Winter 12-19-2001

Developing A New Decision Support System for University Student Recruitment

Terry Mangles

Ahmed El-Ragal

DEVELOPING A NEW DECISION SUPPORT SYSTEM FOR UNIVERSITY STUDENT RECRUITMENT

Terry Mangles
University of Plymouth Business School
Drake Circus,
Plymouth, PL4 8AA
Devon, UK
Tel: +44-01752-232852
Fax: +44-01752-232853
E-mail: terry.mangles@pbs.plym.ac.uk

Ahmed El-Ragal
Arab Academy For Science and Technology
Management Information Systems Department
P.O. Box 1029, AASTMT, Miami
Alexandria, EGYPT
Tel: +2-03-5485473
Fax: +2-03-5566072
E-mail: aelragal@mis.aast.edu

ABSTRACT

This paper investigates the practical issues surrounding the development and implementation of Decision Support Systems (DSS). The paper describes the traditional development approaches analyzing their drawbacks and introduces a new DSS development methodology. The proposed DSS methodology is based upon four modules; needs' analysis, data warehouse (DW), knowledge discovery in database (KDD), and a DSS module.

The proposed DSS methodology is applied to and evaluated using the admission and registration functions in Egyptian Universities. The paper investigates the organizational requirements that are required to underpin these functions in Egyptian Universities. These requirements have been identified following an in-depth survey of the recruitment process in the Egyptian Universities. This survey employed a multi-part admission and registration DSS questionnaire (ARDSSQ) to identify the required data sources together with the likely users and their information needs. The questionnaire was sent to senior managers within the Egyptian Universities (both private and government) with responsibility for student recruitment, in particular admission and registration.

Further, access to a large database has allowed the evaluation of the practical suitability of using a DW structure and knowledge management tools within the decision making framework. 2000 records have been used to build and test the data mining techniques within the KDD process. The records were drawn from the Arab Academy for Science and Technology and Maritime Transport (AASTMT) students' database (DB).

Moreover, the paper has analyzed the key characteristics of DW and explored the advantages and disadvantages of such data structures. This evaluation has been used to build a DW for the Egyptian Universities that handle their admission and registration related archival data. The decision makers' potential benefits of the DW within the student recruitment process will be explored.

The design of the proposed admission and registration DSS (ARDSS) will be developed and tested using Cool: Gen (5.0) CASE tools by Computer Associates (CA), connected to a MS-SQL Server (6.5), in a Windows NT (4.0) environment. Crystal Reports (4.6) by Seagate will be used as a report generation tool. CLUSTAN Graphics (5.0) by CLUSTAN software will also be used as a clustering package.

The ARDSS software could be adjusted for usage in different countries for the same purpose, it is also scalable to handle new decision situations and can be integrated with other systems.

INTRODUCTION

The type of information required by managers is directly related to the level of management and the amount of structure in the decision situations they face. Decisions at the operational level tend to be more structured, those at the tactical level more semi-structured, and those at the strategic level more unstructured. Structured decisions involve situations where the procedures to follow to reach a decision can be specified in advance. The inventory reorder decisions faced by most businesses are a typical example [42].
Unstructured decisions involve decision situations where it is not possible to specify in advance most of the decision procedures to follow. At most, many decision situations are semi-structured. That is, some decision procedures can be pre-specified, but not enough to lead to a definite recommended decision. For example, decisions involved in starting a new line of products or making a major change to employee benefits would probably range from unstructured to semi-structured.

Therefore, information systems must be designed to produce a variety of information products to meet the changing decision needs of managers at different levels of an organization. For example, the strategic management level requires more summarized, ad hoc, unscheduled reports, forecasts, and external intelligence to support its more unstructured planning and policy-making responsibilities. The operational management level, on the other hand, may require more regular internal reports emphasizing detailed current and historical comparisons that support its more structured control of day-to-day operations. Thus, we can generalize that higher levels of management require more ad hoc, unscheduled, infrequent summaries, with a wide, external, forward-looking scope. On the other hand, lower levels of management require more pre-specified, frequently scheduled, and detailed information, with a more narrow, internal, and historical focus [42].
However, managers use different information systems for each class of decisions. For structured decisions Transaction Processing Systems (TPS) and Management Information systems (MIS) are usually used. For the unstructured decisions, Decision Support Systems (DSS) Expert Systems (ES) and Artificial Neural Networks (ANN) are used. Executive Information Systems (EIS) are special

types of information systems that support unstructured decisions [36] [37] [55].

DSS

DSS were first articulated in 70's. DSS are interactive computer-based systems that help decision-makers utilize data and models to solve unstructured problems. DSS are based on models and procedures for processing data and judgments to assist a manager in his decision making [56].

A study [51] evaluated the use of DSS in the banking industry in England. The results of this study showed that the use of DSS has evolved from being a simple system that only accessed the corporate data and reported it, to become knowledge based systems. No details on such knowledge based systems were given i.e. components, technologies used etc.

DSS were defined [32] as computer programs that provide information in a specific application domain by means of analytical decision models and access to databases in order to support decision makers effectively in semi-structured and unstructured decisions. The definition emphasized that end users of a specific DSS are not always known during the development but what is well known is the problem. Also, discussed the data sources of the DSS, ignored completely the use of historical or archival data sources. The study was focused on the types of problems, which the DSS can handle, the data models used, the users of the DSS and their objectives.

Another study [9] suggested a storage structure that will contain the results of any statistical queries which are accessed frequently by the DSS users. The study claimed that storing such queries' results would enhance the DSS performance. The study did not handle the problems associated with storing results of statistical queries while the data sources are being updated. That means that the stored results do not necessarily reflect the most up-to-date data values. This problem is handled by the use of summary tables in data warehousing. However the study did not mention any warehousing components.

A study by [27] examined the organisational strategies for supporting DSS (e.g. traditional application development groups, DSS groups within end-user services). The study mentioned that DSS are increasing exponentially in organisations, but little is known about how these organisations support their DSS users. The study emphasized that support practices affect the DSS groups' ability to provide support services. managers of twenty three DSS groups were interviewed (fifty seven organisations were contacted, twenty three accepted to participate). The results showed that the support characteristics tend to vary when comparing groups in different locations, suggesting that certain strengths and limitations could be associated with organisational support strategies. The factors that affect the DSS support services are; staff-to-user ratio, staff background, use of evolutionary development methodology, and functional area of the users.

A study [45] focused on the data structure on which the results obtained by what-if based DSS are stored. According to the study, this data structure independently updated views (IUVs). IUVs are used to store derived data from the DSS without losing consistency with the original data. SQL was used to create and maintain these IUVs. The study is found redundant, as the IUVs they proposed are basically summary tables that are used to store frequently accesses queries, and it would have been easier if they had employed a DW component to resolve the update problems and this would also allow the system to handle larger number of records than the proposed IUVs.

Another study [8] on the effectiveness of DSS revealed that the research efforts on the DSS outcomes, especially those that take the longitudinal effect, are very few. The study suggested that the effect of DSS on decision outcomes develops over time. Also claimed that the DSS effect is due to two factors; DSS development and the reliance effect (i.e. the decision maker defers the decision because the computer will take it later). The research concluded that both factors affect decision makers, that is, reliance has a short-term effect whilst DSS development has a long-term effect. Finally, DSS aided decision makers outperform non-DSS decision maker in organisations.

Another study [12] found out that there are many ways and various software for organisations to resolve decision problems. The study claimed that database management systems and the other decision technologies (i.e. any kind of computational procedures that have the ability to support decision making) are little used in real world applications because the market that distributes them is not professional enough. Also suggested an electronic market for decision technologies (Electronic market is a market where the enabling medium for transactions between consumers, providers, and services is an information network e.g. World Wide Web. The DecisionNet was introduced which performs functions (user accounts, billing, setting up interfaces) that would have been otherwise need to be developed for each consumer and/or technology.

A study [6] focused on the knowledge-based decision support systems (KBDSS). The study claimed that many KBDSS have been developed, however few systems address the use of knowledge in the decision problem. The study suggested a KBDSS for the telemarketing industry that predicts the probability of a customer disconnecting the telephone line based on previous promotions and customer history. The prediction model used a Bayesian network model linked with an influence diagram. The study was based on the customer's history (i.e. transactions), however their study did not include any DW components.

Another study [33] focused on groups of people taking decisions together, using what is known as group decision support systems (GDSS). GDSS allow a variety of specialists to be assembled whereby each of them is contributing to the solution using his expertise. GDSS could stimulate creative thinking and allow people from different departments to take the decision together. The disadvantage of GDSS is the possible conflict of the people contributing to the solution because each has a departmental view.

However, when all of them share the organisational goals this could encourage them to be committed to the organisation rather than individual departments. The study concentrated on the type of models and decisions taken rather than the components and technologies that should be used to implement GDSS.

A final study [2, p.2] said, "The DSS area is probably the most widely researched in the information systems field and is one that continues to be a focus for information systems researchers."

Data Warehousing

Not all the databases at an organization have the same classification. For example, the running applications have certain database design requirements and for this reason, these types of databases are known as operational databases. They are not designed to respond to spontaneous queries, however, they are optimized for carrying high speed and large numbers of users. Another type of database that can be found in organizations is the DW. This is designed for strategic decision support, and is largely built up from the operational databases. The DW can contain a large amount of data and millions of data records. Smaller, local data warehouses are called data marts [4] [43].
A DW is the means for strategic data usage [10]. In other words, DW is a blend of technologies aimed at effective integration of operational databases into an environment that enables the strategic use of data [3] [7] [17].

Data Warehousing was introduced to build a logically centralized data repository to fulfill the requirements of strategic data usage and prevent local systems from DW users competition [57].

The purpose of a DW is to establish a data repository that makes the operational data accessible in a form that is readily acceptable for the analysis. Only the data required to meet the executives' needs are taken from the operational environment.

DW can be viewed as an information system (IS) foundation that has the following characteristics [10]:

1. It is used intensively for READ type transactions;
2. It includes a large volume of records in a few number of tables;
3. Each query is processed in large data sets using multi-table joins;
4. It contains current and historical data;
5. It is periodically updated;
6. It supports a small number of users;
7. It is a database designed for analytical tasks;
8. It uses data from different databases, from various applications.

A study [50] focused on the creation of a DW using SQL Server. The study includes nothing on how to use front-end tools to utilize the DW. The study did not mention anything about the DW indexing strategy.

The factors that affect the DW success were studied [61]. The study findings showed that resources, user participation, highly skilled project team members increase the likelihood that warehousing projects will end on time and budget constraints and with proper functionality. Also it was found out that the implementation success with organizational and project issues will in turn affect the system quality of the DW.

