GM/Y 5014-2024

涉及身份管理的个人信息保护技术研究

密码行业标准化技术委员会
CRYPTOGRAPHY STANDARDIZATION TECHNICAL COMMITTEE

2024 年 12 月

摘 要

涉及身份管理的个人信息保护技术研究，主要研究国内外先进的个人信息保护技术、标准与政策，并分析我国涉及身份管理功能的系统中个人信息面临的风险与保护技术措施现状。研究并分析身份管理系统中涉及到的个人信息，以及安全需求和密码技术应用需求，并结合我国身份服务领域的国情，设计适用于我国的网络身份服务方的个人信息安全保护技术框架，结合密码技术对身份管理系统中涉及的个人信息处理过程提出相应技术要求，并基于最新国内法律法规提出安全措施建议。保障身份管理系统中个人信息的安全，为用户使用网络身份服务、身份服务提供方部署安全措施提供指导。

关键词：身份服务，个人信息，动态鉴别，数据脱敏，差分隐私

目 录

前 言

本报告按照 GB/T1.1-2020 给出的规则起草。

本报告由密码行业标准化技术委员会提出并归口。

本报告起草单位：中国科学院数据与通信保护研究教育中心、北京数字认证股份有限公司、公安部第一研究所、中国电子技术标准化研究院、中国科学院软件研究所、国民认证科技（重庆）有限公司、山东大学。

本报告主要起草人：彭佳、高能、李敏、刘中、李景华、何延哲、张立武、李俊、孔凡玉等。

涉及身份管理的个人信息保护技术研究

1. 概述

1.1 背景

随着大数据技术的发展，对用户的信息进行收集和处理分析已经成为企业盈利的重要手段，也给用户隐私、企业安全、国家安全带来了挑战。基于大量的用户身份信息和用户网络行为信息，可以利用大数据推理分析出更多的隐私信息或趋势和规律等信息，这些信息犹如宝贵的财富可以用于企业决策、精准营销，也可成为犯罪分子的利器。一方面，用户的身份信息大多敏感，一旦泄露或者被非法获取，将给用户的隐私、甚至生命财产安全带来极大威胁。另一方面，只有流动起来的数据才能更大发挥其价值，信息的共享已经成为趋势。因此个人信息或者身份信息的保护和使用之间的平衡，已经成为业界研究的热点。因此，亟待制定身份管理系统的个人信息安全保护技术要求，保障网络中个人信息的安全性，为用户使用网络身份服务、身份服务提供方部署安全措施提供指导。

1.1.1 个人信息或者身份信息的保护和使用之间的平衡，成为业界研究的热点

对用户的信息进行收集和处理分析已经成为企业盈利的重要手段，也给用户隐私、企业安全、国家安全带来了挑战；用户的身份信息大多敏感，一旦泄露或者被非法获取，将给用户的隐私、甚至生命财产安全带来极大威胁。只有流动起来的数据才能更大发挥其价值，信息共享已经成为趋势。所以通过安全保护等技术手段寻求个人信息的充分共享和利用与个人隐私的安全的平衡点，是当前产业界、学术界要解决的热点问题之一。

1.1.2 掌握大量用户信息的 IT 服务提供方及其数据中心要应对合规性检测

各项与个人信息、身份信息相关的国际国家标准相继出台，包括我国《中华人民共和国网络安全法》、《中华人民共和国数据安全法》及与个人信息保护最为相关的《中华人民共和国个人信息保护法》等相关法律，以及国家标准《个人信息安全规范》、《个人信息去标识化指南》等标准规范。2018 年 5 月，欧盟的 GDPR 正式实施。美国也正在更新“数字身份指南”。由此可见，对个人信息保护的重视程度已经提升到新的高度。国内外的 IT 服务提供方、身份服务提供方、数据中心都面临合规性检测。需要尽可能采用技术和管理手段来更加合规的收集、分析和管理用户信息，并采用数据保护技术、密码技术、管理手段来保护所收集的用户数据，否则将面临违法或巨额的罚款。

1.1.3 应对网络身份服务的多样化、差异化、复杂化发展趋势

个人信息安全保护技术研发是我国身份管理产业自主发展的关键驱动，同时也是我国个人信息安全标准的制定和落地实施的重要支撑。加强我国个人信息安全保护技术研究，包括分布式环境下的数据加密、数据完整性验证、去标识化技术、差分隐私、动态认证、数据脱敏与安全审计等技术，是应对网络身份服务的

多样化、差异化、复杂化发展趋势的重要举措。同时，加强身份服务系统在网络安全防护方面的研究，包括入侵检测、安全态势感知、网络攻击取证、威胁情报分析等，可以抵御针对个人信息的网络攻击威胁。

1.1.4 我国亟需信息保护技术在身份服务系统的应用相关政策与标准体系

为提升身份管理服务产业的安全保障能力，维护网络安全秩序，考虑到个人信息安全问题的泛在性、复杂性、专业性的现状，需要建立涉及身份管理的个人信息安全保护标准体系研究长效机制。建议立足我国国家安全管理要求和有关产业发展现状，借鉴国际国外标准化工作模式和经验，逐步建立我国个人信息安全标准体系。持续规划个人信息安全保护标准有关规范、标准化体系等方面研究，以充分发挥个人信息安全标准的支撑作用，有效引导大数据安全标准化工作科学推进。

身份管理领域涉及大规模个人信息，作为战略性和基础性重点领域，应依法保护公民的个人信息安全，加快相关标准的出台步伐，规范商业利用与数据资源安全和个人信息保护之间的关系，重点针对个人信息的收集和使用环节建立规则，明确身份管理生态中不同主体的责任，促进网络基础设施的发展，加强网络安全与个人信息保护。加快推动数据资源开放共享和开发应用的同时，建立个人信息安全保障体系，推进涉及身份管理的个人信息保护安全标准应用，覆盖标准研制、验证和推广等标准化活动。促进产业界安全标准研制与科技研发的衔接，增强标准与技术环境的适应能力，保障重点领域个人信息安全标准实施应用效果。构筑适应身份管理服务发展的规范制度，健全数据时代个人信息安全新秩序。

1.2 研究目标

本项目旨在研究国内外的个人信息保护政策和标准，按照收集、传输、存储、使用、销毁等个人信息生命周期过程，提出涉及身份管理的个人信息全生命周期的安全需求，并讨论在身份管理系统中可采用的保护技术。结合我国网络应用和可信身份服务体系建设现状，提出涉及身份管理功能的系统中，保护用户的个人信息所采用的技术框架，以满足不同种类个人信息以及众多应用领域的、不同程度的安全保护需求。编写《涉及身份管理的个人信息保护技术研究》标准研究报告，用于指导网络身份服务方根据安全需求部署安全措施、实施个人信息安全保护自查，同时为监管机构和用户管理网络身份服务方的个人信息提供依据。

2. 国内外个人信息保护战略及标准研究现状

2.1 国内外个人信息保护战略相继出台不断提升战略重要性

2.1.1 我国相关法案与规范

我国越发重视个人信息保护，近年来相继出台《中华人民共和国网络安全法》、《中华人民共和国数据安全法》及与个人信息保护最为相关的《中华人民共和国个人信息保护法》。这些法律法规进一步增强法律规范的系统性、针对性和可操作性，在个人信息保护方面形成更加完备的制度、提供更加有力的法律保障。

（1）《中华人民共和国个人信息保护法》颁布

《中华人民共和国个人信息保护法》（以下简称《个人信息保护法》）自 2021 年 11 月 1 日起施行[24]。规范了个人信息的收集、处理和利用，保护自然人个人信息权以及其他合法权益，促进个人信息的合理利用，规范个人信息跨境传输。

《个人信息保护法》中明确规定，个人信息是以电子或者其他方式记录的与已识别或者可识别的自然人有关的各种信息，不包括匿名化处理后的信息。个人信息的处理包括个人信息的收集、存储、使用、加工、传输、提供、公开、删除等。

《个人信息保护法》确立了个人信息处理应遵循的原则，强调处理个人信息应当采用合法、正当的方式，具有明确、合理的目的，限于实现处理目的的最小范围，公开处理规则，保证信息准确，采取安全保护措施等，并将上述原则贯穿于个人信息处理的全过程、各环节。《个人信息保护法》吸取了来自世界各国和各地区的先进经验。与此同时，其结合了我国国情及长期以来面对的各类困难，使得自然人个体、互联网企业、国家安全和公共利益、跨境传输技术要求以及国际上通行的惯例等达到了一定的平衡。对我国国内互联网企业的合规和我国政府监管提出了更高要求。

（2）《中华人民共和国数据安全法》

《中华人民共和国数据安全法》自 2021 年 9 月 1 日起施行[25]，是为了规范数据处理活动，保障数据安全，促进数据开发利用，保护个人、组织的合法权益，维护国家主权、安全和发展利益，制定的法律。

其中，第三十八条提及“国家机关为履行法定职责的需要收集、使用数据，应当在其履行法定职责的范围内依照法律、行政法规规定的条件和程序进行；对在履行职责中知悉的个人隐私、个人信息、商业秘密、保密商务信息等数据应当依法予以保密，不得泄露或者非法向他人提供”；第五十三条提到，“涉及个人信息的数据处理活动，还应当遵守有关法律、行政法规的规定”。这些条款特意指出“个人信息”的重要性，作为一类重要的数据类型，个人信息关乎人民群众生活安宁、生命健康和财产安全等问题，需要得到重视，明确个人信息处理的原则和规则。

（3）《中华人民共和国网络安全法》

我国于 2017 年 6 月 1 日正式实施《中华人民共和国网络安全法》[1]（以下简称《网安法》），是我国首部全面规范网络空间安全管理方面问题的基础性法律。全文共七章七十九条，针对数据安全和隐私保护，大致分为总体要求、分类分级保护、数据安全存储、个人信息安全多个方面，从上述几个方面保障数据安全和

个人信息安全。

总体要求方面，第十条从整体要求的角度，提出采取技术措施和必要措施维护数据完安全。第二十七条从网络运行安全角度，提出不得从事窃取网络数据的活动、不得提供窃取网络数据的程序和工具。

分类分级保护方面，第二十一条指出国家实行网络安全保护制度，采取数据分类、重要数据备份和加密等措施，防止网络数据泄露或者被窃取、篡改。

在数据安全存储方面，明确了境内和境外的重要数据的存储要求，第三十七条规定关键信息基础设施的运营者在境内运营中收集和产生的个人信息和重要数据应当在境内存储。因业务需要，确需向境外提供的,应当按照国家网信部门会同国务院有关部门制定的办法进行安全评估；法律、行政法规另有规定的,依照其规定。第六十六条对违反第三十七条的规定行为给出具体的惩治措施。

个人信息安全方面，第三十七条从个人信息存储角度提出境内外数据存储要求。第四十二条明确网络运营者应当采取技术措施和其他必要措施,确保其收集的个人信息安全，防止信息泄露、毁损、丢失。

（4）《中华人民共和国密码法》

《中华人民共和国密码法》自 2020 年 1 月 1 日起施行，这标志着我国在密码的应用和管理等方面有了专门性的法律保障。密码法设立该制度，是维护国家安全和社会公共利益的需要。

商用密码产品、服务是专业技术性很强的特殊产品和服务，广泛应用于国民经济和社会发展各领域，应用于关键信息基础设施，其质量与安全性直接关系国家安全和社会公共利益，需要通过检测认证的方式对其质量与安全性进行技术把关。

2.1.2 美国身份隐私保护管理法案

美国将可信身份管理上升到了国家战略。2010 年 6 月，美国白宫宣布公布“美国网络空间可信身份国家战略”草案，提出构造“可信的安全身份生态系统”，旨在通过使用特殊 ID 证明用户身份，保证网络空间交易的安全性。2011 年 4 月，美国发布了《网络空间可信身份国家战略》（National Strategy for Trusted Identities in Cyberspace，NSTIC），指出可信身份是改善网络安全的基石，计划用 10 年左右的时间，构建一个网络身份生态体系，推动个人和组织在网络中使用安全、高效、易用的身份解决方案。为此，美国成立了专门的主管办公室，负责协调政府和私人部门的活动，并牵头制定实施路线图。NSTIC 明确身份生态体系必须遵循四个原则：一是身份解决方案应当增强隐私保护并且由公众自愿应用；二是身份解决方案应当是安全、可扩展的；三是身份解决方案应当是互操作的；四是身份解决方案应当是高效且易于应用的。这四个指导原则是任务目标和行动实施的基础。国家战略和法规的制定推进了对网络身份分级的研究。

2019 年 1 月 14 日，信息技术和创新基金会（ITIF）发布关于数据隐私立法的报告（A Grand bargain on data privacy legislation for America），报告比较了世界各地不同的法律和框架如何解决各种数据隐私问题，描述了现有法律、框架和立法提案中包含的 30 个组成部分，并解释了对消费者、企业和数字经济的可能影响。在此基础上，报告呼吁建立制定全面的数据隐私立法，扩展和简化消费者数据隐私权，废除和取代现有的联邦隐私法，提供一套共同的保护措施，降低现有州和联邦法规的合规成本，并为更多数据驱动的创新铺平道路。

2019 年 2 月 27 日，美国参议院提出《数据隐私法案》（Digital Accountability

and Transparency to Advance Privacy Act or the DATA Privacy Act），加强了对美国消费者的数据隐私保护，同时确保企业专注于实施新的数据安全标准以及采用必要的隐私保护措施。法案还增加了对保护美国公民隐私技术研究的规定，并确保小企业免受不必要的监管。

2019 年 3 月 14 日，美国参议院提出《2019 年商业人脸识别隐私法案》提案（Commercial Facial Recognition Privacy Act of 2019），该法案是美国关于人脸识别隐私保护的第一部法案，被认为是联邦监管的重大举措。该法案规定，商业公司在使用人脸识别技术时需要经过用户的明确同意，同时要求对投入市场应用的人脸识别技术进行第三方测试，以确保其符合准确性标准，避免对消费者可能造成的损害。法案禁止在未经用户同意的情况下，将用户数据提供给第三方实体。

2019 年 4 月 9 日，美国参议院提出《2019 年遗传信息隐私法案》（Genetic Information Privacy Act of 2019），该法案禁止基因检测服务在未经其明确同意的情况下披露或使用消费者的个人信息。今年早些时候，美国最大的基因检测公司之一 FamilyTreeDNA 向其用户道歉，未经授权共享他们的 DNA 信息。全国有数百万美国人目前容易受到这种脆弱性的影响。法案加强了用户的隐私权，特别要求用于医学研究的遗传信息符合美国国立卫生研究院的知情同意要求。

2019 年 4 月 10 日，美国参议院提出《2019 算法问责法案》（Algorithmic Accountability Act of 2019 Bill），因为科技公司“自动决策系统”中的算法对美国人的生活和决策的影响越来越大，但这些算法往往依赖而不是消除有偏见的假设或数据。对此，该法案要求大型科技公司评估并消除其“自动决策系统”给个人信息隐私和安全带来的风险，以及因种族、肤色、宗教、政治信仰、性别或其它方面差异带来的歧视性偏见，要求公司纠正他们在影响评估期间发现的任何问题，还要求公司评估其信息系统如何保护消费者个人信息的隐私和安全；

2019 年 4 月 11 日，美国参议院提出《隐私权利法案》（Privacy Bill of Rights Act），该法案旨在从立法层面全面保护个人数据，明确定义了个人数据，还要求联邦贸易委员会出台立法，赋予个人若干权利，包括：接收关于数据控制者收集、使用、存储和共享其个人数据通知的权利；被收集和使用个人信息时的选择加入权；访问个人数据处理活动有关细节的权利；访问可携式电子表中个人数据的权利；纠正不准确的个人数据的权利；要求删除个人数据的权利等。此外，法案还要求联邦贸易委员会颁布立法，禁止数据控制者从事以下类型的活动，包括：因消费者拒绝被收集、使用、存储个人数据，而拒绝为消费者提供服务；提供经济奖励以换取个人同意数据控制者收集、使用其个人数据；在没有某些特定合同条款的情况下，根据书面合同向第三方披露个人数据；将个人数据用于某些“不合理”的目的；

2019 年 5 月 14 日，旧金山通过《停止秘密监视条例》（Stop Secret Surveillance ordinance）的修订[4]，旧金山市监事会于 5 月 14 日投票表决通过了对《停止秘密监视条例》所作的修订，成为美国第一个禁止使用人脸识别技术的城市。条例全面禁止旧金山当地政府部门如警察局、治安官办公室、交管部门等使用人脸识别技术。同时，购买任何类似的新监控设备如自动识别车牌号系统、带有摄像机的无人机等，都需要得到市政府的许可。该条例还要求明确政策，以确定市政府如何使用监控技术。该条例指出，“人脸识别技术危害公民权利和公民自由的倾向大大超过了其声称的好处，这项技术将加剧种族不平等，并威胁到我们不受政府长期监控的生活能力”。

2019 年 6 月 5 日，纽约州奥尔良市参议院民主党多数派通过了《停止黑客

攻击和改进电子数据安全法案》（下文称之为《纽约法案》），要求企业将客户的隐私权置于利润之上，并将个人数据的控制权返还给纽约州人。这是纽约州在ChoicePoint 安全事件后，以 2003 年加州通过的相同法律为范本而制定的新法律。ChoicePoint 泄密事件暴露了 3 万名加州用户的信息，受影响的纽约州用户有 9000 人。该法案早在去年就已经在纽约州众议院通过，但在参院未能得到足够的多数支持。今年，随着一系列重要安全事件的爆发，法案随即得到了州参议院的支持。该法案扩大了数据泄露的定义，同时扩大当前《数据泄露通知法》约束的信息范围，并授权司法部长对侵犯隐私行为提起诉讼。《纽约法案》提高了安全标准，这项立法是隐私和消费者保护的合作方式，将迫使公司最终严肃对待数据安全，保护纽约人最敏感的信息。

2.1.3 欧盟跨界认证研究与 GDPR

欧盟国家和地区加快在信息网络中引入和部署身份管理。在战略层面、技术层面，欧盟为网络身份管理的大范围部署与推广做了充足的准备，开展了 **STORK** 等与身份管理相关一系列的研究。2010 年实现了整个欧盟范围内电子身份的启用，欧盟成员国公民持有电子身份即可在任意欧盟国家享受相应的求职、医疗、保险等一系列社会服务。

（1）跨界安全认证（Secure idenTity acrOss boRders linKed 2.0，STORK 2.0）

该项目致力于欧洲电子身份识别和认证领域，在国家和欧盟的角度，建立不同方法之间的互操作性、个人电子身份、合法实体电子身份，应用领域涵盖电子银行、商业公共服务、学术资格认证、电子医疗领域等。STORK 2.0 将成为实现欧洲 2020 战略（欧洲 2020 战略是欧盟针对当前十年的经济增长和就业问题提出的议程）中开放的、数字化的经济设想的关键环节。

STORK 2.0 中的 QAA（Quality Authentication Assurance）模型将认证保证分为 4 个等级，身份可信保证程度由低到高。STORK QAA level 的具体要求分为两个阶段：注册阶段（线上或线下）和在线电子认证阶段。对于 STORK QAA level 的每一个等级，由组织性因素和技术性因素两个方面决定。注册阶段，需要考虑标识过程质量等级、证书发放实体质量等级、证书发放过程的质量等级；电子认证阶段，需考虑认证机制的安全性等级和证书的类型和稳定性质量等级。这五个方面决定了 STORK QAA 等级。

（2）《通用数据保护条例》

面对愈演愈烈的信息泄露事件，欧盟于 2018 年 5 月份出台《通用数据保护条例》（General Data Protection Regulation，简称 GDPR）[1]，目的在于遏制个人信息被滥用，保护个人隐私，于 2018 年 5 月 25 日正式生效。

欧洲隐私法律的地域适用范围正在通过 GDPR 不断扩大。而正是由于这种适用范围的扩大，位于欧盟境外、不受现行欧洲隐私法律规制的许多组织也将随着 GDPR 的实施而不得不适用欧盟隐私法律。GDPR 要求这些组织及时对 GDPR 的“合规”要求作出回应。

第一，GDPR 适用于在欧盟境内设有业务机构（establishment）的组织，只要这些组织在业务机构在欧盟境内的活动中处理个人数据（而不论此类处理行为是否实际发生在欧盟境内）。因此，如果业务机构（例如分公司或联络代理）的活动与位于欧盟境外的组织进行的个人数据处理密不可分，则应当适用 GDPR。

第二，如某一组织虽不在欧盟境内设立业务机构，但却处理欧盟境内个人的个人数据，并且此类处理行为与向欧盟境内个人提供商品或服务相关，无论该等

商品或服务是否收费，则也应当适用 GDPR。

第三，GDPR 适用于非欧盟组织处理欧盟境内个人的个人数据，只要此类处理行为涉及对这些个人的行为进行监控，且该处理行为发生在欧盟。如果（尤其是）为了作出与这些个人有关的决定或者为了分析或预测其个人喜好、行为和态度，而在互联网上追踪这些个人，且在此过程中使用了处理技术来形成画像等，则构成监控行为。

2.2 国内外标准不断丰富，逐步细化技术要求

2.2.1 我国标准体系现状

（1）GB/T 35273-2020 《信息安全技术 个人信息安全规范》

该标准由全国信息安全标准化技术委员会发布，于 2020 年 10 月 1 日实施，代替 GB/T 35273-2017《信息安全技术个人信息安全规范》。

该标准规定了开展收集、存储、使用、共享、转让、公开披露、删除等个人信息处理活动应遵循的原则和安全要求，适用于规范各类组织的个人信息处理活动，也适用于主管监管部门、第三方评估机构等组织对个人信息处理活动进行监督、管理和评估。

