ICS 03.060
CCS A 11

团　　体　　标　　准

T/BFIA 024—2023

基于交互式风险防控的反欺诈技术指南

Anti-fraud technical guidelines based on interactive risk prevention and control

2023－09－22 发布　　2023－09－22 实施

北京金融科技产业联盟　　发 布

目　次

前　言

本文件按照GB/T 1.1—2020《标准化工作导则　第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由北京金融科技产业联盟归口。

本文件起草单位：中国金融电子化集团有限公司、北京国家金融标准化研究院有限责任公司、蚂蚁科技集团股份有限公司、北京金融科技产业联盟、浙江网商银行股份有限公司、中国工商银行股份有限公司、中国建设银行股份有限公司、建信金融科技有限责任公司、重庆农村商业银行股份有限公司、中金金融认证中心有限公司、北京银联金卡科技有限公司、中国民生银行股份有限公司、中信银行股份有限公司、泰康保险集团股份有限公司。

本文件主要起草人：潘润红、周夕崇、聂丽琴、陆碧波、崔征两、杨倩、白璐、邓琳莹、李俊奎、王维强、赵亮、盛闯、岳汉、彭晋、罗马慧、彭姝雯、吴思捷、应缜哲、陈春宝、徐鹏、姜城、金驰、程佩哲、程元鸿、陈德锋、吴猛、刘洋、何伟明、田成志、卢华玮、李松涛、王洪波、杨阳、杨波、邱晓慧、马艺桂、吕博良、徐晓剑、郭启铭、周星、王允保、宁赋宣、许宝东、陶蓉、李强、洪丹、陈锣斌、孟昌华、张哲、杨重阳、黄本涛、李明艳、李璐、刘昌娟。

基于交互式风险防控的反欺诈技术指南

1 范围

本文件给出了交互式风险防控应用于反欺诈领域的基本原则、框架、功能、模型、技术、实施与监测和应用场景等内容。

本文件适用于金融机构和非银行支付机构及相关机构的风险防控系统应对用户本人操作的欺诈风险（如电信网络诈骗、信贷欺诈、保险欺诈等），本文件同样适用于金融机构和非银行支付机构进行金融消费者风险教育。

2 规范性引用文件

本文件没有规范性引用文件。

3 术语和定义

下列术语和定义适用于本文件。

3.1

反欺诈 anti-fraud

针对欺诈（如电信网络诈骗等）带来的风险和损失采取的一系列事前、事中和事后处理措施。

3.2

风险 risk

不确定性对目标的影响。

注1：影响是指偏离预期，可以是正面的和/或负面的。

注2：目标可以是不同方面（如财务、健康与安全、环境等）和层面（如战略、组织、项目、产品和过程等）的目标。

注3：通常用潜在事件、后果或者两者的组合来区分风险。

注4：通常用事件后果（包括情形的变化）和事件发生可能性的组合来表示风险。

注5：不确定性是指对事件及其后果或可能性的信息缺失或了解片面的状态。

注6：在本文件中，风险主要是指负面影响。

[来源：GB/T 23694—2013,2.1，有修改]

3.3

交互式风险防控 interactive risk prevention and control

针对用户本人操作的欺诈风险、合规风险及其他风险，金融机构和非银行支付机构的风险系统在支付交易时，通过与用户多轮交互的方式识别风险、化解风险的风险防控模式。

注：通过描述功能规范和操作应用实践，提升支付和交易环节的风险防控能力，尤其是反欺诈能力。本标准主要针对欺诈风险的应对。

3.4

机器学习 machine learning

在历史数据中自动发现规律并利用规律对未知数据进行应用（预测）的算法（技术）。

[来源：JR/T 0202—2020,3.6]

4 基本原则

交互式风险防控反欺诈的基本原则包括：

——可行、适用、有效性原则：针对已识别的欺诈风险源，用可操作的风险防控措施，提高效率和效果；

——经济、合理、智能化原则：在欺诈风险防控中充分考虑成本和应用的复杂度，借助智能化的方式实现经济的效果和合理的收益；

——主动、及时、全过程原则：在欺诈风险管理中，主动、及时控制，覆盖交易全过程；

——综合、系统、全方位原则：由于欺诈风险产生的原因复杂，后果影响是多层次、多维度的，风险防控全方位的角度思考，防范死角十分重要。

5 交互式风险防控反欺诈框架

5.1 欺诈风险

欺诈风险指不法分子以非法占有为目的，通过电话、网络和短信等方式，编造虚假信息，通过虚构事实或者假冒身份来实施诈骗，对受害人实施远程、非接触式欺诈，诱使受害人陷入错误认知，最终在错误认知下按照诈骗人的意愿处置资产的风险。欺诈风险是金融领域防范的重要风险之一。常见的欺诈风险分类示例可参见附录A.1。本文件主要针对用户本人操作的欺诈风险防范（如商户合谋、电信网络诈骗等），是对账户盗用、伪冒申请等欺诈场景防范能力的补充。

金融机构和非银行支付机构业务在交易环节面临欺诈风险以及因欺诈带来的合规和其他风险，相似性的合规风险和其他风险（典型示例参考附录A.2和A.3 ）也可参考本文件。支付场景下，基于大数据的支付欺诈风险智能防控可参考JR/T 0202—2020开展。

5.2 交互式风险防控反欺诈框架

交互式风险防控是在风险防控领域内针对用户操作化解风险，结合了人工智能、云计算、大数据等领域的研究和应用，是一种与用户产生互动式链接的主动式智能风险防控体系，是基于风险交互计算等技术对智能风险防控体系的补充。

在当前智能风控引擎和策略中存疑但无法进一步判断是否存在欺诈风险时，可调用交互式风控，进一步识别风险、评估风险，实现更精准的风险防控。交互式风险防控主动发起与用户的风险交互，在互动中获取用户面临的风险信息，在用户合法授权的情况下通过风险模型进行风险识别后，针对性的进行风险处置，帮助用户防范网络诈骗、网络赌博、非面欺诈等风险的识别、评估、触达、交互处置。

交互式风险防控体系通过建设全场景触达渠道能力，打造全链路交互式风险防控能力建设，从而双向传递信息，有效增强风险防控精度和水平。在面对欺诈风险时，通过针对性的欺诈风险交互语料库、

