前　言

依托于云计算、物联网、大数据技术的发展，自动驾驶、人脸识别、智能家居等人工智能技术快速进入了人们的视野，并成为先进科技社会化应用的代表和社会热点.但是，安全问题却为这些技术的广泛应用提出了严峻挑战，没有强大的自主可控安全技术的支撑，人工智能带来的也许不仅仅是便利，更可能是灾难.安全问题可以说是人工智能走向大规模应用的瓶颈和一个关键问题.而作为解决安全问题的核心技术——密码学，如何适应人工智能安全的需要是另一个关键问题.

为推动我国学者在智能安全领域的研究，为人工智能的现实应用提供理论与技术支撑，及时报道我国学者在智能安全理论与技术方面的最新研究成果，《计算机研究与发展》和我们共同策划和组织了“密码学与智能安全研究”专题.本期专题通过公开征文共收到 99 篇普通投稿，4 篇特邀投稿，分别在多个方面阐述了智能安全研究领域具有重要意义的研究成果.本专题严格按照该刊审稿要求进行，特约编委先后邀请了近百位相关领域的专家参与评审，每篇论文邀请至少 3～4 位专家进行评审，历经初审、复审、终审等阶段，整个流程历经一个半月，最终本专题共精选录用文章 26 篇(含 4 篇特邀稿件).这 26 篇文章分别涵盖了智能密码算法、智能隐私保护、智能系统安全等研究内容，在一定程度上反映了当前国内各单位在智能安全研究领域的主要研究方向.由于刊物单期容量所限，本专题分别刊登在 2019 年第 10 期和第 11 期，智能系统安全相关的 7 篇文章将在第 11 期刊登.

1　综　　述

人工智能安全是一个新领域，为了便于更多的读者学习和了解，推动密码学与智能安全的发展，此部分共收录了 6 篇论文，主要内容包括推荐系统的隐私保护、机器学习系统的安全问题、机器学习模型可解释性、安全漏洞自动利用、量子人工智能密码、人工智能系统安全与隐私风险等方面.

智能推荐系统是建立在海量数据挖掘基础之上的一种智能平台，根据用户个人信息与物品特征，利用统计分析和机器学习等人工智能技术建立模型，预测用户对新物品的评价与喜好，从而向用户推荐其可能感兴趣的潜在物品，以实现个性化的信息服务和决策支持.然而，推荐系统的历史数据集、预测模型和推荐结果都与用户的隐私休戚相关，如何能在有效保护用户隐私的前提下，提供正确性可验证的有效推荐结果是一个具有挑战性的问题.“推荐系统的隐私保护研究进展”一文从推荐系统隐私保护的模式、安全模型、轻量级的推荐系统隐私保护一般性构造与推荐结果正确性可验证、可审计等方面，系统阐述了国内外最新研究成果，并在此基础上提出了不依赖公钥全同态加密技术，通过减少公钥加密/解密次数(最优时一次)，在单用户、多数据模型和多用户、多数据模型下，实现推荐系统隐私保护一般性构造，为适用于推荐系统隐私保护的新型加密方案研究及其实用化提供了新思路.

人工智能已经渗透到生活的各个角落，给人类带来了极大的便利.尤其是近年来，随着机器学习中深度学习这一分支的蓬勃发展，生活中的相关应用越来越多.不幸的是，机器学习系统也面临着许多安全隐患，而机器学习系统的普及进一步放大了这些风险.为了揭示这些安全隐患并实现一个强大的机器学习系统，“机器学习系统的隐私和安全问题综述”一文对主流的深度学习系统进行了调查.该文主要侧重在机器学习中的深度学习领域，设计了一个剖析深度学习系统的分析模型，并界定了调查范围.特别地，调查的深度学习系统跨越了 4 个领域——图像分类、音频语音识别、恶意软件检测和自然语言处理.该文提取了 4 种类型的安全隐患，并从复杂性、攻击成功率和破坏等

多个维度对其进行了表征和度量.随后调研了针对深度学习系统的防御技术及其特点.最后通过对这些系统的观察,该文提出了构建健壮的深度学习系统的建议.