该文件明确了“个人信息”的定义，提出了个人信息安全基本原则、从个人信息的收集、个人信息的存储、个人信息的使用、个人信息主体的权利、个人信息的委托处理、共享、转让、公开披露、个人信息安全事件处置以及组织的个人信息安全管理要求等方面提出了要求。该标准对个人信息控制者对个人信息的保存提出时间最小化、去标识化处理、个人敏感信息的传输和存储等具体要求。对个人信息控制者处理个人信息安全事件的方式方法提出安全事件应急处置和报告、安全事件告知等具体要求；对个人信息控制者组织管理提出明确责任部门与人员、数据安全能力、安全审计等具体要求。

（2）GB/T 39335-2020 《信息安全技术 个人信息安全影响评估指南》

全国信息安全标准化技术委员会发布国家标准《信息安全技术 个人信息安全影响评估指南》于 2021 年 6 月 1 日正式实施。标准主要分别针对个人信息安全影响评估的原理和框架、评估流程、评估的具体实施方式进行了详细的说明，资料性附录中给出了评估过程使用的判定准则和工具表。

《个人信息安全影响评估指南》的编制遵循先进性、开放性、适应性、简明性、中立性、一致性的原则，是《个人信息安全规范》的配套标准。《个人信息安全影响评估指南》的核心思路是针对个人信息处理活动，评估可能对个人权益造成的影响以及风险，影响主要包括四个方面：一是，影响个人自主决定权，比如被强迫执行不愿执行的操作、无法更正错误上传的个人信息、无法选择推送广告的种类、被蓄意推送影响个人价值观判断的资讯；二是，引发差别性待遇，比如隐私信息（疾病、婚史、种族等）泄露造成的歧视、故意设置个人福利、资格、权利的差别等；三是，个人名誉受损或遭受精神压力，比如公开不愿为人知的事实（生活习惯、以往经历等），被频繁骚扰、监视追踪等；四是，个人财产受损或遭受人身伤害，比如帐户被盗、遭受诈骗、被勒索恐吓、限制自由等。

《个人信息安全影响评估指南》规定了个人信息安全影响评估的基本概念、框架、方法和流程，并提出了特定场景下进行评估的具体方法。适用于各类组织自行开展个人信息安全影响评估工作。同时为国家主管部门、第三方测评机构等开展个人信息安全监管、检查、评估等工作提供的指导和依据。

（3）GB/T 37964-2019 《信息安全技术 个人信息去标识化指南》

在大数据、云计算、万物互联的时代，基于数据的应用日益广泛，同时也带来了巨大的个人信息安全问题。为了保护个人信息安全，同时促进数据的共享使用，特制定个人信息去标识化指南标准。该标准旨在借鉴国内外个人信息去标识化的最新研究理论，提炼业内当前通行的最佳实践，研究个人信息去标识化的目标、原则、技术、模型、过程和组织措施，提出能科学有效地抵御安全风险、符合信息化发展需要的个人信息去标识化指南。

该标准关注的待去标识化的数据集是微数据（以记录集合表示的数据集，逻辑上可通过表格形式表示）。去标识化不仅仅是对数据集中的直接标识符、准标识符进行删除或变换，而且应当结合后期应用场景考虑数据集被重标识的风险，进而选择恰当的去标识化模型和技术措施，并实施合适的效果评估。对于不是微数据的数据集，可以转化为微数据进行处理，也可以参照该标准的目标、原则和方法进行处理。比如针对表格数据，如果关于同一个人的记录有多条，则可将多条记录拼接成一条，从而形成微数据，其中同一个人的记录只有一条。

该标准描述了个人信息去标识化的目标和原则，提出了去标识化过程和管理措施。针对微数据提供具体的个人信息去标识化指导，适用于个人信息处理相关方，也适用于网络安全相关主管部门、第三方评估机构等组织开展个人信息安全监督管理、评估等工作。

（4）GB/T 35273-2017 《非结构化数据访问接口规范》

该标准规定非结构化数据管理系统的访问接口要求，包括查询语言访问接口、应用程序访问接口和 Web 服务访问接口。该标准适用于非结构化数据管理系统产品的研制、开发和测试。

（5）GB/T 34978-2017 《信息安全技术 移动智能终端个人信息保护技术要求》

该标准规范了全部或部分通过移动智能终端进行个人信息处理的过程，根据移动智能终端个人信息的分类和不同的处理阶段，对相应的个人信息保护提出了技术要求。

该标准对移动智能终端个人信息进行了分类，并提出了保护原则，进一步将个人信息在移动智能终端中分为收集、加工、转移、删除四个阶段，并对每一阶段提出了相应的安全技术要求。

（6）JR/T 0171-2020 《个人金融信息保护技术规范》

该标准针对个人金融信息在收集、传输、存储、使用、删除、销毁等生命周期各环节的安全防护要求，从安全技术和安全管理两个方面对个人金融信息保护提出规范性要求。

本标准旨在加强个人金融信息安全管理，指导各相关机构规范处理个人金融信息，最大程度保障个人金融信息主体合法权益，维护金融市场稳定。适用于提供金融产品和服务的金融业机构，并为安全评估机构开展安全检查与评估工作提供参考。

2.2.2 ISO/IEC 相关标准体系

（1）ISO/IEC 29100:2011 《信息技术 安全技术 隐私保护框架》

该标准为信息与通信技术（ICT）系统内可识别个人信息（PII）的保护提供了一个高层次隐私保护框架。该隐私保护框架规范了通用的隐私保护术语；定义了处理 PII 中的参与者及其角色；描述了隐私保护的考虑事项；为实现由许多国

际组织开发的 11 个隐私保护原则提供指导。11 个隐私保护原则包括同意和选择、意图合法性和规约、收集限制、数据最小化、使用/保留/披露限制、准确和质量、开放/透明/告知、个体参与和访问、可核查性、信息安全、隐私保护合规。该标准适用于涉及规范、获取、构建、设计、开发、测试、维护、管理和运行需要隐私保护控制措施来处理 PII 的 ICT 系统或服务的任何自然人和组织。

（2）ISO/IEC 29101:2013 《信息技术 安全技术 隐私保护体系结构框架》

该标准定义了一个隐私参考体系结构框架，该框架明确提出了处理 PII 的 ICT 系统的关心点，列出了实现这种系统的组件，并提供了将这些组件语境化的体系结构视图。该标准适用于涉及规划、获取、构建、设计、测试、维护、管理和运行处理 PII 的 ICT 系统的实体。

（3）ISO/IEC 29190:2015 《信息技术 安全技术 隐私保护能力评估模型》

该标准为组织评估其管理隐私保护相关过程的能力提供高层指南，规范了确定隐私保护能力的评估过程和评估级别，为评估隐私保护能力的关键过程域及其实现，以及如何将隐私保护能力评估继承到组织运行中提供了指南。

（4）ISO/IEC 29191:2012 《信息技术 安全技术 部分匿名、部分不可链接鉴别要求》

该标准为部分匿名和部分不可链接鉴别要求提供基本框架并建立相应要求。当前的实体身份验证技术要求揭示被验证的实体的可识别信息。在许多类型的事务中，实体倾向于保持匿名和不可链接，这意味着当执行两个事务时，很难区分事务是由同一用户还是两个不同的用户执行的。但是，在某些情况下，在有正当理由下，可以在以后能够重新确定用户身份(例如责任的利益)。“部分匿名，部分不可链接”意味着一个预先指定的触发器，并且只有该指定的触发器可以识别经过身份验证的实体。本标准阐述了一些部分匿名、部分不可链接的身份验证加密技术。

（5）ISO/IEC 27018:2014 《信息技术 安全技术 可识别个人信息（PII）处理者在公有云中保护 PII 的实践指南》

该标准依据 ISO/IEC 29100 给出的隐私保护原则，为在公有云计算环境中保护可识别个人信息（PII），建立了普遍接受的控制目标、控制措施和测量实现指南。特别是，该标准考虑到在公有云提供者的信息安全风险环境下适用的 PII 保护法规要求，基于 ISO/IEC 27002 给出指南。该标准适用于作为 PII 处理者通过云计算提供信息处理服务的所有类型和规模的组织。

（6）ISO/IEC 29151:2017 《信息技术 安全技术 可识别个人信息保护实践指南》

该标准为隐私影响评估（PIA）过程以及 PIA 报告的结构和内容给出指南。该标准适用于所有类型和规模组织。随着个人信息被盗用事件数量的增加，收集或处理个人信息被盗用事件的组织将越来越需要关于如何防止个人信息被盗用的指导，以减少隐私被盗用事件发生的风险，并减少被盗用事件对组织和有关个人的影响。本规范提供了这样的指导。

（7）ISO/IEC 29134:2017 《信息技术 安全技术 隐私影响评估指南》

该标准为满足通过可识别个人信息（PII）保护相关的风险和影响评估而识别的要求，建立了控制目标和控制措施，并提供了控制措施实现指南。该标准考虑到在组织信息安全风险环境下适用的 PII 处理要求，基于 ISO/IEC 27002 给出指南。该标准适用于作为 PII 控制者的所有类型和规模的组织。

（8）ISO/IEC TR 27550:2019 《信息技术 安全技术 系统生命周期过程的

隐私工程》

该标准针对企业如何将隐私保护工程与自身工程实践相结合，给出了相应的体系架构和指南，具体包括以下内容：1）隐私工程和其他工程（如系统工程、安全工程、风险管理）之间的关系；2）与知识管理、风险管理、需求分析、架构设计等关键工程过程相关的隐私工程活动；3）隐私工程附录，如与隐私工程相关的实体，考虑到域管理、供应链、软件开发方法等因素后的隐私工程实践，用于指导隐私工程活动的目录，以及隐私风险分析的案例。

该标准的目标人员是需在系统开发、应用或操作维护过程中考虑隐私工程的专业人员和工程师，也适用于企业中负责隐私工程、系统开发、产品管理、市场和运维的相关人员。

（9）ISO/IEC 27551:2021 《信息安全 网络安全和隐私保护-基于属性的不可链接实体认证要求》

目前实体认证都要求被认证实体提供可识别的身份信息，但在很多交易中，实体更倾向于维持匿名化或非链接性，这就使得完成两笔交易时，很难区分交易是由一个用户还是两个不同的用户完成的。该标准正是针对基于属性的非链接实体认证提出了架构并建立相应要求。

（10）ISO/IEC 29184:2020 《信息技术 安全技术 在线隐私通知和同意指南》

宽带网络等通信基础设施的快速普及、智能手机和可穿戴设备等可收集用户详细信息的终端的广泛应用、信息处理能力的大幅度提升，使得大范围信息收集和分析成为可能。在技术升级给用户带来使用便利性和有吸引力的服务并催生新商机的同时，用户也变得对“隐私”越来越敏感，对在线服务中的 PII（个人识别数据）收集和使用产生的影响越来越存疑。这种质疑通常是由于未对如何使用、处理、存储个人 PII 数据进行明确的解释造成的。

该标准为企业提供了一个基本架构，可向被收集 PII 数据的用户提供明晰、易于理解的基本信息，解释企业将如何处理这些 PII 数据。同时，该标准为如何落实 ISO/IEC 29100 中的两个隐私原则（原则 1：同意和授权；原则 7：开放、透明和通知）提供了详细指南。

（11）ISO/IEC 27701:2019 《隐私信息管理体系》

ISO/IEC 27701 是在隐私保护方面对 ISO/IEC 27001《信息安全管理》以及 ISO/IEC 27002《信息安全控制实用规则》的进一步拓展。该标准针对保护可能受到个人信息收集和处理影响的隐私提供了更多相关指南。设计的目的在于借助更多的要求增强现有信息安全管理体系（ISMS），以建立、实施、维护和持续改进隐私信息管理体系（PIMS）。标准概述了适用于个人身份信息（PII）控制者和 PII 处理者的框架，以有效管理隐私控制，降低个人隐私权面临的风险。

关于个人隐私保护的系统性要求标准还有 BS 10012：2017《数据保护 个人信息管理系统规范》，是特定于英国发布的标准，最初由英国标准协会（BSI）于 2009 年发布，2017 年修订，2018 年按照 GDPR 法案要求做了更新。该标准主要是针对个人信息保护提出了“个人信息保护标准”，规范了 PIMS 要求，提供了一个框架用于维护和改进数据保护的合规性实践。主要为在其内部启动、实施和维护 PIMS 的组织所用，使得内部和外部评估者能够有效地评估数据保护的合规性和最佳实践。这两项标准的详细比较，可参考[28]。

2.2.3 NIST 相关标准化现状

美国国家标准与技术研究院(NIST)发布隐私框架草案，以通过风险管理帮助企业改善个人隐私保护。NIST 表示，隐私框架旨在通过三个事项帮助企业保护个人隐私：在服务和产品中支持道德决策以建立客户信任；履行合规义务；促进与客户和监管机构就隐私实践进行沟通。

该政策遵循网络安全框架的结构，由核心、概况和实施层组成。核心部分旨在促进关于隐私保护运营和期望结果的对话；概况部分推进满足组织使命和隐私价值的活动与结果的优先次序；实施层则对组织处理隐私风险流程的充分性进行沟通和决策提供支持。

（1）NIST SP 1500-4《大数据互操作框架：第 4 册 安全与隐私保护》

该框架由 NIST NBD-PWG 的安全与隐私保护小组负责编写。其中第四部分包含有关大数据的安全性和隐私性主题的探讨。考虑了与大数据有关的安全和隐私的新方面，审查了安全和隐私用例，提出了安全和隐私分类法，提出了 NIST 大数据参考体系结构（NBDRA）的安全和隐私结构的详细信息，并开始进行映射 NBDRA 的安全和隐私用例。

（2）NIST SP 800-63《数字身份指南》

该指南涵盖了通过开放网络与政府 IT 系统交互的用户（例如员工，承包商或私人）的身份证明和身份验证。他们在身份证明，注册，身份验证器，管理过程，身份验证协议，联合身份验证和相关声明的各个领域中定义技术要求。

（3）NIST SP 800-128《信息系统安全配置管理指南》

NIST SP 800-128 假定信息安全是组织整体配置管理的组成部分。该文档的重点是配置管理的信息系统安全性方面的实现，因此，“以安全性为重点的配置管理（SecCM）”用于强调对信息安全性的关注。除了与 SecCM 相关的基本概念之外，还介绍了将 SecCM 应用于信息系统的过程。SecCM 活动的目标是管理和监视信息系统的配置，以实现足够的安全性并最大程度地降低组织风险，同时支持所需的业务功能和服务。

（4）NIST SP 800-37《信息系统和组织的风险管理框架:安全和隐私的系统生命周期方法》

该出版物描述了风险管理框架（RMF），并提供了将 RMF 应用于信息系统和组织的指南。RMF 为管理安全性和隐私风险提供了纪律严明，结构化且灵活的流程，其中包括信息安全性分类；控制选择，实施和评估；系统和通用控制授权；并持续监控。RMF 包括准备组织在适当的风险管理级别执行框架的活动。RMF 还通过实施持续监控流程来促进近乎实时的风险管理和持续的信息系统以及通用控制授权；为高级领导者和高管提供必要的信息，以就支持其任务和业务功能的系统做出有效，具有成本效益的风险管理决策；并将安全性和隐私性纳入系统开发生命周期。执行 RMF 任务将系统级别的基本风险管理流程链接到组织级别的风险管理流程。此外，它为在组织的信息系统内实施并由这些系统继承的控制措施建立责任和问责制。

（5）NIST SP 800-53 Rev 4《联邦信息系统和组织的安全和隐私控制》

该出版物提供了针对联邦信息系统和组织的安全和隐私控制的目录，以及选择用于保护组织运营（包括任务，职能，形象和声誉），组织资产，个人，其他组织以及国家的控件的过程。各种各样的威胁，包括敌对的网络攻击，自然灾害，结构故障和人为错误。这些控件是可自定义的，并且是组织范围内管理信息安全和隐私风险的过程的一部分。这些控制措施满足了联邦政府和关键基础设施的一系列不同的安全和隐私要求，这些要求源自立法，行政命令，政策，指令，法规，

标准和/或任务/业务需求。该出版物还介绍了如何针对特定类型的任务/业务功能，技术或运营环境开发专门的控件集或叠加集。最后，安全控制目录从功能角度（提供的安全功能和机制的强度）和保证角度（对已实现的安全功能的置信度）进行处理。同时解决安全功能和安全保证问题，可以确保信息技术产品以及使用可靠系统和安全工程原理从这些产品构建的信息系统具有足够的信任度。

2.3 发展趋势

个人信息的范围逐步扩大，除了生物特征信息（如人脸、虹膜和指纹）等传统个人信息外，需要将个人的行为特征信息等也纳入考虑范畴。根据生物特征进行身份认证和识别已经成为社会发生的一个主流趋势，例如火车站和机场安检的人脸识别以及商场的人脸验证付款，人们的生物特征已经逐渐成为不法分子企图获取的重要信息。另外，越来越多的互联网公司企图通过人们在互联网上的行为特征去进行用户行为轮廓建模定位和捕捉个人信息，达到推送广告等各种目的。因此，我国在进行个人隐私保护法案制定时应该将这些个人信息也考虑在内，全方位保障用户的个人信息安全，使得用户对自己的个人信息有着绝对的掌控权。

根据业务规模创建合理的个人信息安全要求，并为验证其保护个人信息的措施提供有针对性的手段和依据。不同企业有着不同的业务规模，有些企业的产品可能只是针对一个公司内部的所有用户，而另外一些企业的产品则是面向全国用户，针对这些不同规模的企业应该制定不同的个人安全政策规范，不可一视同仁，否则将导致政策在实际实施时很难真正落地。另外，一些企业的业务场景需要对个人的数据信息进行验证识别，这些企业势必会接触到大量的个人可识别数据，如何让这些企业实施合规的个人信息安全保护措施，并使得监督管理措施有据可依，是国家标准体系建设过程中需满足的切实需求之一。

近年来，**有关个人信息保护的法律法规，标准措施发布、迭代速度明显加快。**标准体系根据技术更迭快速响应，系统级技术框架应有弹性、可扩展，灵活与规范并存。标准制定应密切跟踪新兴安全保护技术发展情况，促进科研领域和企业的技术交流共享，加快跨界融合、系统整合的标准创新业态建设。建立相关标准动态调整机制，发挥企业、联盟的积极作用和认证认可对标准调整的推动作用，建立适应技术更迭和产业变革要求的标准动态调整和快速响应机制。技术框架应具备适应性，在符合规范的前提下，提供多样化的技术选择范围，具体实施时有弹性、可扩展。同时，可为企业提供标准实施的监督指导。

3. 涉及身份管理的个人信息保护需求

3.1 身份管理系统交互模型

身份管理系统是一种支撑网络身份服务的软硬件集合，提供给用户或其他授权的实体对资源的合规正确的访问权限。身份管理的重点是为用户、服务等主客体定义一个数字（电子）身份，为该身份关联属性，并完成用户的身份鉴别，从而使持有该身份的主体得到访问权限，享受权限范围内的服务。

图 3-1 显示了身份管理服务各参与方之间的交互。过程 1-5 显示了普通网络身份认证的基本流程。当一个用户希望使用网络服务（依赖方）1 时，由于网络服务 1 并不记录用户的身份信息，网络服务 1 会连接到身份服务提供方 1，用户在身份服务提供方录入相关个人信息并通过身份管理系统的鉴别后，身份管理系统通过网络向用户发送登录凭证，用户持有凭证即可登录网络服务 1。上述过程对用户是完全透明的。一个依赖方可以接受和使用一个或多个身份服务提供方提供的登录凭证；一个身份服务提供方也可以为多个依赖方提供身份服务。

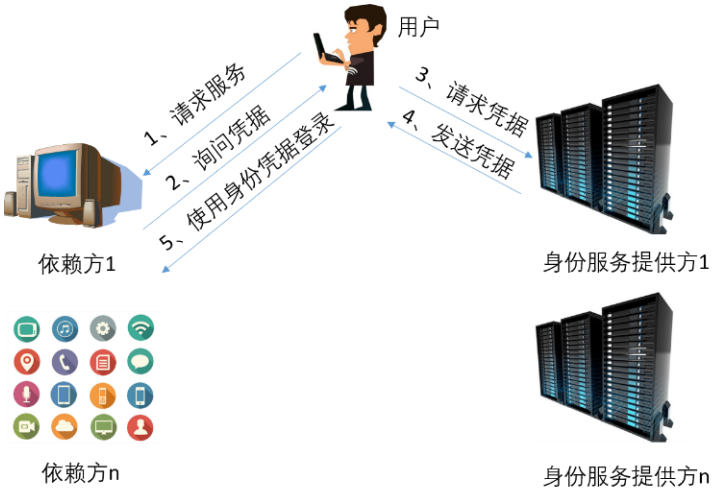

图 3-1 身份管理各参与方的关系

身份管理系统为用户提供身份核验、凭证管理、身份鉴别、权限管理和身份联合等身份服务的软硬件集合，通常是作为单个整体解决方案实施的。本文以身份管理系统为核心视角，将身份管理系统主要分为三个组件：身份管理，鉴别和授权，如图 3-2 所示。身份管理组件主要包括身份核验、注册用户和身份凭证管理等功能；身份鉴别组件主要包括身份鉴别、持续追踪与保持；授权组件为基于鉴别结果执行访问控制策略。

身份管理组件主要包括身份核验、用户注册、凭证生成和持续监控功能。身份管理的目的是定义逻辑上与物理实体相关联的数字身份。物理实体基于申请者的身份证明文件，例如身份证、护照或许可证等进行核验后注册。数字身份是为在系统上建立存在而产生的人工产物。这个数字身份，一旦通过鉴别，授权组件就允许或限制该数字身份的各项权限。同时，身份管理组件还要对管理中的身份进行持续监控，定期更新属性信息。

