Integration of Decision Support Systems and Data Mining for Improved Decision Making

Omar al-Ketbi and Marc Conrad
Institute for Research in Applicable Computing, University of Bedfordshire, Luton, U.K.

Keywords: Data Mining, Decision Support Systems, System Integration, Abu Dhabi Police.

Abstract: A data mining (DM) integrated decision support system (DSS) is suggested to improve the performance of the DSS implemented in a police organisation. A prototype of the suggested system is provided. The paper provides an insight into DM-DSS integrated systems in the literature, and uses the results of the investigation as a basis for a suggested system tailored against the particularity of Abu Dhabi (AD) Police Organisation in the United Arab Emirates. It is believed that in general DM processes are demanding in terms of time and other resources. Hence, it is suggested that the process be reduced in order to accommodate to the particularity of AD Police in terms of size and current system performance. The system is perceived on this basis.

1 INTRODUCTION

Abu Dhabi (AD) Police operate with other UAE police departments through the Ministry of Interior to achieve a safer society. The Police serve four major UAE districts: Abu Dhabi, Al-Ain City, the external region, and the western region. AD Police have several units, which include police patrol, emergency response, crime investigation, and traffic. The primary objective of the Police is to become an intelligence-led, proactive police force that reacts to the needs of society with the highest level of integrity and training.

For this aim, AD Police have constantly undergone development of their information systems in order to integrate a range of processes that include the human side as well as the software and hardware deployment with the ultimate goal to improve the accuracy and quality of their undertaken decisions. AD Police have established the “Decision-Making Support Center” to help the Police explore the future challenges rather than just conduct research on the current phenomena. The Center also helps in quality assessment and control. In 2011, the Police implemented a GIS (Geographic Information System) in order to integrate people and processes for making better decisions.

However, the outcomes have not been very promising: slow and inefficient results have been noticed compared to the set expectations. Decision making is supported by the Quality Department that serves as a hub which provides quantitative and qualitative data received from different departments and organisational bodies countrywide. This system then would serve to support accurate and relevant decisions. The decision process is usually not straightforward and takes several factors into consideration. Typical challenges are associated with data formats, content, validity and reliability. It is also important to acknowledge the human factor involved in the decision making process. According to Hofstede (2001) the Arab culture is characterised by collectivism and high-power distance. Collectivism and high-power distance are two dimensions set by Hofstede in his attempt to establish quantifiable aspects of culture. Collectivism refers to a higher extent to which the group is prioritised over the self-interest as opposed to individualism. Hofstede asserts that cultures with high power distance, such as Arab culture, tend to be more collectivist. This has implications into how training and raising awareness are implemented to lead to effective methods of system development.

Abu Dhabi Police aim to realise the full and potential value of the data that it acquires. By the right framework and strategy, data that are highly utilised will eventually improve quality. Implementing data mining and combining currently separate collections can provide better information and hence knowledge to improve quality. A data
quality strategy will achieve consistent direction towards optimal decision making in the organisation. The improved data quality will then ensure that the organisation is more able to make informed and accurate decisions on policies and strategies.

Decision Support Systems (DSS’s) comprise different aspects of software, hardware, and data, as well as human inputs in order to help decision makers improve and enhance their decisions based on analytical processing of the available information. DSS’s have found uses in different types of organisations and wherever strategic decisions are to be made with high uncertainty involved. On the other hand, Data Mining (DM) allows robust analysis of huge amounts of data in order to discover useful relationships or patterns based on advanced statistical methods.

This paper provides insights into a prototype of an integrated DM-DSS (Data Mining – Decision Support System) solution for AD Police to assist decision makers in the decision making process based on better knowledge of the stored data. The solution can be implemented by integrating a set of DM techniques in the already established DSS in the organisation. The new system is expected to increase the efficiency sought from the DSS by acquiring knowledge of the data fed to the system.

2 BACKGROUND AND RELATED WORK

Public organisations face real challenges of using the correct analysis of huge amounts of data. These data are used for producing statistical analyses and forecasts on economic, social, health and education issues, which are highly related to government planning in aspects such as economic growth, development of interest rates and inflation, household income, education standards, crime trends and climate change are a major input.

Decision support systems (DSS’s) are widely used in businesses around the world for the main aim of helping executives to make better decisions based on advanced levels of data refined and presented to them. According to Hardin and Chhieng (2007) DSS’s refer to a class of computer-based systems that help in the process of decision making. Similarly Liu et al. (2010) report that a DSS is commonly defined in the literature as an interactive computer information system designed to support solutions to problems with taking decisions.

