第 5 章

数智安全技术基础

本章重点介绍数智安全的基础技术，包括密码技术、身份管理技术、访问控制技术、日志及安全审计技术等。这些基础安全技术的典型应用是 4A，即认证（Authentication）、授权（Authorization）、账号（Account）及审计（Audit），是身份和访问管理的四个基本环节。它们构成了一个完整的闭环，确保只有授权的用户才能访问系统或资源，并对用户的操作进行监控和审计。

5.1 密码技术

作为信息安全的核心技术，密码技术被广泛应用于数据加密、消息验证、身份鉴别、访问控制、电子印章、隐私计算、数字防伪、数字版权、区块链等多个应用领域。本节介绍密码学的发展过程、研究内容、基本概念及范畴等内容，以理解密码技术的基本原理。此外，本节还简要介绍密码技术标准化相关内容。

5.1.1 密码技术概述

1. 密码学的发展过程

密码学是指研究密码理论与技术的专门学科。从几千年前神秘性的字谜开始，密码学的发展已有数千年，广泛应用于军事、政治、商业和生活的方方面面。

现代密码学则起源于 1949 年香农（数学家、信息论创始人）发表的论文《保密系统的通信理论》，该论文科学地阐述了保密系统的设计、分析及评价的原理，为密码学奠定了理论基础，密码学从此开始成为一门科学。

自 20 世纪六七十年代开始，现代密码学研究走向公开领域并开始高速发展。1976 年，美国密码学家 Diffie 和 Hellma 发表了论文《密码学的新方向》（New Direction in Cryptography），正式提出了公钥密码体制（Public Key Cryptosystem），允许同时公开密码算法及加密公钥，解决了不可靠信道下密钥交换的难题，改变了人类几千年来的单钥密码体制。如今，密码技术几乎被应用在所有的信息技术产品中，人们生活中常见的身份证、门禁卡、银行卡、电子发票、数字人民币、虚拟货币等都采用了密码技术，密码学已经成为信息化、数字化、智能化发展不可或缺的基石。

2. 密码学的研究内容

根据 GB/T 25069—2022《信息安全技术 术语》的定义，密码学（Cryptology）是研究密码与密码活动本质和规律，指导密码实践的学科，主要探索密码的编制、破译以及管理的一般规律。密码学包括密码编码学（Cryptography）和密码分析学（Cryptanalysis）两部分。密

码编码学主要研究信息的编码，构建各种安全有效的密码算法和协议，用于消息的加密、认证等方面；密码分析学是研究破译密码获得消息，或对消息进行伪造。

密码学中常见的概念包括：明文、密文、加密、解密、密码算法、密钥等，下面进行简要介绍。

(1) 明文(plaintext)：是未加密的原始数据，一般用小写字母 m 或 p 表示。全部的明文集合称为明文空间，一般用大写字母 M 或 P 表示。在信息系统中，明文通常是一段文本，也可以是一个文件，也可以是图片、音视频、网络比特流等。

(2) 密文(ciphertext)：是采用密码算法，将明文变换后的数据，一般用小写字母 c 表示。全部的密文集合被称为密文空间，一般用大写字母 C 表示。

(3) 密码算法(cryptographic algorithm)：是描述密码处理过程的算法。

(4) 加密算法(encryption algorithm)：是将明文转换为密文的算法，一般用大写字母 E 表示加密算法。

(5) 解密算法(decryption algorithm)：是将密文转换为明文的算法，一般用大写字母 D 表示解密算法。

(6) 加密(encipherment/encryption)：是对数据进行密码变换以产生密文的过程，可以用 $E(p)$表示通过加密算法 E 对明文 p 进行加密过程。

(7) 解密(decipherment/decryption)：是与加密过程对应的逆过程，可以用 $D(c)$表示通过解密算法 D 对密文 c 进行解密过程。

(8) 密钥(key)：是在加密或解密算法中实施控制的参数，一般用小写字母 k 表示。全部的密钥集合称为密钥空间，一般用大写字母 K 表示。根据密钥是用在加密算法还是解密算法过程中，可以分为加密密钥和解密密钥。

