ICS 35.030
CCS L 80

中华人民共和国国家标准

GB/T 18238.1—2024
代替 GB/T 18238.1—2000

网络安全技术　杂凑函数
第1部分：总则

Cybersecurity technology—Hash-functions—Part 1：General

（ISO/IEC 10118-1:2016，Information technology—Security techniques—Hash-functions—Part 1：General，MOD）

2024-09-29 发布　　2025-04-01 实施

国家市场监督管理总局
国家标准化管理委员会　发布

目　　次

前　言

本文件按照GB/T 1.1—2020《标准化工作导则　第1部分：标准化文件的结构和起草规则》的规定起草。

本文件是GB/T 18238《网络安全技术　杂凑函数》的第1部分。GB/T 18238已经发布了以下部分：

——第1部分：总则；

——第2部分：采用分组密码的杂凑函数；

——第3部分：专门设计的杂凑函数。

本文件代替GB/T 18238.1—2000《信息技术　安全技术　散列函数　第1部分：概述》，与GB/T 18238.1—2000相比，除结构调整和编辑性改动外，主要技术变化如下：

a) 更改了术语“无碰撞散列函数”为“抗碰撞杂凑函数”（见3.4，2000年版的2.1）；

b) 更改了术语“散列码”为“杂凑值”（见3.3，2000年版的2.3）；

c) 增加了“输出变换”“轮函数”等术语（见第3章）；

d) 增加了符号B_i、D_i、H_i、h、L_1、L_2、n、q、T、ϕ（见第4章）；

e) 增加了“杂凑函数的通用模型”（见第6章）；

f) 更改填充方法2为填充方法1，增加了填充方法2；并删除了填充方法1（见A.3，2000年版的附录B）。

本文件修改采用ISO/IEC 10118-1:2016《信息技术　安全技术　杂凑函数　第1部分：总则》。

本文件与ISO/IEC 10118-1:2016相比做了下述结构调整：

——调整了术语“杂凑函数”和“抗碰撞杂凑函数”的顺序（见第3章）；

——4.1对应ISO/IEC 10118-1:2016的4.1和4.2；

——删除了ISO/IEC 10118-1:2016的附录B，将附录C调整为附录B。

本文件与ISO/IEC 10118-1:2016的技术差异及其原因如下：

——将“范围”中关于杂凑函数的介绍内容移至“引言”；

——增加了规范性引用文件GB/T 25069—2022（见第3章）；

——增加了符号n表示分组密码的分组长度、q表示输入数据位串的分组个数（见第4章）；

——将ISO/IEC 10118-1:2016中3.4的注移至第5章，改为正文；

——删除了规范性附录B“ISO/IEC 10118（所有部分）采纳杂凑函数的原则”，因为该原则是ISO采纳各国算法提案所考虑的原则，并不直接适用于具体各国规范自身的算法。

本文件做了下列编辑性改动：

——为与我国技术标准体系协调，标准名称更改为《网络安全技术　杂凑函数　第1部分：总则》；

——纳入了ISO/IEC 10118-1:2016/Amd.1:2021的内容；

——增加了术语“抗碰撞杂凑函数”“杂凑值”“杂凑函数”“初始化值”及“填充”的来源标准（见第3章）；

——删除了术语“杂凑函数”“杂凑值”及“初始化值”的注，更改了术语“轮函数”的注（见第3章）；

——用资料性引用的GB/T 18238（所有部分）替换了ISO/IEC 10118（所有部分），用GB/T 18238.2和GB/T 18238.3替换了ISO/IEC 10118的其他部分；

——删除了资料性附录B中B.3的示例1（见ISO/IEC 10118-2:2016的附录C）；

——删除了ISO/IEC 10118-1:2016/Amd.1:2021的填充方法3，因为本系列文件规定的杂凑函数

未使用该填充方法。

——更改了参考文献。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由全国网络安全标准化技术委员会(SAC/TC 260)提出并归口。

本文件起草单位:中电科网络安全科技股份有限公司、国家密码管理局商用密码检测中心、中国电子技术标准化研究院、中国电子科技集团公司第十五研究所、中国科学院信息工程研究所、中国科学院软件研究所、中国科学院大学、山东大学、西安西电捷通无线网络通信股份有限公司、北京银联金卡科技有限公司、格尔软件股份有限公司、北京信安世纪科技股份有限公司、山东得安信息技术有限公司、华为技术有限公司、北京江南天安科技有限公司、智巡密码(上海)检测技术有限公司等、北京海泰方圆科技股份有限公司。