Padhy et al., (2012) argue that the value of strategic information systems is easily recognised yet efficiency and speed are not the only factors of competitiveness. The large amounts of data have called for new methods to analyse and understand the relationship in between this data. Conclusions and inferences from these data need special tools and techniques that are able to delve deeper than traditional decision support systems can. Moreover, the rapid development of data digitisation produced data stored in organisations’ data warehouses, which required efficient exploitation and knowledge extraction. Consequently, traditional problem solving DSS’s became less efficient and started to decline in the 1990s (Liu et al., 2010). Liu et al. (2010) classify the main challenges facing DSS’s in supporting decision making. These challenges include:

- Changes in technology from database to data warehouse and on-line analysis processing (OLAP), from mainframe to client/server architecture, and from single user model to World Wide Web access;
- Increasing business interconnections in a more dynamic business environment and intelligence. For this a variety of other information systems have been proposed, such as supply chain management (SCM), enterprise resource planning (ERP), and customer relationship management (CRM);
- The continuous increase in complexity in decision making which requires executives to consider a vast number of inputs and a considerable amount of knowledge.

As Han and Kamber (2006 pp.4-5) point out, organisations are usually data rich but information poor. Data mining techniques help to analyse data and uncover important data patterns, which may contribute to better, knowledge-based strategies. In doing so, data mining helps to bridge the gap between data and information, which usually prevents realising knowledge.

Given the uses and nature of solutions based on data mining and those based on decision support systems, it can be confidently suggested that such solutions can be integrated in order to offer optimal solutions to knowledge-based decision making processes. However, only few examples of using data mining to support management decisions are found in the literature which will be discussed in the following.

For example, Abu-Naser et al., (2011) suggest a DSS based on data mining techniques for optimising e-Learning in educational institutions. The suggested system was not implemented, but according to the
authors, integrating data mining functionality into a single DSS will be promising. The authors believe that such a system will enable the educational institutions to realise the importance of the DSS-produced information in optimising their adopted learning strategies.

El Seddawy et al., (2012) propose a DM-based DSS to support top level management to make a good decision in any time. According to the authors, the proposed system can help decision makers in the banking sector to address decisions related to new investments.

In the application area of health services, Kumar et al., (2011) suggest the use of DM based on decision tree algorithms to classify certain diseases and compare the effectiveness and correction rate among them in order to support decisions on the diagnostic process. According to Kumar et al., (2011), traditional decision support systems developed to assist physicians in the diagnostic process often are based on static data which may be out of date. Hence, a decision support system which can ‘learn’ the relationships between certain parameters would be useful to physicians and hospitals.

Mohemad et al., (2010) argue that traditional support systems that are widely used in the construction industry are not optimal. Despite efforts to integrate and transform the whole construction tendering processes into electronic or digital forms, the use of unstructured documents either in hard copy or digital are still widely present. The authors stress the need to extract and represent information in machine-readable formats, attained by integrating data mining in DSS model, which they believe to be a promising approach.

Liu et al., (2010) have conducted research into integrated decision support systems (IDSS) including DM agent-enhanced integrated DSS to improve decision support performance. The researchers conclude that the main challenges in developing an integrated DSS are the trade-off between loose and tight integration strategies within the integration frameworks and the seamless integration across data, models and processes within the integration frameworks.

Srinivasan et al., (2011) also suggest using DM intelligent agents in DSS for achieving higher work efficiency. They suggest that such a system can provide autonomy, mobility, and collaboration of different agents in order to provide a simple and fast solution. Srinivasan et al. maintain that using agents as data mining techniques, can help decision makers by providing a more robust and quick DSS in resolving issues in any complex situation.

Mladenić et al., (2003) maintain that there has been no systematic attempt to integrate DM and DSS. Reasons behind that are many but mainly include the nature of data mining processes that combine computer science and statistics, which create some confusion on what implementation aspects may be suitable for managerial decisions.

The following table (Table 1) summarises the different approaches to linking DM to DSS:

<table>
<thead>
<tr>
<th>Author</th>
<th>Context</th>
<th>Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abu-Naser et al., (2011)</td>
<td>E-Learning in educational institutions</td>
<td>DM-based DSS for optimised results</td>
</tr>
<tr>
<td>El Seddawy et al., (2012)</td>
<td>The banking sector</td>
<td>DM-based DSS to support top level management to make decisions in any time</td>
</tr>
<tr>
<td>Kumar et al., (2011)</td>
<td>Disease classification</td>
<td>DSS that can learn the relationships between certain parameters</td>
</tr>
<tr>
<td>Srinivasan et al., (2011)</td>
<td>Higher work efficiency</td>
<td>DM intelligent agents in DSS</td>
</tr>
</tbody>
</table>

In an initiative to address the above challenges, the EU sponsored from 1999 over a 39 month period the SoiEuNet project, which comprised a network of expert teams from business and academia to meet client’s Data Mining and Decision Support needs (Mladenić, 2001). The outcomes of the project were promising. The project identified the main objectives to improve collaboration and communication, promote awareness of organisational resources and achievements, and enable organisational learning and dissemination of such knowledge. However, certain difficulties were encountered as detailed below.