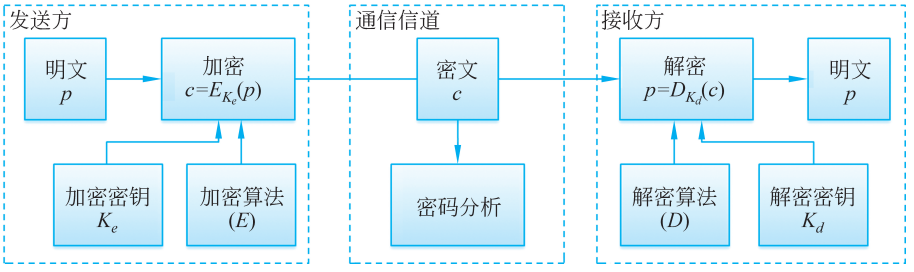

图 5.1　加密与解密过程示意图

图 5.1 描述了加密与解密的基本过程。加密过程可表示为：$c=E_{K_e}(p)$，即发送方的加密过程是使用加密算法(E)基于加密密钥(K_e)将明文(p)转换为密文(c)；密文(c)经通信信道传输给接收方，期间可能受到攻击者的窃听或干扰，攻击者通过密码分析过程试图还原明文或篡改信息；解密过程可表示为：$p=D_{K_d}(c)$，即接收方的解密过程是使用解密算法(D)基于解密密钥(K_d)将密文(c)转换为明文(p)。

3. 密码技术的保护作用

密码技术在数智安全保护中是不可或缺的，能够提供保密性、完整性、真实性、不可否认性等安全保护能力。

(1) 在保密性方面，可通过对称加密、非对称加密、数字信封等密码技术，对抗网络窃听、数据窃取、敏感信息泄露等威胁。

(2) 在完整性方面，可通过哈希函数、消息认证码、数据加密、数字签名等密码技术，对

抗数据篡改、数据破坏、重放攻击等威胁。

(3) 在真实性方面，可通过口令和共享密钥、数字证书、数字签名等密码技术，对抗身份假冒等威胁。

(4) 在不可否认性方面，可通过数字签名等密码技术，对抗数据发送及接收中的否认问题。

5.1.2 密码技术原理

1. 数据加密技术

1) 对称密码算法

对称密码算法的基本特征是用于加密和解密的密钥是一样的，或实质上等同，即从其中一个容易推导出另一个。因此，对称密码算法也被称为对称密钥算法、秘密密钥算法或单钥密码算法。所使用的密钥被称为“对称密钥”。

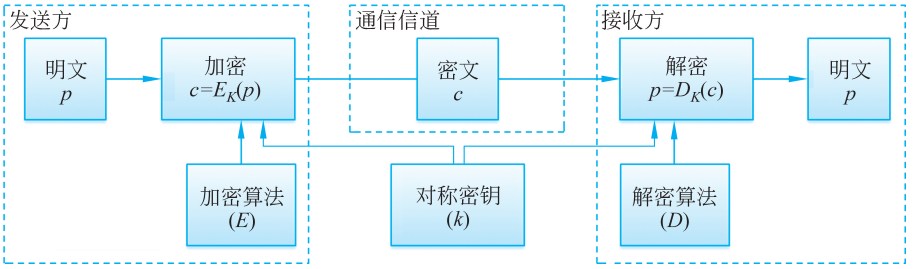

图 5.2 对称密码算法使用示意图

图 5.2 描述了对称密码算法的加密和解密过程，与图 5.1 所描述的过程区别在于加密过程 $c=E_k(p)$ 与解密 $p=D_k(c)$ 所使用的密钥均为对称密钥(k)，且对称密钥(k)应通过安全的方式发送给接收方。

典型的对称密码算法有 SM4 算法、数据加密算法(Data Encryption Standard，DES)、高级加密标准(Advanced Encryption Standard，AES)、国际数据加密算法(International Data Encryption Algorithm，IDEA)等。

对称密码算法可以分为序列密码算法及分组密码算法两种，适用于不同加密需求的应用场景。其中，序列密码算法也称为流密码算法，是将明文消息按字符逐位地加密。序列密码算法适用于流式数据加密，如网络音视频通信。我国的祖冲之算法(ZUC)属于序列密码算法。分组密码算法是将明文按组分成固定长度的块，用同一密钥和算法对每一块加密，每个输入块加密后产生得到一个固定长度的密文输出块。常见的分组密码算法有 SM4、DES、IDEA 等。

