

网络空间安全中的人格研究综述

吴桐^① 郑康锋*^① 伍淳华^① 王秀娟^② 郑赫慈^①

^①(北京邮电大学网络空间安全学院 北京 100876)

^②(北京工业大学信息学部 北京 100124)

摘要: 网络空间是所有信息系统的集合,是人类赖以生存的信息环境。网络空间安全已经从物理域、信息域安全扩展到以人为中心的认知域和社会域安全,对人的安全研究已经成为了网络空间安全的必然趋势。人是复杂多变的综合体,而人格作为人的稳定的心理特征,成为了人的安全相关研究的切入点。该文梳理当前网络空间安全中的人格研究,介绍了网络空间安全和人格的基本概念,提出了网络空间安全中人格的研究框架,涵盖理论研究、技术研究和技术应用3个层次;其中技术研究主要包括人格测量、人格的安全脆弱性和针对人的安全防护3部分;详细探讨了当前人格安全研究现状和存在的问题,最后讨论了未来的研究方向和发展趋势。

关键词: 人格;网络空间安全;人格测量;人格脆弱性;安全防护

中图分类号: TN915.08; TP393.08

文献标识码: A

文章编号: 1009-5896(2020)12-2827-14

DOI: 10.11999/JEIT190806

A Survey on Personality in Cyberspace Security

WU Tong^① ZHENG Kangfeng^① WU Chunhua^①

WANG Xiujuan^② ZHENG Heci^①

^①(School of Cyberspace Security, Beijing University of Posts and Telecommunications, Beijing 100876, China)

^②(Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China)

Abstract: Cyberspace is a collection of all information systems, which refers to the information environment for human survival. Cyberspace security has expanded from physical and information domain security to human-centered social and cognitive domain security. The research on human security has become an inevitable trend of Cyberspace security. The characteristics of human are complex and changeable. The personality, as a stable psychological characteristic, is an appropriate breakthrough point for human security research. This paper investigates and untangles relevant personality research in Cyberspace security. The concepts of Cyberspace security and personality are introduced concisely. A research framework of personality in Cyberspace security is proposed, including theoretical research, technological research and technological application. The technological research mainly includes three parts: personality measurement, personality vulnerability and protection methods in Cyber security. Additionally, this paper discusses the current research status and problems of personality in Cyber security in detail. Finally, the future research directions and development trends of personality in Cyber security are explored.

Key words: Personality; Cyberspace security; Personality measurement; Personality vulnerability; Safety protection

1 引言

网络空间领域可划分为物理域、信息域、认知域和社会域。从网络空间安全的角度来看,物理域

和信息域主要关注以信息技术为核心的网络基础设施安全及网络信息通信安全,而认知域和社会域更关注以人为核心的认知文化等精神层面以及个人与集体相互作用的社会层面。以往传统的信息安全较多地注重信息系统的物理安全和技术安全,即更加关注物理域和信息域,对同样属于网络空间领域的“人”缺乏相应的重视。据统计,在已识别的安全缺陷中,有80%是由人为因素引起的。美国计算机犯罪和安全调查局资料调查同样显示,网络安全事

收稿日期: 2019-10-18; 改回日期: 2020-09-12; 网络出版: 2020-09-22

*通信作者: 郑康锋 zkf_bupt@163.com

基金项目: 国家重点研发计划(2017YFB0802803)

Foundation Item: The National Key Research and Development Program of China (2017YFB0802803)

故中,人为因素占到52%,技术错误仅占到10%。在过去的几十年中,安全领域的专家们主要还是通过升级系统防护软件、修复安全漏洞、加固硬件设备等技术手段保障网络空间安全,达到对抗恶意攻击的目的。而事实上,只有将软硬件技术与人的结合在一起考虑,才能形成真正有效的安全保障体系。

人在网络空间中的行为是复杂多样的,受到安全知识、安全意识等水平的调节,受到思维、情绪等心理活动的支配,也受道德观、人生观、价值观的影响。研究者试图从这些复杂的因素中选择较为基础、稳定的因素来描述人、解释人的行为。在诸多心理学特征中,人格是一个统一的、相对稳定的结构组织,其在不同时间、区域下影响着人的内隐和外显的心理特征和行为模式。一个人的人格在不同时期、不同情境下具有相对的一致性,因此被很多领域的研究者视作一个重要的心理特征,广泛应用于与人有关的研究中。

网络空间安全领域的研究者试图将人格与网络安全行为联系起来。由于人的认知复杂性,人格理论的研究仍没有统一的标准,这导致人格与网络安全行为的关联研究有很大的不确定性。此外,人格与网络安全行为之间的作用关系受到多种因素的影响,并不能明确其中各个因素的作用形式,这也增加了人格在网络空间安全中研究的难度。当前的研究者为了解决这些问题,提出了许多可以借鉴的方法,例如,选用比较经典的大五人格理论开展研究,采用控制变量的方法控制作用因素等。这些方法都不断推进了网络空间安全中人格的相关研究,为网络空间安全认知域研究的发展做出了巨大贡献。

本文将对网络空间安全中的人格研究进行系统性的总结。

2 网络空间安全与人格概述

网络空间中人的因素较以往受到更加广泛的关注,人们认识到“人”是安全体系中潜在的、脆弱的一环,既有可能在技术的运用、安全策略的遵守上出现漏洞,也有被攻击者利用的危险。一些研究者希望运用心理学的理论了解安全体系中“人”的行为,其中人格就是一个备受关注的心理学特征。本节主要介绍网络空间安全及人格的概念,并且对网络空间安全中人格的研究做出简单介绍,提出网络空间安全中的人格研究框架。

2.1 网络空间安全

在19世纪80年代,Gibson^[1]在其短篇科幻小说《燃烧的铬》中创造了Cyberspace一词,意指由计算机创建的虚拟信息空间。2008年,美国第54号总统令对Cyberspace进行了定义: Cyberspace是信息

环境中的一个整体域,它由独立且互相依存的信息基础设施和网络组成,包括互联网、电信网、计算机系统、嵌入式处理器和控制器系统^[2]。Ottis和Lorents^[3]提出网络空间是随时间衍变的、互联的信息系统与人的合集。张焕国等人^[2]提出网络空间是信息时代人们赖以生存的信息环境,是所有信息系统的集合,人在其中与信息相互作用相互影响。

相较于传统信息网络,网络空间更加强调“人”的因素。在网络空间安全体系中,也更注重人的安全。人是安全体系中重要的一环,一个组织中的内部人员,由于具有组织内部系统、网络以及数据的访问权,一旦他们利用自身合法身份对组织信息系统中信息的机密性、完整性以及可用性造成负面影响,就会形成所谓的内部威胁,显然这样的威胁是传统防火墙、入侵检测系统等防护工具所不能抵御的。因此,仅仅关注以信息技术为主的物理域、信息域安全已经不能满足安全体系的要求,需要对认知域、社会域的安全加以研究。

近年来,以社会工程学^[4-8]为代表的,面向“人”的攻击事件频发。社会工程学由来已久,目前已经随着移动通信网、互联网的飞速发展,逐步扩展到了网络空间安全领域,成为网络空间中的重要安全问题。网络空间安全中的社会工程学概念普遍认为由凯文·米特尼克在其《欺骗的艺术》^[4]一书中正式提出。根据书中描述,可以将网络空间中的社会工程学总结为:通过自然的、社会的和制度上的途径,利用人的心理弱点(如人的本能反应、好奇心、信任、贪婪)以及规则制度上的漏洞,在攻击者和被攻击者之间建立起信任关系,获得有价值的信息,最终可以通过未经用户授权的路径访问某些敏感数据和隐私数据^[8]。社会工程学是典型的面向人的认知域和社会域的攻击形式,可以作为研究网络空间认知域和社会域安全的突破口。

2.2 人格理论概述

人格作为人稳定的心理特征,已经成为包括网络空间安全在内的众多领域中与人有关研究的切入点。为了运用人格理论,本文对人格理论做出了一个系统性的概述。

Burger^[9]将人格定义为稳定的行为方式和发生在个体身上的人际过程。彭聃龄^[10]提出人格是构成一个人的思想、情感及行为的特有的统合模式,这个独特模式包含了一个人区别于他人的稳定而统一的心理品质。由以上定义可以看出,人格具有稳定性,可以作为网络空间安全中人的研究的一个要素。目前在心理学领域,就如何描述人格的结构的问题划分了6个流派^[9],分别站在不同的知识领域来解释人格,这6个流派分别是精神分析流派、特

质流派、生物学流派、人本主义流派、行为主义流派以及认知流派。

在精神分析流派的理论中,人格的核心是一个人头脑中的各种事件(即内心事件, intrapsychic event), 这些事件是产生行为的动机。通常情况下人们会意识到这些动机, 然而, 某些动机也会产生在非意识层面。精神分析理论的创始人弗洛伊德认为, 所有的行为都是由动机引发的。精神分析理论^[1]将人格结构分为本我、自我和超我。本我是原始驱动力的源泉, 本我的行动是非理性的, 追求即刻的满足感, 而不考虑现实的可行性。超我负责内化社会价值、道德和观念, 超我的发展也就是一个人价值观逐步建立的过程, 超我和本我经常存在冲突, 本我想做快乐的事, 而超我坚持做应该做的事。自我的行动基于现实, 调和本我冲动和超我需求之间的冲突。