身份鉴别组件承担在信息系统中确认通信方的身份及其附属信息的功能。利用身份鉴别技术，确定通信方的身份，就能够在此基础上完成其它各种安全服务，

包括访问控制、授权、安全审计、数据分发、人员管理等等。在几乎所有的信息系统中，身份鉴别功能都是必须的功能。

授权组件执行访问控制功能，确保特定的数据和资源能够在合适的时间和地点，被合适的主体正确地访问和利用。访问控制技术的发展主要体现在策略和协议的演进，授权策略方面，目前广泛使用的有自主访问控制、基于角色的访问控制，基于属性的访问控制，基于任务、基于策略以及结合密码技术等的细粒度或动态访问控制策略正在研究推进。

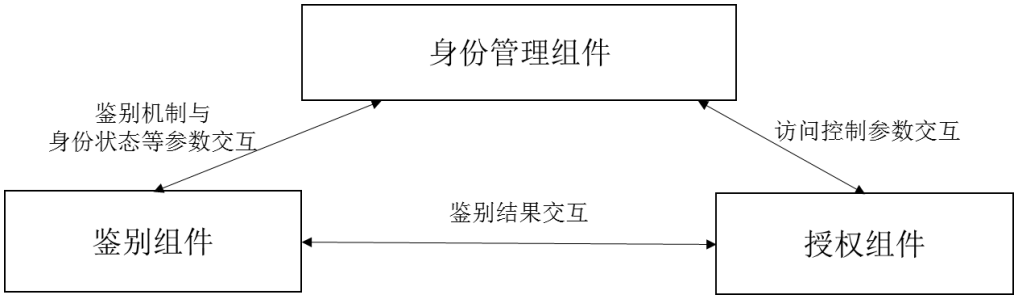

图 3-2 身份管理系统各组件运行过程

3.2 个人信息定义与生命周期

《个人信息保护法》第四条规定，个人信息是以电子或者其他方式记录的与已识别或者可识别的自然人有关的各种信息，不包括匿名化处理后的信息。《网络安全法》第七十六条中明确定义了个人信息的概念，即以电子或者其他方式记录的能够单独或者与其他信息结合识别特定自然人的各种信息，包括自然人的姓名、出生日期、身份证件号码、生物识别信息、住址、电话号码、电子邮箱、健康信息、行踪信息等。除此之外，公安部联合北京网络行业协会以及公安部第三研究所共同发布的《互联网个人信息安全保护指南》中对个人信息的定义进行了进一步的扩展，认为该个人信息还包括通信通讯联系方式、通信记录和内容、账号密码、财产信息、征信信息、行踪轨迹、住宿信息、健康生理信息、交易信息等。

在身份管理系统中，个人信息是由各种电子化记录集合而成，能够识别用户实体身份及相关属性证明，用于为用户提供身份服务所必须收集处理的数据集。《个人信息保护法》第四条指出：个人信息的处理包括个人信息的收集、存储、使用、加工、传输、提供、公开、删除等。在每一个环节中，需要达到一定的安全保护目标：

（1）各环节防止个人信息被恶意使用，防止个人信息泄露；

（2）断开属性与标识间的关联关系，即拿到部分属性信息，无法进行关联，无法窥探某个电子标识对应的全部属性；

（3）断开标识与数字身份之间的对应关系，拿到标识类信息无法确定是哪一个数字身份；

（4）断开数字身份与物理实体间的对应关系，即使攻击者拿到数字身份标识，无法确定出是现实世界中哪一个自然人。

本文聚焦于身份管理系统，首先考虑在身份管理的各个组件中涉及到哪些对个人信息的操作。针对每一个环节，分析安全需求，并给出相应的安全防护技术的使用建议。

3.3 身份管理系统中个人信息的分布

如图 3-3 所示，个人信息主体在使用网络应用时，个人信息主要会被个人信息控制者及其他个人信息控制者接触或掌握。个人信息控制者主要为个人信息主体直接使用的 APP、网站、软硬件产品或服务系统背后的主体。其他个人信息控制者为个人信息主体的产品供应方、合作方或管理方。这些主体会直接或间接的涉及个人信息管理全生命周期过程中。所以，个人信息控制者及其他个人信息控制者均应按照个人信息处理的相关法律规范，合规处理涉及到的所有个人信息，对全生命周期采取保护措施。

个人信息控制者的身份管理系统是最直接的接触到个人信息的。在身份管理组件中，身份核验与用户注册阶段涉及个人信息的收集、使用、更新、传输、存储；在凭证管理阶段涉及个人信息的使用、传输、存储和删除；身份鉴别组件主要涉及个人信息的收集、传输、存储、使用；授权组件主要涉及身份数据的使用、存储、使用等。

身份管理系统中各个组件内部均包括对个人信息的保护需求，每个组件之间的通信也涉及到对相关个人身份凭证的保护，本节将具体介绍，每个组件中涉及到的个人信息，以及安全需求。

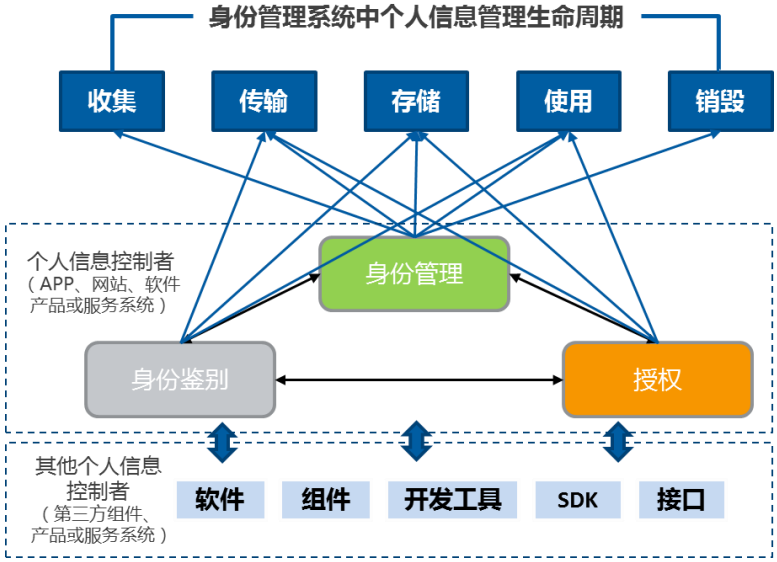

图 3-3 身份管理系统中涉及的个人信息数据生命周期

3.3.1 身份管理组件中的个人信息

3.3.1.1 身份管理组件概述

身份管理组件主要包括身份核验、用户注册、凭证生成和持续监控等功能，基本功能流程如下图所示。身份管理的目的是定义逻辑上与物理实体相关联的数字身份。物理实体基于申请者的身份证明文件，例如身份证、护照或许可证等进行核验后注册。数字身份是为在系统上建立存在而产生的人工产物。这个数字身份，一旦通过鉴别，授权组件就允许或限制该数字身份的各项权限。同时，身份管理组件还要对管理中的身份进行持续监控，定期更新属性信息。

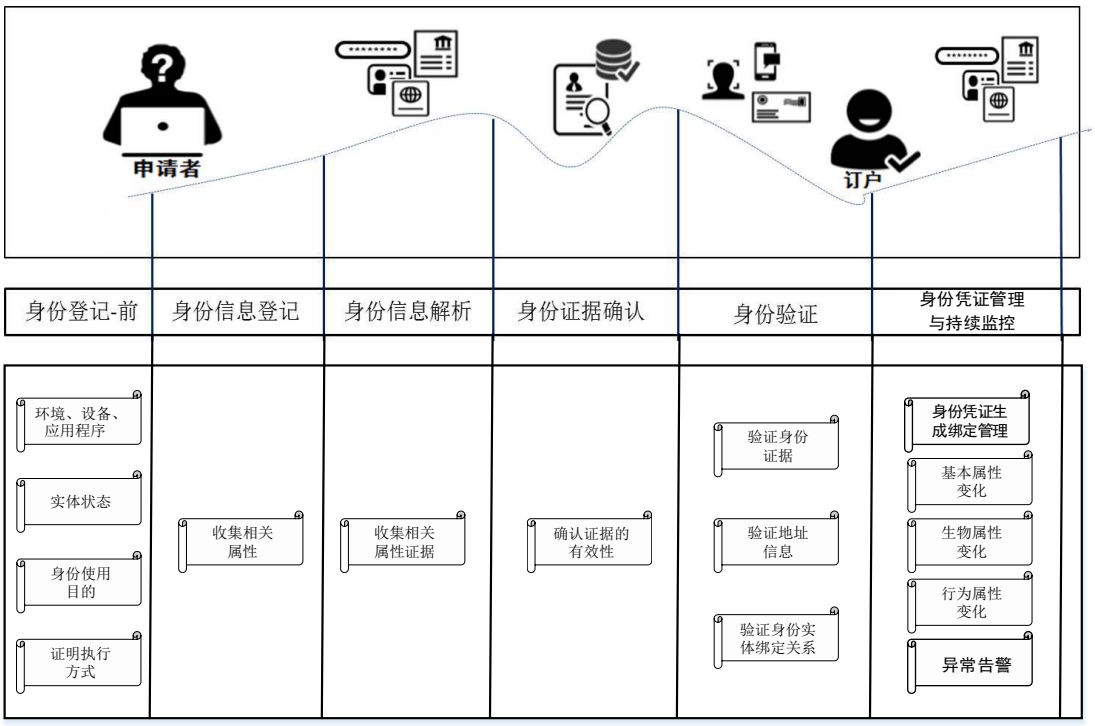

图 3-4 身份管理组件流程

对物理实体身份的信任保证通常与第三方文档的数量和质量有关，而对数字实体的身份的信任保证则与所使用的身份核验的强度和要访问的资源的保护级别有关。设计和维护身份管理系统时，对身份和资源的信任保证均需要考虑。

身份管理组件需要为申请者设置充分的身份证明的要求。一旦身份管理组件确信该申请者具有足够的信息，它将为该申请者创建一个数字身份并注册为具有某种权限的用户，指示系统提供访问权限或可访问的资源位置和方式。身份管理组件完成用户注册后，可以向用户颁发凭证，该凭证允许访问已授权的任何系统。

除了身份外，身份管理组件还须与身份鉴别和授权组件进行通信，以实施数字身份的权利。身份管理组件和身份鉴别之间的通信至少应支持请求许可，撤销和请求确认。如果代表实体的硬件，软件或过程是由身份管理组件授权提供的，则必须在身份管理组件和身份鉴别组件之间协商参数以启用或更新凭证用法。在某些情况下，对于独立的多因素身份鉴别机制，必须同时管理多个身份鉴别凭证，一般由身份管理组件实施。

身份管理还可以直接与授权提供者进行通信，以管理访问控制参数。随着技术变得越来越复杂，可以预见，信任级别可能取决于身份鉴别机制的类型和数量，这可能会导致动态信任级别。这些信任级别和所产生的授权必须通过身份管理组件的管理之后传达给授权提供者。

3.3.1.2 身份管理过程

注册阶段是在系统中注册实体人、设备、服务等信息，收集个人信息的过程。在该过程中身份信息登记过程中会收集由用户提供，具有第三方证明效力的身份证件、用户属性等。

- 身份登记

首先，身份管理方根据申请者的身份用途，告知申请者应提交那些身份属性

信息。申请者根据身份服务提供方的要求，提供所需要的身份属性信息。这些属性信息可能包括：基本属性（姓名、生日、出生地、性别、住址、头像等）、法定属性（身份证、护照、出生证明等）、通信属性（手机号码、固话、邮箱、通信地址等）、社会属性（亲属关系、朋友关系等）、经济属性（收入、资产、银行流水、股票等）、生物属性（人脸、指纹、虹膜等）、环境属性（登录环境）、生理属性（身高、体重、健康状态等）、心理属性（偏好）、行为属性（消费记录、操作习惯、访问记录等）等。

- 身份解析

其次，身份管理方会进行身份信息解析，解析的目的是能够从给定的人群或上下文中唯一区分实体。有效的身份解析方式应使用最小集合的属性来区分实体，这些最小集合的属性是必要的、且能够唯一区分实体。实际上，这是整个身份证明和验证环节的起始点，在该阶段可在一定程度上发现潜在的身份信息伪造等行为。

例如，实时拍摄个人头像与身份证的照片进行比对。在身份信息难以提取的情况下，身份服务提供方也应采用合适的匹配算法，并收集多种格式的身份证据、权威机构的记录或者第三方记录，来进行比对以查找差异。

- 身份证据确认

对申请者提供的身份属性的证据进行真实性、完整性、有效性和准确性验证。首先，检查身份证据的真伪，例如，使用身份证验证器验证身份证本身是否是真的；其次，需要确定身份属性证据是否有足够高质量的信息，以确定所满足的可信程度；随后，确认身份证据中包含的数据是否有效、是否在当前状态合理、是否与该主体真实情况相关。

- 身份验证

验证身份证据。如确认通信属性，可通过发送邮件验证码的方式，确认该邮件地址的有效性，并且确认该邮箱确实被主体所拥有；从公安机关查询该身份证是否被挂失；向公安机关查询该申请实体是否有违法违纪行为记录；验证盖章的银行流水记录，核实银行是否真的发放过该证明文件；向中央银行查询该申请者的信用记录等方式。

执行身份与实体的绑定关系验证。例如，通过询问其结婚日期、教育背景、社会关系等问题，再一次确认所申请的身份和申请者是一致的。

- 身份凭证管理

通过身份验证后，申请者成为身份管理系统的合法用户（订户），身份管理系统为该订户生成、绑定、颁发身份凭证。并对身份凭证的全生命周期进行管理，包括：生成、使用、更新、损坏、丢失、到期、撤销等环节。

- 持续监控

持续的监控是有效的身份证明与验证程序的必要元素。申请者注册身份之后，需要持续地对该身份相关信息进行监控，以识别任何可影响身份可信度的发生变化的信息，从而实时地根据策略调整身份可信度对应的身份级别。

3.3.1.3 身份管理组件的个人信息分布

从上述身份管理组件的主要功能中涉及到个人信息如下：

- 身份注册过程：涉及到个人信息的收集过程、存储过程、删除过程；
- 身份凭证管理过程：涉及到个人信息的使用过程、存储过程、传输过程；
- 身份信息持续监控过程：涉及到个人信息的收集过程、存储过程、传输

过程。

上述收集的个人信息在身份管理系统中，以电子标识符的形式存储，按照个人信息的应用特点，可以划分为直接标识符、准标识符、其他敏感信息、身份凭证等。

- 直接标识符：个人信息中的属性，在特定环境下可以单独识别个人信息主体。例如：姓名、地址、电子邮件地址、电话号码、传真号码、信用卡号码、车牌号码、车辆识别号码、社会保险号码、健康卡号码、病历号码、设备标识符、生物识别码、互联网协议（IP）地址号和网络通用资源定位符（URL）等。
- 准标识符：个人信息中的一个或多个属性，结合其它属性可唯一识别个人信息主体。比如：性别、出生日期或年龄、事件日期（例如入院、手术、出院、访问）、地点（例如邮政编码、建筑名称、地区）、族裔血统、出生国、语言、原住民身份、可见的少数民族地位、职业、婚姻状况、受教育水平、上学年限、犯罪历史、总收入和宗教信仰等。
- 其他敏感信息：一旦泄露、非法提供或滥用可能危害人身和财产安全，极易导致个人名誉、身心健康受到损害或歧视性待遇等的个人信息，包括但不限于身份的鉴别记录（使用的服务类型、服务名称、服务内容、时间、环境信息等）、行为信息（行踪轨迹、交易记录、操作习惯）等。
- 身份凭证：为确定实体所声称的身份而提供的数据，包括但不限于帐户名口令、动态口令、带外验证信息、数字证书、生物识别信息摘要以及在生成凭证过程中使用的秘密信息等。

身份管理过程中涉及到这些携带个人信息的数据的收集过程、传输过程、使用过程、存储过程和删除过程。

3.3.2 鉴别组件中的个人信息

3.3.2.1 鉴别组件概述

身份鉴别是在信息系统中确认通信方的身份及其附属信息的行为过程。利用身份鉴别技术，确定通信方的身份，就能够在此基础上完成其它各种安全服务，包括访问控制、授权、安全审计、数据分发、人员管理等等。在几乎所有的信息系统中，身份鉴别功能都是必须的功能。

鉴别的目的是通过代表实体的硬件，软件或过程来确认数字身份。身份的表示由身份管理组件定义，并与必要的信息一起传达给负责身份鉴别的组件。在成功地对代表实体的硬件，软件或过程完成鉴别后，身份鉴别组件将向授权组件传达确认或拒绝以执行访问控制策略。

身份鉴别需接受身份管理组件的管理。身份管理组件通过提供硬件，软件或过程给身份鉴别组件或请求身份鉴别组件提供硬件，软件或过程来进行管理。

当超过失败尝试阈值时，身份鉴别组件可能不允许进一步的身份鉴别请求。当实体未通过身份鉴别时，身份鉴别组件将决定实体是否必须通过不同的（通常是独立的）过程进行身份鉴别。在每次最终锁定之前，鉴别机制可能会增加每次尝试失败的等待时间。

鉴别组件还需要与授权组件进行通信。虽然访问监督通常由身份管理组件或授权组件来管理，但通常会通过身份鉴别组件将成功或失败的指示提供给授权机制。如果使用多因素身份鉴别，则根据身份鉴别，身份管理和授权的复杂程度，可以分别报告每种机制的结果或将其报告为单个结果。在某些情况下，诸如位置

之类的属性也可以传递给授权组件。

当前的身份鉴别强度取决于具体所用机制的类型：生物识别技术取决于低误报率；密码取决于对密钥不成功的猜测；PKI 的实现依赖于强大的公钥和私钥。然而，不同的身份鉴别机制强度取决于不同的环境因素，因此对于每一个鉴别组件，身份鉴别机制的选择不是最强大或可用性最高的问题。目前对于身份鉴别机制的相对可信度保障尚无方法来比较。

在用户控制下的硬件、软件、生物特征或知识通常称为身份凭证。身份凭证可以采用多种不同的形式，具体取决于身份鉴别过程和所使用的机制。尽管凭证类型不直接决定鉴别强度，但是不同形式的凭证组合可用来增加对鉴别过程的信任程度。

3.3.2.2 鉴别过程

按照鉴别的方向来划分，鉴别过程可分为单向鉴别和双向鉴别，如果面对复杂的应用场景，还需要多层鉴别。

- 单向鉴别

当只有一方需要建立信任关系时（例如，当用户或管理员登录到独立工作站时），将使用单向身份验证。当用户在工作站上拥有一个帐户时，该用户必须提供一组与系统上已设置的帐户相匹配的凭证。用户不对该机器进行鉴别。但是，机器需要确认用户的凭证。

在基于 Web 的系统中，用户无法保证他们已到达正确的机器。在这种情况下，用户不会登录，但是可以使用基于 PKI TLS 的解决方案或类似方法来验证服务器。单向身份鉴别通过验证网站所使用的证书来验证服务器（由浏览器上的图标指示通过鉴别），然后协商安全功能。不需要用户登录即可保持连接，因此服务器不对用户进行鉴别。

- 双向鉴别

双向鉴别通常用于对话中两个实体的相互鉴别。双向鉴别通常使用不同的身份鉴别方法。例如，如果购物者希望从网店购买商品，则需要在两个方向上建立信任。购物者使用帐户名口令作为凭证进行登录，用户则通过验证网站所使用的证书对商店进行鉴别。

双向鉴别也可以使用相同的身份鉴别方法。例如，当员工从公司网络外部访问服务时，企业需要更强的身份鉴别。在这种情况下，他们可能会使用双向 TLS 会话，由于用户获得了由相同或公认的证书颁发机构颁发的证书，因此通常认为该会话具有更高的保证。用户和服务器都具有有效的证书，因此它们可以通过类似 TLS 协议的身份进行相互身份验证。

- 多级身份鉴别

多级身份鉴别是通过单向信任关系和双向信任关系的组合来实现的。例如，使用信用卡进行在线购物会带来一系列非常复杂的多级鉴别，如图 3-5 所示。

该过程具有三种不同身份验证类型的三种信任关系。通常是服务器在用户产生购买行为时使用服务器证书证明自己身份。浏览器通过检查在线商店有效凭证来支持用户对商店的鉴别。但是，商店不知道谁在浏览，除非用户使用某些凭据（例如用户名和口令）登录。除用户外，每个实体中都使用 PKI 证书进行身份验证。为了在网站上进行购买，用户可以使用用户名和口令或用于存储用户信息的类似机制登录，从而增强了用户的便利性并为店主提供了额外的保证。用户的信用卡信息既可以作为用户信息的一部分，也可以作为一种身份验证机制，用于