2) 非对称密码算法

非对称密码体制又称双钥或公钥密码体制，其加密密钥和解密密钥不同，从一个很难推出另一个。其中，一个可以公开，称为公开密钥(public key)，简称公钥；另一个必须保密，称为私有密钥(private key)，简称私钥。一对相关联的公钥和私钥被称为非对称密钥对。典型的非对称密码算法有 SM2、RSA、ECC 和 ElGamal 等。

图 5.3 描述了非对称密码算法的加密和解密过程，其加密过程 $c=E_{PK}(p)$ 所使用的密钥为公开密钥(PK)，解密过程 $p=D_{SK}(c)$ 所使用的密钥为私有密钥(SK)，其中公开密钥(PK)可通过不可靠的通信信道传输。

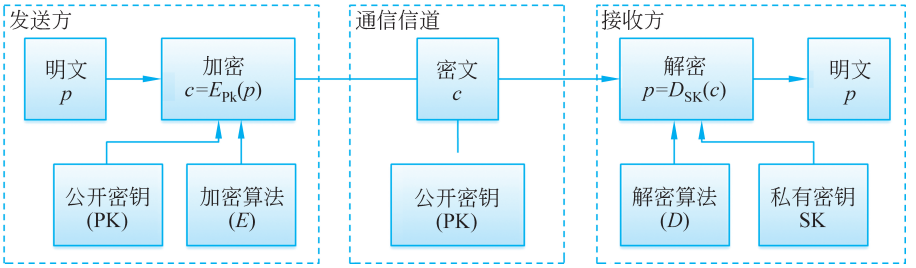

图 5.3　非对称密码算法使用示意图

与对称加密相比，公钥加密的速度较慢，一般适用于短数据的加密，如用于共享密钥交换等场景。

2. 信息认证技术

认证技术主要起到鉴别和确认的作用，一般被用于验证主体的真实性、数据的完整性、访问及操作的不可否认性等。本节将介绍与之相关的杂凑函数、消息鉴别码算法及数字签名等密码技术相关内容。

1）杂凑函数

杂凑函数也称为散列函数、哈希函数。杂凑函数的作用是接收一个消息作为输入，产生固定长度的字串。这个固定长度的字串被称为散列值、哈希值或摘要值。杂凑函数的特点是能够应用到任意长度的数据上，并且能够生成大小固定的输出。对于任意给定的 x，杂凑函数 $H(x)$ 的计算相对简单，易于软硬件实现。安全的杂凑函数需要满足以下性质：

（1）单向性：对任意给定的码 h，寻求 x 使得 $H(x)=h$ 在计算上是不可行的；

（2）弱抗碰撞性：任意给定分组 x，寻求不等于 x 的 y，使得 $H(y)=H(x)$ 在计算上不可行；

（3）强抗碰撞性：寻求任何的 (x,y) 对，使得 $H(x)=H(y)$ 在计算上不可行。

典型杂凑函数有 SM3、MD5、SHA-1 等。

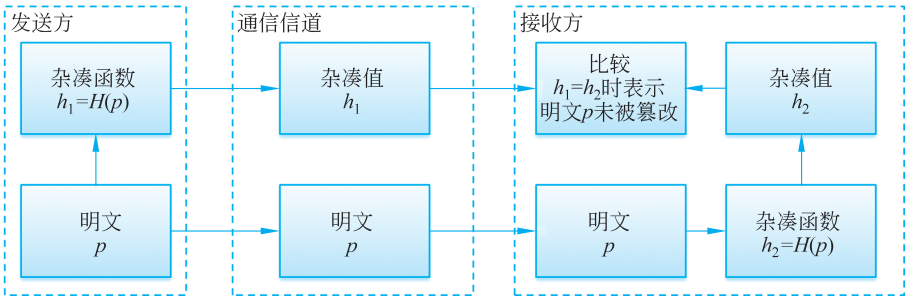

图 5.4　杂凑算法使用示意图

图 5.4 描述了杂凑函数的常见使用过程，其过程 $h_1=H(p)$ 表示发送方通过杂凑函数 (H) 以明文 (p) 作为输入，得到杂凑值 (h_1)；明文 (p) 和杂凑值 (h_1) 经通信信道传输给接收方；接收方通过同样的杂凑函数 (H) 将接收的明文 (p) 作为输入，得到杂凑值 (h_2)；如果杂凑值 (h_1) 与杂凑值 (h_2) 相等，则认为明文 (p) 在传输过程中未被篡改。需要注意的是，杂凑函数 (H) 和明文 (p) 在通信信道传输过程中存在被同时篡改的可能性。

