Data Mining Technology for Structural Health Monitoring

Zhongdong Duan*, Kun Zhang**

Abstract: Data mining (DM) technology provides a potential solution to structural health monitoring (SHM) as a deeper data analysis method. In this paper, a comprehensive introduction to the background, definition, function, process, methods and advantage of DM technology is made. The research and application state of DM technology in various fields is reviewed. Responding to the need of SHM and the characteristics of monitoring data, a DM framework for SHM is proposed. This framework includes task definition, build of data-house, data preprocessing, data mining and scheme evaluation. Data preprocessing is stressed for an effective DM scheme, which may include data selection and cleaning, data reduction and transformation. The DM framework is integrated to a DM platform for SHM and the DM platform is modularized. The main tasks which need be solved in using DM technology for SHM are analyzed. Some examples are illustrated on how to utilize specific DM techniques to mine monitoring data.

Keyword: Structural health monitoring, data mining, framework, platform, modularization

INTRODUCTION

In recent years the frequent occurrence of natural disasters and the continuous enhancement of people’s capability for disaster prevention and reduction bring the need of structural health monitoring (SHM) technology for civil structures. The flying advancement of science and technology promotes the prosperous development of SHM technology. A great number of civil infrastructures, such as long span bridges, large hydraulic structures, nuclear power stations, tall buildings and large space structures, have been instrumented SHM systems to monitor the state and damage of them [1-7]. With the development of data acquisition and transmission technology, the capability of SHM system to collect data is enhancing gradually. Large amount of monitoring data are increasingly becoming available.

However, the explosive growth of monitoring data has outstripped the capability of current data processing techniques to extract useful information to perform SHM and structural damage identification. On the one hand, the structures of civil infrastructures are complex and the external factors influencing civil infrastructures are multivariate, so a large number of structural parameters and environmental parameters need be monitored. And the categories of monitored parameters are various because of their different properties, testing principles, acquisition and storage formats. Moreover, the uncertain factors causing noise and missing data etc are excessive since the SHM systems are inevitable to suffer from the influence of environment and external load. Therefore, the monitoring data of civil infrastructures has such characteristics as large data size, great varieties and low quality. On the other hand, besides the traditional mathematical statistics methods [8-11], some machine learning and artificial intelligence methods, such as artificial neural networks [12], wavelet analysis [13], fuzzy [14], chaos [15], genetic algorithm [16], support vector machine [17,18] etc, have also been adopted to process the data to identify possible structural damage. However the researches on above methods emphasize particularly the usability of methods and pay little attention to the characteristics of monitoring data.

It is the inconsistency between the characteristics of monitoring data and the limitation of identification methods that limits the development of SHM of civil infrastructures. So a data processing technique which can quickly and correctly handle a large quantity of monitoring data is indispensable in SHM. Data mining (DM) technology provides a potential solution to this problem as a deeper data analysis method.

In this paper, a comprehensive introduction of DM technology is made. The research and application state of DM technology in various fields is reviewed. Responding to the need of SHM of civil structures and the characteristics of monitoring data, a DM framework for SHM is proposed and is integrated to a modularized...
INTRODUCTION OF DM TECHNOLOGY

Background of DM

Contributing factors including popular use of computer and internet, advance in data collection and storage tools, etc, have flooded us with a tremendous amount of data. The explosive growth of data has caused information excessive and difficult to identify and digest, so it is urgent to find new technologies and automated tools which can intelligently assist us in transforming the vast amounts of data into useful information and knowledge [19]. DM as a promising technology frontier is applied to solve this problem and has attracted a great deal of attention in the information industry. This has been fostered by the advancements in large-scale database, advanced computer technologies and mature DM methods to meet the need of business management, production control, market analysis, engineering design and science exploration.

Definition of DM

DM has been defined in various ways. Fayyad et al. [20] gives an explanatory definition “the nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data”. Fig.1 gives a visual explanation of this definition.

Function of DM

DM can describe generic characteristics of historical and current data, predict future trend and behavior and make knowledge-based decisions through normatively describing the information in data [21]. DM mainly has the following five functions.

Association analysis

Association analysis identifies relationships or affinities between entities or between variables in a given dataset [22]. Its purpose is to discover associations, correlations, or causal structures showing patterns that occur frequently together. Association rules and sequential patterns are the main techniques of association analysis. Association rules are used to find the correlation or mutual dependence among a piece of things and other things. Sequential patterns focus on analyzing the causal relation between data.