Knowledge Discovery and data mining

Knowledge Discovery in Databases (KDD) means a process of nontrivial extraction of implicit, previously unknown and potentially useful information [20]. This discovered knowledge could be of great value in many areas, foremost among which is decision making [14]. The KDD was defined [41] as a three-step process that involves data selection, data cleaning, and finally knowledge interpretation.

KDD is not a separate body of knowledge that stands for its own. KDD is a multi-discipline branch of science; databases, statistics, visualization techniques, machine learning, and expert systems. They all contributed to the KDD process [25].

On the other hand, data mining gives organisations the tools to navigate through large amounts of data to find trends and patterns which can guide the strategic decision making process. Data mining is not specific to any particular industry- it requires intelligent technology and the willingness to explore the possibility of hidden knowledge that resides in the data. A study results [52] suggest that data warehousing is not a mandatory component in establishing the data mining environment, that is, data mining could be established without using a DW.

There is confusion about the exact meaning of the two terms 'data mining' and 'KDD', with many authors regarding them as synonymous [4] [21]. At the first international KDD conference in Montreal in 1995, it was proposed that the term 'KDD' would be employed to describe the whole process of extraction of knowledge from data. In this context, knowledge means understanding the relationships and patterns between data elements. It was further proposed that the term data mining should be used exclusively for the discovery stage of the KDD process. For the purpose of this paper, data mining techniques will be viewed as a step in the KDD process.

A study [48] on the interestingness in knowledge discovery revealed that the more actionable the discovered patterns are, the more interested the user is and further the user will apply those patterns and update his experience accordingly. The less interesting the patterns the less actionability and it is more likely that the user will discard the discovered pattern.

A different research study [49] discussed the degree of user involvement in the discovery process. The study outlines a spectrum of degrees of user-involvement and presents the data monitoring and discovery-triggering approaches that provide a balanced "division of labor" between the KDD application development and the discovery engine.

The security and privacy implications of data mining were examined by some studies [15]. It was found that data mining technology provides a whole new way of exploring large databases. Also, it was revealed that there is a trade off between restricting the access to data and the advantages of database and network technology in the ability to share data.

Some authors [10] differentiated between the operational data that is used in the traditional information systems, and informational data that is used in the information systems built around the mining techniques.

Another study [40] on knowledge discovery by inspection revealed that simple rules could be discovered by visual inspection of frequency tables.

The interestingness of the patterns resulting from the KDD process was examined in some studies [35]. Results revealed that it is easy to discover a huge number of patterns from any database, however, most of these patterns are useless and uninteresting to the users. Results suggested that to prevent the users from being overwhelmed by the large number of patterns, techniques are required to rank them according to their interestingness. Although data mining is a new strategic tool in the executive manager's hand, the strategic value of data mining is time sensitive, that is the output of data mining changes over time due to changes of the data sources itself [25].

In order to conduct effective data mining, one needs first to know what kind of features and requirements that should be available for the mining process. According to some studies [14], these requirements are:

1. Handling different types of data. Because there are many kinds of databases used in different applications, one might expect that a knowledge discovery system should be able perform data mining on different kinds of data. Since most databases available are Relational Data Base Management systems-RDBMS, it is expected that data mining techniques can find knowledge easily in these RDBMS. Moreover, data mining techniques should find knowledge in other complex data types; for example structured data, hypertext, multi-media data sets, transaction data, and legacy databases;
2. Efficiency and scalability of data mining algorithms. To extract information from a huge amount of data in databases, the knowledge discovery algorithm must be efficient and scaleable to large databases;
3. Meaningful data mining results. The output of the data mining process should be meaningful to the decision makers;
4. Testing the data mining results. The consistency of knowledge obtained from the data mining process can be examined by using different analysis techniques on the same data set;
5. Mining information from different sources of data. Mining knowledge from different formatted and unformatted data sets is a very rich technique that may enhance the results and increase the confidence level in the data mining results. Examples of sources of unformatted data sets are found in the data stored in Wide Area Networks-WAN, Internet, and many other external databases. However, this poses the problems of non-heterogeneous databases - refer to point number 1 in this list;
6. Security and privacy versus data mining. When data can be viewed from different angles, it may threaten other people's privacy and data security mechanisms. To resolve this point, it is important to prevent the disclosure of very sensitive information;
7. Data cleaning. Before in-depth analysis is carried out using data mining techniques, the data should be cleaned, i.e. be error-free, so that we can increase the confidence in the results.

It was emphasized that the combination of data warehousing, decision support, and data mining provides an innovative and totally new approach to information management [4].

FOUNDATION OF THIS RESEARCH

Foundations of this research in literature

This research will introduce a new DSS methodology; further the research will study the existing procedures. Fundamentally the DSS definitions introduced are all narrow as far as they do not present a comprehensive definition that encompasses all DSS characteristics. Further, they do not deal with the new blend of technologies DSS, Data mining techniques, and Data Warehousing [52]. Even those who tried to use the knowledge as DSS component said that it is optional and independent [56]. Managers and decision-makers that are the primary users of DSS have an interest in past data but unfortunately it is either not available or not in a suitable form for direct use [55]. To overcome these issues this research introduces and develops the use of KDD techniques, and DW to the DSS development.

Thus, the research study has three major components; DSS, KDD, and DW. The idea of linking these three components has its foundations in literature as follows:

1. [4, p.1] said, "The combination of data warehousing, decision support, and data mining indicated an innovative approach and totally new approach to information management";
2. [52, p.77] said, "Decision Support is data access targeted to provide the information needed by the decision makers. DBs and DW with reporting and analysis tools optimized to support business decision making are the components of the DSS";
3. [56, p.135] said, "Organisations are recognizing that their data contain a gold mine of information if they can dig it out. Consequently, they are warehousing and data mining for users to obtain information on their own and to establish relationships that were previously unknown";
4. [24, p.1] said, "Data warehouses, OLAP, and KDD are leading to new ways of performing decision support systems and creating executive information systems for data rich environments. Yet, these developments have

received almost no attention from academics either in their research on in teaching";

5. [16, p.566] said, "DW and other advances in IT are now solving the very difficult technical problems. They make it possible to organize, store, and retrieve huge volumes of information and to select critical information for a given decision. However, before organizations can realize that "grand promise" of MIS, most will have to reshape their business processes and decision making cultures to take advantage of the technology's new capabilities. This is a non-trivial transformation".

Based on the pre-stated literature foundations a decision was made to develop a new DSS methodology that will encompass these three major components (i.e. DW, KDD, and DSS). The new development methodology will be applied to the admission and registration functions in Egyptian Universities.

Foundations of the choice of the application domain

The choice of the admission and registration functions in Egyptian Universities to be this research's application domain is based on the following:

1. Reports from the Ministry of Higher Education's annual meeting (different volumes from 1980: 1999) which assured the importance of having a computer based information system that is capable of automating the admission and registration related processes and decisions;
2. The research efforts exerted in the area of admission and registration functions in Egyptian Universities are few and incomplete. Examples are:
    a. [63] Resulted in the development of a Web-based admission and registration information system. However, the research did not include the use of data warehousing and/or any knowledge discovery technique;
    b. Another study [19] resulted in an electronic payment Registration system. Again the research neither includes the admission function, nor the use of data warehousing and/or knowledge discovery techniques;
    c. A third study [38] developed a University Data Warehouse. The study did not include the use of any front-end tools (i.e. DSS or EIS);
3. The increasing number of users' (i.e. Registrars, Admission Officers, Deans, and Associate Deans) complaints on their current admission and registration Information Systems in different Universities;
4. The amount of unsatisfied users' needs (inability of current admission and registration information systems to take decisions, to predict, to profile students, the inaccessibility of the historical data);
5. The need for a computer based admission and registration information system, that meet the users' requirements, has been raised on many occasions:
    a. Annual meeting of the Arab Countries Universities' Registrars (including Egyptian Universities' Registrars);
    b. Annual evaluation of some computer based admission and registration information systems in some Universities e.g. AASTMT.

THE RESEARCH PROBLEM

This part of the research investigates the admission and registration functions taking place in the Egyptian Universities. The Egyptian Universities are classified by funding into two main categories; private and government funded.

All Egyptian Universities follow the same regulations and are controlled by the same authorities; Supreme Council of Universities and The Ministry of Higher Education. However, the admission and registration functions are different in these Universities in the sense of:

1. The Admission function. Private Universities act independently when making decisions about accepting or rejecting students. However, in government Universities these decisions are taken centrally (there is a central board that accepts and distributes students between the different academic institutions in each University);
2. The admission and registration department structure. In the private Universities, the admission and registration functions are centralized in one department for the entire University, whilst in government Universities there is a separate admission and registration department in each academic institute in the University.

Apart from the previous differences, the admission and registration functions in both University types are similar. They handle students' applications, map the students to the relevant academic institutions, doing course Registration and other grading-related jobs, graduation, class scheduling etc.

POPULATION, SAMPLE, AND DATA SOURCES

Population

The population of this questionnaire is the Egyptian Universities. These Universities are classified into two groups government and private-funded. According to the Egyptian Supreme Council of Universities statistics [54], the UNESCO World List of Universities [62], and the British Council Global Education and Training Information Service- Egyptian Universities [22] there are twenty-one Egyptian Universities; eight are private and thirteen are government. Each of the twenty-one universities consists of number of colleges, schools, faculties, and/or higher institutes. The total number of colleges, schools, faculties, and/or higher institutes in the entire population is 354.

This research investigates the admission and registration functions taking place in the Egyptian Universities both private and government. The admission and registration functions in both University types are similar in many areas. That is, they both handle students' applications, map the students to the relevant academic institutions, doing course registration, grading-related jobs, graduation, class scheduling etc.

However, there are differences in some areas between the two University types. For example, private Universities act independently when making decisions about accepting or rejecting students whilst in the government Universities these decisions are taken centrally. Moreover, in the private Universities the admission and registration functions are centralized in one department for the entire University, whilst in the government Universities there is a separate admission and registration department in each academic institute in the University.

The sample

As the number of Academic Institutions in Egyptian Universities is 354 distributed among 21 different Universities a decision was made to target them all and the number of Universities that will accept to participate would be used as a sample. All the twenty-one government and private universities were contacted to send their correspondence information and to notify them that they will receive the questionnaire. Only thirteen universities (six private, seven government) responded positively. The questionnaires were sent to these thirteen universities. A number of 670 questionnaires were sent to 13 universities.

The Response rate

On the University level, out of thirteen Universities contacted, responses from twelve Universities have been received. All six private Universities responded, whilst six out of seven government Universities responded. The overall response rate at the University level is 12/13= 92.3%.

On the respondent level, out of the 670 questionnaires sent, 167 returned, which gives a response rate 167/670= 24.9%. This response rate is adequate in this kind of research that is based on mailed questionnaires [13] [47] [53].