2）消息鉴别码

消息鉴别码（Message Authentication Code，MAC）也被称为消息认证码，它也是将一个

任意长度的消息变换成一个固定长度的、较短的字串。和杂凑函数不同的是，消息鉴别码在计算过程中需要使用一个密钥来生成字串，而验证方在验证消息鉴别码时也需要知道该密钥才能进行计算。

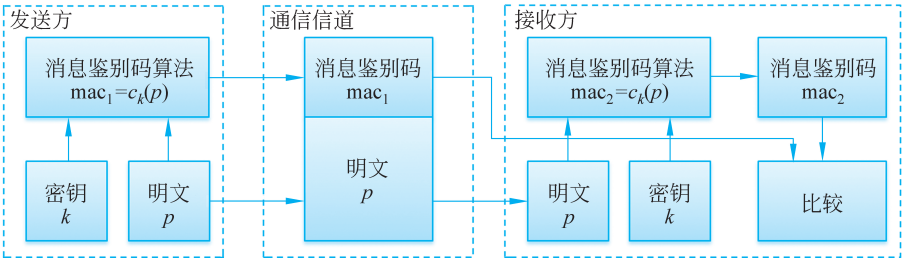

图 5.5　消息鉴别码算法使用示意图

图 5.5 描述了消息鉴别码算法的常见使用过程，其过程$mac_1=C_k(p)$表示发送方通过消息鉴别码算法(C)基于密钥(k)以明文(p)作为输入，得到消息鉴别码(mac_1)；明文(p)和消息鉴别码(mac_1)经通信信道传输给接收方；接收方通过同样的消息鉴别码算法(C)基于相同的密钥(k)将接收的明文(p)作为输入，得到消息鉴别码(mac_2)；如果消息鉴别码(mac_1)与消息鉴别码(mac_2)相等，则认为明文(p)在传输过程中未被篡改。其使用的密钥(k)应通过安全的方式发送给接收方。

3) 数字签名

数字签名(digital signature)，是附加在数据单元上的一些数据，或是对数据单元做密码变换，这种附加数据或密码变换被数据单元的接收者用以确认数据单元的来源和完整性，达到保护数据，防止被人(例如接收者)伪造的目的。可以看作是以数字化形式进行的“签名”，以替代纸质签名、印章等。其基本要求是：签名与所签原始数据的“绑定”，不可篡改且容易验证；签名的不可否认性及不可伪造性。数字签名可保护数据的完整性、不可否认性、真实性，且具备易认证、不可伪造、不可抵赖、不可篡改等特点。