Classification and prediction

Classification is the process of finding a set of models (or functions) to describe and distinguish data classes. The derived model is based on the analysis of a set of training data and it may be presented in various forms, such as classification rules, decision trees, mathematical formulae or neural networks. It can be used to predict the class label of data from history data, find the predictive information, and automatically derive the prediction model which can be used to predict the trend of future data.

Concept description

Concept description describes individual concepts in summarized, concise, and yet precise terms. It can be used to describe the connotation of data and summarize the characteristic of data. These descriptions can be divided into data characterization and data discrimination. The former is realized by summarizing the data of the studied class, but the latter by comparing the target class with one or a set of comparative classes.

Outlier analysis

Outlier analysis detects data which don not comply with the general behavior or model of the data in a database. Its basic principle is to detect the meaningful difference between observation and reference value. And it is useful in some applications such as fraud detection.

Cluster analysis

Unlike classification and prediction, which analyzes class-labeled data, cluster analyzes data without consulting a known class label. Cluster analysis can generate such labels. And then the data are clustered based on these labels so that data within a cluster have high similarity to one another. Cluster analysis can enhance the cognition of people to objective rules and is prerequisite to concept description and outlier analysis when the class labels are unknown [23].

To what attention should be paid is that above five functions often connect each other and function together in the application of DM [24, 25].
Methods of DM

DM is a multidisciplinary field, drawing attention from areas including mathematical statistics, machine learning, artificial intelligence, database technology, high-performance computing and data visualization etc [26], and the corresponding mining methods mainly include statistical methods, machine learning methods, artificial intelligence methods, and database techniques. Statistical methods involve regression analysis, clustering analysis, principal component analysis (PCA), fuzzy set, and rough set etc. Machine learning methods involve support vector machine (SVM), decision tree (DT), case-based reasoning (CBR), genetic algorithm (GA), and Bayes belief networks etc. Artificial intelligence methods involve artificial neural networks (ANNs), particle swarm algorithm, ant colony algorithm and multi-agent technique etc. Database techniques involve visualization and multidimensional data analysis etc. Some DM methods in common use are introduced in detail as following.

Clustering

Clustering, as a branch of statistics, is under vigorous development owing to the huge amount of data collected in databases. In this method, data are divided into different groups according to similar assessing criteria and the comprehensive descriptions of data in the same group are given.

Principal component analysis (PCA)

PCA is a statistical technique that linearly transforms an original set of variables into a substantially smaller set of uncorrelated variables that represents most of the information in the original set of variables [27]. It can be viewed as a classical method of multivariate statistical analysis for achieving a dimensionality reduction. Because a small set of uncorrelated variables is much easier to understand and use in further analysis than a large set of correlated variables, this data compression technique usually is used to deal with large-scale sparse data.

Rough set

Rough set theory was proposed by Pawlak [28] as a new mathematical tool for reasoning about vague, uncertain and imprecise information. It can be used to evaluate the importance of particular attributes in relationships between objects and decisions, reduce all redundant objects and get minimal subsets of attributes, create models of the most representative objects in particular classes of decisions and represent the relationship between objects.

Fuzzy set

This algorithm utilizes fuzzy set theory to carry on fuzzy judgment, fuzzy decision, fuzzy pattern recognition and fuzzy cluster analysis for practical problems [29]. Fuzziness objectively exists. The more complex the system is, the lower the accuracy of system is and the stronger the fuzziness of system is. Fuzzy modeling and fuzzy decision are commonly used in DM.

Support vector machine (SVM)

SVM, a new machine learning method, originates from mathematical statistics and is based on the principle of structural risk minimum. It is renowned for the computational ability to project data into a trillion dimensions and the statistical ability to process a classic overfitting trap. Owing to good popularization nature and better categorized accuracy, SVM can deal with a great deal of problems such as regression, pattern recognition, prediction, comprehensive evaluation and so on.

Decision tree (DT)

Decision tree is one of the most popular classification algorithms in current use of DM. This algorithm uses an entropy-based measure as a heuristic for selecting the attributes, which can best separate the samples into individual classes. This algorithm usually is used for classification in DM with the merits of simple description and fast classification, especially suitable for processing massive data. But it may miss valuable rules when information is sparse.

