

## 9. 词典

排解冲突

平方试探

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

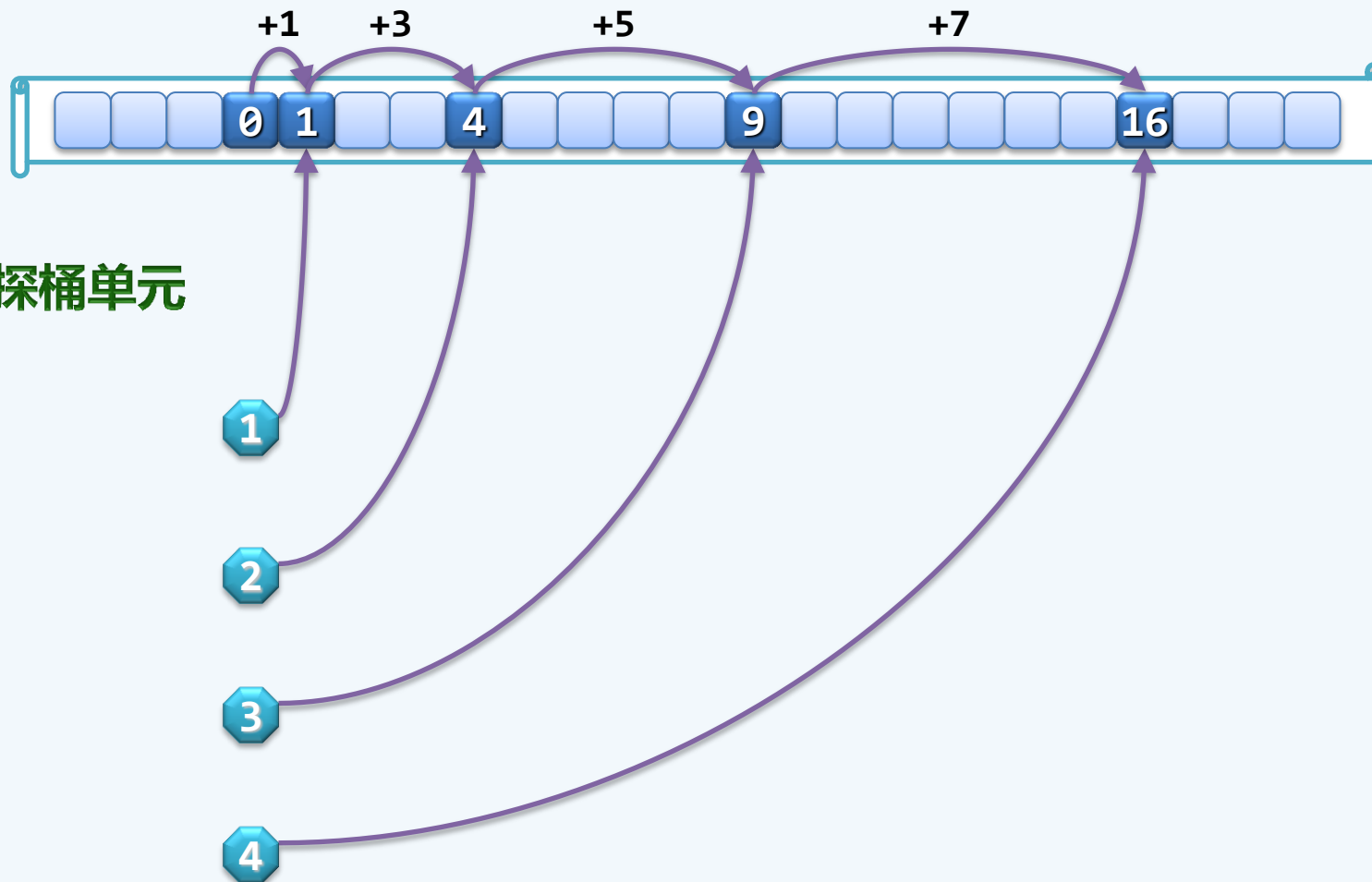
$$[ \text{hash}(\text{key}) + 1^2 ] \% M$$

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

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

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

...



