用户数据及信息保护方案解析

第一章 背景描述

近年来，随着 5G 网络以及物联网设备的日益普及。在加快信息流动的同时也带来了更多的信息安全问题。

2021 年 3 月 15 日，央视 315 曝光三个涉及个人信息安全案例：商家安装摄像头捕捉记录顾客人脸信息，多门店共享并进行综合报价；智联招聘、猎聘等平台简历给钱就可随意下载，大量简历流入黑市；许多针对老年人开发的手机清理 App 背地里不断获取手机信息，并推送带有欺骗套路的内容。

国家相关部门在信息安全方面也出台了多项法规及规定。2017 年以来，国家质检总局和标准化委员会、国家市监总局和国家标准化委员会和全国信息安全标准化技术委员会陆续发布了一系列与个人信息和数据保护相关的国家标准和指南。

其中，国家质检总局和标准化委员会发布的文件共有 10 部，目前均处于生效状态，内容涵盖了云计算服务、移动智能终端个人信息保护、大数据服务、移动终端、移动互联网应用服务器、电子政务移动办公系统、网站身份和系统、网站可信标识等领域中的信息安全技术要求。

国家市监总局和国家标准化委员会颁布的文件共有 21 部，目前均处于生效状态，内容涵盖了金融信息服务、公民网络电子身份标识、电子邮件、射频识别等信息系统的安全运维管理、网络安全检测、网络安全等级保护测试评估及管理、智能卡、物联网等领域中的信息安全技术要求。

全国信息安全标准化技术委员会颁布的文件共有 19 部，目前均处于未生效状态，内容涵盖了网络可信评估、移动应用网络安全评价、金融信息保护、大数据安全管理、公民网络电子身份标识、数据交易服务、数据出境安全评估、网络安全等级保护定级、信息安全风险评估、个人信息安全影响评估、恶意事件预防和处理、关键信息基础设施网络安全保护、网络安全设计和实现、漏洞分类分级、信息安全管理体系审核、政务信息共享、个人信息安全等领域中的信息安全技术要求。

从产品及用户角度，我们需要更多的关注在信息安全方面的投入，以获得更加安全的产品。

第二章 风险分析

2.1 身份认证风险

在物联网系统中， 身份认证包括设备之间的认证、 设备与用户、 设备与平台之间的认证。 传统的身份认证方式中， 普通的用户名和密码认证方式，无法避免弱口令、 撞库攻击、 字典攻击等问题， 大多数用户也没有定期更改密码的习惯；而采用数字证书认证虽然安全， 但使用繁琐， 在终端设备上安装控件或 KEY 驱动和管理程序等， 使用极其不便。 下图是现行的各种身份认证方式存在的弊端分析。

2.2 数据传输风险

物联网系统中数据传输过程中， 一般都是明文形式传输， 尽管有些采取了密钥加密的安全措施， 但通常采用的是对称算法， 加密解密为同一把密钥， 一旦密钥被破解或被内部人员泄露， 灾难将是延续性的， 存在巨大的安全隐患； 若是使用低强度密钥进行加密， 依靠目前发达的计算条件， 通过穷举等方式存在破解的可能， 例如 DES 算法已经证明被破解。在物联网系统中的智能摄像头、 智能家居等应用环境中， 海量终端设备， 网络环境复杂，安全漏洞较多， 在业务交互过程中， 一旦网络被监听， 数据即被窃取或篡改， 造成敏感或重要数据泄露， 将带来巨大的灾难。

2.3 数据存储风险

物联网数据存储风险主要表现在云服务端和终端设备两个方面。 在物联网云服务端， 数据集中和新技术的采用是产生云存储安全问题的根据， 由于云计算的技术特性， 多租户、 资源共享、 分布式存储等这些因素加大了数据保护的难度，增大了数据被滥用和受攻击的可能， 因此云端数据安全保护是必须解决的问题；在终端设备， 由于大量的感知设备暴露在复杂的社会环境中，受攻击的风险大大增加， 遭到暴力的攻击和破解， 终端设备上存储的采集数据、 用户数据、 业务数据一旦被窃取， 将造成用户隐私数据泄露， 带来严重影响。