欺诈风险图谱识别等技术支撑，在交易前中后阶段根据风险评估策略调用交互能力完成欺诈风险的进一步识别、评估、触达、解析和处置。基于交互式风险防控的反欺诈框架见图1。

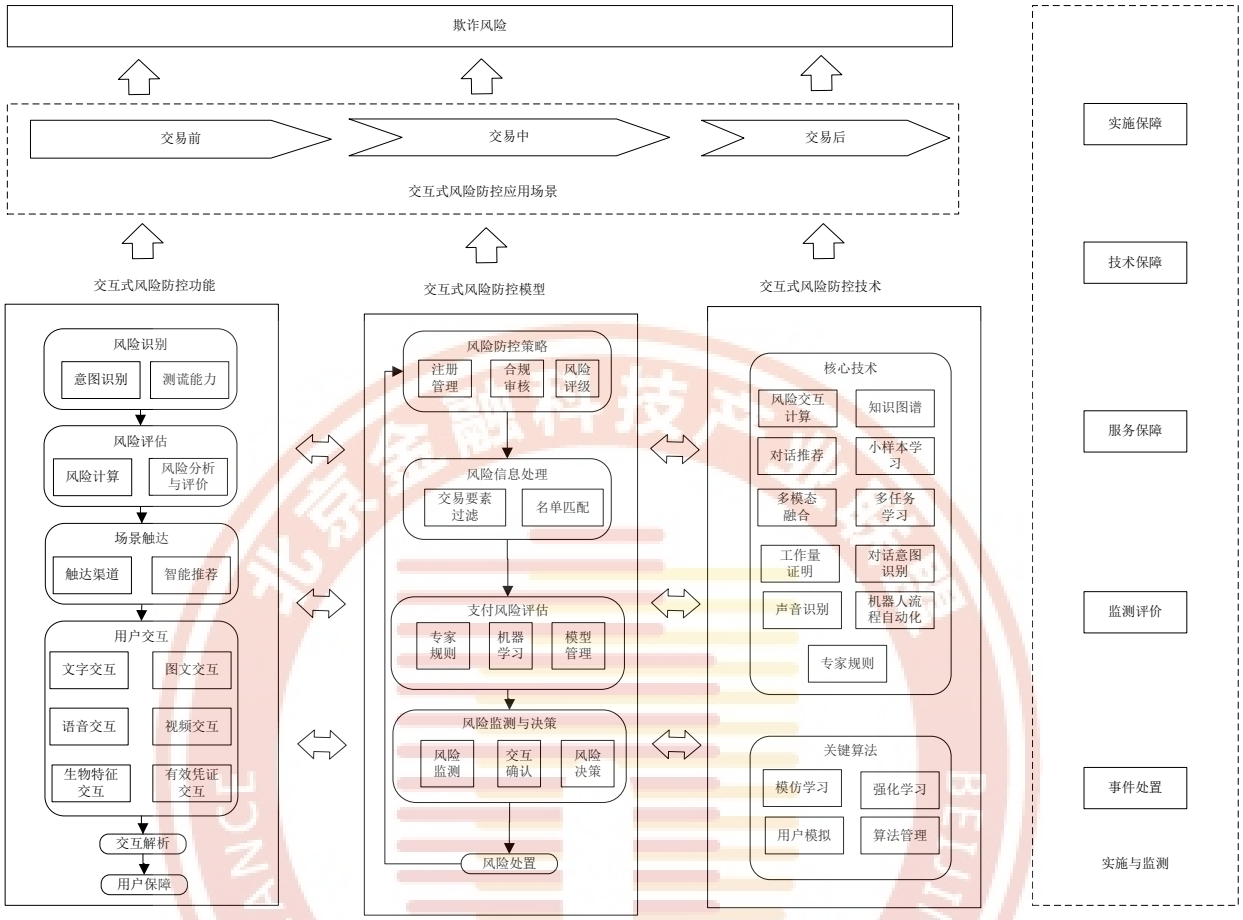

图 1 交互式风险防控反欺诈框架

6 交互式风险防控反欺诈功能

6.1 欺诈风险识别

基于黑样本等已知风险和历史数据表现，通过大数据和人工智能等相关技术建立风险防控模型，在交易前、中、后存在欺诈风险事前预测、事中识别和事后确认。在原有的智能风险防控系统对是否存在欺诈风险信息不足，判断存疑时可调用交互式风控进一步识别。在每一轮交互完成后根据风险评估确定是否进入下一轮交互，进一步开展风险识别。交互式风险防控在欺诈风险识别中最主要的功能是交互时的意图识别和情感识别，通过交互判断交易是否存在主观欺诈意图或被欺诈而产生非常规交易。

——意图识别，指判断用户交易的意图，判断操作者是否存在主观欺诈的意图。根据风险策略识别出的风险标签，通过测谎能力判断交易意图的真实性几率，输出给风险评估。

——情感识别，指采用语音和文本融合的多模态情感识别方案分析用户情感状态及变化。用户被欺诈团伙诱导或者挟持时的交易情绪和正常场景交易情绪存在差异。通过情感识别对用户交易时被欺诈的风险进行识别。

注：黑样本是已经确定为欺诈团伙或确定实施欺诈行为的用户、账号、网址、电话等特征的样本。通过机器学习模型基于黑样本进行训练可以提升风控模型识别率和精准率。

6.2 欺诈风险评估

6.2.1 风险计算方法

风险计算采用定性、定量相结合的计算方法。风险计算方法针对业务面临风险大小进行准确排序，以确定风险的处置策略。风险计算可参考 GB/T 20984—2022 中的附录 F 给出的风险的计算方法。

6.2.2 风险分析与评价

通过风险计算，对欺诈场景的业务实际风险进行综合分析与评价，并对风险计算值进行等级化处理。等级化处理的方法是按照风险评价得分的高低进行等级划分，风险值越高，风险等级越高。

风险分级可参考 GB/T 20984—2022 中 5.4 的风险等级划分方法。

6.3 场景触达

场景触达提供触达渠道和触达方式的灵活配置，根据欺诈风险场景的特点配置调用策略和适应的触达渠道。其功能包括但不限于：

——根据欺诈风险场景需要，设定触达用户的节点、途径。触达渠道包括但不限于电话、APP内消息、短消息、H5（超文本标记语言5.0）、小程序、网站弹窗等；

——根据欺诈风险场景需要，设定触达用户的方式，包括但不限于：智能外呼、智能图文、智能调查、生物特征、有效凭证等；

——根据交互的目的和需求，能在用户面临欺诈风险的环节自动唤醒发起，智能推荐触达方式和触达渠道，触达用户不再局限于交易，全场景可触达；

——提供风险级别与触达渠道的策略配置和智能推荐能力。对于高风险的情形采取语音外呼、生物特征、有效凭证核实等交互方式，对于中低风险可采取短信、邮件、弹窗提示等交互方式。

6.4 用户交互

明确安全交互服务目的，确定交互的形式和内容，从而传递安全信息，获取用户有效信息。交互方式可采取单向和双向交互。

——单向交互主要是对风险客户进行告知提醒，双向交互是双向传递信息，完成客户身份与真实意愿确认，通过多次交互全面保护可提供更好的用户体验。

——双向用户交互根据风险分级开展：通过定义风险等级，对不同级别风险采取不同的交互策略，可以达到在交互式风险防控过程中取得安全与易用的平衡，降低风险的同时减少对用户体验的影响。高风险级别的风险可以采用多因素交互的方式，中低风险可减少交互因素，采用单因素交互。

为了防止欺诈事件对静态固定风险防控交互方式的突破，还可以采取多因素动态组合交互的策略，能够最大程度提升用户体验，同时能够更好的提升风险防控效果。通过智能推荐的形式提供对风险的交互，形式包括不限于文字交互、图文交互等形式。相关功能可组合后在反欺诈场景应用（反欺诈场景应用见第 10 章）。

——文字交互。通过短信、APP文字提示、邮件、问答等文字形式展示、提示或确认风险信息。反欺诈场景应用时通常在交易前组合在图文教育、社交提醒或者交易中的弹窗提醒中，交易后的邮件通知等环节。