Data analysis

Since the research study relies on both primary and secondary data sources, different data analysis methods will be used.

Primary data analysis techniques

The data obtained by the questionnaire will be analyzed using Chi Square (X^2) analysis and Canonical correlation analysis.

Secondary data analysis techniques

The 2000 students' records will be analyzed using the KDD techniques; SQL, visualization, and clustering analysis.

RESEARCH OBJECTIVES

The focus of this research is on how to deliver information and knowledge to a specific category of business managers (i.e. admission and registration managers) to better understand their business problems and hence to improve their decisions. This will be achieved throughout the investigation of the following research objectives:

1. To investigate and critically evaluate the current DSS practices: Before introducing a new DSS methodology and a new DSS definition that is able to conform with the components of the proposed methodology, the current DSS literature and practices will be reviewed and analyzed;
2. Develop a new DSS methodology that consists of the following three components:
    a. DSS;
    b. DW;
    c. KDD. Since there are many techniques used in different contexts for achieving various goals, the techniques that will be used are the following:
        i. SQL;
        ii. Visualization;
        iii. Clustering analysis;
3. Identify the current admission and registration Information Systems in the Egyptian Universities concerning the following:
    a. The managers' perspectives towards computers and their current admission and registration information systems;
    b. Features of these information systems;
    c. Functions of these information systems;
4. Extract the information requirements for a new admission and registration DSS in the Egyptian Universities concerning the following:
    a. The managers' perspectives towards the role of computers and the ideal admission and registration information systems;
    b. The decisions that the DSS can take;
    c. DSS functions;
    d. DSS characteristics;
5. Use the proposed methodology to develop the required admission and registration DSS.

METHODOLOGY AND ITS TOOLS

While the new DSS proposed methodology has three major components (i.e. DW, KDD, and DSS), the methodology has four modules; one for each component and a module zero for the needs' analysis. Table (1) illustrates the methodology and its modules, the expected deliverables of each module and the tools and mechanisms used to accomplish each of which.

Table (1): The new DSS methodology and its modules

Module	Deliverable	Tools and Mechanisms
Module 0:	Needs' Analysis	Questionnaire for user requirements Cool: Gen CASE tools
Module 1:	Building the DW	MS-SQL Server: Star schema structure Crystal Reports: Report generation tool
Module 2:	The KDD process	MS-SQL Server Cool: Gen CASE tools The data mining techniques are:

		SQL Visualization Clustering analysis
Module 3:	Building the DSS	Cool: Gen CASE tools

STUDY RESULTS

- Discussion of the study's first objective No. 1 "Investigate and critically evaluate the current DSS practices".

Discussion of this research objective revealed that past studies on DSS have made significant contributions, however, new contributions are required to address their shortcomings. The past DSS research has focused on many issues, but no integrated approach has been found because each study has tried to narrow the different aspects of the DSS. For example, in the 70's the DSS definitions focused on data processing, and model. During the 80's the definitions focused on CBIS, effectiveness and the knowledge component appeared as part of the DSS. Whilst during the 90's the DSS definitions [51] [46] [42] [38], focused on the models and problem structureness with more attention given to the knowledge component. Recently [37] definitions required the DSS to be interactive and user-friendly. However, no definition has been found to include all the following aspects: the type of data used, the management level, the DSS effect, effectiveness of the DSS, type of knowledge targeted by the DSS. From this analysis it was evident that none of the definitions found was comprehensive, and this conclusion led to the need for a new DSS definition and methodology that are to be comprehensive.

Different DW definitions were illustrated and investigated. However, each of the definitions tried to define the DW taking into consideration certain point of view. Some definitions are technical [11], others are about the use of DW [10], the features [29], or the goals [31]. However, no definition has been found to be comprehensive covering the data sources, front-end, and the purpose of the DW in a business context. This has led to the development of a new DW definition.

The DW proposed definition: "A DW is a group of data extracted from different sources; internal, external, historical, and personal data archived in one or more data stores. The purpose of constructing a DW is to provide the DSS and the decision maker with the necessary data, which when transformed into information, will provide a better understanding of the business problem."

- The DW benefits were also clarified:

1. Increases the decision maker's productivity by providing accessible data in a ready to use format;
2. More cost-effective decision making process by separating the query processing from the operational databases;
3. Enhancing asset and liability management by providing the overall picture of the enterprise purchasing and inventory transactions;
4. Supporting the corporate strategy that positions the clients at the center of all operations which could not be achieved without a DW;
5. Reduces redundant processing, support, and software to enhance DSS applications;
6. Enhancing the work process, which also affects the success of business process reengineering;
7. Improve customer service;
8. Organisations will be able to exceed competitor capabilities and achieve competitive advantages.

The DW database can be built using the relational model or star schema structure. The star schema structure is the best for DW design.

The star schema structure captures the measurements of importance to the business and the parameters by which the business measurements are broken down. The measurements are referred to as facts, whilst the parameters by which a measurement can be viewed are called dimensions.

Data mining is a useful technique for extracting information and knowledge from the DW. Data mining is the process of finding hidden knowledge and unknown facts and trends in data. Data mining is a step of the KDD process.

There are four different types of knowledge; shallow, multi-dimensional, hidden, and deep knowledge [4]. Shallow knowledge can be discovered using SQL, multi-dimensional knowledge can be discovered using OLAP and visualization, hidden knowledge can be discovered using group of data mining techniques, and the deep knowledge can be discovered by employing the entire KDD process.

KDD is the process of fining hidden knowledge, patterns and unknown facts from the data sets. Data mining is a step in the KDD process.

- Various data mining techniques were discussed including:

1. SQL. Query tools are used to extract data that matches search criteria or to represent this data in a way that the user finds easier to handle or interpret. By applying simple SQL users can obtain a wealth of information;
2. Visualization. Visualization techniques depend strongly on the human side of the analysis [10]. Data visualization is emerging as a technology that may allow organizations to process amounts of data and present it in a usable format;
3. OLAP. The key driver for the development of OLAP is to enable the multi-dimensional analysis [44]. Although all the required information can be formulated using relational database and accessed via SQL, the two dimensional relational model of data and SQL have some serious limitations for investigating complex real world problems. Also slow response time and SQL functionality are a source of problems [10]. OLAP is a continuous and iterative process; an analyst can drill down to see much more details and then he can obtain answers to complex questions;
4. Association rules. According to [58] [34] [1] [11] [4] [5], the interest in discovering association rules from large relational tables has been increased recently.

Association rules are focused on finding relationships (i.e. associations) between a certain attribute (i.e. target attribute) that the user is interested in, and the remaining attributes in a relational table. The strength of association rules is that they can efficiently discover a complete set of associations that meet the user's requirements. However, there is no single algorithm that will automatically give the users everything of interest in the database;

5. Cluster analysis. Clustering is basically classifying unclassified data [23] [18]. The data to be classified consists of a set of items (sometimes referred to as objects, fields, or records). Each item is described by a set of characteristics called variables (sometimes referred to as attributes). The target of clustering is to classify the items in the data set into a number of groups (sometimes referred to as classes, or clusters), such that objects within one group have similarities with one another. Where the number of items is n, the maximum number of groups should be n-1 (i.e. the number of expected groups varies from 1 to n-1);
6. Decision trees. A decision tree is a predictive model that provides a means of visualizing complex decision problems where the questions can be posed in sequence;
7. GA. The term is a combination of both biology and computer disciplines, and sometimes referred to as simulated evolution. GA refers to these simulated evolutionary systems, but more precisely these are the algorithms that dictate how populations of organisms should be formed, evaluated and modified [11];
8. ANN. An ANN is a computer programs that implements complex pattern detection and machine learning algorithms to build predictive models from large database(s). In order for the ANN to detect patterns in the data sets, it should learn to detect these patterns and make predictions, in the same way a human does. ANN are widely used in many business applications;
9. Probabilistic graphical dependency technique. These models specify the probabilistic dependencies, which underlie a particular model using a graphical structure. The model specifies which variables are dependent on each other. These models are used for categorical or discrete-valued variables, however, some extensions also allow for the use of real-valued variables [20].

Each data mining technique has its own features and characteristics, hence not all the data mining techniques are applied to a certain application. Rather, each application applies only one or more data mining techniques that best fit with the data to be analyzed and the objective of the entire KDD process.

When applied, each data mining technique should have a goal and a task. The goals of the data mining techniques are prediction and description, whilst the data mining tasks are clustering, classification, summarization, dependency, regression, and change detection.

The spread of the KDD process especially the data mining techniques will be enhanced many folds if used with the DW. Moreover, the DW need a front-end tool (i.e. DSS), and this is the basic idea of this research, combining these three ingredients together.

In order to produce effective decision support systems that are able to enhance the quality of the decision making process research efforts have concentrated on introducing different combinations of three components (i.e. DSS, DW, and KDD), whilst other research efforts have ascertained the importance of linking them together. Moreover, this research has found that there have been very little efforts made to integrate these three components in certain application areas or to explain how these components work together and what tools and mechanisms used. The research efforts can be classified into two groups: the first group includes those who have tried to show the importance of linking two components together [29] [39] [56] [28]. The second group includes those who have assured the importance of linking the three in future research [4] [26] [16].

Due to the shortcomings of the traditional DSS definitions evaluated, and to respond to the new methodology which combines DSS, DW, and KDD a new DSS definition will be introduced: "A DSS is a computer-based information system that deals with semi-structured and unstructured problems facing managers at all management levels. The DSS goal is to enhance the decision quality and the manager effectiveness. To do so, the DSS integrates itself to the strategic data stores which is the DW, and to the KDD process that will find the deep knowledge and hidden patterns in the DW and present them to the DSS user."

The research objective No.2 "Develop a new DSS methodology" was evaluated a new DSS proposed DSS methodology was also introduced. The methodology consists of four modules. In Module 0 the needs' analysis is performed, the module uses a questionnaire for collecting the user requirements and Cool: Gen CASE tools planning and analysis phases to maintain and implement the requirements. In Module 1 the DW will be built, the module uses MS-SQL Server, star schema structure, and Crystal reports. In Module 2 the KDD process is to be applied, the module uses MS-SQL Server, Cool: Gen CASE tools analysis and design phases, and utilizing the following data mining techniques: SQL, visualization, and clustering analysis. In Module 3 the DSS will be built, the module uses Cool: Gen CASE tools design and generation phases.