特质流派采用特质对人格进行描述, 在不同特质上对个体进行等级评定。特质是持久的品质或特征, 这些品质或特征使个体在各种情况下的行为具有一致性。一些特质理论家认为特质是引起行为的先决条件, 但更加保守的理论家仅仅将特质作为描述性唯独, 认为它只是简单地概括了被观测到的行为模式。特质流派的人格理论众多, 其中经过了理论实践验证的包括艾森克的人格层次模型、卡特尔的16因素模型、人格环形分类模型和五因素模型。

生物学流派关注的是影响人格或被人格影响的内在生理因素^[2]。生物学流派将生物学领域与人格结合起来, 生物学流派会研究遗传学与人格的关系, 例如基因如何影响人格, 以及生理系统, 例如大脑或神经系统如何影响人格, 他们还会以达尔文进化论为基础, 研究人格的进化论观点。

人本主义流派从整合个体的个人经验和意识经验及成长潜能的角度理解人格, 其核心是强调自我实现的驱力。马斯洛最早提出了自我实现这一术语, 他的需求层次理论将人类需求从低到高划分为生理需要、安全需要、归属需要、尊重需要和自我实现。在马斯洛之后, 一些研究者提出了自我实现的测量方法。

行为主义流派将人格等同于行为, 认为人格是一个人的条件反射经验所导致的结果。华生在20世纪早期创立行为主义, 主张心理学应研究行为而非意识。社会学习理论者又对行为主义理论进行了补充, 罗特引入期望理论来预测行为, 班杜拉则提出个体因素、行为和环境刺激三者构成的复杂的互动系统^[3], 他指出行为会受到态度、观念、之前的强化经历以及环境中的刺激所影响; 行为也可以对环

境产生作用, 同时人格中某些重要成分受到环境和行为反馈的影响。

认知流派认为, 个体人格的不同是由人们信息加工的方式不同引起的。人格心理学家感兴趣的认知水平包含3个方面^[4], 第1个水平是知觉, 即对感觉器官手机的信息施加指令的过程, 在知觉水平上, 人们看到的世界也有着很大的不同, 这些不同与每个人的人格差异有着联系。第2个水平是解释, 即给生活中各种事件赋予意义, 个体对各种事件的解释会揭示人格。第3个认知水平是意识目标, 即个体形成的评估自己和他人的标准。

人格心理学的6个流派理解人格的角度颇有不同, 人格中的各个流派都强调并且验证了人格的一个重要方面, 例如人本主义学派从内在机制上提供了对人格的解释, 却没有对特定行为进行预测, 行为主义则强调外部环境对未来行为预测的重要性。特质流派则认为每个人类个体都处在各种各样人格特质连续体的某个比较固定的位置上, 从这个角度上看, 特质流派并未对人格给出过多的原理性的解释, 而是给人格提供了一种相对科学可靠的分析和量化的可能。

在网络空间安全的研究中, 主要采用的是特质流派的人格理论, 并且使用的人格模型就是特质流派中最为主流的大五人格模型(big five factors model)^[4,15], 它从神经质(neuroticism)、外向性(extraversion)、开放性(openness)、宜人性(agreeableness)和尽责性(conscientiousness)5个方面描述人格。这5个特质的具体描述如表1所示。

2.3 网络空间安全中的人格研究

本节概括性地提出了网络空间安全中人格的研究框架, 如图1所示。研究框架划分了网络空间安全中人格研究的层次, 包括理论研究、技术研究和

表1 大五人格模型

特质因素	特征
神经质	烦恼对平静
	不安全感对安全感 自怜对自我满意
外向性	好交际对不好交际
	爱娱乐对严肃 感情丰富对含蓄
开放性	富于想象对务实
	寻求变化对遵守惯例 自主对顺从
宜人性	热心对无情
	信赖对怀疑 乐于助人对不合作
尽责性	有序对无序
	谨慎细心对粗心大意 自律对意志薄弱

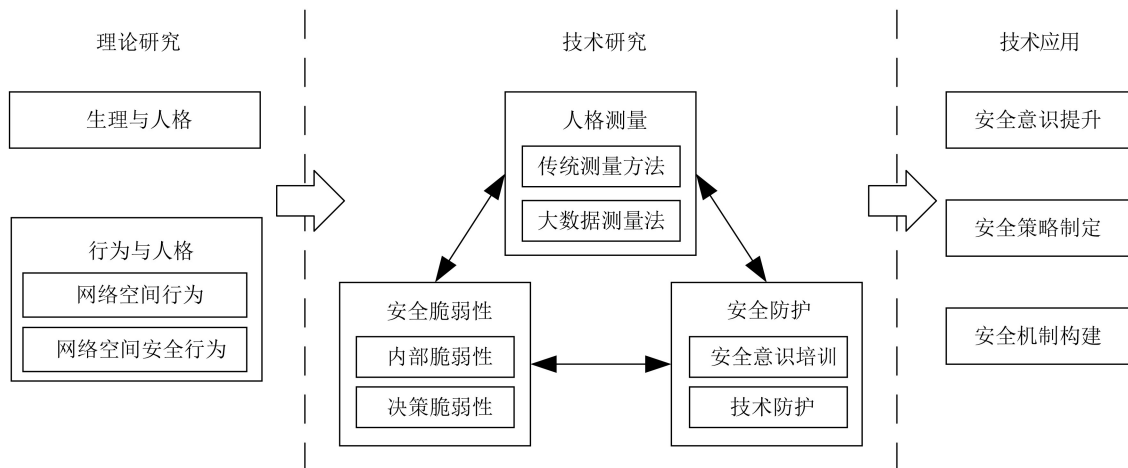


图1 网络空间安全中的人格研究框架

技术应用3个层次。如2.2节所述,人格的理论研究大致可以分为基于生理的研究和基于行为的研究两种,其中网络空间中的人格研究一般是基于网络空间行为的。此外,针对网络空间安全的人格理论研究是建立在网络空间中与安全相关的行为之上的。网络空间安全中人格相关的技术研究建立在理论研究的基础上,从人格测量、安全脆弱性挖掘与安全防护技术3个方面展开,下文将会详细探讨这一层次。网络空间安全中人格研究的技术应用一般体现在安全意识提升、安全策略制定和安全机制构建3方面,目的是提升网络安全等级。

本文重点探讨了技术研究这一层面,在网络空间安全领域,对于人格的技术研究主要包括人格的测量、安全脆弱性的研究、安全防护的研究3个方面。其中人格测量技术是进行其他人格相关研究的基础,安全脆弱性的挖掘是建立人格与网络空间安全关联研究的关键,安全防护技术是进行网络空间安全中人格研究的应用技术基础。这3个方面的技术研究相互支撑,共同构成了网络空间安全中人格研究的技术基础。

人格的测量是进行人格研究的基础。传统上,研究者为了获取人格数据,需要召集一些被试人员,对其进行一系列的人格测验,获取人格信息。这种测量的方法需要被试者的主动参与,难以获得大规模的样本。因此,一些社会计算领域的研究者利用大数据的方法,通过大量收集网络数据,从其中的网络行为、文本内容等信息^[16-18]中推测用户人格,达到大规模获取样本的目的,实现统计有效性。

人格与安全脆弱性的研究是网络空间安全中人格研究的核心。这一部分的研究旨在了解网络空间安全中人格脆弱性的主要表现,以网络空间安全行为为基础,挖掘人格安全脆弱性。总结已有的研

究,可以将脆弱性归为内部脆弱性和决策脆弱性两方面。内部脆弱性主要研究组织内部人员的人格如何影响其安全行为;决策脆弱性主要研究人在受到社会工程学攻击时,人格会如何影响人做出正确决策,避免受到欺骗的能力。

对于人格表现出的安全脆弱性,需要提出相应的安全防护策略。一方面,要根据安全体系中不同的人格差异,进行针对性的安全意识培训,另外,对于人格的安全脆弱性也需要开发有效的技术防护措施。

3 人格测量

人格测量是研究中获取被试者人格数据的方法。在具体的研究过程中,根据实验条件的不同,需要采用适当的人格测量方法,以获取相对准确的人格数据。传统测量方法包括自陈量表、投射测验等,每种方法适用于不同的实验情境。另外,近年来有大量的研究采用大数据方法分析人格,这些研究试图分析文本内容、网络行为、终端设备、肢体动作等数据与人格之间的关联,从而建立起客观模型用于人格的预测。

3.1 传统测量方法

3.1.1 自陈量表

自陈量表^[19,20]是一种客观化测验,被试者需要回答不同情况下与自身的行为、感受、兴趣有关的一系列问题。这些问题通常以陈述句的形式给出,被试者需要评估自身与陈述句给出情况的相符程度。自陈量表包括非结构化测验和结构化测验。

非结构化测验也被称作开放式人格测验。一个重要的人格测验是由库恩和麦克帕特兰编制的20项人格测验^[21],该测验要求被试补充20个以“我是”开头的句子,通过这20个陈述句完成对自身的描述。例如,一个人可能会做出这样的回答:我20岁;我是一个男孩;我是一个胆小的人;我周末