将钱从用户的帐户转移到商家的帐户。此过程采用多级身份鉴别来管理和验证帐户和费用（包括信用卡票据交换所与商户银行和购物者的银行之间的双向鉴别）。

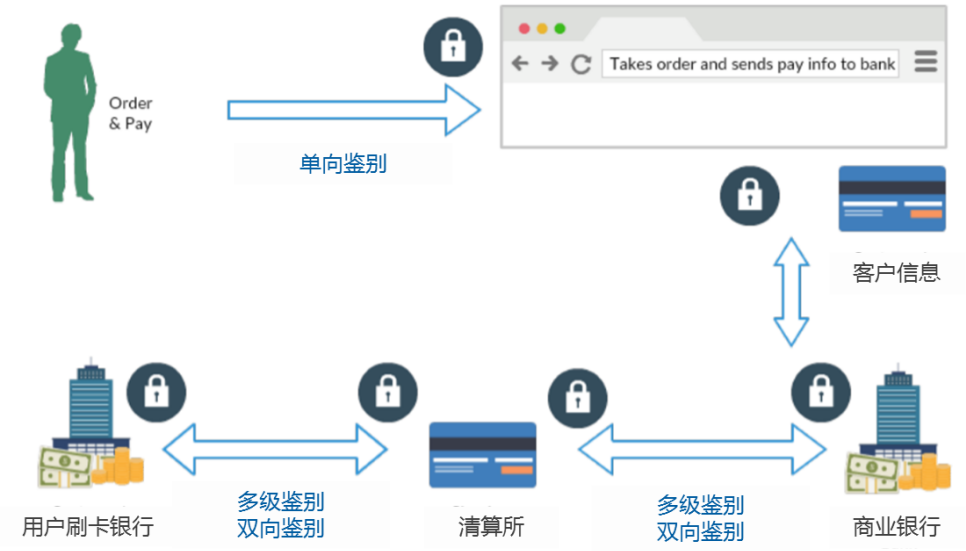

图 3-5 复杂鉴别场景

3.3.2.3 鉴别过程中的个人信息

鉴别过程涉及的个人信息主要包括身份凭证（包括代表实体的硬件，软件或过程）和鉴别结果（令牌、断言）两种，其中身份凭证包括以下几类：

实体所知：即利用实体和鉴别组件所共享的秘密信息进行身份鉴别。例如，口令，个人识别码（PIN），图片和音频等，“记忆秘密”尽管存在多种不同形式，但它们全部仅基于与身份鉴别组件共享的秘密信息来证明实体。许多研究都呼吁取消口令，但是口令仍然是最常用的身份鉴别形式，并且经常被人们喜欢作为附加或替代形式，例如解锁智能卡或作为身份鉴别的备用手段。

实体所有：通常指鉴别器，鉴别器由身份管理服务方提供，其中记录实体的若干身份属性值和身份凭证。

实体特性：通常是指基于生物特征的身份鉴别。常见示例包括指纹，面部，虹膜和语音识别。初始身份核验中使用的生物识别技术可以确定可信度，对生物特征的主动扫描是为了确认与之前收集的生物识别数据来自同一自然人，对生物特征的连续扫描确定合法用户可持续使用该系统。

多因素身份鉴别：为缓解身份鉴别过程中的风险，可使用多种形式的身份鉴别方式。

所以，鉴别过程涉及的一类个人信息为身份凭证信息，包括实体所知道的秘密信息，如口令、PIN 码等；实体所拥有的，如口令卡、Ukey，及其中的密钥、证书、身份属性等信息；实体的特性，即基于生物识别信息生成的摘要信息；多因素鉴别中涉及的则是上述身份凭证的多项组合。

另一类个人信息为鉴别响应中包含的个人信息。鉴别完成后生成的令牌、断言等，其中包含的个人信息，如鉴别结果、主体标识符、令牌接收方、协商的鉴别方式、有效时间等自身属性及其他可能包含的声明等。未通过鉴别则可返回状态参数和错误相应参数，不应返回包含个人信息的其他参数。

在鉴别过程中涉及到这些携带个人信息的数据收集过程、传输过程、使用过

程、存储过程。

3.3.3 授权组件中的个人信息

3.3.3.1 授权组件概述

授权组件的核心功能即访问控制。访问控制是一种保证数据处理系统的资源只能由被授权主体按授权方式进行访问的手段。授权组件可以按用户身份及其所归属的某个定义组来限制用户对某些资源的访问，或限制对某些控制功能的使用。

在从身份鉴别组件收到鉴别成功的报告后，授权组件将允许数字实体访问以执行程序或操纵信息。通常按照不同的授权粒度提供权限，例如只读，执行权限或允许实体编辑信息。访问控制可通过基于角色的访问控制（RBAC）和基于属性的访问控制（ABAC）等技术实现解决。

随着标识、凭证和访问控制技术的发展，目前国际业界已涌现出一系列标准，用于不同网络身份服务系统的互联互通，以及跨域进行访问授权。

组件之间主要传递是否允许数字身份的访问结果。未来可能会出现多个身份鉴别信任级别，并可能给授权技术和权限管理提出更高的要求。

3.3.3.2 授权过程

访问控制是网络安全防范和保护的重要手段，它的主要任务是维护网络系统安全、保证网络资源不被非法使用和非常规访问。通常在技术实现上，包括以下几部分。

- 接入访问控制：接入访问控制为网络访问提供了第一层访问控制，是网络访问的最先屏障，它控制哪些用户能够登录到服务器并获取网络资源，控制准许用户入网的时间和准许他们在哪台工作站入网。例如，互联网服务提供方（ISP）实现的就是接入服务。用户的接入访问控制是对合法用户的鉴别，通常使用用户名和口令的认证方式。一般可分为三个步骤：用户名的识别与验证、用户口令的识别与验证和用户账号的缺省限制检查。
- 资源访问控制：是对客体整体资源信息的访问控制管理。其中包括文件系统的访问控制（文件目录访问控制和系统访问控制）、文件属性访问控制、信息内容访问控制。文件目录访问控制是指用户和用户组被赋予一定的权限，在权限的规则控制许可下，哪些用户和用户组可以访问哪些目录、子目录、文件和其他资源，哪些用户可以对其中的哪些文件、目录、子目录、设备等能够执行何种操作。系统访问控制是指一个网络系统管理员应当为用户指定适当的访问权限，这些访问权限控制着用户对服务器的访问；同时也应服务器控制台设置口令，以防止非法用户修改、删除重要信息或破坏数据；应设定服务器登录时间限制、非法访问者检测和关闭的时间间隔；应对网络实施监控，记录用户对网络资源的访问，对非法的网络访问，能够用图形或文字或声音等形式报警等。文件属性访问控制：当用文件、目录和网络设备时，应给文件、目录等指定访问属性。属性安全控制可以将给定的属性与要访问的文件、目录和网络设备联系起来。
- 网络端口和节点的访问控制：网络中的节点和端口往往加密传输数据，这些重要位置的管理必须防止黑客发动的攻击。对于管理和修改数据，应该要求访问者提供足以证明身份的硬件鉴别器（如智能卡）。

3.3.3.3 授权过程中的个人信息

当系统完成对用户的鉴别后，系统创建一个访问令牌，里面包含登录进程返回的安全标识符（SID）和由本地安全策略分配给用户或用户的安全组的特权列表。以该用户身份运行的所有进程都拥有该令牌的一个拷贝。系统使用访问令牌控制哪些依赖方可以访问哪些用户个人信息，并控制系统进程执行相关操作的能力。

授权过程中涉及到两种需要保护的信息，分别是访问令牌和安全描述符，访问令牌是欲进行访问的进程使用的表明自己身份和特权的信息数据，是身份服务提供方发送给依赖方用于访问用户身份信息的凭据，代表着用户的授权。访问令牌中给出了用户信息的访问范围和访问有效期，访问范围和访问有效期由用户授权同意。访问令牌可作为提取授权信息的标识符，也可自包含授权信息，访问令牌中包含的授权信息可通过数字签名的方式进行验证。

安全描述符是欲被访问的安全对象的相关安全信息，包括谁拥有该对象，以何种方式访问以及何种审查访问类型等信息。如什么样的用户的什么访问请求可以被允许，什么样的用户或者组的什么访问要被拒绝。

在授权过程中涉及到对访问令牌的存储过程、传输过程、使用过程和删除（失效）过程、安全描述符的存储过程、使用过程和删除过程等的保护。

3.3.4 身份管理中个人信息分布总结

下表梳理了身份管理系统中，各个组件提供身份服务时内部及交互过程中涉及到的个人信息，以及在这些个人信息所涉及的生命周期过程。

表 3-1 身份管理中个人信息分布

身份管理系统组件	涉及个人信息	个人信息生命周期				
		收集	传输	存储	使用	删除
身份管理组件	**直接标识符：**个人信息中的属性，在特定环境下可以单独识别个人信息主体。可包括但不限于：姓名、地址、电子邮件地址、电话号码、传真号码、信用卡号码、车牌号码、车辆识别号码、社会保险号码、健康卡号码、病历号码、设备标识符、生物识别信息摘要、互联网协议（IP）地址号和网络通用资源定位符（URL）等；	●	●	●	●	●
	准标识符：个人信息中的一个或多个属性。属性可包括但不限于：性别、出生日期或年龄、婚姻状况、工作经历、教育背景、财产信息、收入情况、征信信息、健康属性（身高、体重、病史等）、环境属性（IP、地点、终端等）、事件日期（例如入院、手术、出院、访问）、地点（例如出生国、建筑名称、地区、住宿信息、族裔血统、语言、犯罪历史和宗教信仰等。	●	●	●	●	●

身份管理系统组件	涉及个人信息	个人信息生命周期				
		收集	传输	存储	使用	删除
	其他敏感信息：一旦泄露、非法提供或滥用可能危害人身和财产安全，极易导致个人名誉、身心健康受到损害或歧视性待遇等的个人信息，包括但不限于身份的鉴别记录（使用的服务类型、服务名称、服务内容、时间、环境信息等）、行为信息（行踪轨迹、交易记录、操作习惯）等；	●	●	●	●	●
	身份凭证-由身份管理方生成绑定：为确定实体所声称的身份而提供的数据，包括但不限于动态口令、数字证书以及在生成凭证过程中使用的秘密信息等；	○	●	●	●	●
	身份凭证-由用户定义并注册：包括但不限于帐户名口令、用户已知的知识、带外验证信息等；	●	●	●	●	●
	身份凭证-生物信息：包括但不限于指纹、虹膜、掌纹等。这些信息是用户所拥有的反映生物特征的信息。一旦被盗用则无法被替换，需受到严格保护，一般不存储生物识别信息，仅存储生物识别信息摘要（作为直接标识符进行保护）。	●	○	○	●	●
鉴别组件	**身份凭证：**为确定实体所声称的身份而提供的数据，包括但不限于帐户名口令、动态口令、带外验证信息、数字证书、生物识别信息摘要以及在生成凭证过程中使用的秘密信息等；	●	●	●	●	●
	鉴别响应：通过鉴别后生成的令牌、断言等，令牌或断言中包含的个人信息，如鉴别结果、主体标识符、令牌接收方、协商的鉴别方式、有效时间等自身属性及其他可能包含的声明等。未通过鉴别则可返回状态参数和错误相应参数，不应返回包含个人信息的其他参数。	○	●	●	●	●
	鉴别记录：鉴别过程日志，管理日志等，可能包含用户使用的服务类型、服务名称、服务内容、时间、环境信息等。	○	●	●	●	●
授权组件	**访问令牌：**用于访问用户身份信息的凭据，代表着用户的授权。其内容包含用户个人信息的访问范围和访问有效期，访问范围和访问有效期由用户授权同意。	○	●	●	●	●
	用户个人信息的安全描述符：安全描述符包括用户自主决定的访问控制表，系统访问控制列表，访问请求的日志和拥有该资源（个人信息）的主体。	○	○	●	●	●

注：●为涉及该环节，○为不涉及该环节

3.4 身份管理系统安全需求

3.4.1 安全收集

《个人信息安全规范》在数据生命周期第一步个人信息收集环节，严格界定了个人信息控制者的权利和义务，规定在收集个人信息前，应当向信息主体明示相关内容并取得同意；涉及间接方式获取个人信息时，应要求个人信息提供方说明来源，并确认合法性。

收集过程中的安全需求如下。

（1）合规性：应遵守《个人信息保护法》、《个人信息安全规范》等相关法律标准条款要求，例如：

- 合法性，要求个人信息控制者在法律法规规定的范围内采用合法的手段和获取信息的渠道，在征得个人信息主体同意的前提下收集个人信息或要求信息主体提供个人信息；
- 最小化，要求个人信息的收集类型、频率和数量应在必要性的最小要求之内，即符合最少够用原则。在能达到所需目的条件下，只处理最少的个人信息类型和数量；
- 授权同意，要求个人信息控制者处理个人信息时的目的、方式、范围以及相关规则，均要经过个人信息主体的授权同意；
- 制定个人信息保护政策，实践中通常的表现形式为制定隐私政策。并采取技术措施，如弹窗、URL 等明显的方式，引导用户查阅隐私政策，获得明示同意后，开展个人信息收集活动。

（2）明确收集方案：按照主动收集、第三方采购和爬取等手段收集用户个人信息的方式分别涉及合规的收集方案，收集前审批，收集中监管，收集后确权，形成收集管理闭环。

（3）收集接口安全需求：应提供规范、安全的访问接口用于信息收集，采用技术手段防止被篡改、重放攻击、DOS 攻击等，如对 URL 进行签名，防止请求参数被篡改，加时间戳或签名的方式防止重放攻击等。

（4）对输入信息进行安全防护：采取技术手段对个人信息的输入进行安全保护，如：对输入信息合法性校验、数据非明文显示等；可采用本地差分隐私技术，对用户属性进行脱敏后收集。

（5）用户个人信息的验证能力：能够追溯个人信息属性及证据的真实有效，确保收集个人信息来源的可追溯性，并采用技术手段防抵赖。

3.4.2 安全传输

个人信息在终端设备、信息系统或系统组件之间传递的过程中，各参与方应保证个人信息及其凭证在传输过程中的安全性、完整性和可用性，具体技术要求如下：

（1）合规性：建立相应的个人信息传输安全策略和规程；采取有效措施，对安全策略进行审核、监控和优化，包括对通道的安全配置、密码算法配置、密钥管理等保护措施；

（2）传输前鉴别：传输个人信息之前，通信双方应通过有效技术手段进行身份鉴别，并通过权限管理；

（3）传输中提供安全措施：采用满足个人信息传输安全策略的安全控制措施，如安全通道、数据加密等技术措施；

（4）选择安全的传输协议：通过公共网络传输时，个人信息应使用加密通道或将传输数据进行加密，保障个人信息传输过程的安全，如使用 HTTPS 作为

传输协议；

（5）传输服务可用性：应采取有效措施，保证数据传输可靠性和网络传输服务可用性，如传输链路冗余等；

（6）传输后，提供完整性校验机制，保障数据的可用性和完整性。

3.4.3 安全存储

存储系统作为个人信息的保存空间，是数据保护的关键一道防线；随着存储系统由本地直连向着网络化和分布式的方向发展，并被网络上的众多计算机共享，使存储系统变得更易受到攻击，相对静态的存储系统往往成为攻击者的首选目标，达到窃取、篡改或破坏数据的目的。存储安全变得至关重要，安全存储主要包括存储安全技术、重复数据删除技术、数据备份及灾难恢复技术等。

在个人信息存储方面具有以下安全需求。

（1）合规性：应依据我国法律法规及上级主管部门的管理措施，建立健全个人信息存储过程中的流程规范与制度保障措施。

（2）时间最小化：要求信息的保存时间应是使用目的所需的最短时间；在超过保存期限后，即应对信息进行删除或匿名化处理。

（3）去标识化：可对信息主体采用脱敏技术进行保护，要求将收集到的信息去除个人信息主体特征，确保该数据后续不能对应到特定实体。

（4）采用技术防护措施对存储的个人信息进行保护：存放个人信息的系统需要对存储个人信息采取相应的安全措施，如加密、审计、脱敏等；对不同类别、不同等级的个人信息采用差异化安全存储；如采用差异化脱敏技术存储、不同加密算法、不同密钥长度等方法，并做好加密算法、脱敏方法的保密。

对于生物识别类信息，应采用技术措施处理后再行存储。该处理方式与去标识化处理方式存在差异，去标识化处理主要是针对信息的个体特征处理，使该信息失去独立识别信息主体的能力，而此处涉及的技术措施是对生物识别信息通过加密技术手段存储或只存储该信息的摘要部分等方法对生物特征信息本身进行保护。

（5）隔离存储：应将去标识化、匿名化后的数据与可用于恢复识别实体对象的个人信息采取逻辑隔离的方式进行存储，确保脱敏后的信息与个人信息不被混用。

（6）密钥管理机制：应提供稳定成熟的密钥管理措施，支持不同策略更新秘密信息、适时回收密钥信息。

（7）设置访问控制策略，定期实施安全风险评估，配置安全基线、部署必要的安全存储技术手段等。

（8）建立完备的数据存储容灾备份和恢复机制：个人信息存储系统应提供完备的数据备份和恢复机制来保障数据的可用性和完整性，一旦发生数据丢失或破坏，可以利用备份来恢复数据，从而保证在故障发生后数据不丢失。应加强数据备份介质的管理，对各类介质进行访问控制和保护。

（9）环境与管理要求：对存储个人信息的设备及基础设施重点做好安全防护，包括落实个人信息存储设备的操作终端安全管控措施及接入机制；根据所承载的个人信息的重要程度对介质实行分类和标识，并实行存储环境专人管理。

3.4.4 安全使用

随着信息处理技术日臻成熟，越来越多的数据操作行为，包括分析、打标签、

用户画像、数据分享等行为均被确立为数据使用行为。在《个人信息安全规范》中，针对个人信息控制者对信息的使用要求也做出了进一步细化：个人信息的访问控制、展示限制、使用限制等。

个人信息的访问遵循最小授权原则，这与收集阶段的信息最少够用和存储阶段的信息保存时间最小化相呼应，强调个人信息处理环节每项操作的必要性。通过设置内部审批流程，做到有源可溯；设置角色分离，控制权限；对超权限处理信息人员记录在册；对敏感信息采用触发授权等方式严格限制访问个人信息的人员范围。通过建立有效的控制机制来切实保护个人信息在使用时的安全性。

针对个人信息主体所提出的请求，《个人信息安全规范》中规定对个人信息控制者应提供相应的服务；其中包括：个人信息的访问、更正、删除、撤回同意、注销帐户以及获取个人信息副本。

所以针对个人信息使用过程中的安全需求包括：

（1）**合规性：**应遵守《个人信息保护法》、《个人信息安全规范》等相关法律标准条款要求，例如：

最小化原则：对被授权访问个人信息的内部数据操作人员，应按照最小授权的原则，使其只能访问职责所需的最少够用的个人信息，且仅具备完成职责所需的最少的数据操作权限。

知情权：针对在使用时超出收集所声称的用户明示同意范围时，需要在此征得个人信息主体的明示同意。

（2）严格访问控制管理：设置角色分离，控制权限；对个人敏感信息的访问、修改等行为，宜在对角色的权限控制的基础上，根据业务流程的需求触发操作授权。例如，因收到客户投诉，投诉处理人员才可访问该用户的相关信息。

（3）去标识化：针对个人信息的展示与共享，需要通过数据脱敏等技术手段对数据进行去标识化处理，避免产生的信息可单独识别个人身份，降低信息泄漏的风险。

（4）有限制的使用：系统不应具备开放式查询能力，应严格限制批量下载，查询等功能。应采取多因素鉴别对导出信息操作人员进行身份鉴别。

（5）缓存要求：服务过程中的个人信息只缓存在受保护的内存空间内（Security Barrier），而不存在于任何持久化存储内。

（6）共享与转让：应有严格的审查监管制度，对共享方和数据接收方的合规资质进行审查，采取措施保障个人信息的主体权益；应采用去标识化技术对共享和转让数据进行脱敏处理，若无法脱敏时，应进行加密处理。

（7）外部组件安全：使用外部嵌入或接入的自动化工具进行对个人信息的使用时，应定期检查评估工具、组件、通道的安全性和可靠性。

（8）审计要求：系统应对所有操作进行细粒度的授权，记录日志，对操作日志、日志访问与全过程行为进行审计。

3.4.5 安全删除

（1）合规性：应依据我国法律法规及上级主管部门的管理措施，建立健全个人信息删除过程中的流程规范与制度保障措施。

（2）用户主动注销：当个人信息主体要求删除个人信息时，身份管理系统应依据国家法律法规、行业主管部门的有关规定以及与信息主体的约定，及时予以响应，进行注销和删除操作。

（3）系统停止运营时：应根据国家法律法规与行业主管部门的有关规定要

求，对所存储的个人信息进行妥善处置，或移交国家与行业主管部门指定的机构继续保存。

（4）采取技术措施保障存储信息彻底删除：应彻底删除保存于介质的个人信息，不应只采用删除索引、删除文件系统的方式进行信息销毁，应通过多次覆写等方式安全的擦除个人信息。

（5）采取技术措施保障存储介质彻底销毁：存储个人信息的介质如不再使用，应采用不可恢复的方式对介质进行销毁处理，如消磁、焚烧、粉碎等。

4. 个人信息安全保护技术

4.1 涉及身份管理的个人信息安全保护技术框架

本章节进一步梳理身份管理系统应对个人信息的全生命周期安全保护需求，可采用的技术措施。

表 4-1 个人信息安全保护技术

生命周期过程	保护需求	技术手段
安全收集	接口安全：应提供规范、安全的访问接口用于信息收集，采用技术手段防止被篡改、重放攻击、DOS 攻击等。	采用时间戳超时机制防御 DOS 攻击； 对 URL 进行签名，防止请求参数被篡改； 加时间戳或签名的方式防止重放攻击。
	对输入信息进行安全防护。	采取技术手段对个人信息的输入进行安全保护，如：对输入信息合法性校验、数据非明文显示等；可采用**本地差分隐私技术**，对用户属性进行脱敏后收集。
	用户个人信息的核验能力：能够追溯个人信息属性及证据的真实有效。	通过权威第三方验证身份证明文件中的身份信息，缓解身份信息伪造、不正确或不完整带来的风险。
	确保收集个人信息来源的可追溯性，并采用技术手段防抵赖。	采用数字证书或电子签章方式防抵赖。
安全传输	传输前鉴别：传输个人信息之前，通信双方应通过有效技术手段进行身份鉴别，并通过权限管理。	使用被正确实施的鉴别协议； 身份凭证的全生命周期保护； **可采用持续身份鉴别技术；** 可采用被正确实施的**访问控制模型**。
	选择安全的传输协议：通过公共网络传输时，个人信息应使用加密通道或将传输数据进行加密，保障个人信息传输过程的安全。	选择安全的传输协议：如 https。
	传输服务可用性：应采取有效措施，保证数据传输可靠性和网络传输服务可用性。	使用传输链路冗余、双机热备等措施保证可用性。
	传输后，提供完整性校验机制，保障数据的可用性和完整	采用完整性校验机制，如：可基于 SM3 算法（GB/T 32905-2016）

