1 SM3 Length Extension Attack 验证

1.1 攻击原理与数学表示

Length Extension Attack（长度扩展攻击）是一种针对 **基于 Merkle–Damgård 结构的哈希函数**的攻击方式，包括 MD5、SHA-1、SM3 等。攻击的核心在于：如果我们已知 Hash(m) 和 m 的长度（但不知道 m 本身），我们可以在不改变哈希内部状态的情况下，计算 Hash($m \parallel \text{padding}(m) \parallel m_{\text{extra}}$)。

1.1.1 Merkle–Damgård 结构回顾

SM3 的压缩过程可抽象为：

$$H_0 = \text{IV}$$

$$H_{i+1} = CF(H_i, B^{(i)})$$

其中 $B^{(i)}$ 为第 i 个 512 位分组，CF 为压缩函数。

最终：

$$\text{SM3}(M) = H_n$$

1.1.2 长度扩展攻击的数学推导

已知：

$$h = \text{SM3}(m) = H_{\text{len}(m)/512}$$

攻击者不知道 m，但知道其长度 l_m，以及 h。

步骤：1. 构造 m 的填充串：

$$\tilde{m} = m \parallel 1 \parallel 0^k \parallel \text{len}(m)_{64}$$

其中 k 满足：

$$(l_m + 1 + k) \equiv 448 \pmod{512}$$

2. 以 h 作为初始向量（IV），将额外消息 m_{extra} 当作新的分组输入压缩函数：

$$h' = \text{SM3}_{\text{IV}=h}(m_{\text{extra}})$$

3. 得到的 h' 即为：

$$h' = \text{SM3}(m \parallel \text{padding}(m) \parallel m_{\text{extra}})$$

1.2 攻击效果

- 攻击者无需知道原消息内容，只需知道其长度。- 只要哈希函数是 Merkle–Damgård 结构，且填充规则可推测，就可实现该攻击。- 在基于哈希的消息认证码（如 `MAC = SM3(key || message)`）中，如果 `key` 固定且可预测长度，则存在被利用风险。

1.3 实现思路

- **已知条件**：原消息长度 l_m、原哈希值 h。
- **步骤 1**：根据 l_m 推算填充串长度，模拟 SM3 的消息填充函数生成 m 的 padding。
- **步骤 2**：将 h 作为 IV 初始化 SM3 状态。
- **步骤 3**：输入额外数据 m_{extra}，继续执行 SM3 消息扩展与压缩函数。
- **步骤 4**：输出新哈希值 h'，该值等于 $\text{SM3}(m \parallel \text{padding}(m) \parallel m_{\text{extra}})$。

1.4 实验验证要点

1. 实现可自定义 IV 的 SM3 压缩函数接口。
2. 通过与直接计算的结果对比验证攻击成功。
3. 测试多种消息长度，确保填充逻辑正确。