Case-based reasoning (CBR)

CBR is the analogy and deduction methods showing systematic knowledge with cases in the knowledge base. According to the similar measurement principle, new cases are matched to cases in the knowledge base to realize assessment. This method, based on experiential knowledge, is in accord with human behavior, so it is universally applied in the field where professional knowledge is difficult to summarize. But the flexibility and adaptability of this method is not satisfying.

Genetic algorithm (GA)

GA is based on Darwin’s theory of “survival of the fittest”. The input of GA is an initial set of possible solutions. The output of GA is the set of “fittest solutions” that will survive in a particular environment.
The process produces the next generation by a cross-over function, evolves solutions by a mutation function and discards weak solutions based on a fitness function. GA often expresses mining task as a kind of search problems to display its optimization and search ability.

**Artificial neural networks (ANNs)**

ANNs are massively parallel computing systems consisting of extremely large number of simple processors (neurons) with many interconnections. They attempt to simulate the way that the human brain works. ANNs are shown to be universal approximators of nonlinear functions, so they are capable of learning very complex nonlinear input-output relationship. They can realize many kinds mining task such as classification, clustering, and feature mining and so on.

**Visualization and multidimensional data analysis**

Visualization is the conversion of data into a visual format so that the characteristics of data and the relationships among data or attributes can be analyzed. Multidimensional data analysis means that adopts various movements such as slicing, stripping, rotating etc. to anatomize the data organized in multidimensional form, in order to ensure users can clearly observe the data from different angles and deeply understand the information implied in data. Visualization and multidimensional data analysis can make the process and results of DM easier to understand and is helpful to check mining results and direct mining process.

**Multi-Agent technology**

Parallel computation is an effective method for improving the utilization efficiency of the large-scale database. Multi-Agent technology is a new developing method in parallel computation. Agent is "a computer system, situated in some environment that is capable of flexible autonomous action in order to meet its design objectives” [30]. Multi-agent system is composed of many agents. According to parallel distributed treatment technology and module design philosophy, the complicated system can be divided into relatively independent agent subsystem and can resolve problems through the cooperation between agents.

From basic DM to complex DM, and then to intelligent DM, DM methods are enriched and improved constantly. No matter which kind of methods is employed, final goals are all to find valuable knowledge and information. General process is common, but each method has its own function, characteristics and application environment, so the choice of methods will influence the efficiency and quality of mining. Usually some algorithms integrate the ideas of several methods to have complementary advantages.

**Process of DM**

The basic process and the detailed process of DM are shown in Fig2. The process of DM mainly includes task definition, data aggregation, data preprocessing, data mining and scheme evaluation.

Task definition confirms the purpose of DM after understanding the knowledge of relevant fields, which is the foundation of the whole DM process, the standard to examine the mining results and the guide for analysts to finish DM.

Data aggregation means that the relevant data are drawn from the databases to form the data-house as requested.

Data preprocessing is in charge of data selection, cleaning, reduction and transformation to extract attribute information in appropriate forms according to the requirement of data mining.

From basic DM to complex DM, and then to intelligent DM, DM methods are enriched and improved constantly. No matter which kind of methods is employed, final goals are all to find valuable knowledge and information. General process is common, but each method has its own function, characteristics and application environment, so the choice of methods will influence the efficiency and quality of mining. Usually some algorithms integrate the ideas of several methods to have complementary advantages.

**Process of DM**

The basic process and the detailed process of DM are shown in Fig2. The process of DM mainly includes task definition, data aggregation, data preprocessing, data mining and scheme evaluation.

Task definition confirms the purpose of DM after understanding the knowledge of relevant fields, which is the foundation of the whole DM process, the standard to examine the mining results and the guide for analysts to finish DM.

Data aggregation means that the relevant data are drawn from the databases to form the data-house as requested.

Data preprocessing is in charge of data selection, cleaning, reduction and transformation to extract attribute information in appropriate forms according to the requirement of data mining.
of data. It is applied to extract data pattern based on the preprocessed data.

Scheme evaluation is an effective demonstration of the DM process, where visualization and expert knowledge are used to present the mined knowledge to the user. Once the scheme is evaluated to be effective, it can be implemented in this problem.

DM process is spiral ascending as shown in Fig.2(a), because numerous circulations are repeated by adjusting the corresponding steps in order to ensure the mining results are consistent with the expected results.