尽管机器学习在许多领域取得了巨大的成功,但缺乏可解释性严重限制了其在现实任务尤其是安全敏感任务中的广泛应用.为了克服这一致命弱点,许多学者对如何提高机器学习模型可解释性进行了深入的研究,并提出了大量的解释方法以帮助终端用户理解模型内部的工作机制.然而,可解释性研究还处于初级阶段,依然还有大量的科学问题尚待解决.并且,不同的学者解决问题的角度不同,对可解释性赋予的含义也不同,所提出的解释方法也各有侧重.迄今为止,学术界对模型可解释性仍缺乏统一的认识,对可解释性研究缺乏科学的指导."机器学习模型可解释性方法、应用与安全性研究综述"一文回顾了机器学习中的可解释性问题,并对现有的研究工作进行了系统的总结和科学的归类.同时讨论了可解释性相关技术的潜在应用,分析了可解释性与可解释机器学习的安全性之间的关系,并且探讨了可解释性研究当前面临的挑战和未来潜在的研究方向,以期进一步推动可解释性研究的发展和应用.

随着安全漏洞数量急剧上升,高效率地评估与修复漏洞面临更大的挑战.目前漏洞的可利用性评估主要依赖人工方法,如何智能化和自动化地进行安全漏洞利用是本领域一个热点研究问题."安全漏洞自动利用综述"一文调研了2006年至今安全漏洞自动利用文献,分析了现状并指出了漏洞利用研究的发展趋势,同时给出了漏洞自动利用的一般框架;分别从漏洞自动利用的信息输入、漏洞类型和利用方法这3个角度对当前研究成果进行了梳理,指出了这3个角度对漏洞自动利用的影响;分析了漏洞自动利用研究的不足与挑战,并对将来的研究趋势进行了展望.

采用人工智能设计出高强度密码和使密码设计自动化是人们长期追求的目标,其中演化算法以其相对数学规划方法更大的优越性成为人工智能领域的研究热点之一."从演化密码到量子人工智能密码综述"一文论述了演化密码的发展:中国学者将密码学与演化计算结合,借鉴生物进化的思想提出演化密码的概念和用演化计算设计密码的方法,得到可变渐强的密码,减少攻击所需搜索空间的量级.研究表明,演化密码已经在对称密码和非对称密码领域均取得了一些成果,已具备人工智能密码的一些特征.该文还介绍了量子计算机设计密码的理论成果和D-Wave 2000Q真实量子计算机密码设计.

人类正在经历着由机器学习技术推动的人工智能浪潮,在某些特定领域中,人工智能已经表现出达到甚至超越人类的工作能力.然而,以往的机器学习理论大多没有考虑开放甚至对抗的系统运行环境,人工智能系统的安全和隐私问题正逐渐暴露出来."人工智能系统安全与隐私风险"一文通过回顾人工智能系统安全方面的相关研究工作,揭示了人工智能系统中潜藏的安全与隐私风险.该文介绍了包含攻击面、攻击能力和攻击目标的安全威胁模型,在此基础上,从人工智能系统的4个关键环节——数据输入(传感器)、数据预处理、机器学习模型和输出,分析了相应的安全隐私风险及相应对策.最后讨论了人工智能系统安全研究方面未来的发展趋势.

2 智能密码算法

智能密码算法部分共收录了7篇论文,主要围绕在人工智能安全方面可以发挥重要作用的属性基密码、轻量级密码和可搜索密码等研究方向展开.

基于密文策略的属性基加密被认为是实现云上数据细粒度访问控制最有效的方法之一."隐藏访问策略的高效CP-ABE方案"一文提出了一种隐藏访问策略的高效CP-ABE方案,它可以使得属性隐藏和秘密共享能够同时应用到"与"门结构中,然后利用合数阶双线性群构造一种基于包含正负及无关值的"与门"的策略隐藏方案.该方案有效地避免了用户的具体属性值泄露给其他第三方,且具有解密时间短、解密效率高的优点.

