## 2020年CCF-绿盟科技“鲲鹏”科研基金重点资助课题

### 1.新型基础设施和通信网络安全方向

#### 1.1面向云原生的异常检测、监测和防护技术研究

服务网格、无服务和云原生等新一代云计算系统具有短生命周期、面向业务等特性，其安全防护技术也有较大的差异。

此外，工业界对云环境虚拟机或者宿主机的审计功能比较完备，但是对容器化环境的审计能力还没有相关成熟的技术和标准，容器化是下一代的大方向，相关的审计能力和标准化也是重中之重。以Docker、Kubernetes为代表的云原生基础设施和编排体系，其内部的权限管理和审计能力较弱，对组件的角色、访问API权限的控制，以及运维开发对环境的访问等，如不能及时发现异常会出现安全隐患。

研究方向包括但不限于：

1. 研究基于人工智能的云原生环境中的异常检测技术，可对容器、微服务、API和业务的行为进行分析，快速发现异常行为或恶意攻击。研究容器环境和微服务架构下的高效防护机制，实现细粒度的隔离、访问控制和入侵防护等能力；
2. 研究面向容器、Kubernetes的逃逸攻击和防护安全机制。设计并实现审计访问调用的检测机制，发现越权等问题，或对操作者操作资源的记录审核。

研究输出成果：

1. 云原生下的安全异常检测、行为监测和安全防护原型系统；
2. 相关研究报告、高水平学术论文、专利等。

#### 1.2下一代新型网络的安全技术研究

随着新基建开展，5G、SDWAN、天地一体化等下一代新型网络的建设逐步加快，关键基础设施的安全是保障国家安全的基础。

研究方向包括但不限于：

1. 研究面向5G、SDWAN、卫星互联网下的安全风险和安全威胁，研究相应的计算、网络等基础设施安全机制；
2. 研究5G无线协议的安全漏洞和防护方案；
3. 设计卫星互联网的安全通信协议和入侵检测机制；
4. 研究边缘计算的安全检测和防护机制；
5. 研究5G核心网的网元异常业务检测机制和网元安全防护机制。

研究输出成果：

相应新型网络的安全防护原型系统，高水平学术论文，相关研究报告、专利、漏洞等。

### 2.工业互联网和物联网安全

#### 2.1物联网蜜罐攻击检测与情报挖掘技术研究

近年来针对物联网设备的攻击日益增多，攻击手段和方式层出不穷，其攻击意图及利用方式错综复杂。绿盟威胁捕获系统已经积累了海量的威胁数据，为了及时构建面向物联网领域的威胁情报预警与攻击团队画像、溯源等特定需求，亟需开展面向物联网蜜罐的攻击检测与情报挖掘技术研究，尤其及时监测针对全球范围内的物联网攻击事件及其情报挖掘，并在有限的条件下对攻击者或者团伙进行攻击手法或者行为画像。

研究方向包括但不限于：

1. 物联网蜜罐样本分析；
2. 物联网蜜罐攻击检测与攻击团伙挖掘；
3. 基于物联网蜜罐攻击数据的情报自动生成技术研究。

研究输出成果：

高水平学术论文，物联网蜜罐威胁分析原型系统，相关研究报告、专利等。

#### 2.2智能推荐的物联网资产识别技术研究

对网络空间中存在的软硬件资产进行准确的探查与识别，是掌握资产暴露面，感知安全态势、进行威胁处置的基础。随着工控、物联网等新型网络技术普及和发展，新型的网络设备、协议和服务层出不穷，对网络探查技术提出了新的挑战。目前资产识别主要通过人工来进行标记，很难发现新出现和小众的物联网设备，需要不断的维护已知物联网设备信息，投入人力成本也比较高。所以本项目需求研究一种基于机器学习智能推荐的物联网设备识别方法，意在减少人工投入，提高物联网资产识别覆盖度，并对新型资产进行发现。