生命周期过程	保护需求	技术手段
	性。	生成消息鉴别码对数据完整性进行校验。
安全存储	时间最小化：要求信息的保存时间应是使用目的所需的最短时间。	在超过保存期限后，即应对个人信息（直接标识符、准标识符、敏感信息等）进行删除或去标识化处理。
	去标识化：将收集到的信息去除个人信息主体特征，确保该数据后续不能对应到特定个体。	对于无需关联到特定个体的个人信息，采用去标识化技术（可参考 GB/T 37964—2019）进行保护。
	采用技术防护措施对存储的个人信息进行保护。	对存储的个人信息进行按列或按表加密，并对密钥进行安全可靠管理，如存储至安全的与个人信息物理隔离的介质中； 对于生物识别类信息，应采用技术措施处理后再行存储。对生物识别信息通过加密技术手段存储或只存储该信息的摘要部分等方法对生物特征信息本身进行保护。
	对不同类别、不同等级的数据采用差异化安全存储。	对个人信息进行分类，并可采用不同的脱敏技术、不同加密算法或不同密钥长度等方法进行保护。
	确保去标识化后的信息不被混用。	将去标识化、匿名化后的数据与可用于恢复识别实体对象的个人信息采取逻辑隔离或者物理隔离的方式进行存储。
	密钥管理机制。	提供稳定成熟的密钥管理措施，支持不同策略更新秘密信息、适时回收密钥信息等。
	防止非授权访问，定期实施安全风险评估。	采用被正确实施的访问控制模型； 访问、修改等关键行为记录日志，并进行审计。
安全使用	防止非授权访问，定期实施安全风险评估。	使用被正确实施的鉴别协议； **可采用持续身份鉴别技术；** 可采用被正确实施的访问控制模型； 对个人敏感信息的访问、修改等行为，宜在对角色的权限控制的基础上，根据业务流程（任务）的需求触发操作授权；

生命周期过程	保护需求	技术手段
		访问、修改等关键行为记录日志，并进行审计。
	个人信息的展示。	可采用**差分隐私**等去标识化技术对数据进行去标识化处理，避免产生的信息可单独识别个人身份，降低信息泄漏的风险，平衡可用性与隐私保护的关系。
	个人信息的共享与转让。	可采用**差分隐私**等去标识化技术对共享和转让数据进行去标识化处理； 采用密码技术对共享或转让的数据进行加密。
	外部组件安全。	使用外部嵌入或接入的自动化工具进行对个人信息进行处理时，应定期检查评估工具、组件、通道的安全性和可靠性。
安全删除	采取技术措施保障存储信息彻底删除。	对个人信息进行去标识化处理； 必要时应通过多次覆写等方式安全的擦除个人信息； 应对数据导入导出通道缓存的数据进行及时清除； 采用深度内容检测技术进行内容监测，对个人信息删除的执行结果进行检查，对违规行为及时处置； 需要彻底删除保存于介质的个人信息时，可采用物理手段保证介质不可用。

以个人信息为中心，按照数据生命周期（收集、传输、存储、使用、删除）等过程中用到的个人信息保护技术进行归纳整理，形成涉及身份管理的个人信息安全保护技术框架如下图所示，

本文建议的涉及身份管理的个人信息保护技术框架应依据法律法规和标准体系，采用持续身份鉴别与授权机制，对身份管理系统中的人、设备进行持续性鉴别与动态授权。采用符合标准的网络安全保护措施，保障身份管理系统的系统级安全。以合规密码算法及规范的密钥管理为底层基础，保障个人信息的机密性、完整性、可用性和操作的不可否认性。对不同类别、不同重要程度的个人信息可采用数据脱敏、合规正确的鉴别与访问控制技术、密码技术等对个人信息的收集、使用、共享、销毁等全生命周期过程实施保护。

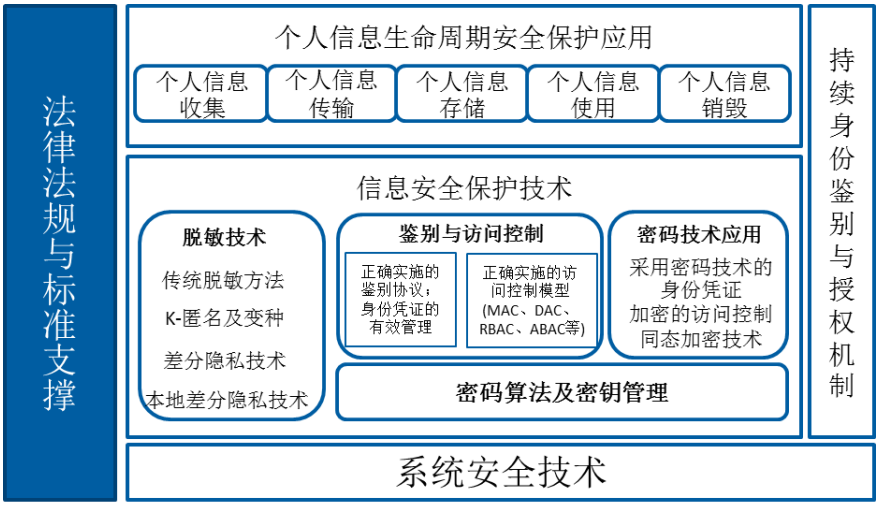

图 4-1 涉及身份管理的个人信息保护技术框架

4.2 身份鉴别技术

4.2.1 持续鉴别与授权技术

2020 年 8 月，美国 NIST 发布了《零信任架构》标准。该标准的目的是通过引入零信任架构策略，对越来越复杂的企业（组织、机构）网络进行安全规划。主要解决了零信任网络如何落地的问题，包括零信任架构包含有哪些设计原则、有哪些组件如何协同工作、如何缓解可能存在于零信任网络中的安全威胁、以及如何进行零信任网络迁移。这在很大程度上说明，零信任作为新兴的安全策略技术架构，不仅在产业界进行了验证推广，也在美国政府层面取得了应用战略地位。

零信任（Zero Trust）架构，其防御边界缩小到单个或更小的资源组，身份管理系统中各环节、各类别，甚至单个实体的个人信息均可以作为单个或更小的资源组，其安全策略是并不根据物理或网络位置对系统授予完全可信的权限，只有当资源确定被需要的时候才授予系统内部人员、程序对个人信息的访问权限，且在连接建立之前会进行鉴别。零信任网络的概念建立在以下5个基本假定：

——网络无时无刻不处于危险的环境中；

——网络中自始至终存在外部或内部威胁；

——网络的位置不足以决定网络的可信程度；

——所有的设备、用户和网络流量都应当经过认证和授权；

——安全策略是动态的，并基于尽可能多的数据源计算而来。

零信任架构的设计和部署遵循以下基本原则：

——所有数据源和计算服务都被视为资源。如果允许个人拥有的设备访问企业拥有的资源，则企业可以决定将其归类为资源；

——网络位置并不意味着信任。换言之，不应对位于企业自有网络基础设施上的设备自动授予任何信任。所有通信应以安全的方式进行（即加密和鉴别）；

——对单个企业资源的访问是基于每个连接授予的。在授予访问权限之前，将对声称者进行鉴别。对一个资源的身份鉴别并不会自动授予对另一个资源的访问权限；

——对资源的访问由策略决定，包括用户身份和请求系统的可观察状态，并且可以包括其他行为属性；

——确保所有拥有的和关联的系统处于尽可能最安全的状态，应建立持续诊断和缓解（CDM）方案，以监测系统状态，并根据需要应用补丁/修复程序；

——在允许访问之前，用户身份鉴别是动态的并且是严格强制的。这是一个不断的访问、扫描和评估威胁、调整和不断验证的循环。根据策略（如基于时间的、请求的新资源、资源修改等）的定义和实施，在用户交互过程中不断进行监视和重新验证，以实现安全性、可用性和成本效率之间的平衡。

4.2.2 基于行为特征的身份鉴别技术

对接触个人信息的用户、管理员、程序进行身份鉴别，是建立健全网络可信身份管理的重要手段。为了应对传统网络身份鉴别可能存在的安全风险高、用户体验差、静态性等缺点，相关组织和研究人员开展了一系列行为特征、认知特征等可用于身份识别的生物学特征的研究。相比于基于传统生物特征（指纹、虹膜等）的身份识别技术，基于行为生物特征的身份优势明显：

——安全风险低，不易被盗用；

——用户体验好，不需用户配合；

——可实现对用户身份的动态连续监测。

4.2.2.1 行为生物特征

针对现有研究现状，根据所收集的用户信息的类型，行为生物特征可分为指定行为和非指定行为，其中指定行为可理解为按照应用要求作出适当的动作，通过比对行为特征一致性以进行身份鉴别，包括但不限于击键规律、鼠标移动轨迹、肢体动作分析等；非指定行为包括但不限于基于文体学的上下文行为、网页访问行为、操作系统交互习惯等。下面介绍几种典型行为特征：（1）击键和鼠标移动轨迹；（2）捕捉眼球运动特征；（3）体态手势识别；（4）文体学（包括基于认知加工时间、认知节奏文体学）；（5）基于用户搜索模式的身份识别；（6）屏幕或操作系统交互行为习惯。

（1） 击键和鼠标移动轨迹

学界对击键动力学和鼠标动力学的研究时间较长。基于击键和鼠标移动轨迹身份识别技术原理：从用户击键规律和鼠标移动轨迹中，捕获用户行为特征，用于连续身份认证。如 Ahmed[26]等利用鼠标移动的速度、移动的距离实现了主机用户身份的认证，Pusara[27]等收集用户在浏览同一网页时的鼠标行为，采用决策树算法对用户的 18 个行为习惯进行了研究。

鼠标运动的主要特征包括:（1）鼠标移动速度和加速度；（2）鼠标水平移动和垂直移动偏移量；（3）鼠标一次移动的持续时间；（4）鼠标点击时间间隔，包括单击和双击时间间隔等。

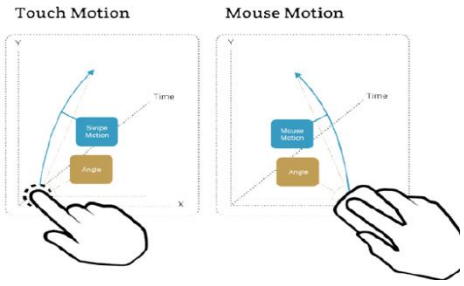

图 4-2 鼠标轨迹分析示意图（触摸屏、鼠标）

基于击键规律的身份认证，利用户敲击键盘所用的时间特性，对用户身份进行认证。击键间隔通常指敲击两个字母的间隔时间。不同的人键盘敲击节奏不同，因此，可以用于身份认证。

常见的击键行为特征有：（1）击键的时间规律统计：击键时间间隔（释放一个键与按下一个键之间的时间），击键持续时间（即压键时间，按下一个键在释放的时间），击键输入速度，击键错误频率（即使用退格键的频率）等，按下一个键与释放该键的但持续时间不同；（2）击键习惯：使用键盘上额外按键的习惯，如输入数字时使用数字键盘；大写字母的输入方式、使用 CapsLock 键还是 Shift 键+字母键等。

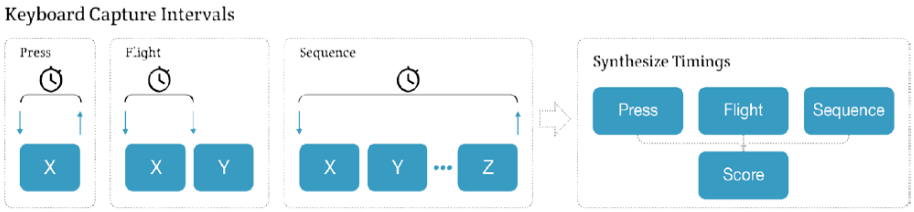

图 4-3 击键规律分析示意图

该方法最大的问题是用户内在可变性，即用户本身击键间隔是可变的。为了解决这个问题，很多研究人员给出了解决方法：研究变化的混乱程度法。同一个用户，尽管敲击两字母的时间可变，但连续敲击两个字母的时间排序的趋势是一致的，利用排序的距离，可判定两种击键模式是否相似。

（2） 眼球的运动

Sluganovia 等人[29]基于眼球运动可以被反射和可预测的出发，开发了交互式的视觉刺激界面来激发用户相应的眼球运动。接触眼睛追踪传感器精准定位眼球运动轨迹，在几秒内提取出眼球行为的生物特征用于鉴别不同该用户，每次鉴别的刺激不同，可达到防止重放攻击的目的。

（3） 步态与手势识别

步态识别一般原理是通过摄像机采集一个人的走路视频，对其进行检测、分割，提取出人在行进中的动态及静态特征，并转化成数字编码，然后和数据库中存储的数据进行比对，得出身份识别的结果。其关键技术包括步态检测、步态分割、步态特征提取与比对[30-31]。

手势识别分为两大类，分别是静态手势、动态手势。手势识别是将输入的特征向量预处理后和建立的模型进行匹配，如果当前手势和模型的相似度高则判断为真用户，否则判断为假用户。本质是通过模型对手势进行分类，常用手势识别方法包括：（1）模板匹配，其原理是将输入的动态手势与训练好的手势模型进行

比对，通过一些衡量标准来计算二者之间的相似度，采用动态事件规划算法，通过测量两个特征事件序列之间的差异，或者预测待测特征时间序列与模板特征时间序列之间的相似程度，通过模板匹配完成身份认证［32］；（2）神经网络，能够解决复杂的非线性问题，如采用反向传播神经网络模型与粒子群优化的径向函数网络模型进行特征训练与用户识别[33]；（3）机器学习，如支持向量机算法，通过核映射方法转化到高维空间解决线性数据可分问题，完成数据模型训练及认证[34]。

（4） 文体学（基于上下文、认知加工、认知节奏）

文体学（Forensical stylometry）是对语言风格的研究，在用户正常构建一篇文档（例如邮件、备忘录）的同时，分析用户上下文行为，从而总结出反映其身份的规律。

常用文体学特征包括：文本的类符/形符比（Type-token ratio）统计特征，不同单词数/总数；文本的语义和语法分析统计特征；平均词汇、句子、段落长度独特词汇的使用标点符号的使用方法等的统计；反映用户的文化程度、专业领域和个性偏好的词；反映认知节奏的文体学特征。自然停顿的频率和时间；用户产生文本的速度；停顿频率和时间；文本修订模式、粘贴模式。

（5） 基于用户搜索模式的身份识别

该技术重点研究基于用户搜索模式的身份识别，部署诱饵文件，由于真实用户不会接触这些文件，攻击者却不知道，通过检测操作者是否搜索诱饵文件，判定是否存在数据窃取的攻击行为，识别是否为真实用户。可用该方法检测是否存在攻击者信息搜集行为。

基于用户搜索模式的身份识别技术的研究已经取得一些成果，包括研发了诱饵分布工具创建诱饵文件；对用户搜索行为进行建模；通过识别与搜索和信息访问活动相关联的操作，构建用户模型；外国语言诱饵文件优势；提出了使用“诱饵”技术保护云中的数据，通过监视云中的数据访问并检测异常数据访问模式。该技术的关键是诱饵文件的部署，诱饵文件应具备如下属性：

——可信性：诱饵文件首要性能之一是具有可信性。诱饵文件应尽可能模仿真实文件的形式和内容，使得诱饵文件看起来像真实的文件；

——吸引性：诱饵文件包含攻击者感兴趣的内容。例如面对财务攻击，诱饵文件中尽可能包含和财务有关的数据；

——显著性：将文件放置于明显的位置，例如用户桌面；

——可检测性：前述三个性质关注诱饵和攻击者之间的关系，可检测性侧重文件被访问时，通知文件拥有者的能力。一个理想的诱饵应满足文件一旦被访问，诱饵系统能发出警报；

——可变性：文件池中的诱饵文件之间具有一定的差异；

——隐秘性：诱饵文件被检索发出警报要保持隐秘性，以免被攻击者发现；

——非干扰性：诱饵文件的部署不能影响操作者的习惯；

——可辨别性：真实用户容易辨别真实文件和诱饵文件；

——有效期：攻击者往往对更新的数据感兴趣，将诱饵文件中添加更新的数据，诱饵文件更易被攻击，随着时间的推移，这些数据将变得不再新，换句话说诱饵文件具有保质期。

（6） 屏幕或操作系统交互行为习惯

有研究使用系统调用模式或与屏幕交互模式，假定用户和计算机交互产生的调用序列可作为标识，来确定计算机用户是否合法。包括操作系统接收的任务指

令，系统调用和操作系统服务请求，电脑屏幕（可视化）的信息处理方式等验证用户身份。主要特征及功能包括：认知能力：通过捕捉屏幕上文本大小，了解浏览文本的能力；运动限制：通过从屏幕检测到的窗口动量，了解用户拖动窗口有多快；主观偏好：通过从屏幕捕捉窗口边缘的叠放情况，了解用户如何对多窗口布局特点；工作模式：通过可以在屏幕捕捉的特定应用程序分布，了解哪些套件是用户常用的。

4.2.2.2 主动身份认证典型过程

对行为生物特征的研究中，美国国防部高级研究计划局（Defense Advanced Research Projects Agency，DARPA）开展的“主动身份认证项目（Active Authentication Program，AA）”最具代表性。有研究表明基于行为特征的身份识别典型过程如下：

（1）常规方式登录，取得基本权限；

（2）监测用户的在线行为，获取并分析用户行为特征信息，例如使用鼠标的规律、创建文档的风格；

（3）随着用户行为确认度提高，访问权限上升，达到预设权限范围；

（4）达不到门限值，系统将重启身份认证过程。

下图给出了身份认证准确率未达到一定门限时的场景：

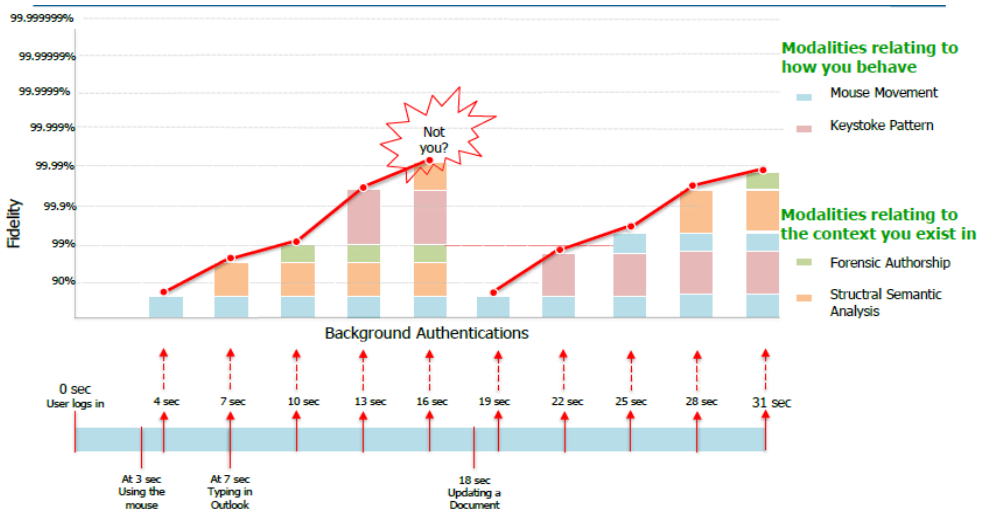

图 4-4 身份认证准确率未达到门限

4.2.3 匿名实体鉴别

在网络空间中进行实体鉴别时对用户隐私进行保护是个人隐私保护的关键步骤。GB/T 34953系列标准规范了实体鉴别隐私保护的技术，用于支持匿名实体鉴别。在匿名实体鉴别机制中，被鉴别实体（声称方）提供证据给验证方，该证据证实声称方知晓秘密且不会泄露声称方的身份给任何未授权实体，也就是说，通过在声称方与验证方之间交互的完整信息，未授权实体不能发现待验证实体（即声称方）的身份。同时，验证方可以通过拥有声称方的确定属性（如预定义的群组成员身份）来保证声称方的真实可信。即使被授权的验证方也不可能被授权去获得被鉴别实体的身份。匿名实体鉴别机制允许被授权方执行打开过程，这个过程使被授权方能够获得产生签名的实体的身份。允许打开的机制称为部分匿名实体鉴别机制。

匿名实体鉴别能够应用在许多场景中，如电子商务、电子投票、电子身份（例如电子驾照、电子健康证明和电子护照）、社交网络、移动支付以及可信计算。在许多这样的服务中，客户的个人身份信息（PII）被透露给服务提供者作为鉴别过程的一部分。限制服务提供者获取PII的一种方法就是使用匿名鉴别机制。匿名实体鉴别的一些用例可参见ISO/IEC 29191附录A。

4.2.4 零知识证明

设P（Prover）表示掌握某些信息并希望证实这一事实的实体，V（Verifier）是验证这一事实的实体。某个协议向V证明P的确掌握，某些信息，但V无法推断出这些信息是什么，我们称P实现了零知识证明。如果V除了知道P能够证明某一事例外，不能够得到其他任何知识，我们称P实现了零知识证明，相应的协议称作零知识协议。零知识证明目前在密码学中得到了广泛的应用，尤其是在认证协议、数字签名方面。零知识证明比较典型的方式是在每轮V都向P发出一询问，P向V做出应答。所有轮执行完后，V根据P是否在每一轮对自己发出的询问都能正确回答，以决定是否接受P的证明。