基于属性加密可以有效保护云服务器中数据的隐私性,但是属性加密中密钥分配、数据加密和

解密过程的计算开销过大,给资源受限的用户造成很大的计算负担.“支持属性撤销的可追踪外包属性加密方案”一文构造了一个将密钥分配与解密工作外包给云服务器的支持属性撤销的属性加密方案,同时该方案可验证外包计算的正确性.该方案使用线上/线下加密,既有效保护用户数据的隐私性,又减少用户的计算开销,提升方案运行效率;其次该方案中使用树形访问策略,以提供更加细粒度的访问控制;同时利用重加密的方法实现细粒度的属性撤销,通过生成重加密密钥更新属性与密文,间接撤销单个属性;最后将用户身份嵌入密钥,达到用户可追踪的性质.

由于云存储密文的静态性特征,密钥泄露成为影响存储数据安全性的重要因素.数据重加密是应对密钥泄露的有效手段,但相应的计算开销以及上传下载的通信开销增加了用户和存储系统的负担.此外,对基于分布式编码的数据存储而言,密文更新需要在还原密文的基础上进行,密文合并过程同样增加了系统的通信及计算开销.“云环境下支持可更新加密的分布式数据编码存储方案”一文提出了一种云环境下支持可更新加密的分布式数据编码存储方案(DDES-UE).避免数据重加密及数据上传、下载、解码、合并带来的计算和通信开销,对于构建支持直接数据更新的安全高效云存储系统有重要意义.周期性密钥更新可有效增加攻击者通过获取密钥破解密文的难度,从而增强了系统的主动安全防御能力.

基于分支条件混淆的代码加密技术,实现秘钥和程序的分离,能够对抗程序静态和动态分析手段,但仅能用于相等条件分支.“一种基于分支条件混淆的代码加密技术”一文通过引入拉格朗日插值法,生成输入处理函数,在保证分支条件混淆安全的前提下,解决了多输入分支条件下通过输入产生秘钥的问题,实现多输入分支下的条件代码加密;把多输入分支下生成唯一秘钥方法应用到等于条件取或分支、大小比较条件分支和复杂条件分支,实现了基于分支条件混淆的代码加密技术从相等条件分支到区间条件分支和复杂条件分支的扩展.

云存储中为保护数据所有者的数据安全性和隐私性,采用数据加密后再提供按需数据服务的方式,可搜索加密技术是解决加密数据接入的关键方法.但搜索时的多关键词不加区别和忽视索引之间的关联性会造成搜索时间长和准确率低等问题.“基于语义扩展的多关键词可搜索加密算法”一文提出了一种基于语义扩展的多关键词可搜索加密算法.首先,区分多关键词的重要性进行语义扩展,并基于依存句法生成多关键词陷门.其次,基于凝聚层次聚类和关键词平衡二叉树,构建索引关联性的索引树结构.最后,引入剪枝参数和相关性得分阈值对索引树进行剪枝,在索引树中过滤掉索引无关的子树.基于真实数据集的理论和实验分析表明,所提算法能够抵抗规模分析攻击,并能提高搜索时间效率和搜索准确率.

当今社会已经进入数据时代,大数据技术已经成为云计算之后信息技术领域的另一个信息产业增长点.但是,大数据在给经济发展带来巨大推动力的同时,也面临着巨大的安全风险,大数据的安全问题受到了高度重视.而密码技术是解决大数据安全的核心技术.与传统 PKI 时代相比,大数据安全提出了更高的密码需求,传统的密码安全解决方案已经无法适应大数据时代的各种新的安全需求,“基于高性能密码实现的大数据安全方案”一文提出了一种高性能的密码安全方案,满足大数据时代的各种安全需求,主要解决 3 个安全问题:海量数据的高速加解密问题、高并发的大规模用户身份认证问题、大数据的隐私保护问题.