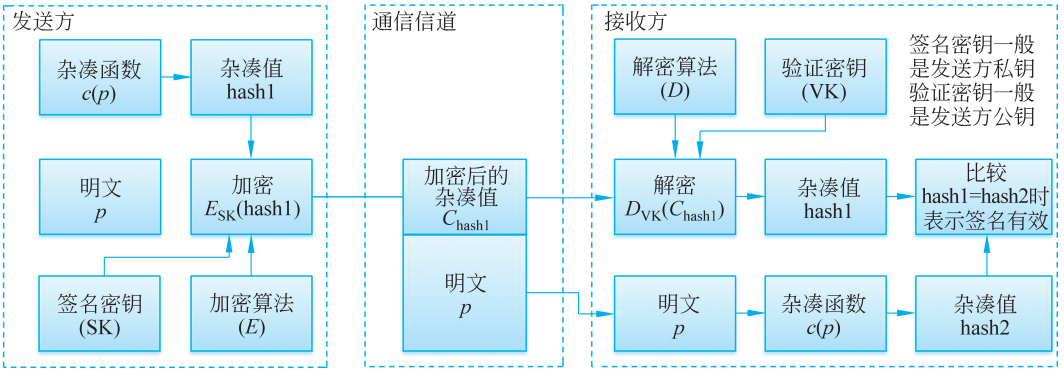

图 5.6　数字签名过程示意图

图 5.6 描述了数字签名的常见使用过程，其加密过程$E_{SK}(hash1)$表示，发送方首先使用杂凑函数(C)以明文(p)作为输入，得到杂凑值(hash1)，并进一步使用加密算法(E)基于签名密钥(SK)以杂凑值($hash_1$)作为输入，得到加密后的杂凑值(C_{hash1})；明文(p)和加密后的杂凑值(C_{hash1})经通信信道传输给接收方；接收方的解密过程$D_{VK}(C_{hash1})$表示，使用解密算法(D)基于验证密钥(VK)将接收到的加密后的杂凑值(C_{hash1})作为输入，得到解密后的

杂凑值(hash1),并使用同样的杂凑函数(C)以明文(p)作为输入,得到杂凑值(hash2),如果杂凑值(hash1)与杂凑值(hash2)相等,则表示签名有效。其中,签名密钥(SK)一般是发送方私钥,验证密钥(VK)一般是发送方公钥。

在数智应用中,尤其是电子商务中通信双方相互之间传递消息时,不可否认性非常重要,它一方面要防止发送方否认曾经发送过的消息,另一方面还要防止接收方否认曾经接收过的消息,以避免通信双方可能存在欺骗和抵赖,数字签名是解决这类问题的有效方法。

3. 数字证书与公钥基础设施

1) 数字证书

数字证书的概念是 1978 年由 Kohnfelder 提出的,也称为公钥证书,其内容包含了用户身份信息、公钥,并以 CA(可信第三方认证机构)数字签名形式确保用户信息及公钥的真实性。数字证书和一对公、私钥相对应,公钥以明文的形式放到数字证书中,私钥为拥有者秘密掌握。CA 确保数字证书中信息的真实性,可以作为终端实体的身份证明。在电子商务和网络信息交流中,数字证书常用来解决相互间的信任问题。数字证书和生活中的身份证、驾驶证等证件的作用相似,都是用来证明身份的,因此数字证书会记录用户身份关联的信息,如:证书所有人名称等。按照 X.509V3 数字证书格式标准,数字证书一般包含:证书的版本信息、唯一序列号、所使用的签名算法、发行机构名称、有效期、所有人名称、所有人的公开密钥、发行者对证书的签名等内容。我国数字证书格式国家标准 GB/T 20518—2018《信息安全技术 公钥基础设施 数字证书格式》兼容于 X.509V3 标准,同时要求支持使用我国的密码算法。

如图 5.7 所示,数字证书生成过程可简要描述为:①证书申请者将主体身份信息、主体公钥提交给权威机构(CA);②CA 在信息中附加上自身名称,并使用自己的私钥对混合的信息进行数字签名,将签名与主体信息生成一份数字证书,并发布到 CA 的目录服务器上;③证书申请者和其他用户均可通过 CA 的目录服务器查询和获取证书;④只有证书申请者具备公钥对应的私钥,可以使用私钥进行签名;⑤其他用户可以使用证书附带的公钥验证签名是否由证书真正的所有者生成,其他人无法伪造。

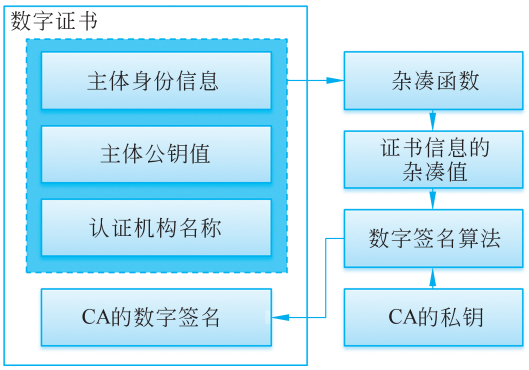

图 5.7　数字证书及其生成过程

数字证书的优点如下。

(1) 数字证书中的公钥不需要保密,其管理、保护成本较低。

(2) 证书本身不易伪造、容易验证,易于通过 CA 的数字签名验证证书的完整性和真实性。

(3) 证书易于使用,证书除了使用 CA 认证中心的目录服务获取外,还可以通过任何文件交换形式传递,如电子邮件、IM 等,且可离线使用,无需 CA 等第三方参与,对使用场景的适应能力强。