建议研究方向包括但不限于：

1）物联网资产特征工程研究。对已经标记的物联网资产数据和指纹数据进行分析，提取可以区分物联网与非物联网资产的特征；

2）基于机器学习的物联网资产聚类推荐模型研究。基于机器学习聚类算法提取数据集中的设备种类，结合物联网资产协议特征找出高置信度的物联网设备的类别；

3）智能推荐人工标记系统研究。设计推荐算法，将机器学习算法生成的物联网资产类别结果推荐给人工标记。通过较小的人工投入达到生成大量资产指纹的目的；

4）互联网上新出现的物联网资产发现、监控和识别研究；

5）其他资产类型识别研究。该技术不仅限于物联网资产的识别，对其他类型资产具有普适性。

研究输出成果：

高水平学术论文，相关指纹库和原型系统，研究报告、专利等。

#### 2.3面向物联网资产的威胁监控研究

物联网威胁层出不穷，越来越多的攻击者开始关注利用物联网服务进行DDoS反射攻击，越来越多的P2P僵尸网络被安全研究人员发现。一个有漏洞的固件可能影响多个厂商的产品，脆弱物联网资产被攻击者利用后可组建僵尸网络，僵尸网络的传播范围决定了后续攻击的严重程度，因此，需要有系统性的方法快速对僵尸网络等安全威胁进行监控。

建议研究方向包括但不限于：

1）新型物联网反射攻击发现方法与影响范围监控研究；

2）物联网P2P僵尸网络监控方法研究；

3）物联网固件相关的漏洞影响范围研究。

研究输出成果：

高水平学术论文，原型系统或工具，相关研究报告、专利等。

#### 2.4物联网轻量级认证技术研究

随着智能硬件的兴起，万物互联的时代已经来临。物联网虽然极大方便了人们的生活，但是也将生活习惯等个人隐私信息暴露出来，成为滋生犯罪分子的温床，确保物联网设备身份安全及数据传输安全在物联网应用发展中至关重要。

基于PKI架构的数字证书认证方式可以有效保证用户的身份安全和设备安全。但该方案对网络资源、计算资源、时延、功耗都有较高的要求，不适用于资源受限的弱物联网设备。因此，对于此类物联网设备，需要更合理的轻量级认证方案。

建议研究内容包括但不限于：

1. 研究适用于物联网设备的轻量级认证技术；
2. 研究的轻量级认证技术可以适用于主流物联网设备，包括强终端、弱终端。

研究输出成果：

相关原型系统，高水平学术论文、专利。

#### 2.5物联网卡诈骗预警分析技术研究

物联网卡目前多被应用在移动支付、智能家居、智慧物流、共享单车等行业中。卡商在运营商那里登记后就可以进行批量购卡，其中有的卡并没有被实名登记，而有些被经过层层转卖流入黑心市场，成为违法犯罪分子进行网络诈骗的工具。最近几年出现了多起物联网卡诈骗案件，已成为威胁公众财产安全和社会稳定的一大公害。该方向研究物联网卡诈骗的预警分析技术，以便于能够尽早的发现违法犯罪分子的不法行为。

建议研究内容包括但不限于：

1. 基于运营商的物联网卡话单数据和上网日志，研究物联网卡诈骗使用的技术手段、诈骗手法、诈骗场景，能够有效发现违法犯罪团队的物联网卡违法行为，分析并提供犯罪团伙画像；
2. 研究从上游的卡商，到接码平台、养号、料商、以及最终的电信网络诈骗各环节，分析物联网卡如何一步步被利用于诈骗，可输出某物联网卡的完整生命周期画像；
3. 基于物联网卡犯罪团队画像，对外提供物联网卡诈骗情报；
4. 提供物联网卡诈骗预警机制，可联合多个部门，包括但不限于运营商、应用提供商（可基于物联网卡提供注册）、公安部门，共同打击物联网卡诈骗违法犯罪分子。

研究输出成果：

高水平学术论文，原型系统、相关研究报告、专利。

#### 2.6被动式工业互联网资产识别技术研究

传统IT系统中可以通过主动扫描等方式来识别相关的软硬件的资产。在工业互联网的环境下，大量的控制类设备和物联网设备通过主动式的方式无法识别到或者扫描后导致系统异常，需要研究通过被动式的方式，基于流量特征或者业务行为特征来对资产进行识别。

