ISSN No. 0976-5697

Volume 8, No. 4, May 2017 (Special Issue)

International Journal of Advanced Research in Computer Science

REVIEW ARTICLE

Available Online at www.ijarcs.info

Overview of the Database Management System

Prabhjot,
MCA Scholar
Computer Application and Information Sciences
Sant Baba Bhag Singh University
Jalandhar, India

Neha Sharma
MCA Scholar
Computer Application and Information Sciences
Sant Baba Bhag Singh University
Jalandhar, India

***Abstract* - Database Management System is a computer software designed for the purpose of managing the database. Database management System. The main aim of a DBMS is to supply a way to store up and retrieve database information that is both convenient and efficient. This paper presents the overview of database management system such as its components, data models which are used for describing how data is to be arranged for a specific purpose. One of the major aim of database is to supply users with an abstract view of data. Normalization of data plays an important role in the database development process. By using this process a developer can eliminate duplication and develop standards by which all data can be measured.**

***Keywords*-DBMS, components, applications, operations, view, relation, entities, redundancy, normalization.**

1. INTRODUCTION

A Database Management System (DBMS) is basically a collection of programs that enables a user to insert, modify, retrieve data and thus make it easy to design, maintain and access databases. The DBMS relieves the user from knowing how the data is stored physically and complex algorithms used for performing operations on the database. It only concentrate on how the operations are to be performed to retrieve the data from the database. The DBMS is also in charge of access, security, storage and host of other functions for the database system. It does through a selection of computer programs. This allows it to manage a large, structured set of data, which makes up the database and provides access to the data for multiple, concurrent users while maintaining the integrity of the data.[1][2].

1.1. Applications Of DBMS

A. *Railway Reservation System:-* Database is required to keep record of ticket booking, train's departure and arrival status. Also if trains get late then people get to know it through database update.

B. *Library Management System:-* There are thousands of books in the library so it is very difficult to keep record of all the books in a copy or register. So DBMS used to maintain all the information relate to book issue dates, name of the book, author and availability of the book.

C. *Banking:-* We make thousands of transactions through banks daily and we can do this without going to the bank. So how banking has become so easy that by sitting at home we can send or get money through banks. That is all possible just because of DBMS that manages all the bank transactions.

D. *Universities and colleges:-* Examinations are done online today and universities and colleges maintain all these records through DBMS. Student's registrations details, results, courses and grades all the information are stored in database.

E. *Credit card transactions:-* For purchase of credit cards and all the other transactions are made possible only by DBMS. A credit card holder knows the importance of their information that all are secured through DBMS.

F. *Social Media Sites:-* We all are on social media websites to share our views and connect with our friends. Daily millions of users signed up for these social media accounts like facebook, twitter, pinterest and Google plus. But how all the information of users are stored and how we become able to connect to other people, yes this all because DBMS.

G. *Telecommunications:-*Any telecommunication company cannot even think about their business without DBMS. DBMS is must for these companies to store the call details and monthly post paid bills.

H. *Finance:-* Those days have gone far when information related to money was stored in registers and files. Today the time has totally changed because there are lots f thing to do with finance like storing sales, holding information and finance statement management etc.

I. *Military:-* Military keeps records of millions of soldiers and it has millions of files that should be keep secured and safe. As DBMS provides a big security assurance to the military information so it is widely used in militaries. One can easily search for all the information about anyone within seconds with the help of DBMS.

J. *Online Shopping:-* Online shopping has become a big trend of these days. No one wants to go to shops and waste his time. Everyone wants to shop from home. So all these products are added and sold only with the help of DBMS. Purchase information, invoice bills and payment, all of these are done with the help of DBMS.

K. *Human Resource Management:-* Big firms have many workers working under them. Human resource management department keeps records of each employee's salary, tax and work through DBMS.

L. *Manufacturing:-* Manufacturing companies make products and sales them on the daily basis. To keep

CONFERENCE PAPERS
National Conference on Emerging Trends on Engineering & Technology (ETET-2017)
On 21st April 2017
University Inst. of Engg. & Tech. & University Inst. of Computer, SBBS University, Punjab (India)

records of all the details about the products like quantity, bills, purchase, supply chain management, DBMS is used.

