

Examination Surveillance TOTAL Database System

نظام قاعدة بيانات لمراقبة الامتحانات مع نموذج للتقييم

Dr. Hany M Harb
Faculty of Engineering, Azhar University
Nasr City, Cairo

الخلاصة - يقدم البحث قاعدة بيانات لنظام مراقبة امتحانات مطبقة على قاعدة بيانات TOTAL مع تصميم لغة بسيطة للتخاطب مع النظام حيث تسمح بالاضافة والتعديل والحذف والاسترجاع مع العلم بأن TOTAL لا تقدم هذا النوع من اللغات الخاصة - والبحث يقدم نموذجاً للتقييم نظم قواعد البيانات فيه تتم المقارنة بين أزمنة البحث والاسترجاع العملاقة والمستنتحة بواسطة نموذج التقييم .

Abstract

This paper introduces a database system for the examination committee surveillances. The system is implemented in TOTAL database management system. The end user interacts easily with the menu-driven TOTAL surveillance system through an application program. The application program acts as a simple query language since TOTAL does not support built-in query language. This query language allows the user to insert, modify and delete data in the database. It also allows the user to check data about examination and staff data sets and outputs surveillance tables for a faculty member, a grade or a date. An evaluation model has been built to evaluate the implemented system performance. search and retrieval times for output lists are measured and contrasted against modeled times under different loads.

Keywords: Database, System Analysis and Design, TOTAL Database, Performance Evaluation.

Table Of Contents

Introduction
The Surveillance System Design and Implementation
The System Evaluation
Conclusion

1. Introduction

Database technology allows an organization's data to be processed as an integral whole. Integration of data offers several important advantages. The integrated data should be compatible. In effect, more information can be obtained from existing data. Information may be considered as knowledge derived from stored data. More advantages of database systems are full (or partial) elimination of data duplication and application program/data independence.

Due to above advantages many practical information systems

are built as database applications. This paper introduces a surveillance of university examination system as an information system to be analyzed, designed and implemented in TOTAL database.

TOTAL represents data in network structure. It consists of master and variable data sets. A master data set may be linked to zero variable data set (independent master) or it may be linked to one or more variables. A variable data set has to be dependent and must be attached to one or more than one master data sets. The logical master records can be accessed directly by respective control keys. The logical variable records are chained in groups. Each variable group is attached to a unique logical master record in a related master data set. The chaining provides the access paths which are stored within data records, hence there is no need to have separate cross references. A master record contains a pointer to the first record of the variable chained group and another pointer to the last record of the same variable group. A variable record has two access paths to the next record and to the previous record in the variable group. A variable record may have more than one control key, one for each group to which it is chained.

This network of data sets (masters and variables) is defined, described and maintained as a unit separate from the application programs by database definition language. TOTAL provides also a data manipulation language (dml) to access and manipulate the data. A user may use dml in an application program which is written in a programming language supporting subroutine call statement to call TOTAL. An application program may work with a portion of the whole database for which schema has to be declared in the application program.

Section 2 describes the surveillance examination system analysis, design and implementation. Section 3 evaluates the implemented system and section 4 concludes the paper. Sample of output tables are given in Appendix A.

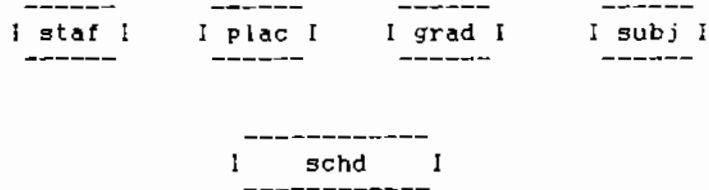
2. The Surveillance System Design and Implementation

The system handles different types of data: staff data, place data, date data, grade data and subject data. The interpretation of the associations between these data types is that grade is examined on subject at date in place under staff observation. An example of faculty of engineering examination is considered as samples of data to be loaded in the database and to be tested for system evaluation. The grade may be considered as group, section or whole rank of students. For our sample of data, we describe the faculty grades as: first year electrical (1elc), second year electrical (2elc), third year civil (3civ), fourth year mechanical (4mec) and so on. It is assumed that there is only one examination period per grade per day.