## 优点、缺点及疑惑

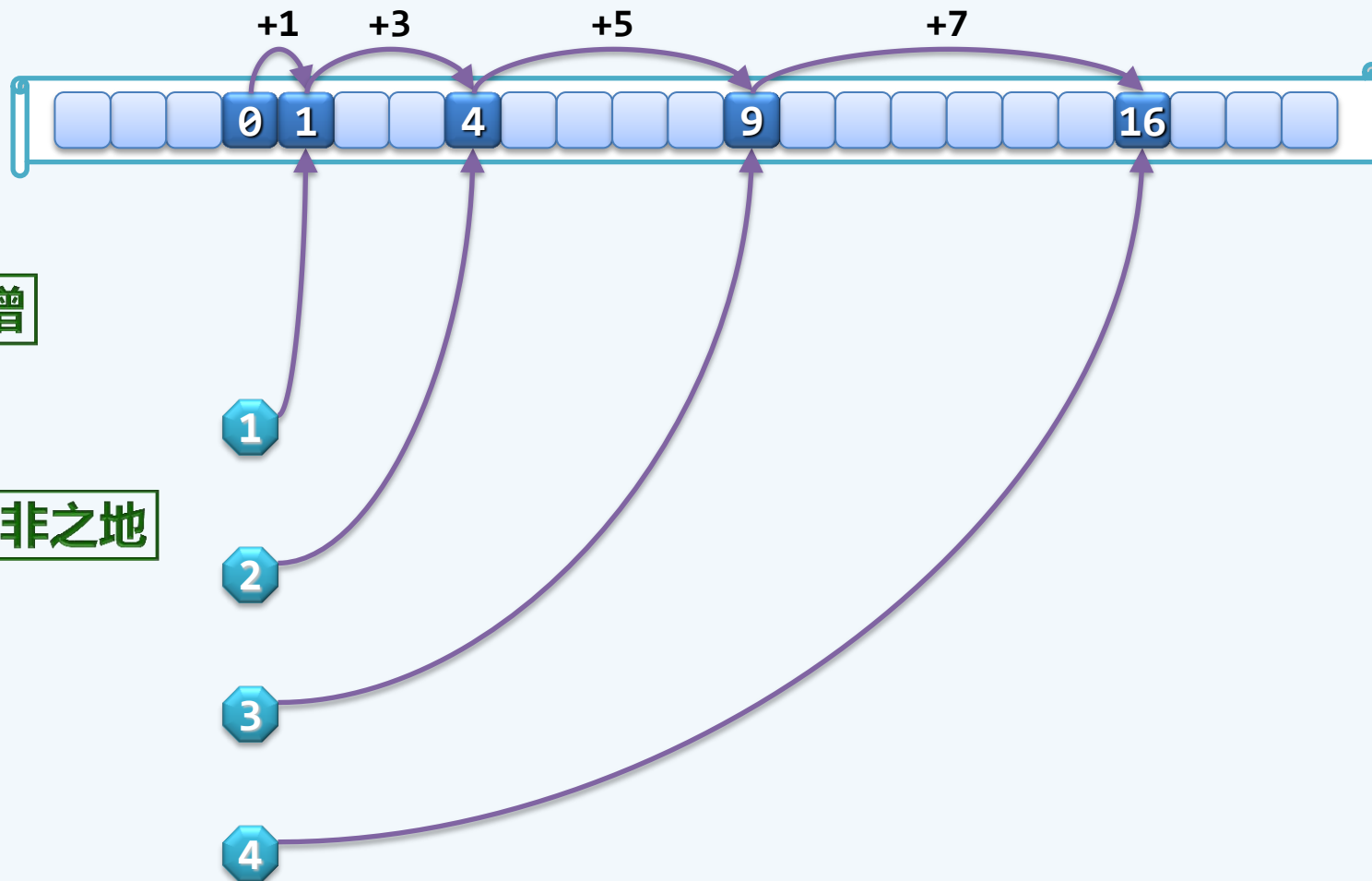
### ❖ 数据聚集现象有所缓解

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**

### ❖ 若涉及 **外存**，I/O 将激增

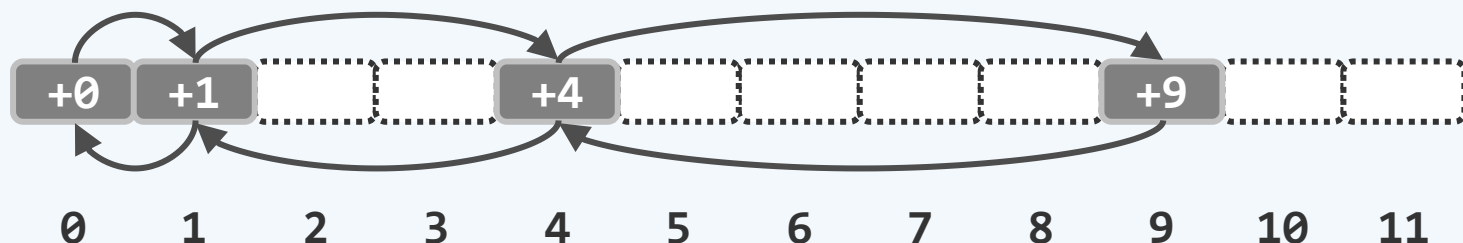
### ❖ **只要有** 空桶，就... **一定能** ...找出来吗？



// 毕竟不是挨个试探

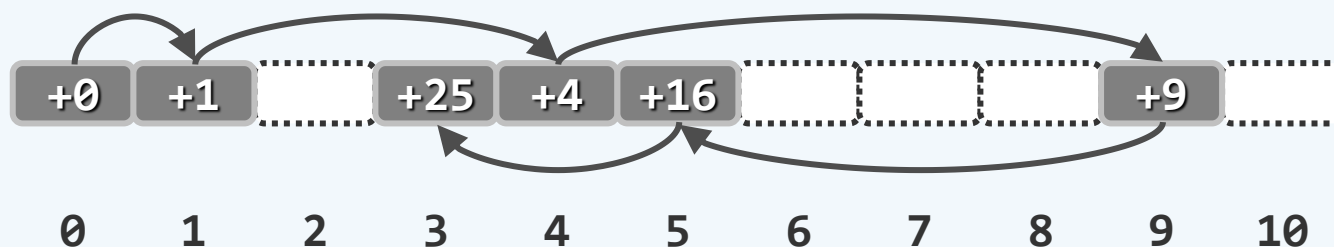
## 装填因子，须足够小！

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



$M$ 若为合数： $n^2 \% M$ 可能的取值必然少于  $\lceil \frac{\mathcal{M}}{2} \rceil$  种——此时，只要对应的桶均非空...

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



$M$ 若为素数： $n^2 \% M$ 可能的取值恰好会有  $\lceil \frac{\mathcal{M}}{2} \rceil$  种——此时，恰由查找链的前  $\lceil \frac{\mathcal{M}}{2} \rceil$  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  $0 \leq a < b < \lceil \frac{\mathcal{M}}{2} \rceil$ ，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  $\mathcal{M}$  的[同余]，亦即

$$a^2 \equiv b^2 \pmod{\mathcal{M}} \qquad 0 < b - a \leq b + a < \mathcal{M}$$

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

❖ 然而： $0 < b - a \leq b + a < \mathcal{M}$  —— 这与  $\mathcal{M}$  为[素数]矛盾

❖ 那么，[另一半]的桶，可否也利用起来呢...