喜欢一个人在家度过；等等。一般而言，这类测试的特点就是被试者要做出一系列无结构的开放式回答，并且针对收集到的测验数据，研究者必须有一套解释以及打分的方法，以便从开放式的回答中获取背后的意义。

结构化测验则更为普遍。与开放式测验不同的是，结构化测试量表的反应选项都是固定的。常用的结构化人格测试量表包括艾森克人格问卷(EPQ)、明尼苏达多项人格测验(MMPI)和卡特尔16因素人格测验(16PF)，以及目前最为流行的大五人格量表(BFI和NEO)。在常用的BFI-44大五人格量表中，每个题目都是一个描述个人特质的陈述句，例如第1题“我是一个健谈的人”。被试者需要用1~5的数字评估自身与该陈述句的符合程度，符合程度1~5逐渐增加，1代表非常不赞同，5代表非常赞同。

自陈量表的优点是容易解释，可操作性强，应用较为广泛。缺点在于被试者可能缺乏准确的自我认识，或者倾向于以积极的方式展现自己，从而造成自陈量表存在局限性。即便如此，自陈量表仍然是人格测量中最为常用的方法。

3.1.2 投射测验

投射测验^[22]也是一类重要的人格数据来源。使用投射测验时，研究者会给被试者提供一种标准化的多义刺激，要求他们短时间内表达自己看到了什么，由于反应时间短，被试者没有充足的时间进行全面的思考，因此常常会把自已最真实的情绪、态度、需要、动机、观点等心理活动表现出来。

一个著名的投射测验就是1921年罗夏克发明的墨迹测验^[23]。在墨迹测验中，研究者展示给被试者多张墨迹的图片，要求被试者描述看到了什么，由此想到了什么。被试者会在不经意间将自己的思想感情和对事物的态度表现出来，研究者再根据这些反应去判断被试者的人格特征。

另一个著名的投射测验^[24-26]就是美国心理学家Murray和Morgen在1935年编制的主体统觉测验。由30张模棱两可的图片和一张空白图片组成。图片内容多为人物，也有部分风景，但每张图片都至少有一个人物。每次给被试呈现一张图片，让被试根据看到的内容编故事。每次被试都必须回答这样4个问题：(1)图中发生了什么事？(2)为什么会出现这种情境？(3)图中的人物正在想什么？(4)故事的结局会怎样？投射测验的原理在于，人们对外界事物的看法，往往受到自己的心理特征的影响，这种现象称之为投射效应。通过分析被测试者对这些材料的解释，便可以探究其内心深处的心理特征。

投射测验的优点在于有机会把被试者最原始的

心理特征表现出来，缺点在于评分机制缺乏客观标准，测验的结果难以解释。

3.2 大数据测量法

传统人格测量方法的缺陷在于样本规模过小，难以实现统计有效性^[27]。部分研究者利用大数据方法分析人格。这类研究的基本方法是通过让用户填写心理量表获取用户的人格属性，同时获取用户在浏览网络或使用终端设备时产生的与用户自身相关的数据，进而利用大数据分析的方法将用户属性与用户相关数据进行关联，发现其中规律，建立预测模型。根据用户相关数据来源的不同，主要可分为基于文本内容、基于社交网络、基于终端设备和基于人体语言的4种人格分析方法。

3.2.1 基于文本内容的人格分析方法

Vinciarelli和Mohammadi^[28]指出，在语言心理学的理论中，一个人对于语言中词语的选择不仅仅取决于词语本身的含义，同样也受情绪、态度以及人格特质等心理现象的影响。因此，结合社会语言学和文本分析技术，从文本内容中推断人格特质的方法也就成为了可能。

利用文本内容分析人格的研究在多年前就已经开展，这一方向最早的一个研究就是在2005年，Argamon等人^[29]召集了1200名学生所进行的实验。该实验收集了这些学生所写的2263篇论文，将这些论文的词语分为4个具有心理学意义的类别，并分别计算每个类别词汇的词频，以此将文章转化为词汇学特征，随后利用支持向量机建立预测模型，在外向性和神经质这两个特质上可获得大约58%的准确率。Mairesse等人^[30]扩充了上述实验的数据集，并且利用语词计量工具LIWC(Linguistic Inquiry and Word Count)^[31]，生成了88个词汇学特征，利用支持向量机训练得到的预测模型在5个人格特质上分别可达到50%~62%的准确率。

从文本数据的选择上看，早期研究者使用的文本数据是论文或是一些作家的文章，这一类文本的差异主要体现在作者的写作风格上。而近期对文本内容分析的研究集中在个人博客^[32-36]及社交网络状态上，这些文本信息不仅体现了写作风格，而且与用户的个人事务及经历更加密切相关，因此在语言中可能会隐藏更多与人格有关的信息。

从分析方法上看，一部分针对文本内容的研究依赖语词计量工具，通过预先定义的词典，对文本中的词汇先进行分类统计，以统计量作为文本特征。LIWC是人格分析中最为常用的语词计量工具^[37]。它可以将文本中不同类别的词语，特别是心理学类的词语进行统计，如因果词、情绪词、认知词等词

类在整个文本中的使用百分比。另外,也有研究不依赖语词计量工具, Schwartz等人^[38]就直接提取了文本中的单词、n-gram短语、主题作为文本特征,并且计算了这些文本特征与人格之间的相关程度,进行了可视化的展示。

3.2.2 基于社交网络的人格分析方法

社交网络是人们进行线上互动的主要渠道,也是用户进行自我展示的重要途径^[39],因此社交网络是进行人格分析的一个理想的平台,利用社交网络分析人格的研究众多^[40-42]。根据社交网络中不同类型的数据,可以产生不同类型的特征,用于分析人格。

社交网络的使用统计数据是用于分析人格的重要数据。人格的差异会影响用户对社交网络的使用,研究者一般会总结人格与社交网络使用行为之间的关系,对用户社交网络的使用情况数据进行统计,并将这些统计量作为特征,建立人格分析模型。Ross等人^[43]的研究是最早利用社交网络进行人格分析的研究,他们采用问卷的方式收集被试者facebook的使用情况数据,例如每天使用facebook的时长、发布的评论数等,将这些使用数据作为特征,分析了它们与大五人格之间的相关性。研究发现高外向性人群朋友数较多,并且会加入更多的群组,高神经质人群会把更多的时间花费在社交平台上,高尽责性人群则会在社交平台上花费更少的时间,这些结果在一定程度上证明了人格特点会反映在用户社交网络的使用情况。Amichai和Vinitzky等人^[44]基于Ross的实验进行了改进,不再采用问卷方式收集facebook的使用数据,而是抓取真实的facebook数据,得到了与Ross的研究相同的结果,并且得到的相关性更加显著。Li等人^[45]则利用微博的数据,提取了静态与动态两类特征,静态特征包括用户的账户信息、隐私设置信息等;动态特征与时间序列有关,包括一段时间内更新微博的频率,更博数量等。利用这两类特征,结合SVM和Pace回归算法训练人格预测模型。研究结果表明,基于微博行为的人格计算模型拥有良好的测量属性。在SVM模型中,微博行为对各人格维度高低得分组被试的分类精度达到84%~92%;而在Pace回归模型中,基于微博行为的人格预测结果与基于自评问卷的人格测验结果之间的相关系数达到0.48~0.54。

另外,社交网络的结构数据也被用于人格分析的研究中。人格是影响一个人社交环境的重要内因,它会影响到社交关系的类型和规模,例如,Swickert等人^[46]的研究中就指出外向性得分与社交

网络规模呈正相关。Friggeri等人^[47,48]详细地解释了自我中心的社交网络的拓扑结构与外向性之间的关系,内向型人群往往会加入数量较少,但结构更加密集的社交群体;外向型人群往往会加入数量更多、结构较为松散、规模较小的社交群体,并且会成为各个社交群体之间的桥梁人物。

值得一提的是,为了进行社交网络中的人格分析, Kosinski等人特意建立了名为myPersonality^[49]的项目,持续地收集了facebook用户的信息以及人格数据,该项目主页目前也开放下载了部分数据,有许多的人格相关研究都基于myPersonality数据集^[38,50,51]。

3.2.3 基于终端设备数据的方法

和社交网络一样,智能手机、可穿戴设备已经深入到人们的日常生活中,人们日常行为活动都会被终端设备以不同的方式记录下来,使得终端设备也是进行人格分析的重要数据来源。Olguin和Gloor等人^[52]利用可穿戴设备收集了67名参与者谈话、移动、地理位置等特征,通过统计学方法研究人格与这些特征之间的关系。Chittaranjan和Blom等人^[53,54]以智能手机为研究对象,收集了117名被试者17个月的手机使用记录,研究这些数据与这些被试者人格之间的关系。Staiano和Pianesi等人^[55]则通过智能手机的数据提取用户的社交网络特征,利用这些特征推断用户人格。