M. *Airline Reservation system:-* Same as railway reservation system, airline also needs DBMS to keep records of flights arrival, departure and delay status[8].

The various commercially available database management system,

- For the small organizations- MS-Access, File Maker Pro, DB Text Works, Superbase etc.
- For the enterprise (client /server)- Oracle, SQL Server, My SQL, Sybase, DB2, Informix,Paradox, Dbase etc.[1]

1.3. Purpose Of Database Management System

DBMS were developed to handle all the drawbacks of typical file-processing systems supported by conventional operating systems. Some of the purpose that attain by DBMS are Data redundancy and inconsistency, Difficulty in accessing data, Data isolation, Integrity problems, atomicity of updates, concurrent access by multiple users and security problems.

- Drawbacks of using file systems to store data:
  - Data redundancy and inconsistency
    - ✓ Multiple file formats, duplication of information in different files.
  - Difficulty in accessing data.
    - ✓ Need to write a new program to carry out each new task.
  - Data isolation — multiple files and formats
  - Integrity problems.
    - ✓ Integrity constraints (e.g. account balance > 0) become part of program code.
    - ✓ Hard to add new constraints or change existing ones
  - Atomicity of updates
    - ✓ Failures may leave database in an inconsistent state with partial updates carried out
    - ✓ E.g. transfer of funds from one account to another should either complete or not happen at all
  - Concurrent access by multiple users
    - ✓ Concurrent accessed needed for performance
    - ✓ Uncontrolled concurrent accesses can lead to inconsistencies–E.g. two people reading a balance and updating it at the same time Security problems
  - Database systems offer solutions to all the above problems. [3]

1.5. Disadvantages of DBMS

Database System has following disadvantages:-

A. *Cost:-* DBMS requires high initial investment for hardware, software and trained staff. A significant investment based upon size and functionality of organization if required. Also organization has to pay concurrent annual maintenance cost.

B. *Complexity:-* A DBMS fulfill lots of requirement and it solves many problems related to database. But all these functionality has made DBMS an extremely complex software. Developer, designer, DBA and End user of database must have complete skills if they want to user it properly. If they don't understand this complex system then it may cause loss of data or database failure.

C. *Database Failure:-* As we know that in DBMS, all the files are stored in single database so chances of database failure become more. Any accidental failure of component may cause loss of valuable data. This is really a big question mark for big firms.

D. *Extra Cost of Hardware:-* A DBMS requires disk storage for the data and sometimes you need to purchase extra space to store your data. Also sometimes you need to a dedicated machine for better performance of database. These machines and storage space increase extra costs of hardware.

E. *Size:-* As DBMS becomes big software due to its functionalities so it requires lots of space and memory to run its application efficiently. It gains bigger size as data is fed in it.

F. *Performance:-* Traditional files system was very good for small organizations as they give splendid performance. But DBMS gives poor performance for small scale firms as its speed is slow. [4]

1.5. Components Of DBMS

DBMS have several components, each performing very significant tasks in the database management system environment. Below is a list of components within the database and its environment.

- *Software:-* This is the set of programs used to control and manage the overall database. This includes the DBMS software itself, the Operating System, the network software being used to share the data among users, and the application programs used to access data in the DBMS.
- *Hardware:-* Consists of a set of physical electronic devices such as computers, I/O devices, storage devices, etc., this provides the interface between computers and the real world systems.
- *Data:-* DBMS exists to collect, store, process and access data, the most important component. The database contains both the actual or operational data and the metadata.
- *Users:-* The users are the people who manage the databases and perform different operations on the databases in the database system. There are three kinds of people who play different roles in database system:
  - *i) Application Programmers:-* The people who write application programs in programming languages (such as Visual Basic, Java, or C++) to interact with databases are called Application Programmer[1].
  - *ii) Database Administrators:-* A person who is responsible for managing the overall database

management system is called database administrator or simply DBA. DBA is a person or group of persons who is responsible for management of the database in the database management system. DBA is a highly skilled person with strong technical background to monitor various operations such as creating, modifying, and maintaining which help in handling three levels of the database. DBA has most of the powers such as defining schemas, storage structures and access method strategies, physical organization, authorization and integrity constraints etc. DBA even grants permission to the users of the database and the stores the profiles of the users in the database. The user profile describes the activities a user can perform on the database that is whether a user can perform a given operation or not. So the DBA has all the power that system can give on all the database objects. DBA is top level authority among all persons connected to the database.[1]

iii) *End-Users:*- The end-users are the people who interact with database management system to perform different operations on database such as retrieving, updating, inserting, deleting data etc.