Justification of the chosen tools and techniques proposed by the new DSS methodology was investigated deeply. For example, Cool: Gen 5.0 by Sterling is the CASE tool employed in this research to capture the information needs at the highest possible level of abstraction and transform them into executable application systems. The choice of Cool: Gen CASE tools is based on the following backgrounds: cost savings achieved by using the Cool: Gen CASE tool, developer productivity increases by up to 300%, dramatic improvements in business processes, higher levels of customer satisfaction, extraordinary flexible and high performance applications, accelerated systems development, and greater growth potential. Also, MS-SQL Server was

chosen to be used in the development of the DW because it supports both the relational and star schema structure models efficiently. Moreover, the data mining techniques were critically evaluated and based on this evaluation SQL, visualization, and clustering analysis were chosen.

- Different proximity measures were critically evaluated including Gower, Jaccard, City Block, Euclidean, Modified Euclidean, and Canberra. The evaluation obtained the following results:

1. The research results are consistent with the previous related work in two findings. Firstly previous related work indicated that Gower is the most suitable measure where the data types are mixture. Secondly previous work also did not recommend Jaccard where the data types are mixture. When applying Gower and Jaccard similar results were obtained; Gower was able to detect similarities and dissimilarities between the different record groups, whilst Jaccard obtained an extreme value (i.e. 0 or 1). This is because Jaccard is based on the number of similarities and dissimilarities and it completely discards the attribute values, it is possible that its value could be 1 or 0;
2. Moreover, City block generated a value of 0.5 for the first group that includes two very similar records. That makes it incomparable with the 0.99 value produced by Gower or the 0.92 produced by modified Euclidean for the same group. In addition to that, according to City block the similarity of the records in the second (i.e. dissimilar records) is .017 which is higher than the similarity of the records in the last group .01 (fairly dissimilar records);
3. Although Euclidean metric measure did not perform consistently well as Gower and modified Euclidean, it was decided to be retained and used with the clustering techniques because of the following reasons: Euclidean was recommended to be used with many clustering techniques (e.g. Nearest Neighbour, Furthest Neighbour, Ward's) by many scholars [60] [23] [18], also the results of the Euclidean measure are good because it was able to distinguish between similar and dissimilar records;
4. Another finding of this research is the results obtained by the modified Euclidean and Canberra metric measures (after transforming their results from distance metric to equivalent similarity coefficients). Previous work shows that little analysis has been carried out to apply the modified Euclidean and Canberra metric measures to a mixture of data types. However, the results of this research indicate that they are in close agreement with the results obtained by Gower's similarity measure.

- Various clustering techniques were studied including hierarchical methods, optimization, density, and clumping. The results obtained revealed the following:

1. The analyst must know that there are many clustering techniques available each of which has its own assumptions and gives different results if applied to the same data sets, the decision on choosing the clustering technique should be made in the light of the advantages and disadvantages of the chosen technique and the type of data to be classified;
2. Deciding on the number of clusters is a problem that is common to most of the clustering techniques. There are a lot of research efforts to handle this problem, however, there is no satisfactory solution for determining the optimum number of clusters [1] [23] [18]. There is no optimum number of clusters, it largely depends on the problem being resolved;
3. Although the concept of hierarchical techniques was developed in biology, this group of techniques is now used in many areas. One advantage is that the question on the optimum number of clusters does not arise since the researcher is interested in the complete hierarchy;
4. The biggest disadvantage associated with these techniques is their inability to reallocate items, which might be misclassified at early stages [18]. However, others [30] said that the Nearest Neighbour (NN) has the greatest mathematical appeal amongst, and would generate suitable results for most application areas;
5. The NN and FN techniques have the problem of chaining [59] [18]. Chaining means that the method tries to accumulate the new records/case on existing cluster(s) rather than creating new clusters. As a result of that, the number of records/cases in each cluster is highly affected by that problem, that is, few clusters retain the majority of records whilst the remaining clusters have quite little number of records;
6. The optimization techniques suffer from a number of problems. The techniques are transformation dependent; that is different results would be obtained from applying the same technique to the same data set. However, the advantages of using optimization techniques are the ability to reallocate misclassified item in further stages, and these techniques also do not assume that all clusters are hyper spherical (i.e. have the same shape). The most serious problem with the optimization techniques is the large amount of computation power they require, which in turn makes them irrelevant for the very large data sets;
7. The density solutions suffer from the problem of sub-optimal solutions; they might be more than one solution for the data sets (i.e. maximum likelihood). TAXMAP also suffer from the problem of containing various parameters that control the technique and arbitrary chosen by the investigator [18];
8. The Clumping techniques rather than their unsuitability for the data sets, they suffer from the problem of optimization techniques that is the computation power they need.

Based on the analysis and evaluation of the various proximity measures and clustering techniques, it was decided that Ward's would be used base on a Euclidean metric proximity measure. Ward's was less affected by chaining than NN and FN. Ward's also obtained a logical and consistent results because similar students came in a consequent order and dissimilar students did not come in a consequent order.

The research problem was defined as the admission and registration function in Universities.

The research population was defined as the Egyptian Universities both government and private.

The response base was identified as Deans, Associate Deans, Admission Officers, Registrars, and Others.

Searching the literature revealed that no research questionnaire was found relevant. Hence, a new research questionnaire (ARDSSQ) was developed and validated. The ARDSSQ consists of seven constructs, each of which investigates one research objective.

The population was contacted and asked to participate and then the ARDSSQ was sent to those who agreed to participate.

The response rate attained was 92.3% on the University level, and 24.9% on the respondent level.

Each of the questionnaire constructs will be analyzed in terms of these three dimensions: University type, Respondent position, and Whether the University uses a CBIS or not. The reason for using these three dimensions is due to the nature of the population, the response base, and the expected effects of using CBIS on the answers. And also to identify areas of commonality and discrepancy between the major segmentations identified within the population.

- Discussion of the research objective No. 3-1 "The managers' perspectives towards computers and their current admission and registration information systems":

1. The managers' perspectives towards computers and their current admission and registration information systems is affected by these two dimensions; the University type and the managers' position;
2. The use of CBIS does not affect the managers' perspectives towards computers and their current admission and registration information systems;
3. The percentage of private Universities that use CBIS (92%) is greater than the percentage of government Universities that use CBIS (35%);
4. The DW component is expected to enhance the decision quality.

- Discussion of the research objective No. 3-2 "Features of these information systems":

1. The features of the current admission and registration IS in the Egyptian Universities are not affected by any of these dimensions; the University type, the respondent position, and the use of CBIS;
2. The admission and registration information systems in the Egyptian Universities have the following features: Printing reports that describe students' records feature, Electronic stores of students' data;
3. The admission and registration information systems in the Egyptian Universities do not have the following features: Predicting the new applicants' performance, and Predicting the current-students' performance.

- Discussion of the research objective No. 3-3 "Functions of these information systems:

1. The functions of the current admission and registration information systems are affected by these two dimensions; the University type and the manager's position;
2. The functions of the current admission and registration information systems are not affected by the use of CBIS dimension;
3. Different management levels require different information needs;
4. The admission and registration information systems in the Egyptian Universities have the following functions: Student description reports, General statistics, Classifying students into similar groups, Using the historical data to describe the Students' history (only in the private Universities);
5. The admission and registration information systems in the Egyptian Universities have do not have the following functions: Student performance prediction, and Finding relationships between a student's data fields.

- Discussion of the research objective No. 4-1 "The managers' perspectives towards the role of computers and the ideal admission and registration information systems":

1. The managers' perspectives towards the role of computers and the ideal admission and registration information system is affected by both the University type and the manager's position dimensions;
2. The use of CBIS does not affect the managers' perspectives towards the role of computers and the ideal admission and registration information system;
3. Respondents from the private Universities have a better understanding to the managers' perspectives towards the role of computers and the ideal admission and registration information system;
4. There is a need for the DSS to be developed for the admission and registration functions. The results of the survey "the admission and registration information system should be able to help managers take decisions" showed that 98% of the private Universities, and 100% of the government Universities think that their admission and registration information system should be able to help managers take decisions.

- Discussion of the research objective No. 4-2 "The decisions that are expected to be taken by DSS":

1. The decisions that the admission and registration DSS is expected to take are affected by these dimensions; the University type, the manager's position, and the use of CBIS;
2. The admission and registration DSS should be able to take the following decisions:
    a. Accept or reject a new applicant;
    b. Predict the new applicants that will join the faculty/college/institute this term/year based on our archival records;
    c. Predict the new applicants that will join the faculty/college/institute this term/year based on government statistics on secondary school students;
    d. Predict the new applicants that will join the faculty/college/institute this term/year based on our archival records besides other records like the government statistics;

e. Based on our archival records we can make an applicant-major match and provide this to the new applicant to help him/her chooses a suitable major;
f. Hold the applicant until the following term/year;
g. Accept or reject the applicant who is transferred from another educational institution;
h. Accept or reject the applicant who is transferred from another educational institution based on our transfer history records;
i. Predict a student's performance based on the students' history we keep;
j. Predict a course's results based on the courses' history we keep;
k. Classifying students into similar groups;
l. Predict a student's performance based on the group that he/she belongs to;
m. Set the student status to "On probation";
n. Make relationships between students' performance and academic departments;
o. Forecast course booking;
p. Decide on Student abandonment.

- Discussion of the research objective No. 4-3 "4-3 DSS functions":

1. Both the University type and the manager's position dimensions affect the ideal admission and registration information system functions, whilst not affected by the use of CBIS dimension;
2. The ideal admission and registration information system should have the following functions:
   a. Predict new applicants' performance (only for the government Universities);
   b. Producing student description reports;
   c. Provide general statistics;
   d. Student classification into groups;
   e. Using historical data;
   f. Being able to use external data;
   g. Finding relationships between students' data fields;
   h. Gives the user the ability to create ad hoc reports.

- Discussion of the research objective No. 4-3 "4-4 DSS characteristics":

1. The ideal admission and registration information system should have the following characteristics:
   a. Easy to use;
   b. Requires minimum training;
   c. User involvement in design of the system;
   d. Able to grow;
   e. Flexible;
   f. Integrated;
   g. Have E-mail facility;
   h. Web-accessible;
   i. And cost effective.

- Statistical analysis using Chi-squared statistic revealed the following:

1. There is a significant relationship between the University type and the use of CBIS by which the number of private Universities who are using CBIS differs from that number in the government ones $\{X^2 = 50.79, S, p < .05\}$;
2. There is a significant relationship between the respondent position and the use of CBIS by which the number of Deans who are using CBIS differs from that number with regard to the Associate Deans, Registrars, Admission Officers, and Others $\{X^2 = 19.19, S, p < .05\}$;
3. There is a significant relationship between the respondent position and acceptance to role of computers as data stores by which the number of Deans who believe that being a data store is the main role of computers differs from that number as for the Associate Deans, Registrars, and Others $\{X^2 = 14.42, S, p < .05\}$;
4. There is no relationship between the respondent positions and the role of computers as being decision makers $\{X^2 = 2.45, NS, p > .05\}$;
5. There is a significant relationship between the respondent position and the ownership of a PC on his desk by which the number of Deans who have a PC's on desks differ from that number for the Associate Deans, Registrars, Admission Officers (Admission Officer and Associate Deans are the same), and Others $\{X^2 = 22.77, S, p < .05\}$.