本文件主要起草人:张立廷、罗鹏、李彦峰、李艳俊、毛颖颖、李世敏、黄晶晶、史丹萍、眭晗、孙思维、王鹏、王薇、王丹丹、杜志强、王提、杨波、郑强、龚晓燕、马洪富、曾光、李雪雁、韩玮、潘文伦、贾世杰、熊云、杨慧慧。

本文件及其所代替文件的历次版本发布情况为:

——2000 年首次发布为 GB/T 18238.1—2000;

——本次为第一次修订。

引　言

杂凑函数使用特定的算法将任意长度(通常设有上限)的位串映射到固定长度的位串。杂凑函数通常用于:

——将消息压缩为摘要,用于数字签名机制的输入;

——向用户承诺一个给定的位串,而不泄露该位串。

注:GB/T 18238(所有部分)中规定的杂凑函数不涉及密钥的使用。但是,这些杂凑函数与密钥搭配使用,以构建消息鉴别码(Message Authentication Code, MAC)。消息鉴别码提供数据源鉴别和消息完整性保护。GB/T 15852.2给出了杂凑函数计算 MAC 的技术。

GB/T 18238《网络安全技术　杂凑函数》由 3 个部分构成。

——第 1 部分:总则。目的在于规定杂凑函数的要求和通用模型,用于指导 GB/T 18238 的其他部分。

——第 2 部分:采用分组密码的杂凑函数。目的在于规定采用分组密码的杂凑函数。

——第 3 部分:专门设计的杂凑函数。目的在于规定专门设计的杂凑函数。

网络安全技术　杂凑函数
第1部分:总则

1　范围

本文件规定了杂凑函数的要求和通用模型,描述了杂凑运算的四个步骤,并给出了通用模型的使用方法。

本文件包含GB/T 18238(所有部分)所共用的定义、符号和要求。

2　规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中,注日期的引用文件,仅该日期对应的版本适用于本文件;不注日期的引用文件,其最新版本(包括所有的修改单)适用于本文件。

GB/T 25069—2022　信息安全技术　术语

3　术语和定义

GB/T 25069—2022界定的以及下列术语和定义适用于本文件。

3.1

杂凑函数　hash-function

将任意长位串映射为定长位串的函数,满足下列性质:

——给定一个输出位串,寻找一个输入位串来产生该输出位串,在计算上不可行;

——给定一个输入位串,寻找另一个不同的输入位串来产生相同的输出位串,在计算上不可行。

[来源:GB/T 25069—2022,3.505]

3.2

数据串　data string

杂凑函数的输入位串。

3.3

杂凑值　hash value

密码杂凑运算的结果。

[来源:GB/T 25069—2022,3.764]

3.4

抗碰撞杂凑函数　collision-resistant hash-function

满足如下性质的杂凑函数:找出映射到同一输出的任何两个不同输入在计算上是不可行的。

注:计算可行性依赖于特定安全要求和环境。

[来源:GB/T 25069—2022,3.322,有修改]

3.5

初始化值　initialization value

在密码变换中,为增强安全性或使密码设备同步而引入的用于数据变换的起始数据。

[来源:GB/T 25069—2022,3.80]

3.6

输出变换　output transformation

对迭代操作的输出进行变换,得到杂凑值。

3.7

填充　padding

向某一数据串附加额外位的操作。

[来源:GB/T 25069—2022,3.598]

3.8

轮函数　round-function

构成杂凑函数的主要部件之一,将两个特定长度的位串映射为一个定长位串的函数,被迭代地用于杂凑函数的计算过程。

注:在该领域的文献中,多个术语具有与轮函数相同或相似的含义。例如:压缩函数和迭代函数。

4　符号

4.1　一般符号

下列符号适用于本文件。

B_i:当 B 是由多个 m 位字构成的序列时,B_i($i \geqslant 0$)表示 B 的第 i 个 m 位字。特别地,当 $m=8$ 时,B_i 是 B 的第 i 个字节。

