Chapter 2

SYSTEM ANALYSIS

2.1 Analysis of Project

Placement Statistics is used by the user to view a particular college's placement details. The system is initially used to add student records placed in various companies, into the file corresponding to the branch. The system can then be used to search, delete, modify or display existing records and, to view the current placement statistics of a college.

2.2 Structure used to Store the Fields and Records

Storing Fields

- **Fixing the Length of Fields:**

In the Placement System, the USN field is a character array that can hold a string value of some maximum size. The size of the array is larger than the longest string it can hold.

- **Separating the Fields with Delimiters:**

We preserve the identity of fields by separating them with delimiters. We have chosen the vertical bar character, as the delimiter here.

Storing Records

- **Making Records a Predictable Number of Fields:**

In this system, we have a fixed number of fields, each with a maximum length, that combine to make a data record. Fixing the number of fields in a record does not imply that the size of fields in the record is fixed. The records are used as containers to hold a mix of fixed and variable-length fields within a record. We have 18 contiguous fields and we can recognize fields simply by counting the fields modulo 18.

- **Using an Index to Keep Track of Addresses:**

We use indexes to keep byte offsets for each record in the original file. The byte offsets allow us to find the beginning of each successive record and compute the length of each record. We look up the position of a record in the index file, and then seek to the record in the data file.

- **Placing a Delimiter at the End of Each Record:**

Our choice of a record delimiter for the data files is the end-of-line (new-line) character.

2.3 Operations Performed on a File

Insertion

The system is initially used to add student records containing the details of each and every placement, into the file. Here company's name is a secondary data, therefore records with duplicate composer's name fields are not allowed to be inserted. The length of the USN is checked to see whether it contains only 10 characters. The name of the student is entered, along with the company name in which he/she is placed and the package which was offered.

Display

The system can then be used to display existing records of all the students that have been placed in a company in that year. The records are displayed based on the ordering of USN maintained in the inverted list, which here is, an ascending order. Only records with references in the index file are displayed. This prevents records marked as deleted (using '$') from being displayed. There is also an option to display the secondary index file.

Search

The system can then be used to search for existing records of all departments. The user is prompted for a usn, which is used as the key in searching for records. The index file is searched to obtain the desired starting byte address, which is then used to seek to the desired data record in any of the semester files. The details of the requested record, if found, are displayed, with suitable headings on the user's screen. If absent, a "record not found" message is displayed to the user.

Delete

The system can then be used to delete existing records from placement records. The reference to a deleted record is removed from index while the deleted record persists in the data file. A '$' is placed in the first byte of the first field (USN) of the record, to help distinguish it from records that should be displayed. The requested record, if found, is

marked for deletion, a "record deleted" message is displayed, and the reference to it is removed from the index file. If absent, a "record not found" message is displayed to the user.

Modify

The user is prompted to enter the usn, whose record is to be modified. If there are no records in the file, a "no records found" message is displayed, and the user is prompted to press any key to return back to the menu screen. If there is at least 1 record in the file, the user is prompted for the option to update the name, department, company name or the package. The user has to give his choice as 'y' or 'Y' to update the record.
After all the values are accepted, a "record updated" message is displayed and the user is prompted to press any key to return back to the menu screen.

2.4 Indexing Used

B^+ TREE

A B^+ TREE of simple indexes on the primary key is used to provide direct access to data records. Each node in the B^+ TREE consists of a primary key with the reference to record. The primary key field is the usn number field while the reference field is the starting byte offset of the matching record in the data file. Each B^+ TREE node can have max of 2 child node. The usn no. is stored in the B^+ TREE, and hence written to an index file. On retrieval the matched usn no. is used, before it is used to seek to a data record, as in the case of requests for a search, delete, modify operation. As records are added, nodes of the B^+ TREE undergo splitting (on detection of overflow), merging (on detection of underflow) or redistribution (to improve storage utilization), in order to maintain a balanced tree structure.

The data files are entry-sequenced, that is, records occur in the order they are entered into the file. The contents of the index files are loaded into their respective B^+ TREEs, prior to use of the system, each time. Each B^+ TREE is updated as requests for the available operations are performed. Finally, the B^+ TREEs are written back to their index files, after every insert, delete and modify operation.

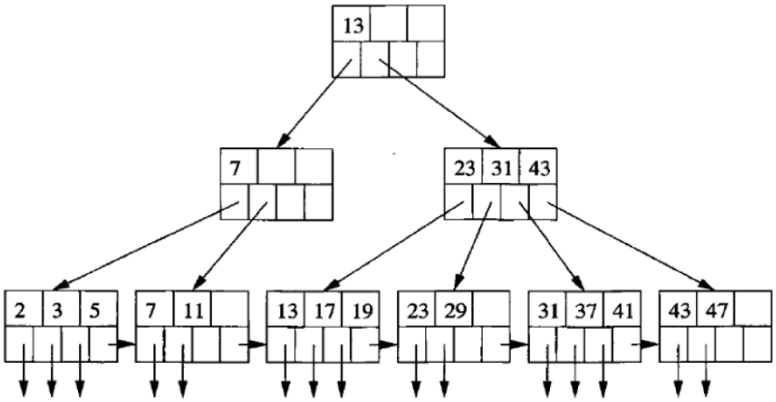

Figure 2.1 Example of a B^+ Tree