——图文交互。通过图片形式将风险内容或特定信息从文字转换为漫画、文字图片等图形，对用户进行提醒、确认交互，提升管控有效率，降低打扰。反欺诈场景应用时可在交易前通过图文教育的方式提示风险，交易中使用图文提醒暂时中断交易。

——语音交互。通过语音和电话触达用户，交互内容为风险防控模型推荐内容，用户交互方式为语音问答，语音内容来自于欺诈风险模型的语料库，由交互式风控系统自动外拨生成。语音交互通常在交易中和交易后配置应用。

——视频交互。通过视频触达用户进行更深一层的交互，用户交互方式为视频问答，内容来自于欺诈风险模型的语料库，由交互式风控系统自动外拨生成。视频交互通常在交易中和交易后配置应用。

——生物特征交互。通过人脸识别、活体识别、指纹、声纹、脉搏等方式对用户真实身份进行确认交互，防控非本人操作、假冒等欺诈风险。生物特征交互一般用于交易前确认是否由本人操纵账户，防范盗用、假冒等欺诈风险。

——有效凭证交互。通过身份证、数字证书、交易密码、动态令牌、短信验证码、小额打款、原始凭证等方式进行多因素确认交互，防控非本人操作、假冒等欺诈风险。有效凭证交互可在交易前历史交易确认中配置应用，也可内置在交易中的智能答题、问答提醒或语音外呼的交互中。

6.5 交互解析

基于意图模型、测谎模型、情感模型等模型算法，对意图、行为和真实性等交互信息，在风险信息识别的基础上对交互风险进行解析和评估。

6.6 用户保障

在用户保障方面，注重产品与服务相结合，服务升级体验和风险防控能力双提升，并根据风险交互解析的结果，进行全方位的用户保障，对面临欺诈的用户和未面临欺诈的用户采取不同的保障方式，具体内容包括：

——对于面临欺诈风险的用户，进行风险处置，保障用户资金；

——对于未面临欺诈风险的用户，快速通行，让交易或操作能继续进行。

7 交互式风险防控反欺诈模型

交易阶段的反欺诈风险防控模型包含风险防控策略、风险信息处理、支付风险评估、风险监测与决策、风险处置等五个模块。通过大数据、机器学习等技术建立满足要求的风控模型，进一步加强对欺诈风险的事前预测和事中识别的能力。同时通过合理引入多个模型、强鲁棒性的模型、抗AI攻击的模型等方式，提高模型评分的稳定性。支持多渠道、多维度的数据整合，形成机构内统一的风控系统。

——风险防控策略作为风险防控的第一道屏障，通过注册管理、合规审核、风险评级等方面的控制，对潜在风险进行初步分辨。

——风险信息处理包含交易要素过滤和名单匹配等，将过滤所得信息输出到支付风险评估模型中。

——支付风险评估从模型方法、模型管理等方面设计模型，完成对潜在欺诈风险的识别、分析和评价。

——风险监测与决策，根据风险模型的计算结果，结合业务要求，采取阻断、挂起、预警、批准等不同的决策行为。

——风险处置在决策的基础上，开展风险调查、关联排查、案件协查等，其结果可以优化风险防控策略。

在支付环节的风险防控模型参见JR/T 0202—2020中4.3，当需要进一步识别和评估欺诈风险时可调用交互式风险防控系统中的相关模型进一步识别。

8 交互式风险防控反欺诈技术

8.1 核心技术

支撑交互式风险防控模型应用于反欺诈场景的核心技术主要包括风险交互计算等多种技术。

a) 风险交互计算：将人机交互应用到风险防控领域，通过风险信息交互，获取所需风险信息，并实时计算出风险等级和风险类型。
b) 知识图谱：通过知识表示、提取、融合、存储以及推理等，解决交互中问答、检索、推荐等难题，确保交互流程不生硬；通过知识库的维护，可以更新风险概念及诈术手段。作为一个对话算法的补充输入源，可以起到快速改进对话效果的作用。
c) 对话推荐：在设定交互完成度、风险浓度、止付率等目标下，对话推荐算法自动进行对话内容的选择和对话路径的推荐;用对话话术选择的算法选择合适的话术进行输出,完成交互的目标。
d) 小样本学习：小样本学习是一类机器学习问题，其经验中仅包含有限数量的监督信息。
e) 多模态融合：将多种模态的信息，包括：文本、图像、视频、音频等数据融合分析的技术。
f) 多任务学习（Multi-Task Learning,MTL）：多任务学习是一个非常重要的机器学习模式，它旨在用其他相关任务来提升主要任务的泛化能力。简单说来多任务学习是一种集成学习方法（ensemble approach），通过对几个任务同时训练而使得多个任务互相影响。
g) 工作量证明（Proof-of-Work，PoW）：一种对应服务与资源滥用、或是拒绝服务攻击的经济对策。
h) 对话意图识别：在对话过程中识别用户的意图，主要有两个作用，第一是对用户的意图进行响应，并完成对话，第二是沉淀下来作为对话结果，并提供给使用方，做后续决策。
i) 声音识别(Automatic Speech Recognition,ASR)与拟人发音（Text To Speech,TTS）：通过机器学习算法将人的声音转化为文本及将文本转化为智能客服语音的算法工作。声音识别效果的好坏可以通过错字率评估。
j) 机器人流程自动化（Robotic Process Automation，RPA）：一种根据预先设定的程序，通过模拟并增强人类与计算机的交互过程，执行基于一定规则的大批量、可重复性任务，实现工作流程自动化的软件或平台。
k) 专家规则：根据专家经验制定检测规则，当交易信息与规则匹配后执行相应的业务策略。

8.2 关键算法

8.2.1 概述

支撑交互式风险防控模型应用于反欺诈场景的关键算法主要包括模仿学习、强化学习、用户模拟和算法管理。在服务平台后通过分析用户提交的交易相关的多维信息为后续的交互提供判断。在不断的交互过程中，算法逐渐收集所需的决策依据，最后输出该笔交易的管控手段。

8.2.2 模仿学习

在对话场景下使用行为克隆来训练模型，从人工对话里学习更好的对话策略。

交互式风险防控使用两阶段行为克隆框架，基于行为克隆的机器人框架，整个系统是个对话系统（Dialogue System，SDS），基于策略克隆和话术克隆的模块完成自然语言理解（Natural Language Understanding,NLU）、对话管理(Dialogue Management,DM）和对话生成。策略克隆模块主要对当前对话上文进行识别，并给出合适的对话策略；基于对话策略，模块会将相应的话术返回给话术克隆模块，并选择得分最高的话术作为当前对话状态下最合适的应答。给定有限动作集合后，可使用语义相似度模型或者文本分类模型，对人工对话每一轮去关联构造出相应的动作。

8.2.3 强化学习

传统训练模型的方式主要是通过人工数据标注的方式教育模型如何响应用户的操作，而实际场景中，人通常无法考虑到足够全面的信息，所以可使用模型自学习探索的方式完成训练，通过强化学习的技术训练动作模型。强化学习会在模型给出所有召回后，开展一个排序工作，针对所有的召回话术提供相应的Q（reward）。

