



Assignment No-1

- Title :- Telephone db (Data Base)
- Objective :-
 - 1] To understand concept of hashing
 - 2] To understand to find record quickly using hash function.
- Problem Statement :-

Consider telephone book database of N clients make use of hash table implementation to quickly look up clients telephone numbers. Make use of collision handling techniques and compare them using number of comparison required to find set of telephone no.
- Outcome :-

Students will be able to learn python programming features and implement concept of hash table.
- Software and hardware requirement :-

Operating System requirement : 64 bit open source
programming tools recommended : open source python.
- Theory :- Hash Function

A hash function is a mathematical function that takes an input & produce a fixed size

string of character. The output often referred as hash value of hash table, is unique representation of input data. Hash functions are commonly used in various applications such as data integrity verification, password storage & digital signatures.

Here are some good properties of good hash function :-

- 1) Deterministic :- For the same input hash function produces same output.
- 2) Efficient :- It is efficient to compute & calculate hash value for any given input.
- 3) Preimage resistance :- Given a hash function value, it should be computationally infeasible to reverse the process.
- 4) Collision resistance :- It should be unlikely for two different inputs to produce the same hash value.

- How to calculate a Hash :-

A hash function takes input data & produces a fixed size string character known as hash value. The process typically involves preprocessing the input then applying a series of mathematical operations in blocks to produce the final hash. The output is a unique

representation of the input. It's important to use established hash function for security purpose.

eg:-

- 1) Input data : "Hello"
- 2) ASCII values : $h=104$, $E=101$, $l=108$, $l=111$
- 3) Sum : $104 + 101 + 108 + 111 = 632$
- 4) Modulus 10 : $632 \% 10 = 2$

• Conclusion:-

In computer science, a collision occurs when two distinct inputs to a hash function produce the same output resulting in collision.

There are two types of collision resolution technique as shown below-

Collision Resolution Technique.

Separate chaining
(open chaining).

Open Addressing
(closed hashing)

Linear Probing

Quadratic Probing

Double hashing

1) Linear Probing :-

In linear probing when a collision occurs, the probing resolves it by searching for the next available (empty) slot in the hash table starting from the collision point & moving sequentially forward until an empty slot is found.

Advantage :-

- 1) Simplicity :- It is straight forward to implement
- 2) Memory efficiency :- It typically requires minimal additional memory overhead beyond the hash table.

Disadvantage :-

- 1) Primary clustering :- It is susceptible to primary clustering, where collision result in long sequence of occupied slots.
- 2) Secondary clustering :- While linear probing can reduce primary clustering by reducing colliding elements into available slot, can still suffer from secondary clustering.

2) Quadratic Probing :-

It is collision resolution technique used in hash table it address the primary clustering issue encountered in linear probing by using quadratic function to probe for alternative slots when collision occurs we probe for i^2 th slots in i th iteration.

Advantage :-

- i) Reduce primary clustering :- It reduces the occurring of primary clustering compared to linear probing by a quadratic function to calculate the probe sequence.
- ii) No additional data structure :- Quadratic probing, like linear probing, does not require additional data structure for collision resolution.

Disadvantage :-

- i) Performance degradation :- In the worst case scenario, where the hash table heavily loaded quadratic probing may experience significant performance degradation due to increased likelihood of longer probe sequence compared to linear probing.

ii) secondary clustering :- Just like linear p, it suffer from secondary clustering.

• Algorithm :-

- 1) start
- 2) Initialize a hash table with given size of set all element to 0.
- 3) Implement a hash function that calculate index using module division handle collision using linear probing.
- 4) Implement a function find to search for an element in the hash table.
- 5) Implement the main program loop to allow the user to perform operation like inserting element using probing, inserting using quadratic probing, searching elements using quadratic probing searching elements & exiting the program.

Conclusion :-

Thus we understand how hashing is an effective method to quickly access data using key value.

Flowchart :-