According to the final report at the project closure, the project team maintained that collecting and managing knowledge is a very hard task that can never be fully accomplished. The project team considered that the set of knowledge management tools to support e-collaboration in data mining and decision support can be considerably extended, improved and further integrated. The team suggested
including in the near future, database and data transformation services, automatically built activity logs for data mining and decision support, a data mining advisor that, given a dataset, suggests data mining algorithms, and a central model evaluation service. Another important improvement suggested by the team is to adopt a standardised description of the knowledge produced in all phases, which would significantly simplify communication among distributed cooperating groups.

3 PROBLEM DEFINITION

The aim of retaining information is not anymore a significant concern of the organisation as it is common that a numerous number of transactions are stored in its warehouses regularly, with only a few of real relevance to the organisation’s decision making process. The amount of information stored in databases increases daily and goes beyond the technical skills and human capacity to interpret that information.

Database management systems have advanced at a faster rate than the techniques used for extracting and utilising information to be used in making decisions (Power, 2008), in the sense of using the trend of the past endeavours to anticipate the future tendency (Lv and Li, 2009). Obtaining, storing and managing information in larger organisations are now ordinary operations of business companies and usually performed automatically by electronic data repositories (Saxena and Rajpoot, 2009).

One of the efficient techniques used for this aim (i.e. extracting useful information) is data mining. Some of the organisation’s information may be in a textual format described in natural languages or does not have a structure like that present in data tables and structured relational databases. This type of information, found mainly in the form of electronic documents and emails, particularly in organisations with limited affordability (technical and financial), cannot be used with traditional data mining tools, and thus minimises their potential. The data miner thus has to prepare the data for mining. This operation, called pre-processing, is complex and may take several months to complete, depending on the size of the project, and require a significant part of the organisational resources, and usually performed only by large enterprises (Calderon et al., 2003). Artificial Intelligence systems such as Neural Networks and Fuzzy Inference algorithms might be appropriate for finding the solution to this problem.

Law enforcement organisations can highly benefit from a DM-DSS integrated system regarding the vast and loosely related information these organisations deal with. According to Oatley et al., (2006), computer technologies that support criminal investigations are wide ranging and include geographical information systems displays, clustering and link analysis algorithms and the more complex use of data mining technology for profiling crimes or offenders and matching and predicting crimes. They also argue that knowledge from disciplines such as forensic psychology, criminology and statistics are essential to the efficient design of operationally valid systems. According to McCue (2003) however, one of the biggest challenges in using data mining and predictive analytics in law enforcement is that most, if not all, data encountered was never intended to be analysed. Therefore, significant challenges associated with data form, content, reliability, and validity must be constantly evaluated and addressed.

The above issue applies as well to AD Police. Given the state-of-the-art DSS implemented in the organisation, little success has been seen in improving the day-by-day tasks. The decision process in the organisation is usually not straightforward and involves many sub-processes. Some of the challenges faced are related to data formats, content, validity and reliability.

4 DATA MINING TECHNIQUES

Recent developments in Information Systems as well as the availability of extensive business data repositories and database management systems, accompanied by the advances in computer systems and algorithms, have provided a gateway to pattern matching and data models (Hand et al., 2001).

Data Mining consists of a set of techniques inferred from Statistics and Artificial Intelligence with the specific aim of discovering new, useful, relevant and non-trivial knowledge that may be hidden in a large mass of data (Markov and Larose, 2007). There have been numerous examples of its uses in areas such as Marketing, Economics, Engineering, Medicine and others. These techniques are briefly described below.

Several techniques mentioned in the literature of data mining such as classification, neural network, genetic algorithm and others have long been known (Segall and Zhang, 2006). Statistics on the contribution of data mining techniques to selecting data and evaluating the results of data mining like clustering are also usually considered. What
distinguishes the perception of data mining is the development of techniques for data mining applications on a large scale databases. These techniques are generally applied to the data on a small scale. In addition, several techniques from the fields of databases for data transformation are also an integral part of the process of data mining.