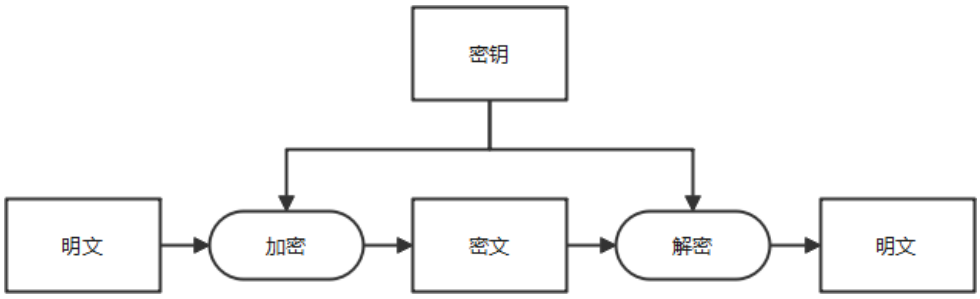
第三章 安全加密介绍

3.1 加密算法

对称加密算法也叫共享密钥加密算法、单密钥加密算法。采用单密钥的加密方法，密钥同时用作信息的加密和解密，即解密算法为加密算法的逆算法。因此在知道了加密算法后也就知道了解密算法。

对称加密算法顾名思义，其解密过程是加密过程的逆过程，在加密及解密过程中使用的是同一个密钥。

对称加密的流程如下图：

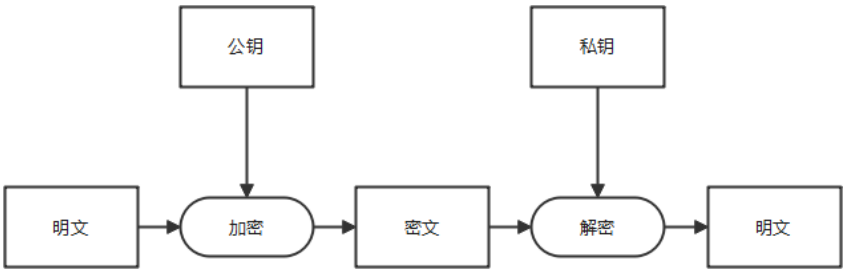


衡量对称加密算法优劣的取决于其密钥的长度。密钥越长，破解需要测试的密钥就越多，破解这种算法的难度就越大。其安全性取决于是否有未经授权的人获得了对称密钥。

常用对称加密算法主要有 DES、3DES、AES。

非对称加密算法 在加密消息传递中存在一方提前将密钥传递给另一方的过程。对称加密算法中的加密和解密使用的是同一个密钥。如果密钥在传递过程中被窃取，那么加密信息将不再安全。由于**密钥配送问题**的存在，衍生出了一些解决此问题的方案。而**非对称密码**（也称公钥密码）就是解决密钥配送问题的方案之一。在对称密码中，加密密钥和解密密钥是相同的，但非对称密码中，加密密钥和解密密钥却是不同的。只要拥有加密密钥，任何人都可以加密，但没有解密密钥时无法解密的。因此，公钥密码的一个重要性质，就是拥有解密密钥的人才能够进行解密。

接收者事先将加密密钥发送给发送者，这个加密密钥即使被窃取或者也没有问题。发送者使用加密密钥对通信内容进行加密并发送给接收者，而只有拥有解密密钥的人才能够进行解密（即接收者本身就拥有解密密钥）。这样一来，就不需要将解密密钥配送给接受者了。



公钥加密大多数是基于求解难题的。也就是说，是很难解决的问题。人们往往把大数字的因子分解或找出一个数的对数之类的问题作为公钥系统的基础。

公钥算法主要有 RSA、ECC（椭圆曲线算法）。

混合密码 公钥密码的处理速度只有对称密码的几百分之一。不适用于处理大量数据的加密。但对称密码又存在着**密钥配送问题。**为了系统的性能及安全性。往往会在系统的不同部分及阶段会将对称密码及公钥密码进行混合使用。即用对称密码提高处理速度，用公钥密码解决密钥配送问题。这样的方式即为混合密码系统。