研究基于业务属性的被动式资产识别技术，可以对工业现场设备、边缘层设备及工业云平台进行识别。

主要研究内容：

1. 研究针对控制器的被动式识别技术，可以识别控制器的类型、版本号、控制器的异常状态等信息；
2. 研究工业中常见的智能传感器的被动式识别技术，可以识别设备类型、设备装备信息，运行状态信息等；
3. 研究针对工业云平台的识别技术，可以基于互联网识别主要的工业云平台厂商信息；
4. 研究常见组态软件的被动识别技术；
5. 可以基于业务特征来预判潜在运行异常所带来的潜在风险。

技术指标：

1. 可以识别不少于5种品牌10种以上控制器，至少包括：西门子、施耐德、AB品牌的PLC不少于5种；
2. 可以识别不少于5种智能传感器；
3. 可以识别不少于5种的组态软件，至少包含：kingview、intouch、WINCC等常见组态软件；
4. 可以识别不少于5种主流的工业云平台；至少包含：indics平台，mindphere平台；
5. 提供基于业务特征的异常判断能力，可以判断系统中潜在的运行风险及提供相关的可度量的风险评价方法。

研究结果输出：

提供原型系统1套、申请发明专利1项、发表不少于2篇高质量论文。

#### 2.7基于数字双胞胎的工业互联网安全研究

数字化双胞胎是指基于现实世界，利用数字化技术营造的与现实世界对称的数字化镜像。数字化双胞胎模型具有模块化、自治性和连接性的特点。利用数字化双胞胎模型，可以从测试、开发、工艺及运维等角度，打破现实与虚拟之间的藩篱，实现产品全生命周期内生产、管理、连接的高度数字化及模块化。针对工业互联网的安全研究需要考虑到安全的部署与业务机理之间的相互影响，目前采用的研究方法采用多控制设备结合的方式，缺乏对于不同被控设备之间关联的分析及业务之间的影响分析，需要通过数字双胞胎的方式来研究不同业务之间的深度分析和处理。

主要研究内容：

1. 研究化工、典型制造业场景下的数字孪生技术在安全领域中的应用；
2. 研究化工、典型制造业的重要工艺段重要指标的数字化，可以对不同参数具可量化的回应结果；
3. 研究融合安全与业务分析的数字双胞胎分析技术，可以复现出安全对业务过程的影响性分析；
4. 研究形成化工、典型制造业业务场景下的安全数字化量化指标结合业务，可以量化分析出安全对业务不同工艺的影响。

技术指标：

1. 形成的数字双胞胎模型系统至少可以适用于化工的两个业务场景，典型制造业的两个应用场景中；
2. 形成业务与安全评价模型，可以适配于化工和典型制造业的场景；

3）形成一套业务与安全融合的评价指标体系；

研究结果输出：

提供1套工业互联网数字双胞胎安全原型系统、申请发明专利2项、发表不少于2篇高质量论文。

### 3.威胁分析和人工智能安全方向

#### 3.1基于告警分析的安全智能决策研究

现代企业通常引入了防火墙、防病毒系统、入侵检测系统等大量安全产品，每天产生的大量告警会发生告警爆炸的问题，分析人员往往应接不暇。此外， APT的攻击过程中也会触发大量告警，而这些告警与常规告警也存在区别，从告警中分析出这些网络攻击行为理论上是可行的。

一般有两种方法应对海量告警：1.提升安全设备的检测能力，减少告警数量，2.对告警进行进一步分析，减少告警数据，并基于告警信息，做出智能决策，助力应急响应。基于告警分析的智能决策是一项前沿的安全技术研究课题，对安全运营的意义深远。