训练一个机器人响应用户的操作，同时训练一个评价角色去评价每一次响应的效果好坏。当这个机器人返回了下一步动作，并给到真实环境时，真实环境的用户会给出一个反馈，交互式风险防控系统从中获得了回馈，在反欺诈风险防控场景中这个回馈通常指用户是否被欺骗，或系统是否控制住了资金损失风险。每一次训练评价角色都会将回馈反馈给这个机器人，让它不断修正自己的算法。

8.2.4 用户模拟

对话模型的学习除了真实样本外，通过用户模拟技术进行学习。具备专家经验的用户模拟模型主要涉及文本生成、决策规则等技术。通过用户模拟技术可以自定义模仿用户的响应动作，扩充训练数据。

8.2.5 算法管理

交互式风险防控技术的应用算法充分考虑算法可能带来的技术风险，从合法性、安全性、可解释性、精准性与性能、伦理等方面对算法进行评估。安全性、可解释性、精准性与性能方面，具体评价标准可参考JR/T 0221—2021的要求，评估达到要求后可上线。具体内容包括算法合法合规等以下多个方面。

——算法合法合规：滥用或恶意使用人工智能算法将会给物理世界和国家社会带来巨大的负面影响。算法应用目标符合国家法律法规要求是十分必要的。

——算法安全性：对常见攻击的防范是至关重要的，如窃取攻击、投毒攻击、对抗攻击、后门攻击、逆向攻击、供应链攻击，算法依赖库安全，算法的可追溯性，算法的内控合理也需要重点考虑。

——算法可解释性：算法可解释性为诊断、发现、修复算法模型内在缺陷提供指导，是算法安全管控的基础。因此，算法以人类可以理解的方式提供对其行为和结果合理性、准确性的解释是算法管理的重要因素。

——算法精准性与性能：算法泛化性良好，保证线下精准性与线上真实场景下精准性的差别在可控范围内是十分必要的。

——人工智能伦理：训练数据失衡、算法设计有误等原因可能导致产生带有偏见歧视的决策，损害国家社会公平正义。因而，算法兼顾各类群体的特征信息是十分必要的，避免对特定人或群体做出带有歧视和偏见的决策，将伦理道德融入人工智能应用的全生命周期，促进公平、公正，避免偏见、歧视等问题。

9 反欺诈实施与监测

9.1 配置实施

9.1.1 概述

对于交互式风险防控，在面对欺诈场景时配置相应的风险防控策略以及交互策略。实时防止服务滥用至关重要，通过控制用户接收安全服务的数量，防止用户产生服务疲劳以及防止对用户进行不必要的打扰。

9.1.2 服务划分

以服务本身的是否应通知用户，将服务类型划分为两种，内容包括普通服务和特殊服务。

——普通服务：应触达用户的服务，包括提醒类、教育类服务等。

——特殊服务：应要求用户接收的服务，包括审理触发、用户主动设置等。

9.1.3 风险防控规则

规则内容示例包括普通服务和特殊服务，具体内容包括普通服务和特殊服务。

——普通服务，具体内容包括：

- 同一产品，同一服务码，每天只会成功触发1次；
- 同一产品，不同服务码，每天可以成功触发3次；
- 反作弊、客户风险、赌博套现这三个域单独计数；
- 不同服务码，每天可以成功触发2次。

——特殊服务：特殊服务没有触发次数限制。策略使用上，可根据服务数据，判断是否需要减少重复输出。

9.2 技术保障

9.2.1 个人金融信息保护

交互式风险防控过程中涉及到用户个人金融信息（包含个人隐私数据）的采集、传输、使用、展示等数据全生命周期的内容，宜参考JR/T 0171—2020、JR/T 0197—2020等相关个人金融数据的安全保护要求。

9.2.2 技术安全保障

交互式风险防控相关系统参考GB/T 22239—2019对三级系统的要求，安全技术要求参考GB/T 22239—2019中8.1.1安全物理环境、8.1.2安全通信网络、8.1.3安全区域边界、8.1.4安全计算环境相关要求以及8.2云计算安全扩展要求、8.3移动互联安全扩展要求、8.4物联网安全扩展要求以及8.5工业控制系统安全扩展要求。

9.3 服务保障

基于交互式风险防控的反欺诈技术体系宜建立相应的服务保障体系，包括但不限于制度保障、组织保障、人员保障，服务宜注意以下内容：

——建立基于交互式风险防控的反欺诈的制度体系，制定相关管理规定、操作规程和技术服务规则，覆盖反欺诈相关领域和流程环节；

——建立由专属部门牵头的组织结构，统筹相关业务和技术的发展和管理。成立专门的团队或部门组织开展相关业务系统的研发、信息安全和科技管理工作；

——组建服务保障专业人才队伍，配备充足的客服人员，并持续开展培训和服务质量管理，保障服务水平。

9.4 监测评价

交互式风险防控体系可建立安全风险监测和评估，定期或在系统运行环境等发生重大变更时开展风险评估，并定期开展安全评估情况抽查，相关情况见表1。

表 1 交互式风险防控监测评估表

名称	评估项	评估要求	评估方法
交互输出风险类别	输出风险的细分级别	1.一级：能从交互识别细分风险，例如刷单可以细分为返利/抽奖，杀猪盘可以细分为色情交友/引导投资等	——

表 1 交互式风险防控监测评估表（续）

名称	评估项	评估要求	评估方法
		2.二级：能从交互识别具体的风险大类，例如盗用/刷单/杀猪盘等 3.三级：仅能从交互识别有无风险	
交互输出风险准确度	输出风险的准确度	1.一级：准确率高于90% 2.二级：准确率高于80% 3.三级：准确率高于60%	可通过人工对对话抽检判断，对人与机器对风险的判断是否一致进行评估
交互对话效果	对话时长和用户配合度	1.一级：能进行平均超过1分钟的对话，1分钟用户配合度高于70% 2.二级：能进行平均超过30秒的对话，30秒用户配合度高于70% 3.三级：能进行平均超过30秒的对话，30秒用户配合度高于50%	用户配合度定义为一定时间内用户没有挂机，不包括未接通的用户
交互对话话术	智能交互系统实际应用包含的话术类型数量	1.一级： 话术类型多于50种 2.二级：话术类型多于30种 3.三级：话术类型多于10种	每1000通呼出电话的话术总类型数，不宜包含：1.系统内存在，但是算法不能推荐播报的话术。2.通过替换姓名等参数产生的新话术
交互对话测谎能力	是否能针对用户的谎言识别出风险	1.一级：能召回50%以上谎言 2.二级：能召回25%以上谎言 3.三级：完全没有谎言识别能力	需要业务有交互后使用真实风险标签进行评估，例如业务中有被欺诈的报案数据
交互后输出数据维度	输出数据丰富程度	1.一级：用户所有表达意图、用户风险情况、整体对话总结 2.二级：用户所有表达意图、用户风险情况 3.三级：无输出数据，知会型对话	——
交互可解释性	算法决策的可解释性	1.一级：可以输出算法对话/风险判断等所有模型决策依据，到用户表达词粒度级别 2.二级：仅有风险情况判断依据，到用户表达句粒度级别 3.三级：没有可解释系统	——
交互可调整性	在交互系统出现遗漏风险时及时调整能力	1.一级：运营可以在1天内调整交互系统，完成新风险的应对 2.二级：运营可以在15天内调整交互系统，完成新风险应对 3.三级：需要技术研发进行开发应对，上线时间超过30天。	——
交互触达率	通过各种手段与用户取得沟通的能力	1.一级：80%以上 2.二级：60%以上 3.三级：40%以上	该指标针对用户群体不同会产生差异，同时不宜为提高触达率降低用户配合度

