生成式人工智能驱动下电信网络诈骗受害风险影响因素量化分析

周胜利[1,2]，徐睿[2]，陈庭贵[3]，汪邵杰[1]
（1. 浙江警察学院信息网络安全学院，浙江 杭州 310053；
2. 杭州电子科技大学网络空间安全学院，浙江 杭州 310018；
3. 浙江工商大学统计与数学学院，浙江 杭州 310018）

摘　要：开展生成式人工智能驱动下电信网络诈骗的受害风险影响因素研究，对于揭示犯罪规律和提升技术防治能力具有重要的理论价值与实践意义。为此，依托真实AI诈骗案件数据开展模拟实验，将犯罪过程解构为伪造信息的生成、传播与影响3个阶段，从中提取生成式人工智能、数据流、数据包、网络行为和受害风险等潜变量，再结合结构方程模型理论构建分析框架，系统量化了不同要素对受害风险的影响路径与作用贡献度。研究结果表明，生成式人工智能对受害风险具有显著的直接效应，且在整体影响效应中占据主导地位；数据流与数据包特征的中介效应较弱，在影响路径中作用不显著。

关键词：生成式人工智能；电信网络诈骗；结构方程模型；量化分析
中图分类号：D924.3
文献标志码：A
doi: 10.11959/j.issn.1000−0801.2025163

Quantitative analysis of influencing factors of victimization risk of telecom network fraud driven by generative artificial intelligence

ZHOU Shengli[1,2], XU Rui[2], CHEN Tinggui[3], WANG Shaojie[1]
1. School of Information Network Security, Zhejiang Police College, Hangzhou 310053, China
2. School of Cyberspace Security, Hangzhou Dianzi University, Hangzhou 310018, China
3. School of Statistics and Mathematics, Zhejiang Gongshang University, Hangzhou 310018, China

Abstract: Conducting research on the factors influencing victimization risks in generative AI (GAI) driven telecom network fraud holds significant theoretical and practical implications for areas such as summarizing patterns of criminal behavior and enhancing technological defense capabilities. For this purpose, simulation experiments were carried out based on real AI fraud case information. The criminal process was deconstructed into three stages: forged information generation, dissemination, and impact of forged information. Latent variables such as GAI, data flow, data pack-

收稿日期：2025−04−02；修回日期：2025−07−02
基金项目：国家社会科学基金项目（No.23BGL272）
Foundation Item: The National Social Science Foundation of China (No.23BGL272)

ets, network behavior, and network risk were extracted. An analytical framework was then constructed by combining structural equation modeling theory to systematically quantify the influence paths and contribution degrees of different elements on victimization risk. The findings revealed that GAI had a significant direct effect on network risk, and the direct effect played a dominant role in the overall effect. The mediating effects of data flow and packet characteristics were weak, and its role in the influence path was not significant.

Key words: generative artificial intelligence, telecom network fraud, structural equation model, quantitative analysis

0 引言

随着数字信息技术的快速发展，电信网络诈骗呈现出高发案率、高额经济损失的显著特征，不仅给受害人带来直接财产损失，更对社会诚信体系造成持续性破坏，已成为当前最为突出的犯罪类型之一[1]。同时，随着自然语言处理等人工智能技术的突破性发展，生成式人工智能（generative artificial intelligence，GAI）成为新型电信网络诈骗犯罪的技术推手。诈骗分子通过生成海量含有虚假信息的文本、视频、图像等媒介[2-3]，甚至快速制作诈骗网站、诈骗语音等犯罪工具，能快速骗取犯罪对象的信任，诱使犯罪对象产生电信网络诈骗受害风险，最终产生财产损失。《中华人民共和国反电信网络诈骗法》将电信网络诈骗定义为“以非法占有为目的，利用电信网络技术手段，通过远程、非接触等方式，诈骗公私财物的行为”，因此诈骗分子利用GAI技术实施的诈骗犯罪（以下简称AI诈骗）本质上是一种新型电信网络诈骗。但相较于传统诈骗手段，AI诈骗犯罪具有欺骗性更强、针对性更强等特点[4]，并且随着GAI技术的普及和开源集成工具的广泛传播，不具备专业技术背景的诈骗分子也能轻松利用AI工具实施诈骗，极大可能导致AI诈骗犯罪的快速蔓延，给网络空间综合治理带来巨大挑战。

本文将构建融合GAI特征、网络流量特征、网络行为特征与受害风险特征之间的结构方程模型，基于相关量化分析理论，揭示GAI技术对电信网络诈骗受害风险的影响路径和程度，为当前电信网络诈骗受害风险识别研究提供数据与理论基础。

1 国内外相关研究

在以GAI技术为代表的信息技术高速发展背景下，电信网络诈骗已成为全球范围内危害公共安全与社会稳定的突出问题。传统事后被动干预的治理模式已难以满足技术驱动的新型犯罪防范需求，亟须通过系统性数据分析犯罪的关键影响因素，降低公众遭受电信网络诈骗的风险[5]。为此，国内外学者已针对电信网络诈骗受害风险的影响因素进行了充分的分析研究，研究视角主要集中在人口属性特征和时空分布规律两个方面。

在人口属性特征方面，早期的日常生活理论、生活方式理论为解释诈骗受害者的社会属性差异提供了基础分析框架[6]。现有研究表明，年龄、性别等人口学特征与教育水平、职业类型等社会属性共同构成受害风险的关键预测因子[7]。学者们发现单纯依赖人口统计学变量难以完全解释个体层面的受害机制差异，为此，部分学者进一步从受害者的心理视角展开研究。例如，刘晓萌[8]从受害者的心理视角出发，揭示了年轻受害者可能存在如对快速获取利益的追求、社会经验不足及风险意识缺乏等心理漏洞；田丽娟等[9]进一步指出，受害者贪财逐利的虚荣心理、向他人看齐的从众心理、同情弱者的支持心理或受到威胁的恐惧心理等心理特征是电信网络诈骗犯罪既遂的主要原因。以上研究主要基于多源实证数据进行涉案人群画像统计分析，剖析了各类人口属性特征对犯罪形成的影响程度[10]，目前已在理论