3.2.4 基于人体语言的分析方法

心理学研究表明,人体语言,或非言语行为,也会体现不同的人格。Biel等人^[56]采集了Youtube视频中人的面部表情,并且分析了这些面部表情表达的情绪,探究了这些情绪与人格之间的关系。Skowron等人^[57]提取了Twitter和Instagram照片中人的面部表情和身体姿态的特征,研究了这些特征与人格之间的关系。Ferwerda等人^[58]认为人格会影响用户在照片中展示自己的方式,因此收集了112名Instagram用户的照片,提取了其中RGB的数值、亮度、对比度、饱和度等图像特征,分析了这些特征和人格之间的关联关系。

4 人格与安全脆弱性

人的安全脆弱性体现在两个方面,一方面是由于安全意识薄弱或行为与意识的不一致造成的安全脆弱性,会无意地破坏内部安全,成为内部威胁;另一方面是由于内在的心理弱点,会使得人在遭受主动攻击,如受到社会工程学攻击时,无法做出正确的决定。本节将会对这两类研究的起源、研究的方法与结果进行归纳与总结。

4.1 人格与内部脆弱性

由于内部特权用户可以利用自身权限绕过系统内部的访问控制机制，因此一旦内部用户对系统操作不当，或是被潜在的攻击者所利用，就有可能造成以合法手段发动内部攻击的情况，而用于防范外部威胁的防火墙、入侵检测系统、漏洞扫描等传统安全产品对内部威胁基本无能为力。这就要求除了发展传统技术手段之外，还要研究内部人员自身的因素对安全行为的影响，从而提出针对人的安全防护措施。人格在许多研究中被视作人自身的主要因素，其中部分研究认为安全意识能够决定人的安全行为，因此将问题转化为人格与安全意识之间的关系。但也有研究发现安全行为与意识之间存在不一致性，因此将人格运用到了意识与行为一致性的研究中来。

4.1.1 安全意识

为了避免潜在的内部威胁，无论是政府、企业、任何组织机构的管理者，还是一些安全领域的研究者，都对内部人员的安全意识极为重视。在许多过往的研究中，良好的安全意识都被当作一个人实施安全行为的首要前提。人与人之间的安全意识存在差异，部分研究者希望找到人格与安全意识之间的联系。

Warkentin和McBride^[59]利用大五人格理论来解释内部人员在安全策略遵守上存在的意识差异，该研究提出了一个评估模型，如图2所示。其总结了以往研究中的保护动机理论和一般威慑理论，认为安全策略遵守的意识是人们根据威胁评估(threat appraisal)，应对评估(coping appraisal)，惩罚评估(sanction perspectives)这三个方面共同作用而形

成的决策。威胁评估是指人们对威胁的认识，包含2个因素的内容，分别为威胁严重性(threat severity)和威胁易感性(threat vulnerability)。威胁严重性是人们对威胁严重性的判断，包括威胁产生的经济损失和社会影响；威胁易感性是人们对于自己违反安全策略或者遭受攻击的可能性的主观判断后形成的主要信念，包括个体对威胁提醒的接受程度。应对评估是评价个体应付和避免威胁的能力，是人们对威胁处理能力的认识，包含有3个因素的问题，分别为：自我效能(self-efficacy)、反应效能(response efficacy)和反应代价(response costs)。自我效能是指个体对自己采取某种保护性能力的知觉，即个体在执行某一行为操作前对自己能够在什么水平上完成该行为活动所具有的信念、判断和主观自我感受，是保护动机理论的核心部分；反应效能指个体对所采取的某种保护性行为是否起作用的知觉。一般而言，人们采取一种行动是因为相信他们将会从这一行动中获益，而且这种益处对个人有意义。反应代价，指个体采取某种保护性行为所付出的社会或者经济方面的代价，是一种阻止人们采取某种行为的障碍或影响保护行为的反作用力。惩罚评估包括惩罚严重性(sanction severity)和确定性(sanction certainty)两方面，都只是与人们的个人感知有关而非实际的惩罚的严重性和检测到的可能性。当个体感知惩罚较重或是惩罚可能性较大时，会趋向于采取某种安全行为或保护性行为。

该研究将大五人格理论加入到评估模型中，研究人格对这3个方面的评估所产生的影响，从而得到人格对个人安全策略遵守意识的影响。研究结果表明，高宜人性人群的自我效能感较高，因此会更

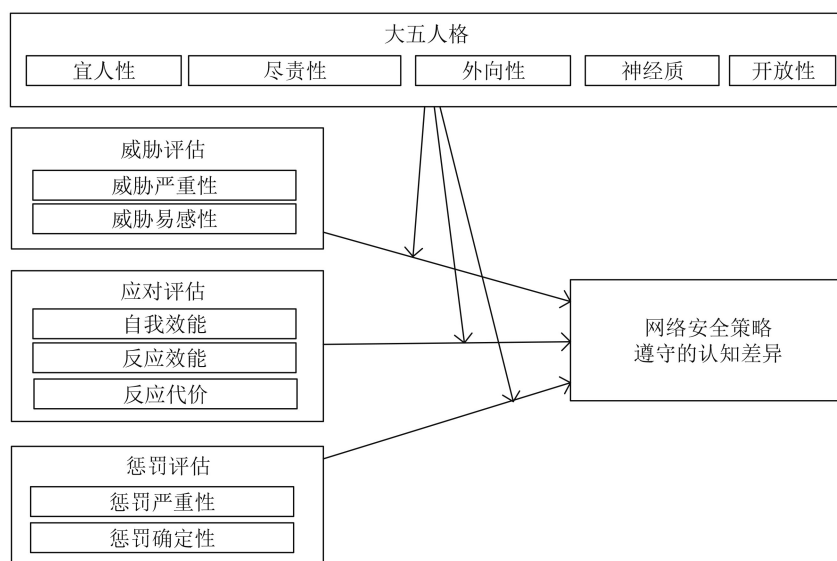


图2 安全意识评估模型^[59]

加积极地遵守安全策略。高尽责性人群相比于低尽责性人群,更会被惩罚评估所影响,因此遵守安全策略的意识更强。而高神经质与高开放性人群更容易关注反应代价,即遵守安全策略所要付出的社会或经济代价,因此相对于低分人群,遵守安全策略的意识较弱。

在其他的一些研究中, Korzaan等人^[60]调查研究人格如何影响个体对隐私的重视程度,发现高宜人性人群对于隐私保护的重视程度较高。Uffen和Breitner^[61]调查了889名被试者的人格数据,并设计问卷用于评估每位被试的安全意识,通过统计学的方法研究人格与安全意识之间的联系,也发现高宜人性人群安全意识较高。Brecht等人^[62]通过收集24名用户对匿名网络的使用数据,研究人格对个体隐私保护意识的影响,得到高神经质人群隐私保护意识高的结论。

这些研究结果表明安全意识存在着个体差异,研究的意义在于:安全意识的培训应当充分考虑人的差异性,针对不同人格的人可以设计更加适合的安全培训计划。例如Warkentin和McBride等人^[59]的研究中就提出了一个针对人格差异的安全培训框架,对各种不同的人格特质定制培训规范。不过目前,诸如此类的个性化安全培训框架还处于概念阶段,还没有研究制定出更为详细具体的内容。

4.1.2 意识与行为一致性

Shropshire和Warkentin等人^[63]的研究指出,内部人员即使有意识去遵守安全策略,仍然会成为安全防护中最脆弱的一环,因为他们真实的行为与他们的意识存在不一致性。基于这样的观点,他们希望利用人格理论来解释安全意识与行为存在的不一致问题,并建立了人格与意识、行为的关联模型,如图3所示。他们借鉴了技术接受模型等理论,认为个体的感知有用性(perceived usefulness)和感知易用性(perceived ease of use)以及组织支持感(perceived organization)与个体安全意识之间存在

正相关,因此对安全意识的评估可以通过这3个因素,同时指出宜人性和神经质两个人格维度会影响意识与行为的一致性,为此他们开发了一个用于评估系统安全等级的软件,评估被试者4周内个人电脑的安全等级,以此对被试者的安全行为进行评估。研究结果表明,安全意识较高的人群中,高尽责性人群和高宜人性人群的行为与意识更加一致。

从传统的观点来看,内部人员的安全意识提高有助于保护组织内部的信息网络安全,但以上的研究表明,即使内部人员的安全意识较高,但由于人格会影响意识与行为的一致性,因此仍然有造成内部威胁的风险。

4.2 人格与决策脆弱性

对于人格与安全脆弱性的另一个重要研究就是人格与决策脆弱性的关系。这里的决策脆弱性是指在社会工程学攻击中,即使目标本身具有良好的安全意识,但社会工程师仍然可以利用目标的心理弱点,让目标做出一些未经过仔细思考的、缺乏逻辑的决策。在受到社会工程学攻击时,做出正确决策的能力存在个体差异,一些研究希望通过人格理论探寻其中的差异,并对不同的个体提出相应的安全防护措施。

