Tecniche di probing (tabelle hash)

È desiderabile che la sequenza di posizioni da controllare generata da ogni chiave abbia la stessa probabilità di essere una delle $m!$ permutazioni di $\{0, \ldots, m-1\}$ (*uniform hashing*). Tuttavia, nessuna delle seguenti funzioni è in grado di restituire più di m^2 sequenze.

Probing lineare

$$h(k, i) = (h'(k) + i) \bmod m$$

Clustering primario: si accumulano slot occupati in un'unica sezione perché uno slot vuoto preceduto da i occupati viene riempito con probabilità $\frac{i+1}{m}$. Le sequenze di elementi con hash uguale sono uguali, quindi ce ne sono soltanto m possibili.

Probing quadratico

$$h(k, i) = (h'(k) + c_1 i + c_2 i^2) \bmod m$$

È necessario scegliere c_1 e c_2 in modo che tutte le posizioni vengano controllate. m possibili sequenze come nel probing lineare (clustering secondario).

Doppio hash

$$h(k, i) = (h_1(k) + i h_2(k)) \bmod m$$

Per generare tutte le posizioni $h_2(k)$ deve sempre essere coprimo rispetto a m, proprietà che si può ottenere ad esempio se l'hash è sempre dispari e m è una potenza di 2, oppure m primo e $h_2(k) < m$. Dal momento che ad ogni coppia di hash corrisponde una sequenza diversa, ce ne sono m^2 possibili.

Esempio:

$$h_1(k) = k \bmod m \qquad h_2(k) = 1 + (k \bmod (m-1)).$$