D:数据串。

D_i:数据 D 经填充后的第 i 个 n 位分组。

H:杂凑值。

H_i:用于存储杂凑运算中间结果的位串,其长度为 L_2。

h:杂凑函数。

IV:初始化值。

L_1:输入到轮函数 ϕ 的两个位串中,第一个位串的位长度。

L_2:输入到轮函数 ϕ 的两个位串中,第二个位串的位长度,也是轮函数 ϕ 输出值的位长度,以及初始值 IV 的位长度。

L_X:位串 X 的位长度。

n:n 位分组密码算法 E 的分组长度。

q:经过填充和分割操作后,输入数据位串 D 的分组个数。

T:输出变换,比如截短。

$X \| Y$:按顺序将位串 X 和 Y 连接所构成的位串。

$X \oplus Y$:位串 X 和位串 Y 的异或(其中 $L_X = L_Y$)。

ϕ:轮函数。

4.2　编码约定

如果需要定义“最高有效位/字节”和“最低有效位/字节”(例如,将位/字节串视为数值),则一个分组的最左边的位/字节被视为最高有效位/字节。

5　要求

实体在使用杂凑函数前,应将数据串表示为统一形式,使得各方操作的位串是完全相同的。

为使得数据串的长度达到要求，GB/T 18238(所有部分)规定的杂凑函数需进行填充操作。具体填充方法按附录 A 中描述的方法。

附录 B 给出了安全性注意事项。

6 杂凑函数的通用模型

6.1 概述

GB/T 18238(所有部分)规定的杂凑函数要求使用轮函数 ϕ。

GB/T 18238.2 和 GB/T 18238.3 规定的杂凑函数输出长度为 L_H 位的杂凑值，其中 L_H 不大于轮函数 ϕ 中的 L_2。

6.2 杂凑运算

6.2.1 通则

在 GB/T 18238.2 和 GB/T 18238.3 规定的杂凑函数使用轮函数 ϕ 和长度为 L_2 的初始化值 IV。对于给定的 ϕ，IV 的值应是固定的。通过以下四个步骤计算数据串 D 的杂凑值 H。

6.2.2 步骤 1(填充)

对数据串 D 进行填充操作，以确保其长度是 L_1 的整数倍。具体方法见附录 A。

6.2.3 步骤 2(分割)

填充后的数据串 D 被分割成多个 L_1 位长的分组 $D_1, D_2, \cdots, D_q$。其中，D_1 表示第一个分组，D_2 表示第二个分组，以此类推。填充和分割过程如图 1 所示。

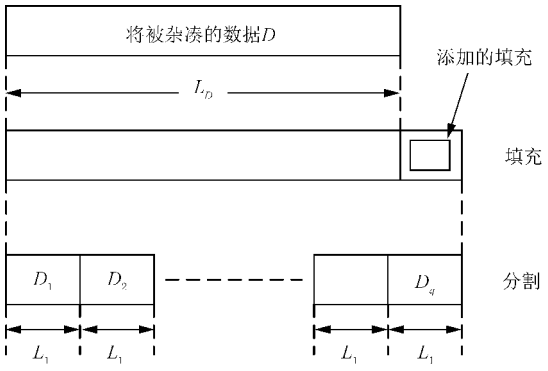

图 1 填充和分割示意图

6.2.4 步骤 3(迭代)

令 $H_0 = IV$，以如下方式迭代计算长度为 L_2 位的串 $H_1, H_2, \cdots, H_q$。

对 $i = 1, \cdots, q$，依次计算：

$H_i = \phi(D_i, H_{i-1})$。

6.2.5 步骤 4(输出变换)

对步骤 3 的输出 H_q 执行变换 T，得到 L_H 位的杂凑值 H。

示例：变换 T 可以是截短操作。

6.3 通用模型的使用

GB/T 18238.2 和 GB/T 18238.3 规定了基于通用模型的杂凑函数的示例。在每个示例中，描述一个具体杂凑函数都需要定义以下内容：

——参数 L_1，L_2；

——填充方法；

——初始化值 IV；

——轮函数 ϕ；

——输出变换 T。

在实际中使用通用模型所定义的杂凑函数还需要选择参数 L_H。

附 录 A
（规范性）
填充方法

A.1 概述

杂凑值的计算需要填充，使得填充后的数据串的位长度为 L_1 的整数倍，填充的位串无需随原消息存储或发送，本附录列出了两种填充方法。

A.2 方法 1

在数据串右侧填充一个位“1”，然后在所得到的位串右侧填充“0”，尽可能少填充（甚至不填充“0”），以达到所要求的长度。

注：方法 1 总是要求填充至少一个位。

A.3 方法 2

选择一个参数 r（其中 $r \leqslant L_1$），例如 $r=64$，以及一种将数据串 D 的位长度 L_D 编码为 r 位的位串的方法。参数 r 的选择限制了可处理的数据串 D 的长度，$L_D < 2^r$。