和方法层面具备一定研究基础。

在时空分布规律方面，当前已有部分学者证实，电信网络诈骗虽然具备虚拟空间作案特征显著的特性，但受基础设施、人口流动规律等因素的影响，仍呈现出一定的空间集聚性、时间分异性等时空规律[11]。关注重点犯罪区域与时段，有助于理解电信网络诈骗的传播路径和作案规律。同时，针对不同的电信网络诈骗热点地区、时段采取差异化的防控策略[12]，可以有效提升电信网络诈骗的预警效果[13]，或进一步研究电信网络诈骗行为运行模式，把握诈骗者行为规律，构建电信网络诈骗的风险预警框架，实现更全面的电信网络诈骗预防[14]，以全局性、精准性特征的“全链条式”打击思路从源头阻断犯罪。

基于数据的电信网络诈骗规律发掘研究为犯罪治理提供了重要的决策依据，但现有针对电信网络诈骗受害风险的研究成果大多依赖人口属性、地理位置、设备信息等静态特征，缺乏针对犯罪过程中通信流、网络行为流等多模态数据的融合分析。整体呈现出以静态特征分析为主、动态过程研究不足的特点，不利于全面、真实地把握风险影响因素和犯罪模式特征。同时，尽管已有学者注意到，诈骗分子正利用GAI技术伪造多模态诈骗内容[15-16]，以进一步骗取犯罪对象的信任，并研究了此类生成内容的特征（如字符级向量化[17]、分节标签嵌入[18]），但对于AI诈骗等新型电信网络诈骗的犯罪模式规律探索仍处于起步阶段。现有针对影响因素的分析研究也主要依靠受害人的信息辨识能力[19]或专家经验[20]展开，未能充分结合AI诈骗的动态交互性特点。因此，亟须从AI诈骗场景中的动态行为维度深化AI诈骗的受害风险影响因素分析研究。

2　面向受害风险影响量化分析的结构方程模型构建

2.1　GAI驱动下电信网络诈骗的犯罪模式

随着GAI技术的快速发展，其被滥用于电信网络诈骗的犯罪风险显著增加。此类技术能够以极低的成本制作高度逼真的伪造信息，打破传统以“眼见为实”为基础的信任认知模式，对公众的生命财产安全构成严重威胁[21]。因此，对GAI驱动下的电信网络诈骗犯罪模式进行分析，在总结犯罪规律、提升风险预警能力等方面具有重要意义。为此，本文基于现有案例进行分析，总结了当前GAI驱动下的电信网络诈骗犯罪模式，如图1所示。

（1）伪造信息生成阶段

在伪造信息生成阶段，犯罪分子依托黑灰产业链准备犯罪的相关物料，一方面通过黑灰产业链的物料商批量采购包含公民身份证号码、生物特征等敏感信息的隐私数据集，另一方面获取深度伪造工具包、语音克隆软件等GAI犯罪工具，用于实施犯罪活动。然后，诈骗分子会搜集冒充对象的照片、语音等生物特征数据，并利用GAI犯罪工具生成用于实施诈骗所需的人工智能生成内容（artificial intelligence generative content，AIGC）伪造信息。通过模型训练迭代生成高度仿真的换脸视频、拟真声线等伪造信息，不仅加

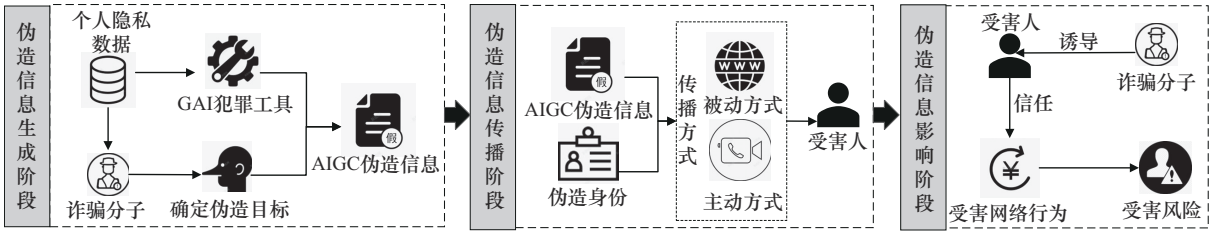

图1　GAI驱动下的电信网络诈骗犯罪模式

剧了公民隐私数据的非法流通风险，更存在着GAI技术向犯罪工具异化的风险。

（2）伪造信息传播阶段

在此阶段，诈骗分子通过网络平台，以主动或被动的方式与受害者产生接触，进行犯罪的初步引流。就被动方式而言，诈骗分子通常会在社交平台上大量散布经过GAI技术处理的虚假图片、新闻信息等内容，并可能附加对应的诈骗媒介（如诈骗网站、App等）链接，诱导受害者主动点击、下载或转发。就主动方式而言，以利用换脸视频实施的AI诈骗犯罪为例，诈骗分子借助GAI换脸工具进行身份伪装，再与受害人进行视频通话，以此来主动骗取受害人的信任。伪造信息经过技术赋能后传播，大幅削弱了公众对虚假信息的辨识阈值，进一步增加了AIGC伪造信息传播及误辨等风险。

（3）伪造信息影响阶段

在此阶段，诈骗分子将AIGC伪造信息和犯罪剧本深度融合，对受害人进行心理诱导和操控。在接收到AIGC伪造信息后，受害人一时间难以辨识接收信息的真伪，很容易对诈骗分子的伪造身份产生错误的信任，进而受制于犯罪分子的指示，并最终产生受害风险，包括点击访问未知的诈骗媒介、向陌生账户进行转账操作等。

2.2 研究问题与假设

随着互联网技术的快速发展与应用，网络环境中用户行为与数据交互的复杂性显著增加，受害风险的形成机制逐渐从单一技术漏洞向多维度、多路径的复合型影响演变。本文聚焦于探索GAI应用场景下网络流量、网络行为与受害风险之间的关联路径，并结合第2.1节分析结果，构建综合性的受害风险影响评估框架，开展特征间相互影响作用的量化分析研究。具体特征包括：GAI特征（对应伪造信息生成阶段）、数据流特征（对应伪造信息传播阶段）、数据包特征（对应伪造信息传播阶段）和用户网络行为特征（对应伪造信息影响阶段）。基于以上论述，本文提出以下研究问题。