在一个安全的身份鉴别协议中，我们希望被鉴别者P能向验证者V证明他的身份，而又不向P泄露她的鉴别信息。较为著名的包括Feige-Fiat-Shamir身份鉴别方案，Guillo-Quisquater身份鉴别方案，Schnorr身份鉴别方案等。Feige-Fiat-Shamir零知识身份识别协议的目的是证明者P向验证者V证明他的身份（私钥），且事后V不能冒充P。证明者P有k个公私密钥对（向量），公钥组（v_1，v_2，…v_k）与与其对应的私钥组（s_1，s_2，…s_k），P欺骗V一次的概率为$1/2^k$，欺骗t次的概率为$1/2^{kt}$。属于一种交互式零知识证明协议。Guillo-Quisquater身份鉴别方案需要三方参与、三次传送，利用公钥体制实现。Schnorr身份鉴别方案是一种基于计算离散对数的困难性的鉴别方案，可以做预计算来降低实时计算量，所需传送的数据量减少，适用于计算能力有限的情况。

4.3 访问控制

4.3.1 自主访问控制

自主访问控制模型，客体的属主（资源拥有者）对客体（资源）进行管理，自主决定是否将资源授予其他主体。每个客体附加一个可以访问的主体的访问控制列表/矩阵，通过该列表/矩阵进行访问控制。其优点在于：自主性强，分布式授权灵活性高，配置细粒度小。缺点在于安全性低，效率低，维护管理工作量大。用户多、数据大时，访问列表不易维护。不保护客体副本：户 A 将对资源 O 的访问权限传递给用户 B，使不具备对 O 访问权限的 B 具有访问权。主体权限大，无意间泄露大量信息。

自主访问控制模型多用于商用、民用系统中，适用于安全要求不高、用户数少的小型 C/S 架构的个人应用，不适用于分布式网络。

4.3.2 强制访问控制

通过划定系统内各部分的安全等级实现了粗粒度的授权。是基于安全标识和信息分级的访问控制，不允许主体干涉。基于安全标签实现，将数据分成不同等级，用户权限类似划分，确定权限对应关系。其优点在于：强调安全性；有形式化安全模型；信息单项流通，若某等级出现安全风险，限制其访问高等级资源。

缺点在于权限划分粗粒度，过于强调保密性导致授权的可管理性不强，实现的工作量大，灵活性不强（安全需求发生变化时需繁琐的授权变动）。

强制访问控制适用于多级安全的军用系统和操作系统的访问控制，而不适用于主体客体频繁更新的场所。

4.3.3 基于角色的访问控制

20 世纪 90 年代提出，通过为用户指派对应的角色，使用户具有相应访问权限。访问控制系统中授权操作的对象是一种角色，代替以往模型中授予具体的用户。以银行环境为实例，角色可为出纳员、分行管理者、顾客、系统管理者、审计员等。

其优点在于：简化了系统管理，减轻了权限分配的重复劳动，明确责任预授权，加强了安全策略，具有灵活性和可管理性。缺点在于：模型依然从系统的角度出发的，没考虑执行的上下文，用户量增多、用户关系复杂度上升时，如何确定系统中的角色成为难点（角色挖掘）。

基于角色的访问控制只适用于角色结构严谨的企业内部访问控制场景。

4.3.4 基于任务的访问控制模型

随着数据库、网络和分布式计算的发展，关注点从静止的主体和客体保护转移到随着任务的执行而进行动态授权的保护上。只有在任务开始执行时才授予且任务结束授权就要被收回，否则可能导致安全泄露。基于上述问题，提出基于任务的访问控制模型，该模型从任务的角度来建立安全模型和实现安全机制，在任务处理的过程中提供了动态实时的安全管理。

其优点在于：动态授权的主动安全模型，防止主体对客体权限的无限期拥有；面向上下文，客体的访问控制权限随着执行任务的上下文环境变化而变化；能够对不同工作流实行不同的访问控制策略，并且能够对同一工作流的不同任务实例实行不同的访问控制策略。其缺点在于：授权管理工作繁杂，不支持角色的层次等级，不支持被动访问控制，很难单独模型化。

在实际应用中往往需要和其他的访问控制模型相结合。适用于工作流、分布式处理、多点访问控制的信息处理以及事务管理系统。运用于云环境时，依据任务和任务状态的不同对权限进行动态管理，极大增强了云中的访问控制的动态性。

4.3.5 基于任务−角色的访问控制模型

结合基于角色的和基于任务的访问控制模型的优点，既支持主动访问控制又支持被动访问控制，通过任务来实现上下文环境的权限分配而使用角色来对访问控制进行等级分层，满足了分布式环境下信息系统的访问控制需求。

基于任务-角色的访问控制模型基于角色模型框架上的扩展，继承角色模型的特点，结合任务模型的动态授权，将传统角色模型的 3 层访问控制模型改为 4 层，将权限通过任务分配给角色，而不是直接分配给角色，实现了动态分配基于任务-角色的访问控制模型。

4.3.6 基于属性的访问控制模型

针对复杂信息系统中细粒度访问控制和大规模用户动态扩展问题，提出基于属性的访问控制模型。模型核心要素包括主体、资源、操作、环境约束，要素统一使用属性、属性值表示。能够将其他模型中权限、安全标签、角色等概念用属

性统一描述。角色模型是基于属性的访问控制模型的一个子集，角色在属性模型中仅仅是用户的一个单一属性。

该模型适用于解决分布式环境下动态大数据访问控制。该模型通过属性来对实体及约束进行描述，能够严格控制访问者取得权限的各种条件，精确设定属性-权限关系，实现最小权限原则，达到细粒度访问控制。属性模型支持自主授权，可为资源拥有者提供策略管理接口，策略无需由管理员统一设定，资源拥有者可以根据自身实际资源保护需求发布、更新、撤销策略，保证资源能够按照资源拥有者的意愿被访问。属性模型依据请求者所具有的属性集合决定是否赋予其访问权限，实现了策略管理和权限判定的分离，且属性的设置与更新具有极大的灵活性和扩展性，可满足不同应用场景需求。属性模型还具备较小的系统开销，用户和资源数量大幅度增加时，传统访问控制模型策略数目将呈指数级增长，系统维护难度及开销将极大增加，属性模型中，策略随用户和资源的增长呈线性增加，当达到一定规模后，系统开销趋平稳。

4.4 脱敏技术

数据脱敏（Data Masking）的概念最先由 Adam 和 Wortmann 于 1989 年提出[2]。相关技术发展到现在，学术界提出了很多算法用于解决数据脱敏相关的问题，如噪声干扰（Noise disturbance），K-匿名（K-anonymous），微聚合（Micro polymerization）等。

脱敏技术贯穿于数据收集,数据存储,数据使用等数据全生命周期各个环节。与数据安全保护主要基于密码学技术实现数据机密性、完整性、可用性和不可否认性的关注点不同，隐私保护技术更关注数据的匿名化特征，防止攻击者将获得的公开数据与个人身份信息进行唯一、确定性的关联[14]。本节主要介绍几种代表性的隐私保护技术，包括：数据共享阶段的数据匿名化技术，差分隐私技术；数据利用阶段的同态加密技术；数据获取阶段的匿名通信技术，本地差分隐私技术等。

4.4.1 传统脱敏技术

传统的数据脱敏方法包括替代、混洗、数值变换、加密、遮挡和空值插入或删除等，这些方法的使用根据脱敏需求以及具体场景进行适当的选择，这些方法也是目前商业场景中所广泛使用的方法。经典的脱敏技术手段还包括但不限于以下几种：

——隐藏：通过将无需公布的敏感属性值置空或者替换为常数值的方式实现数据隐藏；

——置换：基于置换转换表将原始属性值映射到新值中以实现数据隐藏，只有拥有置换表的数据发布机构才能实现数据逆置换，恢复原始数值；

——调换：在不改变数据内容的前提下，通过改变数据的所属个体的方式实现数据的隐藏；

——截断：将属性值的末尾数据或前缀数据进行删除以实现数据隐藏效果；

——扰动：在数据发布之前对数据添加相应噪声，包括添加固定偏移量、随机增减数值等方式干扰攻击者区分真实数据以及带噪声数据；

——数据剪裁：数据剪裁的核心思想是将属于不同用户的数据记录水平或垂直方向进行分组与剪裁，通过将数据分开发布的方式保证攻击者无法从

碎片化的发布数据中找出特定目标对象所对应的敏感信息。

除了以上传统的数据脱敏实现方法以外，近年来，还有不少对上述方法的改进。Sarada 和 Abitha 在[6]中提出了基于最小最大归一化的算法，实现了对数值范围的映射。同时还提出了范围映射的方法，能够实现对数值的双向映射。Vishal A.Gujjary 和 Ashutosh Saxena 在[7]中提出了基于神经网络的自适应脱敏算法，该算法需要实现设置规则并对神经网络进行训练，并集合了替换、置乱、遮蔽以及数值抖动等基本的脱敏方法，通过神经网络进行智能的脱敏选择处理，在保持数据看似真实的前提下有效的避免了词典攻击等传统脱敏方法的一些弊端。

在产业化方面，IBM、Oracle 和 Microsoft 等传统数据库大厂也推出了相应的数据脱敏产品。IBM 将数据脱敏集成到自身的关系型数据库系列产品中[3]。而 Oracle 以插件的方式提供数据脱敏的扩展[4]。Microsoft 的方案可以实现在生产数据库到测试数据库的数据传输过程中部署用户自定义转换函数从而实现数据脱敏操作。

除了数据库厂商外，还有不少数据集成领域的厂商也推出了相应的数据脱敏产品，如 Informatica 和 NET 2000 等，Informatica 的主营业务为数据集成，其数据脱敏产品构建于其数据集成框架上，能够以中间件的方式对外提供服务。NET 2000[5]所提供的产品为工具软件，提供数据脱敏的功能。

随着大数据时代的到来，人们更加关注于对数据的挖掘，传统的方法虽然屏蔽了敏感数据，同时也对数据的价值造成了损坏，这就带来了两个问题。如何能够在保护数据价值的同时对敏感信息进行保护。如何在数据准确性损失最少的情况下实现对隐私的最大保护。S.Vijayarani 和 Dr.A.Tamilarasi 在[8]中提出了两种数据变形方法，一种是数据变形技术，一种是比特位变形技术，均为对数值型数据的变形处理技术，处理后的数据不仅不会破坏统计分析算法的结果，而且能够在一定程度上对分析结果进行改善，有助于数据挖掘。

Yijie Zhou 和 Francesca Dominici 在[9]中提出了基于空间平滑（spatial smooth-ing）的矩阵屏蔽算法。并将其应用在了医疗记录的脱敏当中，该方法基于空间平滑技术，在该技术中，空间权重函数控制屏蔽的格式，平滑参数用于调整屏蔽的程度。这样的话，数据的可用程度和泄漏风险都可以通过屏蔽程度和屏蔽方式计算得出。实际上，平滑权重函数 W 可以是任何的权重函数，不仅限于现有的平滑方法，并且可以由多种权重函数和不同参数的平滑函数组合而成。在地理位置脱敏领域，Marc P. Armstrong 和 Gerard Ruston 在[10]中提出了对医疗数据地理位置的脱敏算法，在保障基于地理位置的医疗数据挖掘分析价值的前提下，对病患的地理位置信息进行脱敏处理。

国内外大多数研究都关注于具体的数据脱敏算法，并且已经取得了一定的成果，在数据脱敏模型方面，Min Li 和 ZheLi Liu 等在[11]中提出了数据脱敏通用模型，该模型在 Oracle 提出的 FAST 模型[4]和 IBM 的 OPTIM 脱敏模型[3]的基础上进行了优化，更加详细的对数据脱敏通用模型进行了设计，并细致的阐述了其中每一个步骤的功能。B.Liver 等人提出了隐私应用服务架构[12]，该架构关注于企业信息系统中不同服务组件与脱敏系统的关系，通过设置脱敏网关的方式部署数据脱敏服务，对所有对外发布的数据自动进行脱敏处理。

目前静态数据脱敏的研究相对成熟，并且已经有不少学者开始研究动态数据脱敏，动态数据脱敏的难点在于无法提前预知整体数据的情况，难以控制整体的信息损失，使得脱敏效果不如静态数据脱敏。流数据脱敏和动态数据脱敏具有一

定的共性，虽然可以用现有的动态数据脱敏方法解决流失数据脱敏领域的问题，但是两者之间并不完全相同，脱敏效果有所欠缺。流数据脱敏领域的相关研究较少，Aleksey I. Baranchikov, Aleksey Yu. Gromov 等在[13]中对动态脱敏技术进行了讨论，通过与传统的静态数据脱敏技术进行对比，阐述了动态脱敏中的关键技术难点并提出了一系列的动态数据脱敏方法。

4.4.2 k-匿名及变种

k-匿名最早是由 Sweeney 等人提出的一种隐私保护模型，由于其直观的模型解释、相对低开销的实施成本被 Google、英特尔等众多国际知名 IT 公司所采用，作为其数据发布、数据共享环节的一种隐私保护技术。

k-匿名模型的核心思想是通过将每条个人记录信息隐藏在一组具有相似属性值的人群记录中来达到隐藏当前个人隐私的目的，避免当前记录所对应的个人被攻击者唯一识别出来。

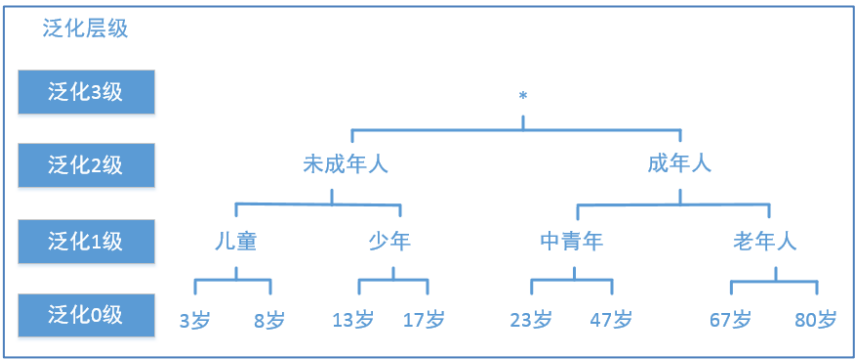

图 4-5 年龄属性多层级泛化

为了实现数据发布与共享过程的隐私保护，k-匿名模型在数据发布之前首先对唯一标识属性进行删除或者模糊化处理，防止基于此类属性的直接个人身份识别。随后针对准标识属性进行泛化处理，根据实际场景下的隐私保护需求，将低级别的具体准标识符信息逐次泛化为更高级别的抽象标识符信息（图 展示了一种针对年龄属性的多层级泛化方式）。通过逐层泛化，当所有准标识属性列中至少有 k 条记录拥有完全相同的泛化后属性值时，即称当前数据集实现了 k-匿名化。

尽管 k-匿名化模型能够保证攻击者无法将单条匿名化信息映射到个人唯一身份标识上，但仍存在多种攻击方式能够从 k-匿名化数据集中提取出个人隐私信息，这些攻击方式包括：

1) 同质化攻击：如果某个泛化后的 k-匿名记录组内所有的敏感属性值是相同的，那么攻击者无需精确识别数据记录与个人标识间的对应关系即可获知目标个人的该项敏感属性值；
2) 背景知识攻击：如果某个泛化后的 k-匿名记录组内所有的敏感属性值并不完全相同，但是攻击者能够根据自身掌握的额外信息排除记录组内少量的差异性敏感属性值，也能够挖掘出个人敏感隐私信息。

为了应对上述针对传统 k-匿名模型的攻击方案，学术界又提出了多种改进模型以实现模糊化程度更高的数据隐私保护，包括 l-多样性模型，t-贴近性模型等。

l-多样性模型的核心思想是在 k-匿名化模型的基础之上进一步要求被发布数据集的每个等价匿名分组中至少存在 l 个不同的敏感属性值，多样性 l 取值越大则攻击者越难以恢复出实际目标个体的准确敏感属性取值，从而能够一定程度上对抗同质化攻击以及背景知识攻击，表 4-2 展示了一个实现 3-匿名与 2-多样性

脱敏处理的病人档案数据。

表 4-2 3-匿名模型与 2-多样性模型处理后的的病例数据

所在区	年龄	疾病
海淀	4*	胃溃疡
海淀	4*	支气管炎
海淀	3*	心脏病
海淀	4*	胃溃疡
海淀	3*	乙肝
海淀	3*	阑尾炎

t-贴近模型则更进一步，通过约束敏感属性值在整个数据集中的分布与每个泛化后 k-匿名等价类中的分布差异不超过阈值 t 来实现更高安全性的反匿名化效果。

4.4.3 差分隐私技术

差分隐私技术最早由微软研究院硅谷实验室首席研究员 Dwork 和哈佛大学教授 Mckay 等人于 2006 年提出，是一种有效限制和量化个人隐私信息泄露的信息保护输出模型。在差分隐私机制的保护下，数据共享机构或组织无需发布其所掌握的原始数据资料，而是通过在用户查询的响应信息中添加随机噪声的方式来保护个人隐私数据。

通过使用随机算法向用户查询结果中添加随机噪声的方式，差分隐私技术能够保证任意个人的数据信息不被泄露。更形式化的说，假设我们将已有的数据库集合记为 D，将与 D 集合仅相差一条数据记录的数据集记为 D’（这里的单条数据记录差异可以由 D 集合中针对任一数据记录的增删改操作产生，称 D’为 D 的邻近数据集合）。如果随机算法针对 D 和 D’所产生的带噪输出结果拥有相近的概率分布，那么攻击者就难以判断当前获得的查询结果来源于哪个具体的数据集合。换言之，对于任意一条数据记录，攻击者所获得的带噪声查询结果可能来源于包含此条数据记录的数据库集合，也可能不是。因此，攻击者无法判断该数据记录是否真实的存在于原始数据库集合 D 中，也即无法从中获取个人的隐私泄露信息。

下面以评选先进党支部为例阐述指数机制下的差分隐私保护运行机制。假设现在需要通过投票的方式评选先进党支部，所有选票的原始数据集合为 D，查询函数 f 的目标是获知票数最高的某个最终获选先进党支部r_i，三个候选党支部的最终得票数情况如表 所示。

表 4-3 党支部得票数情况

党支部	得票数
第一党支部	50
第二党支部	20
第三党支部	30

随后将打分函数$\mathrm{q}(\mathrm{D}, r_i)$定义为各党支部所获得的得票数情况。则打分函数敏感度$\Delta \mathrm{q} = 1$（所有选票集合 D 中任意增加、删除、修改一票只对最终的总得票数影响为 1）。通过将隐私保护预算ε设置为 0，0.1，0.5 三个不同的值，可以计算出相应的带噪声随机函数 M(D)输出分布，如表 所示。

表 4-4 随机函数 M(D)输出分布

党支部	ε=0	ε=0.1	ε=0.5
第一党支部	1/3	0.629	0.993
第二党支部	1/3	0.140	0.001
第三党支部	1/3	0.231	0.006

当隐私保护预算ε为 0 时，带噪输出函数 M 完全不依据实际投票情况进行响应，等概率输出任意一个党支部为最终的优胜党支部，数据查询结果的不具备可用性，隐私保护效果却最优。而当ε取值为 0.5 时，带噪输出函数 M 将以极大概率输出实际获胜的第一党支部为查询结果，查询响应结果可用性较高，但隐私保护效果却不佳。

通过调节隐私保护预算ε，数据共享机构能够找到满足自身隐私保护要求的隐私保护预算值，在用户隐私保护与共享数据可用性两方面中寻找最佳平衡。

4.4.4 本地差分隐私技术

差分隐私技术根据应用场景的不同分为集中式差分隐私以及本地差分隐私两大类。

上面介绍过的集中式差分隐私技术主要关注拥有可信数据管理第三方的场景下如何针对汇聚数据添加相应的噪声扰动，再进行共享与发布，以防止个人信息泄露。本地差分隐私的应用场景则拥有更加苛刻的条件，该场景假定拥有隐私信息的用户在没有可信数据管理第三方或者不相信除自身以外任何第三方的前提条件下仍然能够在自身数据被收集时确保个人隐私信息安全。

Kasiviswanathan 等人于 2008 年的 IEEE FOCS 会议上最先提出了本地差分隐私技术[15]，该技术的一项应用是挖掘统计意义上的用户倾向。举例来说，互联网服务提供方常常希望通过网络问卷调查的方式了解目前大多数用户的关注热点（但不关注任何单一个体的关注热点），以便能够更具针对性的提供定制化应用服务。用户为了获得更优质的服务一方面乐于参与这样的调查活动，另一方面也希望自己的个人调查问卷内容得到保护。本地差分隐私技术为这种热门选项统计应用场景提供了实现可能。

目前差分隐私领域的两大最重要的奠基性成果分别是 2014 年提出的 Rappor 协议[16]以及 2015 年提出的 SH 协议[17]。

Rappor 协议由谷歌的 Erlingsson 等人设计并开发，并已经成功的应用在谷歌浏览器中进行用户隐私信息的收集。Rappor 协议第一步在本地将用户的待收集数据通过布隆过滤器（Bloom filter）映射到定长的二进制序列中，随后基于随机应答协议（Random Response，RR）的两次随机比特翻转过程实现用户终端数据的随机化处理。第二步数据收集者通过汇总所有用户上交的随机化二进制序列，基于本地随机化步骤中 RR 协议预设的随机化概率对统计信息进行随机化修正，从而获得大数据量情况下所有候选项全集的近似投票频数，进而排序获得最热门选项统计结果。