在人格与决策脆弱性的研究中,一部分研究以社会工程学实验为主要研究方法,一般是收集被试者的人格数据,并且向被试者发送钓鱼邮件,观察他们的反馈情况(即选择相信、点击邮件中的链接或者选择不相信)^[64]。利用这些数据探究被试者在收到钓鱼邮件时,人格是否会影响他们做出的反馈。Pattinson和Jerram^[65]就设计了这种社会工程学实验,发现只有开放性这一人格维度影响个体对钓鱼邮件的反馈,开放性得分越高,越不容易相信钓鱼邮件。这一研究结果似乎与人们的直觉相反,在过往的一些研究者中,高外向性与高开放性人群由于更容易相信他人,因此更有可能把钓鱼邮件视作正常邮件。研究人员对这一结果做出的解释是,高开放性人群的学习能力更强,会更加有意识地积

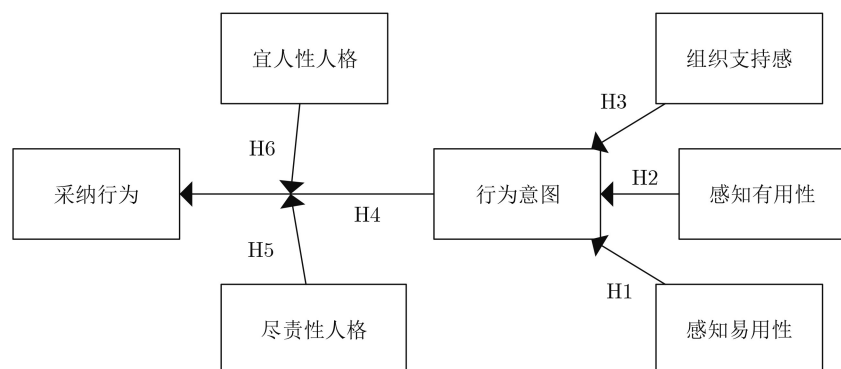


图3 人格与意识、行为关联模型^[63]

累辨别钓鱼邮件的经验。另外，Darwish等人^[66]设计了类似的社会工程学实验，发现高宜人性人群更容易相信钓鱼邮件，高尽责性人群一般不会相信钓鱼邮件，研究者认为这与高宜人性人群容易相信别人、而高尽责性人群更加谨慎的特点有关。Enos等人^[67]在事先不告知实验目的的条件下向100名被试者发送包含中奖信息的钓鱼邮件，统计了被试者的反馈情况，发现高神经质的女性更容易相信钓鱼邮件，研究人员对此的解释是高神经质的女性人群更加关注自身的情感需求，更趋向于相信互联网拥有满足她们需求的能力。之后，Halevi和Memon等人^[68]又采用了类似的实验方法，并且试图利用人格特点针对性地设计钓鱼邮件内容，实现鱼叉式邮件攻击，例如，利用高尽责性员工追求高效率的特点，以管理者的身份向他们发送钓鱼邮件，要求他们尽快填写表格。结果表明，尽责性与点击钓鱼邮件链接这一行为之间存在正相关。

从以往的这些实证研究中可以看出，人格会在一定程度上影响个体对钓鱼邮件的反馈，但这种影响在不同的研究中并不是完全一致的。除了实验本身的局限性之外，一方面的原因是由于任何一个人格维度都囊括了许多性格特征，而相同人格维度下

的性格特征有些会促使个体相信钓鱼邮件，有些反而会帮助个体辨别钓鱼邮件。例如高开放性人群拥有强烈的好奇心，可能会导致他们点击一个钓鱼邮件的链接，但他们对于过往经验积极学习的态度也会帮助他们辨识钓鱼邮件。此外，钓鱼邮件内容的差异也会影响个体做出不同的反应，高尽责性人群做事更加缜密，相比于粗心大意的人可能更容易辨别钓鱼邮件，但是高尽责性人群守时、服从权威的特点也有可能被鱼叉式的钓鱼邮件所利用。因此，利用社会工程学实验的方法探究人格与决策脆弱性的关系的这类研究中，会出现一些存在争议的研究结果。

另外，有一些研究通过总结已有文献，从理论角度提出了一些研究模型。文献^[69]就对以往钓鱼邮件实验的研究做出了总结，提出了一个以大五人格理论为核心的钓鱼易感性研究框架，如图4所示。这一框架主要由个人因素、经验因素、大五人格因素、钓鱼易感性4方面组成，其中个人因素，例如性别、年龄，作为人与生俱来的内因，会决定人的经验因素，个人因素与经验因素又会同时对个体人格的形成产生影响，而人格会影响个体对钓鱼邮件的反馈。文献^[70]指出社会工程学主要通过权威(authority)、承诺与一致(commitment &

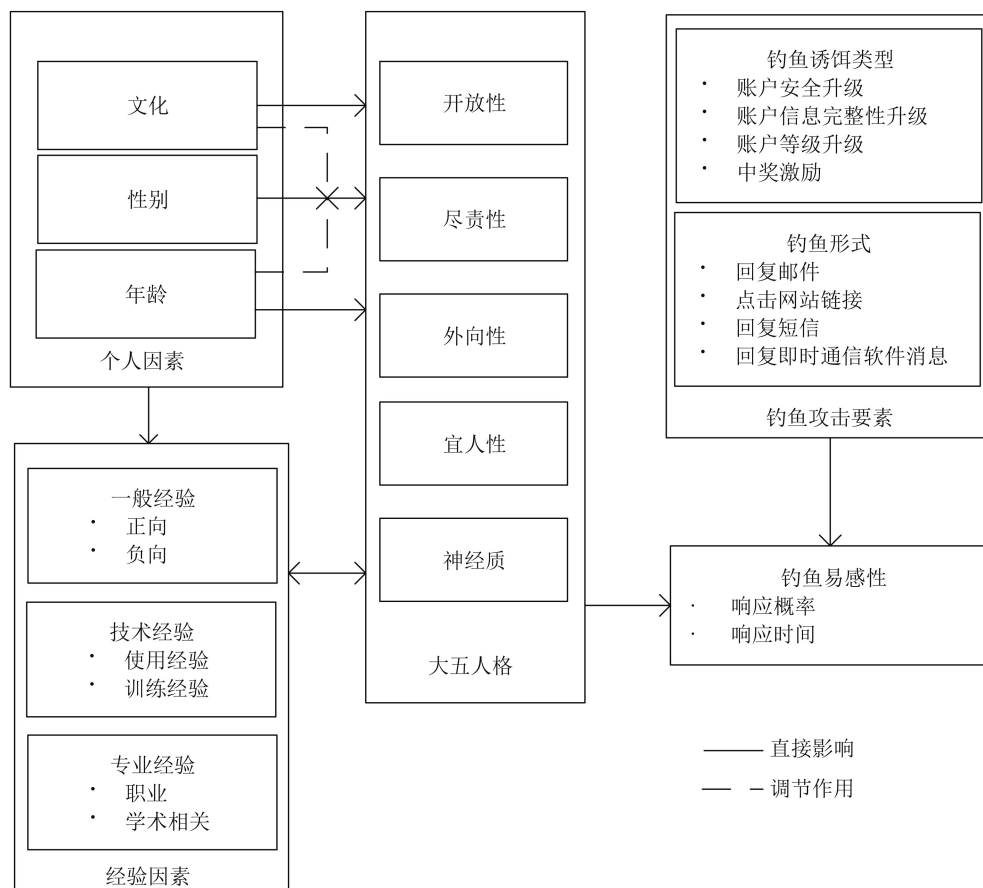


图4 以大五人格理论为核心的钓鱼易感性研究框架^[69]

consistency)、互惠(reciprocity)、喜好(liking)、社会认同(social proof)以及短缺(scarcity)6个原则对受害者进行说服。并且总结了大五人格与影响力原则之间的联系,形成了一个社会工程学-人格框架,如图5所示。其中实线表示某一维度的高分人群更容易受到箭头指向的几类影响力原则的利用,相反,虚线表示这一维度的高分人群不易受到对应的影响力原则的利用。

从一些实验中可以看出,人格因素可能会导致人在受到社会工程学攻击时,做出未经思考、缺乏逻辑的决策。一些恶意攻击者可以发现目标的人格,判断其是否容易被成功攻击,甚至对于特定人格的目标,攻击者可以设计针对性内容的钓鱼邮件,从而提高攻击成功率。除了以实验为基础的研究之外,部分研究者也提出了与人格有关的社会工程学研究框架,希望找出人格影响决策脆弱性更本质的理论基础。

5 人格与安全防护

从第4节的总结中可以看出,人格会在一定程度上影响人的安全意识、安全行为、以及面对攻击时做出正确决策的能力^[71-74]。一旦人格所表现出的安全脆弱性被攻击者所利用,就会造成更大的安全威胁。在讨论人格与安全脆弱性的文献中,研究者对于安全防护的建议一般包含两方面:其一是针对不同人格制定不同的安全训练计划,其二是基于人格的差异设计技术防护措施。