表 1 交互式风险防控监测评估表（续）

名称	评估项	评估要求	评估方法
交互效能自检	既定的算法自我升级能力注：根据呼叫用户不同，呼叫地域不同等。	1.一级：主动发现交互中效能不好的案例并改进 2.二级：主动发现交互中效能不好的案例 3.三级：无效能自检能力	——
新型风险发现	主动发现新型风险的能力	1.一级：主动发现新的风险并输出具体手法 2.二级：能向人工预警可能存在未知的风险 3.三级：无新型风险发现能力	——

9.5 事件处置

9.5.1 事件分类

根据组织的业务情况和面对的风险状况对突发事件进行分类，突发事件包括但不限于：

——系统安全事件：包括但不限于木马病毒、网络攻击、系统入侵、系统漏洞等导致业务中断、影响业务及平台系统安全的事件；

——数据安全事件：包括但不限于信息泄露、数据盗取、个人信息未经授权获取滥用等导致平台或用户产生大额资金损失及业务负面影响、监管问责及处罚等重大影响事件；

——技术风险事件：包括但不限于技术原因导致的系统连续性、系统可靠性以及业务服务能力下降，且短时间内难以恢复，将影响正常运营或造成重大损失的事件；

——物理安全事件：包括但不限于电力中断、网络损坏或硬件设备故障等事件；

——业务安全事件：包括但不限于用户利用虚假身份、虚假资料方式，通过破解、篡改、劫持、伪造等手段，实施账户盗用冒用、交易作弊、欺诈、洗钱等，引起公司或客户资金损失、客户投诉、公司法律诉讼、监管处罚、公司负面舆情的事件。

9.5.2 事件处置方法

对于欺诈事件的处理，宜建立规定统一的应急预案框架，包括启动预案的条件、应急组织构成、应急资源保障、事后教育和培训等内容。根据不同的事件分类建立事件处置预案，制定应急处置流程并定期演练。

10 交互式风险防控反欺诈应用场景

10.1 概述

交互式风险防控反欺诈的总体框架参见第5章。交互式风险防控可应用于反欺诈场景中交易发生的前、中、后的三个阶段，覆盖交易的全流程和全部链路。通过与用户的同步或者异步多次交互，双向传递信息，加强用户风险防范意识，保障用户安全权益。当识别到疑似欺诈风险时，在交易的各阶段根据场景常用的交互式风险防控见图2。交互式风险防控的反欺诈实施案例可参见附录B。

为应对不同类型的欺诈，交互产品根据风险标签，依据服务策略、服务内容、服务模板、服务数据和服务接入进行智能组合，在合适的节点接入，输出给用户，产生合适的互动，相关内容及对应的关键技术点总体参见第8章交互式风险防控反欺诈技术，具体内容如下：

——服务策略：基于对话意图识别、声音识别等技术开展风险交互计算，确定交互式风控服务产品模块的选择和链路；

——服务内容：基于多任务学习、小样本学习等技术对识别欺诈风险和欺诈风险教育的关键点进行学习并不断沉淀知识库；

——服务模板：基于多模态融合和对话推荐等技术定义智能交互的服务模板，根据场景推送适配的交互内容和形式；

——服务数据：基于知识图谱、对话推荐等技术从服务内容中抽取在交互中完成有逻辑的多轮互动内容衔接；

——服务接入：基于专家规则、对话意图识别、风险交互计算等技术选择并完成交互式风控接入的时间和链路。

交互式风险防控反欺诈的整体运行过程由交互式风险防控反欺诈模型阐述，具体内容参见第7章。

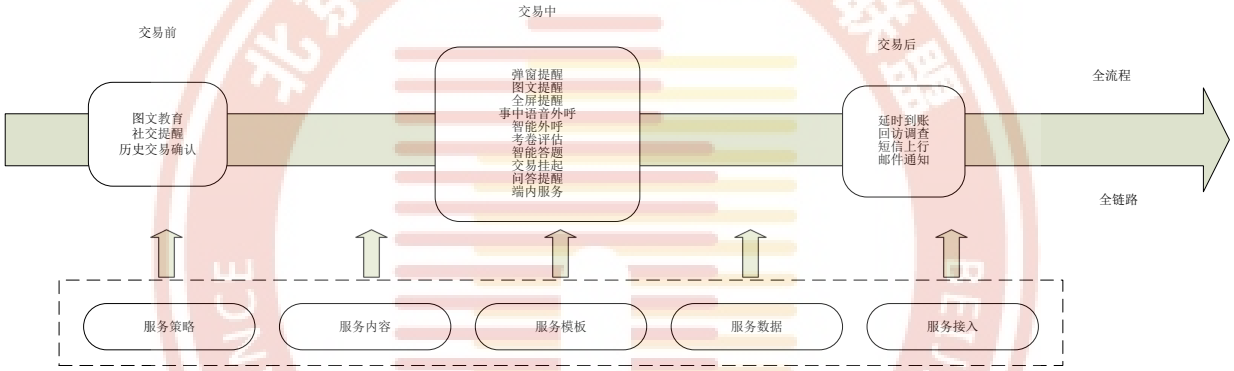

图2 全链路交互式风险防控

10.2 交易前

“交易前”指在用户操作的资金流转发生之前识别到潜在的欺诈风险，此时宜根据风险情况输出合适的事前交互产品，与用户进行双向信息传递，从而完成风险释放。“交易前”交互不影响用户的操作，不需要在线等待用户反馈结果。用户反馈回来的信息，实时应用于此后的风险判断。“交易前”的交互式风险防控的反欺诈功能包括风险识别、风险评估等多个方面，具体内容参见第6章。在交互过程中的实施与监测相关内容，参见第9章，例如在交互环节的服务划分参见9.1.2，风险防控规则参见9.1.3等。

“交易前”交互式风险防控方式包括但不限于：

——图文教育：通过服务提醒、推送等渠道触达用户，引导用户访问个性化的图文教育页面，并反馈相应风险信息，进行风险释放；

——社交提醒：通过APP内社交对话聊天窗触达用户，将风险信息传递给用户；

——历史交易确认：通过服务提醒、push推送等渠道触达用户，引导用户访问风险交易信息页面，并反馈相应风险信息，进行风险释放。

10.3 交易中

“交易中”指在用户操作资金流转发生时识别到交易风险，此时宜输出合适的事中交互产品，与用户进行双向信息传递，从而完成风险释放。“交易中”交互会影响用户的操作，且需要同步在线等待用户反馈结果，需要与用户进行同步交互，用户反馈回来的信息会影响当次风险分析结果。“交易中”的

交互式风险防控的反欺诈功能包括风险识别、风险评估、场景触达等多个方面，具体内容参见第6章。在交互过程中的实施与监测相关内容，参见第9章，例如在交互环节的服务划分参见9.1.2，风险防控规则参见9.1.3等。