Advantages of DM

DM originates from mathematical statistics, machine learning and artificial intelligence, and absorbs many technologies from above three fields. Therefore it has the following predominance.

1) DM is driven by data to find unpredictable mode through heuristic inductive learning, whereas mathematical statistic guides data processing with theory and assumption.

2) DM focuses on all available data and can handle large-scale data set with noise, missing values and high dimension data, whereas machine learning mainly pays attention to how to improve the ability of machine learning with the less data source.

3) DM theory has carried on effective integration of various existing advanced methods and algorithms, and has defined system, function and process of mining, which makes the whole mining process clear, definite and easy to implement. But artificial intelligence lacks unified mining models to process objective information and intact systems to classify and merge various methods.

DM not only relies on the technologies of the above three fields but also needs support of other technologies. But it is an integrated whole, not just a simple association of these technologies. Therefore DM can have more powerful technological advantages with the gradually enhancement of theoretical research and practical application.

RESEARCH AND APPLICATION STATE OF DM

Research state of DM

DM is also popularly referred to as knowledge discovery in databases (KDD). Its essential is applying knowledge discovery in database. The term KDD was coined at the first KDD workshop in 1989 to emphasize that “knowledge” is the end product of a data-driven discovery. It has been popularized in artificial intelligence and machine learning. Up to now, international academic conference on KDD is held every year. Pacific-Asia conference on KDD (PAKDD) is also held every year since 1992. In 1993, Knowledge and Data Engineering association of IEEE takes the lead in publishing KDD technological monograph. In addition, the international academic publications of fields such as database, artificial intelligence, information processing, knowledge engineering etc have published KDD special topic or monograph one after another. The research on DM can be divided into macroscopic research and microscopic research. The former focuses on the necessity, intension and theory overview of DM, but the subjects of the latter include the task, method, technology, language and system of DM. In recent years, the emphasis of DM research is gradually changing from developing methods to systematic application, especially paying more attention to the integration of methods and the mutual infiltration between disciplines [31].

The development of DM systems is an important aspect in DM research, which is the impetus of developing other techniques of DM. Many companies have paid attention to developing DM systems and more than one hundred DM tools have been developed at present [32]. The typical mining systems include Intelligent Miner of IBM company, Clementine of SPSS company, Enterprise Miner of SAS company, MineSet of SGI company, Darwin of Oracle company and so on.

Application of DM

DM is born to meet application such as simple search, query, analysis, systematization and deduction in database. At first, DM is applied to commercial fields including bank, telecommunication, insurance, traffic and retailing etc. where information is intensive. In recent years, DM has a successful application in athletics sports, medical care, judicial authentication, and Internet’s information filter. In addition, DM also has been used to classify the universe picture data, analyze monitoring data of crustal activity gathered by satellites and to construct living model. Whit the improvement of mining methods, the evolution of mining language, the systematization of mining process, the emergence of mining tools and software, DM will serve more and more fields in the society as an industry.

Application of DM in engineering

The application of DM in engineering mainly focuses on mechanical field. In 1992, German
Management Intelligenter Technologien (MIT) GmbH developed DataEngine, a combined software platform, to carry on intellectual data analysis and data mining. This platform has two functions. One function is maintaining and controlling machines by detecting the state of bearing device based on Kohonen network and fuzzy Kohonen network methods. The other function is monitoring and diagnosing machine by identifying failure modes of machine on the basis of knowledge-based fuzzy system.

Research institute of information and technology belonging to National Research Council (NRC) of Canada developed an automatic DM system, Aerospace Data Miner (ADM). At first, this system can utilize DM technology to find out the reasons of element failure and anomalism and realize data visualization. And then, it can search model and trend of data objects in database using decision tree, CBR, rough set, and ANNs etc to provide decision support for the fault prediction and diagnosis of aircraft equipment.

Blair and Shirkhodaie of Tennessee state university in U.S.A. adopted DM and digital visual technique to design Multi-Agent, which is used in the diagnosis of bearing vibration. At first, statistical analysis and signal processing agents are adopted to extract the vibration attributes of bearing. And then, 3D graphics are used to track the reason and state of bearing fault.