2.1 Data Definition

For staff data type, staff identification number (id), name and degree are considered. The staff attendance may be kept track. Place number, place name and place size (in terms of maximum allowable number of seats) are considered for place data set. It is also considered the grades as grade id, grade specialization and the grade total number of students who have the right to attend the exams. The subjects of each grade are analyzed to keep data about exam period for each subject and other comments (like open book exam or not).

We implement the system as four master records and one variable record linked to all of them as in the following figure.



TOTAL record name is restricted to be four characters. The staff, grade, subject and place master records have names as staf, subj, grad, and plac respectively. The control keys for staff, grade, subject and place master records are s_id, g_id, j_id and p_id respectively. schd is the schedule variable record consisting of the four control fields since each logical variable record is chained into four different groups as the staff group, place group, subject group and grade group. The logical variable record structure is shown below.

```

-----
| s_id | p_id | g_id | j_id | date | period | attendance |
-----
    
```

There are three extra fields in schd variable record which are date, period and attendance fields. The date and period correspond to the exam date and period. The attendance field keeps track of the staff attendance on this date and period.

2.2 Application Programs

The end user may communicate the menu-driven system through an application program which is written in Fortran. The application program is independent of data definition. The program supports three types of queries: summary type, report type and yes/no type. The summary type contains queries as:

- . number of observations for a staff
- . number of attendance and absence for staff
- . number of exams for a grade
- . number of busy days for a place
- . what are the examination places for a grade
- . what is the examination date for a subject

The report type includes queries demanding lists as its response. Examples of this type are:

- . a staff surveillance schedule
- . a place examination schedule
- . a grade examination table
- . list of exams held at given date

The yes/no type asks for yes or no response like checking if certain staff was absent at certain date.

The application program may also be used to load and update the database files through Add, Delete, Update functions.

We consider for demonstration and evaluation purposes four report queries in details: list staff surveillance schedule, list place examination schedule, list grade examination table and list date examination schedule. The opening menu of the system has three options: Access, Manipulate and Exit and they are self explanatory. By selecting A for Access the system responds by showing query type menu which contains three types as discussed before: Report, Summary, Yes/no and Upper_menu. By selecting R for Report the system responds by showing reporting menu with four items where each item corresponds to a report. The four reports are listed above. The user may select one of them (by typing s for staff schedule, p for place schedule, g for grade table or d for date schedule) or return to upper menu which is query type menu. By selecting s the system responds by asking the user to input the staff id. By accepting the staff id the system prints error message if it is illegal id (the input staff id does not match any s_id in the staff master records otherwise the system asks the user to select one of two output media (Terminal or Printer) to continue processing the query. The system returns to the reporting menu when the processing is done. The other reports are processed the same way with an appropriate input (grade id for grade table, place id for place schedule and date for date schedule). The flow charts of staff schedule, grade table and date schedule are given in figures 1, 2 and 3 respectively. Sample of output lists are given in Appendix A.

3. The Surveillance System Evaluation

The two considered parameters for the surveillance system evaluation are search time (st) and retrieval time (rt). The search time is the time to locate a master record. The retrieval time is the consumed time to access a list of variable records. The disk characteristics factors affecting the search and retrieval times are transfer time per unit of storage (ut), blocking factor (bf), size of records stored in the block (rs), block random access time (bt), cylinder size (cs), number of blocks in cylinder (cf) and cylinder random access time (ct).

The search time may be approximated to be the cylinder random access time (ct) plus the block random access time (bt) plus the random access time of a master record (rt) plus the multiplication of the transfer time of storage unit (ut) by master record length (ml) as expressed in equation (1). The rt depends on the physical storage policy of TOTAL files.

The retrieval time of a variable list depends on the list length. The list length may be expressed as the number of cylinders (cn) and the number of blocks in a cylinder (bn) on which the list expands. So the retrieval time may be considered as the sum of three parts of time: the time to randomly seek a cylinder multiplied by the number of cylinders on which the list expands (may be called as variable cylinders), the time to randomly access a block multiplied by the number of blocks on which the variable list are stored within a cylinder (may be called variable blocks) for all the variable cylinders (cn) and the time to transfer a unit of storage multiplied by the variable record length (vl) multiplied by the number of variable records per a variable block per a variable cylinder (rn) for each variable block for each variable cylinder.