- *Procedures:*- These are the instructions and rules that assist on how to use the DBMS, and in designing and running the database, using documented procedures, to guide the users that operate and manage it.
- *Database Access Language:*- This is used to access the data to and from the database, to enter new data, update existing data, or retrieve required data from databases. The user writes a set of appropriate commands in a database access language, submits these to the DBMS, which then processes the data and generates and displays a set of results into a user readable form.
- *Query Processor:*- This transforms the user queries into a series of low level instructions. This reads the online user's query and translates it into an efficient series of operations in a form capable of being sent to the run time data manager for execution.
- *Run Time Database Manager:*- Sometimes referred to as the database control system, this is the central software component of the DBMS that interfaces with user-submitted application programs and queries, and handles database access at run time. Its function is to convert operations in user's queries. It provides control to maintain the consistency, integrity and security of the data.
- *Data Manager:*- Also called the cache manger, this is responsible for handling of data in the database, providing a recovery to the system that allows it to recover the data after a failure.
- *Database Engine:*- The core service for storing, processing, and securing data, this provides controlled access and rapid transaction processing to address the requirements of the most demanding data consuming applications. It is often used to create relational databases for online transaction processing or online analytical processing data.
- *Data Dictionary:*- This is a reserved space within a database used to store information about the database itself. A data dictionary is a set of read-only table and views, containing the different information about the data used in the enterprise to ensure that database representation of the data follow one standard as defined in the dictionary.
- *Report Writer:*- Also referred to as the report generator, it is a program that extracts information from one or more files and presents the information in a specified format. Most report writers allow the user to select records that meet certain conditions and to display selected fields in rows and columns, or also format the data into different charts.[5]

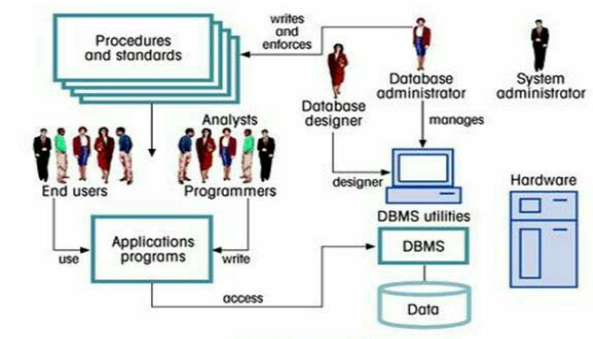


Fig 1. DBMS components and its environment

2. VIEW OF DATA

A Database system is a collection of interrelated data and a set of programs that allow users to access and modify the data. A major purpose of a database system is to provide users with an abstract view of data. That's the system hides certain details of how the data are stored and maintained.

2.1. Data Abstraction

For the system to be usable, it must retrieve data efficiently. The need for efficiency has led designer to use complex data structures to represent data in the database. Since many database system users are not computer trained, developers hide the complexity from users through several levels of abstraction, to simplify user's interaction with the system:

A. *Physical level:-* The lowest level of abstraction describes *how* the data are actually stored and also describes the data structures and access methods to be used by the database.

B. *Logical level :-* The next-higher level (also known as conceptual level) describes *what* data are stored in the database, and what relationships exist among those data. The logical describes the entire database in terms of a small number of relatively simple structures. Although implementation of the simple structures at the logical level may involve complex physical-level structures, the user of the logical level doesn't need to be aware of this complexity. Database administrators, who must decide what information to keep in the database, use the logical level of abstraction.