著名的密码软件 PGP 以及网上的密码通信所使用的的 SSL/TLS 都运用了混合密码系统。

数字摘要算法又称哈希算法、散列算法，是一种单向算法，它通过对数据内容进行散列得到一个固定长度的密文信息（信息是任意长度，而摘要是定长）。即用户可以通过哈希算法对目标信息生成一段特定长度的唯一的 Hash 值，却不能通过这个 Hash 值重新获得目标信息。该算法不可逆。所以理解是与上述的加密方式均有差别，可以看作是**无法从密文（Hash 值）再去获得明文的。**

数字摘要算法有一个输入和一个输出，其中输入称为消息，输出称为散列值。单向散列函数可以根据消息的内容计算出散列值，而散列值就可以被用来检查消息的完整性。

散列值的长度和消息的长度无关。无论消息是 1 比特，还是 100MB，甚至是 100GB，单向散列函数都会计算出固定长度的散列值。以 SHA-1 单向散列函数为例，它所计算出的散列值的长度永远都是 160 比特。

理论上一个 Hash 值是可能对应无数多个原文的，算法把无限的映射成有限，因此可能会有碰撞（两个不同的信息，算出的摘要相同）。但是在现实中，由于原文的长度有限，所以想要出现两段原文对应同一个 Hash 值的是及其困难的，即几乎不会出现碰撞的情况，这就使得该算法得以在现实中可以应用。

单向散列函数的一些应用有：1.检测软件是否被篡改 2.基于口令的加密 3.消息认证码 4.数字签名 5.伪随机数生成器 6.一次性口令

数字摘要算法主要有 MD5、SHA-1、SHA-2、SHA-3。

3.2 硬加密与软加密

加密算法的安全不在于加密算法是否保密，而在于其在数学理论上是不可解或者几乎不可解的问题，或者以现有计算机算力水平计算其密钥不具备时效性（举例来说，如果一个加密算法的密钥需要用几十年才可以计算得出，那么也就认为加密算法是安全的）。

加密芯片是将各种加密算法以硬件逻辑电路的形式集成进芯片本身，同时还会集成如随机数生成模块、侧信道防护模块，物理破解防御模块等。这种对加密算法的应用通常称之为**硬加密**。

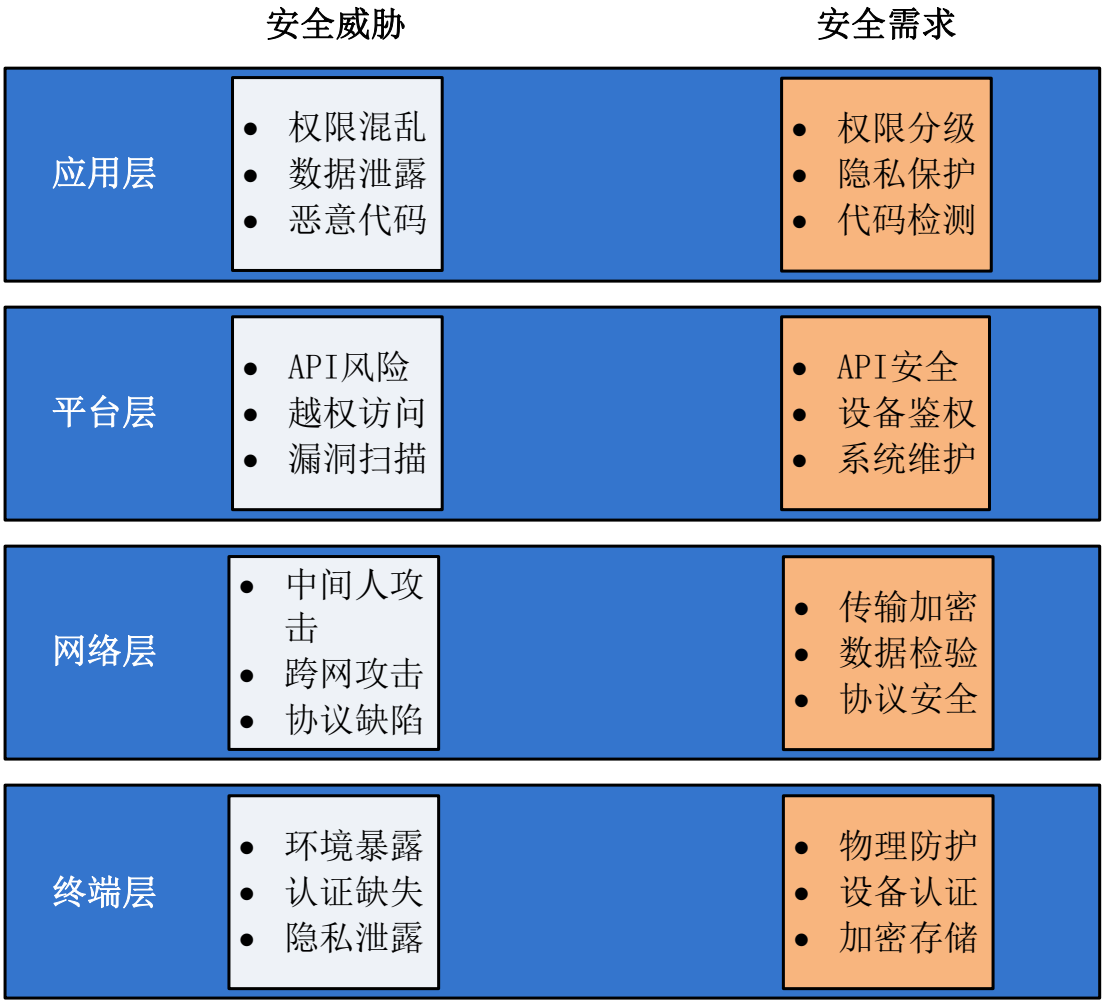
而单纯使用 CPU 或者内核资源进行的加密算法运算，一般称之为**软加密**。
下表为软加密和硬加密的区别对比。

