# 第一周

# 项目前期调研

1. **项目背景**

随着民航信息化建设的发展，民航行业对网络和信息的依赖性越来越强，但相对应的安全管理体制不健全、监管制度不完善、安全顶层规划设计缺失和建设滞后使得应对当前日益复杂的攻防对抗局势乏力。为加大网络与信息安全监管力度，需要基于现有安全监管体系开发民航网络与信息安全管理平台。

1. **安全态势感知和威胁感知**

网络安全态势感知（Network Security Situation Awareness, NSSA）是对网络系统安全状态的认知过程，包括对从系统中测量到的原始数据逐步进行融合处理和实现对系统的背景状态及活动语义的提取，识别出存在的各类网络活动以及其中异常活动的意图，从而获得据此表征的安全态势和该态势对网络系统正常行为的影响的了解[[1]](#endnote-1)。

威胁感知是网络安全态势感知系统的核心功能之一，威胁感知基于一定的判别机制，实现对网络中恶意行为的识别和警报。目前威胁感知的基本方法有两种，一种是基于特征的分析方法，该方法基于已知的攻击方式，建立威胁模型，从而识别恶意行为。这种方式易于实施，但面对未知的攻击方式时，其辨识能力差，且无法及时对防御策略进行更新。另一种方法是基于异常的分析方法，该方法通过建立正常行为模型，并以此为基准判断网络行为是否正常，从而分辨恶意行为。该方法很好地克服了第一种方法的局限性，但若基准设置不精确，则会产生严重的误报问题。

1. **国内外研究现状**

网络态势感知源于空中交通监管(Air Traffic Control, ATC)态势感知，是一个比较新的概念[[2]](#endnote-2)。Time Bass于1999年在文献中首次提出网络安全态势感知的概念，其目的是关联相互独立的IDS以融合攻击信息用于评估网络安全[[3]](#endnote-3)。但是他仅限于理论上的研究，未能实现理论模型的应用。2001年，Information Extraction & Transport在研究攻击的检测方法和攻击对网络安全的影响时，为了检测广域计算机的攻击和评估态势响应，开发了一种SSARE工具，将理论方法付诸应用，但是该工具所用方法过于依赖专家主观经验[[4]](#endnote-4)。卡内基梅隆大学开发的SILK (the System for Internet-Level Knowledge)是一个流量分析工具集合。该系统支持高效的采集、存储及网络流量数据的分析，使网络安全分析师快速地查询大型历史流量数据集[[5]](#endnote-5)。

国内在网络态势感知方面的研究起步较晚，相应的研究成果也较少。2005 年冯毅在《中国信息战》中发表文章《关于我军信息与网络安全的几点思考》，提出从包括网络空间态势感知等四个方面入手，研究网络空间信息安全技术[[6]](#endnote-6)。文中提出网络空间态势感知的关键技术主要是网络空间多传感器数据融合技术和数据挖掘技术。华北电力大学李元诚等申请的专利《一种网络安全态势感知方法》，公开了信息安全技术领域中的一种网络安全态势感知方法[[7]](#endnote-7)。宋一非等人于2016年开发的《基于大数据的信息安全威胁感知平台》利用大数据技术实现了对网络安全的监管。吴小雄则进行了基于关联规则的威胁感知的研究，提出了一种新的数据转化的方法以及新的数据挖掘算法[[8]](#endnote-8)。史海滨提出了一种基于机器学习来进行威胁感知的方法，使用了神经网络框架Caffe来实现算法[[9]](#endnote-9)。

网络安全态势感知作为安全研究领域新热点之一，对网络环境内安全要素进行采集、分析，并预测未来发展趋势。网络安全态势感知已在数据融合、评估方法、体系结构等领域取得相应研究成果，但基本属于理论学术层面偏多，实际应用案例偏少[[10]](#endnote-10)。

1. **研究内容总结**

可以看出，随着网络规模不断增大，结构日益复杂，研究方向逐渐偏向利用以数据作为威胁感知的依据，人的主观经验不再占据主导地位，而仅仅提供方向指导以及辅助检测的作用。一方面，因为传统的威胁感知系统，例如防火墙等，大多基于正则匹配等人为设定的规则，这样的方法本身存在比较严重的误报问题，同时在面对未知威胁时也会由于规则更新不及时而失去感知能力，如今网络中攻击手段逐渐多样化，传统的威胁感知方法已经不能完全满足安全需求。另一方面网络规模的扩张使得日志和流量信息不断增加，在数据量足够大的情况下，数据可靠性有了一定保障，而硬件的传输、存储、计算能力的提升，为大批量的数据处理提供了条件。因此，数据挖掘和机器学习技术成为了研究的主流。