SH 协议则是由 Bassily 等人在 STOC’15 会议上提出的，其与 Rappor 协议的不同点在于，SH 协议仅对用户赞成的候选项结果进行 RR 随机应答，而对其他候选项执行概率为 50%的随机支持响应，有效提高了本地化数据处理的效率。同时后续的改进版本又通过引入 Hash 函数等方式将用户向服务器端发送的信息量压缩为 1 比特，从而极大地提高 SH 协议的信息传输效率。

4.5 基于密码技术的个人信息保护

4.5.1 采用密码技术的身份凭证

密码软件身份凭证是存储在磁盘或其他“软”介质中具有密钥存储和密码计算功能的软件，通过证明对密钥的拥有和控制完成鉴别。可分为单因素密码软件和多因素密码软件。多因素密码软件需通过额外的鉴别因素激活（例如，记忆秘密或生物特征）。

密码设备身份凭证是一种硬件设备，可分为单因素密码设备和多因素密码设备。多因素密码设备是使用受保护的密钥执行密码操作的硬件设备，受保护的密钥应由额外的鉴别因素激活。

密码技术广泛应用于身份凭证用于身份鉴别，它可以大大降低在线的身份窃取、身份伪造等风险。然而，除了需要防范浏览器攻击和中间人攻击的风险外，还需在使用时考虑以下几个方面：

——适宜的密钥长度，降低被破解的威胁；

——服务和管理的质量，包括：确保鉴别器安全地传输到正确的用户；确认正确、真实的个人信息被安全地存储；共享秘密的安全存储和传输等；

——防止主动的欺骗检测，如对失败的鉴别请求的审计。

4.5.2 基于数据加密的访问控制技术

基于数据加密的访问控制技术常用于隐私数据加密存储。加密存储是指在不完全信任身份提供方的情况下，数据拥有者在客户端先对数据加密，再将加密后的数据外包给身份服务方。在加密存储中，存在着三种主要角色：提供数据存储服务的身份提供方、作为数据拥有者的用户和作为共享用户的依赖方。其中，加密数据的密钥由数据拥有者掌控，提供方在没有密钥的情况下不能泄露数据内容。因此，只有授权获得密钥的共享用户才能读取加密数据。访问控制的实施是通过安全管理密钥实现的，从而使得数据拥有者具备对其数据访问的控制能力。因此，在加密存储中，访问控制主要是以客户端密码技术为基础，集成了身份鉴别、加密处理、远程完整性验证、数字签名等技术，通过管理密钥（生成、分发、更新、销毁）来实施访问控制策略。

在具有大规模数据存储需求下，使用对称密码机制加密数据内容是一种有效的解决办法。为了实现数据的共享，访问控制通常可通过安全地分发密钥实现。处理大规模个人信息的系统对个人信息数据安全具有更高的要求，应实现大规模数据存储需求下的细粒度访问控制，并且减轻用户维护密钥的负担。然而，许多的高性能计算系统、云存储系统、安全内容分发网络系统等在处理这些大规模个人信息时，个人信息的机密性对物理隔离的依赖性较强，个人信息存储通常采用明文方式或者简单的加密方式，并且很少有实现细粒度加密策略的方案。

为了减轻密钥维护的负担，基于属性的加密方法（ABE:Attributed-based encryption）被提出，属于一类在不需要为每个文件或文件组维护密钥前提下，保护系统中敏感数据的解决方案。目前，ABE 已经在敏感数据保护领域得到了广泛的研究。基于属性的加密通过对每个用户或每个文件附加一组属性的方式来设计访问控制策略。只有当用户的属性匹配上文件的属性时，用户才能够访问文件，从而使得在多个用户共享访问多个数据资源的情况下，用户只需要维护少量的属性密钥便可以访问授权的资源，从而实现细粒度的访问控制。

4.5.3 同态加密技术

大数据、云计算时代，用户无时不刻都在产生大量的个人信息。从用户主动上传到云端网盘的个人工作、学习资料数据，到日常购物产生的历史订单数据，再到手机 App 时刻记录的用户浏览数据、位置轨迹数据，这些包含个人信息的海量数据，往往最终都通过汇聚到服务提供方建设的各地数据中心进行集中化的存储与管理。用户在关心这些数据如果被更好的利用以便为自身提供各种定制化便捷服务（数据共享、商品推荐、路线规划等）的同时，也十分关注自身上传数据的安全性问题以及隐私保护问题。

传统上由于服务提供方的各类查询功能、分析工具只能针对明文数据开展计算与统计分析，因此用户不得不牺牲自己的隐私信息才能换取服务提供方开放的各种定制化服务。更糟糕的情况是，这种用户隐私的牺牲是完全不可逆的，即使用户最后由于某种原因终止了自己与各类服务提供方之间的服务与被服务关系，在相当长的一段时间内，服务商仍然能够在用户毫不知情的情况下分析、使用这些历史存量数据。为了解决上述困境，学术界对同态加密技术开展了深入研究。

同态加密技术与传统加密技术最大的区别在于，同态加密技术允许直接在加密结果上进行相关计算，密文计算结果与直接针对明文数据进行计算之后再进行加密的结果完全相同。这就意味着用户能够放心的将自己拥有的隐私数据加密后再提交给云端服务商，服务商在不知晓用户隐私数据的前提下直接对密文数据进行计算、分析，并将密文计算结果以及提供的相应服务返回给终端用户，用户在终端解密数据后即可获得正确的数据计算结果，而整个数据流动过程中，用户的个人隐私信息却完全没有泄露给第三方的服务提供方。

图 4-6 显示的是基于同态加密技术，第三方服务提供方能够基于某公司员工提交的加密工资单数据为公司财务提供查询员工平均工资服务的整体流程示意图。

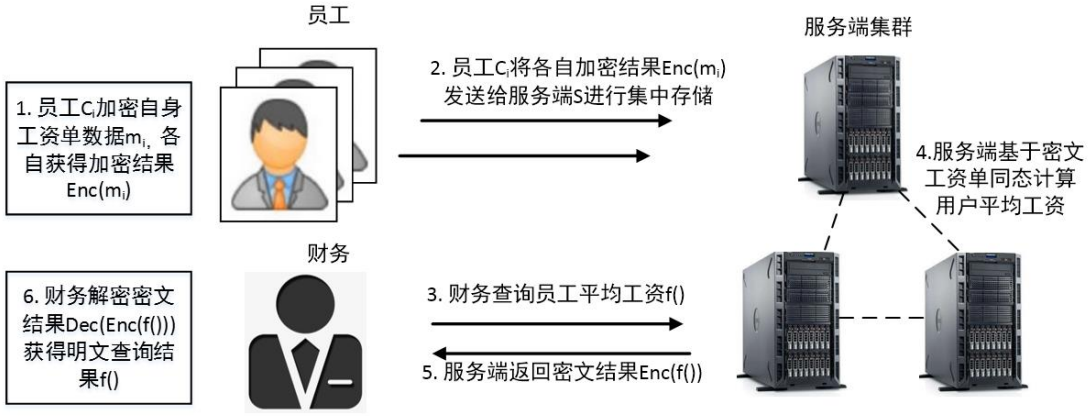

图 4-6 同态加密应用场景示意图

根据对加密数据进行运算的种类以及运算操作次数的限制不同，同态加密方案可以被分为三种类型[18]：

1) 部分同态加密方案（PHE）：该方案只允许对加密数据执行一种操作运算，但运算次数不限；
2) 类同态加密方案（SWHE）：允许对某些类型的操作进行有限次数的操作运算；
3) 全同态加密方案（FHE）：允许无限次数的操作。

一方面，PHE 方案被部署在一些只包括加法或乘法运算的特定应用程序中，如电子投票应用或私人信息检索应用，这些应用受到同态求值运算类型的限制。另一方面，SWHE 方案同时支持加法和乘法。然而，在第一个 FHE 方案提出之前，SWHE 方案中的加密密文的大小会随着每一次同态操作而增长，因此允许的最大同态操作数是有限的。这些问题限制了 PHE 和 SWHE 方案在实际应用中的使用。最终，基于云的服务越来越受欢迎，加速了 FHE 方案的设计，这些方案可以支持任意数量的同态操作和随机函数。Gentry 的 FHE 方案是第一个可行的和可实现的 FHE 方案。它以数学中的理想格为基础，开创性地提出了基于类同态加密方案构造全同态加密方案的构造思路，即自举（Bootstrapping，用同态方案运行自身解密方案）和压缩（Squashing，压缩解密电路深度）。然而，该方案在自举部分，即经过处理的密文中间刷新过程计算成本过高，限制了其在实际应用场景中的可行性。在接下来的几年里又陆续提出了多种后续改进方案，包括 11 年 Brakerski，Gentry 等人构造的 BGV 方案，Gentry，Sahai 等人在 13 年美密会上提出的基于格密码 LWE 问题的全同态方案，2015 年欧密会上 Ducas 和 Micciancio 提出的能够将自举过程的时间压缩到 1s 以内的双层全同态方案等。

4.5.4 安全传输技术

互联网一般采用 HTTP 协议进行数据传输，HTTP 是明文传输的，恶意的中间人和窃听者通过截取用户发送的网络数据包可以拿到用户的敏感信息。尤其是涉及利用身份管理服务实现网上交易，银行转账等操作更是危害极大。为了防止个人属性信息、身份凭证等数据被窃密、篡改和伪造，通过公共网络传输时，应使用加密通道或数据加密的方式进行传输，保障个人息传输过程的安全。

信道加密技术注重解决信息在线路传输过程中的安全问题，并且可很好地控制非法用户的侵入。建议采用 HTTPS 进行数据传输，HTTPS 的核心是 SSL/TLS 安全协议层，该层位于应用层和传输层之间，TLS 对应用层数据加密后再向下交给传输层，以解决 HTTP 安全传输的问题。身份管理系统的各参与方之间数据传输，应遵循 GM/T 0054 的要求采用安全通道的方式对通讯数据进行保护。安全通道应使用合规的加密算法和安全协议保护所有连接，保证传输数据的机密性和完整性，例如使用 SSL/TLS IPSec VPN 等协议。

对传输的数据进行内容加密可采用基于密码技术的加解密功能实现数据内容的机密性保护；使用基于对称密码算法或密码杂凑算法的消息鉴别码机制、基于公钥密码算法的数字签名机制等密码技术实现数据内容的完整性保护；使用基于对称密码算法或密码杂凑算法的消息鉴别码机制、基于公钥密码算法的数字签名机制、动态口令机制等密码技术实现传输数据的真实性保护；使用基于公钥密码算法的数字签名机制等密码技术实现不可否认性。

5. 身份管理产品中的个人信息保护技术

Amazon、IBM、腾讯、阿里、McAfee、Google 等国内外大型互联网公司、安全公司也在积极研发自己的数据安全解决方案或云端数据安全产品，以期能够为广大企业用户提供便捷、高效的数据安全与隐私保护服务。下面仅以亚马逊云、Auth0 的涉及身份管理中个人信息保护的解决方案为例进行阐述。

5.1 亚马逊云服务中的个人信息保护方案

AWS[19]从四个方面对云安全提供了相应的解决方案，包含了身份与访问管理，检测式控制，基础设置保护和数据保护，其中每个方面都提供了一些对应的服务来帮助用户实现云上的安全。其中涉及身份管理的模块包括 IAM，SSO，Cognito 等服务。

使用 AWS Identity and Access Management (IAM)组件，用户可以创建并管理 IAM 用户身份，并授予这些 IAM 用户访问您资源的权限，安全控制 AWS 资源的个人访问权限或组访问权限，也可以为 AWS 以外的用户授权（联合身份服务）。Amazon Cognito 借助 Amazon Cognito，可以快速轻松地为 Web 和移动应用程序添加用户注册/登录和访问控制功能，可将用户规模扩展到数百万，并支持通过 SAML 2.0 使用社交身份提供方（如 Facebook、Google 和 Amazon）以及企业身份提供方进行登录。

AWS Secrets Manager 可以保护访问应用程序、服务和 IT 资源所需的密钥。该服务能够跨整个生命周期轮换、管理和检索数据库凭证、API 密钥和其他密钥。借助 Secrets Manager，可以使用细粒度 AWS Identity and Access Management (IAM) 策略和基于资源的策略管理私密信息的访问权限。例如：创建一个策略以使开发人员仅检索开发环境中使用的特定私密信息。同样的策略可以使开发人员仅在请求来自企业 IT 网络内时才可以检索生产环境中使用的口令。对于数据库管理员，可以创建一个策略以使数据库管理员能够管理所有数据库凭证和权限，以便读取对托管数据库的特定实例执行操作系统级别更改时所需的 SSH 密钥。

AWS Key Management Service (KMS)AWS Key Management Service (KMS) 可以创建和管理加密密钥，并控制密钥在各种 AWS 服务和应用程序中的使用。AWS KMS 是一种安全且有弹性的服务，使用已经通过 FIPS 140-2 验证或正在验证的硬件安全模块来保护密钥。AWS KMS 还能与 AWS CloudTrail 集成，从而为用户提供所有密钥的使用记录，帮助用户满足监管和合规性要求。

AWS Certificate Manager 提供证书服务，可帮助用户预置、管理和部署公有和私有安全套接字层/传输层安全性（SSL/TLS）证书，以便用于 AWS 服务和用户的内部互联资源。SSL/TLS 证书用于保护网络通信的安全，并确认网站在 Internet 上的身份以及资源在私有网络上的身份。

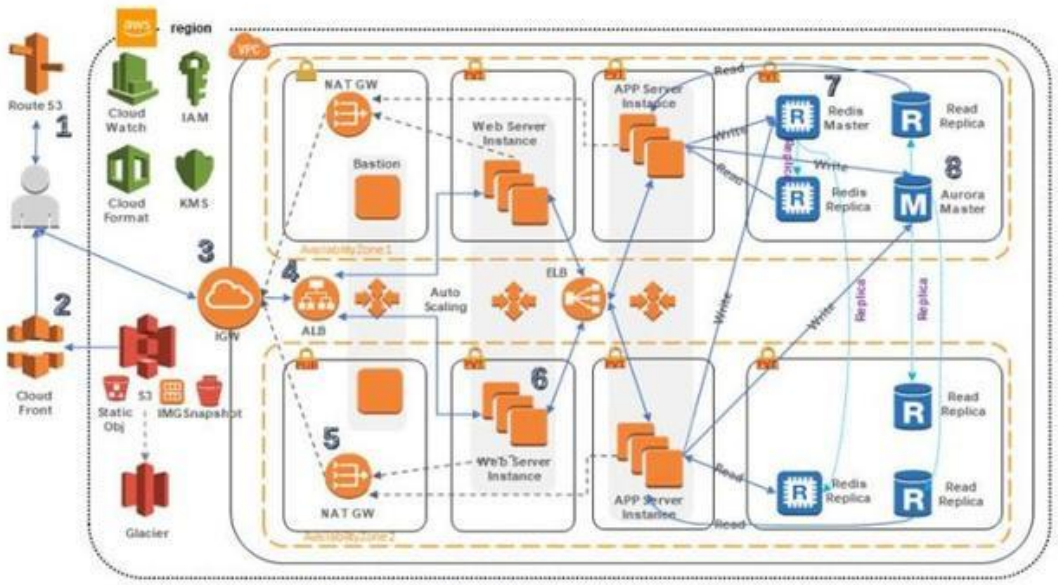

图 5-1 AWS 系统部署经典安全实践

AWS 系统部署经典安全实践如上图所示：基础框架为经典的三层结构，所有的服务器都在内网。每个子网都有自己的 NSG 网络安全组，对外的访问全部通过路由表指向到 NAT 网关。在公有子网中只有 NAT 网关，堡垒机和 ALB 应用程序负载平衡。外来访问都是通过 ALB 再发送到 web 服务器。这样实现了内外阻断。唯一可以从来访问内网的方法就是通过堡垒机来实现。所有的服务器都是多可用区部署，实现了服务的高可用性和数据的冗余。利用 IAM 和 KMS 来实现对访问的管控，内部数据传输和存储的加密。而 Cloudwatch 则用来监控整个系统的运行状况。

5.2 Auth0 身份服务中的个人信息保护

Auth0[20]作为身份与授权管理云服务商，提供了一整套 SDK 和 API 用于实现多种身份与授权管理功能，便于 APP 开发者可以更加关注与业务逻辑实现。主要功能包括授予访问权限、从任意设备提供单点登录、通过多因素身份验证增强安全性、用户身份数据生命周期管理、保护特权帐户等。

Auth0 的用户是个人信息控制者。Auth0 是个人信息处理者。Auth0 维护并满足多个合规性框架和认证的需求，如欧盟 GDPR、美国 HIPAA 和 HITECH 法案等。

Auth0 处理的个人信息包括终端用户的所有数据，数据库链接的用户配置文件信息以及外部身份提供者提供的信息，均位于 Auth0 用户配置文件中。用户配置文件中包含的特定属性因用户实现的不同而不同，并且取决于许多因素，例如连接类型、身份验证期间的用户同意以及用户是否用额外的信息扩充了配置文件等。

存储在 Auth0 中的个人数据仅用于提供其服务的目的，即验证用户。当终端用户的帐户被删除时，他们的用户配置文件(包括元数据)也被删除。

作为数据处理者，Auth0 承担以下责任：

——按照《订阅协议》(SA)和《数据处理附录》(DPA)(适用于企业客户)或《服务条款》(适用于自助服务客户)中所述的数据处理者的说明执行；
——如果终端用户行使作为数据主体的访问、擦除等权力，则通知客户；
——如果收到监管部门的请求，通知客户(除非执法部门禁止)；
——如果发现安全漏洞，通知客户；
——如果其任何子处理器向 Auth0 通知已确认的影响 Auth0 客户数据的数据泄露(除非执法部门禁止)，则通知客户；
——提供隐私政策、服务条款、安全声明、数据保护协议等，以提供有关其政策和实践的信息；
——提供有关数据处理的信息，使客户获得合法处理数据所需的信息；
——定义其服务和功能，数据如何处理，以及客户的权利和义务；
——通过 Auth0 Dashboard 和 Auth0 Management API，客户可以检索、查看、更正或删除客户数据；
——为客户提供显示同意条款和同意协议复选框组件，如果需要更详细的同意方案，客户还可以设计定制的注册和登录表单。

Auth0 个人信息处理方面的安全机制：

——用户同意方面：客户可以使用 Auth0 在注册时征求用户的同意（使用 Lock 或自定义表单），并将此信息保存在用户配置文件中。之后可以使用 Management API 更新此信息；
——访问、更正和删除的权力方面：客户可以手动或使用 API 访问、编辑和删除用户信息；
——数据最小化方面：Auth0 所收集的个人资料限于处理所需的信息，只在必要时保存，在数据处理过程中确保适当的安全性，包括防止未经授权或非法的处理以及防止意外的丢失、破坏或损坏，其他功能如帐户链接、用户资料加密等可保护用户个人信息；
——数据可移植性：可以手动或以编程方式导出存储在 Auth0 用户存储区中的用户数据。Auth0 的原始数据可以导出为 JSON 格式；
——用户个人信息保护措施：提供用户配置文件加密、强制保护、破解口令检测、增强身份验证等方法满足个人信息保护需求。此外，Auth0 建议采用以下做法，帮助确保用户数据及个人信息的安全：

- 保护客户密钥；
- 保护身份凭证，并需要多因素身份验证才能访问主控面板；
- 定期检查管理员列表、租户关联的连接和应用程序列表；
- 确保管理员使用可以在必要时可安全撤销的企业凭据，而不是个人凭据，如个人电子邮件帐户；
- 立即删除被解雇员工的帐户；
- 确保管理员使用强制锁屏的设备；
- 为所有主控面板管理员和开发人员提供关于安全和隐私最佳实践的定期培训。

其他类似的可用于身份管理系统中个人信息的安全产品或解决方案还包括 McAfee 公司的云/数据库/网络/终端数据安全组合产品[21]，阿里云数据安全解决方案[22]，微软数据信息保护方案[23]等。

6. 基于《个人信息保护法》的技术措施建议

为了保护个人信息权益，规范个人信息处理活动，促进个人信息合理利用，我国颁布了《中华人民共和国个人信息保护法》，明确了自然人的个人信息受法律保护，任何组织、个人不得侵害自然人的个人信息权益。在中华人民共和国境内处理自然人个人信息的活动，在中华人民共和国境外处理中华人民共和国境内自然人个人信息的活动均适用于该法。编制组从该法律的具体条款要求出发，对具体法律条款进行分类解析，推荐为满足条款可采取的管理与技术措施，并分析设计形成个人信息保护的技术标准体系。

6.1 总则

总则部分给出了该法律的基本信息，适用范围和个人信息保护基本原则。个人信息处理基本原则包括：

- 合法、正当、必要、诚信原则；
- 目的明确和最小必要原则；
- 公开透明原则；
- 质量及安全保障原则。

身份管理系统主体应采取措施保证符合这些基本原则，并将上述原则贯穿于个人信息处理的全过程、各环节。

表 6-1 总则部分措施建议

措施建议	技术保障	密码技术应用
● 注册前主动告知个人信息主体，包括但不限于以下内容： - 声明合法性与正当性：身份管理服务主体，联系方式，主体证照； - 声明正当性：说明身份管理系统可提供的服务及功能，以及对个人信息的处理需求； - 声明必要性：需收集的所有个人信息，各项个人信息及其处理结果的用途； - 遵守公开透明原则：声明可能对个人信息进行处理的规则、目的、处理方式，处理的个人信息种类、保存期限； - 声明身份管理服务主体对个人信息处理所遵循的法律要求和承担的法律责任； - 声明个人依法行使权利的方式和程序；	● 信息接口保护 ● 会话管理 ● 防注入攻击 ● 采用可靠信道 ● 提供用户个人信息完整性校验机制，保障用户个人信息的完整性； ● 针对个人信息制定访问控制策略并正确合规实施访问控制，防止违规篡改； ● 制定个人信息处理规则、流程，处理后不完整、不准确的信息进行去标识化处理，确保无法关联到特定自然人，不	● 对告知用户信息进行签名，防篡改； ● 用户同意信息可溯源，防抵赖； ● 采用 https 等加密信道传输； ● 保障会话密钥安全； ● 采用密码技术对用户的个人信息进行加密，可对完整性进行验证。