- Statistical analysis using Canonical Correlations revealed the following:

$$-0.035 * Y\_1 + 0.012 * Y\_2 - 1.016 * X\_3 = 0.174 * construct\_1 + 0.752 * construct\_2 + 0.105 * construct\_3 + 0.018 * construct\_4 - 0.002 * construct\_5 - 0.035 * construct\_6 + 0.004 * construct\_7$$

Where the magnitude of the variable represents its contribution to the variate it belongs to. Variables of opposite signs represent inverse relationships to each other's. That is, among the independents' variate the CBIS use (X_3) accounts for the highest effect and works on the same direction as the University type (Y_1) and both are opposite to the Manager's position (Y_2) which has the least effect on the variate. Also the second, first and third constructs (in order) have the highest effect on the variate of the dependent variables. Among the dependent variables only the fifth construct moves in the opposite direction to the remaining (DSS decisions). The reason why the first three constructs have the highest relationship magnitudes is because they representing the current managers' perspective towards CBIS, the current admission and registration IS features, and the current admission and registration IS functions which are highly affected by the three independents, whilst the remaining constructs are about ideal managers' perspectives and ideal DSS decisions, functions, and characteristics where the three dimensions have little impact.

- The research objective No. 5 "Use the proposed methodology to develop the required admission and registration DSS" was investigated. That is, the proposed DSS methodology will be applied to the ARDSS development. The objective has been met by implementing the methodology's four modules.

Module 0: Needs' Analysis

This module was accomplished in four phases as follows:

1. The first phase is the development and validation of a new research questionnaire that is used to define the current admission and registration information systems in the Egyptian Universities and to explore the requirements that are not satisfied by these current systems;
2. The second phase the questionnaire was used to collect data from the admission and registration managers in the Egyptian Universities;
3. The third phase, the managers' information needs that are required to be satisfied have been identified;
4. The last phase, Cool: Gen CASE tools Planning and Analysis phases were utilized to start the development.

Module 1: Building the data warehouse

In this module the University DW was designed and implemented on MS SQL Server. In this module the research objective No. 2-2 "Designing the DW" was achieved by the undertaking the following five steps:

1. Study and evaluate the data sources;
2. Establish the source-to-target fields' matrix as a design validation tool;
3. Build and the DW Star Schema design using MS SQL Server;
4. The Updating strategy of the DW;
5. Design the managers' reports using Crystal Reports.

Module 2: Knowledge from the KDD process

In this module the KDD process was applied to 1800 records. SQL, Visualization, and Clustering analysis techniques have been used as data mining techniques. The techniques have been applied for the following reasons:

1. Finding the knowledge which will be stored in the ARDSS knowledge base;
2. Creating the managers' reports from the DW.

The knowledge base is found in detail in Appendix (1).

The discovered knowledge is considered deep knowledge because of the following:

1. More than one mining techniques was used:
    a. SQL which is able to reach the shallow knowledge;
    b. Visualization which can explore the multi-dimensional knowledge;
    c. Clustering which enables the hidden knowledge to be identified;
2. The entire KDD process was adopted (i.e. from domain understanding till the knowledge consolidation), instead of just applying the techniques;
3. The knowledge rules evaluated by the admission and registration managers were described as deep and new.
4. Same admission and registration managers who were involved in the evaluation of the University DW reports evaluated the discovered knowledge, they reported positively on the validity of the knowledge;
5. The use of the DW added a strategic dimension to the discovered knowledge.

Module 3: Building the ARDSS

In this module the following components have been identified:

1. *Data management component*. In this component the ARDSS DB has been created in Cool: Gen CASE tools and being transferred to MS-SQL Server;
2. *Knowledge management component*. The results of applying the knowledge discovery techniques were stored using the design tool provided by Cool: Gen CASE tools. As new data is added to the DW, the knowledge discovery techniques run again and new knowledge could be found, which could affect the knowledge base by adding new rules, changing or modifying or deleting existing rules;
3. *User Interface*. The ARDSS adopts a GUI environment;
4. *Users*. The users of the ARDSS have been identified Dean, Associate Deans, Registrars, Admission Officers, and Others.

- The ARDSS capable of taking the following decisions:

1. Accept or reject a new applicant;
2. Predict the new applicants that will join the faculty/college/institute this term/year based on our archival records;
3. Based on our archival records we can make an applicant-major match and provide this to the new applicant to help him/her chooses a suitable major;
4. Hold the applicant until the following term/year;
5. Accept or reject the applicant who is transferred from another educational institution;
6. Predict a student's performance based on the students' history we keep;
7. Classifying students into similar groups;
8. Predict a student's performance based on the group that he/she belongs to;
9. Set the student status to "On probation";
10. Make relationships between students' performance and academic departments;
11. Decide on Student abandonment.

Testing the ARDSS

The ARDSS has passed successfully four levels of testing including: consistency check level (provided by Cool; Gen CASE tools software), professionals' level, user level, and model level (using 200 records).

The ARDSS limitation

Following are the limitations of the ARDSS:

1. Restricted to the knowledge stored in its knowledge-base;
2. The ARDSS is able to take only eleven decisions, which is turn means that not all of the admission and registration related decisions are incorporated into the system;
3. The ARDSS is an environment-specific system; that is it requires a Client/Server environment which has MS SQL Server 6.5 RDBMS running on a Windows NT 4 OS, Crystal reports 4.6, CLUSTAN graphics 5.0, and a Windows 95 or 98 on a Pentium machine;

4. The ARDSS is designed for the Egyptian Universities to be used by Deans, Associate Deans, Registrars, Admission Officers, and Others;
5. The discovered knowledge is based on records drawn from the AASTMT students' DB. Although other Universities' managers including both government and private found the majority of the knowledge base relevant and acceptable, this knowledge can only be used for decision making at the AASTMT, and if any other University will use the ARDSS records from this University need to be included in the KDD process.

The management implications of the ARDSS

The following management implications have been identified:

1. Gaining competitive advantages;
2. Managers are more committed and informed;
3. Better-served customers (i.e. students);
4. The benefits of using CASE tools.

CONCLUSION

- The proposed DSS methodology suggested by this research was adopted and applied successfully to the admission and registration functions in the Egyptian Universities. The methodology combines DSS, DW, and KDD together;
- The DW component gives a strategic value to the DSS;
- The KDD techniques enable the different forms of knowledge to be found out and reached.

RECOMMENDATION

- The Ministry of Higher Education and the Egyptian Supreme Council of Universities are currently setting the standards and regulations, which both University types (i.e. government and private) have to follow. It is recommended that these two bodies set the same admission and registration standards and regulations for both University types. The researcher has found some difficulties in some areas and situations whenever a comparison is to be made between the two University types. Examples are the following:
  a. *The respondents*. In some government Universities the admission and registration decision makers are: Deans, Associate Deans, Admission Officers, and Registrars. However, in other government Universities there are no Admission Officers or Registrars, the two positions have been replaced with a position called Director. On the other hand, in the private Universities the decision makers are Deans, Associate Deans, Admission Officers, and Registrars;
  b. *Position responsibilities*. There is also a difference in the definition of the Registrar as a position in both University types. In government Universities the registrar is a college/school level position, whilst in private Universities it is a University level position. The difference in responsibilities would affect their information needs;
  c. *The academic year definition*. For some of the government Universities the academic year starts in September and ends in July, whilst other government and all private Universities have a semester-based academic year; September, February and Summer semesters. This also affects the design of the admission and registration Information System that they may use. For example the Admission procedure happens once a year in some of the government Universities and more in others;
  d. *The grading system.* Some government Universities follow a descriptive grading system (Excellent, Very Good, Good, Satisfactory, and Poor), whilst others follow a GPA grading system (scale of 4). All of the private Universities have the same grading scale (GPA);
- It is advisable that both University types use the ARDSS proposed because of the following:
  a. The current admission and registration systems are incapable of providing the managers with their information needs. For example 69% of respondents in the private Universities and 73% in the government Universities reported that their systems couldn't predict the performance of their current students. However, 69% of respondents in the private Universities and 78% in the government Universities reported that they need their systems to predict the performance of current students. The ARDSS has the performance prediction capability;
  b. The ARDSS fulfills an important design principle "User-involvement during the system development process", on which 86% of respondents in the private Universities and 83% of respondents in the government Universities reported positively;
  c. The ARDSS is enhanced with a DW component that gives the managers access to some reports, which are based on 10-year time span (or more if data is available), and are combining fields/attributes from many tables/entities without affecting the performance of the operational DB. This DW component adds a strategic value to the use of the ARDSS;
  d. The ARDSS is built utilizing KDD techniques (i.e. SQL, visualization, and clustering analysis) which give the system the capability to reach different types of knowledge (shallow, multi-dimensional, hidden, and deep knowledge);
  e. The use of CASE tools in the development of the ARDSS added the following advantages to the system: flexibility, scalability, integrity, complete documentation, business needs and objectives are the system's main drivers, and finally the ARDSS is based on an easy to understand and deal with GUI;
- The Egyptian Universities must pay attention to the wealth of information stored in their admission and registration TPS. From these systems lots of valuable knowledge can be extracted. For example the admission and registration managers are interested in reports that are based on ten years time (and sometimes

more). However, they cannot get these reports from their current systems. Moreover, most of the reports they need are not built on their current systems, as a result of that there has been always a time lag between requesting a report and getting it. Some of these admission and registration managers were interviewed and they described some of the reports that need to be created, the reports have been created based on the data stored in the University data warehouse.

FUTURE WORK

- *Web accessibility*. Implementing the admission and registration DSS on the Web, so that senior managers can log to the Web site and access the system remotely;
- *Systems integration*. Integrating the admission and registration DSS with other systems e.g. the integration with the financial system;
- *Knowledge base comparison*. The knowledge base of the ARDSS is based on records drawn from a private University students' DB. Further research could be undertaken to build the knowledge base based on records drawn from a government University students' DB and then compare the results;
- *Inter Universities Web-Based Data Warehouse*. Develop an inter Universities DW, from which all member Universities can extract useful information and helping them enhancing their business understanding and accordingly the decision quality;
- *Different data mining techniques*. The admission and registration DSS depends on SQL, visualization, and clustering analysis techniques, however, different techniques could be studied and applied to the system where more data is available and the techniques to be implemented are relevant;
- *Comparative research studies*. Carry out the study using the ARDSSQ in different countries and compare the results. It is important to realize that some modifications need to be made to the ARDSSQ to reflect any country's specific educational system environment.