C. *View level:-* The view level also called as external level is the highest level of abstraction. It is the users' view of the database. This level describes that part of the database that is relevant to the user. Therefore, same database can have different views for different users. This level is the one which is closest to the end-users. The view level of abstraction exists to simplify their interaction with the database. The system may provide many views for the same database. [6]

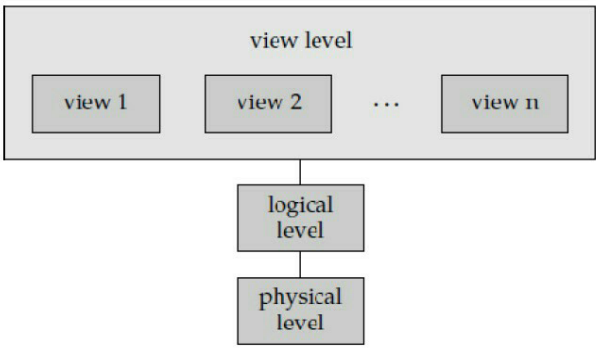
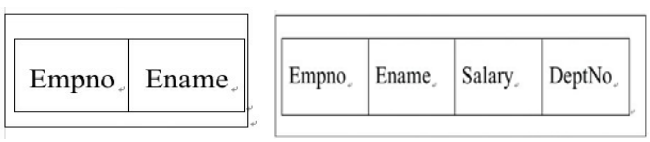


Fig.2. The level of Data Abstraction

Below the internal level is the physical level which is managed by the operating system under the direction of the DBMS. It deals with the mechanism of physically storing data on a device such as disk. So the internal schema does not deal with device specific issues such as cylinder or track size.

External view:-



Global view:-

EMPLOYEE
Empno: Integer(4) key
Ename: String (15)
Salary: String (8)
DeptNo: Integer(4)
Post: String (15)

Internal view:-

STORED_EMPLOYEE record length 60
Empno: 4 decimal offset 0 unique
Ename: String length 15 offset 4
Salary: 8,2 decimal offset 19
DeptNo: 4 decimal offset 27
Post: string length 15 offset 31

Fig. 3 different view of data

2.2. Instances and Schemas in DBMS

Definition of schema:- Design of a database is called the schema. Schema gives the names of the entities and attributes. It specifies the relationship among them are associated. Schema is of three types: Physical schema, logical schema and view schema[7].

- The design of a database at physical level is called *physical schema* , how the data stored in blocks of storage is described at this level.
- Design of database at logical level is called logical schema, programmers and database administrators work at this level, at this level data can be described as certain types of data records gets stored in data structures, however the internal details such as implementation of data structure is hidden at this level (available at physical level).
- Design of database at view level is called view schema. This generally describes end user interaction with database systems.

Consider a sales record database consisting of three files ITEM, CUSTOMER, and SALES. These three files are the schema of the database. The schema diagram for this database is as shown in Fig.

ITEM:

Item_ID	Item_Desc	Item_Cost

CUSTOMER :

Cust_ID	Cust_Name	Cust_Addr	Cust_Bal

SALES:

Cust_ID	Item_ID	Item_Qty	Item_Price

Fig. 4. Schema Diagram

Definition of instance:- The data stored in database at a particular moment of time is called instance of database. Database schema defines the variable declarations in tables that belong to a particular database; the value of these variables at a moment of time is called the instance of that database. Every time any modification(update, insert or delete operation) is performed on the database, the database changes from one state to another.

TABLE 1

INSTANCES OF CUSTOMER RELATION

Cust_ID	**Cust_Name**	**Cust_Addr**	**Cust_Bal**
24101	Ashok	Amritsar	21570
24102	Prabhjot	Hoshiarpur	44562
24103	Kamlesh	Amritsar	40578
24104	Taran	Ludhiana	34000

3. DATA MODELS

A data model [6][8] also known as database model. It is a specification describing how data how data is to be arranged to serve a specific purpose. A data model tells what information is to be contained in a database, how data items in the database will be related to each other and how information will be used. The basic building blocks of data model are entities, attributes, relationships and constraints. There are number of different data models that can be classified into different categories:

A. The Entity-Relationship Model:- The Entity-Relationship (ER) model is a high level conceptual data model developed to facilitate the database design. The E-R model is A high level description of the data and the relationships among the data, rather than how data is stored. It focuses on identifying the relationships among the entities[8]. It is a non-technical method that provides a standard and logical way of visualizing the data and free from ambiguous. It provides basic understanding of The nature of data and how it is used by enterprise.