5.1 安全意识训练

由于不同人格的人会在安全意识上存在差异,

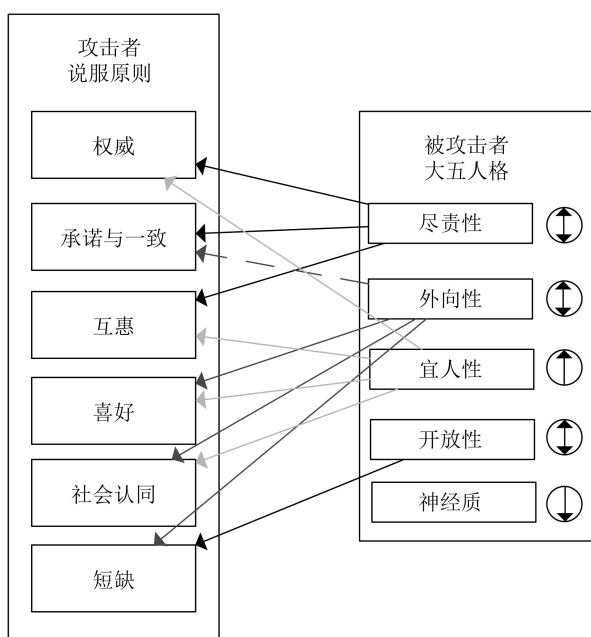


图5 社会工程学-人格框架^[70]

因此安全意识的训练也要充分考虑到人与人之间的差异性。Warkentin和Uebelacker等人^[59]的研究中就提到在不同的工作场景下,可以根据其人格制定不同的安全意识培训计划。Downs等人^[75]认为在对人们进行反钓鱼培训时,最重要的是要让人们了解自身在收到钓鱼邮件时的行为反应,如果人们能够知道是什么因素导致了自己被钓鱼邮件所欺骗,那么针对不同人的安全培训计划才能够进一步的实施。但是,这些研究都没有具体说明如何去制定安全培训计划,这一部分有待进一步的探究。

5.2 技术防护

除了提高人自身的安全意识之外,技术手段的防护仍然是必要的。在对技术升级时,也应当充分考虑人的因素,最大化技术的效用。例如Halevi等人^[76]提出在社交网络中,根据用户的人格,推荐不同的隐私设置,既保护用户隐私,也降低对用户体验的负面影响。在应对鱼叉式钓鱼邮件的攻击时,也可以应用技术手段来弥补人心心理弱点造成的安全漏洞。在4.2节中曾经提到,钓鱼邮件的内容可能会利用人格表现出的安全脆弱性,例如对高尽责性人群发送一个刻意描述紧急情况的邮件,诱使他们受骗。因此,研究者开发了一个可以评估用户人格的钓鱼邮件检测工具,依据用户的人格,找出邮件中容易触发用户心理弱点的词语。一旦发现这类词语,就会对用户提出告警。在当今的网络空间安全领域,防护技术除了关注恶意攻击的特性之外^[77],也应该更多地考虑人的因素,考虑人与人之间存在的差异对防护有效性的影响,从而提高技术防护的成功率。

6 结束语

本文对人格在网络空间安全领域的相关研究做出了总结,首先介绍了网络空间安全 and 人格的基本概念,提出了网络空间安全中人格研究的框架,随后对人格测量、人格与安全脆弱性、人格与安全防护3个方面技术研究分别做出总结。

基于大数据的人格测量方法,不仅会推动社交网络、电商平台等服务提供方更加有效地利用社会数据,提升服务质量,也能够推动心理学、计算机科学等多个学科之间的合作研究。从目前的总结来看,利用大数据推测人格的准确率在50%~70%之间,并且只是针对特定数据集的效果,还没有大规模地利用社会数据去验证,之后的研究可以继续对于人格测量的准确率加以提升。一方面,可以从改进技术入手,例如改进机器学习算法,建立准确率更高、更加稳健的模型。另一方面,应当进一步结合社会科学与计算机科学,根据与人格有关的心理

认知机制和行为机制, 提取数据中的行为特征, 或者提出更加有效的人格测量方法。

另外需要指出, 相较于传统的人格测量方法, 大数据测量人格的方法可以在网络用户不知情的情况下获取其人格信息。而正如本文所提到的, 人格可以作为一个安全脆弱性的因素被攻击者所利用, 因此人格应被视为一个重要的隐私属性加以保护。如何平衡社会数据的可用性和隐私性, 是之后研究的一个重要课题。

本文通过文献调研, 对人格与安全脆弱性的研究进行了总结。通过这些研究, 人们可以更加深入地了解网络空间安全中人格脆弱性的主要表现, 对于内部脆弱性而言, 可以了解何种人格导致了安全意识的缺失, 或是安全意识与行为的不一致, 对于安全培训有着指导意义; 对于决策脆弱性而言, 可以了解人格如何影响面对社会工程学攻击时的决策, 从而也可以针对性地提高防范意识, 或是提出技术防护手段。安全防护的研究是以安全脆弱性的研究为基础, 而目前人格与安全脆弱性的研究主要还只是以一些小规模的实证研究为基础, 还没有一个成熟的研究体系或是理论框架, 这也是未来研究的一个重要方向。

网络空间安全中的人格研究目前在测量方法、安全脆弱性挖掘和防护方法上取得了一定的成果, 但仍然存在多种问题, 有待进一步发展。例如, 网络空间安全中人格研究的模型应用问题, 当前研究一般是采用传统心理学测量方法建立好的有标注的数据集来进行训练和测试, 这种模型的泛化能力一般较差, 没办法适用于大规模的网络空间应用场景; 此外, 一旦开展网络空间安全中人格研究在真实环境下的应用, 模型的效果评估也存在很大的挑战。所以, 如何解决真实环境下人格的量化问题, 形成科学的人格分析模型, 建立完善的评价指标体系, 实现面向人格的网络空间安全防护, 是当前人格研究中需要解决的关键问题。

参 考 文 献

- [1] GIBSON W. *Burning Chrome*[M]. New York: EOS/HarperCollins Publishers, 2003: 1–5.
- [2] 张焕国, 韩文报, 来学嘉, 等. 网络空间安全综述[J]. 中国科学: 信息科学, 2015, 58(1): 1–43.
ZHANG Huanguo, HAN Wenbao, LAI Xuejia, et al. Survey on cyberspace security[J]. *Science China Information Sciences*, 2015, 58(1): 1–43.
- [3] OTTIS R and LORENTS P. Cyberspace: Definition and implications[C]. The 5th International Conference on Information Warfare and Security, Dayton, 2010: 267.
- [4] MITNICK K D and SIMON W L. *The Art of Deception: Controlling the Human Element of Security*[M]. Indianapolis: John Wiley & Sons, 2002: 1–10.
- [5] STEWART JR J H. Social engineering deception susceptibility: Modification of personality traits susceptible to social engineering manipulation to acquire information through attack and exploitation[D]. [Ph. D. dissertation], Colorado Technical University, 2015.
- [6] BULLER D B, BURGOON J K, BUSLIG A, et al. Testing interpersonal deception theory: The Language of interpersonal deception[J]. *Communication Theory*, 1996, 6(3): 268–289. doi: [10.1111/j.1468-2885.1996.tb00129.x](https://doi.org/10.1111/j.1468-2885.1996.tb00129.x).
- [7] DEPAULO B M, LINDSAY J J, MALONE B E, et al. Cues to deception[J]. *Psychological Bulletin*, 2003, 129(1): 74–118. doi: [10.1037/0033-2909.129.1.74](https://doi.org/10.1037/0033-2909.129.1.74).
- [8] IRANI D, BALDUZZI M, BALZAROTTI D, et al. Reverse Social engineering attacks in online social networks[C]. The 8th International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment, Amsterdam, The Netherlands, 2011: 55–74.
- [9] BURGER J M. *Personality*[M]. California: Thomson Learning, 2007: 5–7.
- [10] 彭聃龄. 普通心理学[M]. 北京: 北京师范大学出版社, 2012: 203–210.
PENG Danling. *Common Psychology*[M]. Beijing: Beijing Normal University Press, 2012: 203–210.
- [11] ALMLUND M, DUCKWORTH A L, HECKMAN J, et al. Personality psychology and economics[J]. *Handbook of the Economics of Education*, 2011, 4: 1–181.
- [12] HOGAN R, JOHNSON J, and BRIGGS S R. *Handbook of Personality Psychology*[M]. San Diego: Academic Press, 1997: 750–751.
- [13] GERRIG R J and ZIMBARDO P G. *Psychology and Life*[M]. New York: Pearson, 2012.
- [14] GOLDBERG L R, JOHNSON J A, EBER H W, et al. The international personality item pool and the future of public-domain personality measures[J]. *Journal of Research in Personality*, 2006, 40(1): 84–96. doi: [10.1016/j.jrp.2005.08.007](https://doi.org/10.1016/j.jrp.2005.08.007).
- [15] 郑敬华, 郭世泽, 高梁, 等. 基于多任务学习的大五人格预测[J]. 中国科学院大学学报, 2018, 35(4): 550–560. doi: [10.7523/j.issn.2095-6134.2018.04.019](https://doi.org/10.7523/j.issn.2095-6134.2018.04.019).
ZHENG Jinghua, GUO Shize, GAO Liang, et al. Microblog users' Big-Five personality prediction based on multi-task learning[J]. *Journal of University of Chinese Academy of Sciences*, 2018, 35(4): 550–560. doi: [10.7523/j.issn.2095-6134.2018.04.019](https://doi.org/10.7523/j.issn.2095-6134.2018.04.019).
- [16] KOSINSKI M, BACHRACH Y, KOHLI P, et al. Manifestations of user personality in website choice and behaviour on online social networks[J]. *Machine Learning*, 2014, 95(3): 357–380. doi: [10.1007/s10994-013-5415-y](https://doi.org/10.1007/s10994-013-5415-y).
- [17] HU Jian, ZENG HuaJun, LI Hua, et al. Demographic