“交易中”的交互式风险防控方式包括但不限于：

——弹窗提醒：弹窗展示风险信息；

——图文提醒：利用图片形式将风险内容从文字转换为漫画等图形，对用户进行事中提醒，提升管控有效率，降低打扰；

——全屏提醒：具有分点展示及多个反馈按钮的交互产品；

——事中语音外呼：利用电话触达用户，针对单个风险进行询问，用户按键反馈信息；

——智能外呼：触发智能客服自动电话触达用户，交互内容为模型推荐内容，用户交互方式为语音回答；

——考卷评估：利用考卷形式与用户进行即时互动，同时对用户输出1—6道题目，根据用户回答内容进行用户风险认知分析，从而判断用户是否对风险有清晰认知，能否继续进行支付；

——智能答题：利用问答形式与用户进行即时互动，根据用户每一题的反馈结果，实时决策下一题输出；

——交易挂起：交易有风险，风险防控将交易从立即到账处理成为到交易暂停；

——问答提醒：根据用户反馈的信息，定向进行风险揭示；

——端内服务：利用客服机器人的形式与用户进行深度交互，获取信息。

10.4 交易后

“交易后” 指在用户操作资金流转发生之后识别到交易风险，此时宜即刻输出合适的支付后交互产品，与用户进行双向信息传递，从而完成风险释放。“交易后”交互式风险防控不影响用户的操作，但根据产品不同，会在线同步等待用户反馈结果，也有可能异步回收用户信息。用户反馈回来的信息，有可能实时应用于此后的风险判断，也有可能影响之前交易的判断结果。“支付后”的交互更全面，更深入。通过交互过程进行反欺诈处理后，相关的技术保障内容参见9.2。

“交易后”交互式风险防控方式包括但不限于交易挂起、智能外呼、短信上行和邮件通知。

——延时到账：对用户的资金采取延迟24小时到账的方式，引导用户选择更加安全的转账方式，进行风险防控；

——回访调查：触发智能客服自动电话触达用户，交互内容为模型推荐回访调查内容，用户交互方式为语音回答，根据回访反馈补充黑样本、调剂限额或交易撤销等系列措施；

——短信上行：用户可以按短信提示回复内容，且反馈内容能被实时获取使用；

——邮件通知：利用邮件触达用户，传递风险信息。

附 录 A
（资料性）
风险分类示例

交互式风险防控中存在的各种风险，包括欺诈风险、合规风险和其他风险等类型，同时也存在多种风险交织并存的情况。本标准重点介绍交互式风险防控中欺诈风险，欺诈风险主要包括电信网络诈骗、信贷欺诈、保险欺诈等，同时也介绍合规风险和其他风险，具体分类示例如下。

A.1 欺诈风险

欺诈风险指不法分子以非法占有为目的，通过电话、网络和短信方式，编造虚假信息，通过虚构事实或者假冒身份来实施诈骗，对受害人实施远程、非接触式欺诈，诱使受害人陷入错误认知，最终在错误认知下按照诈骗人的意愿处置资产的风险。以下风险分类中，a）和b）是以诈骗使用的媒介分类，c）至l）以诈骗分子的诈骗手段分类，主要有冒充身份、虚构事实和敲诈勒索等，m）至q）以诈骗的场景分类，具体内容存在一定重叠。

常见欺诈风险场景包括但不限于电信网络诈骗、社交软件诈骗等类型。

a） 电信网络诈骗：电信网络诈骗是指通过电话、网络和短信方式，编造虚假信息，设置骗局，对受害人实施远程、非接触式诈骗，诱使受害人打款或转账的犯罪行为。

b） 社交软件诈骗：社交软件诈骗包括但不限于社交软件冒充公司老总诈骗等类型。

1） 社交软件冒充公司老总诈骗。例如，犯罪分子通过搜索财务人员社交软件群，以发送“会计资格考试大纲文件”等为诱饵发送木马病毒，盗取财务人员使用的社交软件号码，并分析研判出财务人员老板的社交软件号码，再冒充公司老板向财务人员发送转账汇款指令。或者犯罪分子通过技术手段获取公司内部人员架构情况，复制公司老总社交软件昵称和头像图片，伪装成公司老总添加财务人员社交软件实施诈骗。

2） 社交软件伪装身份诈骗。犯罪分子利用社交软件“附近的人”查看周围朋友情况，伪装成“高富帅”或“白富美”，加为好友骗取感情和信任后，随即以资金紧张、家人有难等各种理由骗取钱财。

3） 社交软件发布虚假爱心传递诈骗。犯罪分子将虚构的寻人、扶困帖子以“爱心传递”方式发布在朋友圈里，引起善良网民转发，实则帖内所留联系方式绝大多数为外地号码，打过去不是吸费电话就是电信诈骗。

4） 社交软件点赞诈骗。犯罪分子冒充商家发布“点赞有奖”信息，要求参与者将姓名、电话等个人资料发至社交软件平台，一旦商家套取足够的个人信息后，即以缴纳手续费、公证费、保证金等形式实施诈骗。

5） 社交软件盗用公众账号诈骗。犯罪分子盗取商家公众账号后，发布“诚招网络兼职，帮助网络卖家刷信誉，可从中赚取佣金”的推送消息。受害人信以为真，遂按照对方要求多次购物刷信誉，后发现上当受骗。

c） 虚构亲友紧急事项诈骗：虚构亲友紧急事项诈骗包括但不限于虚构车祸诈骗等类型。

1） 虚构车祸诈骗。犯罪分子以受害人亲属或朋友遭遇车祸，需要紧急处理交通事故为由，要求对方立即转账。当事人因情况紧急，便按照犯罪分子指示将钱款打入指定账户。

2） 虚构手术诈骗。犯罪分子以受害人子女或父母突发疾病需紧急手术为由，要求事主转账方可治疗。遇此情况，受害人往往心急如焚，按照犯罪分子指示转款。

3) 虚构绑架诈骗。犯罪分子虚构事主亲友被绑架，如要解救人质需立即打款到指定账户并不能报警，否则撕票。当事人往往因情况紧急，不知所措，按照犯罪分子指示将钱款打入账户。
4) 社交软件冒充好友诈骗。利用木马程序盗取对方社交软件密码，截取对方聊天视频资料，熟悉对方情况后，冒充该社交软件账号主人对其社交软件好友以“患重病、出车祸”“急需用钱”等紧急事情为由实施诈骗。
5) “猜猜我是谁”诈骗。犯罪分子获取受害人的电话号码和机主姓名后，打电话给受害人，让其“猜猜我是谁”，随后根据受害人所述，冒充熟人身份，并声称要来看望受害人。随后，编造其被治安拘留、交通肇事等理由，向受害人借钱，一些受害人没有仔细核实就把钱打入犯罪分子提供的银行卡内。