In general case bn differs from cylinder to cylinder and rn differs from block to block. The distribution of variable blocks on cylinders and the variable records on blocks depend on the requested list (query) and other factors like the secondary storage allocation policy of the operating system under which the TOTAL dbms is running. Equation (2) shows the general expression to compute the retrieval time and it depends on the query since the variable list is supposed to be query dependent.

For simplicity we may assume that the variable blocks and the variable records are uniformly distributed on the variable cylinders and the variable blocks respectively and we may also assume that the variable records number for all the lists are the same (taken as the total number of the loaded variable records) which means that bn and rn are taken as constants. The retrieval time under above assumptions is expressed in equation (3).

$$st = ct + bt + rt + ut * ml \quad (1)$$

$$rt = (cn * ct) + (bt * \sum_{i=1}^{cn} b_i) + (vl * ut * \sum_{i=1}^{cn} \sum_{j=1}^{bn} r_{ij}) \quad (2)$$

where

b_i number of blocks on which the variable records of the required list expanded in cylinder i .

r_{ij} number of variable records of the required list in block i , cylinder j .

$$rt = cn * ct + cn * bn * bt + cn * bn * rn * vl * ut \quad (3)$$

The system was implemented on interdata 32 bit machine with disk characteristics as shown in table 1.

Table 1: Disk Drive Characteristics

Factor	Value
ut	8 us
bs	512 by
bt	.2 ms
cs	5 mb
ct	.5 ms
ms: millisecond	mb: megabyte
us: microsecond	by: byte

Following the above evaluation model, we can compute the time for processing a database query under the assumption that

forty records are loaded into the database. We only consider the report query type. Consider the schedule staff list as discussed in above section. The user inputs the staff number and then the system locates the master (assuming it is a valid staff id) and then retrieve the variable list of this staff. The staff master record length (ml) is 48 byte and the variable record length is 50 byte. It is assumed that all loaded variable records are stored in one cylinder and they are uniformly distributed on the blocks. Since the block length is 512, the variable record length is 50 and the number of loaded variable records are 40, so there are 10 variable records per block and the forty records are stored in 4 blocks. As discussed before it is assumed that all the forty variable records are chained with the inputted staff id. rt is taken to be .1 ms.

$$\begin{aligned} st &= ct + bt + rt + ut * ml \\ &= .5 \text{ (ms)} + .2 \text{ (ms)} + .1 + 8 \text{ (us/byte)} * 48 \text{ (byte)} \\ &= 1.184 \text{ ms} \end{aligned}$$

$$\begin{aligned} rt &= cn * ct + bn * bt + cn * bn * rn * vl * ut \\ &= 1 \text{ (cylinder)} * .5 \text{ (ms/cylinder)} \\ &\quad + 5 \text{ (block)} * .2 \text{ (ms/block)} \\ &\quad + 1 \text{ (cylinder)} * 4 \text{ (block/cylinder)} \\ &\quad \quad * 10 \text{ (record/block)} * 50 \text{ (byte/record)} \\ &\quad \quad * 8 \text{ (us/byte)} \\ &= .5 \text{ (ms)} + 1 \text{ (ms)} + 16 \text{ (ms)} \\ &= 17.5 \text{ ms} \end{aligned}$$

The processing time for staff schedule list is measured through the application program to be compared with the modeled processing time. Under the same load (40 variable records) the measured results are as following:

$$\begin{aligned} \text{average locate master record} &= .7 \text{ ms} \\ \text{average access variable list} &= 13.4 \text{ ms} \\ \text{total processing time} &= 14.1 \text{ ms.} \end{aligned}$$

The $\sin\alpha$ and $\sin\theta$ times of TOTAL database is not considered. The difference between the measured and modeled processing times are due to the small load. We did the same calculations and the measurements under different load. Fig. 4 shows the relationships between the modeled time and the measured time under different loads. We may notice that the more loaded variable records is the closer between the modeled and the average measured times.