(4) 结合 PKI/CA 体系时,CA 认证中心会对证书申请者进行一定程度的身份审核,证书对应身份的可靠性会更有保障(但不是绝对的);用户间只需通过 CA 就可获取其他用户证书,简化了用户间证书交换的难度。

2) 公钥基础设施

公钥基础设施(Public Key Infrastructure,PKI):是基于公钥密码技术,可用于提供保密性、完整性、真实性及抗抵赖性等安全服务的基础设施。PKI 主要用来解决大规模网络中的网络信任问题,即通过分发和管理数字证书,确保网络中各主体的身份真实性和可验证性。因此,PKI 的主要功能包括了数字证书的生成、管理、存储、分发和撤销等。

公钥基础设施能够用来满足网络空间身份真实性、数据完整性及行为不可抵赖性等安全需求,在网上银行、电子商务、电子政务、互联网金融等多种场景有广泛应用。随着物联网、车联网、工业互联网及各类智能体的普及,机器身份管理成为热点话题,PKI 也可为这些设备、程序、组件提供身份验证及数据安全保护等能力。

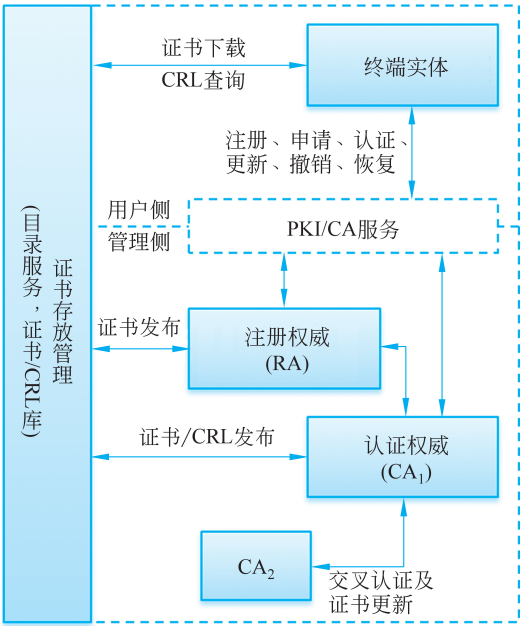

图 5.8 公钥基础设施架构

如图 5.8 所示,公钥基础设施一般由终端实体(证书持有者和应用程序)、注册权威(Registration Authority,RA)、认证权威(Certificate Authority,CA)、证书存放管理(目录服务,证书/CRL 库)等构成。其中:

(1) 终端实体(证书持有者和应用程序):可以是人、设备、进程等,是证书的最终用户和拥有者,拥有公私密钥对和相应公钥。

(2) 注册权威(RA):RA 又称证书注册中心,是数字证书的申请、审核和注册中心,同时也是 CA 认证机构的延伸。在逻辑上 RA 和 CA 是一个整体,主要负责提供证书注册、审核以及发证功能。

(3) 认证权威(CA)：CA 是证书签发权威机构，也称数字证书管理中心，它作为 PKI 管理实体和服务的提供者，管理用户数字证书的生成、发放、更新和撤销等工作。

(4) 证书存放管理(目录服务，证书/CRL 库)：一般通过轻量级目录协议(LDAP)及证书撤销列表(Certificate Revocation List，CRL，也称“证书黑名单”)等方法，负责证书存放管理，提供证书保存、修改、删除和获取等功能。

5.1.3　密码技术标准化

1. 国产密码技术概述

国密算法是国产密码算法的简称，是指我国国家密码管理部门认定的、我国自主研发的密码算法，主要包括对称加密算法、非对称加密算法、杂凑算法等多种密码算法。

对于涉及重要数据或重要应用的信息系统，应当使用国密算法来保护。根据《中华人民共和国密码法》第二十六条中规定，“涉及国家安全、国计民生、社会公共利益的商用密码产品，应当依法列入网络关键设备和网络安全专用产品目录，由具备资格的机构检测认证合格后，方可销售或者提供”；第二十七条对“商用密码应用安全性评估”也给出了明确要求。《关键信息基础设施安全保护条例》第五十条也规定“关键信息基础设施中的密码使用和管理，还应当遵守相关法律、行政法规的规定”。