建议研究方向包括但不限于：

1. 告警评估算法研究。对企业产生告警数据进行风险评估，提炼出能反应真实攻击的TopN告警；
2. 基于告警的安全事件构建机制。将告警信息映射到ATT&CK、Kill Chain等网络攻击模型，展现攻击的更高层视角；
3. 攻击链的自动化构建算法。从告警中还原网络攻击在攻击链中的具体步骤，并且将这些事件连接起来，完成基于时序的攻击链自动化构建；
4. 威胁狩猎。从告警中推断能够辅助安全运营人员寻找网络资产环境中的可疑活动和漏洞，以发现潜在的漏洞利用迹象和恶意的攻击者，进而防止网络资产设备被攻破；
5. 基于事件推理的威胁响应机制研究。在不完整观测的条件下进行建模，进行复杂事件推理，做到对网络攻击的提前防御，即主动防御，提前进行威胁响应。

研究输出成果：

1. 针对细化目标中的需求，给出解决方案，包括说明文档和源代码；
2. 高水平学术论文，原型系统、工具，相关研究报告、专利等。

#### 3.2基于大规模威胁情报与基础数据的挖掘分析

现有威胁情报平台往往累计了大量的威胁情报数据和基础数据，与此同时，基于威胁情报的检测系统也进入了《信息安全技术 网络安全等级保护测评要求》，如何从基础数据中挖掘威胁情报，如何分析运营威胁情报数据和如何在检测中应用威胁情报成为威胁情报生态建设的三大主要问题。

恶意样本相关的威胁情报和基础数据是情报建设的基石，结合恶意样本数据（包括但不限于动态信息，通联IP、域名、URL等；静态信息，控制流图、系统调用序列、PDB路径、DLL信息等）和威胁情报（IP/域名/URL的分类与判定信息），将专家经验特征化，分析流程模型化，评判标准数值化，能够极大减少人工标注的工作量，以及产出能够用于检测的威胁情报IOC和TTP。

建议研究内容包括但不限于：

1. 恶意样本的动静态相似度判定；
2. 海量样本中的相似样本检索技术，能够在10秒内于十亿候选样本集中找到相似样本；
3. 海量样本信息中大规模聚类，且具备可行的准实时更新能力。发现高价值样本家族，跟踪团伙/聚类特性，形成能够用于检测家族/团伙特征。

研究输出成果：

相关原型系统、研究报告（或论文、专利）。

#### 3.3基于终端日志挖掘的威胁检测与溯源

终端侧日志中的溯源数据（Provenance）能够记录系统进程、文件、网络等实体的行为细节和依赖关系，在威胁的细粒度分析中具有关键价值。然而，从大规模终端溯源数据中挖掘APT行为，面临诸多挑战：1）依赖爆炸。因技术瓶颈，目前广泛实际部署的粗粒度溯源信息（Coarse-Grained Provenance）虽能忠实记录所有可能的实体行为依赖关系，但实体所有输出流依赖所有时间较早的输入流，导致信息流的方向的不确定性；2）持续动态演变。终端类型多样，端上服务持续更新，数据规模持续增加，导致传统静态行为分析方案失效；3）大量异常误报。高对抗性的APT行为踪迹淹没在海量背景正常行为数据中，异常事件比例小但绝对数目大，导致威胁检测误报率高。因此需要研究有效的数据分析方法，以基于终端侧数据实现威胁检测与溯源。