Depending on the application domain and user interest, various types of techniques can be identified and applied. Some of these techniques are briefed below in the order of relevance to the DSS of AD Police. For example, dimensionality reduction is useful when the number of involved variables exceeds the capacity of the DSS to perform better, whereas modelling helps improve the DSS operations by feeding it with a stripped-down version of a collection of data which models the entire collection.

**4.1 Dimensionality Reduction**

Dimensionality reduction is a mathematical technique used to reduce the number of random variables involved in a dataset. It uses projection from one vector space onto another one of lower dimension.

**4.2 Correlation**

Correlation is a statistical method used to depict relationships between random variables through analysing potential links and inferring the degree of connectivity. Given the large number of loose data available for the DSS, correlation helps establish useful relationships among seemingly unrelated variable.

**4.3 Modelling**

Modelling or mathematical modelling is a description of observed behaviour, simplified by ignoring certain details to simulate the behaviour of a phenomenon. Modelling allows complex systems to be understood and their behaviour to be predicted within the scope of the generated model.

**4.4 Association**

Finding association is a data mining technique that allows searching for simultaneously occurring items that in the transactions database. Algorithms such as DHP (Dual Heuristic Programming), GSP (Generalised Sequential Pattern) and Apriori (Wang et al, 2006) among others are examples of tools that implement the task of discovery of association.

**4.5 Classification**

Classification is a method that consists of defining a mathematical function that maps a set of records to one another in a predefined set of categorical labels, called classes. This function is applied to predict the class new records fall under.

**4.6 Regression**

Regression includes a search for a function that maps the records from a database to actual values. This method is similar to classification, but being restricted only for numeric values.

**4.7 Clustering**

Clustering is used to separate records in a database into subsets or clusters, such that the elements of a cluster share common properties that distinguish them from other clusters. The objective of this task is to maximise intra-cluster similarity and minimise inter-cluster similarity. Unlike classification task, which has predefined labels, clustering needs to automatically identify the data groups to which the user must assign labels. Some algorithms used for implementing this method are K-Modes, K-means, K-Prototypes, K-Medoids, among others. (Han et al, 2006).

**4.8 Summarisation**

This task is very common in KDD (Knowledge Discovery in Databases), is to seek, identify and indicate common features among data sets. Inductive logic and genetic algorithms are some examples of technologies that can be applied in summarisation.

**4.9 Detection of Deviations**

This technique helps to identify records in the database whose characteristics do not meet the normal standards. Statistics is the main resource provider used by this technique.

**4.10 Discovery Sequence**

It is an extension of the technique of finding associations that are sought frequent by considering several transactions occurring over a period. The technique of association can be adjusted to engage the generalised mining association rules. The post-
processing step includes processing the knowledge gained in data mining. Among the main tasks of the post-processing step are: preparation and organisation and may include the simplification of charts, diagrams or reports demonstrating, in addition to the conversion of the representation of knowledge gained.

5 APPLICATIONS IN SCIENCE AND BUSINESS

Lessons learned from science and business applications of data mining can be transferred to a certain degree to the situation of AD Police. Practitioners have devised methods for obtaining relevant information from a large set of data, such as dimensionality reduction, correlations, modelling and DSS. In their endeavour, engineers have used a range of scientific fields in physics, mathematics, statistics, artificial intelligence and machine learning.

These tools have not found intensive applications so far (Han et al., 2002) that can vigorously utilise their efficiency. This is referred in some extent to that business and IT solution providers are essentially interested in extensibility and automation, and they aim at obtain fast results via combining simple analysis with human expertise (Ganguly and Gupta, 2005).

6 DM-BASED DSS SOLUTION

The vision of the Strategic and Performance Development General Directorate of AD Police is to transform the UAE into one of the safest and most secure countries in the world. In this, they aim to attain a more effective police force which responds to the needs of society with the highest level of integrity and training. For this aim, the organisation is investing in new technologies to assist leaders in decision-making and training, including Decision Support Systems.

The abovementioned aim of the organisation to make the country safer requires making decisions on resource allocation. The available resources are divided into two categories: policemen and financial resources. These resources are to be allocated among assets, such as policing roads and streets, establishing new services, opening new departments and sections, and building new police centres. The decisions taken are concerned with the optimal allocation of the available resources among the assets. The DSS considers many variables including but limited to number of crimes per area, crime evolution, crime types, population per area, population growth rate per area, national population growth rate, available budgets, number of staff members, balanced development, and general guidance of the Directorate. These variables stem from a multitude of data supplied to the system from the different departments of the organisation (Figure 1).