d) 虚构色情、黑灰产业诈骗：虚构色情、黑灰产业诈骗包括但不限于虚构色情服务诈骗等类型。
1) 虚构色情服务诈骗。犯罪分子在互联网上留下提供色情服务的电话，待受害人与之联系后，称需先付款才能上门提供服务，受害人将钱打到指定账户后发现被骗。
2) 重金求子诈骗。犯罪分子谎称愿意出重金求子，引诱受害人上当，之后以缴纳诚意金、检查费等各种理由实施诈骗。
3) 提供考题诈骗。犯罪分子针对即将参加考试的考生拨打电话，称能提供考题或答案，不少考生急于求成，事先将好处费的首付款转入指定账户，后发现被骗。
4) 复制手机卡诈骗。犯罪分子群发信息，称可复制手机卡，监听手机通话信息，不少群众因个人需求主动联系嫌疑人，继而被对方以购买复制卡、预付款等名义骗走钱财。

e) 虚构中奖诈骗：虚构中奖诈骗包括但不限于电子邮件中奖诈骗等类型。
1) 电子邮件中奖诈骗。犯罪分子通过互联网发送中奖邮件，受害人一旦与犯罪分子联系兑奖，犯罪分子即以缴纳个人所得税、公证费、转账手续费等各种理由要求受害人汇钱，达到诈骗目的。
2) 冒充知名企业中奖诈骗。犯罪分子冒充知名企业，预先大批量印刷精美的虚假中奖刮刮卡，通过信件邮寄或雇人投递发送，后以需交手续费、保证金或个人所得税等各种借口，诱骗受害人向指定银行账号汇款。
3) 娱乐节目中奖诈骗。犯罪分子以电视台热播栏目节目组的名义向受害人手机群发短消息，称其已被抽选为幸运观众，将获得巨额奖品，后以需交手续费、保证金或个人所得税等各种借口实施连环诈骗，诱骗受害人向指定银行账号汇款。
4) 兑换积分诈骗。犯罪分子拨打电话，谎称受害人手机积分可以兑换智能手机，如果受害人同意兑换，对方就以补足差价等理由要求先汇款到指定账户，或者发短信提醒受害人信用卡积分可以兑换现金等。如果受害人按照提供的网址输入银行卡号、密码等信息后，银行账户的资金即被转走。
5) 伪基站诈骗。犯罪分子利用伪基站向广大群众发送网银升级、电信运营商商城兑换现金的虚假链接，一旦受害人点击后便在其手机上植入获取银行账号、密码和手机号的木马，从而进一步实施犯罪。
6) 二维码诈骗。犯罪分子以降价、奖励为诱饵，要求受害人扫描二维码加入会员，实则附带木马病毒。一旦扫描安装，木马就会盗取受害人的银行账号、密码等个人隐私信息。

f) 冒充公职人员及公共服务机构人员诈骗：冒充公职人员及公共服务机构人员诈骗包括但不限于冒充公检法工作人员电话诈骗等类型。
1) 冒充公检法工作人员电话诈骗。犯罪分子冒充公检法工作人员拨打受害人电话，以事主身份信息被盗用涉嫌洗钱等犯罪为由，要求将其资金转入国家账户配合调查。

2) 购物退税诈骗。犯罪分子事先获取到事主购买房产、汽车等信息后，以税收政策调整可办理退税为由，诱骗事主到ATM机上实施转账操作，将卡内存款转入骗子指定账户。
3) 包裹藏毒诈骗。犯罪分子以事主包裹内被查出毒品为由，称其涉嫌洗钱犯罪，要求事主将钱转到国家安全账户以便公正调查，从而实施诈骗。
4) 刷卡消费诈骗。犯罪分子群发短信，以事主银行卡消费，可能泄露个人信息为由，冒充银联中心或公安民警连环设套，要求将银行卡中的钱款转入所谓的安全账户或套取银行账号、密码从而实施犯罪。
5) 医保、社保诈骗。犯罪分子冒充医保、社保中心工作人员，谎称受害人医保、社保出现异常，可能被他人冒用、透支，涉嫌洗钱、制贩毒等犯罪，之后冒充司法机关工作人员以公正调查、便于核查为由，诱骗受害人向所谓的安全账户汇款实施诈骗。
6) 补助、救助、助学金诈骗。犯罪分子冒充民政、残联等单位工作人员，向残疾人员、困难群众、学生家长打电话、发短信，谎称可以领取补助金、救助金、助学金，要其提供银行卡号，然后以资金到账查询为由，指令其在自动取款机上进行操作，将钱转走。

g) 冒充运营商诈骗：冒充运营商诈骗包括但不限于电话欠费诈骗等类型。
1) 电话欠费诈骗。犯罪分子冒充通信运营企业工作人员，向事主拨打电话或直接播放电脑语音，以其电话欠费为由，要求将欠费资金转到指定账户。
2) 电视欠费诈骗。犯罪分子冒充广电工作人员群拨电话，称以受害人名义在外地开办的有线电视欠费，让受害人向指定账户补齐欠费，否则将停用受害人本地的有线电视并罚款，部分群众信以为真，转款后发现被骗。

h) 冒充金融机构诈骗：冒充金融机构诈骗包括但不限于贷款诈骗等类型。
1) 贷款诈骗。犯罪分子通过群发信息，称其可为资金短缺者提供贷款，月息低，无需担保。一旦事主信以为真，对方即以预付利息、保证金等名义实施诈骗。
2) 解除分期付款诈骗。犯罪分子通过专门渠道购买购物网站的买家信息，再冒充购物网站的工作人员，声称“由于银行系统错误，买家一次性付款变成了分期付款，每个月都得支付相同费用”，之后再冒充银行工作人员，诱骗受害人到 ATM 机前办理解除分期付款手续，实施资金转账。
3) 金融交易诈骗。犯罪分子以某证券公司名义，通过互联网、电话、短信等方式散布虚假个股内幕信息及走势，获取事主信任后，又引导其在自身搭建的虚假交易平台上购买期货、现货，从而骗取事主资金。
4) 钓鱼网站诈骗。犯罪分子以银行网银升级为由，要求事主登录假冒银行的钓鱼网站，进而获取事主银行账户、网银密码及手机交易码等信息实施诈骗。
5) 冒充保险公司诈骗。冒充保险公司人员以补充保险理赔资料为名，或以快捷支付保险理赔款为名，让保险消费者信以为真，降低防范心理，从而听取他们的指令。诱骗保险消费者按照短信或电话指引通过网上银行或去银行 ATM 机操作，诱导保险消费者输入发送的验证码（实为转账金额），转走其账户资金。

i) 敲诈勒索类诈骗：敲诈勒索类诈骗包括但不限于合成图片实施诈骗等类型。
1) 合成图片实施诈骗。犯罪分子收集公职人员照片，使用电脑合成淫秽图片，并附上收款卡号邮寄给受害人，勒索钱财。
2) 冒充黑社会敲诈类诈骗。犯罪分子先获取事主身份、职业、手机号等资料，拨打电话自称黑社会人员，受人雇佣要加以伤害，但事主可以破财消灾，然后提供账号要求受害人汇款。

j) 冒充房东短信诈骗。犯罪分子冒充房东群发短信，称房东银行卡已换，要求将租金打入其他指定账户内，部分租客信以为真，将租金转出方知受骗。

k) 引诱汇款诈骗。犯罪分子以群发短信的方式，直接要求对方向某个银行账户汇入存款，由于事主正准备汇款，因此收到此类汇款诈骗信息后，往往未经仔细核实，即把钱款打入骗子账户。