建议研究内容为：

1）研究溯源数据图（provenance graph）的图模式设计、图数据组织方法，基于图数据库实现终端与网络侧数据的高效关联检索；

2）针对溯源图模型，设计实现基于图分析算法的终端异常检测算法；

3） 面向事件溯源分析场景，设计实现基于溯源数据的威胁溯源模型及算法。

研究输出成果：

1）终端溯源数据智能挖掘原型系统，支撑准实时终端日志关联检索；支撑自动化检测终端系统异常；支撑根据关键线索快速溯源威胁事件；

2）高水平学术论文，相关研究报告、专利。

#### 3.4基于多源图数据关联分析的风险评估引擎研究

大规模安全日志的接入为高级威胁事件的检测、评估、溯源提供了机会，同时也带来了挑战。在多元、多维、大规模日志中识别高级威胁，存在“Needle-in-the-haystack”的难题。数据驱动的自动化威胁分析方法往往存在高误报、低可用性等问题。因此，亟需研究融合专家经验与数据统计规律的威胁分析方法。网络安全数据具有典型的图结构，图分析技术近年来得到广泛关注。安全领域知识图谱、情报数据图、行为关联图等数据图，融合、浓缩了安全知识库、抽象威胁情报、威胁行为等内容，能够自然衔接数据分析引擎，搭建人机智能协同的自动化分析方法。研究基于多源图结构，融合专家反馈，构建威胁、脆弱性的自适应风险评估流程，能够有效降低大规模无效事件告警及低优先级漏洞的影响，提升安全运营中心的事件处理效率。

建议研究内容为：

1）研究安全领域知识图谱与威胁事件挖掘的数据关联方法，构建数据层次模型，实现知识数据、情报数据与威胁事件等多源数据的快速、多跳关联分析；

2）研究基于多源图数据的自适应风险评估引擎，实现

* 1. 自适应的资产脆弱性评估；
  2. 自适应的事件威胁度评估；
  3. 自适应风险评估框架构建，依据风险指标辅助资产脆弱性管理与威胁事件分析，提供依据风险排序的高风险资产列表与威胁事件列表。

研究输出成果：

1）威胁事件评估推荐引擎原型系统，支撑准实时知识库、威胁、情报、资产等图数据关联检索；可提供大规模威胁事件风险评估排序、大规模资产风险评估排序；

2）高水平学术论文，相关研究报告、专利。

#### 3.5合规背景下的数据安全共享与计算关键技术研究

在欧美GDPR、CCPA，以及我国的《网络安全法》等重磅级的法规强力监管下，数据安全关注热度近年来迅速升温，使得全球数字型企业不得不重新审视数据安全与隐私合规的重要性与急迫性。在大数据时代，数据在流动中产生更大的数据价值，数据在开放、共享、交换、发布等场景需求变得越来越多：企业之间有数据共享和交换的需求，政府部门也有通过发布政府数据，实现国家治理的公开透明和服务社会的需求。目前，一些传统的技术，比如加密、访问控制等难以适应数据利用与数据安全的双重需求，亟需研究数据一批适应企业应用环境的新型数据安全技术，赋能数据共享与计算，在满足合规性的同时帮助企业实现数据价值的最大化。

建议研究方向包括但不限于（研究至少涉及三种关键技术）：

1）基于安全多方计算技术与方法，研究两方以上的安全多方计算任务的实用框架；

2）结合同态加密理论与方法，研究第三方云上的数据安全计算；

3）基于联邦学习与边缘计算技术与方法，研究分布式模型训练与模型合并的安全框架；

4）在区块链环境下，研究高效的零知识证明协议；

5）结合机器学习和深度学习，研究可应用在文档、网页等敏感文本环境的智能数据脱敏技术框架；

6）基于差分隐私技术，研究高效、隐私预算易于管理、且适用多种数据查询与挖掘任务的实用框架。

研究输出成果：

相关原型系统，高水平学术论文，研究报告、专利。

#### 3.6基于知识图谱的网络威胁识别与攻击溯源

随着技术的发展，当前网络环境中的威胁无处不在，攻击技术也越来越趋于隐蔽与复杂。当前企业网络侧检查主要全流量，IPS/IDS, 终端侧主要是EDR,蜜罐等，还有现在流行的UEBA，每天会产生海量告警信息，从而安全运营人员无从下手。而在威胁识别过程中单靠单一的告警并不能够确定攻击事件的发生，必须考虑攻击过程的上下文，这就需要从不同数据层面挖掘有效的语义上下文，以实现精准的威胁评估。攻击溯源面对海量的路径，如何从大量的关联网络中找到有效的攻击路径成为攻击溯源所面临的关键问题。知识图谱的发展，为智能威胁识别与攻击溯源提供了有效的解决方法。