（1）GAI技术的具体应用形态对受害风险的差异化影响

现有研究表明，不同GAI技术的应用会引发差异化的用户行为模式[22]，需要分类讨论。本文将GAI应用特征细分为“生成式人工智能文本图像”（GAI_text_image）、“生成式人工智能视频”（GAI_video）及“未使用生成式人工智能”3类。假设GAI_video的应用可能通过增加网络暴露面（如延长数据流持续时间、使用特定端口传输数据等）间接提升风险水平；GAI_text_image的应用可能通过生成伪造信息数据，诱导用户点击风险链接，直接触发风险事件（如银行账户操作风险、在线金融交易等）。基于以上分析，提出如下假设。

- H1：GAI应用对受害风险存在直接效应，而这又包含两个子假设，分别是H1a和H1b。
- H1a：GAI_video应用通过扩大数据流规模与持续性，对受害风险产生显著正向影响。
- H1b：GAI_text_image应用通过诱导风险行为产生，对受害风险产生显著正向影响。

（2）数据流与数据包特征在GAI应用与受害风险之间发挥中介作用

将数据流特征定义为包含流量字节数、平均负载大等观测变量的潜变量；数据包特征定义为包含端口类型（port）、负载参数等观测变量的潜变量。研究假设GAI应用会显著改变数据流的动态特征，如GAI_video的生成与传输需要持续的高负载传输，这可能导致数据流持续时间延长、字节数激增等；GAI_text_image的交互可能通过高频次、小负载的传输模式，导致产生固定源/目的IP组合、访问特殊端口等特征。因此，数据流与数据包特征既是GAI技术应用的客观结果，也是受害风险传导的关键中介变量。基于以上分析，提出如下假设。

- H2：数据流特征的中介效应假设。数据流的持续时间、字节数等特征在GAI应用与受害风险之间发挥中介作用，改变相关流量特征可间接增加受害风险。
- H3：数据包特征的中介效应假设。数据包的端口特征在GAI应用与受害风险之间发挥中介作用，特殊端口的使用将间接增加受害风险。

（3）用户网络行为模式在风险传导路径中发挥中介效应

将用户在网络空间中进行的网络活动行为模式（以下简称网络行为）定义为涵盖网页浏览、登录行为等观测变量的潜变量，假设GAI应用能通过改变用户不同种类的网络行为频率来影响受害风险程度。如GAI_text_image可能增加网页浏览、登录行为的次数，增加用户对非正规链接的点击率。因此，网络行为既是GAI技术影响用户决策的表现，也是影响受害风险的关键因素。此外，由于可利用网络流量特征识别网络行为特征，故数据流与数据包特征对用户网络行为特征存在直接影响。基于以上分析，提出如下假设。

- H4：网络行为的中介效应假设。GAI应用通过改变用户行为模式影响网络行为特征，进而通过行为特征增强受害风险。
- H5：数据流/数据包特征对网络行为存在直接影响效应，引导产生网络行为。

基于以上问题分析与假设，构建了包含GAI特征、数据流特征、数据包特征、网络行为和受害风险5个维度的研究模型结构，如图2所示。

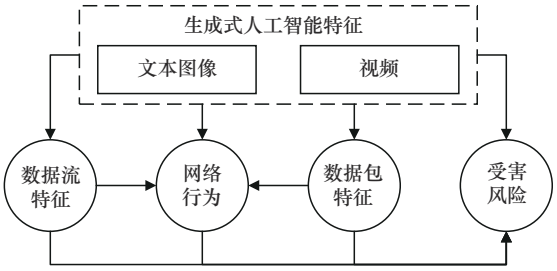

图2　研究模型结构

2.3　结构方程模型的构建与分析方法

基于结构方程模型理论，对GAI应用场景下的受害风险传导路径进行了系统性建模与量化分析。结构方程模型是一种验证性分析模型，其核心目标在于检验理论假设，主要用于探索变量间的复杂关系并验证理论模型的合理性[23]。结构方程模型的构建与分析过程包括数据预处理、潜变量定义、路径关系设定、参数估计及效应分解等多个环节，具体如图3所示。

2.3.1　数据准备

模拟诈骗分子对受害者实施AI诈骗的完整流程，从中采集流量数据包，并利用特征提取脚本抽取对应的特征数据集，特征提取脚本的伪代码如算法1所示。其中，第1～5行为特征提取模块，通过遍历数据包，提取时间信息、包数量、字节数等特征信息。第6～9行为IP地址分析模块，分析网络流量包中的IP地址，确定本机IP地址（出现次数最多的IP地址）和目标IP地址（除本机IP地址外的其他IP地址）列表。第10～14

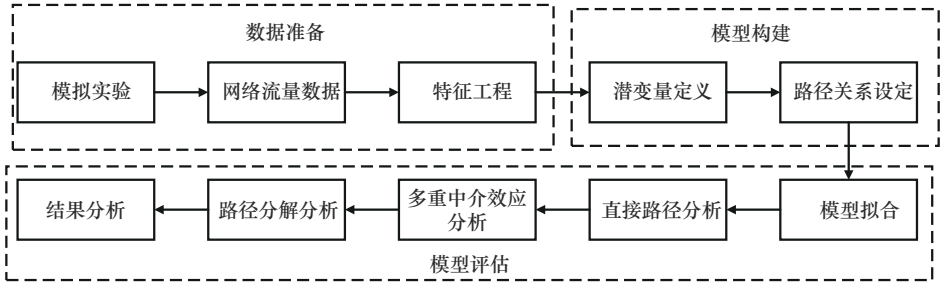

图3　结构方程模型的构建与分析流程

行为特征提取与保存模块，对每个目标IP地址进行特征提取，再将结果保存到CSV文件中，最终生成特征数据集。

算法1 网络流量特征提取

输入：packet #从Wireshark捕获的网络数据包

输出：feature #包含网络流特征的CSV文件

BEGIN

1.Define feature_geter class: #特征提取模块

2.Initialize variables

3.Define get_feature method:

4.For each packet:

#通过遍历数据包，计算数据包、数据流等相关特征

Extract timing info…#提取特征

5.Return feature dictionary

6.For each packet:

#分析流量包中的IP地址

7.Count IP occurrences

8.Determine self IP (most frequent) #获取目标IP列表

9.Get target IP list

10.For each target IP:

#对每个目标IP执行特征提取并保存结果

11.Create filter

#为每个目标IP创建过滤器，筛选数据包

12.Extract features using feature_geter

#使用feature_geter类提取特征

13.Combine features with file info

#将特征与文件信息合并

14.Append results to feature #将结果追加到CSV文件中

END

此外，还需要对特征数据集feature进行处理。

（1）对连续变量进行标准化处理，使用Standard Scaler方法转化为标准正态分布数据。

（2）对分类变量进行编码转换处理，将其转化为类别编码。

（3）通过特征工程构造二值化变量，包括将高频流量（发包个数>50）标记为flow_freq，443端口使用情况标记为port，长时间会话流（持续时长>10 s）标记为time。

通过以上方式增强特征的可解释性。

2.3.2 模型构建

结构方程模型分为测量模型与结构模型两部分。测量模型定义了潜变量与观测变量之间的关系，具体如下。

（1）受害风险潜变量：由银行卡操作与在线交易2个观测变量指标构成，反映金融场景下的具体风险类型。

（2）数据流特征潜变量：由流量字节数、平均负载大小、数据包数量及持续时间构成，刻画流量规模与持续性。

（3）数据包特征潜变量：以端口类型为核心观测变量，侧重传输层协议特征。

（4）网络行为潜变量：涵盖网页浏览、登录注册、在线聊天，表示用户在通信过程中产生的网络行为模式。

结构模型负责设定潜变量间的路径关系，根据第2.2.1节的研究假设可知，共包含4类假设路径，具体如下。

（1）GAI应用的直接效应（H1）：GAI_text_image应用与GAI_video应用直接指向受害风险，将形成“GAI_text_image→受害风险”与“GAI_video→受害风险”2条直接传导路径。

（2）数据流与数据包的中介效应（H2、H3）：GAI应用（包括GAI_text_image与GAI_video）通过影响流量特征间接作用于受害风险，形成“GAI→数据流/数据包→受害风险”的间接传导路径。

（3）网络行为的中介效应（H4）：GAI应用改变用户的网络行为模式，进而提升风险，形成“GAI→网络行为→受害风险”的间接传导路径。

（4）流量特征对行为的传导路径（H5）：数据流与数据包特征直接影响网络行为，形成“GAI→数据流/数据包→网络行为→受害风险”的间接传导路径。

2.3.3　结构方程模型的评估方法

为量化结构方程模型中各变量之间的影响机制，需要进行直接路径分析、多重中介效应分析及路径分解分析，具体如下。

（1）直接路径分析

结构方程模型直接路径分析通过多步骤数据处理流程，实现路径效应的系统化检验，其伪代码如算法2所示。其中，第1～2行实现数据筛选模块，通过操作符过滤提取结构模型路径；第3～4行执行列名转换模块，生成符合中文语义的路径关系表达式；第5行建立假设验证模块，结合路径系数正负性与p值判断假设成立性；第6行完成结果格式化，输出包含路径关系、路径系数、标准误差（S.E.）、临界比值（C.R.）、p值和假设成立情况的表格，为后续分析提供基础数据。

算法2　直接路径分析

输入：模型参数估计结果params_df

输出：结构化路径分析结果structural_model

BEGIN

1. Filter params_df where operator is regression (~)

筛选结构方程路径

2.Select columns: lval, rval, Estimate, Std. Err, *z*-value, *p*-value # 选取关键指标

3.Rename columns to 因变量, 自变量, 路径系数, S.E., C.R., *p* # 中文化列名

4.Generate 路径关系=因变量 + '~' + 自变量

构建路径关系表达式

5.Calculate 假设成立= ($p < 0.05$) # 验证路径显著性

6. Reorder columns for final output # 调整列顺序

END

（2）多重中介效应分析

多重中介效应分析通过全面评估模型中各路径的中介效应，揭示自变量对因变量的影响机制，其伪代码如算法3所示。其中，第1～4行定义了提取直接效应模块，统计从自变量到因变量的直接路径估计值；第5～16行定义了中介路径分析函数，计算自变量通过中介变量到因变量的间接效应（对应第5～8行），同时考虑网络行为的双重中介效应（对应第9～16行），即自变量通过网络行为再到因变量的路径；第17行将所有直接效应和中介效应整理为结构化表格，为机制解释提供量化依据。

算法3　多重中介效应分析

输入：模型参数估计结果params

输出：中介效应分解表mediation_df

BEGIN

1.Initialize effects list

初始化效应存储结构

2.Extract direct effects

提取直接效应模块

3. a. Filter params where 受害风险~AI_text_image

4. b. Filter params where 受害风险~AI_video

5. Define add_indirect_effects(source, mediator):

中介效应计算函数

6. a. Find path1: mediator ~ source

7. b. Find path2: 受害风险 ~ mediator

8. c. If both exist: effect = path1_est * path2_est

9. Loop through all AI_* and mediators:

遍历所有中介路径

10. for source in [AI_text_image, AI_video]:

11. for med in [数据流, 数据包, 网络行为]:

12. execute add_indirect_effects(source, med)

13. Handle dual mediation paths:

处理双重中介路径

14. for source in AI_*:

15. if 网络行为~source and 受害风险~网络行为 exist:

16. effect = path_est1 * path_est2

17. Convert effects list to DataFrame # 结果结构化

END

（3）路径分解分析

路径分解分析通过量化各路径在总效应中的重要性，帮助识别关键路径，其伪代码如算法4所示。其中，第1行计算所有路径效应值的绝对值总和；第2～4行实现单路径贡献计算，通过绝对值反映路径影响力大小；第5行建立贡献度标准化模块，将绝对贡献值转化为百分比形式；第6～7行生成路径贡献度分解表格，包含路径、效应值、绝对值贡献和贡献占比。