我国已形成以 SM1、SM2、SM3、SM4、SM7、SM9、ZUC 等为代表的国产商用密码技术体系。

SM1 算法是对称密码算法，分组长度为 128 位，密钥长度也为 128 位。

SM2 算法是非对称加密算法，其采用 ECC 椭圆曲线密码机制，可以用来实现数字签名、密钥交换以及数据加密应用。

SM3 算法是杂凑算法，适用于应用中的数字签名和验证、消息鉴别码的生成与验证以及随机数的生成，可满足多种密码应用的安全需求。

SM4 算法是国产对称密码算法，属于分组算法。该算法的分组长度为 128 位，密钥长度为 128 位。

SM7 算法是一种分组密码算法，分组长度为 128 位，密钥长度为 128 位。SM7 的算法文本目前没有公开发布。

SM9 是基于一种非对称密码算法。和 SM2 算法不同的是，SM9 算法是一种基于标识的密码算法，即可以直接使用用户的标识(如邮件地址、手机号码、身份证号等)作为公钥。

ZUC 祖冲之算法是一种序列密码算法，也是一种对称加密算法，该算法已经用于 3G、4G 等无线通信领域。

2. 密码标准体系框架

我国密码标准体系由技术维、管理维和应用维 3 个维度刻画，如图 5.9 所示。

其中，技术维包含密码基础类标准、基础设施类标准、密码产品类标准、应用支撑类标准、密码应用类标准、密码检测类标准和密码管理类标准 7 大类密码标准，这 7 类标准的相互关系如图 5.10 所示。

管理维上，我国密码标准可以分为国家标准、行业标准和团体标准 3 种类型；应用维从密码应用领域的视角来刻画密码标准体系。“应用领域”既包括不同的社会行业，如金融、电力、交通等，也包括不同的应用场景，如物联网、云计算等。

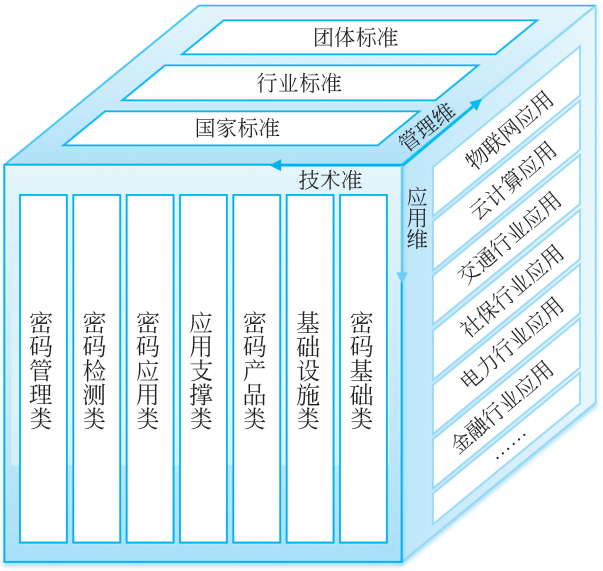

图 5.9　密码标准体系框架

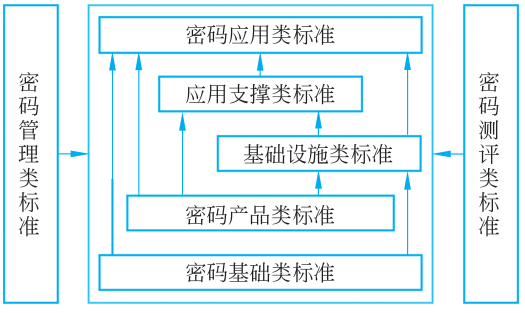

图 5.10　密码标准体系中的技术维

3. 密码算法标准国际化

我国积极推进商用密码算法 SM2、SM3、SM4、SM9、ZUC 等纳入国际标准。其中：

2011 年 9 月，我国设计的祖冲之密码算法（ZUC）被批准成为新一代宽带无线移动通信系统（LTE）国际标准，即 4G 的国际标准。

2017 年 11 月，我国 SM2 和 SM9 数字签名算法正式成为国际标准 ISO/IEC14888-3/AMD1《信息安全技术 带附录的数字签名 第 3 部分：基于离散对数的机制-补篇 1》的内容，由 ISO 正式发布。