APPENDICES

Appendix (1) The Discovered Knowledge Base

The discovered knowledge is will be described in the form of rules in the following sections:

Rule No. 1

How obtained: This rule was obtained by two sources; interviewing some admission and registration managers at the AASTMT and documentation review and partially validated by SQL statement running on the sample data set. The SQL validation was partially because the sample data set does not have all the attributes required to validate the rule e.g. interview.

Decision No. A
"Accept or reject a new applicant**"**

IF major = "Hotels and Tourism" AND interview= "Satisfactory" AND high school percentage >= 60 AND age < 25 AND number of applicants <= batch ceiling
THEN accept applicant = "Y"
ELSE IF major = "Maritime" OR "Marine Eng." AND gender= "Male" AND high school percentage >= 60 AND age < 25 AND number of applicants <= batch ceiling
THEN accept applicant = "Y"
ELSE IF major <> "Maritime" OR "Marine Eng." OR "Hotels and Tourism" AND high school percentage >= 60 AND age < 25 AND number of applicants <= batch ceiling
THEN accept applicant = "Y"
ELSE
accept applicant = "N"
END IF

Rule No. 2

How obtained: This rule was obtained by two sources; interviewing some admission and registration managers at the AASTMT and documentation review.
Decision No. C
"Predict the new applicants that will join the faculty/college/institute this term/year based on our archival records"
IF semester= "September" THEN
applicants predicted = 2000
ELSE IF semester = "February" THEN
applicants predicted = 500
ELSE
applicants predicted = 0
END IF

Rule No. 3

How obtained: This rule was obtained by applying the Ward's clustering technique based on a Euclidean metric measure running on the sample data set.

Decision No. F
"Based on our archival records we can make an applicant-major match and provide this to the new applicant to help him/her chooses a suitable major"
Cluster 1: (38 members)
IF high school certificate="Thanwya Azhar" AND high school origin="Egypt" AND high school percent="66" AND nationality="Libya'' AND gender="Male"
THEN major="BBA Arabic section" AND age on graduation="23" AND GPA="Very Good"
Cluster 2: (81 members)
IF high school certificate="Thanwya Amma- Math" AND high school origin="Egypt" AND high school percent="65" AND nationality="Egypt'' AND gender="Male"
THEN major=" B.Sc. Electronics" AND age on graduation="25" AND GPA="Good"
Cluster 3: (136 members)
IF high school certificate="Thanwya Amma- Science" AND high school origin="Egypt" AND high school percent="65" AND nationality=" Oman'' AND gender="Male"
THEN major=" Bachelor of Hotels and Tourism" AND age on graduation="27" AND GPA="Very Good"
Cluster 4: (50 members)

IF high school certificate=”Thanwya Amma- Science” AND high school origin=”Egypt” AND high school percent=”56” AND nationality=”Egypt’’ AND gender=”Female”
THEN major=” BBA Arabic section” AND age on graduation=”24” AND GPA=”Good”
Cluster 5: (146 members)
IF high school certificate=”Thanwya Azhar” AND high school origin=”Egypt” AND high school percent=”52” AND nationality=”Egypt’’ AND gender=”Male”
THEN major=” Bachelor of Maritime Transport” AND age on graduation=”25” AND GPA=”Pass”
Cluster 6: (60 members)
IF high school certificate=”Thanwya Amma- Science” AND high school origin=”Egypt” AND high school percent=”52” AND nationality=”Egypt’’ AND gender=”Male”
THEN major=” BTech. Electronics” AND age on graduation=”25” AND GPA=”Poor”
Cluster 7: (42 members)
IF high school certificate=”Thanwya Amma- Science” AND high school origin=”Egypt” AND high school percent=”55” AND nationality=”Sudan’’ AND gender=”Male”
THEN major=” Bachelor of Maritime Transport” AND age on graduation=”25” AND GPA=”Good”
Cluster 8: (46 members)
IF high school certificate=”Thanwya Amma- Math” AND high school origin=”Egypt” AND high school percent=”50” AND nationality=”Syria’’ AND gender=”Male”
THEN major=” B.Sc. Electronics” AND age on graduation=”26” AND GPA=”Good”
Cluster 9: (32 members)
IF high school certificate=”Thanwya Amma- Science” AND high school origin=”Egypt” AND high school percent=”59” AND nationality=”Jordan’’ AND gender=”Male”
THEN major=” B.Sc. Marine Eng.” AND age on graduation=”26” AND GPA=”Good”
Cluster 10: (66 members)
IF high school certificate=”Thanwya Amma- Science” AND high school origin=”Egypt” AND high school percent=”58” AND nationality=”Egypt’’ AND gender=”Male”
THEN major=” B.Sc. Electronics” AND age on graduation=”25” AND GPA=”Very Good”
Cluster 11: (206 members)
IF high school certificate=”Thanwya Amma- Science” AND high school origin=”Egypt” AND high school percent=”51” AND nationality=”Egypt’’ AND gender=”Male”
THEN major=” B.Sc. Computers” AND age on graduation=”25” AND GPA=”Good”
Cluster 12: (89 members)
IF high school certificate=”Thanwya Amma- Math” AND high school origin=”Egypt” AND high school percent=”78” AND nationality=”Lebanon’’ AND gender=”Male”
THEN major=” Bachelor of Maritime” AND age on graduation=”24” AND GPA=”Very Good”
Cluster 13: (78 members)
IF high school certificate=”Thanwya Amma- Math” AND high school origin=”Egypt” AND high school percent=”73” AND nationality=”Egypt’’ AND gender=”Male”
THEN major=” B.Sc. Electronics” AND age on graduation=”24” AND GPA=”Very Good”
Cluster 14: (78 members)
IF high school certificate=”Thanwya Amma- Math” AND high school origin=”Egypt” AND high school percent=”66” AND nationality=”Jordan’’ AND gender=”Male”
THEN major=” Bachelor of Maritime Transport” AND age on graduation=”25” AND GPA=”Good”
Cluster 15: (37 members)
IF high school certificate=”Thanwya Amma- Science” AND high school origin=”Yemen” AND high school percent=”68” AND nationality=”Sudan’’ AND gender=”Male”
THEN major=” B.Sc. Electronics” AND age on graduation=”25” AND GPA=”Good”
Cluster 16: (128 members)
IF high school certificate=”Thanwya Amma- Math” AND high school origin=”Libya” AND high school percent=”74” AND nationality=”Jordan’’ AND gender=”Male”
THEN major=” B.Sc. Computers” AND age on graduation=”25” AND GPA=”Very Good”
Cluster 17: (92 members)
IF high school certificate=”Thanwya Amma- Science” AND high school origin=”Egypt” AND high school percent=”80” AND nationality=”Jordan’’ AND gender=”Male”
THEN major=” B.Sc. Electronics” AND age on graduation=”24” AND GPA=”Very Good”
Cluster 18: (58 members)
IF high school certificate=”Thanwya Amma- Math” AND high school origin=”Egypt” AND high school percent=”87” AND nationality=”Libya’’ AND gender=”Male”
THEN major=” B.Sc. Computers” AND age on graduation=”24” AND GPA=”Very Good”
Cluster 19: (107 members)
IF high school certificate=”Thanwya Amma- Math” AND high school origin=”Egypt” AND high school percent=”85” AND nationality=”Saudi’’ AND gender=”Male”
THEN major=” Bachelor of Maritime” AND age on graduation=”25” AND GPA=”Excellent”
Cluster 20: (14 members)
IF high school certificate=”Thanwya Amma- Math” AND high school origin=”Egypt” AND high school percent=”86” AND nationality=”Jordan’’ AND gender=”Male”
THEN major=” B.Sc. Computers” AND age on graduation=”24” AND GPA=”Very Good”
Cluster 21: (133 members)
IF high school certificate=”IGCSE- Old” AND high school origin=”Eritrea” AND high school percent=”76” AND nationality=”Egypt”’ AND gender=”Male”
THEN major=” Bachelor of Hotels and Tourism” AND age on graduation=”25” AND GPA=”Good”
Cluster 22: (9 members)
IF high school certificate=”Thanwya Amma Old- Science” AND high school origin=” Lebanon” AND high school percent=”80” AND nationality=’’ Lebanon” AND gender=”Male”
THEN major=” B.Sc. Computers” AND age on graduation=”25” AND GPA=”Very Good”
Cluster 23: (27 members)
IF high school certificate=”Thanwya Amma Old- Science” AND high school origin=”Libya” AND high school

percent=”63” AND nationality=”Lebanon” AND gender=”Male”
THEN major=” B.Sc. Computers” AND age on graduation=”23” AND GPA=”Excellent”
Cluster 24: (37 members)
IF high school certificate=”Thanwya Amma Old- Math” AND high school origin=”Egypt” AND high school percent=”69” AND nationality=”Libya’ ’ AND gender=”Male”
THEN major=” B.Sc. Computers” AND age on graduation=”25” AND GPA=”Good”

Rule No. 4

How obtained: This rule was obtained by two sources; interviewing some admission and registration managers at the AASTMT and documentation review and partially validated by SQL statement running on the sample data set.

Decision No. G
“Hold the applicant until the following term/year”
IF high school percentage >= 60 AND age < 25 AND number of applicants > batch ceiling
THEN deferred applicant = “Y”
ELSE
deferred applicant = “N”
END IF

Rule No. 5

How obtained: This rule was obtained by two sources; interviewing some admission and registration managers at the AASTMT and documentation review and partially validated by SQL statement running on the sample data set.

Decision No. H
“Accept or reject the applicant who is transferred from another educational institution”
IF major = “Hotels and Tourism” AND interview= “Satisfactory” AND high school percentage >= 60 AND age < 25 AND number of applicants <= batch ceiling AND previous abandonment = “N”
THEN accept applicant = “Y”
ELSE IF major = “Maritime” OR “Marine Eng.” AND gender= “Male” AND high school percentage >= 60 AND age < 25 AND number of applicants <= batch ceiling AND previous abandonment = “N”
THEN accept applicant = “Y”
ELSE IF major <> “Maritime” OR “Marine Eng.” OR “Hotels and Tourism” AND high school percentage >= 60 AND age < 25 AND number of applicants <= batch ceiling AND previous abandonment = “N”
THEN accept applicant = “Y”
ELSE
accept applicant = “N”
END IF

Rule No. 6

How obtained: This rule was obtained and validated by SQL statement running on the sample data set.