算法4 路径分解分析

输入：中介效应分析结果mediation_df

输出：路径贡献度分解表decomp_df

BEGIN

1.Calculate total_effect = sum(abs(效应值))

计算总效应绝对值

2.Initialize decomposition list

创建贡献度存储结构

3.For each row in mediation_df

遍历所有效应路径

4. a. contribution = abs(效应值)

5. b. proportion=contribution/ total_effect*100

6. c. Append dict{路径, 效应值, 贡献度, 占比}

7. Convert list to DataFrame#生成结构化结果

END

3 实验与结果分析

3.1 实验环境

实验环境分为操作系统和算法实现的框架环境，操作系统的环境配置见表1；本实验使用Python编程语言，具体环境配置见表2。

表1 系统环境配置

环境名称	配置信息
系统版本	Microsoft Windows11
GPU驱动版本	31.0.15.5161
CPU	12th Gen Intel(R)Core(TM)i9-12900H@2 500 MHz
内存	16 GB

表2 框架环境配置

名称	版本	功能
Python	3.7.9	执行依赖环境
Scipy	1.2.1	处理数据，进行最优化和统计等操作
NumPy	1.18.1	存储和处理大型矩阵，进行数据科学计算
Matplotlib	3.0.3	对实验数据和计算图进行可视化展示
Sklearn	1.0.2	负责特征项的标准化处理
Pyshark	3.0.9	负责从数据包中提取数据
Semopy	2.3.10	实现结构方程模型的构建

3.2 数据获取实验

通过模拟受害人遭受AI诈骗的全过程来获取相关实验数据。

（1）使用GAI集成工具获取伪造信息：对于换脸视频，使用DeepFaceLive（版本号NVIDIA build 2023）换脸软件生成实时换脸视频；对于虚假文本与图像，利用实战单位提供的诈骗剧本，结合文心一言模型批量生成伪造文本，并配合即梦AI批量生成伪造图片。

（2）诈骗分子模拟方使用QQ（版本号9.7.23）分别进行视频通话（利用OBS studio的虚拟摄像头功能将换脸视频输入QQ的视频通话过程）和文本图像聊天（基于诈骗剧本使用生成的文本配合图片进行通信），并最终发送一个真实诈骗媒介链接（通过电信网络诈骗治理的实战单位获取），以此模拟AI诈骗过程；受害人模拟方通过QQ与诈骗分子进行对应的通信，使用Wireshark捕获整个通信过程中产生的网络流量N。

数据采集的具体步骤如下。

（1）诈骗分子模拟方利用QQ与受害人进行

实时的换脸视频或文本图像聊天通信，并在通信过程中发送一个真实的诈骗媒介链接。

（2）受害人模拟方仅打开 Wireshark，通过 QQ 与诈骗分子方进行通信，并有选择性地点击或不点击对方发送的链接。若点击，则在对应网站中模拟 AI 诈骗受害人在诈骗媒介中进行网络行为活动，然后结束通信流程，保存抓取的数据包 *N*；若不点击则直接结束整个通信流程，保存抓取的数据包 *N*。保存文件的文件名需标注 GAI 技术类型及具体网络诈骗行为标签。

（3）在不使用GAI技术的情况下，通过模拟一般电信网络诈骗的流程，重复步骤（1）、步骤（2），直至完成全部数据采集工作。

通过设置分组实验，分别获取 AI 诈骗数据与非 AI 诈骗数据，每个分组仅负责一类数据的采集任务，同时通信双方均会模拟诈骗分子身份与受害人身份，以此获取不同模拟环境下的数据，并确保实验数据集中不存在同一模拟者同时采集的 AI 诈骗数据与非 AI 诈骗数据，以此消除不同类样本间的关联性。研究最终获取 3 337 条实验数据，实验数据集包含的具体特征参数及其含义见表 3。

根据结构方程模型构建的一般经验，至少需要 100～150 个样本用于构建结构方程模型，并且每存在一个观测变量则需要 5～10 个样本。由于该结构方程模型包含 12 项观测变量，故需要 100～120 个样本。综上所述，研究获取的 3 337 组样本数据能够充分支持结构方程模型的量化分析需求。

3.3　结构方程模型构建结果

本文的主要目标是探究 GAI 特征、数据流特征、数据包特征、网络行为以及受害风险之间的关系。在构建结构方程模型后，需要对模型的拟合度进行检验，以此判断模型与数据的适配程度以及模型对研究假设的解释力，所得的模型拟合度检验结果见表 4。根据参考标准，CFI、GFI、NFI 和 TLI 均大于 0.9，表明模型的拟合度较好；AGFI 接近 0.9，略低于参考标准 0.003，RMSEA 略高于参考标准 0.008；但在大样本量（数据量> 500）的研究中，AGFI 与 RMSEA 仍属可接受范围。综上所述，模型整体拟合度较为理想，能够较好地反映变量之间的关系，即构建的结构方程模型可满足后续量化分析要求。

通过模型拟合度检验后，调用 Semopy 库的 semplot 函数得到结构方程模型路径系数图，如图 4 所示。

3.4　结果分析

3.4.1　路径系数分析

根据第 2.2 节的研究假设，对模型的主体框

表 3　实验数据集包含的特征参数及其含义

特征参数	对应隐变量	特征含义
port	数据包特征	目标端口
flow_number	数据流特征	数据流总包数
flow_byte	数据流特征	数据流总字节数
flow_duration	数据流特征	数据流持续时长
ave_length	数据流特征	数据流中数据包的平均负载数据长度
risk_transaction	受害风险	是否在风险环境中进行在线转账支付行为
risk_bank	受害风险	是否在风险环境中进行操作银行卡行为
browse_web	网络行为	是否存在网页浏览行为
Re_login	网络行为	是否在网站中进行了注册、登录行为
online_chat	网络行为	是否在网站中进行聊天行为
GAI_text_image	GAI	标识文本图像类通信
GAI_video	GAI	标识视频类通信

