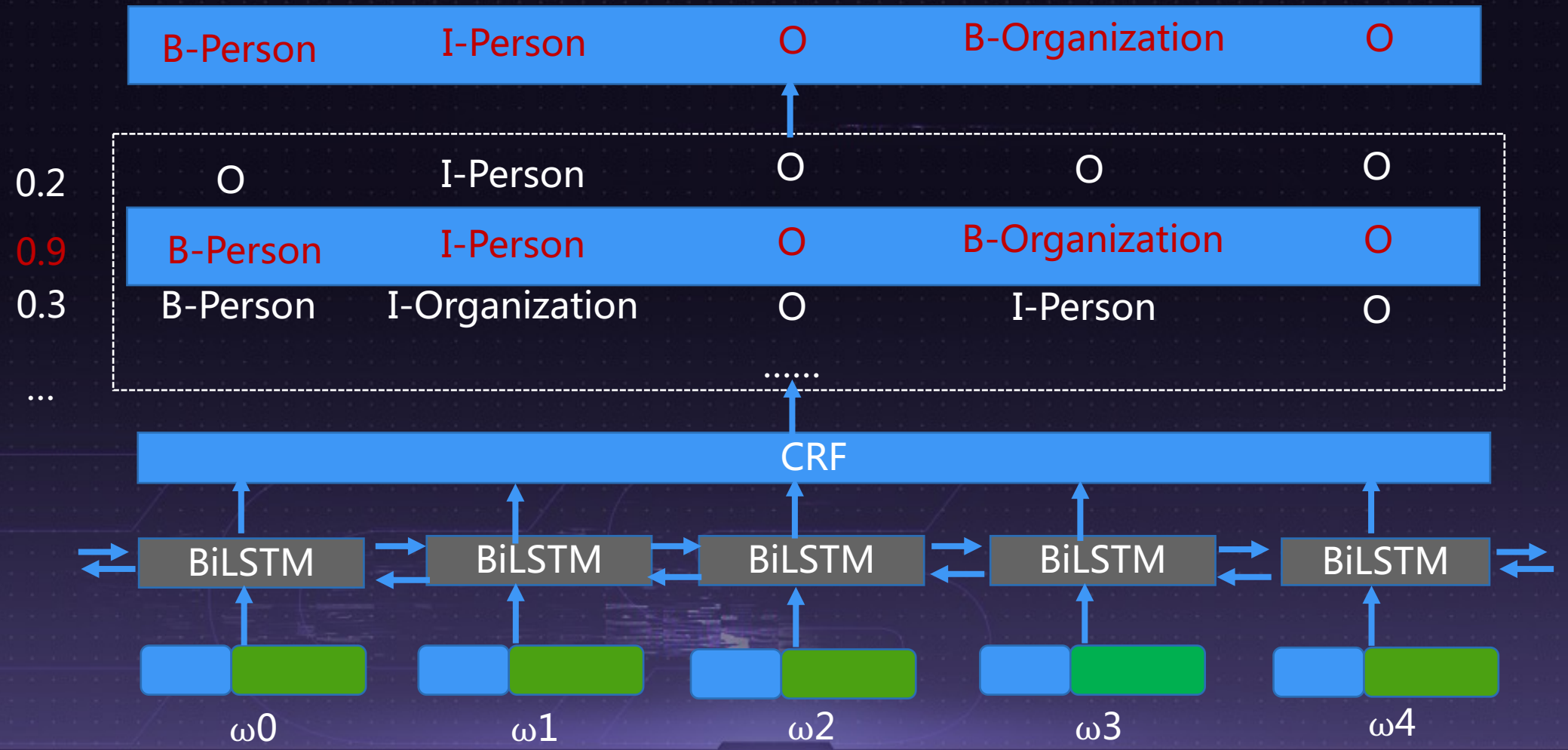
算法

- Lightgbm
- Dbscan
- Iforest

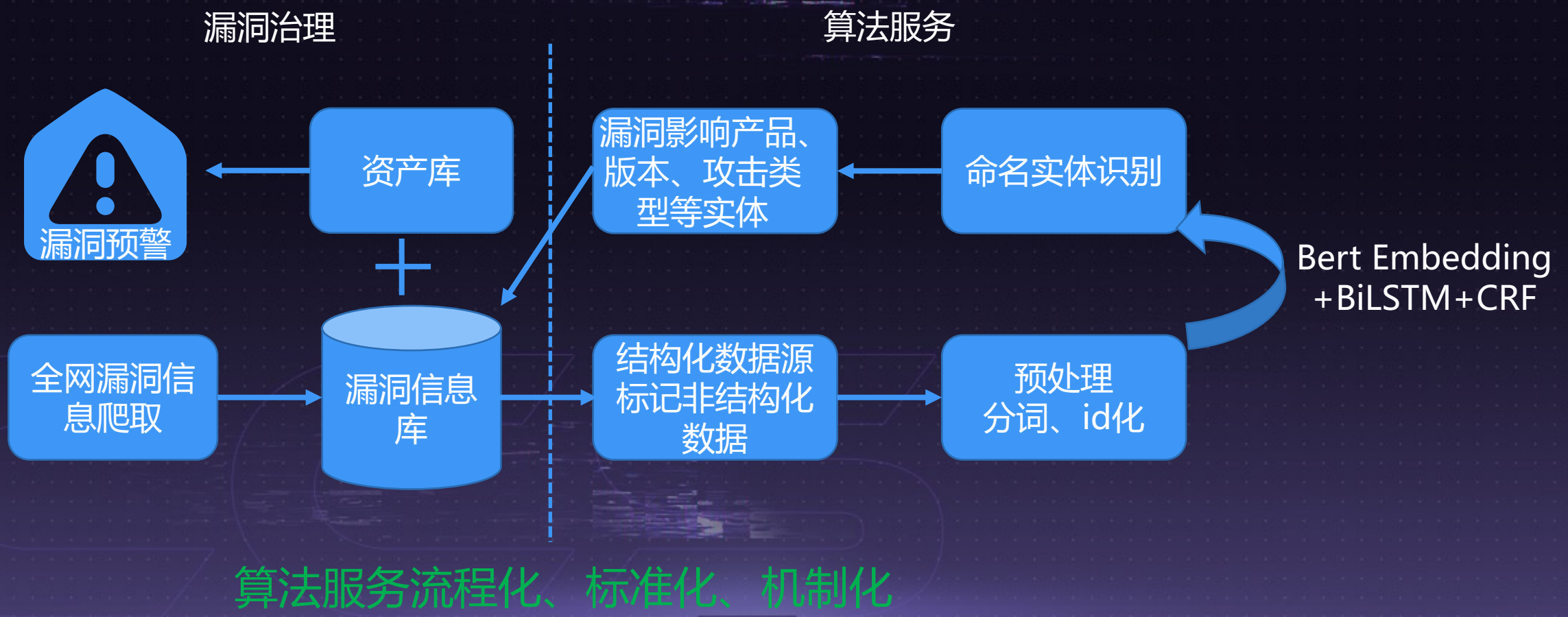
场景

- 对抗机器攻击
- 攻击同源发现
- 用户业务异常行为检测

2. 全网巡检：命名实体识别模型



3. 全网巡检：漏洞情报


漏洞治理
算法服务
漏洞预警
资产库
漏洞影响产品、版本、攻击类型等实体
命名实体识别
Bert Embedding +BiLSTM+CRF
全网漏洞信息爬取
漏洞信息库
结构化数据源标记非结构化数据
预处理 分词、id化
算法服务流程化、标准化、机制化

4. 全网巡检：APT组织情报

Part5 结果及解读

总体效果

维度

- 10+种实体
- 20+种关系

实体

- 身份：数百个攻击者真实身份
- 攻击主体：30+个大型攻击源
- 工具：几十款攻击工具
- ……

底层能力

- 内外部数据：PB级
- 流批一体化计算：千万QPS
- 策略及模型：十余种智能模型

模型效果

内部感知系列模型

- 对抗机器攻击模型：误报率低于0.3%
- 攻击同源发现模型：支撑发现30+个攻击源
- 用户业务异常行为检测模型：历史黑样本重放，威胁排序104/100000+
- 攻击者身份识别策略：准确率>97.5%

全网巡检系列模型

- 漏洞模型：结构化预警时间最快至秒级，召回率有望提升50%

攻击者群体、工具、目的

群体

- 金融级APT组织
- 黑灰产
- 白帽子

最爱的工具

- 机器脚本
- AWVS
- Nuclei
- XRay

目的

- 窃取数据
- 漏洞挖掘

Part6 展望

展望

安全智能化

1. 安全对抗的本质：知识体之间的对抗
2. AI的历史阶段和机会：AI for Science
3. 安全智能化：吸取更多数据，使用更先进生产力，生产更高价值知识

CHENGDU CYBER SECURITY CONFERENCE

CCS 2021成都网络安全大会
CHENGDU CYBER SECURITY CONFERENCE

THANKS

https://4o4notfound.org

root@4o4notfound.org

https://github.com/404notf0und

个人微信

个人公众号