措施建议	技术保障	密码技术应用
-法律、行政法规规定应当告知的其他事项。 ● 区分必选功能与可选功能，声明必选功能对应所需的个人信息与可选功能所需的额外个人信息； ● 得到个人信息主体的授权同意： -提供是否接受身份管理服务必选功能的选择，并获得所需的个人信息的授权； -逐条提供是否接受可选功能服务的选择，并获得所需的个人信息的授权； ● 制定个人信息保护策略规则，并明确所采用的技术措施及标准规范，保障个人信息安全。	会利用处理后不完整、不准确的个人信息为关联用户提供鉴别、授权、公开等服务； ● 对个人信息的全部处理操作记录日志，定期审计。	

6.2 个人信息处理规则

《个人信息保护法》第二章“个人信息处理规则”部分，将个人信息主体的“同意”视为个人信息处理的合法性基础，并增加了处理个人信息的其他合法情形；明确了个人信息主体撤回同意权，并对个人信息处理者响应个人信息主体关于撤回同意的要求也进行了规定；规定了共同处理个人信息的规则，要求共同处理者承担连带责任有利于多个个人信息处理者相互监督、督促对方及时整改违规行为；明确采用自动化决策情形下的处理规则；明确公共场所监控设备的使用规则和限制性规定；提出关于敏感个人信息处理规则。

表 6-2 个人信息处理规则部分措施建议

措施建议	技术保障	相关密码技术应用
● 采用上述主动告知个人信息主体、得到个人信息主体的授权同意等措施； ● 当身份管理系统无法提供正常服务，或主体合法变更后，应主动告知用户相关变更信息，并获得用户授权； ● 当个人信息的处理目的、方式、种类和其他告知内容发生变更的，应重新进行主动告知，并重新获得个人信息主体的授权同意； ● 为个人信息主体提供撤回授权的接口，撤回授权后，不再对未授权的个人信息进行处理； ● 撤回可选功能的授权不影响个人信	● 信息接口保护 ● 会话管理 ● 防注入攻击 ● 采用可靠信道 ● 提供用户个人信息完整性校验机制，保障用户个人信息的完整性； ● 针对个人信息制定访问控制	● 对告知用户信息进行签名，防篡改； ● 用户同意信息可溯源，防抵赖； ● 采用 https 等加密信道传输； ● 会话密钥安全； ● 采用密码技术对用户的个人信息进行加密，

措施建议	技术保障	相关密码技术应用
息主体使用必选功能和其他可选功能的服务； ● 制定个人信息处理规则，并公开告知用户，支持用户查看并支持下载等方式保存； ● 紧急情况下为保护自然人的生命健康和财产安全无法及时向个人告知的，个人信息处理者应当在紧急情况消除后及时告知； ● 应梳理身份管理服务功能中需要使用的个人信息，声明必要性和需要个人信息的期限； ● 个人信息保存期限不超过用户享受身份管理服务功能的时间，且不超过个人信息用于功能实现所需要的时间； ● 超过个人信息保存期限应及时更新授权或彻底删除个人信息； ● 对于其他参与个人信息处理的合作方，应约定好各自的权利和义务； ● 委托第三方处理个人信息的，应当与第三方约定委托处理的目的、期限、处理方式、个人信息的种类、保护措施以及双方的权利和义务等，并对第三方的个人信息处理活动进行监督；委托失效、终止、到期后，第三方应返还个人信息或彻底删除； ● 身份管理服务提供方，如有需利用收集的个人信息进行自动化决策的功能，应作为可选功能，需说明对个人权益的影响和决策过程，并为用户提供是否使用该自动化决策服务的选择功能； ● 自动化决策功能使用的算法不应以交易价格作为最主要特征；应对自动化决策结果进行公平性验证，如使用不同交易价格区间的分组用户进行自动化决策，说明决策结果与交易价格不强相关。 ● 个人信息生命周期管理： - 收集：一般不使用公共场所设备收集的个人信息，若确需使用，需经过个人单独同意； - 收集：可以收集个人自行公开或合法公开的个人信息，通过此途径收集的个人	策略并正确合规实施访问控制，防止违规篡改； ● 制定个人信息处理规则、流程，处理后不完整、不准确的信息进行去标识化处理，确保无法关联到特定自然人，不会利用处理后不完整、不准确的个人信息为关联用户提供鉴别、授权、公开等服务； ● 为敏感个人信息制定专门的个人信息处理规则，采用差异化强度更高的保护措施，如采用双因素鉴别方式对访问敏感个人信息的实体进行鉴别等； ● 对个人信息的全部处理操作记录日志，定期审计。	可对完整性进行验证。

措施建议	技术保障	相关密码技术应用
信息宜经过个人同意，个人明确拒绝的则不可收集； - 公开：个人信息被公开前应征得个人同意，声明公开的目的、途径、方式等； - 存储：身份管理系统掌握的个人信息应存储于境内，需跨境的应当遵守安全评估有关规定。 ● 敏感个人信息生命周期管理： - 在收集、处理敏感个人信息之前，应对处理目的和必要性做出评估和审批； - 确需处理敏感个人信息时，应对处理目的、必要性以及对个人权益的影响进行说明，应依据法律规范取得合规的个人的单独同意； - 应在敏感个人信息全生命周期采取严格防护措施； - 当未满十四周岁的未成年人进行身份注册时，应将全部需告知用户的信息告知该未成年人父母或者其他监护人，提供未成年人父母知晓和授权的权利； - 可为未成年人或失去行为能力等个人，提供身份管理服务代理功能，由其监护人进行代理； - 针对不满十四周岁未成年人个人信息制定专门的个人信息处理规则。		

6.3 个人信息跨境提供规则

《个人信息保护法》在第三章“个人信息跨境提供规则”中，澄清个人信息跨境传输规则。在沿用《网络安全法》监管思路的基础上，扩大了个人信息本地化存储的义务主体范围。具体的出境安全评估办法等问题还有待相关部门、行业出台相应文件进一步明确。

表 6-3 个人信息跨境部分措施建议

措施建议	技术保障	相关密码技术应用
● 身份管理系统中存储的个人信息原则上应存储于境内； ● 确需向境外提供个人信息的，应通过国家网信部门组织的安全评估；或按照国家网信部门的规定经专业	● 信息接口保护； ● 会话管理； ● 提供用户个人信息完整性校验机制，保障用户个人信息的	● 对跨境传输数据进行加密传输，密钥以安全的方式进行管理；

措施建议	技术保障	相关密码技术应用
机构进行个人信息保护认证;或按照国家网信部门制定的标准合同与境外接收方订立合同,约定双方的权利和义务;或满足法律、行政法规或者国家网信部门规定的其他条件。 ● 境外接收方处理个人信息的活动应达到我国法律、行政法规或者国家网信部门规定的条件; ● 确需向境外提供个人信息的,应告知个人境外接收方信息,包括但不限于:境外接收方的名称或者姓名、联系方式、处理目的、处理方式、个人信息的种类以及个人向境外接收方行使本法规定权利的方式和程序等事项,并取得个人的单独同意; ● 向外国司法或者执法机构提供存储于中华人民共和国境内的个人信息,应经中华人民共和国主管机关批准; ● 不得向限制或者禁止个人信息提供清单中的个人或组织提供个人信息;	完整性; ● 审查境外接收方对所需处理的个人信息制定访问控制策略并正确合规实施访问控制; ● 采用签名、加密、可信信道等技术保障个人信息跨境传输安全; ● 可对境外接收方处理个人信息的活动记录日志,并定期审计,保证其处理活动满足我国法律法规和合同要求。 ● 跨境业务失效、终止、到期后,境外第三方应返还个人信息或彻底删除,应对跨境个人信息接收方业务系统进行深度扫描,或通过第三方进行检测,确保个人信息已被删除。	● 提供个人信息完整性验证机制; ● 对跨境传输数据的摘要进行签名,防篡改; ● 采用 https 等加密信道传输; ● 会话密钥安全。

6.4 个人在个人信息处理活动中的权利

《个人信息保护法》从法律层面赋予了个人信息主体关于个人信息保护的相关权利,如个人对其个人信息的处理享有知情权、决定权,有权限制或者拒绝他人对其个人信息进行处理,有权查阅、复制其个人信息,有权请求将个人信息转移至其指定的个人信息处理者(可携带权),有权要求个人信息主体更正、补充、删除其个人信息。

表 6-4 个人权利部分措施建议

措施建议	技术保障	相关密码技术应用
● 采用上述主动告知个人信息主体、得到个人信息主体的授权同意等措施; ● 制定个人信息保护策略规则,并明确所采用的技术措施及标准规范,保障	● 信息接口保护 ● 会话管理 ● 防注入攻击 ● 采用可靠信道	● 对告知用户信息进行签名,防篡改; ● 用户同意信

措施建议	技术保障	相关密码技术应用
个人信息安全； ● 应提供用户了解、询问个人信息处理规则的途径，并针对用户的问题提出相应的解答； ● 应当建立便捷的个人行使权利的申请受理和处理机制，拒绝个人行使权利的请求的，应当说明理由； ● 身份管理系统应提供给用户查看其个人信息的途径； ● 身份管理系统应提供给用户复制其个人信息的途径； ● 身份管理系统应提供给用户将个人信息转移至合规个人信息处理者的途径； ● 身份管理系统应提供给用户更正个人信息的途径； ● 当用户要求时，身份管理系统应提供给用户补充个人信息的途径； ● 身份管理系统应提供给用户删除个人信息的途径； ● 合规删除处理目的已实现、无法实现或者为实现处理目的不再必要的个人信息； ● 个人信息处理者停止提供产品或者服务，或者保存期限已届满时，应合规删除个人信息； ● 用户撤回对个人信息处理的授权，应合规删除个人信息； ● 违法违规或违反约定行为时，应合规删除个人信息； ● 应提供当自然人死亡后，由其近亲属行使对个人信息的权利的途径，并按照死者生前要求，若无则按照其近亲属要求，对其个人信息进行处理。	● 提供用户个人信息完整性校验机制，保障用户个人信息的完整性； ● 针对查看、复制、转移等行为，应对个人进行身份鉴别，确认其为个人信息主体； ● 针对个人信息制定访问控制策略并正确合规实施访问控制，防止违规篡改； ● 必要时采用覆写或物理销毁等方式删除个人信息； ● 用户对个人信息的全部处理操作记录日志，定期审计。	息可溯源，防抵赖； ● 采用 https 等加密信道传输； ● 会话密钥安全； ● 采用密码技术对用户的个人信息进行加密，可对完整性进行验证。

6.5 个人信息处理者的义务

《个人信息保护法》中规定了个人信息处理者的义务，包括安全保障义务，合规审计义务，个人信息保护影响评估义务，安全事件通知义务，平台特殊义务等。个人信息处理者应从制度、管理、技术等多方位保障个人信息处理者的义务。

表 6-5 个人信息处理者义务部分措施建议

措施建议	技术保障	相关密码技术应用
● 制定内部管理制度和操作规程； ● 并定期对从业人员进行安全教育和培训； ● 制定并组织实施个人信息安全事件应急预案； ● 指定个人信息保护负责人，负责对个人信息处理活动以及采取的保护措施等进行监督； ● 在采用公开途径告知用户个人依法行使权利的方式和程序的公告中，公开个人信息保护负责人的联系方式，并将个人信息保护负责人的姓名、联系方式等报送履行个人信息保护职责的部门； ● 在境外进行分析评估境内自然人的个人信息处理者，应在境内设立专门机构或者指定代表，负责处理个人信息保护相关事务，并将有关机构的名称或者代表的姓名、联系方式等报送履行个人信息保护职责的部门； ● 定期对身份管理系统及个人信息处理者进行审计；重点审计敏感个人信息处理过程、利用个人信息进行自动化决策过程、委托处理或对外提供个人信息、公开个人信息、向境外提供个人信息、对个人权益有重大影响的个人信息处理活动等。 ● 当有敏感个人信息处理过程、利用个人信息进行自动化决策过程、委托处理或对外提供个人信息、公开个人信息、向境外提供个人信息、对个人权益有重大影响的个人信息处理等活动时，应在处理事前进行个人信息影响评估，评估内容包括但不限于： - 处理目的及其合法性、正当性、必要性； - 对个人权益影响及安全风险； - 采取的保护措施是否合法、有效、与风险程度相适应。 ● 个人信息保护影响评估报告和处理情况记录，以及审计结果至少保存三	● 信息接口保护 ● 会话管理 ● 防注入攻击 ● 采用可靠信道 ● 防止未经授权的访问以及个人信息泄露、篡改、丢失： ● 制定内部管理制度和操作规程； ● 对个人信息实行分类管理； ● 采取相应的加密、去标识化等安全技术措施； ● 提供用户个人信息完整性校验机制，保障用户个人信息的完整性； ● 针对查看、复制、转移等行为，应对个人进行身份鉴别，确认其为个人信息主体； ● 针对个人信息制定访问控制策略，合理确定个人信息处理的操作权限，并正确合规实施访问控制，防止违规篡改； ● 必要时采用覆写或物理销毁等方式删除个人信息； ● 个人信息处理者对个人信息的全部处理操作记录日志，定期审计； ● 研究开发和推	● 对告知用户信息进行签名，防篡改； ● 用户同意信息可溯源，防抵赖； ● 采用 https 等加密信道传输； ● 会话密钥安全； ● 可采用基于密码技术的加解密功能实现数据内容的机密性保护； ● 使用基于对称密码算法或密码杂凑算法的消息鉴别码机制、基于公钥密码算法的数字签名机制等密码技术实现数据内容的完整性保护； ● 使用基于对称密码算法或密码杂凑算法的消息鉴别码机制、基于公钥密码算法的数字签名机制、动态口令机制等密码技术实现传输数据的真实性保护； ● 使用基于公钥密码算法的数字签名机制等密码技术实现不可否认性。

措施建议	技术保障	相关密码技术应用
年。 ● 发生或者可能发生个人信息泄露、篡改、丢失的，个人信息处理者应当立即采取补救措施，并通知履行个人信息保护职责的部门和个人。通知应当包括下列事项： - 发生或者可能发生个人信息泄露、篡改、丢失的信息种类、原因和可能造成的危害； - 个人信息处理者采取的补救措施和个人可以采取的减轻危害的措施； - 个人信息处理者的联系方式。 ● 提供重要互联网平台服务、用户数量巨大、业务类型复杂的个人信息处理者，应当履行下列义务： - 建立健全个人信息保护合规制度体系，成立主要由外部成员组成的独立机构对个人信息保护情况进行监督； - 定期由第三方机构对个人信息措施的合规实施进行审计； - 遵循公开、公平、公正的原则，制定身份管理系统处理个人信息的规范和保护个人信息的义务； - 对严重违反法律、行政法规处理个人信息的平台内的产品或者服务提供者，停止提供服务； - 定期发布个人信息保护社会责任报告，接受社会监督。 ● 针对小型个人信息处理者、处理敏感个人信息以及人脸识别、人工智能等新技术、新应用，制定专门的合规的个人信息保护规则； ● 配合专业机构开展个人信息保护评估、认证服务，评估或审计后，应当按照要求采取措施，进行整改，消除隐患。	广应用安全、方便的电子身份认证技术。	

7. 总结

本报告主要研究国内外先进的个人信息保护技术、标准与政策，分析我国涉及身份管理功能的系统中个人信息面临的风险与保护措施现状。设计适用于网络身份服务方的个人信息安全保护技术框架和要求。并将密码技术要求融合其中，采用加密、脱敏、动态认证等方法对身份管理系统中维护的用户身份信息等个人信息进行处理，以达到保护关键信息的目的。依据实际安全需求和密码技术应用需求，并结合我国法律法规要求，提出个人信息在收集、传输、存储、使用、删除等全生命周期的安全保护技术建议。为监管方取证用户行踪提供入口，也能为用户信息删除提供接口服务。该报告研究身份管理系统的个人信息安全保护技术，保障网络中个人信息的安全性，为用户使用网络身份服务、身份服务提供方部署安全措施提供指导。

参考文献

[1]《中华人民共和国网络安全法》
http://www.npc.gov.cn/wxzl/gongbao/2017-02/20/content_2007531.htm

[2] N. R. Adam, J. C. Worthmann. Security-control methods for statistical databases: a comparative study[J]. Acm Computing Surveys, 1989, 21(4):515－556

[3] Closing the data privacy gap: Protecting sensitive data in non production environments[EB/OL]. IBM Co., 2009.

[4] Data Masking Best Practices CamouFlage an Oracle White Paper[EB/OL]. Oracle Co., 2010.

[5] Data Masking White Paper: What You Need to Know [EB/OL]. A Net 2000 Ltd . 2010

[6] G. Sarada, N. Abitha, G. Manikandan, et al. A few new approaches for data masking[M]. 2015, 1－4

[7] V. A. Gujjary, A. Saxena. A neural network approach for data masking[J]. Neurocomputing, 2011, 74(9):1497－1501

[8] S. Vijayarani, A. Tamilarasi. An efficient masking technique for sensitive data protection[M]. 2011, 1245 － 1249

[9] Y. Zhou, T. A. Louis. A smoothing approach for masking spatial data[J]. Annals of Applied Statistics, 2010, 4(3):1451－1475

[10] M. P. Armstrong, G. Rushton, D. L. Zimmerman. Geographically masking health data to preserve confidentiality[J]. Statistics in Medicine, 1999, 18(5):497－525

[11] M. Li, Z. Liu, C. Jia, et al. Data Masking Generic Model[M]. 2013, 724－727

[12] B. Liver, K. Tice. Privacy Application Infrastructure: Confidential Data Masking[M]. 2009, 324－332

[13] A. I. Baranchikov, A. Y. Gromov, V. S. Gurov, et al. The technique of dynamic data masking in information systems[M]. 2016, 473－476

[14] 冯登国等. 大数据安全与隐私保护[M]. 北京:清华大学出版社,2018.

[15] Kasiviswanathan S P , Lee H K , Nissim K , et al. What Can We Learn Privately?[J]. Proc.ieee Annual IEEE Symp.on Foundations of Computer Science, 2008, 40(3):793-826.

[16] Erlingsson Ú, Pihur V, Korolova A. RAPPOR: Randomized Aggregatable Privacy- Preserving Ordinal Response[C]// ACM SIGSAC Conference on Computer and Communications Security, ACM, 2014, 1054－1067.

[17] Bassily R, Smith A. Local, Private, Efficient Protocols for Succinct Histograms [C]//ACM Symposium on the Theory of Computing, ACM, 2015: 127-135

[18] Abbas A , Hidayet A , Selcuk U A , et al. A Survey on Homomorphic

Encryption Schemes: Theory and Implementation[J]. Acm Computing Surveys, 2017,51(4):1-35.
[19]亚马逊云安全服务
https://aws.amazon.com/cn/products/security/?nc=sn&loc=2
[20] Auth0 Data Privacy and Compliance.
https://auth0.com/docs/secure/data-privacy-and-compliance
[21]McAfee 数据保护产品
https://www.mcafee.com/enterprise/zh-cn/products/data-protection-products.html
[22] 阿里云数据安全解决方案
https://cn.aliyun.com/solution/security/datasecurity
[23]微软数据信息保护白皮书
https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE4r1mQ
[24] 《中华人民共和国个人信息保护法》
http://www.gov.cn/xinwen/2021-08/20/content_5632486.htm
[25] 《中华人民共和国数据安全法》
http://www.gov.cn/xinwen/2021-06/11/content_5616919.htm
[26] Ahmed A , I Traoré. Detecting Computer Intrusions Using Behavioral Biometrics[C]// Conference on Privacy. DBLP, 2005.
[27] User re-authentication via mouse movements[C]// Workshop on Visualization and Data Mining for Computer Security (VizSEC/DMSEC 2004), 29 October 2004, Washington DC, USA. ACM, 2004.
[28] https://www.bsigroup.com/globalassets/localfiles/en-gb/isoiec-27701-privacy-information-management/resources/mapping-iso-iec-27701.pdf
[29] Sluganovic I , Roeschlin M , Rasmussen K B , et al. Using Reflexive Eye Movements For Fast Challenge-Response Authentication[C]// the 2016 ACM SIGSAC Conference. ACM, 2016.
[30] Zifeng, Wu, Yongzhen, et al. A Comprehensive Study on Cross-View Gait Based Human Identification with Deep CNNs.[J]. IEEE transactions on pattern analysis and machine intelligence, 2016.
[31] Yu S , Chen H , Reyes E , et al. GaitGAN: Invariant Gait Feature Extraction Using Generative Adversarial Networks[C]// Computer Vision & Pattern Recognition Workshops. IEEE, 2017.
[32] Martinez-Diaz M , Fierrez J , Galbally J . Graphical Password-Based User Authentication With Free-Form Doodles[J]. IEEE Transactions on Human-Machine Systems, 2016, 46(4):607-614.
[33] Nader J , Alsadoon A , Prasad P W C , et al. Designing Touch-Based Hybrid Authentication Method for Smartphones[J]. Procedia Computer Science, 2015, 70:198-204.
[34] Antal M , Bokor Z , Szabo L Z . Information revealed from scrolling interactions on mobile devices[J]. Pattern Recognition Letters, 2015, 56(apr.15):7-13.