表4　模型拟合度检验结果

指标	参考标准	实验结果
比较拟合指数（CFI）	>0.9	0.939
格列比斯好比率（GFI）	>0.9	0.937
调整后的格列比斯好比率（AGFI）	>0.9	0.897
拟合指数（NFI）	>0.9	0.937
塔克-刘易斯指数（TLI）	>0.9	0.901
近似误差均方根（RMSEA）	<0.08	0.088

架进行路径系数分析，并根据路径系数结果判断变量之间的影响关系大小及其显著性，最终得到H1～H5的5个假设的标准化路径系数结果，见表5。

（1）GAI应用的直接效应（H1）：GAI_text_image和GAI_video对受害风险均具有显著的正向直接效应，路径系数分别为0.499（p<0.001）和0.48（p <0.001），表明GAI应用能够独立于中介变量，并直接增加受害风险，假设H1成立。

（2）数据流与数据包的中介效应（H2与H3）：数据流对受害风险具有显著的正向影响，路径系数为0.005（p<0.001）；而数据包对受害风险具有显著的负向影响，路径系数为−0.203（p < 0.001）。然而，GAI应用对数据流的影响均不显著（GAI_text_image的路径系数为−0.018，p=0.670；GAI_video的路径系数为−0.045，p=0.272）；仅GAI_text_image对数据包的影响显著，路径系数为0.005，p=0.017。因此，假设H2部分成立，H3不成立。

（3）网络行为的中介效应（H4）：GAI_text_image和GAI_video对网络行为均具有显著的正向影响，路径系数分别为0.860（p <0.001）和0.813（p <0.001）；网络行为对受害风险具有显著的负向影响，路径系数为−0.403（p <0.001）。这表明GAI应用通过改变网络行为间接降低了受害风险，与假设H4相反，因此H4不成立。

（4）流量特征对行为的传导路径（H5）：数据流和数据包对网络行为的影响均不显著，数据流的路径系数为−0.004，p为0.375；数据包的路径系数为−0.285，p为0.954。因此，假设H5不成立。

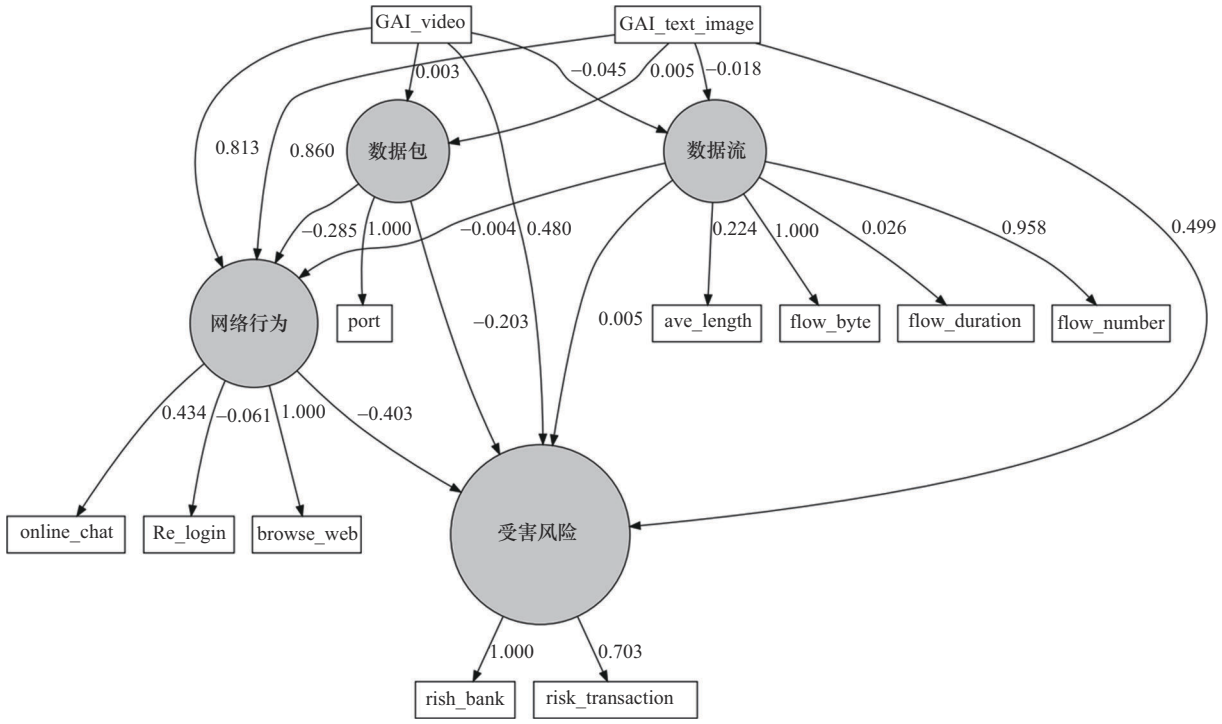

图4　结构方程模型路径系数

表5　结构方程模型的标准化路径系数结果

路径关系	路径系数	S.E.	C.R.	*p*	假设成立
受害风险←GAI_text_image	0.499	0.061	8.21	<0.001	TRUE
受害风险←GAI_video	0.480	0.057	8.359	<0.001	TRUE
受害风险←数据流	0.005	<0.001	17.095	<0.001	TRUE
受害风险←数据包	−0.203	0.02	−10.088	<0.001	TRUE
受害风险←网络行为	−0.403	0.071	−5.699	<0.001	TRUE
网络行为←数据流	−0.004	0.005	−0.887	0.375	FALSE
网络行为←数据包	−0.285	4.981	−0.057	0.954	FALSE
网络行为←GAI_text_image	0.860	0.028	30.559	<0.001	TRUE
网络行为←GAI_video	0.813	0.018	46.283	<0.001	TRUE
数据流←GAI_text_image	−0.018	0.043	−0.426	0.670	FALSE
数据流←GAI_video	−0.045	0.041	−1.098	0.272	FALSE
数据包←GAI_text_image	0.005	0.002	2.371	0.017	TRUE
数据包←GAI_video	0.002	0.002	1.244	0.213	FALSE

综上所述，研究假设H1成立，H2部分成立，而H3、H4和H5均未得到支持。以上结果表明，GAI应用主要通过直接效应和部分中介效应影响受害风险，而数据流特征与数据包特征在传导路径中的作用未达到预期效果。