![Figure 1: Resource allocation among assets at AD Police based on the information provided by the DSS.](image)

Optimising the decision making process requires targeting different stages of the process, including the quality of the supplied data. This paper suggests a framework to improve the DSS implemented in the AD Police organisation in order to optimise the DSS and hence the decisions made in the organisation.

Database management systems have advanced at a faster rate than the techniques used for extracting and utilising information to be used in making decisions (Power, 2008), in the sense of using trends of past endeavours to anticipate the future tendency (Lv and Li, 2009). Obtaining, storing and managing information in larger organisations are now ordinary operations of business companies and usually performed automatically by electronic data repositories (Saxena and Rajpoot, 2009).

One of the efficient techniques used for this aim (i.e. extracting useful information from large data) is data mining. Some of the organisation’s information may be in a textual format described in natural languages; even systematically collected information does not necessarily has a structure like the one used for data tables and structured relational databases. This type of information, found mainly in the form of electronic documents and emails, particularly in organisations with limited affordability (technical and financial), cannot be used with traditional data mining tools, and thus minimises the potential of such organisations. The data miner thus has to prepare the data for mining. This operation, called pre-processing, is complex and may take several months to complete, depending on the size of the
project, and require a significant part of the organisational resources, and usually performed only by large enterprises (Calderon et al, 2003). Artificial Intelligence systems such as Neural Networks and Fuzzy Inference algorithms might be appropriate for finding the solution to this problem.

For police operations, however, the process needs to be reduced in order to accommodate to the particularity of AD Police in terms of size and current system performance. For AD Police, Data Mining Techniques (DMT) and DSS techniques can be used to design and implement custom applications that help in pattern recognition, designing of predictive models, human-error reduction, civil society’s protection, infrastructures maintenance, quality enhancement and economic and environmental sustainability (Ripley, 2008).

What we propose to enhance the currently used DSS as opposed to the currently used DSS in the organisation is improved implementation of DMT based on extensive utilisation of DMT to attain optimal results gained from advanced multi-disciplinary combinations of these techniques. The proposed system will require integration of a set of techniques to be able to deliver the desired outcomes. Those techniques consist of the database system, the business logic and a user interface model. The work will contribute to the development and enhancement process through defining a development approach and setting up all the needed variables for establishing the solution. The proposed DSS system will benefit from data mining processes prior to having to deal with huge amounts of data acquired from the different departments involved. The techniques involved will be selectively deployed on the different sets of data acquired (Figure 2).

Several components will make up the data mining system. The collected data will selectively undergo several processes prior to be sent to the DSS. These processes are data preparation, classification, cluster analysis and association analysis. Once done, the data enter into anomaly detection process where anomalies in the processed data will be detected and flagged up for further investigation. These components are particularly relevant to the AD Police DSS system given the data formats and the other constraints such as time and affordability limitations (Figure 3).

Figure 3: A diagram of the final DM-DSS depicting the data mining processes that take place prior to the DSS implementation stage. The DSS acquires relevant data to support decisions on resource allocation.

7. CONCLUSIONS

This paper looked at the integration of DM-DSS (Data Mining – Decision Support System) in the literature in order to propose a solution for Abu Dhabi Police. It was shown that the integration of DM and DSS based solutions is not widely implemented, although several researchers have suggested the robustness of such integration. Given the particular case of Abu Dhabi Police, the currently used DSS can highly benefit from a DM-DSS-based solution to improve the limited performance and unsatisfactory results attained. The suggested system uses DM as a pre-process to the DSS implementation, all in an integrated DM-DSS. The paper provided a prototype of DM-DSS integrated system as part of the framework for AD Police for in their data quality project. Given the sparse data formats and the limited time and affordability of the organisation, the suggested prototype is tailored to the particularity of AD in terms of size and DM techniques chosen. For the implementation part of the prototype, it is necessary to acknowledge that ‘Western’ processes usually cannot be transferred directly. It is also necessary to acknowledge that decision making process in the West cannot be transferred directly to the case of Abu Dhabi Police given the special aspects of Arab culture. Therefore, cultural differences are significantly influential in any system implementation involving humans.
REFERENCES


El Seddawy, A. B., Khedr, A., & Sultan, T. Adapted Framework for Data Mining Technique to Improve Decision Support System in an Uncertain Situation.


Kumar, D. S., Sathyadevi, G., & Sivanesh, S. Decision Support System for Medical Diagnosis Using Data Mining. *International Journal of Computer Science, 8*.


Srinivasan, S., Singh, J., & Kumar, V. Multi-agent based decision Support System using Data Mining and Case Based Reasoning. *International Journal of Computer Science, 8*.