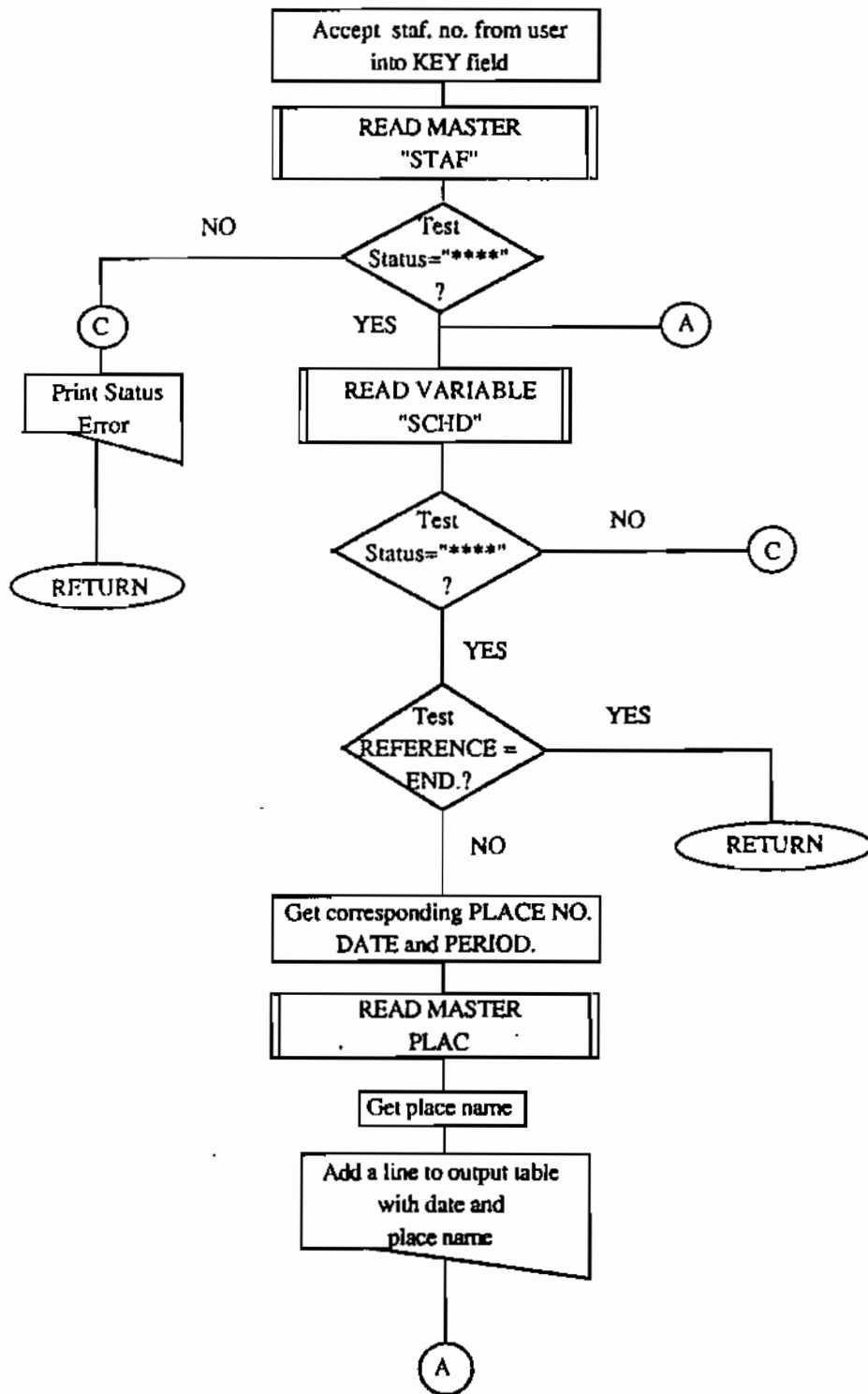


FIG. (1) Staff Surveillance schedule

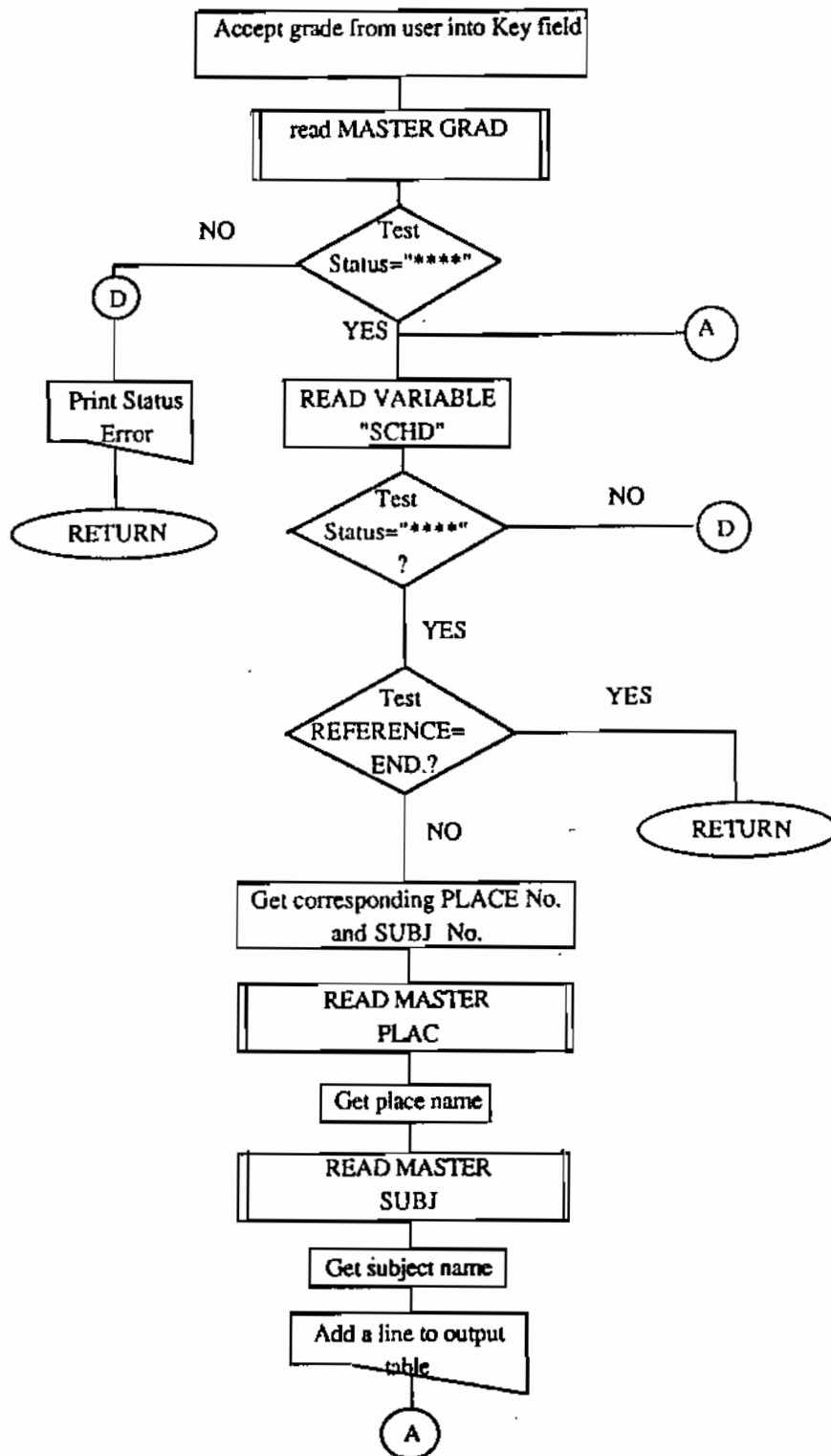


FIG. (2) Grade Examination Table

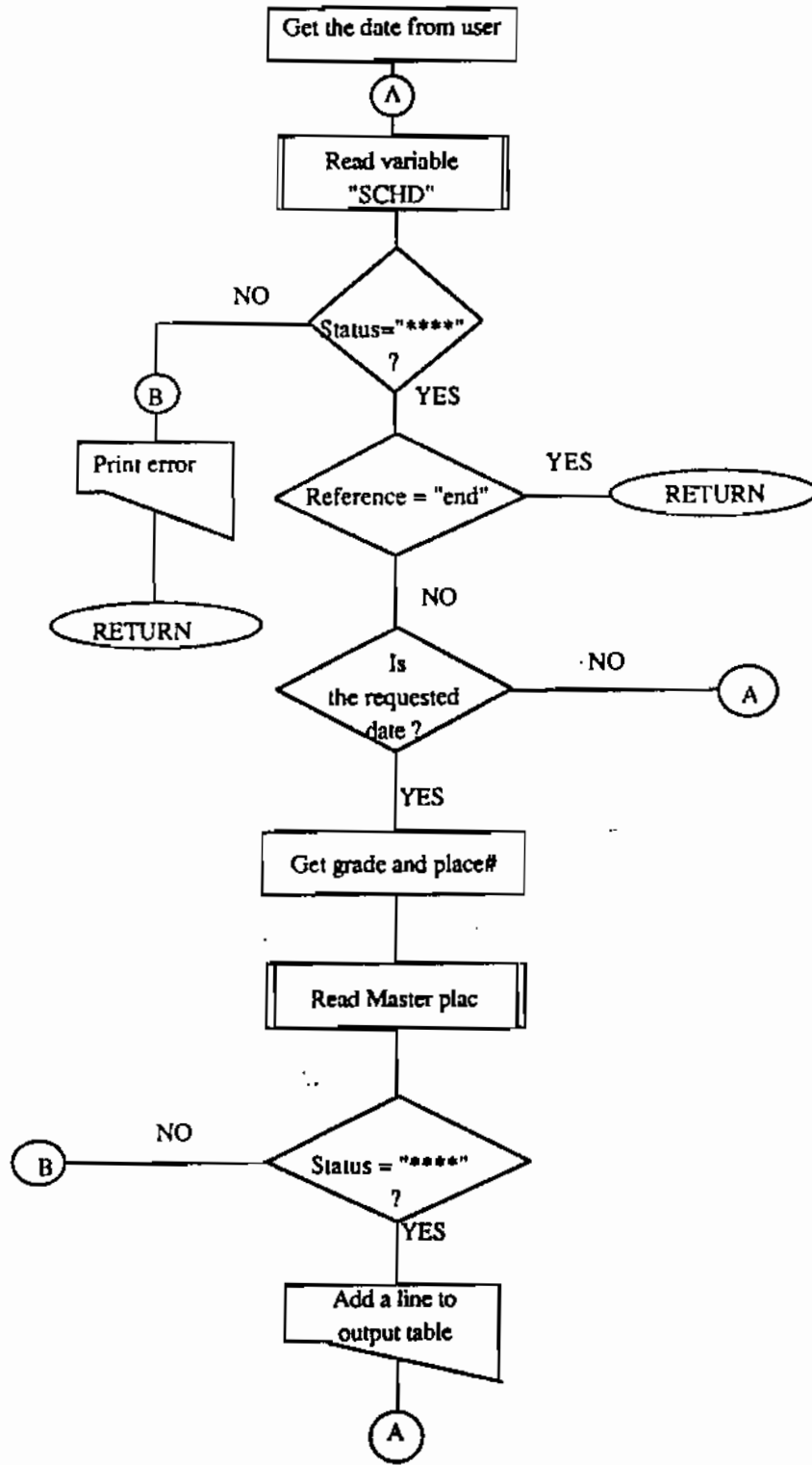


FIG. (3) Date Examination Schedule

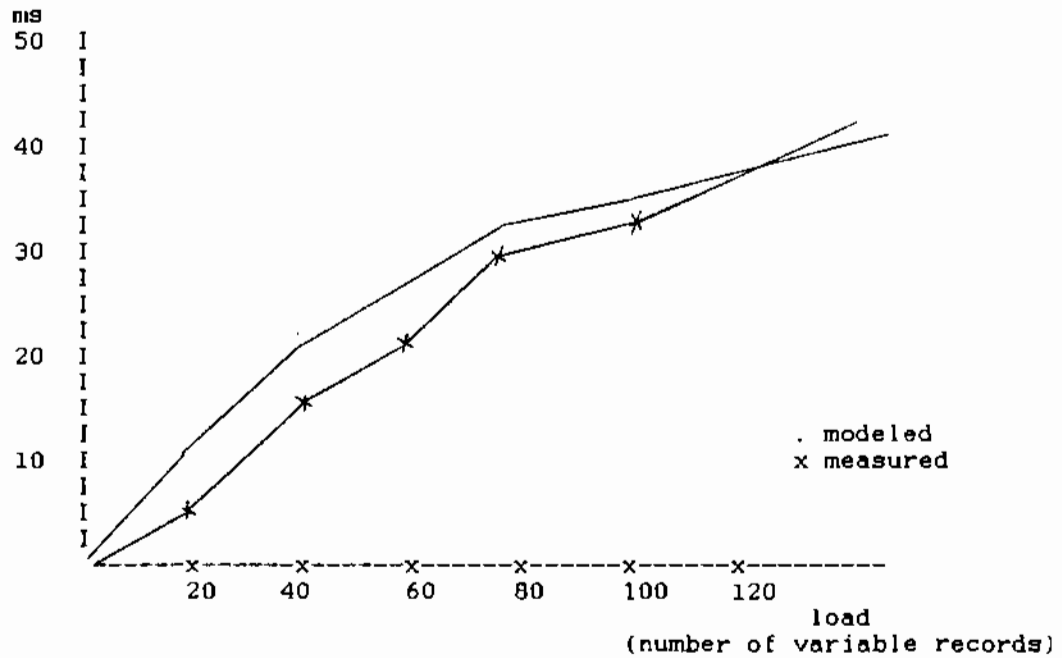


Figure 4: Modeled and Measured times

4. Conclusion

The paper introduced the examination committee surveillances as an information system. The system is implemented in TOTAL database management system on Perkin Elemer 32-bit minicomputer. The end user interacts easily with the TOTAL surveillance system through an application program. The