建议研究内容包括：

1）基于图神经网络的异构、动态图模型特征表示技术；

2）多源图模型的集成技术，主要研究终端数据、网络层数据和情报数据等不同源数据的集成；

3）基于安全知识图谱的已知威胁与未知威胁识别技术；

4）基于安全知识图谱的关系推理技术实现攻击路径推理与溯源。

研究输出成果：

相关原型系统，高水平学术论文，研究报告、专利。

#### 3.7基于无监督学习的TLS加密恶意流量检测

当前互联网已经进入全站HTTPS的时代，在确保通信安全和隐私的同时，TLS加密流量使得恶意行为产生的流量也被加密保护，由于没有明显的可提取的特征字段，难以进行DPI之类的网络监管分析。现有的基于特征签名的DPI技术以及基于机器学习算法的检测技术依赖大量带标签的样本集，一方面高质量的恶意加密样本数据集难以大量获取，不少恶意程序具有沙箱抗性，另一方面特征的设计和提取存在巨大挑战，未公开的恶意行为流量也难以检测，因此无监督学习的TLS恶意流量检测是一项前沿的安全技术研究课题。

建议研究方向包括但不限于：

1）基于无监督学习的加密恶意流量检测技术和方法，可采用自编码器等；

2）结合无监督学习技术，研究加密恶意流量的检测技术、聚类分析技术、以及降维分析及可视化技术；

3）结合无监督学习技术，研究加密恶意流量的细粒度分类方法，如识别加密流量的类型、恶意加密应用、使用加密通信的恶意软件家族等。

研究输出成果：

原型系统，高水平学术论文，相关研究报告、专利等。

### 4.网络空间安全体系和公共、创新技术方向

#### 4.1区块链安全关键技术研究

区块链具有去中心化共识、不可篡改等优良的特性，在电子货币、共享征信、物流信息、法务存证、和供应链等领域得到了广泛的关注。2019年10月，习总书记在中央政治局第十八次集体学习时强调，把区块链作为核心技术自主创新的重要突破口，加快推动区块链技术和产业创新发展。然而，区块链应用可能并不安全，目前已经出现了多种电子货币交易系统、智能合约等漏洞，造成了较大的损失。所以研究区块链安全是必要的。

建议研究方向包括但不限于：

1）研究区块链智能合约的形式化验证，形成智能合约安全检测工具；

2）联盟链安全关键技术研究，尤其是不同于公有链的特有安全研究点；

3）研究区块链在安全场景中的应用，设计并实现基于区块链的应用系统。

研究输出成果：

高水平学术论文，原型系统或工具，相关研究报告、专利。

#### 4.2基于量子密码的新型技术及应用研究

密码是信息安全的核心基础，传统密码学的安全性本质上是建立在数学计算复杂性假设上，然而，随着计算机科学领域的不断发展与创新，计算机的算力也大幅提高，对传统密码学、信息安全、乃至整个数字化环境的安全性造成了严重的威胁。量子密码是一种将量子力学特性与密码学融合的新型密码技术，其安全性建立在量子物理学的基本原理上，理论上可达到无条件的安全性，成为未来信息安全潜在的一个基石。

量子密钥生成、分发、协商、传输、量子秘密分享等量子通信技术近年来在学术界取得了长足的发展。然而多数工作开展的是理论研究，未与具体的应用场景进行结合。基于此，本课题基于现有的量子网络实验床环境、量子SDK 和相关通信接口，开展融合具体的网络安全应用和实际场景的前沿方向研究，以期开展具体场景方案的探索、研究与应用。

建议研究方向包括但不限于：

1）结合区块链应用，研究和设计基于量子密钥生成、分发、传输的区块链密钥管理协议方案；

2）结合工业互联网终端关键数据升级与更新的背景，研究融合了量子保密通信的关键数据传输协议与方案；

3）结合云存储环境，研究和设计基于量子密钥生成、分发、传输的云数据安全存储方案；

4）结合安全多方计算，研究和设计基于量子密钥生成、分发、传输的安全的多方安全计算框架与方案；

5）结合联邦学习，研究和设计基于量子密钥生成、分发、传输的安全的联邦学习框架与方案。

研究输出成果：

高水平学术论文，原型系统或工具、相关研究报告、专利。