2018 年 10 月，我国 SM3 杂凑密码算法成为 ISO/IEC10118-3：2018《信息安全技术 杂凑函数第 3 部分：专用杂凑函数》的内容，由 ISO 正式发布。

2018 年 11 月，作为补篇纳入国际标准的 SM2/SM9 数字签名算法，以正文形式随 ISO/IEC14888-3：2018《信息安全技术 带附录的数字签名 第 3 部分：基于离散对数的机制》最新一版发布。

2020 年 4 月，我国 ZUC 序列密码算法正式成为国际标准 ISO/IEC18033-4/AMD1《信息技术 安全技术 加密算法 第 4 部分：序列算法-补篇 1》的内容，由 ISO 正式发布。

2021 年 2 月，我国 SM9 标识加密算法作为国际标准 ISO/IEC18033-5：2015/AMD1：

2021《信息技术 安全技术 加密算法 第 5 部分：基于标识的密码补篇 1：SM9》，由 ISO 正式发布。

2021 年 6 月，我国 SM4 分组密码算法作为国际标准 ISO/IEC 18033-3：2010/AMD1：2021《信息技术 安全技术 加密算法 第 3 部分：分组密码补篇 1：SM4》，由 ISO 正式发布。

2021 年 10 月，我国 SM9 密钥协商协议正式成为国际标准 ISO/IEC11770-3：2021《信息安全 密钥管理 第 3 部分：使用非对称密码技术的机制》的内容，由 ISO 正式发布。

5.2　身份管理技术

身份管理也是网络安全中最基本、最常见的安全技术之一，是对各类 IT 资源、数据保护的首要方法。而其中身份管理作为主体访问客体的前提，其安全性直接影响信息系统的安全性。本节内容将介绍身份管理技术的基本概念、原理及主要应用场景。

5.2.1　身份管理概述

1. 身份管理的基本概念

身份(identity)是与某一实体相关的一组属性。需要说明的是，这里的实体并不特指人，而是包括人在内的软件、硬件、智能体等任何参与访问过程的实体。在网络中，一个实体可能同时具备多个身份，而多个实体也可能共同拥有同一个身份，如图 5.11 所示。

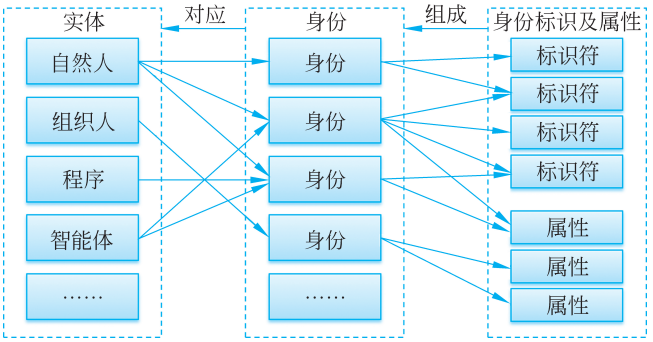

图 5.11　实体、身份及标识的关系示意图

身份的本质是对某一实体的映射或关联，表明是谁、具备哪些特征，而这个映射关系是靠标识、特征描述等信息与实体进行关联的。在生活中的自然人的身份常用如下标识：公民户籍信息、公民身份号码、护照、驾照等，也可以是电话号、银行卡号、社会保险号、车牌号、个人生物特征信息等。

随着数智化的发展，在网络空间中所使用的身份标识可被称为“数字身份”或“网络身份”。数字身份(digital identity)是以数字代码表达的身份，可在网络空间中用于识别和查询。数字身份有助于大幅提高整体社会效率、释放数字经济潜力和价值。常见的数字身份如：网络账号、电子邮件地址、互联网协议(IP)地址号、设备识别码、网络通用资源定位符(URL)等。随着大数据及人工智能的发展，身份标识已经不局限于以上的内容，通过大数据分析及机器学习等方式，可以通过网络操作行为特征、多信息关联等形式推定实体身份，一方面提高了人工智能的识别能力，但另一方面也对实体身份及信息的保护带来了难度。

数字身份的管理方式大致分为中心化数字身份、联盟式数字身份、分布式数字身份三种