

词典

排解冲突：平方试探

11 - C4

三十六计，走为上计

我真的以为，这样何尝不是一种所谓的解脱  
要背负的辛苦又有谁能够清楚，那内心的冲突

邓俊辉

deng@tsinghua.edu.cn

# 平方试探

## ❖ Quadratic Probing

以平方数为距离，确定下一试探桶单元

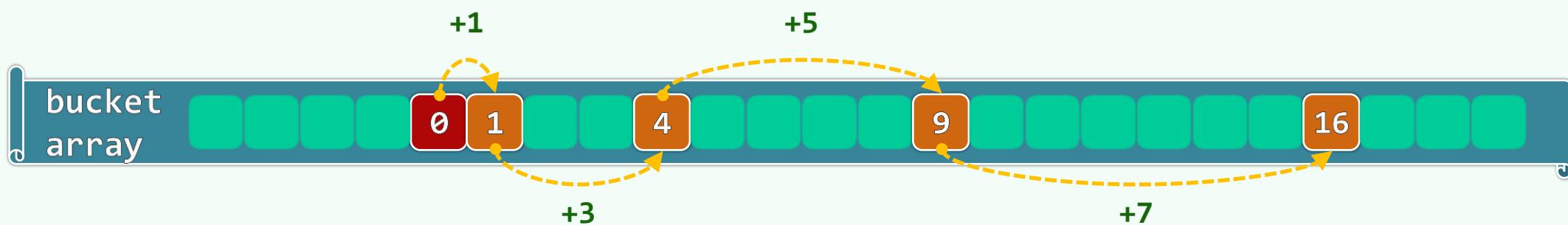
$$[ \text{hash(key)} + 1^2 ] \% M$$

$$[ \text{hash(key)} + 2^2 ] \% M$$

$$[ \text{hash(key)} + 3^2 ] \% M$$

$$[ \text{hash(key)} + 4^2 ] \% M$$

...



## ❖ 数据聚集现象有所缓解

- 查找链上，各桶间距线性递增

- 一旦冲突，可“聪明”地跳离是非之地

❖ 对于大散列表，I/O操作有所增加

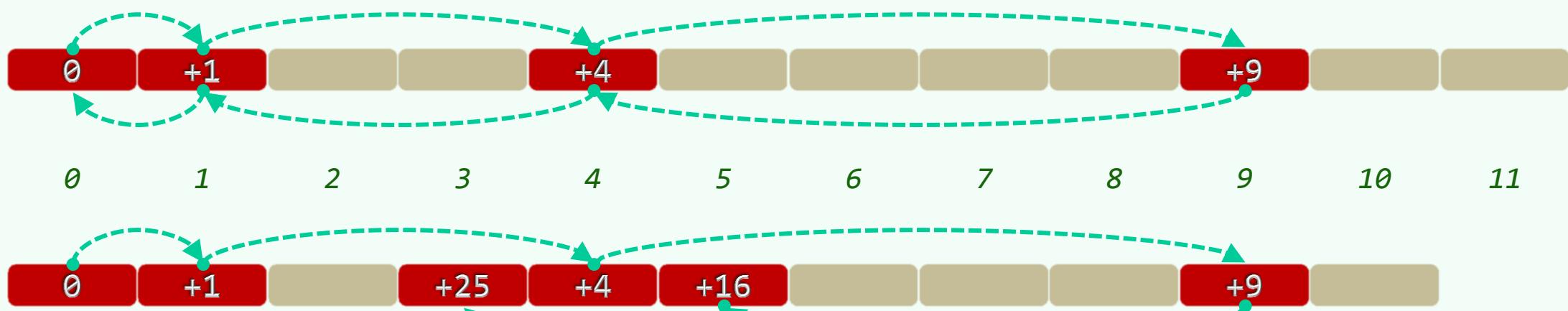
❖ 只要有空桶，就...一定能...找出来吗？

//毕竟不是挨个试探

# 装填因子，须足够小！

❖  $\{0, 1, 2, 3, 4, 5, \dots\}^2 \% 12 = \{0, 1, 4, 9\}$

M若为合数： $n^2 \% M$  可能的取值可能少于  $\lceil M/2 \rceil$  种——此时，只要对应的桶均非空...



❖  $\{0, 1, 2, 3, 4, 5, \dots\}^2 \% 11 = \{0, 1, 4, 9, 5, 3\}$

M若为素数： $n^2 \% M$  可能的取值恰好会有  $\lceil M/2 \rceil$  种，且由查找链的前  $\lceil M/2 \rceil$  项取遍

❖ 定理：若M是素数，且  $\lambda \leq 0.5$ ，就一定能够找出；否则，不见得

# 查找链前缀，必足够长！

❖ 反证：假设存在  $0 \leq a < b < \lceil \mathcal{M}/2 \rceil$ ，使得

沿着查找链，第  $a$  项和第  $b$  项彼此冲突

❖ 于是： $a^2$  和  $b^2$  自然关于  $\mathcal{M}$  同余，亦即

$$a^2 \equiv b^2 \pmod{\mathcal{M}}$$

$$b^2 - a^2 = (b + a) \cdot (b - a) \equiv 0 \pmod{\mathcal{M}}$$

❖ 然而， $0 < b - a \leq b + a < \lceil \mathcal{M}/2 \rceil + (\lceil \mathcal{M}/2 \rceil - 1) \leq \lceil \mathcal{M}/2 \rceil + \lfloor \mathcal{M}/2 \rfloor = \mathcal{M}$

无论  $b - a$  还是  $b + a$  都不可能整除  $\mathcal{M}$

❖ 那么，另一半的桶，可否也利用起来呢...