	软加密	**硬加密（加密芯片）**
内核算力资源占用	占用	独立，不占用
物理防护	不具备	具备
侧信道防护	不具备	具备
随机数	伪随机数	真随机数
安全强度	低	高

第四章 解决方案

在互联网或者物联网中，信息在终端层、网络层、平台层及应用层等不同层会受到不同形式的安全威胁。如在最底层的终端层中，设备如果明文存储，则可能会遇到直接泄露的威胁。如在最上层的应用层，则可能被非合法用户侵犯权限。各种威胁，不一而足。

本文所描述解决方案主要针对**终端层**面对的各种安全威胁而实施。

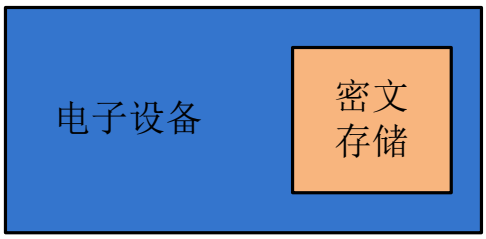


措施一：加密存储

通常的电子产品，数据和代码是以明文的形式存储于存储介质（如 Flash 或者 EEPROM 等中）。以现有的芯片破解手段来看，明文存储的数据和代码都是可以十分轻易被第三方获取的，用户的隐私以及生产厂商的知识产权完全无法得到保护。

因此**加密存储**是对数据进行保护的一种有效手段。对于大量的明文数据进行加密，通常使用**对称加密算法**，相较于非对称加密算法或哈希算法，对称加

密算法运算速度更快，更适用于大量数据的加密场景。

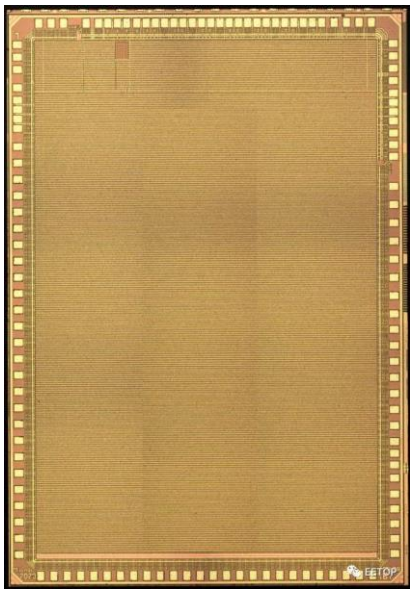


管理者在使用经过密文存储的设备过程中，需要注意的是对**密钥的管理，**通常密钥管理会有相应的平台及管理机制。如果密钥泄露，那么密文存储也就失去了其实际的意义。

措施二：物理防护

密文存储是对存储在芯片内部的数据或代码的直接保护，而物理防护则是针对于芯片本体而言。一种常规的物理破解手段是通过腐蚀芯片外部封装，从而将芯片的晶圆暴露出来，继而再通过探针技术或者 FIB 技术。