- prediction based on user's browsing behavior[C]. The 16th International Conference on World Wide Web, Banff, Canada, 2007: 151–160. doi: [10.1145/1242572.1242594](https://doi.org/10.1145/1242572.1242594).
- [18] RENTFROW P J and GOSLING S D. The do re mi's of everyday life: The structure and personality correlates of music preferences[J]. *Journal of Personality and Social Psychology*, 2003, 84(6): 1236–1256. doi: [10.1037/0022-3514.84.6.1236](https://doi.org/10.1037/0022-3514.84.6.1236).
- [19] 郑日昌, 蔡永红, 周益群. 心理测量学[M]. 北京: 人民教育出版社, 1999: 155–170.
ZHENG Richang, CAI Yonghong, and ZHOU Yiqun. Psychological Testing[M]. Beijing: People Education Press, 1999: 155–170.
- [20] GREGORY R J. Psychological Testing: History, Principles, and Applications[M]. Beijing: Peking University Press, 2013: 25.
- [21] KUHN M H and MCPARTLAND T S. An empirical investigation of self-attitudes[J]. *American Sociological Review*, 1954, 19(1): 68–76. doi: [10.2307/2088175](https://doi.org/10.2307/2088175).
- [22] 朱廷劭, 李昂, 宁悦, 等. 网络社会中个体人格特征及其行为关系[J]. 兰州大学学报: 社会科学版, 2011, 39(5): 44–51.
ZHU Tingshao, LI Ang, NING Yue, *et al.* Individual personality and its behavior in the cyber society[J]. *Journal of Lanzhou University: Social Sciences*, 2011, 39(5): 44–51.
- [23] WEINER I B. The Corsini Encyclopedia of Psychology[M]. 4th ed. Hoboken: John Wiley & Sons, 2010.
- [24] MORGAN C D and MURRAY H A. A method for investigating fantasies: The thematic apperception test[J]. *Archives of Neurology & Psychiatry*, 1935, 34(2): 289–306. doi: [10.1001/archneurpsyc.1935.02250200049005](https://doi.org/10.1001/archneurpsyc.1935.02250200049005).
- [25] MURRAY H A and DAN M A. Explorations in personality[J]. *Journal of Projective Techniques & Personality Assessment*, 1938(2): 283–285.
- [26] MURRAY H A. Explorations in personality: A clinical and experimental study of fifty men of college age[J]. *American Journal of Sociology*, 1938, 4(4): 576–583.
- [27] MCCRAE R R and WEISS A. Observer Ratings of Personality[M]. ROBINS R W, FRALEY R C, and KRUEGER R F. Handbook of Research Methods in Personality Psychology. New York: The Guilford Press, 2007: 259–272.
- [28] VINCIARELLI A and MOHAMMADI G. A survey of personality computing[J]. *IEEE Transactions on Affective Computing*, 2014, 5(3): 273–291. doi: [10.1109/TAFFC.2014.2330816](https://doi.org/10.1109/TAFFC.2014.2330816).
- [29] ARGAMON S, DHAWLE S, KOPPEL M, *et al.* Lexical predictors of personality type[C]. 2005 Joint Annual Meeting of the Interface and the Classification Society of North America, 2005.
- [30] MAIRESSE F, WALKER M A, MEHL M R, *et al.* Using linguistic cues for the automatic recognition of personality in conversation and text[J]. *Journal of Artificial Intelligence Research*, 2007, 30(1): 457–500. doi: [10.1613/jair.2349](https://doi.org/10.1613/jair.2349).
- [31] TAUSCZIK Y R and PENNEBAKER J W. The psychological meaning of words: LIWC and computerized text analysis methods[J]. *Journal of Language and Social Psychology*, 2010, 29(1): 24–54. doi: [10.1177/0261927x09351676](https://doi.org/10.1177/0261927x09351676).
- [32] OBERLANDER J and NOWSON S. Whose thumb is it anyway?: Classifying author personality from weblog text[C]. The COLING/ACL on Main Conference Poster Sessions, Sydney, Australia, 2006: 627–634.
- [33] NOWSON S and OBERLANDER J. Identifying more bloggers: Towards large scale personality classification of personal weblogs[C]. The International Conference on Weblogs and Social Media, Boulder, USA, 2007: 115–127.
- [34] MINAMIKAWA A and YOKOYAMA H. Blog tells what kind of personality you have: Egogram estimation from Japanese weblog[C]. 2011 ACM Conference on Computer Supported Cooperative Work, Hangzhou, China, 2011: 217–220. doi: [10.1145/1958824.1958856](https://doi.org/10.1145/1958824.1958856).
- [35] MINAMIKAWA A and YOKOYAMA H. Personality estimation based on Weblog text classification[C]. The 24th International Conference on Industrial Engineering and other Applications of Applied Intelligent Systems Conference on Modern Approaches in Applied Intelligence, Syracuse, USA, 2011: 89–97.
- [36] GILL A J, NOWSON S, and OBERLANDER J. What are they blogging about? Personality, topic and motivation in Blogs[C]. The 3rd International AAAI Conference on Weblogs and Social Media, San Jose, USA, 2009: 18–25.
- [37] HAMBURGER Y A and BEN-ARTZI E. The relationship between extraversion and neuroticism and the different uses of the Internet[J]. *Computers in Human Behavior*, 2000, 16(4): 441–449. doi: [10.1016/S0747-5632\(00\)00017-0](https://doi.org/10.1016/S0747-5632(00)00017-0).
- [38] PARK G, SCHWARTZ H A, EICHSTAEDT J C, *et al.* Automatic personality assessment through social media language[J]. *Journal of Personality and Social Psychology*, 2015, 108(6): 934–952. doi: [10.1037/pspp0000020](https://doi.org/10.1037/pspp0000020).
- [39] WEI Xiangyu, XU Guangquan, WANG Hao, *et al.* Sensing users' emotional intelligence in social networks[J]. *IEEE Transactions on Computational Social Systems*, 2020, 7(1): 103–112. doi: [10.1109/TCSS.2019.2944687](https://doi.org/10.1109/TCSS.2019.2944687).
- [40] LUYCKX K and DAELEMANS W. Using syntactic features to predict author personality from text[C]. Digital Humanities 2008, Oulu, Finland, 2008: 146–149.
- [41] MAO Yu, ZHANG Dongmei, WU Chunhua, *et al.* Feature analysis and optimisation for computational personality recognition[C]. The IEEE 4th International Conference on Computer and Communications (ICCC), Chengdu, China,