B. Object -Based Data Model:- The object oriented model is based on collection of objects, attributes, and relationships which together form the static properties. It also consists of integrity rules over objects and dynamic properties such as operations or rules defining new database states. An object is a collection of data and methods. When different object of same type are grouped together, they form a class. This model is used basically for multimedia applications as well as data with complex relationships. The object model is represented graphically with objects diagrams containing object classes. Classes are arranged into hierarchies sharing common structure and behavior and are associated with other classes. Example of software that follow this model is FastObject.Net etc.

C. Relational Model:- Relational Model is today the primary data model for commercial data-processing applications. It has attained its primary portion because of its simplicity. A relational database consists of a collection of tables, each of which is assigned a unique name. A row in a table represents a relationship among a set of values. Informally, a table is an entity set and a row is an entity. Since a table is a collection of such relationships. The operations performed on the database in a relational model are insertion, deletion, updation very easily. The relational database provides flexibility that allows changes to database structure to be easily accommodated. Ability to easily take advantage of new hardware technology which make things easy for the users. Relational model is not suitable for huge databases but suitable for small databases. The user need not to know the complex physical data storage.

TABLE 2

RELATION OF STUDENT

CONFERENCE PAPERS
National Conference on Emerging Trends on Engineering & Technology (ETET-2017)
On 21st April 2017
University Inst. of Engg. & Tech. & University Inst. of Computer, SBBS University, Punjab (India)

primary key column

Rno	Name	Age	Class	Marks	St_Id
1	Prabh	15	10th	68	1001
2	Neha	15	10th	70	1002
3	Taran	14	9th	60	1003
4	Sajan	13	9th	88	1004
5	Manav	14	10th	89	1005

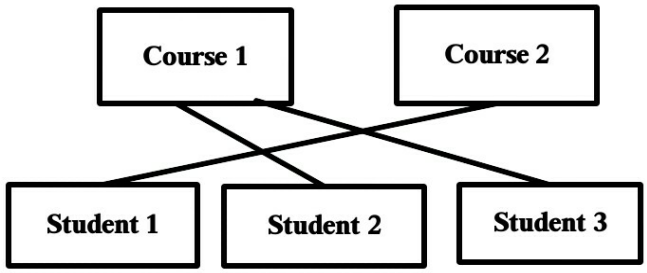


Fig. 6. Network Model

Which data model to use?

Data models are necessary as they provide access methods and data structure for defining a DBMS. Data model describe the organization of data along with operations that manipulate the data. The data model which is best for the organization depends upon the following factors.

i. Primary goals and requirements of the organization

ii. The volume of daily transactions that will be done.

iii. The estimated number of queries that will be made by the organization.

Among the transitional data models, the most preferred one is the relational data model because it can be used for representing most of the real world objects and the relationships among them. Security and integrity are maintained easily by the relational data model. It increases the productivity of the application programs. Moreover, relational tables show only The logical relationship. End-users need to know the exact physical structure of a table.

Network model is also free from anomalies but due to its complex nature it is not a preferred model. Since, hierarchical model suffers from lot of anomalies it is useful for those cases which are hierarchical in nature.

D. *Hierarchical Model:-* Hierarchical model is a record based data model that are used in describing data at the conceptual level. This model is based on the Bill of Materials (BOM). The hierarchical data model organizes the data in the form of tree structure that represents the parent-child relationship. The requirement of this model is that each child record can be linked to only be one parent and child can only be reached through its parent. In this, lack of declarative querying facilities. It generates only one-to-one and one-to-many relationships. It is very difficult to modify and design of the database. Insertion, updation , and deletion is very difficult.