为计算杂凑值，需按以下方式填充数据串 D：

a） 在数据串 D 右侧填充一个位“1”；

b） 如果上一步得到的位串的长度与 $L_1 - r$ 模 L_1 同余，则不填充；否则在上一步得到的位串右侧填充位“0”，尽可能少填充，使填充后的位串长度与 $L_1 - r$ 模 L_1 同余，即填充后的位串长度比 L_1 位的整数倍少 r 位，如果 $r = L_1$，则填充后的位串长度等于 L_1 位的整数倍；

c） 使用选定的编码方法在上述结果后面添加 r 位编码的 L_D，得到填充后的数据串 D。

附 录 B
（资料性）
安全性注意事项

B.1 密码攻击目标

与杂凑函数相关的密码攻击目标有多种(GB/T 15852.2 给出了示例)。以下几点尤为重要。

碰撞攻击:密码攻击目标是寻找两个不同的数据串M_1,M_2,满足 $h(M_1)=h(M_2)$。

原像攻击:给定适合长度的位串 H,密码攻击目标是找到数据串 M,满足 $h(M)=H$。

第二原像攻击:给定数据串 M,密码攻击目标是找到另外一个数据串 M',满足 $h(M')=h(M)$且 $M'\neq M$。

长度延长攻击:给定位串 $h(M)$,其中 M 为未知的非空数据串,密码攻击目标是找到任意数据串 M'以及$h(M \parallel M')$。

注:当前针对杂凑函数的密码分析涉及很多种密码攻击目标,包括但不限于上述目标。标准化过程会考虑这些目标,但仅作为参考。此外,在应用中通常并不要求杂凑函数能抵抗所有密码攻击目标。一般仅考虑一部分特定目标。

B.2 通用攻击

通用攻击是一种适用于所有杂凑函数且不依赖于杂凑函数具体构造的攻击。

示例:暴力搜索原像攻击。给定一个杂凑值,攻击者尝试所有可能的数据串 M,计算 $h(M)$的值,并将结果与给定的杂凑值进行比较,如果两者匹配,则完成原像搜索目标。

B.3 密码算法攻击的影响

在 GB/T 18238(所有部分)中,杂凑函数抵抗潜在攻击的能力是根据攻击达到目标的“计算不可行性”来衡量的。正如定义所示,计算不可行性的含义取决于特定的安全需求和环境。一种经常被安全从业人员使用的含义是,当完成一个任务需要的计算资源超过了通常可用的资源,则称该任务具有计算不可行性。

一种更严格的方法是在相同密码攻击目标下,比较一个具体的攻击方法与通用攻击的效率。对于一个给定的密码攻击目标,如果所有已知密码攻击的效率都不高于相应的通用攻击,则称该杂凑函数能够抵抗该密码攻击目标。然而,如果存在一种比相应的通用攻击的效率高得多的密码攻击,则称该杂凑函数被攻破了。

密码算法攻击的效率取决于三个参数:计算复杂度、存储需求和成功概率。

密码算法攻击的计算复杂度由调用轮函数的次数定义,以确定它们相对于通用攻击的复杂性。这种标准化的复杂性可能因攻击性质而有所不同。在大多数情况下,只能估计密码算法攻击的复杂度。

示例:考虑一个具有 256 位杂凑值的杂凑函数。如果存在原像攻击,该攻击调用约 2^{192}次轮函数,实际存储需求约 2^{20}字节,成功概率接近 1,则对于原像攻击,该杂凑函数被攻破。

参 考 文 献

[1] GB/T 15852.2 网络安全技术 消息鉴别码 第2部分:采用专门设计的杂凑函数的机制

[2] GB/T 18238.2 网络安全技术 杂凑函数 第2部分:采用分组密码的杂凑函数

[3] GB/T 18238.3 网络安全技术 杂凑函数 第3部分:专门设计的杂凑函数

[4] ISO/IEC 9797-2 Information security—Message authentication codes (MACs)—Part 2: Mechanisms using a dedicated hash-function

[5] Preneel B. Analysis and Design of Cryptographic Hash Functions, Doctoral Dissertation, Katholieke Universiteit Leuven, 1993.

中 华 人 民 共 和 国

国 家 标 准

网络安全技术　杂凑函数

第 1 部分:总则

GB/T 18238.1—2024

*

中国标准出版社出版发行

北京市朝阳区和平里西街甲 2 号(100029)

北京市西城区三里河北街 16 号(100045)

网址:www.spc.net.cn

服务热线:400-168-0010

2024 年 9 月第一版

*

书号:155066·1-77561

GB/T 18238.1-2024