而我们的加密芯片在晶圆上层 MESH 防护网模块。MESH 防护网会监控网格线路的短路和开路，一旦触发，会导致存储器复位或者清零，从而保护存储在芯片内部的数据或代码。如下图为 MESH 防护网举例，非本芯片实物图。



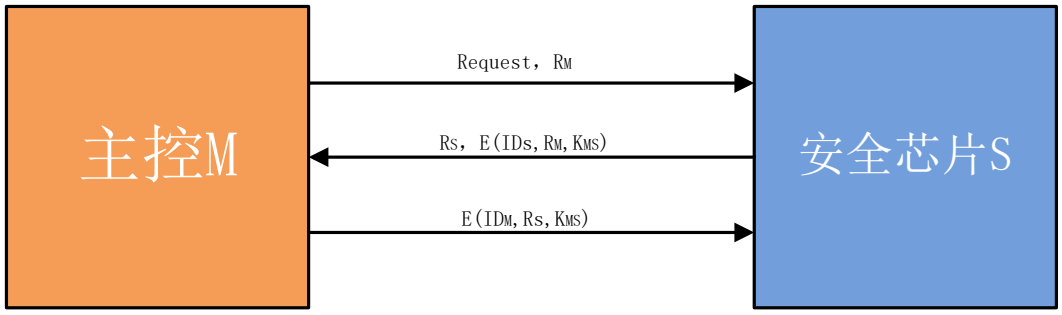
措施三：设备认证

加密芯片具有唯一 ID，可以用于生成密钥以及设备身份认证。联网设备通常需要一种可信身份标识，具备不可篡改、不可伪造的安全属性，这是实现设备互联以及信息流转的关键基础。设备可以和芯片的唯一 ID 进行绑定，从而在进行身份认证时使用该 ID，以避免非法设备侵入系统网络。

第四章 案例说明

在用户数据及用户信息保护的措施中，通过安全芯片与主控的配对、认证及识别是常见做法之一。

示意流程如下：



图示方案是利用对称加密进行安全芯片与主控的配对、认证，从而达到保护用户信息、用户数据及版权等。在这个方案中，有两个安全的“根”。

（1） 在生产阶段，需要将对称密钥灌入进主控 M 以及安全芯片 S 中，需要保证密钥的机密性。

（2） 安全芯片 S 的安全性需要足够保障方案的安全等级。例如安全芯片 S 需要可以应对物理入侵、侧信道攻击以及具备常用加密算法等。

以上方案可以防护如**重放攻击**（延迟后重放信息）、**明文攻击**以及**中间人攻击**等入侵方式。

具体流程如下：

（1） 在生产环节中，对主控 M 及安全芯片 S 进行密钥 KMS 灌入及唯一 ID 绑定。密钥由加密机等安全设备生成。

（2） 主控 M 在启动及运行过程中，向安全芯片 S 发送认证请求，并附加一次性信息 R_M，此一次性信息可以由随机数模块生成。

（3） 安全芯片 S 在收到认证请求后，生成安全芯片的一次性信息 R_S，并通过相应的对称加密算法计算明文（包含主控 M 的唯一 ID 的哈希值和一次性信息 R_M）的密文 MSG_S，安全芯片 S 将两条信息一起返回主控芯片 M。

（4） 主控芯片 M 对安全芯片返回的密文 MSG_S 进行校验，如信息无误，则生成主控芯片的密文 MSG_M（利用 R_S 及主控的唯一 ID 哈希值），并

返回给安全芯片 S，安全芯片进行相应校验。

（5） 双向校验完成后，主控 M 及安全芯片 S 认为对方为合法器件，可以进行正常工作及通信。

总结

以上基于对称加密算法对用户数据及信息的保护作了简要阐述，此方案可以有效的抵御如**重放攻击**（延迟后重放信息）、**明文攻击**以及**中间人攻击**等入侵方式。除此之外还可以利用非对称加密方式进行数据及信息保护，本文中不作赘述。