Ogilvie et al. of Belfast Queen University in Britain used DM to monitor and optimize the performance of steaming system for thermal power plant. First of all, Apriori algorithm is adopted to mine association rules from historical databases of power plant and accurate model for describing the state of equipment is built. And then, model and equipment are used simultaneously. Through comparing the input-output relationship of equipment and that of model, the state of equipment is monitored. Finally, the association rules between components of equipment are further mined and are used to update the model.

A few applications of DM in SHM of civil engineering are launched. Sandhu and Lazarevic et al. of Minnesota University in U.S.A. made some exploratory work on utilizing DM techniques to detect the damage of complicated mechanical structures. In 2001, DM techniques including Decision tree and ANNs were used to realize the failure prediction of a plane frame structure and a three dimensional electrical transmission tower model as shown in Fig.3 (a) and (b). In 2002, a hierarchical clustering was adopted to detect the damage of airplane model as shown in Fig.3 (c). In 2004, a localized clustering-regression was utilized in damage detection of more complicated electrical transmission tower model as shown in Fig.3 (d).

In recent years, DM technology also has been adopted in some engineering fields such as the data processing of dam and slip monitoring and the data analysis of wind response of skyscrapers and so on.
DEVELOPMENT AND APPLICATION OF DM PLATFORM FOR SHM

DM framework for SHM

Responding to the need of SHM of civil structures and the characteristics of structural monitoring data, a framework for DM is proposed as shown in Fig.4. In this framework, firstly, mining task is defined according to the characteristics of SHM; then data-house is built based on collected monitoring data, and data selection and cleaning are applied with the purpose of creating data market; thirdly, data reduction and transformation are used to extract attribute information, and DM methods are chosen to constitute mining model and realize the state monitoring and damage evaluation of structures. Finally, visualization technology and expert examination are used in scheme evaluation. Powerful metadata, as scheduling strategy, is adopted to realize the organic integration of database with the steps of data selection, cleaning, reduction, transformation and the flexible joint of various DM methods.

Fig.4 shows that the DM framework for SHM is consistent with the process of DM. It is extensible because various methods for data preprocessing, data mining or scheme evaluation can be embedded in the corresponding steps of this framework. And this framework can be integrated to a high-efficient DM platform for SHM as shown in Fig.5 (a). This platform can be used to finish all analytical tasks and find the best solution to problems within the shortest time.

In order to make the proposed DM platform for SHM easy to analyze, set up and utilize, it is modularized into four modules including data-house build module, data preprocessing module, data mining module and scheme evaluation module, as shown in Fig.5 (b).

Data-house build module is mainly in charge of collecting all available data related to structural state and setting up the data-house suitable for SHM.

Data preprocessing module can improve the quantity of data by transforming the database with original and massive monitoring data into the data market suitable for mining, and consequently improve the efficiency of mining. It realizes extracting attribute information of structural state according to the requirement of mining methods. This module is composed of four steps including data selection, data cleaning, data reduction and data transformation. Data selection merges multiple sources into a coherent data set.

Data cleaning routines work to “clean” the data by smoothing noisy data, identifying or removing outliers, resolving inconsistencies and filling in missing values. Data reduction reduces the data size by aggregating, eliminating redundant features, or clustering etc. Data transmission extracts attribute information by reducing data dimension. The essential of data preprocessing is the integration of various data processing techniques. It is one of the most important parts of the whole mining process and most time-consuming as shown in Fig.6 [43].

Responding to the purpose and task of DM and the characteristics of attribute information obtained by data preprocessing, data mining module is applied to choose appropriate mining methods to build model and carry on mining. Finally, the optimum method is chosen by comparing the mining results of these methods.

Scheme evaluation module is responsible for checking and proving the validity of the mining process. If the scheme is evaluated to be unsuccessful, the mining process need restart and adjust the corresponding steps. Therefore, this platform is dynamic and circulatory as illustrated in Fig.7. Because various methods can be adopted in the modules of data preprocessing and data mining, the integration of methods is a key problem and the optimum combination of these methods can be found by constant comparison and adjustment. In addition, new methods can be constantly adopted in this platform to improve the efficiency of DM.
Main tasks and implementation scheme in building DM platform for SHM

Build of four modules in the DM platform can be regarded as the foundation of building the DM platform for SHM. In this part, the main tasks need be solved in defining tasks and building four modules are analyzed separately. Some examples are illustrated on how to utilize specific DM techniques to mine monitoring data.