Decision No. J
“Predict a student’ s performance based on the students’ history we keep”
IF high school percentage <= 65 THEN
predicted student performance = “Between 0-2”
ELSE IF high school percentage > 65 THEN
predicted student performance = “Between 3-6”
ELSE
predicted student performance = “Undefined”
END IF

Rule No. 7

How obtained: This rule was obtained by applying the Ward’ s clustering technique based on a Euclidean metric measure running on the sample data set.

Decision No. L
“Classifying students into similar groups”
Cluster 1: (44 members)
high school type= “Thanwya Amma- Science” AND high school year= “90” AND year in= “91” AND high school percent= “73” AND major= “Bachelor of Maritime Transport” AND graduation date=”96” AND GPA= “Very Good” AND DOB= “72”AND nationality= “Libya” AND gender= “Male”
Cluster 2: (141 members)
high school type= “Thanwya Amma- Math” AND high school year= “91” AND year in= “91” AND high school percent= “75” AND major= “B.Sc. Electronics” AND graduation date=”97” AND GPA= “Very Good” AND DOB= “73”AND nationality= “Egypt” AND gender= “Male”
Cluster 3: (128 members)
high school type= “Thanwya Amma- Math” AND high school year= “88” AND year in= “88” AND high school percent= “80” AND major= “B.Sc. Computers” AND graduation date=”94” AND GPA= “Very Good” AND DOB= “70”AND nationality= “Jordan” AND gender= “Male”
Cluster 4: (65 members)
high school type= “Thanwya Amma- Science” AND high school year= “88” AND year in= “88” AND high school percent= “73” AND major= “B.Sc. Marine Eng.” AND graduation date=”95” AND GPA= “Very Good” AND DOB= “70”AND nationality= “Jordan” AND gender= “Male”
Cluster 5: (116 members)
high school type= “Thanwya Amma- Math” AND high school year= “90” AND year in= “90” AND high school percent= “88” AND major= “B.Sc. Computers” AND graduation date=”96” AND GPA= “Very Good” AND DOB= “72”AND nationality= “Lebanon” AND gender= “Male”
Cluster 6: (84 members)
high school type= “Thanwya Amma- Science” AND high school year= “88” AND year in= “89” AND high school percent= “88” AND major= “B.Sc. Electronics” AND graduation date=”94” AND GPA= “Excellent” AND DOB= “70”AND nationality= “Jordan” AND gender= “Male”
Cluster 7: (79 members)
high school type= “Thanwya Amma- Science” AND high school year= “91” AND year in= “91” AND high school

percent= “64” AND major= “BBA Arabic section” AND graduation date=”96” AND GPA= “Very Good” AND DOB= “75”AND nationality= “Egypt” AND gender= “Female”
Cluster 8: (148 members)
high school type= “Thanwya Amma- Math” AND high school year= “91” AND year in= “91” AND high school percent= “63” AND major= “B.Sc. Computers” AND graduation date=”97” AND GPA= “Good” AND DOB= “73”AND nationality= “Egypt” AND gender= “Male”
Cluster 9: (103 members)
high school type= “Thanwya Amma- Science” AND high school year= “92” AND year in= “92” AND high school percent= “57” AND major= “B.Sc. Computers” AND graduation date=”97” AND GPA= “Good” AND DOB= “74”AND nationality= “Egypt” AND gender= “Male”
Cluster 10: (147 members)
high school type= “Thanwya Amma- Science” AND high school year= “91” AND year in= “91” AND high school percent= “56” AND major= “BBA Arabic section” AND graduation date=”96” AND GPA= “Very Good” AND DOB= “73”AND nationality= “Egypt” AND gender= “Male”
Cluster 11: (115 members)
high school type= “Thanwya Azhar” AND high school year= “91” AND year in= “91” AND high school percent= “51” AND major= “BTech. Electronics” AND graduation date=”97” AND GPA= “Good” AND DOB= “73”AND nationality= “Egypt” AND gender= “Male”
Cluster 12: (92 members)
high school type= “Thanwya Amma- Science” AND high school year= “90” AND year in= “90” AND high school percent= “55” AND major= “Bachelor of Hotels and Tourism” AND graduation date=”97” AND GPA= “Good” AND DOB= “72”AND nationality= “Syria” AND gender= “Male”
Cluster 13: (45 members)
high school type= “Thanwya Amma- Science” AND high school year= “87” AND year in= “88” AND high school percent= “55” AND major= “Bachelor of Maritime Transport” AND graduation date=”94” AND GPA= “Poor” AND DOB= “69”AND nationality= “Libya” AND gender= “Male”
Cluster 14: (60 members)
high school type= “Thanwya Amma- Math” AND high school year= “88” AND year in= “88” AND high school percent= “53” AND major= “B.Sc. Electronics” AND graduation date=”95” AND GPA= “Good” AND DOB= “70”AND nationality= “Egypt” AND gender= “Male”
Cluster 15: (43 members)
high school type= “Thanwya Amma- Math” AND high school year= “89” AND year in= “90” AND high school percent= “74” AND major= “Bachelor of Hotels and Tourism” AND graduation date=”96” AND GPA= “Good” AND DOB= “70”AND nationality= “Oman” AND gender= “Male”
Cluster 16: (21 members)
high school type= “Thanwya Amma- Math” AND high school year= “88” AND year in= “89” AND high school percent= “57” AND major= “Bachelor of Maritime” AND graduation date=”95” AND GPA= “Good” AND DOB= “66”AND nationality= “Eritrea” AND gender= “Male”
Cluster 17: (75 members)
high school type= “Thanwya Amma- Science” AND high school year= “90” AND year in= “90” AND high school percent= “68” AND major= “BTech. Marine Eng.” AND graduation date=”96” AND GPA= “Good” AND DOB= “72”AND nationality= “Sudan” AND gender= “Male”
Cluster 18: (42 members)
high school type= “Thanwya Amma- Science” AND high school year= “87” AND year in= “87” AND high school percent= “58” AND major= “B.Sc. Marine Eng.” AND graduation date=”93” AND GPA= “Good” AND DOB= “68”AND nationality= “Jordan” AND gender= “Male”
Cluster 19: (74 members)
high school type= “Thanwya Amma- Science” AND high school year= “87” AND year in= “87” AND high school percent= “66” AND major= “B.Sc. Marine Eng.” AND graduation date=”94” AND GPA= “Good” AND DOB= “68”AND nationality= “Jordan” AND gender= “Male”
Cluster 20: (95 members)
high school type= “Thanwya Amma- Science” AND high school year= “87” AND year in= “87” AND high school percent= “65” AND major= “B.Sc. Electronics” AND graduation date=”94” AND GPA= “Very Good” AND DOB= “68”AND nationality= “Egypt” AND gender= “Male”
Cluster 21: (9 members)
high school type= “Thanwya Amma New- Science” AND high school year= “90” AND year in= “90” AND high school percent= “76” AND major= “Bachelor of Hotels and Tourism” AND graduation date=”97” AND GPA= “Good” AND DOB= “72”AND nationality= “Yemen” AND gender= “Male”
Cluster 22: (30 members)
high school type= “Thanwya Amma Old- Science” AND high school year= “91” AND year in= “91” AND high school percent= “79” AND major= “B.Sc. Electronics” AND graduation date=”97” AND GPA= “Very Good” AND DOB= “74”AND nationality= “Palestine” AND gender= “Female”
Cluster 23: (34 members)
high school type= “Thanwya Amma Old- Science” AND high school year= “91” AND year in= “91” AND high school percent= “62” AND major= “B.Sc. Computers” AND graduation date=”97” AND GPA= “Good” AND DOB= “73”AND nationality= “Libya” AND gender= “Male”
Cluster 24: (10 members)
high school type= “Thanwya Amma Old- Math” AND high school year= “91” AND year in= “91” AND high school percent= “69” AND major= “B.Sc. Computers” AND graduation date=”97” AND GPA= “Good” AND DOB= “73”AND nationality= “Libya” AND gender= “Male”

Rule No. 8

How obtained: This rule was obtained by applying the Ward’s clustering technique based on a Euclidean metric measure running on the sample data set.

Decision No. M
“Predict a student’s performance based on the group that he/she belongs to”
Depends on Decision No. *L*.
For example: cluster 1 will be represented as:

IF high school type= "Thanwya Amma- Science" AND high school year= "90" AND year in= "91" AND high school percent= "73" AND major= "Bachelor of Maritime Transport" AND graduation date="96" AND DOB= "72"AND nationality= "Libya" AND gender= "Male" THEN GPA= "Very Good"

Rule No. 9

How obtained: This rule was obtained by two sources; interviewing some admission and registration managers at the AASTMT and documentation review and validated by SQL statement running on the sample data set.

Decision No. N
"Set the student status to ' On probation' "
IF student GPA < 2.0 THEN
student status = "On probation"
ELSE
student status = "Normal"
END IF

Rule No. 10

How obtained: This rule was obtained and validated by SQL statement running on the sample data set.

Decision No. P
"Make relationships between students' performance and academic departments"
m = major
students performance = AVERAGE graduation grade for m WHERE 12 > m > 0
IF 1> students performance >= 0 THEN
major performance = "Poor"
ELSE IF 2> students performance >= 1 THEN
major performance = "Satisfactory"
ELSE IF 3> students performance >= 2 THEN
major performance = "Good"
ELSE IF 4> students performance >= 3 THEN
major performance = "V. Good"
ELSE IF 5> students performance >= 4 THEN
major performance = "Excellent"
ELSE IF 6> students performance >= 5 THEN
major performance = "V. Good, Honor"
ELSE IF students performance = 6 THEN
major performance = "Excellent, Honor"
ELSE
major performance = "Undefined"
END IF
END WHILE

Rule No. 11

How obtained: This rule was obtained by two sources; interviewing some admission and registration managers at the AASTMT and documentation review and partially validated by SQL statement running on the sample data set.

Decision No. R
"Decide on Student abandonment"
IF penalty = "Y" AND student GPA < 2.0 THEN
student abandoned = "Y"
ELSE IF penalty = "Y" AND student GPA >= 2.0 THEN
student abandoned = "W"
ELSE
student abandoned = "N"
END IF

REFERENCES

[1] Aas, K., Husbey, R., and Thune, M. (1999), *Data Mining: A Survey*, unpublished report, Norwegian Computing Center, Norway.

[2] Adam, F., Fahy, M., and Murphy, C. (1997), ' A Framework For The Classification Of DSS Usage Across Organisations' , *Decision Support Systems*, vol. 22, pp. 1-13.

[3] Adamson, C., and Venerable, M. (1998), *Data Warehouse Design Solutions*, John Wiley and Sons, Inc., New York.

[4] Adriaans, P., and Zantinge, D. (1996), *Data Mining*, Addison Wesley Longman Limited, Harlow.