- 2018: 2410–2414, doi: [10.1109/CompComm.2018.8780801](https://doi.org/10.1109/CompComm.2018.8780801).
- [42] WOLFRADT U and DOLL J. Motives of adolescents to use the Internet as a function of personality traits, personal and social factors[J]. *Journal of Educational Computing Research*, 2001, 24(1): 13–27. doi: [10.2190/ANPM-LN97-AUT2-D2EJ](https://doi.org/10.2190/ANPM-LN97-AUT2-D2EJ).
- [43] ROSS C, ORR E S, SISIC M, *et al.* Personality and motivations associated with Facebook use[J]. *Computers in Human Behavior*, 2009, 25(2): 578–586. doi: [10.1016/j.chb.2008.12.024](https://doi.org/10.1016/j.chb.2008.12.024).
- [44] AMICHAH-HAMBURGER Y and VINITZKY G. Social network use and personality[J]. *Computers in Human Behavior*, 2010, 26(6): 1289–1295. doi: [10.1016/j.chb.2010.03.018](https://doi.org/10.1016/j.chb.2010.03.018).
- [45] LI Lin, LI Ang, HAO Bibo, *et al.* Predicting active users' personality based on micro-Blogging behaviors[J]. *PLoS One*, 2014, 9(1): e84997. doi: [10.1371/journal.pone.0084997](https://doi.org/10.1371/journal.pone.0084997).
- [46] SWICKERT R J, ROSENTERTER C J, HITTNER J B, *et al.* Extraversion, social support processes, and stress[J]. *Personality and Individual Differences*, 2002, 32(5): 877–891. doi: [10.1016/s0191-8869\(01\)00093-9](https://doi.org/10.1016/s0191-8869(01)00093-9).
- [47] FRIGGERI A, LAMBIOTTE R, KOSINSKI M, *et al.* Psychological aspects of social communities[C]. 2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing, Amsterdam, Netherlands, 2012: 195–202. doi: [10.1109/SocialCom-PASSAT.2012.104](https://doi.org/10.1109/SocialCom-PASSAT.2012.104).
- [48] LAMBIOTTE R and KOSINSKI M. Tracking the digital footprints of personality[J]. *Proceedings of the IEEE*, 2014, 102(12): 1934–1939. doi: [10.1109/JPROC.2014.2359054](https://doi.org/10.1109/JPROC.2014.2359054).
- [49] University of Cambridge. MyPersonality project[EB/OL]. <http://mypersonality.org/wiki/doku.php>, 2010.
- [50] RIFE S C, CATE K L, KOSINSKI M, *et al.* Participant recruitment and data collection through Facebook: The role of personality factors[J]. *International Journal of Social Research Methodology*, 2016, 19(1): 69–83. doi: [10.1080/13645579.2014.957069](https://doi.org/10.1080/13645579.2014.957069).
- [51] KOSINSKI M, STILLWELL D, and GRAEPEL T. Private traits and attributes are predictable from digital records of human behavior[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2013, 110(15): 5802–5805. doi: [10.1073/pnas.1218772110](https://doi.org/10.1073/pnas.1218772110).
- [52] OLGUÍN D O, GLOOR P A, and PENTLAND A. Capturing individual and group behavior with wearable sensors[C]. 2009 AAAI Spring Symposium on Human Behavior Modeling, Stanford, USA, 2009: 68–74.
- [53] CHITTARANJAN G, BLOM J, and GATICA-PEREZ D. Who's who with big-five: Analyzing and classifying personality traits with smartphones[C]. The 15th Annual International Symposium on Wearable Computers, San Francisco, USA, 2011: 29–36. doi: [10.1109/ISWC.2011.29](https://doi.org/10.1109/ISWC.2011.29).
- [54] CHITTARANJAN G, BLOOM J, and GATICA-PEREZ D. Mining large-scale smartphone data for personality studies[J]. *Personal and Ubiquitous Computing*, 2013, 17(3): 433–450. doi: [10.1007/s00779-011-0490-1](https://doi.org/10.1007/s00779-011-0490-1).
- [55] STAIANO J, LEPRI B, AHARONY N, *et al.* Friends don't lie: Inferring personality traits from social network structure[C]. 2012 ACM Conference on Ubiquitous Computing, Pennsylvania, USA, 2012: 321–330. doi: [10.1145/2370216.2370266](https://doi.org/10.1145/2370216.2370266).
- [56] BIEL J I, TEIJEIRO-MOSQUERA L, and GATICA-PEREZ D. FaceTube: Predicting personality from facial expressions of emotion in online conversational video[C]. The 14th ACM international conference on Multimodal Interaction, Santa Monica, USA, 2012: 53–56. doi: [10.1145/2388676.2388689](https://doi.org/10.1145/2388676.2388689).
- [57] SKOWRON M, TKALČIČ M, FERWERDA B, *et al.* Fusing social media cues: Personality prediction from twitter and instagram[C]. The 25th International Conference Companion on World Wide Web, Montréal Québec, Canada, 2016: 107–108. doi: [10.1145/2872518.2889368](https://doi.org/10.1145/2872518.2889368).
- [58] FERWERDA B, SCHEDL M, and TKALČIČ M. Using instagram picture features to predict users' personality[C]. The 22nd International Conference on Multimedia Modeling, Miami, USA, 2016: 850–861. doi: [10.1007/978-3-319-27671-7_71](https://doi.org/10.1007/978-3-319-27671-7_71).
- [59] WARKENTIN M, MCBRIDE M, CARTER L, *et al.* The role of individual characteristics on insider abuse intentions[C]. The 18th Americas Conference on Information Systems, Seattle, USA, 2012: 4833–4842.
- [60] KORZAAN M L and BOSWELL K T. The influence of personality traits and information privacy concerns on behavioral intentions[J]. *Journal of Computer Information Systems*, 2008, 48(4): 15–24.
- [61] UFFEN J and BREITNER M H. Management of technical security measures: An empirical examination of personality traits and behavioral intentions[C]. The 46th Hawaii International Conference on System Sciences, Wailea, USA, 2013: 4551–4560. doi: [10.1109/HICSS.2013.388](https://doi.org/10.1109/HICSS.2013.388).
- [62] BRECHT F, FABIAN B, KUNZ S, *et al.* Communication anonymizers: Personality, internet privacy literacy and their influence on technology acceptance[C]. The 20th European Conference on Information Systems, Barcelona, Spain, 2012: 214.
- [63] SHROPSHIRE J, WARKENTIN M, and SHARMA S. Personality, attitudes, and intentions: Predicting initial adoption of information security behavior[J]. *Computers & Security*, 2015, 49: 177–191. doi: [10.1016/j.cose.2015.01.002](https://doi.org/10.1016/j.cose.2015.01.002).
- [64] MODIC D and LEA S E G. How neurotic are scam victims,

- really? The Big Five and internet scams[EB/OL]. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2448130, 2014.
- [65] PATTINSON M, JERRAM C, PARSONS K, *et al*. Why do some people manage phishing e-mails better than others?[J]. *Information Management & Computer Security*, 2012, 20(1): 18–28. doi: [10.1108/09685221211219173](https://doi.org/10.1108/09685221211219173).
- [66] DARWISH A, EL ZARKA A, and ALOUL F. Towards understanding phishing victims' profile[C]. 2012 International Conference on Computer Systems and Industrial Informatics, Sharjah, United Arab Emirates, 2012: 1–5. doi: [10.1109/ICCSII.2012.6454454](https://doi.org/10.1109/ICCSII.2012.6454454).
- [67] ENOS F, BENUS S, CAUTIN R L, *et al*. Personality factors in human deception detection: Comparing human to machine performance[C]. The 9th International Conference on Spoken Language Processing, Pittsburgh, USA, 2006: 813–816.
- [68] HALEVI T, LEWIS J, and MEMON N. Phishing, personality traits and facebook[J]. 2013, arXiv: 1301.7643.
- [69] PARRISH JR J L, BAILEY J L, and COURTNEY J F. A personality based model for determining susceptibility to phishing attacks[R]. University of Arkansas at Little Rock Technique Report, Arkansas, USA, 2009: 285–296.
- [70] CHO J H, CAM H, and OLTRAMARI A. Effect of personality traits on trust and risk to phishing vulnerability: Modeling and analysis[C]. 2016 IEEE International Multi-disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), San Diego, USA, 2016: 7–13. doi: [10.1109/COGSIMA.2016.7497779](https://doi.org/10.1109/COGSIMA.2016.7497779).
- [71] JENSEN-CAMPBELL L A, KNACK J M, and GOMEZ H L. The psychology of nice people[J]. *Social and Personality Psychology Compass*, 2010, 4(11): 1042–1056. doi: [10.1111/j.1751-9004.2010.00307.x](https://doi.org/10.1111/j.1751-9004.2010.00307.x).
- [72] JOHN O P, NAUMANN L P, and Soto C J. Paradigm Shift to the Integrative Big Five Trait Taxonomy: History, Measurement and Conceptual Issues[M]. JOHN O P, ROBINS R W, and PERVIN L A. Handbook of Personality: Theory and Research. 3rd ed. New York, NY: Guilford, 2008.
- [73] CROSSLER R E, JOHNSTON A C, LOWRY P B, *et al*. Future directions for behavioral information security research[J]. *Computers & Security*, 2013, 32: 90–101. doi: [10.1016/j.cose.2012.09.010](https://doi.org/10.1016/j.cose.2012.09.010).
- [74] PFLEEGER S L, PREDD J B, HUNKER J, *et al*. Insiders behaving badly: Addressing bad actors and their actions[J]. *IEEE Transactions on Information Forensics and Security*, 2010, 5(1): 169–179. doi: [10.1109/TIFS.2009.2039591](https://doi.org/10.1109/TIFS.2009.2039591).
- [75] DOWNS J S, HOLBROOK M, and CRANOR L F. Behavioral response to phishing risk[C]. The Anti-Phishing Working Groups 2nd annual eCrime Researchers Summit, Pittsburgh, USA, 2007: 37–44. doi: [10.1145/1299015.1299019](https://doi.org/10.1145/1299015.1299019).
- [76] HALEVI T, LEWIS J, and MEMON N. A pilot study of cyber security and privacy related behavior and personality traits[C]. The 22nd International Conference on World Wide Web, Rio de Janeiro, Brazil, 2013: 737–744. doi: [10.1145/2487788.2488034](https://doi.org/10.1145/2487788.2488034).
- [77] 冯登国, 张敏, 叶宇桐. 基于差分隐私模型的位置轨迹发布技术研究[J]. 电子与信息学报, 2020, 42(1): 74–88. doi: [10.11999/JEIT190632](https://doi.org/10.11999/JEIT190632).
- FENG Dengguo, ZHANG Min, and YE Yutong. Research on differentially private trajectory data publishing[J]. *Journal of Electronics & Information Technology*, 2020, 42(1): 74–88. doi: [10.11999/JEIT190632](https://doi.org/10.11999/JEIT190632).
- 吴桐: 女, 1992年生, 博士生, 研究方向为网络安全、社会工程学。
郑康锋: 男, 1975年生, 教授, 博士生导师, 研究方向为网络安全。
伍淳华: 女, 1979年生, 讲师, 研究方向为人工智能、网络安全。
王秀娟: 女, 1979年生, 讲师, 研究方向为人工智能、网络安全。
郑赫慈: 男, 1995年生, 硕士生, 研究方向为网络安全。

责任编辑: 余蓉