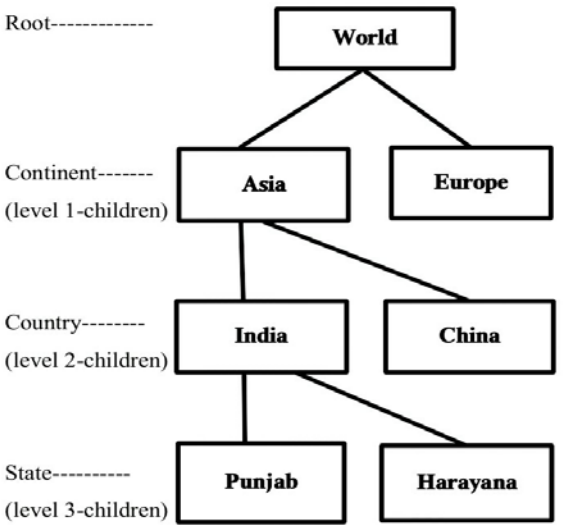


Fig. 5. A Hierarchical Model

E. *Network Model:-* Network Model removes the limitations of the hierarchical model with a graph. In this, we link all the records by using a chain or pointers. The data access is easier and flexible than the hierarchical model. It has the ability to handle many-to-many relationships. This model does not suffer from insertion anomalies, update anomalies and deletion anomalies, also the retrieval operation is symmetric, as compared to hierarchical model, but the main disadvantage is its complexity.

4. DBMS LANGUAGES

A. A DBMS must provide appropriate languages and interfaces for each category of users to express database queries and updates. Database Languages are used to create and maintain database on computer . There are large numbers of database languages like Oracle, MYSQL, MS Access, dBase, FoxPro etc. SQL statements commonly used in Oracle and MS Access can be categorized as data definition language (DDL), data control language (DCL) and data manipulation language (DML).[7]

B. *Data Definition Language (DDL)* :- It is a language that allows the users to define data and their relationship to other types of data. It is mainly used to create files, databases, data dictionary and tables within databases. It is also used to specify the structure of each table, set of associated values with each attribute, integrity constraints, security and authorization information for each table and physical storage structure of each table on disk.

CONFERENCE PAPERS
National Conference on Emerging Trends on Engineering & Technology (ETET-2017)
On 21st April 2017
University Inst. of Engg. & Tech. & University Inst. of Computer, SBBS University, Punjab (India)

C. *Data Manipulation Language (DML*) :It is a language that provides a set of operations to support the basic data manipulation operations on the data held in the databases. It allows users to insert, update, delete and retrieve data from the database. The part of DML that involves data retrieval is called a query language.

D. *Data Control Language (DCL*):- DCL statements control access to data and the database using statements such as GRANT and REVOKE. A privilege can either be granted to a User with the help of GRANT statement. The privileges assigned can be SELECT, ALTER, DELETE, EXECUTE, INSERT, INDEX etc. In addition to granting of privileges, you can also revoke (taken back) it by using REVOKE command.

E. *Transaction Control (TCL)* : Statements are used to manage the changes made by DML statements. It allows statements to be grouped together into logical transactions. COMMIT - save work done. SAVEPOINT - identify a point in a transaction to which you can later roll back. ROLLBACK - restore database to original since the last COMMIT. SET TRANSACTION - Change transaction options like isolation level and what rollback segment to use

In practice, the data definition and data manipulation languages are not two separate languages. Instead they simply form parts of a single database language such as Structured Query Language (SQL). SQL represents combination of DDL and DML, as well as statements for constraints specification and schema evaluation.

5. NORMALIZATION

Database normalization is the process of organizing data into tables in such a way that the results of using the database are always unambiguous and as intended. Such normalization is intrinsic to relational database theory. It may have the effect of duplicating data within the database and often results in the creation of additional tables.

While data normalization rules tend to increase the duplication of data, it does not introduce data redundancy, which is unnecessary duplication. Database normalization is typically a refinement process after the initial exercise of identifying the data objects that should be in the relational database, identifying their relationships and defining the tables required and the columns within each table .

TABLE 3

Customer Relation

Customer	Item	Price
Javed	Shirt	500
Sahibjit	Shoes	650
Abhishek	Shirt	500
Manraj	Trousers	250

If this table is used for the purpose of keeping track of the price of items and you want to delete one of the customers, you will also delete the price. Normalizing the data would mean understanding this and solving the problem by dividing this table into two tables, one with information about each customer and the product they bought and the second with each product and its price. Making additions or deletions to either table would not affect the other.