3.4.2　多重中介效应分析与路径分解分析

基于算法3与算法4的设计思路，进行了多重中介效应分析与路径分解分析，相关结果见表6。

（1）直接效应：GAI_text_image和GAI_video对受害风险的直接效应值分别为0.499和0.480，贡献占比分别为35.488%和35.498%。表明GAI应用对受害风险具有显著的直接影响，且直接效应在整体影响中占据主导地位。

（2）数据流与数据包的中介效应：数据流与数据包在GAI应用与受害风险之间的中介效应均不显著。其中，“GAI_text_image→数据流→受害风险”和“GAI_video→数据流→受害风险”的贡献占比仅为0.004%和0.010%；“GAI_text_image→数据包→受害风险”和“GAI_video→数据包→受害风险”的贡献占比分别为0.044%和0.022%。表明数据流与数据包在GAI应用与网络风险之间的中介效应不显著。

（3）网络行为的中介效应：网络行为在GAI应用与受害风险之间的中介效应显著，但影响方向与预期相反。其中，“GAI_text_image→网络行为→受害风险”和“GAI_video→网络行为→受害风险”的贡献占比分别为14.872%和14.062%。这表明GAI应用通过改变网络行为间接降低了受

表6　中介效应分析与路径分解分析结果

路径	效应值	贡献占比	类型
GAI_text_image→受害风险	0.499	35.488%	直接
GAI_video→受害风险	0.480	35.498%	直接
GAI_text_image→数据流→受害风险	<0.001	0.004%	中介
GAI_text_image→数据包→受害风险	−0.001	0.044%	中介
GAI_video→数据流→受害风险	<0.001	0.010%	中介
GAI_video→数据包→受害风险	−0.001	0.022%	中介
GAI_text_image→网络行为→受害风险	−0.346	14.872%	中介
GAI_video→网络行为→受害风险	−0.328	14.062%	中介

害风险，与直接效应的方向相反。

（4）路径分解分析：根据路径分解结果可知，GAI应用对受害风险的总影响主要由直接效应和网络行为的中介效应构成。其中，直接效应的贡献占比最高（GAI_text_image占35.488%，GAI_video占35.498%），而网络行为的中介效应贡献占比次之（GAI_text_image占14.872%，GAI_video占14.062%）。数据流与数据包的中介效应贡献占比极低，几乎可以忽略。

研究结果显示，GAI技术对受害风险的直接效应贡献占比超过70%，证实了其在新型电信网络诈骗犯罪链条中的核心地位。GAI技术不仅作为犯罪工具参与伪造信息生产，更通过改变用户认知与交互模式诱导触发风险事件。同时，研究假设中“GAI→数据流/数据包特征→受害风险”传导路径的贡献度不足0.1%，即在GAI深度介入的新型电信网络诈骗场景下，单纯依赖流量监测、端口特征分析等技术指标将难以有效识别受害风险。综上所述，GAI应用对网络风险的影响主要通过直接效应实现，网络行为的中介效应虽然显著但影响方向与预期相反，而数据流与数据包的中介效应未发挥显著作用。因此，亟须建立“技术特征-用户行为-风险后果”的全链路监管框架，尤其需要针对文本生成工具的诱导性信息传播机制、视频生成技术的高仿真性欺骗特性设计专项防控策略，以此提升受害风险预警能力。

4 结束语

本文重点分析了GAI技术在电信网络诈骗犯罪中的犯罪模式及其对受害风险产生的影响。首先，对GAI驱动下的电信网络诈骗犯罪模式进行深入分析，总结了GAI驱动下的电信网络诈骗犯罪模式的3个阶段，并揭示了GAI技术在电信网络诈骗犯罪中的作用机制。其次，构建了GAI应用、数据流、数据包、网络行为和受害风险潜变量的结构方程模型。最后，对构建的结构方程模型进行了拟合度检验、路径分析等分析。研究结果表明，GAI应用对网络风险具有显著的直接效应，且直接效应在整体影响中占据了主导地位，而数据流与数据包特征的中介效应未达到预期效果，在传导路径中的作用不显著。在未来的研究中需进一步扩大数据样本，特别是要引入更多犯罪类型的真实案例数据。此外，还要持续跟踪GAI技术的发展动态，不断更新和优化当前结构方程模型，进一步提升电信网络诈骗受害风险影响因素的量化分析效果。

参考文献：

[1] 刘艳红. 网络犯罪向数字犯罪的迭代升级与刑事法应对[J]. 比较法研究, 2025(1): 1-15.
LIU Y H. The evolution of cyber crime to digital crime and the response from criminal law[J]. Journal of Comparative Law, 2025(1): 1-15.

[2] GUPTA M, AKIRI C, ARYAL K, et al. From ChatGPT to ThreatGPT: impact of generative AI in cybersecurity and privacy[J]. IEEE Access, 2023, 11: 80218-80245.

[3] GRESSEL G, PANKAJAKSHAN R, MIRSKY Y. Discussion paper: exploiting LLMs for scam automation: a looming threat[C]// Proceedings of the 3rd ACM Workshop on the Security Implications of Deepfakes and Cheapfakes. New York: ACM Press, 2024: 20-24.

[4] WU T Y, HE S Z, LIU J P, et al. A brief overview of ChatGPT: the history, status quo and potential future development[J]. IEEE/CAA Journal of Automatica Sinica, 2023, 10(5): 1122-1136.

[5] 陈德良, 王火剑, 赵加俊. 电信网络诈骗犯罪现状分析与治理对策研究: 以浙江省为例[J]. 浙江警察学院学报, 2023(6): 67-73.
CHEN D L, WANG H J, ZHAO J J. Research on the current status analysis and the countermeasures of telecom and Internet fraud: taking Zhejiang Province as an example[J]. Journal of Zhejiang Police College, 2023(6): 67-73.

[6] DELIEMA M. Elder fraud and financial exploitation: application of routine activity theory[J]. The Gerontologist, 2018, 58(4): 706-718.