MIBS 密码是在 2009 年的 CANS 会议上提出的一种轻量级算法,它具有较高的软硬件实现效率,并且能够抵抗差分分析、线性分析等传统密码分析方法,适合运行在资源受限,并有一定安全要求的物联网环境中.“物联网中 MIBS 轻量级密码的唯密文故障分析”一文提出了一种针对 MIBS 密码的新型唯密文故障攻击,即利用新型双重“与”故障模型、新型 Parzen-HW 和 Parzen-HW-MLE 区分器对中间状态进行分析,进而破译 MIBS 密码.该方法最少使用 72 个故障即可破译出原始密钥,并且成功率不低于 99%,可以进一步降低导入的故障数和时间,有效地提高了攻击效率.

3 智能隐私保护

智能隐私保护部分共收录了6篇论文,主要围绕智能电网、机器学习、无人驾驶和物联网等研究方向展开.

在抗量子计算领域,基于格的密码学是备受瞩目的.针对智能电网中细粒度的用户能耗相关数据带来的安全与隐私方面的挑战,“后量子的智能电表隐私保护方案”一文提出了一种基于格的可链接环签名来构造抗量子的保护用户隐私的智能电表数据采集方案.作者选择了一个较为先进的基于格的在one-out-of-many证明之上构造的次线性大小的环签名方案,并为其增添可链接性以期为抗量子的隐私保护系统提供异常用户监测和追踪功能.利用后量子签名方案,该系统可以支持动态的用户加入和撤销,拥有更好的灵活性与实用性.

逻辑回归是机器学习重要算法之一,为解决集中式训练方式不能保护隐私问题,“基于数据纵向分布的隐私保护逻辑回归”一文提出了一种隐私保护的逻辑回归解决方案.此方案适用于数据以特征维度进行划分,纵向分布在两方情况下,两方进行协作式训练学习到共享的模型结构.两方在本地数据集上进行训练,通过交换中间计算结果而不直接暴露私有数据,利用加法同态加密算法在密文下进行运算保证计算安全,保证在交互中只能获取到对方的零知识.同时,提供隐私保护的预测方法,保证模型部署服务器不能获取询问者的私有数据.经过分析与实验验证,在几乎不损失精度的前提下,该方案可以在两方均是半诚实参与者情况下提供隐私保护.

基于深度学习的JPEG数字图像隐写分析模型检测能力已超越基于人工设计特征隐写分析模型,但检测能力仍存在提升空间.“基于卷积神经网络的JPEG图像隐写分析参照图像生成方法”一文以进一步提升JPEG隐写分析模型的检测能力为目标,借助深度学习方法,为基于深度学习的JPEG隐写分析模型提供辅助信息,从数据输入角度,探索进一步提升隐写分析模型检测能力的途径.基于卷积神经网络,构建隐写分析参照图像生成模型,对待检测图像进行变换,从而获得对应参照图像.之后,将待检测图像与对应参照图像作为隐写分析模型的输入数据,进一步挖掘待检测图像中存在的隐写分析相关信息.

随着云计算的发展与普及,云计算环境下的安全问题日益突出.云取证技术作为事后追责与惩治技术手段,对维护云计算环境安全具有重大意义.云取证技术研究发展尚不完善,云取证面临电子证据不完整、取证开销较大、取证过程智能化不足等难题.为此,“一种基于软件定义安全和云取证趋势分析的云取证方法”一文提出了一种基于软件定义安全和云取证趋势分析的智能云取证方法.首先,提出一种基于软件定义安全的云取证架构,通过软件定义安全分层理念实现云网络与云计算平台协同实时取证.其次,提出基于隐马尔可夫模型的云取证趋势分析算法和基于改进告警质量的IDS告警选择算法,指导云取证架构中的取证策略决策和取证资源调度,实现智能云取证.

实时地图在无人驾驶车辆导航中发挥着至关重要的作用,基于群智感知的实时地图更新方法具有成本低且准确性高的优势.然而,此方法会增加数据及用户身份泄露的风险.为保证上传数据的机密性和用户的匿名性,“一种安全高效的无人驾驶车辆地图更新方案”一文提出了一种安全高效的无人驾驶车辆地图更新方案.在该方案中,利用签密和代理重加密技术,车辆用户对感知数据进行签密,将加密的数据存储在车辆雾节点中,当地图公司希望访问数据时,雾节点将加密的数据发送给云服务平台,云服务平台重新加密数据发送给地图公司,同时,云服务平台无法获得任何有关数据的明文信息.利用聚合签名技术,降低了计算开销.通过对车辆用户的信誉管理,提高了数据的可靠性.

