**File Management and Processing**

**IS321**

**Mid-Term Task**

**FCAI-CU**

**Tawfik Yasser Tawfik**

**20180075**

**IS-3**

**Topic**

**Why binary search is not efficient and even not a practical approach to be used with large file?**

**For Dr. Ali Zidan**

1. What is binary search algorithm:

Is an algorithm to search for a specific value in a set of values, it is used when we have a given value and a sorted array, we need to find the index of that value such that: array[index] = value.

If we found the value, we are searching for in the array that contains the file data, so we will return its index, otherwise we will return NULL or -1, which NULL indicates that the desired value not found.

The algorithm works as follows:

If we have the RRN (Relative Record Number) of the record we are searching for by its key-value, We can go directly to the record using seek() function.

But usually, we do not have the RRN for the records, so we must search for the entire file, and if the file is not sorted then we must look at all the records till finding the desired record.

So, if we have a file sorted by the key-value we can use binary search to find the desired record.

But let us imagine a large file need to be sorted, so it is too big to put it into main memory, so we will use the keys only to sort them and put the keys in the memory, sort them, and finally we will write them back to the hard disk.

Binary Search Procedure:

* Read the file sequentially into main memory once.
* Sort the keys array in main memory.
* Go a long all records in random order to write the sorted array in new file.
* Write each record once sequentially to replace the original keys.

Binary search time cost: O(log2n)

1. Cases where binary search should be the most reasonable choice

Binary search is very good choice if the data file is small so we can put it in the main memory and sort it then perform the algorithm on it.

Also, if the keys file small as most as possible, also we it can fit into main memory, then sorted, finally rewritten into the original file again.

1. Cases where binary search should not be a choice

The cases where binary search not good for and should not use, is when we have large amount of data, so the file will not fit into the main memory also the keys file will not fit into main memory and we can not perform the sorting process to apply the binary search which working only on a sorted data.