数据挖掘技术通过对数据进行人为的特征提取，进一步对威胁数据建立模型，最后从数据中挖掘出一定的关联规则，基于关联规则和威胁模型进行来判断攻击行为。机器学习是一种十分灵活的分类、回归方法，主要通过已经明确分类的数据来建立模型，但也能够应用一些聚类算法来进行初始分类，而根据应用领域以及数据结构的不同，又可以选择使用效果最佳的算法。但近几年的研究表明，其中效果最好的算法，如神经网络，并不具有良好的可解释性，这意味着在追求最佳感知性能的情况下，这种方法不能帮助加强人对攻击行为的理解，而目前来看，若需要进一步加强系统的感知能力，人为的指导是十分关键的，所以在实际应用中还需要做出权衡。另一个问题在于，虽然基于机器学习的威胁感知研究成果不断丰富，但大多数结果来自于公共数据集或实验室环境，在民航网络中的应用需要深入探究。

1. **对项目的个人理解**

根据本人对项目的了解，将主要开发工作内容总结为以下几个部分：

1. 数据采集：对相应数据源进行采集存储，形成原始的安全数据仓库。重点在于数据传输和存储，由于数据量巨大，对采集、传输、存储系统的性能要求较高，并且需要采取妥善的管理措施，以便于后续的开发和数据调用。
2. 数据处理：数据的处理工作可以分为三步，第一步是对原始数据进行清洗，去除无效数据。第二步是进行数据的统一，确定数据的标准后，去除无用信息，并对空缺部分补充默认值等，再将数据用统一格式重新保存。第三步是对数据的丰富，对数据的相关信息进行补充。
3. 数据呈现：利用所得数据对用户和管理员进行反馈，包括威胁感知和警报、数据可视化、资产管理、安全状况评估、安全策略管理等。
4. **个人研究意向**

**目前机器学习已经广泛应用于网络安全的研究中，但针对民航网络实际情况的研究却并不深入，主要原因是数据的匮乏。本次项目进行民航网络的数据采集和融合，并建立数据仓库，这将成为理想的研究资源。本人希望参加本次项目的开发工作，加深对数据的认识和理解，并利用所得部分数据完成进一步的分析研究，主要方向是基于机器学习的威胁感知，内容包括民航网络中恶意行为的识别、分类、预测等。**

1. 龚俭, 臧小东, 苏琪,等. 网络安全态势感知综述[J]. 软件学报, 2017, 28(4):1010-1026. [↑](#endnote-ref-1)
2. 贾美娟, 于海涛, 蒋庆丰,等. 态势感知技术在网络安全中的研究[J]. 科技创新与应用, 2012(9):56-56. [↑](#endnote-ref-2)
3. BASS T. Intrusion detection systems & multisensor data fusion:creating cyberspace situational awareness[J]. Communications of the ACM, 1999, 43(4):99-105. [↑](#endnote-ref-3)
4. DAmbrosio B，Takikawa M，Upper D，Fitzgerald J，Mahoney S.Security situation assessment and response evaluation[C].Proceedings of the DARPA Information Survivability Conf，&Exposition II，Anaheim，California，USA，2001：387-394. [↑](#endnote-ref-4)
5. Carnegie Mellon's SEI.System for Internet Level Knowledge (SILK) .http://silktools. sourceforge.net,2005. [↑](#endnote-ref-5)
6. 冯毅, 《中国信息战》我军信息与网络安全的思考. http://www.laocanmou.net /Html/20056194115-1.html, 2005. [↑](#endnote-ref-6)
7. 李元诚, 井经涛. 一种网络安全态势感知方法:, CN 102098180 B[P]. 2013. [↑](#endnote-ref-7)
8. 吴小雄. 基于关联规则的安全威胁感知方法研究[D]. 南京理工大学, 2016. [↑](#endnote-ref-8)
9. 史海滨. 基于机器学习的具备威胁感知和反馈环的自反击蜜罐[J]. 网络安全技术与应用, 2017(8):29-30. [↑](#endnote-ref-9)
10. 吴鹏, 皇甫涛. 基于APT攻击链的网络安全态势感知[J]. 电信工程技术与标准化, 2015, 28(12):43-47. [↑](#endnote-ref-10)