针对物联网中设备资源受限、连接数量大、动态性强等特点,传统的集中式访问控制技术已不完全适用,如何在物联网环境中实现安全高效的访问控制授权成为亟待解决的关键问题.对此,“物联网中基于智能合约的访问控制方法”一文提出了一种基于层级区块链的物联网分布式体系架构.

在该架构中以 ABAC 模型为基础,采用智能合约的方式实现对物联网设备基于属性的域内和跨域的灵活、动态、自动化的访问控制.同时,在属性度量中增加信任值与诚实度动态评估不同域间和设备间的信任关系,保证实体能够履行合约的信用能力和稳定性.

4 智能系统安全

智能系统安全部分共收录了 7 篇论文,主要围绕漏洞挖掘、入侵检测和静态代码分析等研究方向展开.

漏洞是系统安全与攻防对抗的核心要素,漏洞的自动发现、分析、利用是长期以来研究的热点和难点,现有研究主要集中在模糊测试、污点分析、符号执行等方面.针对有限资源条件下的漏洞自动挖掘与利用问题,“有限资源条件下的软件漏洞自动挖掘与利用”一文建立了 Weak-Tainted 程序运行时漏洞模型,提出了一套面向漏洞自动挖掘、分析、利用的完整解决方案;提出了污点传播分析优化方法和基于输出特征反馈的输入求解方法等有限资源条件下的分析方案,提升了漏洞挖掘分析与利用生成能力;论文实现了漏洞自动挖掘和利用原型系统,单台服务器设备可并发运行 25 个漏洞挖掘与分析任务.

为了使开发者能安全准确的使用第三库接口,库设计者提供了各种类型的安全规约,进而保护应用程序免受因库函数的误用而造成的安全攻击.目前,已有的安全规约由于不精确的描述、误导性的代码示例、错误的默认设置、碎片化以及缺少强制性检查等原因而大大影响了其在实际运用中的有效性.为了使开发者能更好地遵守安全规约,“TipTracer:基于安全提示的安卓应用通用漏洞检测框架”一文提出了一个自动化的通用漏洞分析框架 TipTracer.TipTracer 主要包含 2 个部分.首先,TipTracer 定义了一个能形式化描述安全规约的安全性语言,并利用该语言对已知的安全规约进行形式化表述.其次,TipTracer 实现了一个静态代码分析器,用于检查应用程序是否满足安全规约.

入侵检测技术旨在有效的检测网络中异常的攻击,对网络安全至关重要.针对传统的入侵检测方法难以从工业控制系统通信数据中提取有效数据特征的问题,“基于相关信息熵和 CNN-BiLSTM 的工业控制系统入侵检测”一文提出了一种基于相关信息熵和 CNN-BiLSTM 的入侵检测模型.该模型将基于相关信息熵的特征选择和融合的深度学习算法相结合,因此能够有效去除噪声冗余,减少计算量,提高检测精度.首先针对不平衡样本等问题进行相应预处理,并通过基于相关信息熵的算法进行特征选择,达到去除噪声数据和冗余特征的目的.然后分别运用卷积神经网络(CNN)和双向长短期记忆神经网络(BiLSTM)从时间和空间维度提取数据特征,通过多头注意力机制进行特征融合,进而得出最终检测结果,最后通过单一变量原则和交叉验证方式获得最优的模型.