l) 钓鱼类诈骗。黑客利用欺骗性的电子邮件和假冒的 Web 站点来进行诈骗活动。该诈骗活动诱骗访问者提供一些个人信息，如信用卡号、账户号和口令、社保编号等内容（通常主要是和财务、账号有关的信息）。

m) 订票诈骗：订票诈骗包括但不限于机票改签诈骗等类型。

1) 机票改签诈骗。犯罪分子冒充航空公司客服，以“航班取消、提供退票、改签服务”为由，诱骗购票人员多次进行汇款操作，实施连环诈骗。

2) 订票诈骗。犯罪分子利用门户网站、旅游网站、搜索引擎等投放广告，制作虚假的网上订票公司网页，发布订购机票、火车票等虚假信息，以较低票价引诱受害人上当。随后，再以“身份信息不全”“账号被冻结”“订票不成功”等理由要求事主再次汇款，从而实施诈骗。

n) 购物诈骗：购物诈骗包括但不限于退款诈骗等类型。

1) 退款诈骗。犯罪分子冒充淘宝等公司客服，拨打电话或者发送短信，谎称受害人拍下的货品缺货，需要退款，要求购买者提供银行卡号、密码等信息，实施诈骗。

2) 网络购物诈骗。犯罪分子开设虚假购物网站或淘宝店铺，一旦事主下单购买商品，便称系统故障，订单出现问题，需要重新激活。随后，犯罪分子通过社交软件发送虚假激活网址，受害人填写好淘宝账号、银行卡号、密码及验证码后，卡上金额即被划走。

3) 低价购物诈骗。犯罪分子通过互联网、手机短信发布二手车、二手电脑、海关没收的物品等转让信息，一旦事主与其联系，即以缴纳定金、交易税手续费等方式骗取钱财。

4) 快递签收诈骗。犯罪分子冒充快递人员拨打事主电话，称其有快递需要签收但看不清具体地址、姓名，需提供详细信息便于送货上门。随后，快递公司人员将送上物品（假烟或假酒），一旦事主签收后，犯罪分子再拨打电话称其已签收必须付款，否则讨债公司或黑社会将找麻烦。

5) 社交软件假冒代购诈骗。犯罪分子在社交软件朋友圈假冒正规微商，以优惠、打折、海外代购等为诱饵，待买家付款后，又以“商品被海关扣下，要加缴关税”等为由要求加付款项，一旦获取购货款则失去联系。

o) 收藏诈骗。犯罪分子冒充各类收藏协会，印制邀请函邮寄各地，称将举办拍卖会并留下联络方式。一旦事主与其联系，则以预先缴纳评估费、保证金、场地费等名义，要求受害人将钱转入指定账户。

p) 套现类诈骗。犯罪分子通过诈骗，诱使受害人以不合法的方式进行信用卡套现、贷款套现的行为。

q) 高薪招聘诈骗。犯罪分子通过群发信息，以月工资数万元高薪招聘某类专业人士为幌子，要求事主到指定地点面试，随后以缴纳培训费、服装费、保证金等名义实施诈骗。

A.2 合规风险

合规风险指商业银行、非银行支付机构、特约商户及第三方专业化服务机构等支付业务参与方因未能遵循法律法规、监管要求、业务规则及内部规范等，可能遭受法律制裁、监管处罚、违规约束进而引发财务或声誉损失的风险，包括但不限于洗钱风险、挪用客户备付金风险、非法集资风险、赌博风险等类型。

a) 洗钱风险。洗钱风险指将通过各种手段掩饰违法所得，隐瞒违法来源，使其在形式上合法化，常见于毒品犯罪、黑社会性质的组织犯罪、恐怖活动犯罪、走私犯罪、贪污贿赂犯罪、破坏金

融管理秩序犯罪、金融诈骗犯罪等违法犯罪过程，洗钱风险指机构未能识别洗钱行为，违反反洗钱管理规定的风险。

b) 挪用客户备付金风险。挪用客户备付金风险指机构未能履行合规管理义务，致使不法分子通过网络攻击，编造虚假交易和信息，伪冒商户等方式挪用、占用、借用客户备用金而造成持卡人或账户所有人、商户和机构资金和声誉损失的风险。

c) 非法集资风险。非法集资指单位或个人未依照法定程序经有关部门批准，以发行股票、债券、彩票、投资基金证券或者其他债权凭证的方式向社会公众筹集资金，并承诺在一定期限内以货币、实物以及其他方式向出资人还本付息或给予回报的行为，非法集资风险指服务机构未能识别非法集资行为，为非法集资提供支付服务等便利条件的风险。

d) 赌博风险。赌博风险指以营利为目的，实施聚众赌博、网络赌博或者以赌博为业，服务机构在支付业务中未能识别赌博资金转移的风险。

A.3 其他风险

除了欺诈风险和合规风险之外，不同机构、不同业务场景可能存在其他的风险类型，如资金清算风险、用户道德风险等。

附 录 B
（资料性）
交互式风险防控反欺诈实施案例

B.1 理财APP反欺诈应用案例

用户在理财APP上发布转账请求。基于该转账请求，发现该APP环境存疑，初步判定操作存在风险。基于该操作的场景可知其操作场景为理财中的炒股场景，通过用户身份模型了解到该用户为60岁以上老年人，存在个人敏感信息泄漏和用户财产损失风险。基于用户60岁以上老年人的身份，宜使用智能语音外呼交互方式，AI智能客服与老年人沟通：

第一轮：

“您是如何获取这个理财APP的？”

您知道投入的资金转往哪家公司吗？

如反馈为：通过群聊中的“专家”指点，投入资金转入xx公司等，可构建用户的序列行为链条，还可关联查询该理财APP的环境是否正常，其所属公司的经营范围等等。

第二轮：

“专家”身份是否知悉？理财APP疑似存风险APP？

您是否知晓其所属公司主营业务？

如回答“专家”为知名私募大鳄，其称所属公司为知名投资公司，群内另有多位用户获利颇丰等相似内容，自动识别后判断真伪，由系统根据策略实时调整，也可由客服人员手动设置风险级别。

第三轮：

根据用户意图判断为理财入资，产品为指数基金类产品对老年人风险较高，主动提醒其“投资有风险，养老钱要谨慎”并提供24小时延迟到账选项。

第四轮：

如客户付款后发现被骗，银行帮助被骗客户对名下账户进行保护性止付、提供交易记录及提示客户风险信息，必要时报警求助。

针对高风险操作，以该用户确认该操作为高风险操作并将要释放风险为多轮交互的停止条件。

参 考 文 献

[1] GB/T 20984—2022 信息安全技术 信息安全风险评估方法
[2] GB/T 22239—2019 信息安全技术 网络安全等级保护基本要求
[3] GB/T 23694—2013 风险管理 术语
[4] JR/T 0171—2020 个人金融信息保护技术规范
[5] JR/T 0197—2020 金融数据安全 数据安全分级指南
[6] JR/T 0202—2020 基于大数据的支付风险智能防控技术规范
[7] JR/T 0221—2021 人工智能算法金融应用评价规范
[8] 人民公安报.公安机关公布48种常见电信诈骗手法[N]，2016年02月25日