- *First normal form (1NF)*[10] This is the "basic" level of database normalization, and it generally corresponds to the definition of any database, namely: It contains two-dimensional tables with rows and columns. Each column corresponds to a sub-object or an attribute of the object represented by the entire table. Each row represents a unique instance of that sub-object or attribute and must be different in some way from any other row (that is, no duplicate rows are possible). All entries in any column must be of the same kind. For example, in the column labeled "Customer," only customer names or numbers are permitted.
- *Second normal form (2NF).* At this level of normalization, each column in a table that is not a determiner of the contents of another column must itself be a function of the other columns in the table. For example, in a table with three columns containing the customer ID, the product sold and the price of the product when sold, the price would be a function of the customer ID (entitled to a discount) and the specific product.
- *Third normal form (3NF)* [11] At the second normal form, modifications are still possible because a change to one row in a table may affect data that refers to this information from another table. For example, using the customer table just cited, removing a row describing a customer purchase (because of a return, perhaps) will also remove the fact that the product has a certain price. In the third normal form, these tables would be divided into two tables so that product pricing would be tracked separately.

Extensions of basic normal forms include the domain/key normal form, in which a key uniquely identifies each row in a table, and the Boyce-Codd normal form, which refines and enhances the techniques used in the 3NF to handle some types of anomalies.

Database normalization's ability to avoid or reduce data anomalies, data redundancies and data duplications, while improving data integrity, have made it an important part of the data developer's toolkit for many years.

6. CONCLUSION

Finally, like all software systems, the DBMS has both advantages and disadvantages. A database management system (DBMS) is a computer application program designed for the efficient and effective storage, access and update of large volumes of information. The DBMS can help users a lot on managing the database. Moreover, it comes to make things easier for users to control their database. The DBMS acts as middle software which resides between computer programs and databases, and also the DBMS interacts with the users' application program and the database at the same time. For its benefits, the first one is that the DBMS helps users to control data redundancy. That means some duplicated data are to be removed by the DBMS. Although those duplicated data are not entirely eliminated, the amounts of redundancy data are to be controlled. Without the DBMS, either programmer or the one who implemented the database will have much more work to do to arrange data not only when inserting data into the database but also when retrieving data out from the database.

REFERENCE

[1] [1] Er.Anshuman Sharma, Mr. Anurag Gupta, Jagmohan Mago (2011). *Fundamentals of DBMS.* Lakhanpal Publishers, fourth edition, 2011.

[2] [2] Database - Definition of database by Merriam-Webster.

[3] [3] Introduction.http://codex.cs.yale.edu/avi/db-book/db4/slide-dir/ch1-2.pdf

[4] [4] http://tutorialink.com/dbms/advantage-and-disadvantages-of-dbms.dbms

[5] [5]http://ecomputernotes.com/fundamental/what-is-a-database/components-of-dbms

[6] [6] A. B. Silberschatz, H. F. Korth, and S. Sudarshan. *Database System Concepts.* McGraw-Hill, sixth edition, 2011.

[7] [7] A.P.G. Brown, "Modelling a Real-World System and Designing a Schema to Represent It", in Douque and Nijssen (eds.), Data Base Description , North-Holland, 1975, ISBN 0-7204-2833-5 .

[8] [8] Beynon-Davies, Paul (2004). Database Systems . Basingstoke, UK: Palgrave: Houndmills. ISBN 1403916012.

[9] [9] J. D. Ullman and J. Widom. A First Course in Database Systems. Prentice Hall, third edition, 2008.

[10] [10] Codd, E.F. (June 1970). "A Relational Model of Data for Large Shared Data Banks" .Communications of the ACM . 13 (6): 377–387.doi : 10.1145/362384.362685.

[11] [11] C.J. Date. An Introduction to Database Systems . Addison-Wesley (1999), p. 290

CONFERENCE PAPERS
National Conference on Emerging Trends on Engineering & Technology (ETET-2017)
On 21st April 2017
University Inst. of Engg. & Tech. & University Inst. of Computer, SBBS University, Punjab (India)