目前,恶意代码的产生越来越简单、各种变体越来越多,已经成为网络安全的主要威胁.如何自动、准确检测恶意代码一直是各大反病毒厂商和研究人员的研究热点.“一种基于概率主题模型的恶意代码特征提取方法”一文提出了一个全新的恶意代码检测框架,将最简单的概率主题模型——潜在狄立克雷分布(latent Dirichlet allocation, LDA)应用到恶意代码样本的特征提取中,获得反汇编文件中汇编指令的潜在“文档-主题”、“主题-词”的分布,构造恶意样本的新的特征提取办法,给出汇编指令标准化规则以及最优“主题”数目的快速评价和确定方法,同时解析了“文档-主题”、“主题-词”在恶意代码汇编指令中的语义可解释性.

作为物联网技术的典型应用,智能家居平台正逐步走进千家万户,但其存在的安全问题阻碍了其进一步的部署.“面向智能家居平台的信息物理融合系统安全”一文回顾了当前智能家居平台的典型架构,并分析了其各个组成部分存在的攻击接口.在信息接口安全方面,分析了存在于如智能摄像头的图像接口与语音控制系统的对抗样本攻击问题;在云端后台安全方面,分析了执行云端智能应用时对安全规则的破坏,以及造成的隐私泄露等问题.针对存在于智能家居中的恶意应用问

题,则提出了一种基于无线流量分析的第三方恶意软件检测系统,能够在不修改智能家居平台的情况下,实现对恶意应用的精准检测.

智能环境下数据存储和处理的方式正在不断改变,其中安全和效率是2个重要的因素.就安全而言,在数据共享的前提下保护隐私势在必行.就效率而言,智能环境中存在诸多资源受限的设备,针对这些设备如何设计高效的算法或协议直接决定其可行性.为满足以上2个需求,"适用于智能环境的高效安全云辅助模式匹配协议"一文首次在双云服务器辅助的安全两方计算模型下给出安全模式匹配协议的功能函数,并基于茫然传输(oblivious transfer, OT)给出协议的具体构造.假设云服务器和参与方之间不合谋,协议在半诚实敌手模型下是安全的.协议需要4轮交互,模式方仅需要执行少量的异或操作,而复杂的OT协议主要集中在数据库方和云服务器之间.

分布式生物特征认证系统因不依赖弱口令或硬件标识物而获得高的可靠性、安全性和便利性,但也因生物特征存在永久失效和隐私泄露的风险而面临更多的安全威胁.基于同态加密技术的生物特征认证方案允许特征向量在密文域匹配以保护向量安全和用户隐私,但也因此要在密文域执行昂贵的乘法运算,而且还可能因为向量封装不当而遭受安全攻击."一种基于同态加密的分布式生物特征认证协议"一文以Brakerski等人的同态加密方案为基础,提出了一种安全向量匹配方法,并设计了一个口令辅助的生物特征同态认证协议.该协议无需令牌等硬件标识物,注册时只需将带有辅助向量的特征模板密文和口令作用生成的辅助向量哈希值外包存储,认证时服务器使用辅助向量匹配法完成模板向量和请求向量的相似性评估即可实现用户身份认证.

承蒙各位作者、审稿专家和编辑部等方面的全力支持,本专题得以顺利出版.目前密码学与智能安全研究涉及领域十分广泛,这给审稿人及特邀编辑的审稿、选稿工作带来了巨大挑战.由于投稿数量大、主题广泛、时间安排紧张、专题容量有限等原因,本专题仅选择了部分有代表性的研究工作予以发表,无法全面体现密码学与智能安全领域所有的最新研究工作.部分优秀稿件无法列入发表,敬请谅解.

我们要特别感谢《计算机研究与发展》编委会和编辑部,从专题的立项到征稿启事的发布,从审稿专家的邀请到评审意见的汇总,以及最后的定稿、编辑和出版工作,都凝聚了他们辛勤的汗水.本专题的出版期望能给广大相关领域研究人员带来启发和帮助.在审稿过程中难免出现不尽人意之处,希望各位作者和读者包容谅解,同时也请各位同行不吝批评指正.

最后,再次衷心感谢各位作者、审稿专家、编辑部和特邀编委的辛勤工作.

曹珍富 (华东师范大学)
徐秋亮 (山东大学)
张玉清 (中国科学院大学)
董晓蕾 (华东师范大学)

2019年9月