Task definition

Objects investigated in the DM platform for SHM include bridges, offshore structures, skyscrapers, and benchmark structures etc instrumented with SHM systems. The available monitoring data for mining include structural response (displacement, acceleration, stress, and strain etc), load (wind and vehicle load etc) and environmental factors (temperature, humidity etc). The purpose of mining is to identify the abnormality occurred in structures, evaluate structural damage and state by studying the impacts of wind, vehicle, temperature, humidity etc and their coupling on structures.

Build of data-house module

Two schemes are adopted in building data-house. One scheme is to build a simulated data-house by setting up a numerical simulation platform. Structural analysis software such as ANSYS etc is used to model structures and simulate the structural response on the effect of wind, vehicle, temperature, humidity etc alone and their coupling. Data-house including structural response, load, environmental factors etc is built. Another alternative scheme is to set up an actual data-house consisted of monitoring data collected by SHM systems in engineering fields. The two data-houses can be combined to enrich available monitoring data.

Data preprocessing module

In data preprocessing module, data selection is the first step. For example, when damage identification is performed based on vibration signals, the structural dynamic response (displacement, velocity, acceleration etc) and the external factors (vehicle load, wind load, temperature and humidity etc) that have effect on structural dynamic response should be chosen. The second step is data cleaning. Noise is smoothed by data smoothing techniques such as binning by consulting “its neighborhood”, clustering with the principle to discriminate the values outside of the set of clusters as outliers, combining computer and human inspection. Missing values in actual data-house are filled in by predicted values obtained by the finite element model, and insensitive DM methods to missing values are adopted to reduce their influence. The third step is data reduction where redundant data are deleted by correlation analysis, PCA and clustering etc. The last step is data transformation where data are transformed or consolidated into the forms appropriate for mining by discretization, normalization, wavelet analysis, clustering and so on.

Data mining module

In the data mining module, mining task is defined firstly according to the purpose of data mining and the type of available data. For examples:

1. When coupling of the effects of load and environmental factors on structural responses, or influence of one damage case to other damage cases are studied, the mining task can be defined as association analysis.

2. When signal aberration is detected to predict structural possible damage by analyzing the historical monitoring data, or damage is forecasted by comparing the response of actual structure and finite element model, the mining task can be defined as outlier analysis.

3. When structural state is predicted in a period of time according to the change tendency of historical monitoring data, the mining task can be defined as classification and prediction.

Another task of data mining module is to determine
mining methods based on the determined mining tasks, and choose the optimum method by analyzing the mining results. For example:

(1) When removal of the effect of environment factors is made, PCA etc. can be used to extract and process the principal factors.

(2) When classification of damage cases is made, decision tree, ANNs and SVM etc can be used. A more excellent method among them can be determined according to the accuracy and speed of mining.

(3) When division of sub-structures is made, clustering etc can be adopted.

Scheme evaluation module

In scheme evaluation module, the knowledge and the experience about structures are applied to evaluate the validity of mining scheme. If the object investigated is a finite element model, mining scheme can be evaluated directly by the state of model that is known in advance. If the object investigated is an actual structure, mainly two methods can be used. One is by observing the deviation of actual results from simulated results of the finite element model. Another is by the in-situ examination by structure experts.

CONCLUSIONS

In this paper, the basic theories of DM technology are systematically introduced, the research and application state of DM technology is reviewed and the application prospect of DM technology in SHM is analyzed. The following conclusions are obtained.

(1) Increasing aggregation of SHM data is urging enabling technologies to handle large amount of data automatically. SHM community may benefit from the computational artificial intelligence on this point for a viable technology for SHM of civil infrastructures.

(2) DM technology as a deeper data analysis method has been successfully applied in mechanical fault diagnosis, which makes it possible that monitoring data of civil structures with characteristics of large data size, large redundancy, great varieties and low quality can be sufficiently processed by DM technology.

(3) The research scheme of building DM platform provides a framework for the application of DM technology in SHM for civil structures. This framework is expected to bridge the DM technology and the SHM for civil structures.

ACKNOWLEDGMENTS

This research is financial supported by the National Science Foundation of China under the grant number 50579008.

REFERENCES


[26] Han, J.W. and Kamber, M., Data Mining for Data Warehouses: Uncovering Hidden Patterns, 1995


[33] http://www.daedalus.es/ProdDataEngine-E.php