application program is written in Fortran and it acts as a simple query language since TOTAL does not support built-in query language. This query language supports three types of queries: the summary, the yes/no and the report types. It also allows the user to insert, modify and delete data in the databasa. A model for the system evaluation was introduced. Search and retrieval times for the system are measured and contrasted against modeled times under different load levels.

References

- [1] TOTAL Reference Manual, Interdate, MA, USA, 1971.
- [2] David Krneke, "Data Processing," Science Research Associates, NY, 1977.
- [3] K.F. Siler, "A Stochastic Evaluation Model for Database Organization in Data Retrieval Systems," Communication of ACM, Vol. 19, No. 2, Feb. 1976, pp. 85-95.
- [4] C.J. Date, "An Introduction to Database Systems," Vol. 1, Fourth Edition, Addison-Wesley Publishing Company, 1986.
- [5] James A. Senn, "Analysis and Design of Information Systems," McGraw-Hill Company, 1984.
- [6] A. Dessoki, M. Shoky, "Personal Computer Software for Use in Student Graduate Affair and Examination System," Thirteen Congress for Statistics and Computer Science, Ain-Shams University, Cairo, Egypt, Vol. 15, March 1988, pp. 319-336.

Appendix A

Output Samples

Surveillance Schedule for Staff: Hoda Mohamed

Date	Period	place name
01/05/88	9:00 - 12:00	upper hall
04/05/88	9:00 - 12:00	cafteria
13/05/88	9:00 - 13:00	cafteria
04/06/88	9:00 - 12:00	resturant

Examination Table at Date: 04/05/88

place name	grade	period
cafteria	1elc	9:00 - 12:00
cafteria	.3civ	13:00 - 16:00
upper hall	2mec	9:00 - 12:00

Examination Table for Grade: 1elc

date	period	subject	place name
04/05/88	9:00 - 12:00	logic	cafteria
07/05/88	9:00 - 12:00	electric	cafteria
11/05/88	9:00 - 12:00	math	cafteria
04/06/88	9:00 - 12:00	drawing	resturant