[5] Agrawal, R., Mannila, H., Srikant, R., Toivonen, H., and Verkamo, A. (1996), ' Fast Discovery Of Association Rules' , In Fayyad, U., Piatetsky, G., and Smyth, P. (Eds), *Advances In Knowledge Discovery And Data Mining*, AAAI Press/ The MIT Press, California, pp. 307-328.

[6] Ahn, J., and Ezawa, K., J. (1997), ' Decision Support For Real-Time Telemarketing Operation Through Bayesian Network Learning' , *Decision Support Systems*, vol. 21, pp. 17-27.

[7] Barquin, R. (1997), ' A Data Warehousing Manifesto' , In Barquin, R., and Edlestein, H. (Eds), *Planning And Designing The Data Warehouse*, Prentice Hall, New Jersey, pp. 3-16.

[8] Barr, S. and Sharda, R. (1997), ' Effectiveness Of Decision Support Systems: Development Or Reliance Effect' , *Decision Support Systems*, vol. 21, pp. 133-146.

[9] Barron, T., and Saharia, A. (1995), ' Data Requirements In Statistical Decision Support Systems: Formulation And Some Results In Choosing Summaries' , *Decision Support Systems*, vol. 15, pp. 375-388.

[10] Berson, A. (1996), *Client/Server Architecture*, McGraw-Hill, New York.

[11] Berson, A., and Smith, S. (1997), *Data Warehousing, Data Mining, and OLAP*, McGraw-Hill, New York.

[12] Bhargava, H., Krishnan, R., and Muller, R. (1997), ' Decision Support on Demand: Emerging Electronic Markets For Decision Technologies' , *Decision Support Systems*, vol. 19, pp. 193-214.

[13] Chan, Y., Huff, S., and Copeland, D. G. (1998), ' Assessing Realized Information Systems Strategy' , *Journal of Strategic Information Systems*, vol. 6, pp. 273-298.

[14] Chen, M., Han, J., and Yu, P. S. (1996), ' Data Mining: An Overview From A Database Perspective' , *IEEE Knowledge And Data Engineering*, vol. 8, no 6, pp. 866-883.

[15] Clifton, C., and Marks, D. (1996), ' Security And Privacy Implications Of Data Mining' , In:

Proceedings Of ACM SIGMOD '96 International Conference On Management Of Data, June 4-6, Montreal, Quebec.

[16] Cooper, B., Watson, H., Wixom, B., and Goodhue, D. L. (2000), 'Data Warehousing Supports Corporate Strategy At First American Corporation', *MIS Quarterly*, vol. 24, no 4, pp. 547-567.

[17] Devlin, B. (1997), *Data Warehouse: From Architecture To Implementation*, Addison-Wesley Longman, Inc., Reading, Massachusetts.

[18] Everitt, R. (1980), *Cluster Analysis*, 2nd ed., Halsted press- A division of John Wiley and Sons, Inc., London.

[19] Fady, R. (2000), *Electronic Payment Student Registration Systems*, unpublished MBA Thesis, The Advanced Management Institute (AMI), Alexandria.

[20] Fayyad, U., Piatetsky, G., and Smyth, P. (1996), 'From Data Mining To Knowledge Discovery: An Overview', In Fayyad, U., Piatetsky, G., and Smyth, P. (Eds), *Advances In Knowledge Discovery And Data Mining*, AAAI Press/ The MIT Press, California, pp. 1-34.

[21] Ganti, V., Gehrke, J., and Ramakrishnan, R. (1999), 'Mining Very Large Databases', *IEEE Computer*, vol. 32, no 8, pp. 38-45.

[22] *Global Education and Training Information Service-Egyptian Universities*. (1999), The British Council, London.

[23] Gordon, A. D. (1981), *Monographs On Applied Probability And Statistics Classification*, The University Press, Cambridge.

[24] Gray, P., and Watson, H., J. (1999), *The New DSS: Data warehouse, OLAP, MDD, and KDD*, Unpublished report, Georgia State University, Georgia.

[25] Guyon, I., Matic, N., and Vapnik, V. (1996), 'Discovering Informative Patterns And Data Cleaning', In Fayyad, U., Piatetsky, G., and Smyth, P. (Eds), *Advances In Knowledge Discovery And Data Mining*, AAAI Press/ The MIT Press, California, pp. 181-203.

[26] Han, J., Lakshmanan, L., and Ng, R., T. (1999), 'Constraint-Based, Multidimensional Data Mining', *IEEE Computer*, vol. 32, no 8, pp. 46-50.

[27] Hawk, S., and Bariff, M., L. (1995), 'An Examination Of Organizational Strategies For Supporting DSS', *Information & Management*, vol. 28, pp. 77-88.

[28] Humphries, M., Hawkins, M., and Dy, M., C. (1999), *Data Warehousing Architecture And Implementation*, Prentice Hall PTR, New Jersey.

[29] Inmon, W., and Hackathorn, R., D. (1994), *Using The Data Warehouse*, John Wiley and Sons, Inc., New York.

[30] Jardine, N., and Sibson, R. (1968), 'The Construction Of Hierarchic And Non-Hierarchic Classifications', *Comp. J.*, vol. 11, pp. 117-184.

[31] Kimball, R. (1996), *The Data Warehouse Toolkit: Practical Techniques For Building Dimensional Data Warehouses*, John Wiley and Sons, Inc., Toronto.

[32] Klein, M., and Methlie, L., B. (1995), *Knowledge-Based Decision Support Systems With Applications In Business*, 2nd ed., John Wiley and Sons Ltd., West Sussex.

[33] Lee, D., Newman, P., and Price, R. (1999), *Decision Making In Organisations*, Financial Times Management, London.

[34] Liu, B., Hsu, W., and Ma, Y. (2000), 'Pruning And Summarizing The Discovered Associations', In: *Proceedings Of ACM SIGMOD '2000 International Conference On Management Of Data*, May 16-18, Dallas, Texas, pp. 125-134.

[35] Liu, B., Hsu, W., Mun, L., and Lee, H., Y. (1999), 'Finding Interesting Patterns Using User Expectations', *IEEE Transactions On Knowledge And Data Engineering*, vol. 11, no 6, pp. 817-832.

[36] Long, L. (1989), *Management Information Systems*, Prentice Hall, New Jersey.

[37] Long, L., and Long, N. (2001), *Computers*, 8th ed., Prentice Hall, Upper Saddle River, New Jersey.

[38] Makram, H. (2000), *Data Warehousing For Student Recruitment Systems*, unpublished MBA Thesis, RITI, Cairo.

[39] Mattison, R. (1997), *Data Warehousing And Data Mining For Telecommunications*, Artech House, Inc., Norwood, Massachusetts.

[40] McSherry, D. (1997), 'Knowledge Discovery By Inspection', *Decision Support Systems*, vol. 21, pp. 43-47.

[41] Moulet, M., and Kodratoff, Y. (1995), 'From Machine Learning Towards Knowledge Discovery In Databases', In: *Colloquium On Knowledge Discovery In Databases*, The Institute of Electrical Engineers (IEE), London, pp. 5/1-5/3.

[42] O'brien, J. (1996), *Introduction To Information Systems*, Irwin, Chicago.

[43] Onder, J., and Nash, T. (1999), 'The Approach To Building A Business Data Warehouse', SYSIX [online], URL: http://www.sysix.com

[44] Pyle, D. (1999), *Data Preparation For Data Mining*, Morgan Kaufmann Publishers, Inc., California.

[45] Ramirez, R., Kulkarni, U., and Moser, K. (1996), 'Derived Data For Decision Support Systems', *Decision Support Systems*, vol. 17, pp. 119-140.

[46] Reynolds, G. (1995), *Information Systems For managers*, West Publishing company, St. Paul, MN.

[47] Saunders, M., Lewis, P., and Thornhill, A. (1997), *Research Methods For Business Students*, Financial Times-Pitman Publishing, London.

[48] Silberschatz, A., and Tuzhilin, A. (1995), 'On Objective Measures Of Interestingness In Knowledge Discovery', In: *Proceedings Of KDD Annual Conference*, Montreal.

[49] Silberschatz, A., and Tuzhilin, A. (1996), 'User Assisted Knowledge Discovery: How Much Should The User Be Involved', In: *Proceedings Of ACM SIGMOD '96 International Conference On Management Of Data*, June 4-6, Montreal, Quebec.

[50] Sørensen, J., and Alnor, K. (1999), 'Creating Data

Warehouse Using SQL Server', In: *Proceedings Of The International Workshop On Design And Management Of Data Warehouses DMDW' 99*, Heidelberg, pp.10-1:10-9.

[51] Stevens, P. (1991), 'Decision Support Systems In Banking', In: *Colloquium On Decision Support Systems*, The Institute of Electrical Engineers (IEE), London, pp. 3/1-3/4.

[52] Taha, Y., Helal, A., and Ahmed, K. (1997), 'Data Warehousing: Usage, Architecture, And Research Issues', *ISMM Microcomputer Application Journal*, vol. 16, no 2, pp. 70-80.

[53] Teo, T., and King, W., R. (1996), 'Assessing The Impact Of Integrating Business Planning And IS Planning', *Information & Management*, vol. 30, pp. 309-321.

[54] *The Egyptian Supreme Council Of Universities Statistics*. (1999), The Egyptian Supreme Council of Universities, Cairo.

[55] Turban, E. (1993), *Decision Support And Expert Systems, Management Support Systems*, Macmillan.

[56] Turban, E., and Aronson, J. (1998), *Decision Support Systems And Intelligent Systems*, 5th ed., Prentice Hall, New Jersey.

[57] Widom, J. (1995), 'Research Problems In Data Warehousing', In: *Proceeding Of The 4th Int'l Conference On Information and Knowledge Management (CIKM) 95 ACM*, pp. 25-30.

[58] Wijsen, J. (2001), 'Trends in Databases: Reasoning and Mining', *IEEE Transactions On Knowledge And Data Engineering*, vol. 13, no 3, pp. 426-438.

[59] Wishart, D. (1998), 'Efficient Hierarchical Cluster Analysis For Data Mining And Knowledge Discovery', *Computing Science And Statistics*, vol. 30, pp. 257-263.

[60] Wishart, D. (2000), *FocalPoint Clustering User Guide*, Clustan Limited, Edinburgh.

[61] Wixom, B. H., and Watson, H. J. (2001), 'An Empirical Investigation Of The Factors Affecting Data Warehousing Success', *MIS Quarterly*, vol. 25, no 1, pp. 17-41.

[62] *World List of Universities and other Institutions of Higher Education*. (2000), 22nd ed., Macmillan Reference Ltd., London.

[63] Yossef, M. (1998), *Web-Based Student Registration Systems*, unpublished MBA Thesis, The Advanced Management Institute (AMI), Alexandria.