[7] 刘晓倩. 电信诈骗受害者的特征分析与预防策略[J]. 法制与社会, 2021(7): 140-141.
LIU X Q. Characteristics analysis and preventive strategies of

telecom fraud victims[J]. Legal System and Society, 2021(7): 140-141.

[8] 刘晓萌. 电信诈骗年轻化趋势: 受害者心理分析与防范策略[J]. 西部学刊, 2024(24): 64-67.
LIU X M. The trend of younger victims in telecommunication fraud: psychological analysis of victims and prevention strategies[J]. Journal of Western, 2024(24): 64-67.

[9] 田丽娟, 钱虹. 电信诈骗事件中大学生受害者的心理特征分析及防范路径[J]. 西部学刊, 2023(11): 131-134.
TIAN L J, QIAN H. Analysis of psychological characteristics of college students' victims in telecom fraud and its prevention path[J]. Journal of Western, 2023(11): 131-134.

[10] 刘鑫悦, 范超云, 李辉. 电信网络诈骗被害人群体特征及防范对策[J]. 云南警官学院学报, 2022(4): 111-122.
LIU X Y, FAN C Y, LI H. Characteristics and preventive countermeasures of telecom and network fraud victims [J]. The Journal of Yunnan Police College, 2022(4): 111-122.

[11] 柳林, 吴林琳, 张春霞, 等. 接触型与非接触型犯罪时空稳定性对比及其联合防控: 以盗窃和电信网络诈骗为例[J]. 地理研究, 2022, 41(11): 2851-2865.
LIU L, WU L L, ZHANG C X, et al. Comparison of spatio-temporal stability between contact crime and non-contact crime and their joint prevention and control: a study of theft and telecommunication network fraud[J]. Geographical Research, 2022, 41(11): 2851-2865.

[12] 张帆锐, 石拓, 尤慧. 电信网络诈骗被害人群空间统计分布及风险特征研究:以刷单返利类为例[J]. 调研世界, 2024: 1-9.
ZHANG F R, SHI T, YOU H. Research on the spatial statistical distribution and risk characteristics of the telecommunications fraud victims: take for example the category of rebates on brush orders[J]. The World of Survey and Research, 2024: 1-9.

[13] 董安阳, 胡晓光. 我国电信网络诈骗犯罪的时空分布及社会人口特征[J]. 江苏警官学院学报, 2024, 39(1): 109-115.
DONG A Y, HU X G. Analysis of spatio-temporal characteristics and socio-demographic characteristics of telecom network fraud crime[J]. Journal of Jiangsu Police Institute, 2024, 39(1): 109-115.

[14] NI P F, YU W. A victim-based framework for telecom fraud analysis: a Bayesian network model[J]. Computational Intelligence and Neuroscience, 2022, 2022(1): 7937355.

[15] DAS R, AHMED W, SHARMA K, et al. Towards the development of an explainable e-commerce fake review index: an attribute analytics approach[J]. European Journal of Operational Research, 2024, 317(2): 382-400.

[16] GUO Z Y. Online disinformation and generative language models: motivations, challenges, and mitigations[C]//Proceedings of the Companion Proceedings of the ACM Web Conference 2024. New York: ACM Press, 2024: 1174-1177.

[17] FAGNI T, FALCHI F, GAMBINI M, et al. TweepFake: about detecting deepfake tweets[J]. PLoS One, 2021, 16(5): e0251415.

[18] AYOOBI N, SHAHRIAR S, MUKHERJEE A. The looming threat of fake and LLM-generated LinkedIn profiles: challenges and opportunities for detection and prevention[C]//Proceedings of the 34th ACM Conference on Hypertext and Social Media. New York: ACM Press, 2023: 1-10.

[19] 毛太田, 汤淦, 马家伟, 等. 人工智能生成内容（AIGC）用户采纳意愿影响因素识别研究: 以 ChatGPT 为例[J]. 情报科学, 2024, 42(7): 126-136.
MAO T T, TANG G, MA J W, et al. Factors influencing user adoption intention of artificial intelligence generated content(AIGC): a study on ChatGPT[J]. Information Science, 2024, 42(7): 126-136.

[20] 钱洪伟, 王旭, 高宁. 生成式人工智能 ChatGPT 风险形成机理与防范策略研究[J]. 中国应急管理科学, 2024(11): 112-122.
QIAN H W, WANG X, GAO N. Research on the formation mechanism and prevention strategies of risks in generative artificial intelligence ChatGPT[J]. Journal of China Emergency Management Science, 2024(11): 112-122.

[21] 刘仁文. 生成式人工智能财产犯罪的特点与应对[J]. 国家检察官学院学报, 2025, 33(2): 10-14.
LIU R W. Characteristics and responses to generative artificial intelligence property crimes[J]. Journal of National Prosecutors College, 2025, 33(2): 10-14.

[22] 蔡梦虹. 直播带货情境下 AI 虚拟主播对消费者冲动购买行为的影响: 基于技术可供性视角[J]. 商业经济研究, 2024(23): 81-84.
CAI M H. The influence of AI virtual anchor on consumers' impulse buying behavior in live broadcasting with goods: based on the perspective of technical availability[J]. Journal of Commercial Economics, 2024(23): 81-84.

[23] 胡剑, 戚湧. 开源创新社区用户知识共享行为影响因素研究: 基于 SEM 与 fsQCA 的实证分析[J]. 情报科学, 2023, 41(9): 59-68, 77.
HU J, QI Y. The influencing factors of users' knowledge sharing behavior in the open-source innovation community: an empirical analysis based on SEM and fsQCA[J]. Information Science, 2023, 41(9): 59-68, 77.

[作者简介]

周胜利（1982– ），男，博士，浙江警察学院信息网络安全学院教授，杭州电子科技大学网络空间安全学院硕士生导师，主要研究方向为网络安全、网络空间治理。

陈庭贵（1979– ），男，浙江工商大学统计与数学学院教授、博士生导师，主要研究方向为网络舆情演化分析。

徐睿（2000– ），男，杭州电子科技大学网络空间安全学院硕士生，主要研究方向为网络安全、机器学习。

汪邵杰（2004– ），男，浙江警察学院在读，主要研究方向为网络安全。